Individual differences in replicated multi-product experiments with Thurstonian mixed models for binary paired comparison data

Often sensory discrimination tests are performed with replications for the assessors. In this paper, we suggest a new way of analyzing data from a discrimination study. The model suggested in this paper is a Thurstonian mixed model, in which the variation from the assessors is modelled as a random effect in a generalized linear mixed model. The setting is a multi-product discrimination study with a binary paired comparison. This model makes it possible to embed the analyses of products into one analysis rather than having to do an analysis for each product separately. In addition, it is possible to embed the model into the Thurstonian framework obtaining d-prime interpretations of the estimates. Furthermore, it is possible to extract information about the assessors, even across the products. More specifically, assessor specific d-prime estimates are obtained providing a way to get information about the panel. These estimates are interesting because they make it possible to investigate if the assessors are assessing in a specific way.

Muscle fibre morphology and microarchitecture in cerebral palsy patients obtained by 3D synchrotron X-ray computed tomography

Background: Synchrotron X-ray computed tomography (SXCT) allows for three-dimensional imaging of objects at a very high resolution and in large field-of-view. Purpose: The aim of this study was to use SXCT imaging for morphological analysis of muscle tissue, in order to investigate whether the analysis reveals complementary information to two-dimensional microscopy. Methods: Three-dimensional SXCT images of muscle biopsies were taken from participants with cerebral palsy and from healthy controls. We designed morphological measures from the two-dimensional slices and three-dimensional volumes of the images and measured the muscle fibre organization, which we term orientation consistency. Results: The muscle fibre cross-sectional areas were significantly larger in healthy participants than in participants with cerebral palsy when carrying out the analysis in three dimensions. However, a similar analysis carried out in two dimensions revealed no patient group difference. The present study also showed that three-dimensional orientation consistency was significantly larger for healthy participants than for participants with cerebral palsy. Conclusion: Individuals with CP have smaller muscle fibres than healthy control individuals. We argue that morphometric measures of muscle fibres in two dimensions are generally trustworthy only if the fibres extend perpendicularly to the slice plane, and otherwise three-dimensional aspects should be considered. In addition, the muscle tissue of individuals with CP showed a decreased level of orientation consistency when compared to healthy control tissue. We suggest that the observed disorganization of the tissue may be induced by atrophy caused by physical inactivity and insufficient neural activation.

Real-time interferometric refractive index change measurement for the direct detection of enzymatic reactions and the determination of enzyme kinetics

Back scatter interferometry (BSI) is a sensitive method for detecting changes in the bulk refractive index of a solution in a microfluidic system. Here we demonstrate that BSI can be used to directly detect enzymatic reactions and, for the first time, derive kinetic parameters. While many methods in biomedical assays rely on detectable biproducts to produce a signal, direct detection is possible if the substrate or the product exert distinct differences in their specific refractive index so that the total refractive index changes during the enzymatic reaction. In this study, both the conversion of glucose to glucose-6-phosphate, catalyzed by hexokinase, and the conversion of adenosine-triphosphate to adenosine di-phosphate and mono-phosphate, catalyzed by apyrase, were monitored by BSI. When adding hexokinase to glucose solutions containing adenosine-triphosphate, the conversion can be directly followed by BSI, which shows the increasing refractive index and a final plateau corresponding to the particular concentration. From the initial reaction velocities, $K_M$ was found to be 0.33 mM using Michaelis–Menten kinetics. The experiments with apyrase indicate that the refractive index also depends on the presence of various ions that must be taken into account when using this technique. This study clearly demonstrates that measuring changes in the refractive index can be used for the direct determination of substrate concentrations and enzyme kinetics.
An Epigenome-Wide DNA Methylation Map of Testis in Pigs for Study of Complex Traits

Epigenetic changes are important for understanding complex trait variation and inheritance in pigs that are also a valuable biomedical model for human health research. Testis is the main organ for reproduction and boar taint in pigs; however, there have been no studies to-date on adult pig testis epigenome. The main objective of this study was to establish a genome-wide DNA methylation map of pig testis that would help identify candidate epigenetic biomarkers and methylated genes for complex traits such as male reproduction, fertility or boar taint. Reduced Representation Bisulfite Sequencing (RRBS) was used to study methylation levels of cytosine in nine pig testis samples. The results showed that genome-wide methylation status of nine samples overlapped greatly and their variation among pigs were low. The methylation levels of promoter, exon, intron, cytosine and guanine dinucleotide (CpG) islands and CpG island shores regions were 0.15, 0.47, 0.55, 0.39, and 0.53, respectively. Cytosines binding to CpG islands showed different methylation levels between exon and intron regions. All methylation levels of CpG islands were lower than CpG island shores in different genic features. The distribution of 12,738 differentially methylated cytosines (DMCs) within CpG islands, CpG island shores and other regions was 36.86, 21.65, and 41.49%, respectively, and was 0.33, 1.71, 5.95, and 92.01% in promoter, exon, intron and intergenic regions, respectively. Methylation levels of DMCs in promoter, exon and intron regions were significantly different between CpG islands and CpG island shores (P < 0.05). A total of 898 genes with 2089 DMCs were enriched in 112 Gene Ontology (GO) terms. Fifteen methylated genes from our study were associated with fertility or boar taint traits. Our analysis revealed the methylation patterns in different genic features and CpG island regions of testis in pigs, and summarized several candidate genes associated with DMCs and the involved GO terms. These findings are helpful to understand the relationship between DNA methylation and genic CpG islands, to provide candidate epigenetic regions or biomarkers for pig production and welfare and for translational epigenomic studies that use pigs as an animal model for human research.

General information
Publication status: Published
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science
Contributors: Wang, X., Kadarmideen, H.
Number of pages: 13
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Frontiers in Genetics
Volume: 10
Article number: 405
ISSN (Print): 1664-8021
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
Original language: English
Keywords: Pig, Testis, Epigenome, DNA methylation, RRBS, DMC
Electronic versions:
fgene_10_00405.pdf
DOI:
10.3389/fgene.2019.00405
Source: FindIt
Source-ID: 2446835129
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review

Characteristics of Canteens at Elementary Schools, Upper Secondary Schools and Workplaces that Comply with Food Service Guidelines and Have a Greater Focus on Food Waste

Policy actions to improve the nutritional environment include the provision of official food service guidelines. This study aimed to examine compliance with food service guidelines for hot meals as well as self-evaluated focus on food waste reduction across settings, i.e., elementary schools, upper secondary schools and workplaces, and different canteen characteristics. The same five criteria for hot meals were applied for all settings with regard to serving of fruit and vegetables, fish, wholegrain product and high fat meat and dairy products. A self-administered questionnaire survey was conducted as a cross-sectional study among 680 Danish canteens. Canteens having a high degree of organic food procurement were more likely to comply with the five criteria for hot meals combined (OR 2.00 (CI 1.13,3.53)). Also, the use of organic food together with having a meal policy was associated with reported focus on food waste reduction (OR 1.91 (CI 1.12,3.25) and 1.84 (CI 1.31,2.59), respectively). Compliance with individual criteria varied across settings with elementary schools being more likely to comply with criteria on, e.g., maximum serving of non-wholegrain products, whereas workplaces were more likely to comply with criteria on, e.g., minimum fruit and vegetable content and serving of fish. In addition, specific characteristics, e.g., serving system, were found to predict compliance with some of the criteria. These findings highlight the need to address differences in canteen characteristics when planning implementation support for both guideline and food waste reduction initiatives.
Detection of co-expressed pathway modules associated with mineral concentration and meat quality in Nelore cattle

Meat quality is a complex trait that is influenced by genetic and environmental factors, which includes mineral concentration. However, the association between mineral concentration and meat quality, and the specific molecular pathways underlying this association, are not well explored. We therefore analyzed gene expression as measured with RNA-seq in Longissimus thoracis muscle of 194 Nelore steers for association with three meat quality traits (intramuscular fat, meat pH, and tenderness) and the concentration of 13 minerals (Ca, Cr, Co, Cu, Fe, K, Mg, Mn, Na, P, S, Se, and Zn). We identified seven sets of co-expressed genes (modules) associated with at least two traits, which indicates that common pathways influence these traits. From pathway analysis of module hub genes, we further found an over-representation for energy and protein metabolism (AMPK and mTOR signaling pathways) in addition to muscle growth, and protein turnover pathways. Among the identified hub genes FASN, ELOV5, and PDE3B are involved with lipid metabolism and were affected by previously identified eQTLs associated to fat deposition. The reported hub genes and over-represented pathways provide evidence of interplay among gene expression, mineral concentration, and meat quality traits. Future studies investigating the effect of different levels of mineral supplementation in the gene expression and meat quality traits could help us to elucidate the regulatory mechanism by which the genes/pathways are affected.
Genomic analyses suggest adaptive differentiation of Northern European native cattle breeds

Native domestic breeds represent important cultural heritage and genetic diversity relevant for production traits, environmental adaptation and food security. However, risks associated with low effective population size, such as inbreeding and genetic drift, have elevated concerns over whether unique within-breed lineages should be kept separate or managed as one population. As a conservation genomic case study of the genetic diversity represented by native breeds, we examined native and commercial cattle (Bos taurus) breeds including the threatened Danish Jutland cattle. We examined population structure and genetic diversity within breeds and lineages genotyped across 770K single nucleotide polymorphism loci to determine (a) the amount and distribution of genetic diversity in native breeds, and (b) the role of genetic drift vs. selection. We further investigated the presence of outlier loci to detect (c) signatures of environmental selection in native vs. commercial breeds, and (d) native breed adaptation to various landscapes. Moreover, we included older cryopreserved samples to determine (e) if cryopreservation allows (re)introduction of original genetic diversity. We investigated a final set of 195 individuals and 677K autosomal loci for genetic diversity within and among breeds, examined population structure with principal component analyses and a maximum likelihood approach, and searched for outlier loci suggesting artificial or natural selection. Our findings demonstrate the potential of genomics for identifying the uniqueness of native domestic breeds, and for maintaining their genetic diversity and long-term evolutionary potential through conservation plans balancing inbreeding with carefully designed outcrossing. One promising opportunity is the use of cryopreserved samples, which can provide important genetic diversity for populations with few individuals, while helping to preserve their traditional genetic characteristics. Outlier tests for native vs. commercial breeds identified genes associated with climate adaptation, immunity and metabolism, and native breeds may carry genetic variation important for animal health and robustness in a changing climate.

General information
Publication status: Accepted/In press
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Aalborg University
Contributors: Stronen, A. V., Pertoldi, C., Iacolina, L., Kadarmideen, H. N., Kristensen, T. N.
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Evolutionary Applications
ISSN (Print): 1752-4563
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
Original language: English
Keywords: Animal health, artificial selection, Bos taurus, climate adaptation, conservation genomics, environmental selection, production traits, single nucleotide polymorphism
Electronic versions:
Stronen_et_al_2019_Evolutionary_Applications.pdf
DOIs: 10.1111/eva.12783
Source: FindIt
Source-ID: 2444113462
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review

Improving genomic predictions by correction of genotypes from genotyping by sequencing in livestock populations

Background: Genotyping by sequencing (GBS) is a robust method to genotype markers. Many factors can influence the genotyping quality. One is that heterozygous genotypes could be wrongly genotyped as homozygotes, dependent on the genotyping depths. In this study, a method correcting this type of genotyping error was demonstrated. The efficiency of this correction method and its effect on genomic prediction were assessed using simulated data of livestock populations.

Results: Chip array (Chip) and four depths of GBS data was simulated. After quality control (call rate ≥ 0.8 and MAF ≥ 0.01), the remaining number of Chip and GBS SNPs were both approximately 7,000, averaged over 10 replicates. GBS genotypes were corrected with the proposed method. The reliability of genomic prediction was calculated using GBS, corrected GBS (GBSc), true genotypes for the GBS loci (GBSr) and Chip data. The results
showed that GBSc had higher rates of correct genotype calls and higher correlations with true genotypes than GBS. For genomic prediction, using Chip data resulted in the highest reliability. As the depth increased to 10, the prediction reliabilities using GBS and GBSc data approached those using true GBS data. The reliabilities of genomic prediction using GBSc data were 0.604, 0.672, 0.684 and 0.704 after genomic correction, with the improved values of 0.013, 0.009, 0.006 and 0.001 at depth = 2, 4, 5 and 10, respectively.

Conclusions: The current study showed that a correction method for GBS data increased the genotype accuracies and, consequently, improved genomic predictions. These results suggest that a correction of GBS genotype is necessary, especially for the GBS data with low depths.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University
Contributors: Wang, X., Lund, M. S., Ma, P., Janss, L., Kadarmideen, H., Su, G.
Number of pages: 10
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Journal of Animal Science and Biotechnology
Volume: 10
Issue number: 8
ISSN (Print): 1674-9782
Original language: English
Keywords: Genomic prediction, Genotype correction, Genotyping by sequencing, Simulation

Electronic versions:
Wang_et_al_JASB_2019.pdf
DOIs:
10.1186/s40104-019-0315-z
Source: PublicationPreSubmission
Source-ID: 166244365
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review

Laccase activity measurement by FTIR spectral fingerprinting
Laccases (EC 1.10.3.2) are enzymes known for their ability to catalyze the oxidation of phenolic compounds using molecular oxygen as the final electron acceptor. Laccase activity is commonly determined by monitoring spectrophotometric changes (absorbance) of the product or substrate during the enzymatic reaction. Fourier Transform Infrared Spectroscopy (FTIR) is a fast and versatile technique where spectral evolution profiling, i.e. assessment of the spectral changes of both substrate and products during enzymatic conversion in real time, can be used to assess enzymatic activity when combined with multivariate data analysis. We employed FTIR to monitor enzymatic oxidation of monolignols (sinapyl, coniferyl and p-coumaryl alcohol), sinapic acid, and sinapic aldehyde by four different laccases: three fungal laccases from Trametes versicolor, Trametes villosa and Ganoderma lucidum, respectively, and one bacterial laccase from Meiothermus ruber. By coupling the FTIR measurements with Parallel Factor Analysis (PARAFAC) we established a quantitative assay for assessing laccase activity. By combining PARAFAC modelling with Principal Component Analysis we show the usefulness of this technology as a multivariate tool able to compare and distinguish different laccase reaction patterns. We also demonstrate how the FTIR approach can be used to create a reference system for laccase activity comparison based on a relatively low number of measurements. Such a reference system has potential to function as a high-throughput method for comparing reaction pattern similarities and differences between laccases and hereby identify new and interesting enzyme candidates in large sampling pools.

General information
Publication status: Published
Organisations: Department of Biotechnology and Biomedicine, Section for Protein Chemistry and Enzyme Technology, Enzyme Technology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Number of pages: 10
Pages: 64-73
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Enzyme and Microbial Technology
Volume: 122
Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors

Birth weight variation is influenced by fetal and maternal genetic and non-genetic factors, and has been reproducibly associated with future cardio-metabolic health outcomes. In expanded genome-wide association analyses of own birth weight (n = 321,223) and offspring birth weight (n = 230,069 mothers), we identified 190 independent association signals (129 of which are novel). We used structural equation modeling to decompose the contributions of direct fetal and indirect maternal genetic effects, then applied Mendelian randomization to illuminate causal pathways. For example, both indirect maternal and direct fetal genetic effects drive the observational relationship between lower birth weight and higher later blood pressure: maternal blood pressure-raising alleles reduce offspring birth weight, but only direct fetal effects of these alleles, once inherited, increase later offspring blood pressure. Using maternal birth weight-lowering genotypes to proxy for an adverse intrauterine environment provided no evidence that it causally raises offspring blood pressure, indicating that the inverse birth weight-blood pressure association is attributable to genetic effects, and not to intrauterine programming.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Queensland, Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia, Institute of Biomedical and Clinical Science, College of Medicine and Health, University of Exeter, Royal Devon and Exeter Hospital, Exeter, UK, RIKEN Centre for Integrative Medical Sciences, Laboratory for Endocrinology, Metabolism and Kidney Diseases, Yokohama, Japan, MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK, KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway, Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, Department of Obstetrics and Gynecology, Sahlgrenska University Hospital Östra, Gothenburg, Sweden, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA, Statens Serum Institut, Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway, Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA, Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA, Department of Preventive Medicine, Division of Biostatistics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA, University of Copenhagen, Department of Public Health, Amsterdam Public Health Research Institute, Amsterdam UMC, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands, Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA, The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands, Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands, Department of Computational Biology, University of Lausanne, Lausanne, Switzerland, Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland, Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, Department of Clinical Chemistry, Fimalab Laboratories, Tampere, Finland, Section of Epidemiology and Biostatistics, School of Population Health, University of Auckland, Auckland, New Zealand, Department of Twin Research, King's College London, St. Thomas' Hospital, London, UK, Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA, Department of Genetics, University of North Carolina, Chapel Hill, NC, USA, Institute of Epidemiology, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany, School of Medicine and Public Health, Faculty of Medicine and Health, The University of Newcastle, Newcastle, New South Wales, Australia, Center for Genomic Regulation, Barcelona Institute of Science and Technology, Barcelona, Spain, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK, QIMR Berghofer Medical Research Institute, Royal Brisbane Hospital, Brisbane, Queensland, Australia, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK, Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway, Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland, Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands, Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland, ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain, Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK, Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK, CIBER Epidemiología y Salud Pública, Sidra Medicine Research Department, Sidra Medicine, Doha, Qatar, Division of Obstetrics and Gynaecology, The University of Western Australia, Perth, Western Australia, Australia, Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany, Department of Life Sciences and Biotechnology, Genetic Section, University of Ferrara, Ferrara, Italy, Fifth Department of Medicine, University Medical Centre Mannheim.
Multiblock PLS: Block dependent prediction modeling for Python

Partial Least Squares (PLS) regression is a statistical method for supervised multivariate analysis. It relates two data blocks X and Y to each other with the aim of establishing a prediction model. When deployed in production, this model can be used to predict an outcome y from a newly measured feature vector x. PLS is popular in chemometrics, process control and other analytic fields, due to its striking advantages, namely the ability to analyze small sample sizes and the ability to handle high-dimensional data with cross-correlated features (where Ordinary Least Squares regression typically fails). In addition, and in contrast to many other machine learning approaches, PLS models can be interpreted using its latent variable structure just like principal components can be interpreted for a PCA analysis.

Multi-phase image segmentation with the adaptive deformable mesh

This paper proposes a method for image segmentation using a deformable triangle mesh in the image domain. We define a piecewise constant function by labeling the mesh triangles with different phases, each representing a segment of an image. Our method finds the optimal mesh configuration and triangle labeling that minimize the piecewise constant Mumford-Shah functional. Contributions of this paper include a force model that moves mesh vertices towards the solution, and an adaptivity model that further adapts the mesh by introducing or removing vertices. The results demonstrate the advantages of our method over traditional methods like snakes and level set. Our approach supports multi-phase segmentation incurring no particular overhead. Furthermore, the use of an adaptive mesh facilitates accurate segmentation with a very compact representation. The biggest challenge of deformable meshes, changes to the topology of the segments, is handled by employing Deformable Simplical Complex (DSC), a method for explicit interface tracking.
Order statistics and multivariate discrete phase-type distributions.
In this thesis, we present three different topics of research which are related to the theory of phase-type distributions. Those topics are explained next. The first research work is on order statistics from matrix-geometric distributions in the case of a sample of independent and non-identically distributed random variables. We prove that order statistics from matrix-geometric distributions are matrix-geometric distributed and we provide representations for their distributions. The second research work is a study of the discrete version of multivariate phase-type distributions introduced by V. G. Kulkarni. We give an expression for the joint probability-generating function in the similar way than in the continuous time case and under this base we make an analysis of this class of distributions and present examples that are commonly found in the literature. The third research work presented came out with the aim of relating the last two topics. That is, we found a problem which relates the concept of order statistics and multivariate phase-type distributions introduced by V. G. Kulkarni, the last in the case of continuous time. Thus, we present a research on concomitants of phase-type distributions. We provide a procedure to calculate the density function of concomitants of phase-type distributions and we prove that concomitants of phase-type distributions are phase-type distributed.

Pharmacogenomics of Methylphenidate Treatment of Danish Children With ADHD
Background : The first choice drug for treatment of Attention Deficit Hyperactivity Disorder (ADHD) in Denmark and several other countries is Methylphenidate (MPH). More than 30% of the children treated with MPH have been reported to switch to another drug due to lack of efficacy or adverse reactions. Most likely, a significant fraction of this individual variation in treatment outcome is caused by genetic differences. Previous studies have not identified genes or genetic variants conclusively implicated in the response to MPH. Methods : We performed a pharmacogenomics study of the efficacy and adverse reactions of MPH in a sample of 207 Danish children with ADHD, who were drug-naïve at baseline and were monitored for 12 weeks. Each week the ADHD symptom severity was scored on the ADHD-RS questionnaire, and any adverse reactions were recorded. Results : None of the GWAS resulted in genome-wide associated hits. However, among the top SNPs for the inattention sub-score were some that fell in: RAS guanyl releasing protein 1
For the hyperactivity/impulsive sub-score one of the top SNPs is in the vicinity of gamma-aminobutyric acid type A receptor gamma3 subunit (GABRG3). Discussion: Despite the lack of genome-wide significant SNPs, some of the top hits from the GWAS reside in or near genes that could modulate the response to MPH. Some of these genes have previously been implicated in psychiatric disorders (RASGRP1, NKAIN2 and MYT1L), are involved in neurotransmission (GABRG3, KCNQ5, KCNS3 and NKAIN2) or have been implicated in substance abuse, e.g. TTC12 has been implicated in heroin dependence and SYT17 in alcoholism.

Predicting and mapping human risk of exposure to Ixodes ricinus nymphs using climatic and environmental data, Denmark, Norway and Sweden, 2016

Background: Tick-borne diseases have become increasingly common in recent decades and present a health problem in many parts of Europe. Control and prevention of these diseases require a better understanding of vector distribution. Aim: Our aim was to create a model able to predict the distribution of I. ricinus nymphs in southern Scandinavia and to assess how this relates to risk of human exposure. Methods: We measured the presence of I. ricinus tick nymphs at 159 stratified random lowland forest and meadow sites in Denmark, Norway and Sweden by dragging 400 m transects from August to September 2016, representing a total distance of 63.6 km. Using climate and remote sensing environmental data and boosted regression tree modelling, we predicted the overall spatial distribution of I. ricinus nymphs in Scandinavia. To assess the potential public health impact, we combined the predicted tick distribution with human density maps to determine the proportion of people at risk. Results: Our model predicted the spatial distribution of I. ricinus nymphs with a sensitivity of 91% and a specificity of 60%. Temperature was one of the main drivers in the model followed by vegetation cover. Nymphs were restricted to only 17.5% of the modelled area but, respectively, 73.5%, 67.1% and 78.8% of the human populations lived within 5 km of these areas in Denmark, Norway and Sweden. Conclusion: The model suggests that increasing temperatures in the future may expand tick distribution geographically in northern Europe, but this may only affect a small additional proportion of the human population.
Process characterization for molding of paper bottles using computed tomography and structure tensor analysis

Packaging products find their significance in almost all classes of consumer goods and products. The use of plastic and metal based packaging for beverages is highly dominant. However, there is a constant urge for development of eco-friendly packaging alternatives. The article focuses on characterizing an inflatable core assisted paper bottle molding process with respect to the obtained fiber distribution in the bottle. Distribution of paper fibers affect product characteristics such as thickness and mechanical strength of the bottle. Assessment of fiber orientation using structure tensor analysis is therefore performed. The results confirmed non-uniform fiber compaction in the paper bottle. This gives rise to non-conformities such as non-uniform thickness distribution. The approach discussed in the work can be utilized as a Non Destructive Testing technique to evaluate the quality of paper bottles.

General information
Publication status: Published
Organisations: Manufacturing Engineering, Department of Mechanical Engineering, Department of Physics, Neutrons and X-rays for Materials Physics, Image Analysis & Computer Graphics, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science
Number of pages: 7
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: E-Journal of Nondestructive Testing & Ultrasonics
Volume: 24
Issue number: 3
ISSN (Print): 1435-4934
Original language: English
Keywords: Paper bottle, Fiber orientation, Molded paper products, Structure tensor analysis, Molding process
Electronic versions:
Research output: Contribution to journal › Conference article – Annual report year: 2019 › Research › peer-review

Puncture of an import gasoline pipeline – spray effects may evaporate more fuel than a Buncefield-type tank overfill event

This paper is concerned with evaporation of moderately volatile liquids, gasoline in particular, due to spray generation, liquid fragmentation and fountain effects following accidental puncture of a pressurized pipeline. Hazard analysis predicts that extensive evaporation will take place. The paper examines a typical fuel depot receiving gasoline from a ship at a nearby port via an above-ground pipeline. For comparative purposes, two types of accidental release during import are considered: 1) The receiving tank overflows in a worst-case Buncefield-type event (baseline). 2) The import pipeline is punctured and a jet of liquid discharges upwards. The paper examines pipeline import of three substances, hexane, octane and winter gasoline. Hazard analysis using the PHAST software suite indicates that the amount of fuel evaporated from the pipeline puncture scenarios greatly exceeds the amount evaporated in a tank overfill event for all three substances, gasoline in particular. Proper modelling of evaporation of wide-range multi-component mixtures such as gasoline is challenging however. PHAST's simplified thermodynamic modelling of properties of mixtures may be a source of error. A PHAST-based stand-alone spray evaporation model with advanced thermodynamic capability is developed. Results indicate that PHAST does indeed overestimate evaporation of mixtures. Still, model output shows that evaporation following pipeline puncture may exceed the evaporation from a Buncefield-type tank overfill event by a factor of two or more. This finding is significant as evaporation from pipeline puncture scenarios appear largely overlooked in hazard analysis. The finding may lead to a fundamental re-appraisal of the hazard potential of fuel depots and pipelines.
Staff optimization for time-dependent acute patient flow

The emergency department is a key element of acute patient flow, but due to high demand and an alternating rate of arriving patients, the department is often challenged by insufficient capacity. Proper allocation of resources to match demand is, therefore, a vital task for many emergency departments. Constrained by targets on patient waiting time, we consider the problem of minimizing the total amount of staff-resources allocated to an emergency department. We test a matheuristic approach to this problem, accounting for both patient flow and staff scheduling restrictions. Using a continuous-time Markov chain, patient flow is modeled as a time-dependent queueing network where inhomogeneous behavior is evaluated using the uniformization method. Based on this modeling approach, we recursively evaluate and allocate staff to the system using integer linear programming until the waiting time targets are respected in all queues of the network. By comparing to discrete-event simulations of the associated system, we show that this approach is adequate for both modeling and optimizing the patient flow. In addition, we demonstrate robustness to the service time distribution and the associated system with multiple classes of patients.

General information
Publication status: Published
Organisations: Department of Management Engineering, Management Science, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Pages: 94-105
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: European Journal of Operational Research
Volume: 272
ISSN (Print): 0377-2217
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
Original language: English
Keywords: OR in health services, Queueing, Markov chain, Heuristics, Stochastic optimization
DOIs:
10.1016/j.ejor.2018.06.015
Source: FindIt
Transcriptomics–genomics data integration and expression quantitative trait loci analyses in oocyte donors and embryo recipients for improving in vitro production of dairy cattle embryos

In this paper we first provide a brief review of main results from our previously published studies on genomewide gene expression (transcriptomics) in donor and recipient cattle used in in vitro production (IVP) of embryos and embryo transfer (ET). Then, we present novel results from applying integrative systems genomics and biological analyses where transcriptomics data are combined with genomic data in both donor and recipient cattle to map expression quantitative trait loci (eQTLs). The eQTLs are genetic markers that can regulate or control the expression of genes in the entire genome, via complex molecular mechanisms, and thus can act as a powerful tool for genomic and gene-assisted selection. We identified significant eQTLs potentially controlling the expression of 13 candidate genes for donor cow quality (IVP parameters; e.g. cyclin B1 (CCNB1), outer dense fiber of sperm tails 2 like (ODF2L)) and 19 candidate genes for recipient cows quality (endometrial receptivity; e.g. ER membrane protein complex subunit 9 (EMC9), mannosidase beta (MANBA), peptidase inhibitor 16 (PI16)). Annotation and colocation of detected eQTLs show that some of the eQTLs are in the same genomic regions previously reported as QTLs for reproduction-related traits. However, eQTLs and the candidate genes identified should be further validated in larger populations before implementation as genetic markers or used in genomic selection for improving IVP and ET performance.

General information
Publication status: Published
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Department of Health Technology
Contributors: Kadarmideen, H., Mazzoni, G.
Pages: 55–67
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Reproduction, Fertility and Development
Volume: 31
ISSN (Print): 1031-3613
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
Original language: English
Keywords: Candidate genes, Cattle IVF, Genomic breeding, Omics data
DOIs: 10.1071/RD18338

Zonohedral Approximation of Spherical Structuring Element for Volumetric Morphology
Performing dilation and erosion using large structuring elements can be computationally slow – a problem especially pronounced when processing volumetric data. To reduce the computational complexity of dilation/erosion using spherical structuring elements, we propose a method for approximating a sphere with a zonohedron. Since zonohedra can be created via successive dilations/erosions of line segments, this allows morphological operations to be performed in constant time per voxel. As the complexity of commonly used methods typically scales with the size of the structuring element, our method significantly improves the run time. We use the proposed approximation to detect large spherical objects in volumetric data. Results are compared with other image analysis frameworks demonstrating constant run time and significant performance gains.

General information
Publication status: Accepted/In press
Organisations: Image Analysis & Computer Graphics, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Technical University of Denmark
Number of pages: 12
Publication date: 2019

Host publication information
Title of host publication: Proceedings of Scandinavian Conference on Image Analysis 2019
Publisher: Springer
(Lecture Notes in Computer Science).
Keywords: Morphology, Computational Efficiency, Zonohedra
Monitoring batch processes with dynamic time warping and k-nearest neighbours

A novel data driven approach to batch process monitoring is presented, which combines the k-Nearest Neighbour rule with the dynamic time warping (DTW) distance. This online method (DTW-NN) calculates the DTW distance between an ongoing batch, and each batch in a reference database of batches produced under normal operating conditions (NOC). The sum of the k smallest DTW distances is monitored. If a fault occurs in the ongoing batch, then this distance increases and an alarm is generated. The monitoring statistic is easy to interpret, being a direct measure of similarity of the ongoing batch to its nearest NOC predecessors and the method makes no distributional assumptions regarding normal operating conditions. DTW-NN is applied to four extensive datasets from simulated batch production of penicillin, and tested on a wide variety of fault types, magnitudes and onset times. Performance of DTW-NN is contrasted with a benchmark multiway PCA approach, and DTW-NN is shown to perform particularly well when there is clustering of batches under NOC.

New insights into the spatial genetic structure of the Indian riverine buffalo populations

Several programmes for the genetic improvement and conservation Indian buffalo breeds have been implemented by various agencies since 1970. These programmes may have shaped the present genetic diversity and thereby the evolutionary potential of buffalo populations across India. In this study, we attempted to provide an insight into the contemporary genetic structure of buffaloes by simultaneously analyzing the genetic and geospatial data. A battery of 100 microsatellite markers was genotyped across 1055 buffalo samples representing different geographical regions of the country. Bayesian and multivariate approaches were used for analyses. Although the clustering results from the different approaches were not convergent, these methods were able to identify sub-structuring within the UP buffalo population. In our study BAPS was able to distinguish most of the buffalo breeds with better spatial distribution. The genetic structuring detected in our study is mainly characterised by isolation by distance and also reflects the effect of improvement programmes. Our analysis suggested genetic connectivity of Murrah and Nili-Ravi with several buffalo breeds, as they have been used for breed improvement. The information may be useful in identifying ecologically distinct or connected populations for future improvement and conservation programmes.
**Hvidvaskning med rejsekort**


---

**General information**

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 15 - 1, sektion
Publication date: 28 Sep 2018
Peer-reviewed: Unknown

---

**Publication information**

Journal: Weekendavisen
Issue number: 39
ISSN (Print): 0106-4142
Original language: Danish

Electronic versions:
Hvidvask.pdf
Source: PublicationPreSubmission
Source-ID: 154488107

Research output: Contribution to journal › Contribution to newspaper - Comment/debate – Annual report year: 2018 › Communication

---

**Breakfast in Denmark. Prevalence of Consumption, Intake of Foods, Nutrients and Dietary Quality. A Study from the International Breakfast Research Initiative**

Breakfast is considered by many to be the most important meal of the day. This study examined the intake of nutrients and foods at breakfast among Danes and the relation to the overall dietary quality. Data were derived from the Danish National Survey on Diet and Physical Activity 2011-2013, a cross-sectional national food consumption study. A total of 3680 participants aged 6-75 years were included in the analyses of breakfast consumption. The Nutrient Rich Food Index 9.3 method was used to examine the overall dietary quality of the diet. The intake of nutrients and foods at breakfast were compared across dietary quality score tertiles by ANCOVA adjusted for energy and socio economic status. Breakfast was eaten frequently by children and adults and contributed with 18-20% of total energy intake. Breakfast was relatively high in dietary fibre, B vitamins, calcium and magnesium and low in added sugar, total fat, sodium, vitamin A and D. A decrease in the intake of added sugar, total fat and saturated fat and an increase in the intake of dietary fibre and most micronutrients were seen across tertiles of dietary quality scores. Commonly consumed foods provided at breakfast in Denmark included bread, breakfast cereals and dairy products as well as water, coffee and juice, while intakes of fruits, vegetables, cakes and soft drinks were low.

---

**General information**

Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Fagt, S., Matthiessen, J., Thyregod, C., Kørup, K., Biltoft-Jensen, A. P.
How a 10-day heatwave impacts barley grain yield when superimposed onto future levels of temperature and CO₂ as single and combined factors

Heatwaves pose a threat to crop production and are predicted to increase in frequency, length and intensity as a consequence of global warming. Future heatwaves will occur in addition to the ongoing increase of mean temperature and CO₂. To test effects of heatwaves superimposed to future climate scenarios, 22 barley accessions were cultivated with elevated temperature (+5 °C) and CO₂ (700 ppm) as single factors and in combination. The control treatment mimicked ambient Scandinavian early summer conditions (19/12 °C, day/night; 400 ppm CO₂). Around flowering a 10-day heatwave of 33/28 °C (day/night) was superimposed to all treatments. The lowest average grain yield was observed when the heatwave was superimposed onto the combined elevated temperature and CO₂ treatment. Here the yield decreased by 42% compared to no heatwave and 52% compared to ambient conditions. When the heatwave was superimposed onto ambient conditions the average grain yield decreased by 37% compared to no heatwave. There was no significant difference between the relative grain yield decrease caused by the heatwave in the ambient and future climate scenarios. In contrast, the vegetative aboveground biomass increased upon heatwave exposure, leading to a strong decline in the harvest index. Our results strongly emphasize the need to produce heatwave resilient cultivars.
Transcription factor co-expression networks of adipose RNA-Seq data reveal regulatory mechanisms of obesity

Background: Transcription Factors (TFs) control actuation of genes in the genome and are key mediators of complex processes such as obesity. Master Regulators (MRs) are the genes at the top of a regulation hierarchy which regulate other genes. Objective: To elucidate clusters of highly co-expressed TFs (modules), involved pathways, highly interconnected TFs (hub-TFs) and MRs leading to obesity and leanness, using porcine model for human obesity. Methods: We identified 817 expressed TFs in RNA-Sequencing dataset representing extreme degrees of obesity (DO; lean, obese). We built a single Weighted Transcription Factor Co-expression Network (WTFCN) and TF sub-networks (based on the DO). Hub-TFs and MRs (using iRegulon) were identified in biologically relevant WTFCNs modules. Results: Single WTFCN detected the Red module significantly associated with DO (P < 0.03). This module was enriched for regulation processes in the immune system, e.g.: Immune system process (Padj = 2.50E-06) and metabolic lifestyle disorders, e.g. Circadian rhythm - mammal pathway (Padj = 2.33E-11). Detected MR, hub-TF SPI1 was involved in obesity, immunity and osteoporosis. Within the obese sub-network, the Red module suggested possible associations with immunity, e.g. TGF-beta signaling pathway (Padj = 1.73E-02) and osteoporosis, e.g. Osteoclast differentiation (Padj = 1.94E-02). Within the lean sub-network, the Magenta module displayed associations with type 2 diabetes, obesity and osteoporosis e.g. Notch signaling pathway (Padj = 2.40E-03), osteoporosis e.g. hub-TF VDR (a prime candidate gene for osteoporosis). Conclusion: Our results provide insights into the regulatory network of TFs and biologically relevant hub TFs in obesity.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Corresponding author: Kadarmideen, H. N.
Contributors: Skinkyte-Juskiene, R., Kogelman, L. J., Kadarmideen, H. N.
Pages: 289-299
Publication date: 1 May 2018
Peer-reviewed: Yes

Publication information
Journal: Current Genomics
Volume: 19
Issue number: 4
ISSN (Print): 1389-2029
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Keywords: Gene networks, Obesity, RNA-Seq, Transcription factors, Transcriptomics, WGCNA
DOIs:
10.2174/1389202918666171005095059
Source: Scopus
Source-ID: 85045920209
Research output: Contribution to journal ➔ Journal article – Annual report year: 2018 ➔ Research ➔ peer-review

Approximation of ruin probabilities via Erlangized scale mixtures
In this paper, we extend an existing scheme for numerically calculating the probability of ruin of a classical Cramér–Lundbergprocess having absolutely continuous but otherwise general claim size distributions. We employ a dense class of distributions that we denominate Erlangized scale mixtures (ESM) that correspond to nonnegative and absolutely continuous distributions which can be written as a Mellin–Stieltjes convolution Π⋆G of a nonnegative distribution Π with an Erlang distribution G. A distinctive feature of such a class is that it contains heavy-tailed distributions. We suggest a simple methodology for constructing a sequence of distributions having the form Π⋆G with the purpose of approximating the integrated tail distribution of the claim sizes. Then we adapt a recent result which delivers an explicit expression for the probability of ruin in the case that the claim size distribution is modeled as an Erlangized scale mixture. We provide simplified expressions for the approximation of the probability of ruin and construct explicit bounds for the error of approximation. We complement our results with a classical example where the claim sizes are heavy-tailed.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Liverpool, University of Queensland
Corresponding author: Peralta, O.
Contributors: Peralta, O., Rojas-Nandayapa, L., Xie, W., Yao, H.
Pages: 136-156
Rare-events classification - An approach based on genetic algorithm and voronoi tessellation

Classification is a major constituent of the data mining tool kit. Well-known methods for classification are either built on the principle of logic or on statistical reasoning. For imbalanced and noisy cases, classification may however fail to deliver on basic data mining goals, i.e., identifying statistical dependencies in data. In this article, we propose a novel strategy for data mining based on partitioning of the feature space through Voronoi tessellation and Genetic Algorithm, where the latter is applied to solve a combinatorial optimization problem. We apply the suggested methodology to a range of classification problems of varying imbalance and noise and compare the performance of the suggested method with well-known classification methods such as (SVM, KNN, and ANN). The results obtained indicate the proposed methodology to be well suited for data mining tasks in case of highly imbalanced classes and significant noise.

Added Value of Individual Flexibility Profiles of Electric Vehicle Users For Ancillary Services

Vehicle-Grid Integration (VGI) research may serve to limit the self-induced adverse effects of electric vehicles (EVs) in terms of additional grid loading, but also as to make the EV an active asset in supporting a stable, economic power system based on renewable energy. Any use of the vehicle for grid services requires an accurate understanding of the user’s driving needs. This paper proposes the introduction of a user profile, describing the energy requirements for driving in terms of an energy deadline, target and minimum. To explore the use of such a profile, the paper analyses data from a Danish pilot project where the driving patterns of ten electric Nissan e-NV200 vans are investigated in terms of leave times and energy consumption. It is shown that the data can be fitted with a log-normal distribution that can be used to establish a per user profile which provides a certain statistical probability of fulfilling the driving needs while allowing an aggregator to optimize earnings. Initially, aggregators may apply similar driving assumptions across an entire fleet. Considering that the driving needs of individual EV owners are different, statistical representations of the individual behaviour may result in more flexibility, and thereby time, for providing grid services. The paper quantifies the
value of such added flexibility based on the Danish market for frequency containment reserves.

**General information**
Publication status: Published
Organisations: Distributed Energy Resources, Center for Electric Power and Energy, Department of Electrical Engineering, Energy Analytics and Markets, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science
Contributors: Andersen, P. B., Sousa, T., Thingvad, A., Berthou, L. S., Kulahci, M.
Number of pages: 6
Publication date: 2018

**Advances of matrix-analytic methods in risk modelling**
This work is concerned with the study of matrix-analytic methods with novel applications to the area of risk theory. First, we review some topics of Applied Probability such as phase-type distributions, matrix-exponential distributions, Markovian arrival processes, Rational arrival processes, fluid flow processes and risk models. With these tools in hand, we propose a method to approximate the probability of ruin of any Cramér-Lundberg process using the theory of phase-type distributions, providing an error bound for such an approximation. With the goal of studying risk models with dependencies, we construct a class of bivariate distributions with given phase-type-distributed marginals and given Pearson's correlation coefficient, which are later used to define different kinds of dependent Sparre-Andersen processes. Later on, we give an explicit formula for the probability of Parisian and cumulative Parisian ruin for a class of risk processes which are based on the theory of fluid flow processes. Next, we study some excursion properties of spectrally negative Lévy processes whenever they are inspected at an independent matrix-exponential time. Finally, inspired by the generalisation of the Markovian arrival process to the Rational arrival process, we construct a novel generalisation of the fluid flow process without Brownian components and study its first passage probabilities.

**General information**
Publication status: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Peralta Gutierrez, O.
Number of pages: 215
Publication date: 2018

**A Methodology for Development of a Pedagogical Simulation Tool used in Fermentation Applications**
In this work, a holistic methodology was developed and implemented in a computer-aided framework with the purpose of establishing a pedagogical simulation tool for bioprocess applications. The methodology integrates the development of a process simulator with the use of gamification elements to improve the pedagogical experience. Moreover, model development and reuse was incorporated into the framework through creating a generic model template. The application of the methodology was demonstrated by the simulation of a batch fermentation process of Saccharomyces cerevisiae cultivated under aerobic conditions with glucose as substrate, thereby, demonstrating mixed metabolism dynamics and considering the inhibition of ethanol consumption due to the presence of glucose with an integrated learning design and gamification elements. The learning experience, done with beta users with different previous levels of knowledge about fermentation processes, confirmed an interest for using gamification for training chemical engineering, biochemical engineering, and biotechnology students.

**General information**
Publication status: Published
Organisations: PROSYS - Process and Systems Engineering Centre, Department of Chemical and Biochemical Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
A multimodal data-set of a unidirectional glass fibre reinforced polymer composite

A unidirectional (UD) glass fibre reinforced polymer (GFRP) composite was scanned at varying resolutions in the micro-scale with several imaging modalities. All six scans capture the same region of the sample, containing well-aligned fibres inside a UD load-carrying bundle. Two scans of the cross-sectional surface of the bundle were acquired at a high resolution, by means of scanning electron microscopy (SEM) and optical microscopy (OM), and four volumetric scans were acquired through X-ray computed tomography (CT) at different resolutions. Individual fibres can be resolved from these scans to investigate the micro-structure of the UD bundle. The data is hosted at https://doi.org/10.5281/zenodo.1195879 and it was used in [1] to demonstrate that precise and representative characterisations of fibre geometry are possible with relatively low X-ray CT resolutions if the analysis method is robust to image quality.

Analysis of MRI by fractals for prediction of sensory attributes: A case study in loin

This study investigates the use of fractal algorithms to analyse MRI of meat products, specifically loin, in order to determine sensory parameters of loin. For that, the capability of different fractal algorithms was evaluated (Classical Fractal Algorithm, CFA; Fractal Texture Algorithm, FTA and One Point Fractal Texture Algorithm, OPFTA). Moreover, the influence of the acquisition sequence of MRI (Gradient echo, GE; Spin Echo, SE and Turbo 3D, T3D) and the predictive technique of data mining (Isotonic regression, IR and Multiple Linear regression, MLR) on the accuracy of the prediction was analysed. Results on this study firstly demonstrate the capability of fractal algorithms to analyse MRI from meat product. Different combinations of the analysed techniques can be applied for predicting most sensory attributes of loins adequately (R > 0.5). However, the combination of SE, OPFTA and MLR offered the most appropriate results. Thus, it could be proposed as an alternative to the traditional food technology methods.
Sensory discrimination tests are used to gain information about products by using the human senses to evaluate the samples. More specifically, a sensory discrimination study is conducted when the aim is to investigate whether products are perceptibly different. Such studies are often considered for food, beverages as well as personal care products. An example is when a company gets a new supplier of an ingredient in one of their products. It is of high importance to investigate how this change of the ingredient affects the product. Even though the chemical composition of the product changes, it does not necessarily mean that people can detect the difference. These days, people become more and more interested in how to improve their health. This is also reflected in the companies desire to make their products healthier without changing how the product is perceived by their customers. Therefore, it is important to conduct sensory discrimination tests when ingredients are changed. This thesis is concerned with the analysis of product and individual differences in sensory discrimination testing.

Sensory discrimination tests become more and more advanced raising a need for new types of analysis of sensory discrimination data. This thesis contributes with the development of Thurstonian models and how these can be aligned with well-known statistical models. Generalized linear mixed models are used in many applications. However, it is not common to consider such complicated models when considering sensory discrimination tests. Actually, sensory discrimination tests are often analyzed by too simplistic methods, ignoring important variables, such as individuals, that affect the results of the analysis. One focus of this project is to propose a way to incorporate such effects in the models when analyzing data from sensory discrimination studies. These models, including random effects, are called Thurstonian mixed models. Considering generalized linear mixed models for sensory discrimination studies opens up for many possibilities. It becomes possible to gain information about the individuals, the so-called assessors, as well as making more proper conclusions regarding the products. Moreover, the estimates of product and individual differences are obtained on the d-prime scale. Often multiple sensory attributes are considered in a discrimination study. These can be analyzed individually by the Thurstonian mixed models we are introducing. This thesis is presenting a multivariate analysis to gain knowledge about the product and individual differences across the sensory attributes. This is achieved by analyzing the product and individual differences, on the d-prime scale, by principal component analysis.

Sensory discrimination tests are sometimes conducted to investigate the performance of sensory panels or to compare different laboratories. In such tests, multiple d-prime values can be obtained. For sensory discrimination tests, which lead to binomially distributed responses, we propose a new test statistic for the comparison of multiple d-prime values. The test statistic we suggest is an improved way of analyzing multiple d-prime values compared to a previous suggested test statistic.
Application of data clustering to railway delay pattern recognition

K-means clustering is employed to identify recurrent delay patterns on a high traffic railway line north of Copenhagen, Denmark. The clusters identify behavioral patterns in the very large ("big data") data sets generated automatically and continuously by the railway signal system. The results reveal where corrective actions are necessary, showing where recurrent delay patterns take place. Delay profiles and delay-change profiles are generated from timestamps to compare different train runs, and to partition the set of observations into groups of similar elements. K-means clustering can identify and discriminate different patterns affecting the same stations, which is otherwise difficult in previous approaches based on visual inspection. Classical methods of univariate analysis do not reveal these patterns. The demonstrated methodology is scalable and can be applied to any system of transport.

A Rollercoaster to Model Touch Interactions during Turbulence

We contribute to a project introducing the use of a large single touch-screen as a concept for future airplane cockpits. Human-machine interaction in this new type of cockpit must be optimised to cope with the different types of normal use as well as during moments of turbulence (which can occur during flights varying degrees of severity). We propose an original experimental setup for reproducing turbulence (not limited to aviation) based on a touch-screen mounted on a rollercoaster. Participants had to repeatedly solve three basic touch interactions: a single click, a one-finger drag-and-drop, and a zoom operation involving a 2-finger pinching gesture. The completion times of the different tasks as well as the number of unnecessary interactions with the screen constitute the collected user data. We also propose a data analysis and statistical method to combine user performance with observed turbulence, including acceleration and jerk along the different axes. We then report some of the implications of severe turbulence on touch interaction and make recommendations as to how this can be accommodated in future design solutions.
**Assessing spatial correlations of sea surges around Copenhagen**

**General information**
- Publication status: Published
- Organisations: Department of Applied Mathematics and Computer Science, National Food Institute, Research group for Risk Benefit, Department of Environmental Engineering, Urban Water Systems, Statistics and Data Analysis
- Contributors: Georgiadis, S., Sørup, H. J. D., Nielsen, B. F., Arnbjerg-Nielsen, K.
- Pages: 36-36
- Publication date: 2018

**Host publication information**
- Title of host publication: Danish Water Forum Annual Water Conference 2018 - abstract book
- Place of publication: Lyngby, Denmark
- Publisher: Danish Water Forum
- Abstract book

**Associations Between DrugStar App Events and Perceived benefits of Diabetes Patients**

**General information**
- Publication status: Published
- Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
- Contributors: Lyndgaard, C. B.
- Number of pages: 48
- Publication date: 2018

**Publication information**
- Publisher: DTU Compute
- Original language: English

**Behavior of passenger train delays and estimation of network effects**

**General information**
- Publication status: Published
- Organisations: Operations Management, Transport DTU, Management Science, Department of Management Engineering, Transport Modelling, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Banedanmark, Ansaldo STS
Benefits of spatio-temporal modelling for short term wind power forecasting at both individual and aggregated levels

The share of wind energy in total installed power capacity has grown rapidly in recent years. Producing accurate and reliable forecasts of wind power production, together with a quantification of the uncertainty, is essential to optimally integrate wind energy into power systems. We build spatio-temporal models for wind power generation and obtain full probabilistic forecasts from 15 minutes to 5 hours ahead. Detailed analysis of the forecast performances on the individual wind farms and aggregated wind power are provided. The predictions from our models are evaluated on a data set from wind farms in western Denmark using a sliding window approach, for which estimation is performed using only the last available measurements. The case study shows that it is important to have a spatio-temporal model instead of a temporal one to achieve calibrated aggregated forecasts. Furthermore, spatio-temporal models have the advantage of being able to produce spatially out-of-sample forecasts. We use a Bayesian hierarchical framework to obtain fast and accurate forecasts of wind power generation at wind farms where recent data are available, but also at a larger portfolio including wind farms without recent observations of power production. The results and the methodologies are relevant for wind power forecasts across the globe as well as for spatial-temporal modelling in general.

Biadditive Mixed Models - Advancing Computational Methods and Applications

Analysis of Variance (ANOVA) is widely used to analyze data in most scientific areas. However, the standard linear ANOVA models are not always adequate to describe the structures in a data set sufficiently. This means that an improved inference and a better insight might be obtained by extending the linear ANOVA models. An example of this, is the multiplicative models also named biadditive models, which arise when one or more multiplicative terms are added to a linear ANOVA model. These models are especially popular within agriculture to analyze genotype-by-environment data, but they are also used in e.g. sensometrics to analyse sensory profile data or in medicine to analyse data from method comparison studies. In general, these models are relevant whenever an interaction between two factors is not completely unstructured, but can be described either fully or partly by a linear regression, where one of the variables in the multiplicative term can be interpreted as the regressor and the other variable as the slope.

In this thesis, the main focus is on a specific version of the multiplicative mixed models, with the multiplicative term being a product of a fixed effect and a random effect, where the fixed effect is a part of the mean structure. This means that this fixed effect enters both the mean and the variance structure, which complicates the estimation of the model. One main goal of this work was to develop a user-friendly open-source software tool to fit this kind of models. For this purpose, R-package mumm was created, which is now available on CRAN. The thesis demonstrates how to use the package, which was found to be faster than the commercial alternative. Another aim of this thesis was to investigate the advantages obtained by using the multiplicative mixed model, instead of a simple linear mixed ANOVA model. By simulation studies, it
was demonstrated that the power to detect a significant fixed effect increases by using a multiplicative mixed model instead of a two-way mixed ANOVA model, when the 'multiplicative effect' is present in the data.

In sensometrics, a linear approximation to the multiplicative mixed model was presented in Brockhoff et al. (2015). This model is named the Mixed Assessor Model (MAM), and is used to analyze sensory profile data, where a panel of judges assesses different products and scores them in relation to a specific characteristic. In this thesis it was showed that also by using the MAM instead of a standard two-way mixed ANOVA model, an increased power to detect significant product differences can be obtained. The power gain by using the MAM was, however, found to be smaller than when using the multiplicative mixed model. Therefore we have proposed a new F-test, which increases the power of the MAM. In Brockhoff et al. (2015), a method to produce confidence intervals for product differences, based on the MAM, was suggested. This method has been evaluated, and we demonstrated that it works well, resulting in confidence intervals that behave similar to the profile likelihood based confidence intervals estimated by the mumm package. Further, the thesis gives an overview of the different biadditive (mixed) model versions, including a literature review and a description of their applications. Finally, it was demonstrated how these models can be estimated by the R package TMB (Template Model Builder).
Building school-based social capital through 'We Act - Together for Health' - a quasi-experimental study

Social capital has been found to be positively associated with various health and well-being outcomes amongst children. Less is known about how social capital may be generated and specifically in relation to children in the school setting. Drawing on the social cohesion approach and the democratic health educational methodology IVAC (Investigation - Vision - Action - Change) the aim of this study was to examine the effect of the Health Promoting School intervention 'We Act - Together for Health' on children's cognitive social capital. A quasi-experimental controlled pre- and post-intervention study design was conducted with 548 participants (mean age 11.7 years). Cognitive social capital was measured as: horizontal social capital (trust and support in pupils); vertical social capital (trust and support in teachers); and a sense of belonging in the school using questions derived from the Health Behaviour in School Children study. A series of multilevel ordinal logistic regression analyses was performed for each outcome to estimate the effect of the intervention. The analyses showed no overall significant effect from the intervention on horizontal social capital or vertical social capital at the six-month follow-up. A negative effect was found on the sense of belonging in the school. Gender and grade appeared to be important for horizontal social capital, while grade was important for sense of belonging in the school. The results are discussed in relation to We Act's implementation process, our conceptual framework and methodological issues and can be used to direct future research in the field. The study finds that child participation in health education can affect the children's sense of belonging in the school, though without sufficient management support, this may have a negative effect. With low implementation fidelity regarding the Action and Change dimension of the intervention at both the school and class level, and with measurement issues regarding the concept of social capital, more research is needed to establish a firm conclusion on the importance of the children's active participation as a source for cognitive social capital creation in the school setting. https://www.isrctn.com/ISRCTN85203017.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Research group for Risk Benefit, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of London, Steno Diabetes Centre
Number of pages: 13
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: B M C Public Health
Volume: 18
Issue number: 1
Article number: 1141
ISSN (Print): 1471-2458
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Original language: English
Keywords: Children, Health promoting school, Intervention, Multi-level logistic regression, Quasi-experimental design, School, Sense of belonging, Social capital
Electronic versions:
s12889_018_6026_0.pdf
DOI:
10.1186/s12889-018-6026-0
Source: FindIt
Source-ID: 2439624371
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

Carbon dioxide – a potentially explosive smoldering silo fire suppressant

As counter-intuitive as it may sound, there are explosion hazards associated with releasing liquid carbon dioxide (CO₂) into environments where an ignitable atmosphere may exist. For instance when attempting to suppress a smouldering fire in a pellet silo. A recent paper by Dr Frank Huss Hedlund, Risk Expert at Danish engineering consultants Cowi and External Associate Professor at the Technical University of Denmark (DTU), explains why CO₂ is a safe inert purge gas but may be unsafe for inerting.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Carbon dioxide not suitable for extinguishment of smouldering silo fires: static electricity may cause silo explosion

Smouldering fires in wood pellet silos are not uncommon. The fires are often difficult to deal with and extinguishment is a lengthy process. Injection of inert gasses to prevent oxygen from reaching the smouldering fire zone and suppress combustion is a new firefighting strategy. This article argues that injection of inert carbon dioxide into the silo headspace is unsafe. Carbon dioxide is generally available as a liquid under high pressure. When discharged, small particles of dry ice are formed. The rapid flow of particles can generate considerable amounts of static electricity, which can act as a source of ignition if ignitable pyrolysis gasses are present. This article discusses a serious wood pellet smouldering fire and silo explosion in Norway in 2010, which took place when firefighters discharged portable CO2 fire extinguishers into the headspace. The attempt to suppress the fire may have ignited pyrolysis gasses. The article examines selected guidelines, standards, popular wood pellet handbooks and other literature and argues that the electrostatic hazard is widely under-appreciated. In the past, major explosions have been attributed to electrostatic ignition of flammable vapours during the release of CO2 for fire prevention purposes. There is evidence to suggest that those early lessons learned have at least partly passed out of sight.

Comparison of different image analysis algorithms on MRI to predict physico-chemical and sensory attributes of loin

Computer vision algorithms on MRI have been presented as an alternative to destructive methods to determine the quality traits of meat products. Since, MRI is non-destructive, non-ionizing and innocuous methods. The use of fractals to analyze MRI could be another possibility for this purpose. In this paper, a new fractal algorithm is developed, to obtain features from MRI based on fractal characteristics. This algorithm is called OPFTA (One Point Fractal Texture Algorithm). Three fractal algorithms (Classical Fractal Algorithm –CFA-, Fractal Texture Algorithm –FTA- and OPFTA) and three classical texture algorithms (Grey level co-occurrence matrix –GLCM-, Grey level run length matrix –GLRLM- and Neighbouring grey level dependence matrix –NGLDM-) were tested in this study. The results obtained by means of these computer vision algorithms were correlated to the results obtained by means of physico-chemical and sensory analysis. CFA reached low relationship for the quality parameters of loins, the remaining algorithms achieved correlation coefficients higher than 0.5 noting OPFTA that reached the highest correlation coefficients in all cases except for the L* coordinate color that GLCM obtained the highest correlation coefficient. These high correlation coefficients confirm the new algorithm as an alternative to the other computer vision approaches in order to compute the physico chemical and sensory
parameters of meat products in a non-destructive and efficient way.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Extremadura, University of Copenhagen
Contributors: Caballero, D., Caro, A., Dahl, A. B., Ersbøll, B. K., Amigo, J. M., Pérez-Palacios, T., Antequera, T.
Number of pages: 10
Pages: 54-63
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Chemometrics and intelligent laboratory systems
Volume: 180
ISSN (Print): 0169-7439
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Original language: English
Keywords: Fractal, Texture features, Algorithms, Data mining, Food technology
DOIs: 10.1016/j.chemolab.2018.04.008
Source: FindIt
Source-ID: 2405741616
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

Computer Aided Identification of Motion Disturbances Related to Parkinson's Disease
We present a framework for assessing which types of simple movement tasks are most discriminative between healthy controls and Parkinson's patients. We collected movement data in a game-like environment, where we used the Microsoft Kinect sensor for tracking the user's joints. We recruited 63 individuals for the study, of whom 30 had been diagnosed with Parkinson's disease. A physician evaluated all participants on movement-related rating scales, e.g., elbow rigidity. The participants also completed the game task, moving their arms through a specific pattern. We present an innovative approach for data acquisition in a game-like environment, and we propose a novel method, sparse ordinal regression, for predicting the severity of motion disorders from the data.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen, Zealand University Hospital
Number of pages: 8
Pages: 1-8
Publication date: 2018

Host publication information
Title of host publication: International Workshop on PRedictive Intelligence In MEdicine
Publisher: Springer
ISBN (Print): 978-3-030-00319-7
(Lecture Notes in Computer Science, Vol. 11121).
Keywords: Game-aided diagnosis, Kinect, Parkinson’s disease, Sparse, Ordinal, Classification
DOIs: 10.1007/978-3-030-00320-3_1
Source: FindIt
Source-ID: 2439219911
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2018 › Research › peer-review

Computing segmentations directly from x-ray projection data via parametric deformable curves: Paper
We describe an efficient algorithm that computes a segmented reconstruction directly from x-ray projection data. Our algorithm uses a parametric curve to define the segmentation. Unlike similar approaches which are based on level-sets, our method avoids a pixel or voxel grid; hence the number of unknowns is reduced to the set of points that define the curve, and attenuation coefficients of the segments. Our current implementation uses a simple closed curve and is capable of separating one object from the background. However, our basic algorithm can be applied to an arbitrary
topology and multiple objects corresponding to different attenuation coefficients in the reconstruction. Through systematic tests we demonstrate a high robustness to the noise, and an excellent performance under a small number of projections.

**General information**
Publication status: Published
Number of pages: 16
Publication date: 2018
Peer-reviewed: Yes

**Publication information**
Journal: Measurement Science and Technology
Volume: 29
Issue number: 1
Article number: 014003
ISSN (Print): 0957-0233

**Dangers of releasing CO₂ to fight fires in the cargo hold of seagoing bulk carriers**
On seagoing general cargo vessels, the cargo is stored in bulk in the holds. Fire protection for cargo holds comprises detection and firefighting capability. Detection normally incorporates a smoke sampling system that continuously draws air from each cargo hold and passes it to a smoke detector cabinet. The fire can be fought by flooding the cargo hold with inert carbon dioxide. The carbon dioxide is stored in its liquid form at pressures in excess of 50 bar and kept in multiple vertical steel cylinders arranged in a battery. For firefighting to be effective, SOLAS regulations require that a large fraction (80-90 percent) of the gas is delivered to the space of the cargo hold over a few minutes. When discharged, the liquid carbon dioxide undergoes a change of phase to a mixture of gas and solid ("dry ice"). The sliding action of particles of dry ice can produce electrostatic discharges with sufficient energy to ignite flammable fuel/air mixtures. Because no vaporizer is present, considerable generation of static electricity is likely upon activation of the CO₂ extinguishment system and the large flow rate. Certain biological materials carried in bulk, in particular wood pellets, can self-ignite and burn as an oxygen-deficient smoldering fire. Such fires produce flammable pyrolysis gases that can travel, accumulate and enter the explosive range. If carbon dioxide is released at this stage, a gas explosion in the cargo hold may result. NFPA 12 on carbon dioxide extinguishing systems endorses the application of CO₂ to deep-seated fires involving solids subject to smoldering, but without identifying or alerting the reader to the potential presence of explosive pyrolysis gases. NFPA 12 appears to presume that electrostatic discharges will dissipate safety if metal nozzles are used and the entire system is grounded to earth. Lessons learned from past serious explosions appear to contraindicate this expectation. In conclusion, the application of carbon dioxide is excellent for extinguishing a fire with flames but unsuitable for quenching a deep-seated smoldering fire without flame. If detection of fire in a cargo hold is based merely on presence of smoke, not detection of fire with flames, an activation of the CO₂ system may lead to explosion.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, Danish Maritime Accident Investigation Board (DMAIB)
Contributors: Hedlund, F. H., Jarleivson Hilduberg, Ø.
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Event: Abstract from 8th International Symposium on Energy, Aberdeen, United Kingdom.
Keywords: Bulk carrier, Wood pellets, Fire protection, CO₂, Static electricity, Explosion
Dangers of using CO₂ to quench wood pellet silo fires

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, Industrial Fire Prevention, LLC
Contributors: Hedlund, F. H., Nichols, J.
Publication date: 2018

Publication information
Media of output: Canadian biomass
Year: 2018
Original language: English
Electronic versions:
URLs:
https://www.canadianbiomassmagazine.ca/pellets/hazards-of-using-co2-to-quench-silo-fires-6715

Bibliographical note
https://www.canadianbiomassmagazine.ca/pellets/hazards-of-using-co2-to-quench-silo-fires-6715
Last modified: 12/02/2018
Source: PublicationPreSubmission
Source-ID: 145756000
Research output: Other contribution › Net publication - Internet publication – Annual report year: 2018 › Communication

Data driven quantification of the temporal scope of building LCAs
In the construction sector, LCAs typically apply an approach based on fixed or partially fixed building lifespans/service lives/reference study period. The temporal scopes applied in building LCAs are hence typically not reflecting that the timeframes buildings can provide the service they are intended to provide, are (highly) dependent on numerous factors e.g.: building location, materials used to construct the building, energy supply and the use of the building. Inaccurate estimation of the temporal scope of a building LCA will lead to incorrect quantification of the environmental impacts of buildings. Incorrect quantification of the environmental performance of buildings may, in the worst case, derange/decelerate the development within the building sector towards more sustainable buildings. In this paper, a data set consisting of 20999 Danish buildings, demolished between 2009 and 2015, is analyzed. A multiple linear regression model is derived and used to quantify the temporal scope (often referred to as the reference study period) of building LCAs in an attempt to improve the accuracy of sustainability assessment of buildings, taking several influencing factors into account. The results obtained from the derived model are subsequently compared with several fixed/partially fixed building lifespan/service life/reference study period quantification approaches. The regression model proved to estimate the lifespan with lower errors (compared to observed values) than the prevailing approach relying on a single fixed value for all building locations, uses and building materials. The application of model based site, use, and/or material specific etc. temporal scope quantification in LCA is new and provides a mean to reduce the uncertainty of LCA results; however, the approach needs to be formalized.

General information
Publication status: Published
Organisations: Department of Civil Engineering, Section for Structural Engineering, Department of Management Engineering, Quantitative Sustainability Assessment, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Contributors: Østergaard, N., Thorsted, L., Miraglia, S., Birkved, M., Rasmussen, F. N., Birgisdottir, H., Kalbar, P., Georgiadis, S.
Pages: 224-229
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Procedia C I R P
Volume: 69
ISSN (Print): 2212-8271
Ratings:
Deformable Curves for Outlining Objects Directly From Projections

General information
Publication status: Published
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Keywords: Tomographic reconstruction, Deformable models, Segmentation, Meshing
Electronic versions: Scandem_2018_abstract_1_.pdf

Bibliographical note
Source: PublicationPreSubmission
Source-ID: 147226783
Research output: Contribution to conference abstract for conference – Annual report year: 2018 › Research › peer-review

Delayed insulin absorption correlates with alterations in subcutaneous depot kinetics in rats with diet-induced obesity: Obesity and insulin injection depot kinetics

Objective: Obesity is associated with delayed insulin absorption upon subcutaneous (s.c.) dosing in humans. The aim of this study was to investigate whether alterations in depot structure and kinetics of the s.c. injection depot could contribute to this delay. Methods: Rats fed a high-fat diet (HFD) and low-fat diet (LFD) were included in a series of insulin pharmacokinetic and imaging studies. Injection depots were visualized with μCT imaging upon s.c. dosing with insulin aspart mixed with the contrast agent iomeprol, and insulin aspart exposure was measured by means of Luminescent Oxygen Channeling Immunoassay. Results: Body weight and fat mass was increased in rats fed a HFD vs. LFD (p<0.05), whereas the lean mass was not. The HFD group exhibited delayed insulin absorption from the s.c. tissue (p<0.001). This delay was associated with smaller injection depots upon s.c. dosing (p<0.05) and correlated with a slower depot disappearance from the s.c. tissue (p<0.05) compared to the LFD group. Depot disappearance from the s.c. tissue was inversely correlated with body fat mass (p<0.05). Conclusions: Alterations in s.c. injection depot structure and kinetics may play a role in the obesity-associated delay in insulin absorption.

General information
Publication status: Accepted/In press
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen, Novo Nordisk AS
Number of pages: 19
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Obesity Science & Practice
ISSN (Print): 2055-2238
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Dimensioning a Product Development Project Portfolio Using a Closed Queueing Network

New product development processes are difficult to model as they are subject to high variability and are hard to monitor. In multi-project environments, one particular problem is to quantify the effects of changing the number of projects that run concurrently. In this paper, a closed queueing network model of the flow of projects in a product development system is presented. The model is evaluated using discrete-event simulation. Using a relation between a project's monetary value and the time it spends in the system, an optimal number of concurrent projects is determined. Application of the model and results are shown for a case study of a Danish biotechnology company. Results show, that the value generated by the system can be increased by 17.5 percent by adjusting the total number of concurrent projects and by terminating low-value projects midway. Though the model is tailored to a specific system, it contains many features relevant for new product development systems in general.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Novozymes AS
Contributors: Andersen, J. F., Lauridsen, C. L., Nielsen, B. F.
Pages: 622-626
Publication date: 2018

Distribution Fitting for Very Large Railway Delay Data Sets with Discrete Values

Modern railway signal systems allow the collection of very large data sets (more than a thousand values). These data sets are often rounded by the signal technology, so that the values are effectively discrete. This paper reviews other literature on fitting distributions to large data sets, and then shares the experience of distribution fitting to a large data set from the Danish railways.

General information
Publication status: Published
Organisations: Department of Management Engineering, Management Science, Transport DTU, Operations Management, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Ansaldo STS
Contributors: Harrod, S., Pournaras, G., Nielsen, B. F.
Number of pages: 16
Publication date: 2018
Peer-reviewed: Yes
Source: PublicationPreSubmission
Source-ID: 160573003
Research output: Contribution to conference › Paper – Annual report year: 2018 › Research › peer-review
Effect of DrugStars App Use on Life Quality and Attitudes of Diabetes Patients

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Lyndgaard, C. B.
Number of pages: 68
Publication date: 2018

Publication information
Publisher: DTU Compute
Original language: English

Bibliographical note
FORTROLIG RAPPORT. Må ikke være offentlig tilgængelig
Source: PublicationPreSubmission
Source-ID: 149066940
Research output: Book/Report → Report – Annual report year: 2018 → Research

Effects of Lifestyle on Muscle Strength in a Healthy Danish Population

Background: Life style is expected to influence muscle strength. This study aimed at assessing a possible relationship between smoking, alcohol intake and physical activity, and muscle strength in a healthy Danish population aged 20-79 years. Population study based on data collected from The Copenhagen City Heart Study (CCHS) and measurements of Isokinetic muscle strength from a sub-study of randomly selected healthy participants from CCHS.

Methods: 126 women and 63 men were studied. All participants completed a questionnaire regarding their lifestyle, including physical activity, alcohol intake and smoking habits. Isokinetic muscle strength was measured over the upper extremities (UE), trunk, and lower extremities (LE). Multivariate analyses including all of the variables were carried out.

Results: The level of daily physical activity during leisure was positively correlated to muscle strength in the lower extremities (p = 0.03) for women, and lower extremities (p = 0.03) and trunk (p = 0.007) for men. Alcohol Intake was in general not correlated to muscle strength. No clear effect of smoking was seen on muscle strength. Conclusions: Our results show that physical activity during leisure is associated with a positive effect on muscle strength in both sexes. When keeping alcohol intake within the recommended limits, alcohol does not seem to affect muscle strength negatively. No effect of smoking on muscle strength was found in our group of healthy subjects. The findings are of importance when considering recommendation on life style when wishing to keeping fit with age to be able to carry out daily activities.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Copenhagen University Hospital, University of Copenhagen, Copenhagen University Hospital Frederiksberg and Bispebjerg
Contributors: Bartels, E. M., Robertson, S., Danneskiold-Samsøe, B., Appleyard, M., Stockmarr, A.
Publication date: 2018
Peer-reviewed: Yes
Estimating functions for jump–diffusions

Asymptotic theory for approximate martingale estimating functions is generalised to diffusions with finite-activity jumps, when the sampling frequency and terminal sampling time go to infinity. Rate-optimality and efficiency are of particular concern. Under mild assumptions, it is shown that estimators of drift, diffusion, and jump parameters are consistent and asymptotically normal, as well as rate-optimal for the drift and jump parameters. Additional conditions are derived, which ensure rate-optimality for the diffusion parameter as well as efficiency for all parameters. The findings indicate a potentially fruitful direction for the further development of estimation for jump–diffusions.

Fighting pellet silo fires

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, Industrial Fire Prevention, LLC
Contributors: Hedlund, F. H., Nichols, J.
Pages: 9
Publication date: 2018
Peer-reviewed: Yes
Gene co-expression networks from RNA sequencing of dairy cattle identifies genes and pathways affecting feed efficiency

Background: Selection for feed efficiency is crucial for overall profitability and sustainability in dairy cattle production. Key regulator genes and genetic markers derived from co-expression networks underlying feed efficiency could be included in the genomic selection of the best cows. The present study identified co-expression networks associated with high and low feed efficiency and their regulator genes in Danish Holstein and Jersey cows. RNA-sequencing data from Holstein and Jersey cows with high and low residual feed intake (RFI) and treated with two diets (low and high concentrate) were used. Approximately 26 million and 25 million pair reads were mapped to bovine reference genome for Jersey and Holstein breed, respectively. Subsequently, the gene count expressions data were analysed using a Weighted Gene Co-expression Network Analysis (WGCNA) approach. Functional enrichment analysis from Ingenuity® Pathway Analysis (IPA®), ClueGO application and STRING of these modules was performed to identify relevant biological pathways and regulatory genes. Results: WGCNA identified two groups of co-expressed genes (modules) significantly associated with RFI and one module significantly associated with diet. In Holstein cows, the salmon module with module trait relationship (MTR) = 0.7 and the top upstream regulators ATP7B were involved in cholesterol biosynthesis, steroid biosynthesis, lipid biosynthesis and fatty acid metabolism. The magenta module has been significantly associated (MTR = 0.51) with the treatment diet involved in the triglyceride homeostasis. In Jersey cows, the lightsteelblue1 (MTR = −0.57) module controlled by IFNG and IL10RA was involved in the positive regulation of interferon-gamma production, lymphocyte differentiation, natural killer cell-mediated cytotoxicity and primary immunodeficiency. Conclusion: The present study provides new information on the biological functions in liver that are potentially involved in controlling feed efficiency. The hub genes and upstream regulators (ATP7b, IFNG and IL10RA) involved in these functions are potential candidate genes for the development of new biomarkers. However, the hub genes, upstream regulators and pathways involved in the co-expressed networks were different in both breeds. Hence, additional studies are required to investigate and confirm these findings prior to their use as candidate genes.
Harvest time prediction for batch processes

Batch processes usually exhibit variation in the time at which individual batches are stopped (referred to as the harvest time). Harvest time is based on the occurrence of some criterion and there may be great uncertainty as to when this criterion will be satisfied. This uncertainty increases the difficulty of scheduling downstream operations and results in fewer completed batches per day. A real case study is presented of a bacteria fermentation process. We consider the problem of predicting the harvest time of a batch in advance to reduce variation and improving batch quality. Lasso regression is used to obtain an interpretable model for predicting the harvest time at an early stage in the batch. A novel method for updating the harvest time predictions as a batch progresses is presented, based on information obtained from online alignment using dynamic time warping.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Chr. Hansen AS
Contributors: Spooner, M. P., Kold, D., Kulahci, M.
Number of pages: 10
Pages: 32-41
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Computers and Chemical Engineering
Volume: 117
ISSN (Print): 0098-1354
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Original language: English
Keywords: Batch process, Dynamic time warping, Lasso regression, Partial least squares, Prediction
DOI:
10.1016/j.compchemeng.2018.05.019
Source: FindIt
Source-ID: 2434769442
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

Identification of Expression QTLs Targeting Candidate Genes for Residual Feed Intake in Dairy Cattle Using Systems Genomics

Background: Residual feed intake (RFI) is the difference between actual and predicted feed intake and an important factor determining feed efficiency (FE). Recently, 170 candidate genes were associated with RFI, but no expression quantitative trait loci (eQTL) mapping has hitherto been performed on FE related genes in dairy cows. In this study, an integrative systems genetics approach was applied to map eQTLs in Holstein and Jersey cows fed two different diets to improve identification of candidate genes for FE. Methods: Liver RNA-seq transcriptomics data from nine Holstein and ten Jersey cows that had been fed control (C) or high concentrate (HC) diets were integrated with genomic data (from 777k BovineHD Illumina BeadChip) by using the Matrix eQTL R package. A total of 170 previously identified candidate genes for FE (89 differentially expressed genes (DEGs) between high and low RFI groups and 81 hub genes (HG) in a group of co-expressed genes) were used in the data integration analysis. Results: From the 241,542 SNPs used in the analysis, we identified 20 significant (FDR < 0.05) local-eQTLs targeting seven candidate genes and 16 significant (FDR < 0.05) local-eQTLs targeting five candidate genes related to RFI for the C and HC diet group analysis, respectively, in a breed-specific way. Conclusions: Interestingly, Holstein and Jersey cows appear to rely on different strategies (lipid and cholesterol metabolism versus immune and inflammatory function) to achieve low RFI. The eQTLs overlapped with QTLs previously associated with FE trait (e.g. dry matter intake, longevity, body weight gain and net merit). The eQTLs and biological pathways identified in this study improve our understanding of the complex biological and genetic mechanisms that determine FE traits in dairy cattle. The identified eQTLs/genetic variants can potentially be used in new genomic selection...
methods that include biological/functional information on SNPs

General information
Publication status: Published
Organisations: Department of Bio and Health Informatics, Integrative Systems Biology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen, Aarhus University
Contributors: Salleh, M., Mazzoni, G., Nielsen, M. O., Levendahl, P., Kadarmideen, H.
Number of pages: 14
Publication date: 2018
Peer-reviewed: Yes

Publication Information
Journal: Journal of Genetics and Genome Research
Volume: 5
Issue number: 1
ISSN (Print): 2378-3648
Original language: English
Keywords: eQTL, RNA-seq, Genotype, Data integration, Systems genomics, Feed efficiency, Residual feed intake
Electronic versions:
JGenetGenomicsRes_2018_5_035.pdf
DOIs:
10.23937/2378-3648/1410035
Source: FindIt
Source-ID: 2438638635
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

Incomplete understanding of biogas chemical hazards – Serious gas poisoning accident while unloading food waste at biogas plant

At a biogas plant, a truck driver was overcome by toxic fumes while unloading food waste slurry collected at sites that operate a shredder and tank system. Trucks unload their cargo into a feedstock pit. While unloading, hydrogen sulfide, carbon dioxide and possibly other toxic gases were released. Ventilation systems were overwhelmed. The driver narrowly escaped death. Several would-be rescuers collapsed. Under marginally different conditions, this accident could have led to five fatalities. The case offers a rich account of the actions of fellow workers who with stubborn determination sought to extract the unconscious driver. Even after having collapsed themselves, been extracted and recovered, they re-entered the building - amazingly still without realizing the gas danger. The article argues that long retention time in the logistics chain may alter the hazard profile of food waste slurry. This emerging risk appears overlooked and under-studied. The case exposes insufficient attention to safety in the rapidly expanding biogas sector. The biogas plant was unprepared. The investigation was quick, superficial and woefully inadequate. Root causes were not identified. Broad information sharing is limited or non-existent.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H., Madsen, M.
Pages: 13-21
Publication date: 2018
Peer-reviewed: Yes

Publication Information
Journal: Journal of Chemical Health and Safety
Volume: 25
Issue number: 6
ISSN (Print): 1074-9098
Original language: English
Electronic versions:
biogas_poisioning_accident_preprint.pdf
DOIs:
10.1016/j.jchas.2018.05.004
Source: PublicationPreSubmission
Source-ID: 148654803
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review
Investigation of a Monturaqui Impactite by Means of Bi-Modal X-ray and Neutron Tomography

X-ray and neutron tomography are applied as a bi-modal approach for the 3D characterisation of a Monturaqui impactite formed by shock metamorphism during the impact of an iron meteorite with the target rocks in the Monturaqui crater (Chile). The particular impactite exhibits structural heterogeneities on many length scales: its composition is dominated by silicate-based glassy and crystalline materials with voids and Fe/Ni-metal and oxihydroxides particles generally smaller than 1 mm in diameter. The non-destructive investigation allowed us to apply a novel bi-modal imaging approach that provides a more detailed and quantitative understanding of the structural and chemical composition compared to standard single mode imaging methods, as X-ray and neutron interaction with matter results in different attenuation coefficients with a non-linear relation. The X-ray and neutron data sets have been registered, and used for material segmentation, porosity and metallic content characterization. The bimodal data enabled the segmentation of a large number of different materials, their morphology as well as distribution in the specimen including the quantification of volume fractions. The 3D data revealed an evaporite type of material in the impactite not noticed in previous studies. The present study is exemplary in demonstrating the potential for non-destructive characterisation of key features of complex multi-phase objects such as impactites.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Image Analysis & Computer Graphics, Department of Physics, Neutrons and X-rays for Materials Physics, University of Copenhagen, Laboratoire Léon Brillouin, European Spallation Source ESS AB, Technical University of Denmark
Number of pages: 24
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Journal of Electronic Imaging
Volume: 4
Issue number: 5
Article number: 72
ISSN (Print): 1017-9909
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Keywords: Neutron imaging, X-ray imaging, Multimodal imaging, Bimodal imaging, Computed tomography, Impactite, Monturaqui
Electronic versions:
jimaging_04_00072_v2.pdf
DOIs:
10.3390/jimaging4050072

Bibliographical note
This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. (CC BY 4.0).
Source: FindIt
Source-ID: 2434671940
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

Kage årsag til iltmangel og indlæggelse på hospital
Ambulance bragte en 58-årig forretningsrejsende til akutmodtagelsen med åndenød, svimmelhed og kvalme. Denne og andre cases viser, at man skal have respekt for kuldioxid.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 12-15
Publication date: 2018
Peer-reviewed: No

Publication information
Journal: Kjemi
Volume: 78
Kuldioxid er uegnet til kvælning af glødebrande

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 16-19
Publication date: 2018
Peer-reviewed: Unknown

**Publication information**
Journal: Dansk Kemi
Volume: 99
Issue number: 1
ISSN (Print): 0011-6335
Original language: English
Electronic versions:
2018_01_DKemi_Kuldioxid_er_uegnet_til_kvældning_af_glødebrande.pdf
Source: PublicationPreSubmission
Source-ID: 143939573
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Communication

Land-use planning risk estimates for a chemical industrial park in China - A longitudinal study
A chemical industrial park (CIP) can centralize the management of companies and facilitate mutual communication between different businesses. Due to these advantages, an increasing number of chemical companies are forced into CIP, especially in developing countries such as China. Thus, the land-use planning of CIP associated with safety becomes an important issue. To illustrate the importance of the continuous risk supervision and give more experiences to other similar changing CIP, we apply a simplified quantitative risk assessment procedure to estimate the risk to a Chinese chemical industrial park (19 x 10 km²) near a rather densely populated off-site region for the years 2014 and 2017. Estimated levels of individual risk and societal risk are compared with Chinese risk acceptance criteria for land-use planning. Off-site risk levels have increased significantly from 2014 to 2017. While off-site risk levels are still low and within the acceptable region, the study concludes that the authorities should review carefully and monitor the risk level in case of future development activities around and within the chemical industrial park, e.g. preserving a buffer zone should be considered. For future chemical park design, it is highly recommended to proactively include QRA analysis especially to deal with risk in an efficient way.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, Department of Chemical and Biochemical Engineering, PROSYS - Process and Systems Engineering Centre, China University of Mining And Technology, COWI AS, China Academy of Safety Science and Technology
Pages: 124-133
Publication date: 2018
Peer-reviewed: Yes

**Publication information**
Journal: Process Safety Progress
Volume: 37
Issue number: 2
ISSN (Print): 1066-8527
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Keywords: Chemical industrial park, Major hazard potential, Quantitative risk assessment, Risk acceptance criteria, Land-use planning
Layered Surface Detection for Virtual Unrolling

We present a method for virtual unrolling of a thin rolled object. From a volumetric image of the rolled object we obtain a flat image of the object’s surface, which allows visual inspection of the object and has a number of applications. Our method exploits the geometric constrains of the problem and detects a single rolled surface. For surface detection we adapt a solution to an optimal net surface problem, previously used for terrain-like and tubular surfaces. We present our approach on an example of a rolled sheet of microelectronic, which has a layer of flexible polymer substrate and a thin metal layer lithographically coated onto the polymer. Our approach is automatic and robust. The unrolled image is undistorted, and the surface structures may be accurately quantified making our approach a good candidate for an industrial application of virtual unrolling.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Image Analysis & Computer Graphics, Department of Physics, Neutrons and X-rays for Materials Physics
Number of pages: 7
Pages: 3074-3080
Publication date: 2018

Linear, Transfinite and Weighted Method for Interpolation from Grid Lines Applied to OCT Images

When performing a line scan using optical coherence tomography (OCT), the distance between the successive scan lines is often large compared to the resolution along each scan line. If two sets of such line scans are acquired orthogonal to each other, intensity values are known along the lines of a square grid, but are unknown inside each square. To view these values as an image, intensities need to be interpolated at regularly spaced pixel positions. In this paper we evaluate three methods for interpolation from grid lines: linear, transfinite and weighted. The linear method does not preserve the known values along the grid lines. The transfinite method, known from mesh generation, preserves the known values but might cause artifacts further away from the grid lines. The weighted method, which we propose, is designed to combine the desired properties of the transfinite method close to grid lines and the stability of the linear method further away. An important parameter influencing the performance of the interpolation methods is the upsampling rate. We perform an extensive evaluation of the three interpolation methods across a range of upsampling rates. Our statistical analysis shows significant difference in the performance of the three methods. We find that the transfinite interpolation works well for small upsampling rates and the proposed weighted interpolation method performs very well for all upsampling rates typically used in practice. On the basis of these findings we propose an approach for combining two OCT scans, acquired such that the lines of the second scan are orthogonal to the first.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Lindberg, A. W., Jørgensen, T. M., Dahl, V. A.
Pages: 293-302
Publication date: 2018
Methods and tools for the statistical data analysis of large datasets collected from bio-based manufacturing processes

In bio-manufacturing, biological systems are harnessed for the production of useful organic materials to be used in, for example, the food, medicine or agricultural industries. The most common mode of production in this sector is through batch processes. In a batch process a reactor vessel is filled with raw materials such as bacteria culture, water and sugar. It is then subjected to controlled conditions for a finite duration during which its contents undergo transformation, and finally the end-product is harvested from the reactor. Typically, a variety of sensors measure conditions in the reactor throughout each batch, such as temperature, pressure and concentration. With advances in sensor technology, and computational power, the volume of data collected in this way is ever increasing. The goal of the thesis is to contribute new techniques for utilising this data to improve process understanding and product quality. The existing literature on statistical monitoring, and quality prediction, for batch processes is reviewed, highlighting the challenges presented by batch process data. These include its three dimensional structure (conventionally represented as I batches × J variables×K time-points) comprising highly multivariate, cross-correlated, autocorrelated and non-stationary variable trajectories for each batch. An aspect of the data which leads to a number of contributions in the thesis is the variation in the time dimension often present in batch processes, meaning that comparable events occur at different times in different batches, so that the shapes and features in the resulting variable trajectories are not synchronised. In addition, the overall duration of different batches in a process may vary leading to different numbers of observations, complicating the application of standard bi-linear or tri-linear methods. Dynamic time warping (DTW) has previously been applied to synchronise batch process data and address these issues. The DTW algorithm identifies an optimal warping function, which stretches and compresses each batch in order to synchronise the variable trajectories. The warping function obtained for each batch may be interpreted as the progress signature of the batch. Using a case study of a bacteria culture batch process from Chr. Hansen, the advantages of including local constraints in the DTW algorithm, so that the warping function is a more realistic representation of batch progress, are demonstrated, and a method for selecting the local constraint is presented. In another case study using data from Chr. Hansen, a novel method is developed for predicting the harvest time of a batch at an early stage, whilst the batch is in progress. The method utilises lasso regression for selection of important variables for making the prediction, and combines the prediction with the progress information contained in the warping function from online alignment with DTW. Early harvest time prediction can contribute to scheduling of down-stream resources. In a third industrial case study, lasso regression is again utilised to obtain quality predictions for batches of pectin produced by CP Kelco. The approach is contrasted with partial least squares models, and comparable estimated prediction error is obtained using lasso regression, in addition to a more parsimonious and interpretable model. Finally, the ability of DTW to quantify similarity between time series is exploited to develop a method for monitoring batch processes online to detect if a fault occurs. This method is based on the nearest neighbour principle, comparing a non going batch to its k nearest neighbours in a database of successful batches, according to the DTW distance. If the distance to the k nearest neighbours increases too quickly, an alarm is signalled to indicate that a fault has occurred. The method is demonstrated using a simulated dataset, representing batch production of penicillin, which contains a wide variety of fault types, magnitudes and onset times. The performance of the novel method is contrasted with a benchmark principle component analysis based approach, and shown to have a higher detection rate and faster detection speed when there is clustering of batches in the reference dataset.
Monthly variation in the probability of presence of adult Culicoides populations in nine European countries and the implications for targeted surveillance

BACKGROUND: Biting midges of the genus Culicoides (Diptera: Ceratopogonidae) are small hematophagous insects responsible for the transmission of bluetongue virus, Schmallenberg virus and African horse sickness virus to wild and domestic ruminants and equids. Outbreaks of these viruses have caused economic damage within the European Union. The spatio-temporal distribution of biting midges is a key factor in identifying areas with the potential for disease spread. The aim of this study was to identify and map areas of negligible adult activity for each month in an average year. Average monthly risk maps can be used as a tool when allocating resources for surveillance and control programs within Europe.

METHODS: We modelled the occurrence of C. imicola and the Obsoletus and Pulicaris ensembles using existing entomological surveillance data from Spain, France, Germany, Switzerland, Austria, Denmark, Sweden, Norway and Poland. The monthly probability of each vector species and ensembles being present in Europe based on climatic and environmental input variables was estimated with the machine learning technique Random Forest. Subsequently, the monthly probability was classified into three classes: Absence, Presence and Uncertain status. These three classes are useful for mapping areas of no risk, areas of high-risk targeted for animal movement restrictions, and areas with an uncertain status that need active entomological surveillance to determine whether or not vectors are present.

RESULTS: The distribution of Culicoides species ensembles were in agreement with their previously reported distribution in Europe. The Random Forest models were very accurate in predicting the probability of presence for C. imicola (mean AUC = 0.95), less accurate for the Obsoletus ensemble (mean AUC = 0.84), while the lowest accuracy was found for the Pulicaris ensemble (mean AUC = 0.71). The most important environmental variables in the models were related to temperature and precipitation for all three groups.

CONCLUSIONS: The duration periods with low or null adult activity can be derived from the associated monthly distribution maps, and it was also possible to identify and map areas with uncertain predictions. In the absence of ongoing vector surveillance, these maps can be used by veterinary authorities to classify areas as likely vector-free or as likely risk areas from southern Spain to northern Sweden with acceptable precision. The maps can also focus costly entomological surveillance to seasons and areas where the predictions and vector-free status remain uncertain.
Multi-Parameterised Matchmaking: A Framework

The competitive scene in online video games is becoming more and more prominent and player satisfaction is of key importance when it comes to a good user experience and a successful game. As such it is important to have efficient skill rating and matchmaking systems in order to provide a proper match experience. We propose a mathematical framework for the analysis of matchmaking systems. The mathematical model addresses the estimated skill or rating, calculation of winning probabilities based on the estimated skill, and the updating of the estimated skill upon completion of a game. We will briefly apply the framework to the ELO skill rating system. Next we will use the framework to analyse the robustness of the TrueSkill algorithm and discuss some of the findings. We have used simulated data to test the robustness of the TrueSkill algorithm. All of the data processing has been done in Python using our own code, built-in functions and Python packages. The code has primarily been used to make the simulations of matches and customise updating functions.

General information
Publication status: Published
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Technical University of Denmark
Contributors: Christiansen, A. H., Gensby, E., Nielsen, B. F.
Number of pages: 4
Publication date: 2018

Multi-phase Volume Segmentation with Tetrahedral Mesh

In life science and material science, it is often desirable to segment a volumetric data set in such a way that multiple materials (phases) are segmented and a tetrahedral mesh representation is obtained for each segment for downstream applications. Unfortunately, obtaining a mesh, typically from CT or MRI scan, is challenging, especially in 3D. This paper proposes a novel approach for volume segmentation using a tetrahedral mesh. Our method employs a deformable model that minimizes the Mumford-Shah energy function. We apply our method to several CT data sets in order to demonstrate its advantages: multi-phase support, robustness to noise, and adaptive resolution outputs. Our method is based on the Deformable Simplicial Complex (DSC) method for tracking deformable interfaces which is designed specifically to deal with topology changes.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Number of pages: 13
Publication date: 2018

Notat om fremskrivning af andelen af multisyge i Danmark frem til 2025

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Oil production monitoring and optimization from produced water analytics; a case study from the Halfdan chalk oil field, Danish North Sea

Produced water analysis is a direct source of information to the subsurface processes active in an oil field. The information is, however, complex and requires a multidisciplinary approach and access to multiple data types and sources to successfully unlock and decode the processes. We apply data analytics on a combined data set of water chemistry and oil and gas production data measured in the production stream from five wells in the Halfdan field. The field is produced applying extensive water injection to ensure the most efficient water sweep of the reservoir. Relationships between daily production data and water chemistry are examined with Principal Component Analysis (PCA), and systematics with respect to predictability of daily changes in the oil production from water chemistry are examined with partial least square (PLS) regression models. For each well, the water chemistry provides a high degree of predictability with respect to daily oil cut in the production stream. The results have potential for application within prediction of sweep efficiency, by-passed oil and for prediction of water break-through. Full potential, however, depend on successful implementation of water chemistry-oil production analytics into other data domains such as seismic (4D) data and well work-over data.

Parisian types of ruin probabilities for a class of dependent risk-reserve processes

For a rather general class of risk-reserve processes, we provide an exact method for calculating different kinds of ruin probabilities, with particular emphasis on variations over Parisian type of ruin. The risk-reserve processes under consideration have, in general, dependent phase-type distributed claim sizes and inter-arrivals times, whereas the movement between claims can either be linear or follow a Brownian motion with linear drift. For such processes, we provide explicit formulae for classical, Parisian and cumulative Parisian types of ruin (for both finite and infinite time horizons) when the clocks are phase-type distributed. An erlangization scheme provides an efficient algorithmic methods for calculating the aforementioned ruin probabilities with deterministic clocks. Special attention is drawn to the construction of specific dependency structures, and we provide a number of numerical examples to study its effect on probabilities.
Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: A smart card data analysis

Waiting time at public transport stops is perceived by passengers to be more onerous than in-vehicle time, hence it strongly influences the attractiveness and use of public transport. Transport models traditionally assume that average waiting times are half the service headway by assuming random passenger arrivals. However, research agree that two distinct passenger behaviour types exist: one group arrives randomly, whereas another group actively tries to minimise their waiting time by arriving in a timely manner at the scheduled departure time. This study proposes a general framework for estimating passenger waiting times which incorporates the arrival patterns of these two groups explicitly, namely by using a mixture distribution consisting of a uniform and a beta distribution. The framework is empirically validated using a large-scale automatic fare collection system from the Greater Copenhagen Area covering metro, suburban, and regional rail stations thereby giving a range of service headways from 2 to 60 min. It was shown that the proposed mixture distribution is superior to other distributions proposed in the literature. This can improve waiting time estimations in public transport models. The results show that even at 5-min headways 43% of passengers arrive in a timely manner to stations when timetables are available. The results bear important policy implications in terms of providing actual timetables, even at high service frequencies, in order for passengers to be able to minimise their waiting times.

General information
Publication status: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Pontificia Universidad Catolica de Chile, Technical University of Denmark
Contributors: Ingvarsson, J. B., Nielsen, O. A., Raveau, S., Nielsen, B. F.
Pages: 292-306
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Transportation Research. Part C: Emerging Technologies
Volume: 90
ISSN (Print): 0968-090X
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Original language: English
Keywords: Public policy, Smart cards, Subways, Automated fare collection, Waiting-time, Scheduling
DOIs: 10.1016/j.trc.2018.03.006
Source: FindIt
Source-ID: 2398202659
Predicting and mapping human risk of exposure to Ixodes ricinus nymphs in northern Europe using climatic and environmental data

In recent years, focus on tick-borne diseases has increased as diseases such as Lyme disease and tick-borne encephalitis have become more common and represent a health problem in many parts of Scandinavia. More effective prevention of infections requires a better understanding of the factors affecting the vector abundance as well as human exposure to the vectors. Hence, there is a great need for analyses and models that can predict how vectors and their associated diseases are distributed now and possibly in the future.

As a part of the ScandTick Innovation project, we surveyed tick nymphs at 159 sites (forests and meadows) in Denmark, southern Norway and south-eastern Sweden. At each site we measured presence/absence, and used the data obtained along with environmental data from satellite images to run Boosted Regression Tree machine learning algorithms to predict overall distribution in southern Scandinavia. Together with the predicted distribution maps, we used human density maps to identify and plot areas with high risk of exposure to ticks.

The predicted distribution and the spatial variation found corresponded well with known distributions of ticks in Scandinavia (sensitivity: 91%, specificity: 60%), and we found that the model was predominantly temperature-driven. Because presence was strongly correlated with forested habitats the risk areas were much larger in Sweden and Norway compared to Denmark. When combining these distribution maps with human population density maps, we were able to quantify the proportion of people living in areas with tick presence in Scandinavia. We found that although tick nymphs were restricted to a small proportion of the modelled area, high proportions of the human populations (67-79%) lived within these same areas. The model suggests that a potential future range expansion of I. ricinus in Scandinavia is likely but may only affect a relatively small additional proportion of the human population.

General information
Publication status: Published
Organisations: Epidemiology, National Veterinary Institute, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Bacteriology & Parasitology, Norwegian Institute of Public Health, University of Agder, Norwegian University of Life Sciences, University of Agricultural Sciences
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Event: Abstract from Workshop on Arthropod-Borne Diseases, Isle of Riems, Germany.
Electronic versions: AB_diseases2018_final.pdf
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2018 › Research › peer-review

Predicting Clinical Control from DrugStars App Medication Reviews

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Lyndgaard, C. B.
Number of pages: 18
Publication date: 2018

Publication information
Publisher: DTU Compute
Original language: English

Bibliographical note
FORTROLIG RAPPORT. Må ikke være offentlig tilgængelig
Source: PublicationPreSubmission
Source-ID: 149066974
Research output: Book/Report › Report – Annual report year: 2018 › Research

Probabilistic approach for assessing cancer risk due to benzo[a]pyrene in barbecued meat: Informing advice for population groups
Consumption of meat prepared by barbecuing is associated with risk of cancer due to formation of carcinogenic compounds including benzo[a]pyrene (BaP). Assessment of a population's risk of disease and people's individual
probability of disease given specific consumer attributes may direct food safety strategies to where impact on public health is largest. The aim of this study was to propose a model that estimates the risk of cancer caused by exposure to BaP from barbecued meat in Denmark, and to estimate the probability of developing cancer in subgroups of the population given different barbecuing frequencies. We developed probabilistic models applying two dimensional Monte Carlo simulation to take into account the variation in exposure given age and sex and in the individuals' sensitivity to develop cancer after exposure to BaP, and the uncertainty in the dose response model. We used the Danish dietary consumption survey, monitoring data of chemical concentrations, data on consumer behavior of frequency of barbecuing, and animal dose response data. We estimated an average extra lifetime risk of cancer due to BaP from barbecued meat of $6.8 \times 10^{-5}$ (95% uncertainty interval $2.6 \times 10^{-7} - 7.0 \times 10^{-4}$) in the Danish population. This corresponds to approximately one to 4,074 extra cancer cases over a lifetime, reflecting wide uncertainty. The impact per barbecuing event on the risk of cancer for men and women of low body weight was higher compared to high body weight. However, the difference due to sex and bodyweight between subgroups are dwarfed by the uncertainty. This study proposes a model that can be applied to other substances and routes of exposure, and allows for deriving the change in risk following a specific change in behaviour. The presented methodology can serve as a valuable tool for risk management, allowing for the formulation of behaviour advice targeted to specific sub-groups in the population.

**General information**
Publication status: Published
Organisations: National Food Institute, Research group for Risk Benefit, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, Research group for Genomic Epidemiology, Research group for Analytical Food Chemistry, National Institute for Public Health and the Environment, Technical University of Denmark
Number of pages: 20
Publication date: 2018
Peer-reviewed: Yes

**Publication information**
Journal: PLOS ONE
Volume: 13
Issue number: 11
Article number: e0207032
ISSN (Print): 1932-6203
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Electronic versions:
journal.pone.0207032_1_.pdf
DOIs:
10.1371/journal.pone.0207032
Source: FindIt
Source-ID: 2441446396
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

**Quantifying fibre reorientation during axial compression of a composite through time-lapse X-ray imaging and individual fibre tracking**
The sudden compressive failure of unidirectional (UD) fibre reinforced composites at loads well below their tensile strengths is a cause of practical concern. In this respect and more generally, analytical and numerical models that describe composite behaviour have been hard to verify due to a lack of experimental observation, particularly in 3D. The aim of this paper is to combine fast in-situ X-ray computed tomography (CT) with advanced image analysis to capture the changes in fibre orientation in 3D during uninterrupted progressive loading in compression of a UD glass fibre reinforced polymer (GFRP). By analysing and establishing correspondence between a sequence of time-lapse X-ray CT images of the composite, we are able for the first time to follow each fibre and quantify the progressive deflection that takes place during axial compression in the steps leading up to fibre micro-buckling and kinking. Even at just 25% of the failure load, fibres have started to tilt in approximately the direction of the ultimate kink band. The rate of tilting increases as the composite approaches the collapse load. More generally, our approach can be applied to investigate the behaviour of a wide range of fibrous materials under changing loading conditions.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Manchester
Pages: 47-54
Queueing and Optimization Models for Hospital Patient Flow

Various organizations claim that increasing attention should be put on an efficient use of healthcare resources. The internationally rising life expectancy and population size is accompanied by hospitals that are relying more on short admissions, and thus on limited bed capacity. The International World Health Report published by the World Health Organization shows that 20-40% of all health care resources are not being sufficiently utilized. Thus, tools that benefit an efficient healthcare system is greatly relevant to the present society. The goal of this thesis is to expand methods in the field of modeling and optimizing hospital patient flow with a view to provide management and planners with a range of decision tools for improving the utilization of hospital resources. We elaborate on a number of relevant hospital optimization problems which relate to decision making on both the strategic, tactical and operational level. In addition, we focus on various types of patient flow, from inpatient to acute and surgical admissions, which has led to four different research studies. Methodologically we mainly focus on evaluating the different instances of patient flow based on Markov chain modeling, and employing these models in heuristic search procedures to optimize the configuration of the related hospital resources. We employ this general approach in three studies. Additionally, the fourth study elaborates on a simulation-based Markov decision process. All four studies have been validated with patient data from Danish hospitals. The thesis consists of seven chapters which have been divided into four different parts. The first part consists of two chapters, where Chapter 1 introduces the reader to the concept of hospital patient flow, and presents the motivation for modeling and optimizing the processes that are related hereto. Next, Chapters 2 prepares the reader for the methods that have been employed in our research with particular focus on Markov chain modeling and heuristic optimization. Part II and III contain our contribution to the literature and comprise two chapters each. In Part II we focus exclusively on inpatient flow. Here, Chapter 3 presents a Markov chain model for evaluating the flow of inpatients, and a heuristic search procedure for deriving an improved distribution of the hospital's bed resources. By employing a heuristic statistical test we find that our approach adequately reflects the behavior of inpatient flow for a specific hospital case, and through additional tests that patient relocations can be reduced by 11.8% by redistributing resources that are already available to the hospital. Next, in Chapter 4 we extend the application of the Markov chain model by introducing patient preferences for room types into the optimization problem. That is, our goal is to maximize the number of patient preference-matches by changing the configuration of room types for the hospital wards. To achieve this we employ a randomized and interpolated search procedure, where solutions are sampled based on an interpolation between the currently known solutions in the search space. Numerical experiments show that this approach is able to derive near-optimal solutions usually within a 1% relative gap from the optimum. In Part III we focus on both acute and surgical patient flow. Chapter 5 presents a method for optimizing emergency department staffing by evaluating the patient flow as a Markov chain model. We employ this model in a search procedure that exploits integer linear programming to minimize the total amount of staff by simultaneously accounting for the patient waiting time. Simulation experiments indicate that our approach is fairly robust to our model assumptions, and that the solutions perform well in emergency departments with multiple triage-classes of patients. Next, in Chapter 6 we present an approach for minimizing the long-term costs related to day-to-day scheduling of surgical patients. Here, we account for the inherent rolling horizon in the problem by employing a simulation-based Markov decision process. By using data from a hospital case, we validate the approach through various simulation experiments, which indicate that distinct improvements can be achieved by employing our approach rather than performing patient scheduling manually. Finally, Part IV comprises a single chapter, namely Chapter 7, where we summarize the findings from each of our studies in a final conclusion to the thesis. In relation hereto, we provide the reader with our reflections and suggestions for future work.
Real-time fault detection and diagnosis using sparse principal component analysis

With the emergence of smart factories, large volumes of process data are collected and stored at high sampling rates for improved energy efficiency, process monitoring and sustainability. The data collected in the course of enterprise-wide operations consists of information from broadly deployed sensors and other control equipment. Interpreting such large volumes of data with limited workforce is becoming an increasingly common challenge. Principal component analysis (PCA) is a widely accepted procedure for summarizing data while minimizing information loss. It does so by finding new variables, the principal components (PCs) that are linear combinations of the original variables in the dataset. However, interpreting PCs obtained from many variables from a large dataset is often challenging, especially in the context of fault detection and diagnosis studies. Sparse principal component analysis (SPCA) is a relatively recent technique proposed for producing PCs with sparse loadings via variance-sparsity trade-off. Using SPCA, some of the loadings on PCs can be restricted to zero. In this paper, we introduce a method to select the number of non-zero loadings in each PC while using SPCA. The proposed approach considerably improves the interpretability of PCs while minimizing the loss of total variance explained. Furthermore, we compare the performance of PCA- and SPCA-based techniques for fault detection and fault diagnosis. The key features of the methodology are assessed through a synthetic example and a comparative study of the benchmark Tennessee Eastman process.
Silo response and safety: The dangers of using carbon dioxide to quench silo fires

The dangers of using carbon dioxide to quench silo fires.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, Industrial Fire Prevention, LLC
Contributors: Hedlund, F. H., Nichols, J.
Pages: 50-54
Publication date: 2018
Peer-reviewed: No

Silo response - The dangers of using carbon dioxide to quench silo fires

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, Industrial Fire Prevention, LLC
Contributors: Hedlund, F. H., Nichols, J.
Pages: 16-17
Publication date: 2018
Peer-reviewed: Unknown

Simulated physiological oocyte maturation has side effects on bovine oocytes and embryos

Oocyte maturation is a complex process involving nuclear and cytoplasmic modulations, during which oocytes acquire their ability to become fertilized and support embryonic development. The oocyte is apparently "primed" for maturation during its development in the dominant follicle. As bovine oocytes immediately resume meiosis when cultured, it was hypothesized that delaying resumption of meiosis with cyclic nucleotide modulators before in vitro maturation (IVM) would allow the oocytes to acquire improved developmental competence. We tested the Simulated Physiological Oocyte
Maturation (SPOM) system that uses forskolin and 3-isobutyl-1-methylxanthine for 2 h prior to IVM against two different systems of conventional IVM (Con-IVM). We evaluated the ultrastructure of matured oocytes and blastocysts and also assessed the expression of 96 genes related to embryo quality in the blastocysts. In summary, the SPOM system resulted in lower blastocyst rates than both Con-IVM systems (30 ± 9.1 vs. 35 ± 8.7; 29 ± 2.6 vs. 38 ± 2.8). Mature SPOM oocytes had significantly increased volume and number of vesicles, reduced volume and surface density of large smooth endoplasmic reticulum clusters, and lower number of mitochondria than Con-IVM oocytes. SPOM blastocysts showed only subtle differences with parallel undulations of adjacent trophoderm plasma membranes and peripherally localized ribosomes in cells of the inner cell mass compared with Con-IVM blastocysts. SPOM blastocysts, however, displayed significant downregulation of genes related to embryonic developmental potential when compared to Con-IVM blastocysts. Our results show that the use of the current version of the SPOM system may have adverse effects on oocytes and blastocysts calling for optimized protocols for improving oocyte competence.

General information
Publication status: Published
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, University of Copenhagen, University of São Paulo, Aarhus University, EmbryoTrans Biotech
Pages: 413–424
Publication date: 2018
Peer-reviewed: Yes

Publication Information
Journal: Journal of Assisted Reproduction and Genetics
Volume: 36
Issue number: 3
ISSN (Print): 1058-0468
Ratings:
BFI (2018): BFI-level 1
Original language: English
Keywords: Bovine blastocyst, Gene expression, In vitro maturation, SPOM, Ultrastructure
DOIs: 10.1007/s10815-018-1365-4
Source: FindIt
Source-ID: 2441596658
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

Sparse Classification - Methods & Applications
With increasing number of more sophisticated tools to acquire data, we are faced with the important question of what matters in the sea of information at hand. This challenge is becoming more prevalent across virtually all scientific disciplines. Improvements over state of the art methods for analysing such data carry the potential to revolutionize tasks such as medical diagnostics where often decisions need to be based on only a few high-dimensional observations. This explosion in data dimensionality has sparked the development of novel statistical methods. In contrast, classical statistics build upon the assumption that we have more samples than variables, and the main asymptotic results, such as the central limit theorem, reflect that. As the assumption of having many samples does not hold for modern datasets, we need new tools and methods to find the signal within the dataset which is predictive of the relevant response variable. The focus in this thesis is on sparse methods where sparse implies that the method selects only a few variables. Different types of data call for different methods. In this thesis the sparse methods we study concern settings where the response variable is ordinal. Such ordinal labeling is common in many fields, for example, medical doctors often summarize their observations into a single class of disease severity, which is known as a medical rating score. Automation offers the potential to improve both the reliability and objectivity of such tasks. To demonstrate the effectiveness of the sparse methods developed in this thesis, they were applied to both challenging and diverse real-world problems: Predicting the severity of motion disorders from Parkinson’s patients, generating short summaries of content from hundreds of online user reviews and detecting foreign objects from Multispectral X-ray scans. It may be noted, that to achieve these results, novel optimization approaches and open-source software were implemented.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Einarsson, G.
Number of pages: 283
Publication date: 2018

Publication information
SpaSM: A MATLAB Toolbox for Sparse Statistical Modeling

Applications in biotechnology such as gene expression analysis and image processing have led to a tremendous development of statistical methods with emphasis on reliable solutions to severely underdetermined systems. Furthermore, interpretations of such solutions are of importance, meaning that the surplus of inputs has been reduced to a concise model. At the core of this development are methods which augment the standard linear models for regression, classification and decomposition such that sparse solutions are obtained. This toolbox aims at making public available carefully implemented and well-tested variants of the most popular of such methods for the MATLAB programming environment. These methods consist of easy-to-read yet efficient implementations of various coefficient-path following algorithms and implementations of sparse principal component analysis and sparse discriminant analysis which are not available in MATLAB. The toolbox builds on code made public in 2005 and which has since been used in several studies.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Rector's office, Image Analysis & Computer Graphics, EXINI Diagnostics AB
Contributors: Sjöstrand, K., Clemmensen, L. H., Larsen, R., Einarsson, G., Ersbøll, B. K.
Number of pages: 37
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Journal of Statistical Software
Volume: 84
Issue number: 10
ISSN (Print): 1548-7660
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Original language: English
Keywords: Least angle regression-elastic net, LASSO, Elastic net, Sparse principal component analysis, Sparse discriminant analysis, MATLAB

Spatial and temporal variation in the abundance of Culicoides biting midges (Diptera: Ceratopogonidae) in nine European countries

Biting midges of the genus Culicoides (Diptera: Ceratopogonidae) are vectors of bluetongue virus (BTV), African horse sickness virus and Schmallenberg virus (SBV). Outbreaks of both BTV and SBV have affected large parts of Europe. The spread of these diseases depends largely on vector distribution and abundance. The aim of this analysis was to identify and quantify major spatial patterns and temporal trends in the distribution and seasonal variation of observed Culicoides abundance in nine countries in Europe. We gathered existing Culicoides data from Spain, France, Germany, Switzerland, Austria, Denmark, Sweden, Norway and Poland. In total, 31,429 Culicoides trap collections were available from 904 ruminant farms across these countries between 2007 and 2013. The Obsoletus ensemble was distributed widely in Europe and accounted for 83% of all 8,842,998 Culicoides specimens in the dataset, with the highest mean monthly abundance recorded in France, Germany and southern Norway. The Pulicaris ensemble accounted for only 12% of the specimens and had a relatively southerly and easterly spatial distribution compared to the Obsoletus ensemble. Culicoides imicola Kieffer was only found in Spain and the southernmost part of France. There was a clear spatial trend in the accumulated annual abundance from southern to northern Europe, with the Obsoletus ensemble steadily increasing from 4000 per year in southern Europe to 500,000 in Scandinavia. The Pulicaris ensemble showed a very different pattern, with an increase in the accumulated annual abundance from 1600 in Spain, peaking at 41,000 in northern Germany and then decreasing again toward northern latitudes. For the two species ensembles and C. imicola, the season began between January and April, with later start dates and increasingly shorter vector seasons at more northerly latitudes. We present the first maps of seasonal Culicoides abundance in large parts of Europe covering a gradient from southern Spain to
northern Scandinavia. The identified temporal trends and spatial patterns are useful for planning the allocation of resources for international prevention and surveillance programmes in the European Union.

General information
Publication status: Published
Organisations: National Veterinary Institute, Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Oldenburg, CIRAD, Université de Strasbourg, EID Méditerranée, University of the Balearic Islands, University of Zaragoza, University of Zurich, Avia-GIS NV, Aarhus University, Roskilde University, National Veterinary Institute, Bernhard Nocht Institute for Tropical Medicine, Friedrich-Loeffler-Institute, National Veterinary Research Institute, National Veterinary Institute, Institute for Veterinary Public Health
Number of pages: 18
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Parasites & Vectors
Volume: 11
Article number: 112
ISSN (Print): 1756-3305
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Keywords: Culicoides abundance, Culicoides distribution, Europe, Seasonal abundance, Spatial pattern, Temporal trend, Vector season, Vector-borne disease
Electronic versions:
Cuellar_et_al_2018_Parasites_Vectors.pdf
DOIs:
10.1186/s13071-018-2706-y

Bibliographical note
This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Source: FindIt
Source-ID: 2396754311
Research output: Contribution to journal > Journal article – Annual report year: 2018 > Research > peer-review

Statistical Learning with Applications in Biology
Statistical methods are often motivated by real problems. We consider methods inspired by problems in biology and medicine. The thesis is in two parts.

In the first part we consider data in the form of graphs (or networks). These occur naturally in many contexts such as social and biological networks. We specifically consider the setting where we have multiple graphs on the same set of nodes. We propose a model in this setting called the multiple random dot product graph model. Fitting the model is an optimization problem which we solve efficiently using a new alternating minimization algorithm. A hypothesis test in the model framework for whether two graphs are drawn from the same distribution is also proposed. Both the fitting algorithm and test are evaluated in simulation studies. The model is also generalized to weighted graphs where we specifically consider Poisson and normally distributed weights. Similar hypothesis tests are proposed in these settings and again we evaluate the performance through simulation studies.

The second part of the thesis considers prediction of disease progression. We compare three common approaches for disease prediction and apply them to a diabetes data set. In this data, the time until a patient goes on to insulin treatment is of interest - especially whether progression is fast or slow. The methods are: A Cox proportional hazards model, a random forest method for survival data, and a neural network approach. The prediction performance, and the pros and cons of the methods are discussed.

General information
Publication status: Submitted
Statistical validation of individual fibre segmentation from tomograms and microscopy

Imaging with X-ray computed tomography (CT) enables non-destructive 3D characterisations of the micro-structure inside fibre composites. In this paper we validate the use of X-ray CT coupled with image analysis for characterising unidirectional (UD) fibre composites. We compare X-ray CT at different resolutions to optical microscopy (OM) and scanning electron microscopy (SEM), where we characterise fibres by their diameters and positions. In addition to comparing individual fibre diameters, we also model their spatial distribution, and compare the obtained model parameters. Our study shows that X-ray CT is a high precision technique for characterising fibre composites and, with our suggested image analysis method for fibre detection, high precision is also obtained at low resolutions. This has great potential, since it allows larger fields of view to be analysed. Besides analysing representative volumes with high precision, we demonstrate that based on our methodology for individual fibre segmentation it is now possible to study complete bundles at the fibre scale and reveal inhomogeneities in the physical sample.

Stort udslip af giftig gas ved aflæsning af madaffald


Stort udslip af giftig gas ved aflæsning af madaffald

Stort udslip af giftig gas ved aflæsning af madaffald (Major release of toxic gas while unloading food waste at biogas plant)


General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 36-40
Publication date: 2018
Peer-reviewed: Unknown

Publication information
Journal: Dansk Kemi
Volume: 99
Issue number: 5
ISSN (Print): 0011-6335
Original language: English
Electronic versions:
2018_DK_Stort_udslip_giftig_gas_madaffald_biogasanl_g_Dansk_Kemi_august_2018_nr_5.pdf
Source: PublicationPreSubmission
Source-ID: 152112401
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Communication

Stort utsläpp av giftig gas vid avlastning av matavfall

Vid avlastning av matavfall i en biogasanläggning fylldes byggnaden bl a med giftigt svavelväte. Under lite annorlunda omständigheter hade fem personer kunnat omkomma. Exakt vilka gaser som frigjordes är okänt liksom orsaken till gasutvecklingen.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 30-33
Publication date: 2018
Peer-reviewed: No

Publication information
Journal: Kemivärlden Biotech med Kemisk Tidsskrift
Volume: 2018
Issue number: 7
ISSN (Print): 1650-0725
Original language: English
Electronic versions:
2018_11_KB_Stort_utslipp_giftig_gas_vid_avlastning_av_matavfall.pdf
Source: PublicationPreSubmission
Source-ID: 161298851
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research
Svovlbrinte stinker, men dræber uden lugt

Svovlbrinte stinker, men dræber uden lugt
Den stank af rådne æg, som svovlbrinte har, forsvinder ved store mængder. Fra Danmark til Kina har man eksempler på, at mennesker er døde i gasulykker, helt uden at have 'lugtet lunten'.

Svovlbrinte, men dræber uden lugt
Svovlsyres fortyndingsvarme
Det er god latin, at syrer fortyndes ved at hælde syren i vand, aldrig omvendt. Der udvikles betydelige mængder varme, og kogning kan slyngø syre og vand ud af beholderen. Selvom svovlsyre anvendes i enorme mængder, er det overraskende vanskeligt at finde data for fortyndingsvarmen. Det rådes der bod på med denne artikel.

Svovlysyre fortyndingsvarme

[English summary] Although sulfuric acid is one of the most fundamentally important heavy industrial chemicals and produced in very large quantities, there is little practical guidance in open sources on how to estimate the heat of dilution. This article provides a simple method to estimate the increase in temperature upon diluting sulfuric acid with water. simple method to estimate the increase in temperature upon diluting sulfuric acid with water.
The Influence of Volume and Anatomic Location of Optic Disc Drusen on the Sensitivity of Autofluorescence

Optic disc drusen (ODD) are acellular deposits in the optic nerve head. ODD can be diagnosed using different imaging modalities, including enhanced depth imaging optical coherence tomography (EDI-OCT) and autofluorescence (AF). It is unknown which factors determine the sensitivity of AF. The aim of this study was to investigate the effect of volume and anatomic location of ODD on the sensitivity of AF. Cross-sectional study. A total of 38 patients (75 eyes) with ODD were included. In 12 of 75 eyes (16%) and in 11 of 38 patients (29%), EDI-OCT detected ODD that were not detected by AF. In 24 distinctly solitary ODD, both increase in ODD volume (P = 0.0388) and a more superficial ODD location (P <0.0001)
increased the possibility of AF detection of ODD, when performing a multivariate analysis. EDI-OCT is superior to AF in the diagnosis of ODD. Volume and anatomic location of ODD have a significant impact on the sensitivity of AF.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen
Contributors: Loft, F. C., Malmqvist, L., Lindberg, A. W., Hamann, S.
Number of pages: 5
Publication date: 2018
Peer-reviewed: Yes

**Publication information**
Journal: Journal of Neuro-Ophthalmology
ISSN (Print): 1070-8022
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
DOIs:
10.1097/WNO.0000000000000654
Source: FindIt
Source-ID: 2411801732
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review

**The Mixed Assessor Model and the multiplicative mixed model**
A novel possibility for easy and open source based analysis of sensory profile data by a formal multiplicative mixed model (mumm) with fixed product effects and random assessor effects is presented by means of the generic statistical R-package mumm. The package is using likelihood principles and is utilizing newer developments within Automatic Differentiation by means of the Template Model Builder R-package. We compare such formal likelihood based analysis with the Mixed Assessor Model (MAM) analysis, where MAM is a linear approximation of the multiplicative mixed model. We use real sensory data as examples together with simulated data. We found that the formal mumm approach for hypothesis testing more resembles the MAM than the standard 2-way mixed model, and that both the mumm approach and the MAM give a higher power to detect product differences than the 2-way mixed model, when a “scaling effect” is present. We also validated that the novel contrast confidence limit method suggested previously for the MAM performs well and in line with the formal likelihood based confidence intervals of the mumm. Finally, the likelihood based mumm approach suggests that the more proper test for product difference would be a test that has a “joint product and scaling effect” interpretation.

**General information**
Publication status: Published
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, National Institute of Aquatic Resources, Section for Marine Living Resources, Technical University of Denmark
Contributors: Pødenphant, S., Truong, M. H., Kristensen, K., Brockhoff, P. B.
Pages: 38-48
Publication date: 2018
Peer-reviewed: Yes

**Publication information**
Journal: Food Quality and Preference
Volume: 74
ISSN (Print): 0950-3293
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Original language: English
Keywords: Sensory profile data, Analysis of variance, Multiplicative mixed model, Scaling differences, Disagreement, Template Model Builder
DOIs:
10.1016/j.foodqual.2018.11.006
Source: FindIt
Source-ID: 2441445001
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research › peer-review
The transmission spectrum of sound through a phononic crystal subjected to liquid flow

The influence of liquid-flow up to 7 mm/s is examined on transmission spectra of phononic crystals, revealing a potential use for slow liquid-flow measurement techniques. It is known that transmission of ultrasound through a phononic crystal is determined by its periodicity and depends on the material characteristics of the crystal's constituents. Here, the crystal consists of metal rods with the space in between filled with water. Previous studies have assumed still water in the crystal, and here, we consider flowing liquid. First, the crystal bandgaps are investigated in still water, and the results of transmission experiments are compared with theoretical band structures obtained with the finite element method. Then, changes in transmission spectra are investigated for different speeds of liquid flow. Two situations are investigated: a crystal is placed with a principal symmetry axis in the flow direction (ΓX) and then at an angle (ΓM). The good stability of the bandgap structure of the transmission spectrum for both directions is observed, which may be of importance for the application of phononic crystals as acoustic filters in an environment of flowing liquid. Minor transmission amplitude changes on the other hand reveal a possibility for slow liquid flow measurements.

Phononic crystals (PCs) are generally formed by a periodic arrangement of materials (scatterers) with elastic properties different from those of the homogeneous matrix in between the scatterers, typically scaled at the wavelengths of interest and giving rise to the emergence of transmission bandgaps. The concept was studied by Yablonovitch1 in optics for a photonic crystal in the ultraviolet microwave regime, where he shows that bandgaps in the spectrum exist as a result of interferences between direct and reverberated paths of waves. A similar behavior of acoustic waves in phononic crystals (PCs) has been observed. Additionally, ultrasonic waves in a periodic structure are used for sensing purposes, such as acoustic waveguides and acoustic lenses, to control, direct, and manipulate sound.2,3 The reported experiments are as follows: ultrasound is emitted by a transducer, and it travels through the PC, thereby probing its acoustic properties (density, viscosity, speed of sound, ..., speed of water flow). A specific transmission spectrum, including bandgaps, emerges, and its specific characteristics are determined by the physical properties of the PC. Over the last decade, PCs have been introduced as a platform for (still) liquid sensing purposes,4-9 based on significant spectral changes induced by composition changes of the liquid mixture.10 Many works discuss the application of PCs for fluid characterization such as viscosity, density, and concentration measurement of liquid solutions. However, no study of possible flow-speed influence on PC filter characteristics has been reported. For the case in which fluid-flow measurements without the presence of a PC is considered, we can cite, for example, Nennimura et al.11 for measuring the small open channel fluid flow using pulse-echo signals scattered from the particles in a pipe. From the slope of the correlation peak amplitude with the variation in pulse-echo excitation time, the authors estimate the flow-speed of the medium, for speeds much higher than what is studied in the current paper. Here, we study the band structure and its stability and explore minor effects in actual transmission amplitudes to the flow-speed. The low speeds involved are comparable to what one may expect on a large scale in tidal water currents for example. The phononic crystal under study consists of a square lattice arrangement of 169 steel rods, each having a diameter of 1.2 mm and a length of 150 mm. A photography of the crystal is shown in Fig. 1(a). The rods were aligned using two supporting plates that had been machined to have periodic arrays of holes, and Fig. 1(b) shows the square lattice pattern of the cylinders and the directions of the highest symmetry, referred to as ΓX and ΓM. The lattice constant, being the distance between the centers of any two adjacent rods, was measured: a = 2.52 mm. The crystal made of cylinders is submerged in water, such that the water in between the cylinders acts as the crystal matrix. Assuming a sound speed in water of 1480 m/s, incident ultrasound with a wavelength corresponding to the lattice constant would have a frequency on the order of 1 MHz. Steel (rods) and water (host medium) were chosen here as the constituent materials of the crystal due to the large contrast in their densities and elastic constants, as this has been shown to be an effective approach for the formation of bandgaps in other studies on phononic crystals.12–14 To study effects of liquid flow on the transmission spectrum, that spectrum was first determined using through-transmission experiments using an emitting and a receiving transducer, namely, two Valpey-Fisher ISO104GP transducers with a nominal center frequency of and a beamwidth of approximately 10 mm. Two types of experiments have been performed on the crystal: through-transmission measurements in the ΓX direction and in the ΓM direction; the results are shown in Fig. 2.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Georgia Institute of Technology
Contributors: Declercq, N. F., Chehami, L., Moiseyenko, R. P.
Number of pages: 5
Publication date: 2018
Peer-reviewed: Yes

Publication information
Volume: 112
Issue number: 2
Article number: 024102
ISSN (Print): 0003-6951
Ratings: BFI (2018): BFI-level 2
Uncertainty Management and Sensitivity Analysis

Uncertainty is always there and LCA is no exception to that. The presence of uncertainties of different types and from numerous sources in LCA results is a fact, but managing them allows to quantify and improve the precision of a study and the robustness of its conclusions. LCA practice sometimes suffers from an imbalanced perception of uncertainties, justifying modelling choices and omissions. Identifying prevalent misconceptions around uncertainties in LCA is a central goal of this chapter, aiming to establish a positive approach focusing on the advantages of uncertainty management. The main objectives of this chapter are to learn how to deal with uncertainty in the context of LCA, how to quantify it, interpret and use it, and how to communicate it. The subject is approached more holistically than just focusing on relevant statistical methods or purely mathematical aspects. This chapter is neither a precise statistical method description, nor a philosophical essay about the concepts of uncertainty, knowledge and truth, although you will find a little bit of both. This chapter contains (1) an introduction of the essential terminology and concepts of relevance for LCA; (2) a discussion of main sources of uncertainty and how to quantify them; (3) a presentation of approaches to calculate uncertainty for the final results (propagation); (4) a discussion of how to use uncertainty information and how to take it into account in the interpretation of the results; and finally (5) a discussion of how to manage, communicate and present uncertainty information together with the LCA results.

Understanding UD Fibre-reinforced Polymers through X-ray Imaging and Individual Fibre Tracking

X-ray computed tomography (CT) is a powerful tool for characterising materials for its ability to reveal their internal structure in a non-destructive manner. The recent advances in X-ray imaging have brought high-resolution X-rays to laboratory sources, making this tool available to a broader public. Additionally, thanks to the developments in ultra-fast X-ray imaging at synchrotron beamlines, it is now possible to capture the very fast structural changes inside materials under realistic working conditions, e.g. in operation or under loading. There is a need for advanced image analysis methods that can exploit the information contained in these 3D and 4D data-sets of high spatial and temporal resolution, which often contain image artefacts and noise. We have developed a method to characterise the geometry of materials reinforced with long fibres [1], such as glass and carbon fibre reinforced polymers. The method is based on segmenting individual fibres and the task is specially challenging when the image is noisy and its resolution is limited, because the fibres are densely packed. A limited spatial resolution might arise from the need of performing fast scans, to capture the sudden micro-structural changes that happen when reaching the composite's collapse load, and will facilitate scanning large fields of view containing many fibres, necessary to ensure representative characterisations of a material's micro-structure. Due to the robustness of our method to image quality [2], we have been able to characterise fibre orientations and diameter distributions in complete bundles, relevant for investigating the effect of the design and manufacturing processes on the mechanical properties of the materials. Moreover, we have applied our methodology to study the behaviour of a fibre composite under compressive loading. Following the changes in each individual fibre under progressive loading conditions, and correlating these with the initial structure of the material, can reveal the precursors to the very complex
damage mechanisms that affect fibre composites.

General information
Publication status: Published
Organisations: Image Analysis & Computer Graphics, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Wind Energy, Waseda University, University of Manchester
Publication date: 2018
Peer-reviewed: Yes
Electronic versions: Scandem_2018_abstract_monj.pdf
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2018 › Research

Un tanque de levadura excedente falla con consecuencias catastróficas
Un gran tanque de levadura excedente voló por los aires, y solo quedó la placa de base y el contenido del tanque. Aunque el tanque no había sido concebido para sobrepresión, se mantenía a "una sobrepresión muy ligera" para evitar la formación de espuma no deseada. La cervecería no conocía los riesgos relacionados con el aire comprimido.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, COWI AS
Contributors: Hedlund, F. H., Selig, R. S.
Pages: 76-79
Publication date: 2018
Peer-reviewed: No

Publication information
Journal: BRAUWELT En Español
Volume: 22
Issue number: May
ISSN (Print): 1619-537X
Original language: Spanish
Source: PublicationPreSubmission
Source-ID: 148654698
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research

Using game design for learning

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Weitze, C. L.
Pages: 54-63
Publication date: 2018

Host publication information
Title of host publication: Enhancing Learning and Teaching with Technology: What the Research Says
Editor: Luckin, R.
Source: PublicationPreSubmission
Source-ID: 147732471
Research output: Chapter in Book/Report/Conference proceeding › Book chapter – Annual report year: 2018 › Research

Very short-term spatio-temporal wind power prediction using a censored Gaussian field
Wind power is a renewable energy resource, that has relatively cheap installation costs and it is highly possible that will become the main energy resource in the near future. Wind power needs to be integrated efficiently into electricity grids, and to optimize the power dispatch, techniques to predict the level of wind power and the associated variability are critical. Ideally, one would like to obtain reliable probability density forecasts for the wind power distributions. We aim at contributing to the literature of wind power prediction by developing and analysing a spatio-temporal methodology for wind power production, that is tested on wind power data from Denmark. We use anisotropic spatio-temporal correlation models
to account for the propagation of weather fronts, and a transformed latent Gaussian field model to accommodate the probability masses that occur in wind power distribution due to chains of zeros. We apply the model to generate multi-step ahead probability predictions for wind power generated at both locations where wind farms already exist but also to nearby locations.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Cyprus
Contributors: Baxevani, A., Lenzi, A.
Pages: 931-948
Publication date: 2018
Peer-reviewed: Yes

**Publication information**
Journal: Stochastic Environmental Research and Risk Assessment
Volume: 32
Issue number: 4
ISSN (Print): 1436-3240
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
Original language: English
Keywords: Covariance function, Gaussian transformed model, Kriging equations, Spatio-temporal model, Wind power, Electric power generation, Energy resources, Forecasting, Gaussian distribution, Probability, Probability density function, Probability distributions, Renewable energy resources, Weather forecasting, Gaussians, Kriging, Probability densities, Spatio-temporal models, Spatiotemporal correlation, Wind power distribution, Wind power predictions, Electric load dispatching, Environment, Earth Sciences, general, Probability Theory and Stochastic Processes, Statistics for Engineering, Physics, Computer Science, Chemistry and Earth Sciences, Computational Intelligence, Waste Water Technology / Water Pollution Control / Water Management / Aquatic Pollution
DOIs: 10.1007/s00477-017-1435-7
Source: FindIt
Source-ID: 2396956706
Research output: Contribution to journal > Journal article – Annual report year: 2018 > Research > peer-review

**Analysis of aggregated functional data from mixed populations with application to energy consumption**
Understanding energy consumption patterns of different types of consumers is essential in any planning of energy distribution. However, obtaining individual-level consumption information is often either not possible or too expensive. Therefore, we consider data from aggregations of energy use, that is, from sums of individuals’ energy use, where each individual falls into one of C consumer classes. Unfortunately, the exact number of individuals of each class may be unknown due to inaccuracies in consumer registration or irregularities in consumption patterns. We develop a methodology to estimate both the expected energy use of each class as a function of time and the true number of consumers in each class. To accomplish this, we use B-splines to model both the expected consumption and the individual-level random effects. We treat the reported numbers of consumers in each category as random variables with distribution depending on the true number of consumers in each class and on the probabilities of a consumer in one class reporting as another class. We obtain maximum likelihood estimates of all parameters via a maximization algorithm. We introduce a special numerical trick for calculating the maximum likelihood estimates of the true number of consumers in each class. We apply our method to a data set and study our method via simulation.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of British Columbia, University of Campinas
Contributors: Lenzi, A., de Souza, C. P. E., Dias, R., Garcia, N. L., Heckman, N. E.
Number of pages: 34
Publication date: 2017
Peer-reviewed: Yes

**Publication information**
Journal: Environmetrics
Volume: 28
Issue number: 2
Article number: e2414
ISSN (Print): 1180-4009
Ratings:
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 1.36 SJR 1.014 SNIP 0.875
Web of Science (2017): Impact factor 1.321
Web of Science (2017): Indexed yes
Original language: English
DOIs:
10.1002/env.2414
URLs:
Source: FindIt
Source-ID: 2348722959
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Research › peer-review

Analysis of the Data Using the R Package sensR

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science , Statistics and Data Analysis
Contributors: Brockhoff, P. B., Linander, C. B.
Pages: 303-344
Publication date: 2017

Host publication information
Title of host publication: Discrimination Testing in Sensory Science
Publisher: Woodhead Publishing
(Woodhead Publishing Series in Food Science, Technology and Nutrition).
Source: PublicationPreSubmission
Source-ID: 140538643
Research output: Chapter in Book/Report/Conference proceeding › Book chapter – Annual report year: 2017 › Research › peer-review

Applying LCA in decision making- the need and the future perspective

General information
Publication status: Published
Organisations: Department of Management Engineering, Quantitative Sustainability Assessment, Department of Civil Engineering, Centre for oil and gas – DTU, Transport DTU, Transport Modelling, Department of Applied Mathematics and Computer Science , Statistics and Data Analysis, Department of Environmental Engineering, Urban Water Systems, National Food Institute, Research group for Genomic Epidemiology, Section for Structural Engineering
Contributors: Dong, Y., Miraglia, S., Manzo, S., Georgiadis, S., Sørup, H. J. D., Boriani, E., Hald, T., Thöns, S., Hauschild, M. Z.
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Electronic versions:
AbstrApplying_LCA_in_policy_deciison_making_Final
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2017 › Research › peer-review

A probabilistic approach to urban flooding from sea surges in Copenhagen

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science , Statistics and Data Analysis, Department of Environmental Engineering, Urban Water Systems
Contributors: Georgiadis, S., Sørup, H. J. D., Nielsen, B. F., Arnbjerg-Nielsen, K.
Pages: 658-662
Publication date: 2017
A Probabilistic Framework for Curve Evolution

In this work, we propose a nonparametric probabilistic framework for image segmentation using deformable models. We estimate an underlying probability distribution of image features from regions defined by a deformable curve. We then evolve the curve such that the distance between the distributions is increasing. The resulting active contour resembles a well studied piecewise constant Mumford-Shah model, but in a probabilistic setting. An important property of our framework is that it does not require a particular type of distributions in different image regions. Additional advantages of our approach include ability to handle textured images, simple generalization to multiple regions, and efficiency in computation. We test our probabilistic framework in combination with parametric (snakes) and geometric (level-sets) curves. The experimental results on composed and natural images demonstrate excellent properties of our framework.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Image Analysis & Computer Graphics
Pages: 421-32
Publication date: 2017

A procedure for grouping food consumption data for use in food allergen risk assessment

Food allergic subjects need to avoid the allergenic food that triggers their allergy. However, foods can also contain unintended allergens. Food manufacturers or authorities need to perform a risk assessment to be able to decide if unintended allergen presence constitutes a risk to food allergic consumers. One of the input parameters in risk assessment is the amount of a given food consumed in a meal. There has been little emphasis on how food consumption data can be used in food allergen risk assessment. To do this, a two-step method was developed. First, based on initial groups of similar food items, the homogeneity of consumption was evaluated using a customized clustering method. Then, the risk was calculated for each initial food group and its subgroups to verify if it also represents a relevant difference in risk. Forty-eight food groups were designated in Denmark (53 in the Netherlands, 54 in France). Finally, summary statistics and names for each food group for the Danish data illustrate the results when applying the procedure.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Research Group for Gut Microbiology and Immunology, Division of Risk Assessment and Nutrition, Netherlands Organisation for Applied Scientific Research - TNO, ANSES - French Agency for Food, Environmental and Occupational Health & Safety
Contributors: Birot, S., Madsen, C. B., Kruizinga, A. G., Christensen, T., Crépet, A., Brockhoff, P. B.
Pages: 111-123
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Journal of Food Composition and Analysis
Är sprängplattan vänd åt rätt håll?
Sprängplattor används för att skydda mot övertryck. Hur tillförlitliga de är beror helt och hållet på om sprängplattan är vänd åt rätt håll.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 22-23
Publication date: 2017
Peer-reviewed: No

Publication information
Journal: Kemivärlden Biotech med Kemisk Tidsskrift
Issue number: 3
ISSN (Print): 1650-0725
Original language: Swedish
Electronic versions:
2017_05_KB_Ar_sprangplattan_vand_at_ratt_hall.pdf
Source: PublicationPreSubmission
Source-ID: 152013174
Research output: Contribution to journal › Journal article – Annual report year: 2018 › Research

A Simulation-based Markov Decision Process for the Scheduling of Operating Theatres

General information
Publication status: Published
Organisations: Department of Management Engineering, Management Science, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Contributors: Andersen, A. R., Nielsen, B. F., Stidsen, T. J. R., Reinhardt, L. B.
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from European Conference on Stochastic Optimization 2017, Rom, Italy
Electronic versions:
Abstract.pdf
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2017 › Research › peer-review

Automatic Segmentation of Abdominal Fat in MRI-Scans, Using Graph-Cuts and Image Derived Energies
For many clinical studies changes in the abdominal distribution of fat is an important measure. However, the segmentation of abdominal fat in MRI scans is both difficult and time consuming using manual methods. We present here an automatic and flexible software package, that performs both bias field correction and segmentation of the fat into superficial and deep subcutaneous fat as well as visceral fat with the spinal compartment removed. Assessment when comparing to the gold standard - CT-scans - shows a correlation and bias comparable to manual segmentation. The method is flexible by tuning the image-derived energies used for the segmentation, allowing the method to be applied to other body parts, such
as the thighs.

**General information**

**Publication status:** Published

**Organisations:** Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Rector's office, Statistics and Data Analysis, University of Copenhagen

**Contributors:** Christensen, A. N., Larsen, C. T., Mandrup Jensen, C. M., Petersen, M. B., Larsen, R., Conradsen, K., Dahl, V. A.

**Pages:** 109-120

**Publication date:** 2017

**Host publication information**

**Title of host publication:** Image Analysis

**Volume:** 10270

**Publisher:** Springer

**ISBN (Print):** 9783319591285

**(Lecture Notes in Computer Science, Vol. 10270).**

**Keywords:** Computer Science, Image Processing and Computer Vision, Pattern Recognition, Artificial Intelligence (incl. Robotics), Computer Graphics, Data Mining and Knowledge Discovery

**DOIs:**

10.1007/978-3-319-59129-2_10

**Source:** FindIt

**Source-ID:** 2372493663

**Research output:** Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2017 › Research › peer-review

---

**Beskedent overtryk gav spektakulære følger**

For at undertrykke generende opskumning blev en tank sat under et såkaldt ”meget beskedent overtryk”. Pludselig svigtede samlingen i bunden. Tanken nåede en højde på 30 m og faldt ned og knuste en varevogn. Uheldet viser, at et stort volumen af gas under lavt tryk indeholder en ikke uvæsentlig mængde energi.

**General information**

**Publication status:** Published

**Organisations:** Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis

**Contributors:** Hedlund, F. H.

**Pages:** 16-18

**Publication date:** 2017

**Peer-reviewed:** Unknown

**Publication information**

**Journal:** Kjemi

**Volume:** 77

**Issue number:** 1

**ISSN (Print):** 0023-1983

**Original language:** Danish

**Bibliographical note**

Artikkelen er tidligere trykket i Dansk Kemi 97, nr. 4, 2016

**Source:** PublicationPreSubmission

**Source-ID:** 130116512

**Research output:** Contribution to journal › Journal article – Annual report year: 2017 › Communication

---

**Big Data Analytics for Industrial Process Control**

Today, in modern factories, each step in manufacturing produces a bulk of valuable as well as highly precise information. This provides a great opportunity for understanding the hidden statistical dependencies in the process. Systematic analysis and utilization of advanced analytical methods can lead towards more informed decisions. In this article we discuss some of the challenges related to big data analysis in manufacturing and relevant solutions to some of these challenges.

**General information**

**Publication status:** Published

**Organisations:** Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University

**Contributors:** Khan, A. R., Schioler, H., Kulahci, M., Knudsen, T. S.
Biting midges (Diptera, Ceratopogonidae) of the genus Culicoides are important vectors of pathogens causing diseases in free living and production animals and can lead to large economic losses in many European countries. In Europe, Culicoides imicola and the Obsoletus group are considered to be the main vectors of bluetongue virus that mostly affects ruminants such as cattle and sheep. Spatio-temporal modelling of vector distribution and abundance allows us to identify high risk areas for virus transmission and can aid in applying effective surveillance and control measures.

We used presence-absence and monthly abundance data of Culicoides from 1005 sites across 9 countries (Spain, France, Denmark, Poland, Switzerland, Austria, Poland, Sweden, Norway) collected between the years 2007 and 2013. The dataset included information on the vector species abundance (number of specimens caught per night), GPS coordinates of each trap, start and end dates of trapping. We used 120 environmental predictor variables together with Random Forest machine learning algorithms to predict the overall species distribution (probability of occurrence) and monthly abundance in Europe. We generated maps for every month of the year, to visualize the abundance of C. imicola and Obsoletus group in Europe as well as distribution maps showing the probability of occurrence.

We were able to create predictive maps of both Culicoides sp. occurrence and abundance using Random Forest models, and although the variance was large, the predicted abundance values for each site had a positive correlation with the observed abundance. We found relatively large spatial variations in probability of occurrence and abundance for both C. imicola and the Obsoletus group. For C. imicola probability of occurrence and abundance was higher in southern Spain, where as the Obsoletus group had higher probability of occurrence and abundance in central and northern Europe such as France and Germany. Temporal variation was also observed with higher abundance occurring during summer months and low or no abundance during winter months for both C. imicula and the Obsoletus group, although abundance was generally higher for a longer period of time for C. imicula than for the Obsoletus group.

Using machine learning techniques, we were able to model the spatial distribution in Europe for C. imicola and the Obsoletus group in terms of abundance and suitability (probability of occurrence). Our maps corresponded well with the previously reported distribution for C. imicola and the Obsoletus group. The observed seasonal variation was also consistent with reported population dynamics for Culicoides, as it depends on environmental factors such as temperature and rainfall. Longer seasonal abundance for C. imicula compared to the Obsoletus group can be explained by the species distribution, as C. imicula is limited to the southern parts of Europe where the warm season lasts longer, whereas the Obsoletus group is found further north. The outputs obtained here will be used as input for epidemiological models and can be helpful for determining high risk areas for disease transmission.

General information
Publication status: Published
Organisations: National Veterinary Institute, Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Zaragoza, University of the Balearic Islands, Avia-GIS, Aarhus University, Roskilde University, National Veterinary Institute, Bernhard Nocht Institute for Tropical Medicine, National Veterinary Research Institute, Norwegian Veterinary Institute, Institute for Veterinary Public Health, Centre de cooperation Internationale en Recherche Agronomique pour le Développement, Universite de Strasbourg, EID Méditerranée, University of Zurich
Number of pages: 2
Publication date: 2017

Host publication information
Title of host publication: NKVet Symposium 2017 - abstract book
Place of publication: Oslo, Norway
Biomass accident investigations – missed opportunities for learning and accident prevention

The past decade has seen a major increase in the production of energy from biomass. The growth has been mirrored in an increase of serious biomass related accidents involving fires, gas explosions, combustible dust explosions and the release of toxic gasses. There are indications that the number of bioenergy related accidents is growing faster than the energy production. This paper argues that biomass accidents, if properly investigated and lessons shared widely, provide ample opportunities for improving general hazard awareness and safety performance of the biomass industry. The paper examines selected serious accidents involving biogas and wood pellets in Denmark and argues that such opportunities for learning were missed because accident investigations were superficial, follow-up incomplete and information sharing absent. In one particularly distressing case, a facility saw a repeat accident, this time with fatal outcome, still without any learning taking place. The paper presents some information on other biomass accidents in Denmark, mostly involving biogas from anaerobic digestion. Details are lacking however, precisely because the accidents were insufficiently investigated and results not communicated. The biomass industry needs to pay more attention to safety. Utmost care should be taken to avoid so-called mediashifting i.e. that the resolution of a problem within one domain, the environmental, creates a new problem in another, the workplace safety domain.

Blygsamt övertryck fick spektakulära följder

För att minska besvärande skumning sattes en tank under "mycket blygsamt" övertryck. Plötsligt brast botten. Tanken for 30 meter upp i luften, föll ned och krossade en varubil. Olyckan visar att en stor gasvolym under lågt tryck innehåller en väsentlig mängd energi.
**Cache-mesh, a Dynamics Data Structure for Performance Optimization**

This paper proposes the cache-mesh, a dynamic mesh data structure in 3D that allows modifications of stored topological relations effortlessly. The cache-mesh can adapt to arbitrary problems and provide fast retrieval to the most-referred-to topological relations. This adaptation requires trivial extra effort in implementation with the cache-mesh, whereas it may require tremendous effort using traditional meshes. The cache-mesh also gives a further boost to the performance with parallel mesh processing by caching the partition of the mesh into independent sets. This is an additional advantage of the cache-mesh, and the extra work for caching is also trivial. Though it appears that it takes effort for initial implementation, building the cache-mesh is comparable to a traditional mesh in terms of implementation.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Nguyen, T. T., Dahl, V. A., Bærentzen, J. A.
Pages: 193-205
Publication date: 2017
Peer-reviewed: Yes

**Publication information**
Journal: Procedia Engineering
Volume: 203
ISSN (Print): 1877-7058
Ratings:
Scopus rating (2017): CiteScore 0.89
Web of Science (2017): Indexed yes
Original language: English
Keywords: Cache, Geometry processing, Dynamics structure, Data Structure, Performance optimization
Electronic versions:
1_s2.0_S1877705817343680_main.pdf
DOIs:
10.1016/j.proeng.2017.09.807
Source: FindIt
Source-ID: 2392237215

**Cellular shear stiffness reflects progression of arsenic-induced transformation during G1**
Cancer cells consistently exhibit decreased stiffness, however the onset and progression of this change has not been characterized. To study the development of cell stiffness changes we evaluated the shear stiffness of populations of cells during transformation to a carcinogenic state. Bronchial epithelial cells were exposed to sodium arsenite to initiate early stages of transformation. Exposed cells were cultured in soft agar to further transformation and select for clonal populations exhibiting anchorage independent growth. Shear stiffness of various cell populations in G1 was assessed using a novel non-invasive assay that applies shear stress with fluid flow and evaluates nano-scale deformation using quantitative phase imaging (QPI). Arsenic treated cells exhibited reduced stiffness relative to control cells, while arsenic clonal lines, selected by growth in soft agar, were found to have reduced stiffness relative to control clonal lines, which were cultured in soft agar but did not receive arsenic treatment. The relative standard deviation of the stiffness of Arsenic clones was reduced compared to control clones, as well as to the arsenic exposed cell population. Cell stiffness at the population level exhibits potential to be a novel and sensitive framework for identifying the development of cancerous cells.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, New York University School of Medicine, Duke University, University of Copenhagen
Contributors: Muñoz, A., Eldridge, W. J., Jakobsen, N. M., Sørensen, H., Wax, A., Costa, M.
Pages: 109-117
Publication date: 2017
Peer-reviewed: Yes

**Publication information**
Journal: Carcinogenesis
Volume: 39
Issue number: 2
Characterization of absorption enhancers for orally administered therapeutic peptides in tablet formulations - Applying statistical learning

To develop a successful oral formulation of insulin for treatment of type-2 diabetes patients would be a great milestone in terms of convenience. Besides protecting insulin from enzymatic cleavage in the small intestine, the formulation must overcome the intestinal epithelia barrier. Absorption enhancers are needed to ensure even a few percent of insulin are taken up. In thesis article 1, various methods to measure the effect of absorption enhancement and enzyme stability of insulin were applied. The major class of absorption enhancers is surfactant-like enhancers and is thought to promote absorption by mildly perturbing the epithelial membranes of the small intestine. The Caco-2 (Carcinoma Colon) cells can grow an artificial epithelial layer, and are used to test the potency of new absorption enhancers. This project was aimed to identify new absorption enhancers, that are both potent and sufficiently soluble. Quantitative structural activity relationship (QSAR) modeling is an empiric approach to learn relationships between molecular formulas and the biochemical properties using statistical models. A public data set testing the potency of absorption enhancers in Caco-2 was used to build a QSAR model to screen for new potent permeation enhancers. Thesis article 2 contains likely the first QSAR model to predict absorption enhancement. The model was verified by predicting molecules not tested before in Caco-2. The Caco-2 model overestimates the clinical effect of lipophilic permeation enhancers. In the Caco-2 model all reagents are pre-dissolved, and therefore the assay cannot predict critical solubility issues and bile salt interactions in the final tablet formulation. A QSAR solubility model was built to foresee and avoid slow tablet dissolution. Due to enzyme kinetics, slow tablet dissolution will allow most insulin to be deactivated by intestinal enzymes. The combined predictions of potency and solubility, will likely provide a more useful in-silico screening of potential permeation enhancers.

Random forest was used to learn relationships between molecular descriptors and potency or solubility. However, unlike multiple linear regression, the explicitly stated random forest model is complex, and therefore difficult to interpret and communicate. Any supervised regression model can be understood as a high dimensional surface connecting any possible combination of molecular properties with a given prediction. This high dimensional surface is also difficult to comprehend, but for random forests, it was discovered that a method, feature contributions, was especially useful to decompose and visualize model structures. The visualization technique was named forest floor and could replace the otherwise widely use technique partial dependence plots, especially in terms of discovering interactions in the model structure. Thesis article 3 describes the forest floor method. An R package forestFloor was developed to compute feature contributions and visualize these according to the ideas of thesis article 3. Better interpretation of random forest models is an exciting interdisciplinary field, as it allows investigators of many backgrounds to find fairly complicated relationships in data sets without in advance specifying what parameters to estimate. Forest floor was used to explain how potency and solubility were predicted by random forest models.
Cloud-shadow removal for Unmanned Aerial System multispectral imagery based on tensor decomposition methods

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Environmental Engineering, Water Resources Engineering
Contributors: Baum, A., Wang, S., Garcia, M.
Number of pages: 1
Publication date: 2017

Host publication information
Title of host publication: Book of Abstracts Sustain 2017
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Article number: Sustain Abstract D-7
Electronic versions:
SustainAbstracts2017c.compressed_40.pdf

Research output: Chapter in Book/Report/Conference proceeding › Conference abstract in proceedings – Annual report
year: 2017 › Research › peer-review

Development of a New Fractal Algorithm to Predict Quality Traits of MRI Loins
Traditionally, the quality traits of meat products have been estimated by means of physico-chemical methods. Computer vision algorithms on MRI have also been presented as an alternative to these destructive methods since MRI is non-destructive, non-ionizing and innocuous. The use of fractals to analyze MRI could be another possibility for this purpose. In this paper, a new fractal algorithm is developed, to obtain features from MRI based on fractal characteristics. This algorithm is called OPFTA (One Point Fractal Texture Algorithm). Three fractal algorithms were tested in this study: CFA (Classical fractal algorithm), FTA (Fractal texture algorithm) and OPFTA. The results obtained by means of these three fractal algorithms were correlated to the results obtained by means of physico-chemical methods. OPFTA and FTA achieved correlation coefficients higher than 0.75 and CFA reached low relationship for the quality parameters of loins. The best results were achieved for OPFTA as fractal algorithm (0.837 for lipid content, 0.909 for salt content and 0.911 for moisture). These high correlation coefficients confirm the new algorithm as an alternative to the classical computational approaches (texture algorithms) in order to compute the quality parameters of meat products in a non-destructive and efficient way.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Extremadura, University of Copenhagen
Contributors: Caballero, D., Caro, A., Amigo, J. M., Dahl, A. B., Ersbøll, B. K., Pérez-Palacios, T.
Number of pages: 11
Pages: 208-218
Publication date: 2017

Host publication information
Title of host publication: Computer Analysis of Images and Patterns
Volume: 10424
Publisher: Springer
ISBN (Print): 9783319646893
(Lecture Notes in Computer Science, Vol. 10424).
Keywords: Computer Science, Image Processing and Computer Vision, Artificial Intelligence (incl. Robotics), Information Systems and Communication Service, Special Purpose and Application-Based Systems, Probability and Statistics in Computer Science, MRI, Fractal, Algorithms, Quality traits, Iberian loin
DOIs: 10.1007/978-3-319-64689-3_17
Source: FindIt
Source-ID: 2372752209
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report
year: 2017 › Research › peer-review

DTU Employee Satisfaction Survey 2017 - importance analysis

General information
Publication status: Published
Effects of Exposure to Carbon Dioxide and Bioeffluents on Perceived Air Quality, Self-assessed Acute Health Symptoms and Cognitive Performance

The purpose of this study was to examine the effects on humans of exposure to carbon dioxide (CO₂) and bioeffluents. In three of the five exposures, the outdoor air supply rate was high enough to remove bioeffluents, resulting in a CO₂ level of 500 ppm. Chemically pure CO₂ was added to this reference condition to create exposure conditions with CO₂ at 1,000 ppm or 3,000 ppm. In two further conditions, the outdoor air supply rate was restricted so that the bioeffluent CO₂ reached 1,000 ppm or 3,000 ppm. The same twenty-five subjects were exposed for 255 minutes to each condition. Subjective ratings, physiological responses and cognitive performance were measured. No statistically significant effects on perceived air quality, acute health symptoms or cognitive performance were seen during exposures when CO₂ was added. Exposures to bioeffluents with CO₂ at 3,000 ppm reduced perceived air quality, increased the intensity of reported headache, fatigue, sleepiness and difficulty in thinking clearly, and reduced speed of addition, the response time in a redirection task and the number of correct links made in the cue-utilisation test. This suggests that moderate concentrations of bioeffluents, but not pure CO₂, will result in deleterious effects on occupants during typical indoor exposures.
Exploring the Use of Design of Experiments in Industrial Processes Operating Under Closed-Loop Control: Experimentation in closed-loop systems

Industrial manufacturing processes often operate under closed-loop control, where automation aims to keep important process variables at their set-points. In process industries such as pulp, paper, chemical and steel plants, it is often hard to find production processes operating in open loop. Instead, closed-loop control systems will actively attempt to minimize the impact of process disturbances. However, we argue that an implicit assumption in most experimental investigations is that the studied system is open loop, allowing the experimental factors to freely affect the important system responses. This scenario is typically not found in process industries. The purpose of this article is therefore to explore issues of experimental design and analysis in processes operating under closed-loop control and to illustrate how Design of Experiments can help in improving and optimizing such processes. The Tennessee Eastman challenge process simulator is used as a test-bed to highlight two experimental scenarios. The first scenario explores the impact of experimental factors that may be considered as disturbances in the closed-loop system. The second scenario exemplifies a screening design using the set-points of controllers as experimental factors. We provide examples of how to analyze the two scenarios.
Explosion i rötkammare med biogas
En biogasreaktor exploderade 1990 på Vejle Centrala reningsverk. Nu, 24 år senare, när biogas är populärare än någonsin, har händelsen helt försvunnit från tillgängliga, öppna källor

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 20-23
Publication date: 2017
Peer-reviewed: Unknown

Publication information
Journal: Kemivärlden Biotech med Kemisk Tidsskrift
Issue number: 4
ISSN (Print): 1650-0725
Original language: English
Electronic versions:
2017_06_KB_Explosion_i_r_tkammare_med_biogas.pdf
Source: PublicationPreSubmission
Source-ID: 134614807
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Communication

Fighting Smoldering Fires in Silos – A Cautionary Note on Using Carbon Dioxide to Inert
This communication seeks to draw attention to the hazards of releasing liquid carbon dioxide into environments where an ignitable atmosphere may exist. Static discharges have sufficient energy to ignite flammable vapors and an internal explosion may result when fighting smoldering fires using this approach.
A recent article in Biomass and Bioenergy examines an explosion in a Norwegian wood pellet silo when attempting to suppress a smoldering fire with CO₂. The article argues that the electrostatic hazard of CO₂ is widely under-appreciated and incidents like this are avoidable.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Publication date: 2017

Publication information
Year: 2017
Original language: English
Electronic versions:
Fighting_Smoldering_Fires_in_Silos_A_Cautionary_Note_on_Using_Carbon_Dioxide_to_Inert.pdf
URLs:
http://www.mydustexplosionresearch.com/smoldering-fires-carbon-dioxide/

Bibliographical note
Senest ændret: 27/11/2017
Source: PublicationPreSubmission
Source-ID: 139939624
Research output: Other contribution › Net publication - Internet publication – Annual report year: 2017 › Communication

First hitting probabilities for semi markov chains and estimation
We first consider a stochastic system described by an absorbing semi-Markov chain with finite state space and we introduce the absorption probability to a class of recurrent states. Afterwards, we study the first hitting probability to a
subset of states for an irreducible semi-Markov chain. In the latter case, a nonparametric estimator for the first hitting probability is proposed and the asymptotic properties of strong consistency and asymptotic normality are proven. Finally, a numerical application on a five-state system is presented to illustrate the performance of this estimator.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Georgiadis, S.
Number of pages: 23
Pages: 2435-2446
Publication date: 2017
Peer-reviewed: Yes

**Publication information**
Journal: Communications in Statistics: Theory and Methods
Volume: 46
Issue number: 5
ISSN (Print): 0361-0926
Ratings:
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 0.47 SJR 0.352 SNIP 0.746
Web of Science (2017): Impact factor 0.353
Web of Science (2017): Indexed yes
Original language: English
Keywords: Semi-Markov chains, Absorption probability, First hitting probability, Nonparametric estimator, Asymptotic properties
DOI: 10.1080/03610926.2015.1045080
Source: PublicationPreSubmission
Source-ID: 126841162
Research output: Contribution to journal > Journal article – Annual report year: 2016 > Research > peer-review

Foderkvalitet og andre faktorer af betydning for forbruget af antibiotika på minkgårde
Antibiotikaforbruget hos mink er steget gradvist gennem det seneste årti, og forbruget på den enkelte gård påvirkes af en række faktorer. Tidligere studier har vist at forekomsten af diarre og ordination af antibiotika er influeret af foderleverandøren. Formålet med dette studie (Jensen et al. 2017) var at undersøge for eventuelle effekter af de foderparametre, som indgår i den frivillige foderkontrol, på ordination af antibiotika til mink. Studiet inkluderede alle undersøgte foderbatch fra 12 fodercentraler og tilknyttede 1472 minkgårde i perioden 2012–2014. De undersøgte foderparametre omfattede både kemiske parametre og mikrobiologiske parametre. Data blev analyseret ved multivariat variansanalyse i to modeller. I den første model var responsvariablen ordination af antibiotika på en given gård i tidsrum på 3,5 eller 7 dage fra dagen efter udfodring af den givne batch. I den anden model var responsvariablen en proportion af de gårde der var tilknyttet en given fodercentral, dvs. andelen af gårde som fik antibiotika i tilknytning til udfodring af en given batch. Analyserne korrigerede for kendte faktorer med indflydelse på antibiotikaforbruget (p < 0.0001), herunder årstids variation, tidstrends, besætningsstørrelse og udbrud af Pseudomonas aeruginosa. I alle modeller var kimtallet for fæcale coccer signifikant (p < 0.0001) relateret til ordination af antibiotika.

**General information**
Publication status: Published
Organisations: National Veterinary Institute, Epidemiology, Diagnostic & Development, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Kopenhagen Fur
Contributors: Jensen, V. F., Chriél, M., Sommer, H. M., Struves, T., Clausen, J.
Pages: 95-98
Publication date: 2017
Peer-reviewed: Yes

**Publication information**
Journal: Faglig Årsberetning
Volume: 2017
ISSN (Print): 2445-9437
Original language: English
Keywords: Mink (neovison vison), Foderkvalitet, Antibiotika, Faktorer
Electronic versions:
Foderkvalitet_og_andre_faktorer_af_betydning_for_forbruget_af_antibiotika_p_minkg_rde.pdf
URLs:
Foreign object detection in multispectral X-ray images of food items using sparse discriminant analysis

Non-invasive food inspection and quality assurance are becoming viable techniques in food production due to the introduction of fast and accessible multispectral X-ray scanners. However, the novel devices produce massive amount of data and there is a need for fast and accurate algorithms for processing it. We apply a sparse classifier for foreign object detection and segmentation in multispectral X-ray. Using sparse methods makes it possible to potentially use fewer variables than traditional methods and thereby reduce acquisition time, data volume and classification speed. We report our results on two datasets with foreign objects, one set with spring rolls and one with minced meat. Our results indicate that it is possible to limit the amount of data stored to 50% of the original size without affecting classification accuracy of materials used for training. The method has attractive computational properties, which allows for fast classification of items in new images.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Copenhagen Center for Health Technology, Danish Technological Institute
Contributors: Einarsson, G., Jensen, J. N., Paulsen, R. R., Einarsdottir, H., Ersbøll, B. K., Dahl, A. B., Christensen, L. B.
Pages: 350-361
Publication date: 2017

Host publication information
Title of host publication: SCIA 2017
Volume: 10269
Publisher: Springer
ISBN (Print): 9783319591254
(Lecture Notes in Computer Science, Vol. 10269).
Keywords: Theoretical Computer Science, Computer Science (all), Foreign object detection, Multispectral, Sparse classification, X-ray, Discriminant analysis, Image analysis, Object recognition, Quality assurance, X rays, Acquisition time, Classification accuracy, Computational properties, Fast and accurate algorithms, Fast classification, Multi-spectral, Sparse classifiers, Object detection
DOIs: 10.1007/978-3-319-59126-1_29
Source: FindIt
Source-ID: 2371753494
Research output: Chapter in Book/Report/Conference proceeding – Article in proceedings – Annual report year: 2017 – Research – peer-review

Formulating and testing a method for perturbing precipitation time series to reflect anticipated climatic changes

Urban water infrastructure has very long planning horizons, and planning is thus very dependent on reliable estimates of the impacts of climate change. Many urban water systems are designed using time series with a high temporal resolution. To assess the impact of climate change on these systems, similarly high-resolution precipitation time series for future climate are necessary. Climate models cannot at their current resolutions provide these time series at the relevant scales. Known methods for stochastic downscaling of climate change to urban hydrological scales have known shortcomings in constructing realistic climate-changed precipitation time series at the sub-hourly scale. In the present study we present a deterministic methodology to perturb historical precipitation time series at the minute scale to reflect non-linear expectations to climate change. The methodology shows good skill in meeting the expectations to climate change in extremes at the event scale when evaluated at different timescales from the minute to the daily scale. The methodology also shows good skill with respect to representing expected changes of seasonal precipitation. The methodology is very robust against the actual magnitude of the expected changes as well as the direction of the changes (increase or decrease), even for situations where the extremes are increasing for seasons that in general should have a decreasing trend in precipitation. The methodology can provide planners with valuable time series representing future climate that can be used as input to urban hydrological models and give better estimates of climate change impacts on these systems.

General information
Publication status: Published
Organisations: Department of Environmental Engineering, Urban Water Systems, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Ramboll Group AS
Contributors: Sørup, H. J. D., Georgiadis, S., Gregersen, I. B., Arnbjerg-Nielsen, K.
Pages: 345-355
Publication date: 2017
Geometrical Characterisation of Individual Fibres From X-Ray Tomograms

We have developed an image analysis pipeline\textsuperscript{1} that can extract individual fibre tracks from low contrast X-ray tomograms of unidirectional composites with high fibre volumefraction. Measuring individual fibre tracks opens up the possibility of modelling the empirical data in a statistical manner. Thus, allowing to analyse the spatial distribution of the parameters characterising the orientation and curvature of these individual fibres, which can also provide insights on the interactions amongst the individual fibres. Finite element models (FEMs) can be built from the extracted geometry to simulate the performance of the scanned fibre structure under realistic conditions. Moreover, aspects of the fibre architecture that influence the macroscopic behaviour of the composite can be quantified. Examples are 2D FEMs to predict the transverse stiffness or the quantification of fibre orientations to estimate the compression strength.\textsuperscript{1} And last but not least, already developed analytical and numerical models to describe the composite's behaviour can be validated against the observed data.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Number of pages: 1
Pages: 59
Publication date: 2017

Subject
Keywords: X-ray tomography, Individual Fibres, Unidirectional Composites, Modelling
Electronic versions: GEOMETRICAL_CHARACTERISATION.pdf

Host publication information
Title of host publication: Proceedings of the 30th Nordic Seminar on Computational Mechanics (NSCM-30)
Editors: Høsberg, J., Pedersen, N.
Keywords: X-ray tomography, Individual Fibres, Unidirectional Composites, Modelling
Electronic versions: GEOMETRICAL_CHARACTERISATION.pdf
Research output: Chapter in Book/Report/Conference proceeding » Conference abstract in proceedings – Annual report year: 2017 » Research » peer-review

Haringsvær - Vejledning for kolonne 3-virksomheder - Sikring af risikovirksomheder

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Number of pages: 38
Publication date: 2017
Høringssvar - Vejledning om definition større uheld 2017

How to foster a High-Tech entrepreneurial mind-set – A multidisciplinary engineering course for Bachelor students

Imaging for monitoring downstream processing of fermentation broths
In relation to downstream processing of a fermentation broth coagulation/flocculation is a typical pretreatment method for separating undesirable particles/impurities from the wanted product. In the coagulation process the negatively charged impurities are destabilized by adding of a clarifying agent thereby neutralizing the charges on the particles. Particles thus agglomerate. Larger agglomerates are formed in the flocculation process by adding a polymer, which forms bridges between the particles. The operation of coagulators, flocculators and clarifiers requires trained operators implying the human factor to play a major risk with regard to performance. Better process monitoring will provide the means for improved control giving higher yield, better quality, and minimize the consumption of water. In particular, the optimal separation of biomass from a soluble enzyme phase is often dependent on an initial coagulation of the biomass and a final flocculation of the solids just prior to separation. We investigate flocculation processes at Novozymes facilities so that the response time and risk of error is minimized. We use oCelloScope [1], an automated microscope, for imaging samples
from the flocculation process and subsequently we extract image features for qualitative and quantitative image characterization. The processing include image morphology, image segmentation and image quantification. The aim is to correlate image information to “quality” of the separation process. Here we report our initial finding. [1] M.Fredborg et al. Journal of Clinical Microbiology Vol 51 Number 7 p. 2047–2053 (2013); http://www.biosensesolutions.dk

Interpolation from Grid Lines: Linear, Transfinite and Weighted Method
When two sets of line scans are acquired orthogonal to each other, intensity values are known along the lines of a grid. To view these values as an image, intensities need to be interpolated at regularly spaced pixel positions. In this paper we evaluate three methods for interpolation from grid lines: linear, transfinite and weighted. Linear method does not preserve the known values along the grid lines. Transfinite method, known from mesh generation, preserves the known values but might cause overshoot. The weighted method, which we propose, is designed to combine the desired properties of transfinite method close to grid lines, and the stability of the linear method. We perform an extensive evaluation of the three interpolation methods across a range of upsampling rates for two data sets. Depending on the upsampling rate, we show significant difference in the performance of the three methods. We find that the transfinite interpolation works well for small upsampling rates and the proposed weighted interpolation method performs very well for all relevant upsampling rates.

Investigating the influence of product perception and geometric features
Research in emotional design and Kansei Engineering has shown that aesthetics play a significant role in the appeal of a product. This paper contributes to establishing a methodology to identify the relationships between perceptions, aesthetic features, desire to own and background of consumers. Surveys were conducted with 71 participants to gather their perceptions of 11 vase concepts. Advanced statistical analyses, including mixed models, were applied to allow generalisation of the results beyond the data sample. Significant relations between the desire to own a product and how the product is perceived were found (the desire to own was found to be related to beautiful, expensive, elegant, exciting, feminine, common and dynamic vases), as well as between the perceptions and the parameters describing the form of the vases (a vase was perceived as beautiful if it had many curved lines and was simple and tall). An automated mixed model analysis was conducted and revealed that general rules can be found between aesthetic features, perceptions and ownership, which can apply across gender and culture. The findings include design rules that link aesthetic features with perceptions. These contribute to research as guidelines for design synthesis and can either be implemented via shape
grammars or parametric modelling approaches. These rules are also interesting for 3D printing applications, especially important when the consumer is the designer. Some of these design rules are linked to the desire to own a product, they have implications for industry, and they offer guidelines to creating attractive products that people want to own.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Management Engineering, Technology and Innovation Management, Imperial College London, University of Tokyo
Contributors: Perez Mata, M., Ahmed-Kristensen, S., Brockhoff, P. B., Yanagisawa, H.
Pages: 357–379
Publication date: 2017
Peer-reviewed: Yes

**Publication information**
Journal: Research in Engineering Design
Volume: 28
Issue number: 3
ISSN (Print): 0934-9839
Ratings:
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 2.83 SJR 1.024 SNIP 2.139
Web of Science (2017): Indexed yes
Original language: English
Keywords: Emotional design, Kansei Engineering, Aesthetics, Perception, Product form and geometry
Electronic versions:
melju_Investigating_the_influence_of_product_perception_and_geometric_features_MPM_SAK_PBB_HY_Final_1_.pdf.

**Lederånsundersøgelse - Lønberegner til BUPL**

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Andersen, J. F., Thyregod, C., Ersbøll, B. K.
Number of pages: 8
Publication date: 2017

**Publication information**
Publisher: DTU Compute
Original language: English

**lmerTest Package: Tests in Linear Mixed Effects Models**
One of the frequent questions by users of the mixed model function lmer of the lme4 package has been: How can I get p values for the F and t tests for objects returned by lmer? The lmerTest package extends the 'lmerMod' class of the lme4 package, by overloading the anova and summary functions by providing p values for tests for fixed effects. We have implemented the Satterthwaite’s method for approximating degrees of freedom for the t and F tests. We have also implemented the construction of Type I - III ANOVA tables. Furthermore, one may also obtain the summary as well as the anova table using the Kenward-Roger approximation for denominator degrees of freedom (based on the KRmodcomp function from the pbkrtest package). Some other convenient mixed model analysis tools such as a step method, that performs backward elimination of nonsignificant effects - both random and fixed, calculation of population means and multiple comparison tests together with plot facilities are provided by the package as well.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Managerial implications for improving continuous production processes

Data analytics remains essential for process improvement and optimization. Statistical process control and design of experiments are among the most powerful process and product improvement methods available. However, continuous process environments challenge the application of these methods. In this article we highlight SPC and DoE implementation challenges described in the literature for managers, researchers and practitioners interested in continuous production process improvement. The results may help managers support the implementation of these methods and make researchers and practitioners aware of methodological challenges in continuous process environments.

Matrix-Exponential Distributions in Applied Probability

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Universidad Nacional Autonoma de Mexico
Contributors: Bladt, M., Nielsen, B. F.
Number of pages: 736
Publication date: 2017

Publication information
Modelling allergic risk

Up to 20 million Europeans suffer from food allergies. Due to the lack of knowledge about why food allergies developed or how to protect allergic consumers from the offending food, food allergy management is mainly based on food allergens avoidance. The iFAAM project (Integrated approaches to Food Allergen and Allergy Management) aims at developing strategies for food allergies based on evidences.

Especially, food allergen risk assessment helps food producers or authorities to make decisions on withdrawing a food product from the market or adding more information on the label when allergen presence is unintended. The risk assessment method has three different kinds of input. The exposure is calculated from the product consumption and the allergen contamination in the food product. The exposure is then compared to the thresholds to which allergic individuals react in order to calculate the chance of allergic reaction in the population.

In allergen risk assessment, the emphasis was on the threshold data, and no effort was made on consumption data. Moreover, no pan-European consumption data suitable for allergen risk assessment are available. A procedure for grouping food products automatically across countries is proposed. Thus, the allergen risk assessment can be performed cross-nationally and for the correct food group.

Then the two probabilistic risk assessment methods usually used were reviewed and compared. First order Monte-Carlo simulations are used in one method [14], whereas the other one combines second order Monte-Carlo simulations with Bayesian inferences [13]. An alternative method using second order Monte-Carlo simulations was proposed to take into account the uncertainty from the inputs. The uncertainty propagation from the inputs to the risk of allergic reaction was also evaluated for all the methods using uncertainty analysis [11].

The recommended approach for the allergen risk assessment was implemented in a Shiny application with the R software. Thus, allergen risk assessment can be performed easily by non-statisticians with the interactive application.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Birot, S.
Number of pages: 204
Publication date: 2017

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
(DTU Compute PHD-2016; No. 441).
Electronic versions:
phd441_Birot_S.pdf

Modelling Dietary Exposure to Chemical Components in Heat-Processed Meats

Several chemical compounds that potentially increase the risk of developing cancer in humans are formed during heat processing of meat. Estimating the overall health impact of these compounds in the population requires accurate estimation of the exposure to the chemicals, as well as the probability that different levels of exposure result in disease. The overall goal of this study was to evaluate the impact of variability of exposure patterns and uncertainty of exposure data in burden of disease estimates. We focus on the first phase of burden of disease modelling, i.e. the estimation of exposure to selected compounds in the Danish population, based on concentration and consumption data. One of the challenges that arises in the probabilistic modelling of exposure is the presence of “artificial” zero counts in concentration data due to the detection level of the applied tests. Zero inflated models, e.g. the Poisson-Lognormal approach, are promising tools to address this obstacle. The exposure estimates can then be applied to dose-response models to quantify the cancer risk.
Multi-phase Volume Segmentation with Tetrahedral Mesh

Volume segmentation is efficient for reconstructing material structure, which is important for several analyses, e.g. simulation with finite element method, measurement of quantitative information like surface area, surface curvature, volume, etc. We are concerned about the representations of the 3D volumes, which can be categorized into two groups: fixed voxel grids [1] and unstructured meshes [2]. Among these two representations, the voxel grids are more popular since manipulating a fixed grid is easier than an unstructured mesh, but they are less efficient for quantitative measurements. In many cases, the voxel grids are converted to explicit meshes, however the conversion may reduce the accuracy of the segmentations, and the effort for meshing is also not trivial. On the other side, methods using unstructured meshes have difficulty in handling topology changes. To reduce the complexity, previous methods only represent the surfaces, thus they only segment a single region without exterior or interior information (e.g. holes). Finally, yet importantly, previous methods of both representations have issues with multi-material segmentation, where vacuum and overlapping between surfaces occur. This paper proposes a method for volume segmentation using a tetrahedral mesh. The compelling advantages of our method include: natural multi-material support; output is tetrahedral mesh that can be utilized for simulation and analysis directly; and the ability to control the resolution for compact meshes. We are also experimenting to prove our advantages on high accuracy; and the potentiality to accompany shape prior information during segmentation.

Nguyen Dahl Bærentzen 2017 Multi-phase Volume Segmentation with Tetrahedral Mesh.pdf
Research output: Contribution to conference › Paper – Annual report year: 2017 › Research › peer-review

New approach for validating the segmentation of 3D data applied to individual fibre extraction

We present two approaches for validating the segmentation of 3D data. The first approach consists on comparing the amount of estimated material to a value provided by the manufacturer. The second approach consists on comparing the segmented results to those obtained from imaging modalities that provide a better resolution and therefore a more accurate segmentation. The imaging modalities used for comparison are scanning electron microscopy, optical microscopy and synchrotron CT. The validation methods are applied to the asses the segmentation of individual fibres from X-ray microtomograms.


On the structure of dynamic principal component analysis used in statistical process monitoring

When principal component analysis (PCA) is used for statistical process monitoring it relies on the assumption that data are time independent. However, industrial data will often exhibit serial correlation. Dynamic PCA (DPCA) has been suggested as a remedy for high-dimensional and time-dependent data. In DPCA the input matrix is augmented by adding time-lagged values of the variables. In building a DPCA model the analyst needs to decide on (1) the number of lags to add, and (2) given a specific lag structure, how many principal components to retain. In this article we propose a new analyst driven method to determine the maximum number of lags in DPCA with a foundation in multivariate time series analysis. The method is based on the behavior of the eigenvalues of the lagged autocorrelation and partial autocorrelation matrices. Given a specific lag structure we also propose a method for determining the number of principal components to retain. The number of retained principal components is determined by visual inspection of the serial correlation in the squared prediction error statistic, Q (SPE), together with the cumulative explained variance of the model. The methods are illustrated using simulated vector autoregressive and moving average data, and tested on Tennessee Eastman process data.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Luleå University of Technology
Contributors: Vanhatalo, E., Kulahci, M., Bergquist, B.
Number of pages: 11
Pages: 1-11
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Chemometrics and intelligent laboratory systems
Volume: 167
ISSN (Print): 0169-7439
Ratings:
  BFI (2017): BFI-level 2
  Scopus rating (2017): CiteScore 2.82 SJR 0.672 SNIP 1.222
  Web of Science (2017): Impact factor 2.701
  Web of Science (2017): Indexed yes
Original language: English
Keywords: Dynamic principal component analysis, Vector autoregressive process, Vector moving average process, Autocorrelation, Simulation, Tennessee Eastman process simulator
Electronic versions:
DOIs:
10.1016/j.chemolab.2017.05.016
Source: FindIt
Source-ID: 2358758302
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Research › peer-review

Optimistic and pessimistic self-assessment of own diets is associated with age, self-rated health and weight status in Danish adults

The aim of this study was to analyse concordance between Danish adults’ recorded diet quality and their own assessment of the healthiness and to examine socio-demographic, health and behavioural characteristics associated with an optimistic or pessimistic self-assessment. Data were derived from The Danish National Survey of Diet and Physical Activity 2011-2013 and included a random sample of 3014 adults (18-75 y). Diet quality was evaluated on the basis of seven-day pre-coded food diaries and categorised 'unhealthy', 'somewhat healthy' and 'healthy'. Self-assessment of the healthiness of own diets was registered via personal interviews and categorised healthy enough 'to a high degree', 'to some degree' or 'not at all/only partly'. Highly and somewhat optimistic self-assessment, respectively, were defined as assessing own diets
as healthy enough to a high degree or to some degree while having unhealthy diets. Highly and somewhat pessimistic self-assessment, respectively, were defined as assessing own diets as not healthy enough or healthy enough to some degree while having healthy diets. Multiple logistic regression models were used to examine characteristics associated with optimistic and pessimistic self-assessments, respectively. Among individuals with unhealthy diets, 13% were highly optimistic and 42% somewhat optimistic about the healthiness of their diets. Among individuals with healthy diets, 14% were highly pessimistic and 51% somewhat pessimistic about the healthiness of their diets. Highly optimistic self-assessment was associated with increasing age, excellent self-rated health, normal weight and a moderate activity level. Highly pessimistic self-assessment was associated with decreasing age, good self-rated health and being overweight or obese. The findings indicate that people seem to use personal health characteristics as important references when assessing the healthiness of their diets.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Research group for Risk Benefit, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Contributors: Sørensen, M. R., Matthiessen, J., Holm, L., Knudsen, V. K., Andersen, E. W., Tetens, I.
Pages: 15-22
Publication date: 2017
Peer-reviewed: Yes

Optimization of hospital ward resources with patient relocation using Markov chain modeling
Overcrowding of hospital wards is a well-known and often revisited problem in the literature, yet it appears in many different variations. In this study, we present a mathematical model to solve the problem of ensuring sufficient beds to hospital wards by re-distributing beds that are already available to the hospital. Patient flow is modeled using a homogeneous continuous-time Markov chain and optimization is conducted using a local search heuristic. Our model accounts for patient relocation, which has not been done analytically in literature with similar scope. The study objective is to ensure that patient occupancy is reflected by our Markov chain model, and that a local optimum can be derived within a reasonable runtime. Using a Danish hospital as our case study, the Markov chain model is statistically found to reflect occupancy of hospital beds by patients as a function of how hospital beds are distributed. Furthermore, our heuristic is found to efficiently derive the optimal solution. Applying our model to the hospital case, we found that relocation of daily arrivals can be reduced by 11.7% by re-distributing beds that are already available to the hospital.

General information
Publication status: Published
Organisations: Department of Management Engineering, Management Science, Operations Research, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Contributors: Andersen, A. R., Nielsen, B. F., Reinhardt, L. B.
Pages: 1152-1163
Publication date: 2017
Peer-reviewed: Yes
Outlier Detection in End-User Performance Monitoring - Smart Innovation.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Andersen, J. F., Stockmarr, A., Thyregod, C., Erbsøll, B. K.
Publication date: 2017

Publication information
Publisher: Technical University of Denmark (DTU)
Original language: English

Bibliographical note
Confidential report, not accessible to the public.
Research output: Book/Report > Report – Annual report year: 2017 > Research

Pillar-type acoustic metasurface
We theoretically investigate acoustic metasurfaces consisting of either a single pillar or a line of identical pillars on a thin plate, and we report on the dependence on the geometrical parameters of both the monopolar compressional and dipolar bending modes. We show that for specific dimensions of the resonators, bending and compressional modes may be simultaneously excited. We study their interaction with an antisymmetric Lamb wave, whether or not they occur at the same frequency, with particular consideration for the amplitude and phase of waves emitted by the pillars at resonance. Especially, the analysis of both the amplitude and the phase of the wave at the common resonant frequency downstream from a line of pillars demonstrates that the reemitted waves allow for the transmission with phase shift of $\pi$.
Pilot Demonstration of Alarm Management in Oil & Gas Operations. - Decision Support from Functional Modelling.

General information
Publication status: Published
Organisations: Centre for oil and gas – DTU, Statistics and Data Analysis, Department of Applied Mathematics and Computer Science
Contributors: Jørgensen, T. M.
Pages: 31-31
Publication date: 2017

Host publication information
Title of host publication: DHRTC Research Abstracts
Electronic versions:
DHRTC_Presentation_abstracts.pdf
Research output: Chapter in Book/Report/Conference proceeding – Conference abstract in proceedings – Annual report year: 2018 › Research › peer-review

Prediction of pork quality parameters by applying fractals and data mining on MRI
This work firstly investigates the use of MRI, fractal algorithms and data mining techniques to determine pork quality parameters non-destructively. The main objective was to evaluate the capability of fractal algorithms (Classical Fractal algorithm, CFA; Fractal Texture Algorithm, FTA and One Point Fractal Texture Algorithm, OPFTA) to analyse MRI in order to predict quality parameters of loin. In addition, the effect of the sequence acquisition of MRI (Gradient echo, GE; Spin echo, SE and Turbo 3D, T3D) and the predictive technique of data mining (Isotonic regression, IR and Multiple linear regression, MLR) were analysed. Both fractal algorithm, FTA and OPFTA are appropriate to analyse MRI of loins. The sequence acquisition, the fractal algorithm and the data mining technique seems to influence on the prediction results. For most physico-chemical parameters, prediction equations with moderate to excellent correlation coefficients were achieved by using the following combinations of acquisition sequences of MRI, fractal algorithms and data mining techniques: SE-FTA-MLR, SE-OPFTA-IR, GE-OPFTA-MLR, SE-OPFTA-MLR, with the last one offering the best prediction results. Thus, SE-OPFTA-MLR could be proposed as an alternative technique to determine physico-chemical traits of fresh and dry-cured loins in a non-destructive way with high accuracy.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science , Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Extremadura, University of Copenhagen
Contributors: Caballero, D., Pérez-Palacios, T., Caro, A., Amigo, J. M., Dahl, A. B., Ersbøll, B. K., Antequera, T.
Pages: 739-747
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Food Research International
Volume: 99
ISSN (Print): 0963-9969
Ratings:
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 3.9 SJR 1.472 SNIP 1.467
Web of Science (2017): Impact factor 3.52
Web of Science (2017): Indexed yes
Original language: English
Keywords: Acquisition sequences, Image analysis, Loin, MLR, Non-destructive analysis, Quality traits, Data mining, Forecasting, Fractals, Linear regression, Quality control, Correlation coefficient, Multiple linear regressions, Physico-chemical parameters, Physico-chemical traits, Predictive techniques, Parameter estimation
Quantitatively Measured Anatomic Location and Volume of Optic Disc Drusen: An Enhanced Depth Imaging Optical Coherence Tomography Study

Optic disc drusen (ODD) are found in up to 2.4% of the population and are known to cause visual field defects. The purpose of the current study was to investigate how quantitatively estimated volume and anatomic location of ODD influence optic nerve function. Anatomic location, volume of ODD, and peripapillary retinal nerve fiber layer and macular ganglion cell layer thickness were assessed in 37 ODD patients using enhanced depth imaging optical coherence tomography. Volume of ODD was calculated by manual segmentation of ODD in 97 B-scans per eye. Anatomic characteristics were compared with optic nerve function using automated perimetric mean deviation (MD) and multifocal visual evoked potentials. Increased age (P = 0.015); larger ODD volume (P = 0.002); and more superficial anatomic ODD location (P = 0.007) were found in patients with ODD visible by ophthalmoscopy compared to patients with buried ODD. In a multivariate analysis, a worsening of MD was significantly associated with larger ODD volume (P <0.0001). No association was found between MD and weighted anatomic location, age, and visibility by ophthalmoscopy. Decreased ganglion cell layer thickness was significantly associated with worse MD (P = 0.025) and had a higher effect on MD when compared to retinal nerve fiber layer thickness. Large ODD volume is associated with optic nerve dysfunction. The worse visual field defects associated with visible ODD should only be ascribed to larger ODD volume and not to a more superficial anatomic ODD location.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen
Pages: 2491-2497
Publication date: 2017
Peer-reviewed: Yes

Quantitative measure of optic disc drusen location in enhanced depth imaging optical coherence tomography scans

Purpose: A quantitative measure of anatomical optic disc drusen (ODD) location in the optic nerve head can be an important parameter in the investigation of ODD formation and the development of visual field defects. We propose a method for defining a quantitative measure of ODD location relative to Bruch’s membrane.

Methods: Optic disc drusen in high resolution enhanced depth imaging optical coherence tomography scans were manually segmented using ITK-SNAP by a trained ophthalmologist. To quantify a location of each ODD, we need a reference. Bruch’s membrane serves as an excellent reference, but does not exist within the optic nerve head, where the ODD are located. Therefore we performed a semi-automatic graph based segmentation of Bruch’s membrane at the margin in each B-scans. From the segmentation we obtained two landmarks per B-scan. Based on the landmarks in each B-scan, we defined a reference surface relative to Bruch’s membrane. The Euclidean distance from the center of mass of each manually segmented ODD to the defined reference surface gave a quantitative measurement of each druse location.
Furthermore, the quantitative measure was signed which indicated whether the ODD was located above or below the reference surface.

Results: We computed a reference surface (Fig. 1) based on Bruch’s membrane segmented at the margin in 97 B-scans per patient for 37 patients in total (Fig. 2). The average number of ODD was 4.4 (± 5.78) per patient and the average distance from the center of mass for each ODD to the defined reference surface was 0.19 mm (± 0.3 mm).

Conclusions: A defined reference surface based on Bruch’s membrane in ODD patients resulted in a quantitative measure of ODD location. The quantitative measure indicates whether the ODD is located above or below the reference surface. The quantitative measure of anatomical ODD location can act as an important parameter in future ODD research.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Copenhagen University Hospital
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Investigative Ophthalmology & Visual Science
Volume: 58
Issue number: 8
ISSN (Print): 0146-0404
Ratings:
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 3.26 SJR 2.058 SNIP 1.222
Web of Science (2017): Impact factor 3.388
Web of Science (2017): Indexed yes
Original language: English
Source: FindIt
Source-ID: 2435410474
Research output: Contribution to journal › Conference abstract in journal – Annual report year: 2018 › Research › peer-review

Risici ved midlertidige oplag af farligt gods.

[English summary] Following a major fireworks accident in Seest (2004), the statutory order implementing the EU Seveso directive in Denmark was enlarged in 2005 to cover also temporary storage of dangerous substances at transportation hubs such as marine terminals and railway yards. The rationale is quite sensible – the risk posed to nearby communities can be permanent even though the individual cargo is present temporarily only. The enlargement of order was ill conceived however, and implementation of the new measures has been troubled, in particular because transportation companies only have information on dangerous goods as provided in ADR/RID/IMO transportation documents, and the nature of this information is incompatible with the Seveso legislation’s definition of dangerous substances. The paper provides a critical commentary on the rule-making and subsequent consultation processes and examines the dismal implementation status 12 years later (2017).

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 20-23
Publication date: 2017
Peer-reviewed: Unknown

Publication information
Journal: Dansk Kemi
Volume: 98
Issue number: 10
ISSN (Print): 0011-6335
Scenario-Neutral Simulation of Flood Risk for Multiple Drivers

General information
Publication status: Published
Organisations: Department of Environmental Engineering, Urban Water Systems, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Monash University, Cooperative Research Centre for Water Sensitive Cities
Contributors: Löwe, R., Urich, C., Kulahci, M., Radhakrishnan, M., Deletic, A., Arnbjerg-Nielsen, K.
Number of pages: 2
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from 7th International Conference on Flood Management (ICFM7), Leeds, United Kingdom.

Selecting local constraint for alignment of batch process data with dynamic time warping
There are two key reasons for aligning batch process data. The first is to obtain same-length batches so that standard methods of analysis may be applied, whilst the second reason is to synchronise events that take place during each batch so that the same event is associated with the same observation number for every batch. Dynamic time warping has been shown to be an effective method for meeting these objectives. This is based on a dynamic programming algorithm that aligns a batch to a reference batch, by stretching and compressing its local time dimension. The resulting "warping function" may be interpreted as a progress signature of the batch which may be appended to the aligned data for further analysis. For the warping function to be a realistic reflection of the progress of a batch, it is necessary to impose some constraints on the dynamic time warping algorithm, to avoid an alignment which is too aggressive and which contains pathological warping. Previous work has focused on addressing this issue using global constraints. In this work, we investigate the use of local constraints in dynamic time warping and define criteria for evaluating the degree of time distortion and variable synchronisation obtained. A local constraint scheme is extended to include constraints not previously considered, and a novel method for selecting the optimal local constraint with respect to the two criteria is proposed. For illustration, the method is applied to real data from an industrial bacteria fermentation process.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Chr. Hansen AS
Contributors: Spooner, M. P., Kold, D., Kulahci, M.
Pages: 161-170
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Chemometrics and intelligent laboratory systems
Volume: 167
ISSN (Print): 0169-7439
Ratings:
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 2.82 SJR 0.672 SNIP 1.222
Web of Science (2017): Impact factor 2.701
Web of Science (2017): Indexed yes
Original language: English
Keywords: Batch process, Dynamic time warping, Local constraint, Global constraint, Step pattern, Trajectory synchronisation
DOIs: 10.1016/j.chemolab.2017.05.019
Source: Findit
Source-ID: 2370692164
Selection of non-zero loadings in sparse principal component analysis

Principal component analysis (PCA) is a widely accepted procedure for summarizing data through dimensional reduction. In PCA, the selection of the appropriate number of components and the interpretation of those components have been the key challenging features. Sparse principal component analysis (SPCA) is a relatively recent technique proposed for producing principal components with sparse loadings via the variance-sparsity trade-off. Although several techniques for deriving sparse loadings have been offered, no detailed guidelines for choosing the penalty parameters to obtain a desired level of sparsity are provided. In this paper, we propose the use of a genetic algorithm (GA) to select the number of non-zero loadings (NNZL) in each principal component while using SPCA. The proposed approach considerably improves the interpretability of principal components and addresses the difficulty in the selection of NNZL in SPCA. Furthermore, we compare the performance of PCA and SPCA in uncovering the underlying latent structure of the data. The key features of the methodology are assessed through a synthetic example, pitprops data and a comparative study of the benchmark Tennessee Eastman process.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of California
Contributors: Gajjar, S., Kulahci, M., Palazoglu, A.
Pages: 160-171
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Chemometrics and intelligent laboratory systems
Volume: 162
ISSN (Print): 0169-7439
Ratings:
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 2.82 SJR 0.672 SNIP 1.222
Web of Science (2017): Impact factor 2.701
Web of Science (2017): Indexed yes
Original language: English
Keywords: Genetic algorithm, Pitprops data, Principal Component Analysis (PCA), Sparse Principal Component Analysis (SPCA), Tennessee Eastman process
DOIs:
10.1016/j.chemolab.2017.01.018
Source: FindIt
Source-ID: 2352285467

Selection of objective function for imbalanced classification: an industrial case study

In this article we discuss the issue of selecting suitable objective function for Genetic Algorithm to solve an imbalanced classification problem. More precisely, first we discuss the need of specialized objective function to solve a real classification problem from our industrial partner and then we compare the results of our proposed objective function with commonly used candidates to serve this purpose. Our comparison is based on the analysis of real data collected during the quality control stages of the manufacturing process.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Contributors: Khan, A. R., Schiøler, H., Kulahci, M.
Number of pages: 4
Pages: 1-4
Publication date: 2017

Host publication information
Title of host publication: Proceedings of 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation
Publisher: IEEE
ISBN (Print): 9781509065059
DOIs:
Setup for Scenario-free Modelling of Urban Flood Risk in Non-stationary Climate and Urban Development Conditions

An early flood warning system has been developed for urban catchments and is currently running in online operation in Copenhagen. The system is highly dependent on the quality of rainfall forecast inputs. An investigation of precipitation inputs from Radar Nowcast (RN), Numerical Weather Prediction (NWP) with assimilation of radar and cloud data (RA3), and Ensemble NWP with 25 members (S05) is conducted by comparing against rain gauge measurements and flood extent. Despite lower spatial and temporal resolution, the ensemble product seems promising for forecasting extreme events. A combination of the three forecast products is expected to yield the optimal input for flood warning.

Sikring af risikovirksomheder

Nu skal risikovirksomheder udføre en sårbarhedsvurdering. Det er en følge af regeringens terrorhandlingsplan. En vejledning har netop været i høring.

Simulering af Lokalebehov i 2025

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 22-25
Publication date: 2017
Peer-reviewed: Unknown

Publication information
Journal: Dansk Kemi
Volume: 98
Issue number: 6/7
ISSN (Print): 0011-6335
Original language: English
Electronic versions:
2017_06_Dkemi_Sikring_af_risikovirksomheder.pdf
Source: PublicationPreSubmission
Source-ID: 133000165
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Communication

Simulering af Lokalebehov i 2025

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Andersen, J. F., Ersbøll, B. K., Thyregod, C.
Number of pages: 22
Publication date: 2017

Publication information
Social disparities in the prevalence of multimorbidity - A register-based population study: A register-based population study

Prevalences of multimorbidity vary between European studies and several methods and definitions are used. In this study we examine the prevalence of multimorbidity in relation to age, gender and educational attainment and the association between physical and mental health conditions and educational attainment in a Danish population. A cross-sectional design was used to study the prevalence of multimorbidity, defined as two or more chronic conditions, and of comorbid physical and mental health conditions across age groups and educational attainment levels among 1,397,173 individuals aged 16 years and older who lived in the Capital Region of Denmark on January 1st, 2012. After calculating prevalence, odds ratios for multimorbidity and mental health conditions were derived from logistic regression on gender, age, age squared, education and number of physical conditions (only for odds ratios for mental health conditions). Odds ratios for having multimorbidity and mental health conditions for each variable were adjusted for all other variables. Multimorbidity prevalence was 21.6%. Half of the population aged 65 and above had multimorbidity, and prevalence was inversely related to educational attainment: 26.9% (95% CI, 26.8-26.9) among those with lower secondary education versus 13.5% (95% CI, 13.5-13.6) among people with postgraduate education. Adjusted odds ratios for multimorbidity were 0.50 (95% CI, 0.49-0.51) for people with postgraduate education, compared to people with lower secondary education. Among all population members, 4.9% (95% CI, 4.9-4.9) had both a physical and a mental health condition, a proportion that increased to 22.6% of people with multimorbidity. Physical and mental health comorbidity was more prevalent in women (6.33%; 95% CI, 6.3-6.4) than men (3.34%; 95% CI, 3.3-3.4) and approximately 50 times more prevalent among older persons than younger ones. Physical and mental health comorbidity was also twice as prevalent among people with lower secondary education than among those with postgraduate education. The presence of a mental health condition was strongly associated with the number of physical conditions; those with five or more physical conditions had an adjusted odds ratio for a mental health condition of 3.93 (95% CI, 3.8-4.1), compared to those with no physical conditions. Multimorbidity prevalence and patterns in the Danish population are comparable to those of other European populations. The high prevalence of mental and physical health conditions highlights the need to ensure that healthcare systems deliver care that takes physical and mental comorbidity into account. Further, the higher prevalence of multimorbidity among persons with low educational attainment emphasizes the importance of having a health care system providing care that is beneficial to all regardless of socioeconomic status.
Sparse supervised principal component analysis (SSPCA) for dimension reduction and variable selection

Principal component analysis (PCA) is one of the main unsupervised pre-processing methods for dimension reduction. When the training labels are available, it is worth using a supervised PCA strategy. In cases that both dimension reduction and variable selection are required, sparse PCA (SPCA) methods are preferred. In this paper, a sparse supervised PCA (SSPCA) method is proposed for pre-processing. This method is appropriate especially in problems where, a high dimensional input necessitates the use of a sparse method and a target label is also available to guide the variable selection strategy. Such a method is valuable in many Engineering and scientific problems, when the number of training samples is also limited. The Hilbert Schmidt Independence Criteria (HSIC) is used to form an objective based on minimization of a loss function and an L1 norm is used for regularization of the Eigen vectors. While the proposed objective function allows a sparse low rank solution for both linear and non-linear relationships between the input and response matrices, other similar methods in this case are only based on a linear model. The objective is solved based on penalized matrix decomposition (PMD) algorithm. We compare the proposed method with PCA, PMD-based SPCA and supervised PCA. In addition, SSPCA is also compared with sparse partial least squares (SPLS), due to the similarity between the two objective functions. Experimental results from the simulated as well as real data sets show that, SSPCA provides an appropriate trade-off between accuracy and sparsity. Comparisons show that, in terms of sparsity, SSPCA performs the highest level of variable reduction and also, in terms of accuracy it is one of the most successful methods. Therefore, the Eigen vectors found by SSPCA can be used for feature selection in various high dimensional problems. © 2017 Elsevier Ltd. All rights reserved.

Spatial models for probabilistic prediction of wind power with application to annual-average and high temporal resolution data

Producing accurate spatial predictions for wind power generation together with a quantification of uncertainties is required to plan and design optimal networks of wind farms. Toward this aim, we propose spatial models for predicting wind power generation at two different time scales: for annual average wind power generation, and for a high temporal resolution (typically wind power averages over 15-min time steps). In both cases, we use a spatial hierarchical statistical model in which spatial correlation is captured by a latent Gaussian field. We explore how such models can be handled with stochastic partial differential approximations of Matérn Gaussian fields together with Integrated Nested Laplace Approximations. We demonstrate the proposed methods on wind farm data from Western Denmark, and compare the results to those obtained with standard geostatistical methods. The results show that our method makes it possible to obtain fast and accurate predictions from posterior marginals for wind power generation. The proposed method is applicable in scientific areas as diverse as climatology, environmental sciences, earth sciences and epidemiology.
Statistical modelling of space-time processes with application to wind power.  
Short-term wind power forecasts together with a quantification of uncertainties are required for the reliable operation of power systems with significant wind power penetration. A challenge for utilizing wind power as a source of energy is the intermittent and hardly predictable nature of wind. This thesis aims at contributing to the wind power literature by building and evaluating new statistical techniques for producing forecasts at multiple locations and lead times using spatio-temporal information. By exploring the features of a rich portfolio of wind farms in western Denmark, we investigate different types of models and provide several forms of predictions. Starting with spatial prediction, we then extend the methodology to spatio-temporal prediction of individual wind farms and aggregated wind power at monitored locations as well as at locations where recent observations are not available. We propose spatial models for predicting wind power generation at two different time scales: for annual average wind power generation and for a high temporal resolution (typically wind power averages over 15-min time steps). In both cases, we use a spatial hierarchical statistical model in which spatial correlation is captured by a latent Gaussian field. We explore how such models can be handled with stochastic partial differential approximations of Matérn Gaussian fields together with integrated nested Laplace approximations. We show that complex hierarchical spatial models are well suited for wind power data and provide results in reasonable computational time. Moreover, the hierarchical approach for obtaining predictions at a high temporal resolution is found to produce accurate predictions with improved performance compared to a standard geostatistical method at a small additional computational cost. The use of the integrated nested Laplace approximations is motivated by the desire to produce forecasts on large data sets with hundreds of locations, which is critical during periods of high wind penetration. Subsequently, the extension from spatial to spatio-temporal models is given. Three different hierarchical models are developed for obtaining probabilistic wind power forecasts. First, a time series model consisting of an autoregressive process with a location specific intercept is considered. This approach gives satisfactory results for individual forecasts but fails to generate calibrated aggregated forecasts. The second approach has a common intercept for all farms and a spatio-temporal model that varies in time with first order autoregressive dynamics and has spatially correlated innovations given by a zero mean Gaussian process. The third model, which also has a common intercept as well as an autoregressive process to capture the local variability and the spatio-temporal term from the second approach, is able to produce reliable individual and aggregated forecasts for multiple lead times. Finally, very-short-term wind power forecasting is considered. Probabilistic forecasts from 15 minutes up to two hours ahead are produced by using anisotropic spatio-temporal correlation models to account for the propagation of weather fronts and a transformed latent Gaussian field is used to accommodate the probability masses that occur in wind power distribution due to chains of zero measurements. Using what is called kriging equations, even the simplest proposed covariance model is able to produce calibrated spatio-temporal predictions of wind power production.
Stochastic Greybox Modeling of Slugging Flows

General information
Publication status: Published
Organisations: Dynamical Systems, Department of Applied Mathematics and Computer Science, Centre for oil and gas – DTU, Statistics and Data Analysis
Contributors: Goranovic, G., Møller, J. K., Jørgensen, T. M., Madsen, H.
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from Danish Hydrocarbon Research and Technology Centre Technology Conference 2017, Lyngby, Denmark.
Electronic versions:
Pages_from_DHRTC_Poster_abstracts_2017_at_registrere_31.pdf
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2018 › Research › peer-review

Svigt af inertgas purge medførte eksplosion i beholder

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis
Contributors: Hedlund, F. H.
Pages: 12-15
Publication date: 2017
Peer-reviewed: Unknown

Publication information
Journal: Dansk Kemi
Volume: 98
Issue number: 1/2
ISSN (Print): 0011-6335
Original language: Danish
Electronic versions:
2017_01_DKemi_Svigt_af_inertgas_purge.pdf
Source: PublicationPreSubmission
Source-ID: 128742996
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Communication

Tank für Überschusshefe versagt – mit katastrophalen Folgen
Ein großer Tank für Überschusshefe flog in die Luft und nur noch die Bodenplatte und der Tankinhalt blieben zurück. Obwohl der Tank nicht für Überdruck ausgelegt war, wurde er auf "sehr geringem Überdruck" gehalten, um störende Schaumentwicklung zu verhindern. Der Brauerei waren die Gefahren durch Druckluft nicht bekannt. Der in diesem Artikel beschriebene Unfall soll aufzeigen, dass Vorsicht geboten ist, wenn ein ursprünglich für atmosphärischen Druck ausgelegter Tank auf Betrieb mit leichtem Überdruck abgeändert wird

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, COWI AS
Contributors: Hedlund, F. H., Selig, R. S.
Pages: 229-232
Publication date: 2017
Trellis plots as visual aids for analyzing split plot experiments

The analysis of split plot experiments can be challenging due to a complicated error structure resulting from restrictions on complete randomization. Similarly, standard visualization methods do not provide the insight practitioners desire to understand the data, think of explanations, generate hypotheses, build models, or decide on next steps. This article demonstrates the effective use of trellis plots in the preliminary data analysis for split plot experiments to address this problem. Trellis displays help to visualize multivariate data by allowing for conditioning in a general way. They can also be used after the statistical analysis for verification, clarification, and communication.

Unidirectional Fibre Composite Characterisation from X-ray Tomography

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Department of Wind Energy, Composites Mechanics and Materials Mechanics, Statistics and Data Analysis, University of Manchester
Number of pages: 1
Publication date: 2017
Peer-reviewed: No
Event: Poster session presented at TMS 2017, San Diego, United States.
Electronic versions: posterTMSconference_monj_final.pdf
Source: PublicationPreSubmission
Source-ID: 130858539
Unraveling fermentation data – a Novozymes case study

Industrial fermentation processes are monitored using a variety of sensors. Typically, measurements are taken throughout the entire production process. Production may be carried out under supervision of different operators (operator variation), on different sites (global variation), in different buildings and/or in different tanks (local variation). However, up to now processes are mainly controlled according to traditional recipes and experience.

General information
Publication status: Published
Contributors: Baum, A., Vermue, L., Moiseyenko, R., Jørgensen, T. M., Devantier, R.
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from Recent Advances in Fermentation Technology (RAFT 2017), Florida, United States.
Source: PublicationPreSubmission
Source-ID: 138594282
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2017 › Research › peer-review

剩余酵母罐出现故障-灾难性的后果
建议慎重：一个盛放剩余酵母的大罐子飞向空中，只剩下底板和罐子中的内容物质，虽然罐子不是针对过压设计的，但它也处于“非常低的过压”下，以防形成过多的泡沫。啤酒厂不清楚压缩空气的危害，本文描述的事故应该告诉人们，如果原本针对大气压力设计的罐子改用在轻微过压环境中，必须特别谨慎。

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, COWI AS
Contributors: Hedlund, F. H., Selig, R. S.
Pages: 9-11
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Brauwelt International
Volume: 2
ISSN (Print): 0934-9340
Original language: Chinese
Source: PublicationPreSubmission
Source-ID: 131572047
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Research › peer-review

An optimized outlier detection algorithm for jury-based grading of engineering design projects

This work characterizes and optimizes an outlier detection algorithm to identify potentially invalid scores produced by jury members while grading engineering design projects. The paper describes the original algorithm and the associated adjudication process in detail. The impact of the various conditions in the algorithm on the false positive and false negative rates is explored. A response surface design is performed to optimize the algorithm using a data set from Fall 2010. Finally, the results are tested against a data set from Fall 2011. It is shown that all elements of the original algorithm (the base rule and the three additional conditions) play a role in the algorithm's performance and should be included in the algorithm. Because there is significant interaction between the base rule and the additional conditions, many acceptable combinations that balance the FPR and FNR can be found, but no true optimum seems to exist. The performance of the best optimizations and the original algorithm are similar. Therefore, it should be possible to choose new coefficient values for jury populations in other cultures and contexts logically and empirically without a full optimization as long as the algorithm assumptions are valid and the limitations for its use are well understood.

General information
Publication status: Published
Organisations: Department of Mechanical Engineering, Manufacturing Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Technical University of Denmark
Contributors: Thompson, M. K., Espensen, C., Clemmensen, L. K. H.
Pages: 172-184
Automatic measurement of orbital volume in unilateral coronal synostosis

Premature fusion of the coronal suture on one side of the calvaria (unilateral coronal synostosis, UCS) results in asymmetric craniofacial development and the deformation of the orbits. Often this necessitates surgery, where CT scanning is employed to obtain measures of the bony orbit. These measures are typically computed by guided procedures that require expert time. We propose a method with higher degree of automation based on finding an optimal smooth closed surface. CT scans of 17 infants with UCS are included in our experimental validation, where we compare our method to expert guided segmentations. We obtain similar measures, as well as high Dice scores, compared to the experts. The run time for the proposed approach with a prototype implementation is around 3 minutes on a standard laptop, making the method suitable for rapid evaluation of orbital volume in UCS.
Causal Analysis of Railway Running Delays

Operating delays and network propagation are inherent characteristics of railway operations. These are traditionally reduced by provision of time supplements or "slack" in railway timetables and operating plans. Supplement allocation policies must trade off reliability in the service commitments against service transit times and railway asset productivity. Methods to investigate the quality of supplement time allocation are necessary to reduce the behavioral response and the waste of resources. This is a preliminary study that investigates train delay data from the year 2014 supplied by Rail Net Denmark (the Danish infrastructure manager). The statistical analysis of the data identifies the minimum running times and the scheduled running time supplements and investigates the evolution of train delays along given train paths. An improved allocation of time supplements would result in smaller overall aggregate timetable supplement, reduced transport travel times, and higher productive utilization of train rolling stock. The study results will lead eventually to both better allocation of time supplements in timetable structures, and identification of areas that should be a high priority for correction.

General information
Publication status: Published
Organisations: Traffic modelling and planning, Department of Management Engineering, Department of Transport, Management Science, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Cerreto, F., Nielsen, O. A., Harrod, S., Nielsen, B. F.
Number of pages: 7
Publication date: 2016
Peer-reviewed: Yes
Event: Paper presented at 11th World Congress on Railway Research (WCRR 2016), Milan, Italy.
Electronic versions: 790.pdf
Research output: Contribution to conference ★ Paper – Annual report year: 2016 ★ Research ★ peer-review

Changes in sensory characteristics and their relation with consumers' liking, wanting and sensory satisfaction: Using dietary fibre and lime flavour in Stevia rebaudiana sweetened fruit beverages

The beverage industry has long revolved around sugar reduction as a response to heightened calorie and health awareness. More recently dietary fibre has also garnered attention to meet the consumer's demands for low calorie and yet more satiating food and beverages. From a health perspective the sweetener Stevia rebaudiana and the fibre β-glucans seem like very good solutions, as stevia is the only natural non-nutritive sweetener and β-glucans have been related to various health benefits besides increasing satiety. However, both also have distinctive perceptual effects on the sensory characteristics of the products they are added to. To gain knowledge on the sensory characteristics of fruit based beverages sweetened with S. rebaudiana and added β-glucans and lime flavour, and how consumers respond to the products, sensory descriptive analysis and a consumer study were conducted. The sensory characteristics of the fruit drinks were affected by stevia and the addition of β-glucans. However, the addition of lime flavour was able to mask the side effect of the aftertaste caused by S. rebaudiana. Further, by adding lime flavour to the fruit beverages, the side effects of increased fibre concentration "Unfresh odour" and "Metallic odour" could also be counteracted to such an extent that the β-glucans containing fruit beverages were evaluated just as favourably in terms of liking, wanting and sensory satisfying as the fruit beverages not containing β-glucans.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University
Contributors: Mielby, L. H., Andersen, B. V., Jensen, S., Kildegaard, H., Kuznetsova, A., Eggers, N., Brockhoff, P. B., Byrne, D. V.
Pages: 14-21
Concurrent elevation of CO₂, O₃ and temperature severely affects oil quality and quantity in rapeseed

Plant oil is an essential dietary and bio-energy resource. Despite this, the effects of climate change on plant oil quality remain to be elucidated. The present study is the first to show changes in oil quality and quantity of four rapeseed cultivars in climate scenarios with elevated [CO₂], [O₃] and temperature (T) combined and as single factors. The combination of environmental factors resembled IPCC's 'business as usual' emission scenario predicted for late this century. Generally, the climate scenarios reduced the average amounts of the six fatty acids (FAs) analysed, though in some treatments single FAs remained unchanged or even increased. Most reduced was the FA essential for human nutrition, C18:3-ω3, which decreased by 39% and 45% in the combined scenarios with elevated [CO₂]+T+[O₃] and [CO₂]+T, respectively. Average oil content decreased 3–17%. When [CO₂] and T were elevated concurrently, the seed biomass was reduced by half, doubling the losses in FAs and oil content. This corresponded to a 58% reduction in the oil yield per hectare, and C18:3-ω3 decreased by 77%. Furthermore, the polyunsaturated FAs were significantly decreased. The results indicate undesirable consequences for production and health benefits of rapeseed oil with future climate change. The results also showed strong interactive effects of CO₂, T and O₃ on oil quality, demonstrating why prediction of climate effects requires experiments with combined factors and should not be based on extrapolation from single factor experiments.
Correlations between fatigue and disease duration, disease activity, and pain in patients with rheumatoid arthritis: a systematic review

OBJECTIVES: Rheumatoid arthritis (RA) patients suffer from disabling fatigue but the causes of this condition are unknown. Our aim was to assess which of the variables disease activity, disease duration, and pain is associated with fatigue.

METHOD: We conducted a systematic literature search in MEDLINE and EMBASE, followed by selection of studies according to set criteria, data extraction, and statistical analyses of the relationships in RA between fatigue and the following covariates: disease duration, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), the 28-joint Disease Activity Score (DAS28), swollen to tender joint count ratio (STR), and pain. Linear regression analyses of fatigue regressed on each of the six covariates, and a multiple regression analysis where fatigue was regressed on the six covariates through a forward selection procedure was carried out with construction of correlation measures between fatigue and the covariates.

RESULTS: A total of 121 studies were included in the analyses, including > 100 000 RA patients. A high level of fatigue was seen even in well-treated patients, demonstrating fatigue as a major problem in RA. Fatigue was found to be positively correlated with pain, CRP, DAS28, and ESR but not with the STR or disease duration, with pain as the overall domineering factor.

CONCLUSIONS: Fatigue has a substantial influence on the lives of RA patients, independent of disease duration. Pain is the domineering factor in the experience and degree of fatigue. Disease activity is positively correlated to fatigue but does not contribute substantially when pain is considered. Optimal pain relief is therefore an important part of the treatment to improve fatigue in RA.
mixed models, where in general the relevant error terms for the fixed effects are not the pure residual error, it is suggested to base the d-prime-like interpretation on the residual error. The methods are illustrated on a multifactorial sensory profile data set and compared to actual d-prime calculations based on Thurstonian regression modeling through the ordinal package. For more challenging cases we offer a generic "plug-in" implementation of a version of the method as part of the R-package SensMixed. We discuss and clarify the bias mechanisms inherently challenging effect size measure estimates in ANOVA settings.

Determinants of meal satisfaction in a workplace environment
Workplace lunches are recurrent meal occasions that can contribute to the general well-being of employees. The objective of our research was to study which factors influence consumers' satisfaction with these meals by exploring the relative role of food-related, personal, situational factors. Using a longitudinal approach, we monitored a total of 71 participants compiled and experienced 519 meals from their workplace canteen buffet during a three-month period; in addition the composed lunches were photographed. Before and after the lunch choice period respondents filled in a questionnaire on several meal-related variables. A mixed modelling approach was used to analyse the data. Meal satisfaction was directly associated with a positive ambience and a positive evaluation of both the quality of the food eaten and the buffet assortment, whereas the meal's energy content did not contribute to meal satisfaction. Additionally, meal satisfaction was associated with a more positive mood, lower hunger level as well as feeling less busy and stressed after lunch. The buffet assortment, a more positive mood before lunch and mindful eating contributed to the perceived food quality, but not associated with the hunger level before lunch. Time available, mindful eating and eating with close colleagues were positively associated with perceived ambience. The results indicate that consumers' satisfaction with workplace meals can be increased by putting emphasis on the quality of food served, but equally important is the ambience in the lunch situation. Most of the ambience factors were related to available time and mental resources of the participants and the possibility to share the meal with close colleagues. These are factors that can be facilitated by the service provider, but not directly influenced.
Does correlated color temperature affect the ability of humans to identify veins?

In the present study we provide empirical evidence and demonstrate statistically that white illumination settings can affect the human ability to identify veins in the inner hand vasculature. A special light-emitting diode lamp with high color rendering index (CRI 84–95) was developed and the effect of correlated color temperature was evaluated, in the range between 2600 and 5700 K at an illuminance of 400 lx on the ability of adult humans to identify veins. It is shown that the ability to identify veins can, on average, be increased up to 24% when white illumination settings that do not resemble incandescent light are applied. The illuminance reported together with the effect of white illumination settings on direct visual perception of biosamples are relevant for clinical investigations during the night. © 2015 Optical Society of America

General information
Publication status: Published
Organisations: Department of Photonics Engineering, Diode Lasers and LED Systems, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Argyraki, A., Clemmensen, L. K. H., Petersen, P. M.
Pages: 141-148
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Journal of the Optical Society of America A
Volume: 33
Issue number: 1
ISSN (Print): 0740-3232
Ratings:
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 1.54
Web of Science (2016): Impact factor 1.621
Web of Science (2016): Indexed yes
Original language: English
Electronic versions:
Josa33_1_141_distribute_version.pdf
DOIs:
10.1364/JOSAA.33.000141
Source: PublicationPreSubmission
Source-ID: 119459629
Research output: Contribution to journal › Journal article – Annual report year: 2016 › Research › peer-review

Er sprængpladen vendt korrekt?

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, COWI AS
Contributors: Hedlund, F. H.
Pages: 41-42
Publication date: 2016
Peer-reviewed: No

Publication information
Journal: Dansk Kemi
Evaluation of temporal surveillance system sensitivity and freedom from bovine viral diarrhea in Danish dairy herds using scenario tree modelling

The temporal sensitivity of the surveillance system (TemSSe) for Bovine Viral Diarrhea (BVD) in Danish dairy herds was evaluated. Currently, the Danish antibody blocking ELISA is used to test quarterly bulk tank milk (BTM). To optimize the surveillance system as an early warning system, we considered the possibility of using the SVANOVIR ELISA, as this test has been shown to detect BVD-positive herds earlier than the blocking ELISA in BTM tests. Information from data (2010) and outputs from two published stochastic models were fed into a stochastic scenario tree to estimate the TemSSe. For that purpose we considered: the risk of BVD introduction into the dairy population, the ELISA used and the high risk period (HRP) from BVD introduction to testing (at 90 or 365 days). The effect of introducing one persistently infected (PI) calf or one transiently infected (TI) milking cow into 1 (or 8) dairy herd(s) was investigated. Additionally we estimated the confidence in low (PLow) herd prevalence.
on microbiological feed quality was obtained from the Voluntary Feed Control under the Mink producers Organization, and data on herd size and the relation between farm and feed producer was obtained from the registers at Kopenhagen Fur, based on yearly reporting from the mink producers. Descriptive analysis showed a clear significant effect of season on antimicrobial use, with a peak in "treatment proportions", TP (defined daily doses per kg biomass-days) in May, around the time of whelping, and a high level in the following months. In autumn, a minor peak in antimicrobial use occurred throughout the study period. From 2007 to 2011, a 102% increase in annual antimicrobial TP was noted; on herd level, the increase was associated with an increasing frequency of prescription, and a decrease in the amounts prescribed in months with prescription. A binomial model showed that on herd level, the annual number of months with antimicrobial prescription was significantly (p<0.01) affected by feed producer, veterinarian, disease (specific laboratory diagnosis) infection, herd size and year, with an interaction between feed producer and year. A log-normal model showed that in months with antimicrobial use, the TP on herd level was significantly (p<0.001) affected by year, month (season), feed producer, feed quality score, veterinarian, herd size and laboratory confirmed diagnosis of specific infections; additionally the interaction terms year×feed producer and herd size×month were significant (p<0.001). In conclusion, antimicrobial use on herd level was significantly associated with the microbiological food quality, the feed producer, and the veterinarian. The prescription patterns varied significantly between veterinarians, and some veterinarians were associated with both larger and more frequent prescriptions of antimicrobials at herd level. Herd size is associated with different prescription patterns. Finally, infection with Pseudomonas aeruginosa, astrovirus, influenza virus and Salmonella spp. was associated with an increase in antimicrobial use.

Fatal Accidents During Marine Transport of Wood Pellets Due to Off-gassing – Experiences from Denmark

The atmosphere in unventilated wood pellet storage confinements, such as the cargo hold of marine vessels transporting pellets in solid bulk, can be severely oxygen deficient and contain deadly concentrations of harmful gasses, of which the most feared is the poisonous and odour-less carbon monoxide. The hazard has been known for over a decade and has been responsible for many accidents. We examine three fatal accidents on marine vessels in or near Danish waters and argue that they share strikingly similar aetiologies, if not repetitive patterns. It is generally recognized that accidents should be thoroughly investigated and lessons learned shared widely in order to minimize the number of times the same lessons have to be learned. The three Danish cases suggest that this learning process is deeply troubled for the solid biomass segment. The International Maritime Organization IMO/SOLAS has recently revised its guidance on entering enclosed spaces aboard ships in response to the ongoing problem of confined space incidents. We argue that the interpretation of the concept of an "enclosed space" is of utmost importance because accidents take place in rooms that are not considered enclosed by the crew.

General information
Publication status: Published
Organisations: National Veterinary Institute, Section for Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Kopenhagen Fur
Contributors: Jensen, V. F., Sommer, H. M., Struve, T., Clausen, J., Chriél, M.
Pages: 170-182
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Preventive Veterinary Medicine
Volume: 126
ISSN (Print): 0167-5877
Ratings:
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.2 SJR 1.249 SNIP 1.361
Web of Science (2016): Impact factor 1.987
Web of Science (2016): Indexed yes
Original language: English
Keywords: Antimicrobials, Mink, Feed quality, Risk factor
Electronic versions:
1_s2.0_S016758771630040X_main.pdf
DOIs: 10.1016/j.prevetmed.2016.01.023
Source: FindIt
Source-ID: 2291755559
Research output: Contribution to journal › Journal article – Annual report year: 2016 › Research › peer-review
From university research to innovation Detecting knowledge transfer via text mining

Knowledge transfer by universities is a top priority in innovation policy and a primary purpose for public research funding, due to being an important driver of technical change and innovation. Current empirical research on the impact of university research relies mainly on formal databases and indicators such as patents, collaborative publications and license agreements, to assess the contribution to the socioeconomic surrounding of universities. In this study, we present an extension of the current empirical framework by applying new computational methods, namely text mining and pattern recognition. Text samples for this purpose can include files containing social media contents, company websites and annual reports. The empirical focus in the present study is on the technical sciences and in particular on the case of the Technical University of Denmark (DTU). We generated two independent text collections (corpora) to identify correlations of university publications and company webpages. One corpus representing the company sites, serving as sample of the private economy and a second corpus, providing the reference to the university research, containing relevant publications. We associated the former with the latter to obtain insights into possible text and semantic relatedness. The text mining methods are extrapolating the correlations, semantic patterns and content comparison of the two corpora to define the document relatedness. We expect the development of a novel tool using contemporary techniques for the measurement of public research impact. The approach aims to be applicable across universities and thus enable a more holistic comparable assessment. This rely less on formal databases, which is certainly beneficial in terms of the data reliability. We seek to provide a supplementary perspective for the detection of the dissemination of university research and hereby enable policy makers to gain additional insights of (informal) contributions of knowledge dissemination by universities.
three fast food restaurants from different parts of Denmark among randomly selected customers (aged 15 or above). The customers were approached after having ordered their meal. They filled out a questionnaire on reasons for their actual fast food meal selection and purchase intentions in relation to four hypothesized burger menus, including a regular beef burger menu, a wholegrain beef burger menu, a nutrition labeled beef burger menu and a nutrition labeled chicken burger menu. Results showed that the majority of the fast food customers expressed a wish for healthier menus (55% males vs. 64% females agree or strongly agree, p < 0.001) and more sustainable menus in terms of environmental impact (43% males vs. 52% females agree or strongly agree, p < 0.001), however only 7% of the participants’ meals included healthier food choices (n = 740). Habits, taste and price were the main drivers among both genders for the actual meal selection. Compared with women, more men expressed that actual food choice was based on offers and promotions (p < 0.001), and on food perceived as the most satiating (p = 0.001). With regard to purchase intentions, the majority of men preferred a beef burger menu (healthier or regular) over a healthier chicken burger menu or a wholegrain burger menu, whereas the majority of women responded positively to either of the healthier-labeled burger menus (p < 0.001). In conclusion, the study shows that having a focus on gender differences is of particular importance in order to improve the food nutrition environment and support healthier food selections among fast food customers.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Research group for Risk Benefit, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Number of pages: 7
Pages: 123-129
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Food Quality and Preference
Volume: 47
ISSN (Print): 0950-3293
Ratings:
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 4.21 SJR 1.17 SNIP 1.681
Web of Science (2016): Impact factor 3.199
Web of Science (2016): Indexed yes
Original language: English
Keywords: Gender differences, Food environment, Health promotion, Food choice, Public private partnership
DOI:
10.1016/j.foodqual.2015.06.011
Source: PublicationPreSubmission
Source-ID: 113064419
Research output: Contribution to journal › Journal article – Annual report year: 2016 › Research › peer-review

Grain protein concentration and harvestable protein under future climate conditions. A study of 108 spring barley accessions
In the present study a set of 108 spring barley (H. vulgare L.) accessions were cultivated under predicted future levels of temperature and [CO2] as single factors and in combination (IPCC, AR5, RCP8.5). Across all genotypes, elevated [CO2] (700 ppm day/night) slightly decreased protein concentration by 5%, while elevated temperature (+5 °C day/night) substantially increased protein concentration by 29%. The combined treatment increased protein concentration across accessions by 8%. This was an increase less than predicted from strictly additive effects of the individual treatments. Despite the increase in grain protein concentration, the decrease in grain yield at combined elevated temperature and elevated [CO2] resulted in 23% less harvestable protein. There was variation in the response of the 108 accessions, which might be exploited to at least maintain if not increase harvestable grain protein under future climate change conditions.

General information
Publication status: Published
Organisations: Department of Chemical and Biochemical Engineering, Department of Environmental Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University
Number of pages: 8
Pages: 2151-2158
Publication date: 2016
Peer-reviewed: Yes

Publication information
Grating-based X-ray tomography of 3D food structures
A novel grating-based X-ray phase-contrast tomographic method has been used to study how partly substitution of meat proteins with two different types of soy proteins affect the structure of the formed protein gel in meat emulsions. The measurements were performed at the Swiss synchrotron radiation light source using a grating interferometric set-up.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Contributors: Miklos, R., Nielsen, M. S., Einarsdottir, H., Lametsch, R.
Pages: 1-5
Publication date: 2016

Host publication information
Title of host publication: Proceedings of the 19th International ESAFORM Conference on Material Forming (ESAFORM 2016)
Publisher: American Institute of Physics
Article number: 180001
Electronic versions:
1.4963604.pdf. Embargo ended: 01/11/2017
DOIs:
10.1063/1.4963604
Source: Findit
Source-ID: 2347937524
Research output: Chapter in Book/Report/Conference proceeding -> Article in proceedings – Annual report year: 2016 -> Research -> peer-review

Har tanken et svagt tag?

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Statistics and Data Analysis, COWI AS
Contributors: Hedlund, F. H., Eisfeld Linde, B.
Pages: 20-21
Publication date: 2016
Peer-reviewed: No

Publication information
Journal: Dansk Kemi
Volume: 97
High-resolution rainfall time series for future climate

General information
Publication status: Published
Organisations: Department of Environmental Engineering, Urban Water Systems, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Sørup, H. J. D., Georgiadis, S., Gregersen, I. B., Arnbjerg-Nielsen, K.
Pages: 42-42
Publication date: 2016

Host publication information
Title of host publication: 4th Nordic Conference on Climate Change Adaptation: Book of Abstracts
Place of publication: Bergen, Norway
Article number: 53

Identifying Drug–Drug Interactions by Data Mining: A Pilot Study of Warfarin-Associated Drug Interactions
Background—Knowledge about drug–drug interactions commonly arises from preclinical trials, from adverse drug reports, or based on knowledge of mechanisms of action. Our aim was to investigate whether drug–drug interactions were discoverable without prior hypotheses using data mining. We focused on warfarin–drug interactions as the prototype.

Methods and Results—We analyzed altered prothrombin time (measured as international normalized ratio [INR]) after initiation of a novel prescription in previously INR-stable warfarin-treated patients with nonvalvular atrial fibrillation. Data sets were retrieved from clinical work. Random forest (a machine-learning method) was set up to predict altered INR levels after novel prescriptions. The most important drug groups from the analysis were further investigated using logistic regression in a new data set. Two hundred and twenty drug groups were analyzed in 61,190 novel prescriptions. We rediscovered 2 drug groups having known interactions (β-lactamase-resistant penicillins [dicloxacillin] and carboxamide derivatives) and 3 antithrombotic/anticoagulant agents (platelet aggregation inhibitors excluding heparin, direct thrombin inhibitors [dabigatran etexilate], and heparins) causing decreasing INR. Six drug groups with known interactions were rediscovered causing increasing INR (antiarrhythmics class III [amiodarone], other opioids [tramadol], glucocorticoids, triazole derivatives, and combinations of penicillins, including β-lactamase inhibitors) and two had a known interaction in a closely related drug group (oripavine derivatives [buprenorphine] and natural opium alkaloids). Antipropulsives had an unknown signal of increasing INR.

Conclusions—We were able to identify known warfarin–drug interactions without a prior hypothesis using clinical registries. Additionally, we discovered a few potentially novel interactions. This opens up for the use of data mining to discover unknown drug–drug interactions in cardiovascular medicine.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Image Analysis for X-ray Imaging of Food

X-ray imaging systems are increasingly used for quality and safety evaluation both within food science and production. They offer non-invasive and nondestructive penetration capabilities to image the inside of food.

This thesis presents applications of a novel grating-based X-ray imaging technique for quality and safety evaluation of food products. In this effort the fields of statistics, image analysis and statistical learning are combined, to provide analytical tools for determining the aforementioned food traits.

The work demonstrated includes a quantitative analysis of heat induced changes in microstructure of meat products. A segmentation framework is presented, from which geometrical parameters are assessed. The grating-based method embraces the complicated microstructure of the meat products, allowing for an analysis of the full three dimensional structure. The results illustrate that the combination of grating-based X-ray imaging and advanced analysis provides a valuable tool for microstructure analysis. Thus, the method can be considered as an alternative to other existing imaging techniques.

Furthermore, the thesis presents the application of grating-based X-ray imaging for novelty and defect detection in food. Compared to the complex three dimensional analysis of microstructure, here two dimensional images are considered, making the method applicable for an industrial setting. The advantages obtained by grating-based imaging are compared to conventional X-ray imaging, for both foreign object and defect detection. The results further emphasize the applicability of grating-based imaging for evaluation of food quality and food safety.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Einarsdottir, H.
Number of pages: 142
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
(DTU Compute PHD-2016; No. 416).
Electronic versions:
Thesis.pdf

Kronik: Virksomheder med elitesmiley flyver under radaren

General information
Publication status: Published
Large Steel Tank Fails and Rockets to Height of 30 meters - Rupture Disc Installed Incorrectly

At a brewery, the base plate-to-shell weld seam of a 90-m³ vertical cylindrical steel tank failed catastrophically. The 4 ton tank "took off" like a rocket leaving its contents behind, and landed on a van, crushing it. The top of the tank reached a height of 30 m. The internal overpressure responsible for the failure was an estimated 60 kPa. A rupture disc rated at <50 kPa provided overpressure protection and thus prevented the tank from being covered by the European Pressure Equipment Directive. This safeguard failed and it was later discovered that the rupture disc had been installed upside down. The organizational root cause of this incident may be a fundamental lack of appreciation of the hazards of large volumes of low-pressure compressed air or gas. A contributing factor may be that the standard piping and instrumentation diagram (P&ID) symbol for a rupture disc may confuse and lead to incorrect installation. Compressed air systems are ubiquitous. The medium is not toxic or flammable. Such systems however, when operated at "slight overpressure" can store a great deal of energy and thus constitute a hazard that ought to be addressed by safety managers.
Layered Surface Detection in Micro-CT Tetra Pak Data

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Physics, Neutrons and X-rays for Materials Physics, Image Analysis & Computer Graphics, Tetra Pak Packaging Solutions AB
Number of pages: 3
Publication date: 2016
Peer-reviewed: Yes
Electronic versions:
TetraPack_abstract.pdf
Source: PublicationPreSubmission
Source-ID: 123970129
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2016 › Research › peer-review

Lederundersøgelse - Lønberegning

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Thyregod, C., Ersbøll, B. K.
Number of pages: 8
Publication date: 2016
Publication information
Original language: Danish
Research output: Book/Report › Report – Annual report year: 2016 › Research

Monitoring Animal Well-being
In recent years, animal well-being in industrial slaughterhouses has become a significant concern for consumers, farmers, and meat producers. Different groups have different interpretations of animal well-being. For the majority of consumers, animal well-being is highly influenced by their values and experiences. Meat producers are interested in the stress animals endure because it affects meat quality.

Pigs that arrive at slaughterhouses are more sensitive than usual for several reasons. In some cases, pigs are transported for long distances. Not all animals are used to transportation. Upon their arrival, it is common to mix pigs from different farmers in one area. Such mixing can cause fights between pigs, which can lead to additional stress or the animals being harmed. The unfamiliar environment also increases the animals' stress levels. In some industrial slaughterhouses, up to 62,000 pigs per week are handled. Ensuring the well-being of such large numbers of pigs using only personnel is a complicated task.

Video surveillance of humans has been widely used to ensure safety and order in multiple situations. Methods have been developed to detect individual actions or abnormal behavior in small groups and dense crowds. In recent years, surveillance has also been used to monitor animals. Research has mainly focused on monitoring laboratory animals and farm animals. In both cases, animals are usually in constrained environment and cameras are used to cover all areas where animals are present. To obtain better results, non-intrusive markers or extracted features are used for tracking. Laboratory environments can be highly controlled; thus, no light and shadow noise are present in videos.

In slaughterhouses, the main focus is on monitoring large groups of animals in locations where additional markers cannot be used and pigs can leave or enter the surveilled area. In addition, pigs have a specific walking pattern; thus, motion analysis is not straightforward. The first aim of this thesis is to monitor the movement of pigs without using any additional markers or feature extraction in an unconstrained environment.

In video surveillance, the behavior of humans and animals is monitored based on extremes: event is present/event is not present, objects behave normally/objects behave abnormally, action 1/actions 2/actions 3, etc. In nature, the motion of humans and animals is continuous with transitions from one action to another. The second aim of this thesis is to propose a method to monitor motion as a continuous process using common classification methods.
Monitoring pig movement at the slaughterhouse using optical flow and modified angular histograms

We analyse the movement of pig herds through video recordings at a slaughterhouse by using statistical analysis of optical flow (OF) patterns. Unlike the previous attempts to analyse pig movement, no markers, trackers nor identification of individual pigs are needed. Our method handles the analysis of unconstrained areas where pigs are constantly entering and leaving. The goal is to improve animal welfare by real-time prediction of abnormal behaviour through proper interventions. The aim of this study is to identify any stationary pig, which can be an indicator of an injury or an obstacle. In this study, we use the OF vectors to describe points of movement on all pigs and thereby analyse the herd movement. Subsequently, the OF vectors are used to identify abnormal movements of individual pigs. The OF vectors, obtained from the pigs, point in multiple directions rather than in one movement direction. To accommodate the multiple directions of the OF vectors, we propose to quantify OF using a summation of the vectors into bins according to their angles, which we call modified angular histograms. Sequential feature selection is used to select angle ranges, which identify pigs that are moving abnormally in the herd. The vector lengths from the selected angle ranges are compared to the corresponding median, 25th and 75th percentiles from a training set, which contains only normally moving pigs. We show that the method is capable of locating stationary pigs in the recordings regardless of the number of pigs in the frame.
Multivariate Time Series Estimation using marima

A computer program, called marima, written in the open source language, R, has been developed. Some of marima’s facilities and ideas are presented in the following.

Muscle wound healing in rainbow trout (Oncorhynchus mykiss)

We followed the progression of healing of deep excisional biopsy punch wounds over the course of 365 days in rainbow trout (Oncorhynchus mykiss) by monitoring visual wound healing and gene expression in the healing muscle at regular intervals (1, 3, 7, 14, 38 and 100 days post-wounding). In addition, we performed muscle texture analysis one year after wound infliction. The selected genes have all previously been investigated in relation to vertebrate wound healing, but only few specifically in fish. The selected genes were interleukin (IL)-1β, IL-6, transforming growth factor (TGF)-β1 and -β3, matrix metalloproteinase (MMP) -9 and -13, inducible nitric oxide synthase (iNOS), fibronectin (FN), tenascin-C (TN-C), prolyl 4-hydroxylase α1-chain (P4Hα1), lysyl oxidase (LOX), collagen type I α1-chain (ColIα1), CD41 and CD163.

Wound healing progressed slowly in the presented study, which is at least partially due to the low temperature of about 8.5 °C during the first 100 days. The inflammation phase lasted more than 14 days, and the genes relating to production and remodeling of new extracellular matrix (ECM) exhibited a delayed but prolonged upregulation starting 1-2 weeks post-wounding and lasting until at least 100 days post-wounding. The gene expression patterns and histology reveal limited capacity for muscle regeneration in rainbow trout, and muscle texture analyses one year after wound infliction confirm that wounds heal with fibrosis. At 100 dpw epidermis had fully regenerated, and dermis partially regenerated. Scales had not regenerated even after one year.

CD163 is a marker of "wound healing"-type M2c macrophages in mammals. M2 macrophage markers are as yet poorly described in fish. The pattern of CD163 expression in the present study is consistent with the expected timing of presence of M2c macrophages in the wound. CD163 may thus potentially prove a valuable marker of M2 macrophages - or a subset hereof - in fish.

We subjected a group of fish to bathing in an immunomodulatory β-glucan product during wound healing, but found this to have very limited effect on wound healing in contrast to a previously published study on common carp.
Nonparametric Estimation of Interval Reliability for Discrete-Time Semi-Markov Systems
In this article, we consider a repairable discrete-time semi-Markov system with finite state space. The measure of the interval reliability is given as the probability of the system being operational over a given finite-length time interval. A nonparametric estimator is proposed for the interval reliability, and the asymptotic properties of the strong consistency and the asymptotic normality for this estimator are proved. A numerical application concerning a four-state semi-Markov system is also presented.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Universite de Technologie de Compiegne
Contributors: Georgiadis, S., Limnios, N.
Pages: 20-39
Publication date: 2016
Peer-reviewed: Yes

Nonparametric Estimation of Interval Reliability for Discrete-Time Semi-Markov Systems
In this article, we consider a repairable discrete-time semi-Markov system with finite state space. The measure of the interval reliability is given as the probability of the system being operational over a given finite-length time interval. A nonparametric estimator is proposed for the interval reliability, and the asymptotic properties of the strong consistency and the asymptotic normality for this estimator are proved. A numerical application concerning a four-state semi-Markov system is also presented.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Universite de Technologie de Compiegne
Contributors: Georgiadis, S., Limnios, N.
Pages: 20-39
Publication date: 2016
Peer-reviewed: Yes

Novelty detection of foreign objects in food using multi-modal X-ray imaging
In this paper we demonstrate a method for novelty detection of foreign objects in food products using grating-based multimodal X-ray imaging. With this imaging technique three modalities are available with pixel correspondence, enhancing organic materials such as wood chips, insects and soft plastics not detectable by conventional X-ray absorption radiography. We conduct experiments, where several food products are imaged with common foreign objects typically found in the food processing industry. To evaluate the benefit from using this multi-contrast X-ray technique over
conventional X-ray absorption imaging, a novelty detection scheme based on well known image- and statistical analysis techniques is proposed. The results show that the presented method gives superior recognition results and highlights the advantage of grating-based imaging.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Technische Universität München
Contributors: Einarsdottir, H., Emerson, M. J., Clemmensen, L. K. H., Scherer, K., Willer, K., Bech, M., Larsen, R., Ersbøll, B. K., Pfeiffer, F.
Pages: 39-47
Publication date: 2016
Peer-reviewed: Yes

**Publication information**
Journal: Food Control
Volume: 67
Issue number: September
ISSN (Print): 0956-7135
Ratings:
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.86 SJR 1.492 SNIP 1.709
Web of Science (2016): Impact factor 3.496
Web of Science (2016): Indexed yes
Original language: English
Keywords: X-ray radiography, Dark-field imaging, Phase-contrast imaging, Foreign object detection, Novelty detection, Texture analysis
DOIs: 10.1016/j.foodcont.2016.02.023
Source: PublicationPreSubmission
Source-ID: 121892008
Research output: Contribution to journal › Journal article – Annual report year: 2016 › Research › peer-review

**Obesity Prevention in the Nordic Countries**
Previous studies have shown that mean BMI and prevalences of overweight/obesity and obesity have increased over the last decades in the Nordic countries, despite highly regulated societies with a focus on obesity prevention. We review recent overweight/obesity and obesity prevention initiatives within four of the five Nordic countries: Sweden, Denmark, Finland, and Iceland. Moreover, we analyze the current situation based on monitoring data on BMI collected in 2011 and 2014, and obtain overall estimates of overweight/obesity and obesity prevalences for the Nordic Region. Data analysis shows that obesity in adults has increased from 2011 to 2014, while no significant changes were found for children. No significant increases were found for mean BMI and overweight/obesity prevalence. Obesity prevention initiatives among the Nordic countries are highly similar although minor differences are present, which is rooted in transnational Nordic cooperation and comparable societal structures.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Division of Risk Assessment and Nutrition, Danish Health Authority
Contributors: Stockmarr, A., Hejgaard, T., Matthiessen, J.
Pages: 156-165
Publication date: 2016
Peer-reviewed: Yes

**Publication information**
Journal: Current Obesity Reports
Volume: 5
Issue number: 2
ISSN (Print): 2162-4968
Ratings:
Scopus rating (2016): CiteScore 1.41 SJR 1.215 SNIP 0.04
Web of Science (2016): Indexed yes
Original language: English
Keywords: Obesity, Overweight, Prevention, BMI, Nordic countries, Nordic model
DOIs:
Objective measures of meal variety lacking association with consumers' perception of variety with self-selected buffet meals at work

Food variety has been linked to higher diet quality and increased food intake, but what constitutes variety for consumers is underexposed. The aim of the study was twofold: first to explore the relationship between objective measures of meal variety and subjective post-meal ratings of perceived variety, and second to explore the associations between subjective meal variety and decision-making rules and individual eating styles. Data consist of 510 meals compiled from workplace lunch buffets by 71 respondents over 31 optional days. Meals were photographed and coded according to the number of components (dishes served), food groups, colours, size and shape of food on the plates. A mixed model approach was used to analyse data due to the repetitive structure of the data. Results show that subjective variety was marginally associated with the number of food groups, but there was no association with other objective measures, such as the number of components or any of the visual cues of the meal. Subjective meal variety was linked with the decision-making rule of having many dishes when compiling buffet lunches. Participants with higher scores on uncontrolled eating and food neophobia were found to perceive their meals less varied than those with lower scores. Moreover, the rule of having many dishes was positively associated with uncontrolled eating and negatively associated with cognitive restraint. Consumers' perception of within-meal variety seems to be more linked to their idea of how to compose their meal and individual tendencies towards food and eating rather than the objective measures applied.

On the need for integrating LCA into decision making

The need for sustainable solutions has gained attention both in academia and industry research due to increasing demands of human beings, which are incompatible with limitations in resources availability. Several methods, such as Life Cycle Assessment (LCA), were developed in the past decades to assess the environmental profile of products and services. However, when decision makers have several alternatives at hand to solve a problem, environmental performance is not the only criterion for choosing the best alternative. Other criteria such as risks and economical costs and benefits that are associated with the alternatives will also influence the final choice. Sometimes the most environmentally sustainable alternative may not be the safest or cheapest one. How to make a balanced decision considering environmental performance together with other criteria is not straightforward.

Decision analysis is broadly used to help decision makers identify the best solution among alternatives. The decision is based on expected utility generation, which incorporates consequences (or impacts) associated with each alternative. Depending on the research field and goal of the study, the included consequences can be e.g. environmental impacts, property damages from natural hazards and/or human health impacts. We examined the current decision analysis practice as it is applied in different research fields. The review shows that generally environmental impacts are considered less often than the other consequences. Meanwhile, LCA has been applied in many research fields to assess a wide range of
environmental impacts associated with products or services. There is a huge potential for integrating LCA into other decisions analysis tools to include assessments of the environmental profile of alternatives. This will provide the possibility of systematical inclusion of environmental considerations in the decision making process, thus facilitating a more holistic decision. However, due to different scopes and purposes of LCA and other decision analysis tools, the integration is not straightforward. The lack of consistency in e.g. system boundaries and handling of uncertainty needs to be carefully managed.

General information
Publication status: Published
Organisations: Department of Management Engineering, Quantitative Sustainability Assessment, Department of Civil Engineering, Section for Structural Engineering, Transport DTU, Transport Modelling, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Environmental Engineering, Urban Water Systems, National Food Institute, Research group for Genomic Epidemiology
Contributors: Dong, Y., Miraglia, S., Manzo, S., Georgiadis, S., Sørup, H. J. D., Boriani, E., Thöns, S., Hauschild, M. Z.
Number of pages: 1
Publication date: 2016
Peer-reviewed: Yes
URLs:
http://www.sustain.dtu.dk/

Bibliographical note
Sustain Abstract A-3
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2016 › Research › peer-review

On the use of functional calculus for phase-type and related distributions
The area of phase-type distributions is renowned for its ability to obtain closed form formulas or algorithmically exact solutions to many complex stochastic models. The method of functional calculus will provide an additional tool along these lines for establishing results in terms of functions of matrices. Functional calculus, which is a branch of operator theory frequently associated with complex analysis, can be applied to phase-type and matrix-exponential distributions in a rather straightforward way. In this article we provide a number of examples of how to execute the formal arguments.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Autonomous University of Mexico
Contributors: Bladt, M., Navarro, A. C., Nielsen, B. F.
Pages: 1-19
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Stochastic Models
Volume: 32
Issue number: 1
ISSN (Print): 1532-6349
Ratings:
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 0.52 SJR 0.356 SNIP 0.809
Web of Science (2016): Impact factor 0.38
Web of Science (2016): Indexed yes
Original language: English
Keywords: Complex analysis, Functional calculus, Matrix-exponential, Mellin transform, Phase-type distributions
DOIs:
10.1080/15326349.2015.1064773
Source: FindIt
Source-ID: 2279750686
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

Optimistic self-assessments of unhealthy diets are associated with positive indicators of health and health behaviours in Danish adults.

General information
Possibilities and barriers for e-learning in primary school in Denmark

How much are we using e-learning in primary school in Denmark? What are the barriers? What are the benefits? Why do we not use e-learning even more? These questions have been the focus for a one year national investigation conducted by ATV, The Danish Academy of Technical Sciences. The investigation included interviews with Danish researchers, persons from the ministry, teachers, students, and e-learning companies. The overall purpose was to make recommendations on how to accelerate the digital transformation of the Danish primary and high schools education system. In this paper, we combine some of the preliminary findings from the interviews with answers to a small questionnaire sent out to 19 Danish e-learning companies. We also add our own observations and visions for how e-learning could be used and which potentials we see for the near future. One major observation is that generally the attitudes to e-learning is positive: Denmark has political goals about access to machines, infrastructure, and internet at every school. There is dedicated substantial funding for schools buying e-learning material, and local successful e-learning companies such as Area9, EduLab, and Clio Online with international footprint and covering 90% of all schools in Denmark. Despite the many positive factors, which contribute to accelerating the use of e-learning, we also found others, which slow down the transformation: Lack of evidence of the benefits from e-learning tools, doubt about whether it is possible to gather evidence in learning, and higher demands for evidence for e-learning tools than for other educational tools. There is also a missing trust and missing communication among stakeholders. We also observed a lack of vision on how digitalization can go beyond "PDFing" a book, and, finally, we met a fear that using Big Data for personalization of the teaching/learning process will be used to stereotype education, or will only be used to save costs.

Probabilistic modelling of sea surges in coastal urban areas

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Contributors: Alstrup, S., Rootzén, H.
Publication date: 2016
Peer-reviewed: Yes
Research output: Contribution to conference › Paper – Annual report year: 2016 › Research › peer-review
Process Knowledge Discovery Using Sparse Principal Component Analysis

As the goals of ensuring process safety and energy efficiency become ever more challenging, engineers increasingly rely on data collected from such processes for informed decision making. During recent decades, extracting and interpreting valuable process information from large historical data sets have been an active area of research. Among the methods used, principal component analysis (PCA) is a well-established technique that allows for dimensionality reduction for large data sets by finding new uncorrelated variables, namely principal components (PCs). However, it is difficult to interpret the derived PCs, as each PC is a linear combination of all of the original variables and the loadings are typically nonzero. Sparse principal component analysis (SPCA) is a relatively recent technique proposed for producing PCs with sparse loadings via the variance sparsity trade-off. We propose a forward SPCA approach that helps uncover the underlying process knowledge regarding variable relations. This approach systematically determines the optimal sparse loadings for each sparse PC while improving interpretability and minimizing information loss. The salient features of the proposed approach are demonstrated through the Tennessee Eastman process simulation. The results indicate how knowledge and process insight can be discovered through a systematic analysis of sparse loadings.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Beijing University of Chemical Technology, University of California
Contributors: Gao, H., Gajjar, S., Kulahci, M., Zhu, Q., Palazoglu, A.
Pages: 12046-12059
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Industrial and Engineering Chemistry Research
Volume: 55
Issue number: 46
ISSN (Print): 0888-5885
Ratings:
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 3.1 SJR 0.95 SNIP 1.155
Web of Science (2016): Indexed yes
Original language: English
Keywords: ENGINEERING,, ROTATION, PCA, DATABASES
DOI:
10.1021/acs.iecr.6b03045
Source: FindIt
Source-ID: 2347934767
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Research › peer-review

Quantitative tumor heterogeneity assessment on a nuclear population basis

Immunohistochemistry (IHC) Ki-67 stain is widely used for visualizing cell proliferation. The common method for scoring the proliferation is to manually select and score a hot spot. This method is time-consuming and will often not give reproducible results due to subjective selection of the hotspots and subjective scoring. An automatic hotspot detection and proliferative index scoring would be time-saving, make the determination of the Ki-67 score easier and minimize the
uncertainty of the score by introducing a more objective and standardized score.

Tissue Micro Array (TMA) cores stained for Ki-67 and their neighbor slide stained for Pan Cytokeratin (PCK) were aligned and Ki-67 positive and negative nuclei were identified inside tumor regions. A heatmap was calculated based on these and illustrates the distribution of the heterogenous response of Ki-67 positive nuclei in the tumor tissue. An automatic hot spot detection was developed and the Ki-67 score was calculated. All scores were compared with scores provided by a pathologist using linear regression models.

No significant difference was found between the Ki-67 scores guided by the developed heatmap and the scores provided by a pathologist. For comparison, scores were also calculated at a random place outside the hot spot and these scores were found to be significantly different from the pathologist scores.

A heatmap visualizing the heterogeneity in tumor tissue expressed by Ki-67 was developed and used for an automatic identification of hot spots in which a Ki-67 score was calculated. The Ki-67 scores did not differ significantly from scores provided by a pathologist.
Spatial distribution and abundance of culicoides imicola and obsolesces group in Europe

General information
Publication status: Published
Organisations: National Veterinary Institute, Section for Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Roskilde University, National Veterinary Institute, Bernhard Nocht Institute for Tropical Medicine, Norwegian Veterinary Institute, Institute for Veterinary Public Health, Centre de coopération Internationale en Recherche Agronomique pour le Développement, Universite de Strasbourg, EID Méditerranée, Aarhus University, National Veterinary Research Institute
Pages: 65-66
Publication date: 2016

Host publication information
Title of host publication: 3rd Conference on Neglected Vectors and Vector-Borne Diseases (EurNegVec): with MC and WG Meeting of the COST Action TD1303: Abstract book
Place of publication: ZARAGOZA
Publisher: European Science Foundation
Electronic versions:
abstractszaragoza.pdf
Source: PublicationPreSubmission
Source-ID: 124141015
Research output: Chapter in Book/Report/Conference proceeding » Conference abstract in proceedings – Annual report year: 2016 » Research » peer-review

Split-plot designs for multistage experimentation
Most of today's complex systems and processes involve several stages through which input or the raw material has to go before the final product is obtained. Also in many cases factors at different stages interact. Therefore, a holistic approach for experimentation that considers all stages at the same time will be more efficient. However, there have been only a few attempts in the literature to provide an adequate and easy-to-use approach for this problem. In this paper, we present a novel methodology for constructing two-level split-plot and multistage experiments. The methodology is based on the Kronecker product representation of orthogonal designs and can be used for any number of stages, for various numbers of subplots and for different number of subplots for each stage. The procedure is demonstrated on both regular and nonregular designs and provides the maximum number of factors that can be accommodated in each stage. Furthermore, split-plot designs for multistage experiments with good projective properties are also provided.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Norwegian University of Science and Technology, Luleå University of Technology
Contributors: Kulahci, M., Tyssedal, J.
Pages: 493-510
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Journal of Applied Statistics
Volume: 44
Issue number: 3
ISSN (Print): 0266-4763
Ratings:
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 0.64 SJR 0.404 SNIP 0.766
Web of Science (2016): Impact factor 0.664
Web of Science (2016): Indexed yes
Original language: English
Keywords: Kronecker product, Mirror image pairs, Projectivity, Restrictions on randomization, Two-level designs
DOIs:
10.1080/02664763.2016.1177497
Surplus yeast tank failing catastrophically
GOOD REASON FOR CAUTION I A large surplus yeast tank shot into the air leaving the floor plate and the contents behind. Although not designed for overpressure, the tank was kept at “very slight overpressure” to suppress nuisance foaming. The brewery was unaware of the hazards of compressed air. The accident described in this article serves to illustrate that care should be taken if a tank originally designed for atmospheric pressure is modified to operate at slight overpressure.

Sustainable flood risk management – What is sustainable?
Sustainable flood risk management has to be achieved since flood protection is a fundamental societal service that we must deliver. Based on the discourse within the fields of risk management and sustainable urban water management, we discuss the necessity of assessing the sustainability of flood risk management, and propose an evaluation framework for doing so. We argue that it is necessary to include quantitative sustainability measures in flood risk management in order to exclude unsustainable solutions. Furthermore, we use the concept of absolute sustainability to discuss the prospects of maintaining current service levels without compromising future generation’s entitlement of services. Discussions on the sustainability of different overall flood risk schemes must take place. Fundamental changes in the approaches will require fundamental changes in the mind-sets of practitioners as well as lawmakers, politicians and the general public, which inevitably will take some time. Right now, the importance lies in setting an agenda where sustainability is important and needs to be quantified and assessed when managing flood risk.
Symptoms and quality of life in patients with chronic obstructive pulmonary disease treated with aclidinium in a real-life setting

**Introduction:** Chronic obstructive pulmonary disease (COPD) is a progressive disease with symptoms that can have a major impact on patients' physical health. The aim of this study was to evaluate quality of life (QoL), symptom severity and dyspnoea in COPD patients treated with aclidinium up to 24 weeks.

**Methods:** In this prospective non-interventional multicentre study (198 centres in Sweden, Denmark, and Norway), COPD patients (age ≥40 years) who started treatment with aclidinium (initial therapy, change of treatment, or add-on therapy) could be included. Health-related QoL was obtained by COPD assessment test (CAT). Symptoms were evaluated on a 6-point Likert scale. The modified Medical Research Council (mMRC) Dyspnoea Scale was used as a simple grading system to assess the level of dyspnoea/shortness of breath from 0 to 4. Patients on treatment with aclidinium who completed baseline and at least one follow-up visit (week 12 or 24) were included in the study population.

**Results:** Overall, 1,093 patients were enrolled (mean 69 years, 54% females), one-third had ≥1 exacerbation the year prior to baseline. At enrolment, 48% were LAMA naïve. Mean (standard deviation, SD) CAT score decreased from 16.9 (7.7) at baseline to 14.3 (7.3) at week 24 (p<.01) with a decrease in all individual CAT items (p<.05). Mean difference in morning and night-time symptoms from baseline to week 24 was -0.60 (SD 2.51) and -0.44 (SD 2.48), respectively (both p<.001). Mean (SD) mMRC Dyspnoea Scale changed from 1.6 (1.0) at baseline to 1.5 (1.0) at week 24 (p<.001).

**Conclusion:** In this observational study of a Nordic real-life COPD population, treatment with aclidinium was associated with a clinically important improvement in QoL and morning and night-time symptoms, most pronounced in the LAMA naïve group. However, there is still room for improvement in the management of symptomatic COPD patients.
The effect of skatole and androstenone on consumer response towards fresh pork from m. longissimus thoracis et lumborum and m. semimembranosus

Consumer liking was assessed for boneless chops (m. longissimus thoracis et lumborum) and schnitzels (m. semimembranosus) from castrates and entire male pigs with an androstenone content of up to 9.4 ppm and a skatole content of up to 0.92 ppm in the back fat. Skatole affected both odour and flavour as assessed by trained sensory panelists (P < 0.05–P < 0.001), while androstenone particularly affected flavour (P < 0.01–P < 0.001). Furthermore, the skatole compound seemed to be more important in explaining the presence of boar taint in the chops, while androstenone seemed to be more important for the schnitzels. For the chops, tenderness was the most important attribute for consumer liking independently of both the androstenone and skatole contents (P < 0.001). Furthermore, increasing contents of both androstenone (P = 0.05) and skatole (P = 0.04) resulted in a decreased liking of the chops. Skatole was the most important factor regarding consumer response towards the schnitzels (P = 0.03). The very low liking scores for both chops and schnitzels were mainly attributable to the androstenone content.
The fundament of food, crop protein production, is threatened by climate change

**General information**
Publication status: Published
Organisations: Department of Chemical and Biochemical Engineering, Department of Environmental Engineering, Atmospheric Environment, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University
Publication date: 2016

**Publication information**
Media of output: Atlas of Science
Year: 2016
Original language: English
Electronic versions:
Ingvordsen_2016_the_fundament_of_food_crop_protein_production_is_threatened_by_climate_change.pdf
URLs:

**Bibliographical note**
Senest ændret: 03/05/2016
Source: PublicationPreSubmission
Source-ID: 123632088
Research output: Other contribution › Net publication - Internet publication – Annual report year: 2016 › Research › peer-review

The Traveling Optical Scanner – Case Study on 3D Shape Models of Ancient Brazilian Skulls
Recovering detailed morphological information from archaeological or paleontological material requires extensive hands-on time. Creating 3D scans based on e.g. computed tomography (CT) will recover the geometry of the specimen, but can inflict bimolecular degradation. Instead, we propose a fast, inoffensive and inexpensive 3D scanning modality based on structured light, suitable for capturing the morphology and the appearance of specimens. Benefits of having 3D models are manifold. The 3D models are easy to share among researchers and can be made available to the general public. Advanced morphological modelling is possible with accurate description of the specimens provided by the models. Furthermore, performing studies on models reduces the risk of damage to the original specimen. In our work we employ a high resolution structured light scanner for digitalizing a collection of 8500 year old human skulls from Brazil. To evaluate the precision of our setup we compare the structured light scan to micro-CT and achieve submillimetre difference. We analyse morphological features of the Brazilian skulls using manual landmarks, but a research goal is to automate this, fully utilize the dense 3D scans, and apply the method to many more samples.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen
Pages: 398-405
Publication date: 2016

**Host publication information**
Title of host publication: Proceedings of the 7th International Conference on Image and Signal Processing (ICISP 2016)
Publisher: Springer
ISBN (Print): 978-3-319-33617-6
ISBN (Electronic): 978-3-319-33618-3
(Lecture Notes in Computer Science, Vol. 9680).
DOIs:
10.1007/978-3-319-33618-3_40
Source: FindIt
Source-ID: 2304853338
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2016 › Research › peer-review

Translation of risk factor estimates into on-farm interventions and their effect on Campylobacter broiler flock prevalence
Before deciding upon interventions to control Campylobacter in broiler flocks, it would be useful to estimate the potential effects of different interventions. Certain previously identified risk factors for colonization of broiler flocks with Campylobacter may seem to have large impact on the broiler flock prevalence. Nevertheless, interventions related to
these risk factors may have only limited effect on the overall prevalence estimate, since in practice only a relatively small fraction of farms are actually amenable for an intervention related to a given risk factor. We present a novel method for the risk assessor that predicts effects of interventions at the farm, based on results from a risk factor study that included data from six European countries (Denmark, Netherlands, Norway, Poland, Spain and United Kingdom). In the present study, five previously identified risk factors, which had shown to have significant impact on Campylobacter flock prevalence, were translated into practical on-farm interventions. Given the implementation of these interventions the population prevalence was predicted by developing and using a statistical method anchored in the ideas behind standardized population estimations using logistic regression. To obtain population estimates per country, the predicted prevalence values were multiplied by the frequencies (no. of farms) in a reference population based on data from the risk factor study and a large questionnaire. The latter was included to improve the representativeness of the reference population. Population prevalence estimates were calculated before and after implementation of a given intervention in the six countries. Results showed that if biosecurity was not accounted for, some individual interventions resulted in a limited reduction of the population prevalence. The reduction differed between countries depending on the current farm management practices and the actual flock prevalence level. In general, the most effective interventions were "building new houses with strict biosecurity for all houses older than 15 years" and "apply drinkers with nipples without cups". In conclusion, the novel method translates results from risk factor studies into effects of on-farm interventions for the reduction of the prevalence of Campylobacter in broiler flocks. The method is very useful for providing the basis for risk management decisions. The usefulness would improve further when the results are integrated with costs of interventions in a cost effectiveness study. The approach was developed for Campylobacter in broiler flocks, but it can also be applied to other pathogens and other farm animals, given that the required data are available.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Research group for Risk Benefit
Contributors: Sommer, H. M., Nauta, M., Rosenquist, H.
Number of pages: 11
Pages: 27-37
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Microbial Risk Analysis
Volume: 2-3
ISSN (Print): 2352-3522
Original language: English
Keywords: European study, Intervention, Logistic regression, On-farm Campylobacter control, Standardized population estimate, Translation model
DOIs: 10.1016/j.mran.2016.06.001
Source: FindIt
Source-ID: 2305408551
Research output: Contribution to journal › Journal article – Annual report year: 2016 › Research › peer-review

Understanding Liking in Relation to Sensory Characteristics, Consumer Concept Associations, Arousal Potential and "Appropriateness for Use" Using Apple Juice as an Application: Understanding Liking
It is crucial to understand influential parameters for acquisition of consumer liking to ensure successful product introduction and competitiveness in the marketplace. This article aims to study and understand liking in relation to sensory characteristics, consumer concept associations, arousal potential and appropriateness for use using apple juices as an application. First, a laboratory panel (n=15; F=10, M=5) determined the sensory profile of the apple juices using the methods Partial Napping and Ultra Flash Profiling based on taste and flavor. Next, consumers (n=196; F=136, M=60) evaluated key apple juice parameters. The basic tastes sweet and sour were key properties and played a central role in liking acquisition. Apple juices having a sweet/sour balance were most liked. The importance of balance in sensory properties was underlined by the fact that consumer liking was related to the concept balanced. Additionally, the consumers liked the apple juices when associated with exclusive, National identity, interesting, complex and unique. High levels of perceived complexity, surprising and novelty increased the consumers' liking. The most liked apple juices received the highest ratings in appropriateness for use regardless of use situation. Practical Applications: Food producers can use the information to understand how a food product's sensory characteristics, consumer concept associations, arousal potential and appropriateness for use influence liking. This information is useful in new food product development and marketing of the products.

General information
Publication status: Published
Organisations: Statistics and Data Analysis, Department of Applied Mathematics and Computer Science, University of Copenhagen, Aarhus University
Contributors: Stolzenbach, S., Bredie, W. L. P., Christensen, R. H. B., Byrne, D. V.
Use of Sparse Principal Component Analysis (SPCA) for Fault Detection
Principal component analysis (PCA) has been widely used for data dimension reduction and process fault detection. However, interpreting the principal components and the outcomes of PCA-based monitoring techniques is a challenging task since each principal component is a linear combination of the original variables which can be numerous in most modern applications. To address this challenge, we first propose the use of sparse principal component analysis (SPCA) where the loadings of some variables in principal components are restricted to zero. This paper then describes a technique to determine the number of non-zero loadings in each principal component. Furthermore, we compare the performance of PCA and SPCA in fault detection. The validity and potential of SPCA are demonstrated through simulated data and a comparative study with the benchmark Tennessee Eastman process.

Validation of Reported Whole-Grain Intake from a Web-Based Dietary Record against Plasma Alkylresorcinol Concentrations in 8- to 11-Year-Olds Participating in a Randomized Controlled Trial
BACKGROUND: Whole-grain (WG) intake is important for human health, but accurate intake estimation is challenging. Use of a biomarker for WG intake provides a possible way to validate dietary assessment methods. OBJECTIVE: Our aim was to validate WG intake from 2 diets reported by children, using plasma alkylresorcinol (AR) concentrations, and to investigate the 3-mo reproducibility of AR concentrations and reported WG intake. METHODS: AR concentrations were analyzed in fasting blood plasma samples, and WG intake was estimated in a 7-d web-based diary by 750 participants aged 8-11 y in a 2 school meal × 3 mo crossover trial. Reported WG intake and plasma AR concentrations were compared when children ate their usual bread-based lunch (UBL) and when served a hot lunch meal (HLM). Correlations and cross-classification were used to rank subjects according to intake. The intraclass correlation coefficients (ICCs) between subjects’ measurements at baseline and after the UBL were used to assess reproducibility. RESULTS:
Correlations between reported WG wheat + rye intake and plasma AR were 0.40 and 0.37 (P <0.001) for the UBL and the HLM diets, and 78% and 77% were classified in the same or adjacent quartiles for the UBL and HLM diets, respectively. The ICC over 3 mo was 0.47 (95% CI: 0.38, 0.55) for plasma total ARs and 0.64 (95% CI: 0.58, 0.70) for reported WG intake. Correlations were higher when using the AR C17:0 homolog as a biomarker, reflecting rye intake instead of plasma total ARs [UBL: r = 0.47; HLM: r = 0.43, P <0.001; ICC = 0.51 (95% CI: 0.43, 0.59)]. CONCLUSIONS: Self-reported WG wheat + rye intake among children showed moderate correlations with plasma AR concentrations. Substantial intraindividual variation was found in WG intake and plasma AR concentrations. The AR homolog C17:0 may be used as a biomarker for WG intake when the WG intake primarily comes from rye as in the present study. This trial was registered at clinicaltrials.gov as NCT01457794.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Research group for Risk Benefit, University of Copenhagen, Karolinska Institutet, Swedish University of Agricultural Sciences
Pages: 377-383
Publication date: 2016
Peer-reviewed: Yes

Publication information
Journal: Journal of Nutrition
Volume: 146
Issue number: 2
ISSN (Print): 0022-3166
Ratings:
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 3.93 SJR 2.025 SNIP 1.336
Web of Science (2016): Impact factor 4.145
Web of Science (2016): Indexed yes
Original language: English
Electronic versions:
ansa_J._Nutr._2016_Biltoft_Jensen_377_83.pdf
jn.115.222620.full.pdf
DOIs:
10.3945/jn.115.222620
Source: FindIt
Source-ID: 2290431689
Research output: Contribution to journal > Journal article – Annual report year: 2016 > Research > peer-review

Visualizing multifactorial and multi-attribute effect sizes in linear mixed models with a view towards sensometrics
In Brockhoff et al (2016), the close link between Cohen's d, the effect size in an ANOVA framework, and the so-called Thurstonian (Signal detection) d-prime was used to suggest better visualizations and interpretations of standard sensory and consumer data mixed model ANOVA results. The basic and straightforward idea is to interpret effects relative to the residual error and to choose the proper effect size measure. For multi-attribute bar plots of F-statistics this amounts, in balanced settings, to a simple transformation of the bar heights to get them transformed into depicting what can be seen as approximately the average pairwise d-primes between products. For extensions of such multi-attribute bar plots into more complex models, similar transformations are suggested and become more important as the transformation depends on the number of observations within factor levels, and hence makes bar heights better comparable for factors with differences in number of levels. For mixed models, where in general the relevant error terms for the fixed effects are not the pure residual error, it is suggested to base the d-prime-like interpretation on the residual error. The methods are illustrated on a multifactorial sensory profile data set and compared to actual d-prime calculations based on ordinal regression modelling through the ordinal package. A generic “plug-in” implementation of the method is given in the SensMixed package, which again depends on the lmerTest package. We discuss and clarify the bias mechanisms inherently challenging effect size measure estimates in ANOVA settings.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aalborg University
Publication date: 2016
Media of output: Video
**Vitamin D3 and 25-hydroxyvitamin D3 in pork and their relationship to vitamin D status in pigs**

The content of vitamin D in pork produced in conventional systems depends on the vitamin D concentration in the pig feed. Both vitamin D3 and 25-hydroxyvitamin D3 (25(OH)D3) are essential sources of dietary vitamin D; however, bioavailability assessed by serum 25(OH)D3 concentration is reported to be different between the two sources. Furthermore, the relationship between serum 25(OH)D3 level and the tissue content of vitamin D3 and 25(OH)D3 is unknown. The objective of this study was to investigate the potential of increasing the content of vitamin D in different pig tissues by increasing the levels of vitamin D3 and 25(OH)D3 in the pig feed for 49 d before slaughter. Concurrently, the 25(OH)D3 level in serum was investigated as a biomarker to assess the content of vitamin D3 and 25(OH)D3 in pig tissues. Adipose tissue, white and red muscle, the liver and serum were sampled from pigs fed feed containing either vitamin D3 or 25(OH)D3 at 5, 20, 35 or 50 µg/kg feed for 7 weeks before slaughter. The tissue 25(OH)D3 level was significantly higher in the pigs fed 25(OH)D3 compared with those fed vitamin D3, while the tissue vitamin D3 level was higher in the pigs fed vitamin D3 compared with those fed 25(OH)D3. The content of 25(OH)D3 in the different tissues fully correlated with the serum 25(OH)D3 level, whereas the correlation between the tissue content of vitamin D3 and serum 25(OH)D3 was dependent on the source of the ingested vitamin D3.
X-ray based micromechanical finite element modeling of composite materials

This is a study of a uni-directional non-crimp fabric reinforced epoxy composite material typically used as the load carrying laminate in wind turbine blades. Based on a 3D x-ray tomography scan, the bundle and fibre/matrix structure of the composite is segmented. This segmentation is used in a multi-scale finite element model bridging the gap from the individual fibers organized in bundles to the stitched non-crimp fabric used for building up the load carrying laminates.

Fremtidens samfund mangler kvindelige professorer

Uden specielle tiltag for kvinder får mændene de største fordele.
A 10-day heatwave at flowering superimposed on climate change conditions strongly affects production of 22 barley accessions

Extreme climate events are projected to be among the future most challenging constraints to plant development. Heatwaves as well as floods and droughts cause acute changes in the growth environment determining our primary production (Collins et al., 2013). Europe experienced extreme heatwaves in 2003 and 2006. In 2003, a 21 % decrease in the French wheat production was found from temperatures up to 6 °C above long-term means and precipitation being less than 50 % of the average (Cliais et al., 2005). One strategy to mitigate the this decrease from heatwaves is to identify resilient cultivars and incorporate them in breeding programs.

AB0346 Correlation of Disease Duration, Disease Activity and Pain to Fatigue – a Systematic Review

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Copenhagen University Hospital Frederiksberg and Bispebjerg
Contributors: Groth Madsen, S., Danneskiold-Samsøe, B., Stockmarr, A., Bartels, E.
Pages: 1009
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Annals of the Rheumatic Diseases
Volume: 74
Issue number: Supplement 2
ISSN (Print): 0003-4967
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 7.4 SJR 4.537 SNIP 2.998
Accuracy of self-reported intake of signature foods in a school meal intervention study: comparison between control and intervention period

Bias in self-reported dietary intake is important when evaluating the effect of dietary interventions, particularly for intervention foods. However, few have investigated this in children, and none have investigated the reporting accuracy of fish intake in children using biomarkers. In a Danish school meal study, 8- to 11-year-old children (n 834) were served the New Nordic Diet (NND) for lunch. The present study examined the accuracy of self-reported intake of signature foods (berries, cabbage, root vegetables, legumes, herbs, potatoes, wild plants, mushrooms, nuts and fish) characterising the NND. Children, assisted by parents, self-reported their diet in a Web-based Dietary Assessment Software for Children during the intervention and control (packed lunch) periods. The reported fish intake by children was compared with their ranking according to fasting whole-blood EPA and DHA concentration and weight percentage using the Spearman correlations and cross-classification. Direct observation of school lunch intake (n 193) was used to score the accuracy of food-reporting as matches, intrusions, omissions and faults. The reporting of all lunch foods had higher percentage of matches compared with the reporting of signature foods in both periods, and the accuracy was higher during the control period compared with the intervention period. Both Spearman's rank correlations and linear mixed models demonstrated positive associations between EPA+DHA and reported fish intake. The direct observations showed that both reported and real intake of signature foods did increase during the intervention period. In conclusion, the self-reported data represented a true increase in the intake of signature foods and can be used to examine dietary intervention effects.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Research group for Risk Benefit, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen, University of Waterloo
Number of pages: 10
Pages: 635-644
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: British Journal of Nutrition
Volume: 114
Issue number: 4
ISSN (Print): 0007-1145
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 3.52 SJR 1.587 SNIP 1.18
Web of Science (2015): Impact factor 4.051
Web of Science (2015): Indexed yes
Original language: English
Keywords: Web-based food diaries, n-3 fatty acids, Direct observation, Matches, Intrusions, Omissions, Faults
Electronic versions:
S0007114515002020a.pdf
DOIs:
10.1017/s0007114515002020
Source: Findit
Source-ID: 2279829473
Research output: Contribution to journal → Journal article – Annual report year: 2015 → Research → peer-review

Analysis of data from the MariCare Smartfloor at Skovhuset Care Home
In this project data is analysed from a smartfloor which is installed in an elderly care home. Two lines of investigation are carried out. The first uses “event data” from the smartfloor. This data contains every event (bathroom visits, people entering/exiting the room etc.) registered by the floor over a one year period. Control charts are used to investigate a link between the event data and UTI incidence. A clear predictor for UTI is not found, but the value of control charts in this setting is demonstrated. In the second line of investigation “position data” is analysed. The position data is more limited
than the event data in that it is extracted manually using image analysis on the smartfloor user interface program. Using the position data, the trajectories traced by a resident moving about their room are visualised and properties such as direction and speed are investigated. A method is found for comparison of trajectories to determine their degree of similarity and this method can identify unusual trajectories in the dataset.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Spooner, M. P., Jørgensen, T. M., Thyregod, C., Ershub, B. K.
Number of pages: 48
Publication date: 2015

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Electronic versions:
2016_04_20_Skovhuset_Report.pdf

**Bibliographical note**
In cooperation with Charlotte Kock Petersen of Skovhuset Care Home and Palle Stevn of MariCare

**Antibiotika til mink fra 2007-2012**

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Division of Epidemiology and Microbial Genomics, National Veterinary Institute, Copenhagen Fur
Contributors: Jensen, V. F., Sommer, H. M., Stuve, T., Clausen, J., Chréil, M.
Pages: 34-37
Publication date: 2015
Peer-reviewed: Yes

**Publication information**
Journal: Dansk Pelsdyravl
Issue number: 2
ISSN (Print): 0011-6424
Original language: Danish
Electronic versions:
Medicin_til_mink_Dansk_pelsdyravl_2015.pdf
URLs:
http://ipaper.ipaperoms.dk/KopenhagenFur/DanskPelsdyravlFebruar2015/?Page=34
Source: PublicationPreSubmission
Source-ID: 106132030
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

**Artifact removal in Differential Phase Contrast X-ray Computed Tomography**

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Algarve
Contributors: Jørgensen, T. M., AL-Tam, F., Dahl, V. A.
Pages: 105-108
Publication date: 2015

**Host publication information**
Title of host publication: Farm animal imaging
Publisher: European Cooperation in Science and Technology
Editors: Maltin, C., Craigie, C., Bünger, L.
ISBN (Print): 978-0-9931063-1-6
Electronic versions:
FA2744_FAIM_REPORT_IV_AW_09_15.pdf
Automated mixed ANOVA modeling of sensory and consumer data

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Groupe ESA
Contributors: Kuznetsova, A., Christensen, R. H. B., Bavay, C., Brockhoff, P. B.
Number of pages: 8
Pages: 31-38
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Food Quality and Preference
Volume: 40
ISSN (Print): 0950-3293
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 3.92 SJR 1.042 SNIP 1.781
Web of Science (2015): Indexed yes
Original language: English
DOIs:
10.1016/j.foodqual.2014.08.004
Source: FindIt
Source-ID: 270405270
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

Back scattering interferometry revisited – A theoretical and experimental investigation
A refractive index based detector based on so called back scattering interferometry (BSI) has been described in the literature as a unique optical method for measuring biomolecular binding interactions in solution. In this paper, we take a detailed look at the optical principle underlying this technique to understand fully the constituents and behaviour of the fringe patterns generated. The simulated results are compared and validated with experimental measurements. Hereby, we show that BSI does not operate as a resonant cavity as often stated in the literature. Recently, we have questioned the claims made that BSI in general can be used to measure molecular bindings. Here we explore this topic further in three cases using fluorescence spectroscopy as a reference method. Finally, we explore whether refractive index sensing can be used to measure the enzymatic phosphorylation of glucose to glucose-6-phosphate.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Photonics Engineering, Aalborg University, Aalborg University Hospital
Contributors: Jørgensen, T. M., Jepsen, S. T., Sørensen, H. S., di Gennaro, A., Kristensen, S. R.
Pages: 1328-1337
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Sensors and Actuators B: Chemical
Volume: 220
ISSN (Print): 0925-4005
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 4.84 SJR 1.305 SNIP 1.472
Web of Science (2015): Impact factor 4.758
Web of Science (2015): Indexed yes
Original language: English
DOIs:
Battling Bluetongue and Schmallenberg virus: Local scale behavior of transmitting vectors

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Veterinary Institute, Section for Epidemiology
Contributors: Stockmarr, A., Kirkeby, C., Bødker, R.
Publication date: 2015
Media of output: PowerPoint

Event information
Event: World Congress on Infectious Diseases 2015
Location: London, United Kingdom
Electronic versions:
Infectious_Diseases_1008_2015_Anders_Stockmarr.pdf
Research output: Non-textual form » Sound/Visual production (digital) – Annual report year: 2015 » Research

Battling Bluetongue and Schmallenberg virus: Local scale behavior of transmitting vectors

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Stockmarr, A.
Pages: 43
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Journal of infectious disease and therapy
Volume: 3
Issue number: 4
Original language: English
Electronic versions:
Abstract_Anders_Stockmarr_et_al_Infectious_Diseases_2015.pdf
2332_0877.S1.002_004_1.pdf
DOIs: 10.4172/2332-0877.S1.002
Research output: Contribution to journal » Conference abstract in journal – Annual report year: 2015 » Research » peer-review

Big Data fra jord til bord
Danske landmænd og virksomhederne i fødevaresektoren har gode forudsætninger for at drage nytte af den rivende udvikling inden for indsamling og bearbejdning af data:
• Danmark har en stærk fødevaresектор. Det skyldes bl.a., at alle dele af værdikæden arbejder tæt sammen. Fra primærproducererne, over forarbejdning, agroindustrien til videns- og forskningsmiljøerne. Effektiv ressourceudnyttelse og fokus på optimering i hele værdikæden gør sektoren i stand til at konkurrere på verdensmarkedet.
• Danske fødevarevirksomheder har altid været gode til at opdyrke nye forretningsmodeller og finde nye innovative veje til øget værdiskabelse. For eksempel gennem smarere måder at producere på, levere produkterne på eller at indarbejde større værdi i produkterne, så de kan sælges med større forfremme.
• Dansk landbrug og hele værdikæden i fødevaresektoren producerer store mængder af data. Det skyldes bl.a. et højt automationsniveau og myndighedernes krav til dokumentation af fødevarekvaliteten, når de danske producenter leverer fødevarer til forbrugerne verden over. Der er imidlertid et stort spring fra at råde over store mængder af data til at bruge dem aktivt i forretningsudviklingen. Denne rapport viser, hvordan Big Data kan være ét af omdrejningspunkter

General information
Publication status: Published
Organisations: Office for Innovation & Sector Services, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Division of Risk Assessment and Nutrition, Research group for Analytical Food Chemistry, National Veterinary Institute, Epidemiology, Department of Management Engineering,
Calculation of ruin probabilities for a dense class of heavy tailed distributions

In this paper, we propose a class of infinite-dimensional phase-type distributions with finitely many parameters as models for heavy tailed distributions. The class of finite-dimensional phase-type distributions is dense in the class of distributions on the positive reals and may hence approximate any such distribution. We prove that formulas from renewal theory, and with a particular attention to ruin probabilities, which are true for common phase-type distributions also hold true for the infinite-dimensional case. We provide algorithms for calculating functionals of interest such as the renewal density and the ruin probability. It might be of interest to approximate a given heavy tailed distribution of some other type by a distribution from the class of infinite-dimensional phase-type distributions and to this end we provide a calibration procedure which works for the approximation of distributions with a slowly varying tail. An example from risk theory, comparing ruin probabilities for a classical risk process with Pareto distributed claim sizes, is presented and exact known ruin probabilities for the Pareto case are compared to the ones obtained by approximating by an infinite-dimensional hyper-exponential distribution.

Challenges for bovine viral diarrhoea virus antibody detection in bulk milk by antibody enzyme-linked immunosorbent assays due to changes in milk production levels

Background: Bovine viral diarrhoea (BVD) is considered eradicated from Denmark. Currently, very few (if any) Danish cattle herds could be infected with BVD virus (BVDV). The Danish antibody blocking enzyme-linked immunosorbent assay (ELISA) has been successfully used during the Danish BVD eradication program, initiated in 1994. During the last decade, the cattle herd size has increased while the prevalence of BVDV has decreased. In this study, we investigated how these changes could affect the performance of the Danish blocking ELISA and of the SVANOVIR® BVDV-Ab indirect ELISA. The latter has successfully been used to eradicate BVD in Sweden. Data (2003–2010) on changes in median herd size and milk production levels, occurrence of viremic animals and bulk milk surveillance were analysed. Additionally, the Danish blocking ELISA and the SVANOVIR ELISA were compared analyzing milk and serum samples. The prevalence of...
antibody positive milking cows that could be detected by each test was estimated, by diluting positive individual milk samples and making artificial milk pools. Results: During the study period, the median herd size increased from 74 (2003) to 127 cows (2010), while the prevalence of BVDV infected herds decreased from 0.51 to 0.02 %. The daily milk yield contribution of a single seropositive cow to the entire daily bulk milk was reduced from 1.61 % in 2003 to 0.95 % in 2010 due to the increased herd size. It was observed that antibody levels in bulk milk decreased at national level. Moreover, we found that when testing bulk milk, the SVANOVIR® BVDV-Ab can detect a lower prevalence of seropositive lactating cows, compared to the Danish blocking ELISA (0.78 % vs. 50 %). Values in the SVANOVIR® BVDV-Ab better relate to low concentrations of antibody positive milk (R2 = 94-98 %), than values in the blocking ELISA (R2 = 23-75 %). For sera, the two ELISAs performed equally well. Conclusions: The SVANOVIR ELISA is recommended for analysis of bulk milk samples in the current Danish situation, since infected dairy herds e.g. due to import of infected cattle can be detected shortly after BVDV introduction, when only few lactating cows have seroconverted. In sera, the two ELISAs can be used interchangeably.

General information
Publication status: Published
Organisations: National Veterinary Institute, Department of Applied Mathematics and Computer Science , Statistics and Data Analysis, Danish Cattle Federation
Contributors: Foddai, A., Enøe, C., Stockmarr, A., Krogh, K., Uttenthal, Å.
Number of pages: 8
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Acta Veterinaria Scandinavica (Online)
Volume: 57
Issue number: 32
ISSN (Print): 0044-605X
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 0.98 SJR 0.425 SNIP 0.814
Web of Science (2015): Impact factor 1.23
Original language: English
Keywords: Bovine viral diarrhoea, Bulk milk, Antibody ELISA, Surveillance
Electronic versions:
art_3A10.1186_2Fs13028_015_0125_z.pdf
DOIs:
10.1186/s13028-015-0125-z

Bibliographical note
© 2015 Foddai et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Source: PublicationPreSubmission
Source-ID: 111793405
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

Characterization of the bacterial gut microbiota of piglets suffering from new neonatal porcine diarrhoea
Background: In recent years, new neonatal porcine diarrhoea (NNPD) of unknown aetiology has emerged in Denmark. NNPD affects piglets during the first week of life and results in impaired welfare, decreased weight gain, and in the worst-case scenario death. Commonly used preventative interventions such as vaccination or treatment with antibiotics, have a limited effect on NNPD. Previous studies have investigated the clinical manifestations, histopathology, and to some extent, microbiological findings; however, these studies were either inconclusive or suggested that Enterococci, possibly in interaction with Escherichia coli, contribute to the aetiology of NNPD. This study examined ileal and colonic luminal contents of 50 control piglets and 52 NNPD piglets by means of the qPCR-based Gut Microbiotassay and 16 samples by 454 sequencing to study the composition of the bacterial gut microbiota in relation to NNPD. Results: NNPD was associated with a diminished quantity of bacteria from the phyla Actinobacteria and Firmicutes while genus Enterococcus was more than 24 times more abundant in diarrhoeic piglets. The number of bacteria from the phylum Fusobacteria was also doubled in piglets suffering from diarrhoea. With increasing age, the gut microbiota of NNPD affected piglet and control piglets became more diverse. Independent of diarrhoeic status, piglets from first parity sows (gilts) possessed significantly more bacteria from family Enterobacteriaceae and species E. coli, and fewer bacteria from phylum Firmicutes. Piglets born to gilts had 25 times higher odds of having NNPD compared with piglets born to multiparous sows. Finally, the co-occurrence of genus Enterococcus and species E. coli contributed to the risk of having NNPD. Conclusion: The results of this study support previous findings that points towards genus Enterococcus and species E. coli to be involved in the pathogenesis of NNPD. Moreover, the results indicate that NNPD is associated with a disturbed bacterial composition and
larger variation between the diarrhoeic piglets.

**General information**
Publication status: Published
Organisations: National Veterinary Institute, Section for Bacteriology, Pathology and Parasitology, Section for Immunology and Vaccinology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Chemical and Biochemical Engineering, Center for BioProcess Engineering, Danish Genome Institute, Danish Agriculture and Food Council
Number of pages: 19
Publication date: 2015
Peer-reviewed: Yes

**Publication information**
Journal: B M C Veterinary Research
Volume: 11
Issue number: 139
ISSN (Print): 1746-6148
Ratings:
- BFI (2015): BFI-level 1
- Scopus rating (2015): CiteScore 1.86 SJR 0.952 SNIP 0.943
- Web of Science (2015): Impact factor 1.643
- Web of Science (2015): Indexed yes
Original language: English
Keywords: NNPD, Neonatal, Piglet, Diarrhoea, qPCR, Microbiota, Gut Microbiota assay, 454 sequencing
Electronic versions:
DOIs:
10.1186/s12917-015-0419-4

**Bibliographical note**
© 2015 Hermann-Bank et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

**Characterizing Digital Light Processing (DLP) 3D Printed Primitives**
The resolution and repeatability of 3D printing processes depends on a number of factors including the software, hardware, and material used. When printing parts with features that are near or below the nominal printing resolution, it is important to understand how the printer works. For example, what is the smallest unit shape that can be produced? And what is the reproducibility of that process? This paper presents a method for automatically detecting and characterizing the height, width, and length of micro scale geometric primitives produced via a digital light processing (DLP) 3D printing process. An upper limit, lower limit, and best estimate for each dimension is reported for each primitive. Additionally, the roughness, rectangularity, and tilt of the top of each primitive is estimated. The uncertainty of the best estimate is indicated using standard deviations for a series of primitives. The method generalizes to unseen primitives, and the results illustrate that the dimension estimates converge as the size of the primitives increases. The primitives’ rectangularity also increases as the size increases. Finally, the primitives specified with 5 to 68μm varying heights have been estimated to group into five different heights with fairly low variance of the best estimates of the heights. This reflects how the requested geometry is parsed and produced by the printer.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Mechanical Engineering, Manufacturing Engineering, Technical University of Denmark
Pages: 302-313
Publication date: 2015

**Host publication information**
Title of host publication: Proceedings of the 19th Scandinavian Conference on Image Analysis, SCIA 2015
Combining analysis of variance and three-way factor analysis methods for studying additive and multiplicative effects in sensory panel data

Data from descriptive sensory analysis are essentially three-way data with assessors, samples and attributes as the three ways in the data set. Because of this, there are several ways that the data can be analysed. The paper focuses on the analysis of sensory characteristics of products while taking into account the individual differences among assessors. In particular, we will be interested in considering the multiplicative assessor model, which explicitly models the different usage of scale. A multivariate generalization of the model will be proposed, which allows to analyse the differences in the use of the scale with reference to the existing structure of relationships between sensory descriptors. The multivariate assessor model will be tested on a data set from milk. Relations between the proposed model and other multiplicative models like parallel factor analysis and analysis of variance will be clarified. Copyright © 2014 John Wiley & Sons, Ltd.

Computer-aided diagnosis of pulmonary diseases using x-ray darkfield radiography

In this work we develop a computer-aided diagnosis (CAD) scheme for classification of pulmonary disease for grating-based x-ray radiography. In addition to conventional transmission radiography, the grating-based technique provides a dark-field imaging modality, which utilizes the scattering properties of the x-rays. This modality has shown great potential for diagnosing early stage emphysema and fibrosis in mouse lungs in vivo. The CAD scheme is developed to assist radiologists and other medical experts to develop new diagnostic methods when evaluating grating-based images. The scheme consists of three stages: (i) automatic lung segmentation; (ii) feature extraction from lung shape and dark-field image intensities; (iii) classification between healthy, emphysema and fibrosis lungs. A study of 102 mice was conducted with 34 healthy, 52 emphysema and 16 fibrosis subjects. Each image was manually annotated to build an experimental dataset. System performance was assessed by: (i) determining the quality of the segmentations; (ii) validating emphysema and fibrosis recognition by a linear support vector machine using leave-one-out cross-validation. In terms of segmentation quality, we obtained an overlap percentage (Ω) 92.63 ± 3.65%, Dice Similarity Coefficient (DSC) 89.74 ± 8.84% and Jaccard Similarity Coefficient 82.39 ± 12.62%. For classification, the accuracy, sensitivity and specificity of diseased lung recognition was 100%. Classification between emphysema and fibrosis resulted in an accuracy of 93%, whilst the sensitivity was 94% and specificity 88%. In addition to the automatic classification of lungs, deviation maps
created by the CAD scheme provide a visual aid for medical experts to further assess the severity of pulmonary disease in the lung, and highlights regions affected.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Technische Universität München, Helmholtz Zentrum München, Ludwig-Maximilians-University Hospital Munich
Pages: 9253-9268
Publication date: 2015
Peer-reviewed: Yes

**Publication information**
Journal: Physics in Medicine and Biology
Volume: 60
ISSN (Print): 0031-9155
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 3.31 SJR 1.577 SNIP 1.648
Web of Science (2015): Impact factor 2.811
Web of Science (2015): Indexed yes
Original language: English
Keywords: X-ray radiography, Dark-field imaging, Lung segmentation, Active appearance model, Pulmonary disease, Grating based interferometry
DOIs:
10.1088/0031-9155/60/24/9253
Source: PublicationPreSubmission
Source-ID: 118220160
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

**Design and analysis of biomedical studies**
Biomedicine is a field that has great influence on the majority of mankind. The constant development has considerably changed our way of life during the last centuries. This has been achieved through the dedication of biomedical researchers along with the tremendous resources that over time have been allocated this field. It is utterly important to utilize these resources responsibly and efficiently by constantly striving to ensure high-quality biomedical studies. This involves the use of a sound statistical methodology regarding both the design and analysis of biomedical studies. The focus of this project is on statistical aspects that arise within the field of biomedicine.

Two types of errors are frequently accentuated within the framework of statistics, namely type I and type II errors. Type I errors occur when a null hypothesis erroneously is rejected. An acceptable type I error rate is specified prior to conducting the statistical analysis. However, all statistical models make assumptions and if violated the actual type I error rate may deviate from the pre-specified type I error rate. Type II errors occur when we fail to reject a false null hypothesis. On contrary to the type I error rate, the type II error rate is not explicitly specified during the statistical analysis and this entails that assessment of the type II error rate in practice is at risk of being neglected altogether. Concerns regarding type I errors, type II errors and adherence (or lack thereof) to model assumptions for biomedical studies are a recurring theme in this thesis.

Data collected in some biomedical studies are positively skewed; hence methods relying on the normal distribution are not directly applicable. We investigated how data from one of these studies are suitably analyzed. We extracted 23 different summary statistics from data gathered from eleven studies. The degree of adherence to the model assumptions evaluated for each of these summary statistics form basis for our conclusions.

Hierarchically structured data are frequently encountered in biomedical studies. For one type of studies entailing such data we have conducted a literature study strongly indicating that this structure commonly is neglected in the statistical analysis. Based on this closed-form expressions for the approximate type I error rate are formulated. The type I error rates are assessed for a number of factor combinations as they appear in practice and in all cases the type I error rates are demonstrated to be severely inflated.

Prior to conducting a study it is important to perform power and sample size determinations to ensure that reliable conclusions can be drawn from the statistical analysis. We have formulated closed-form expressions for the statistical power of studies with a hierarchical structure to guide biomedical researchers designing future studies of this type. Upon model fitting it is important to examine if the model assumptions are met to avoid that spurious conclusions are drawn. While the range of diagnostic methods is extensive for models assuming a normal response it is generally more limited for non-normal models. An R package providing diagnostic tools suitable for examining the validity of binomial regression models have been developed. The binomTools package is publicly available at the CRAN repository.
Development of LED Light Sources for Improved Visualization of Veins: a statistical approach
The present statistical study investigates the difference of diffuse reflectances between skin and vein (defined as contrast indicator) under different visible wavelengths of a population of 39 adult participants. The purpose of the study is to examine if there is a group of wavelengths-color combinations that could explain most of the variance (of the contrast indicator) in the data set. Moreover the effect of gender and age on the contrast indicator is explored.

Do different spectral domain OCT hardwares measure the same? Comparison of retinal thickness using third-party software
Purpose
Spectral-domain optical coherence tomographies (OCTs) from different companies do not give identical retinal thicknesses. The purpose of this study was to evaluate if differences in thickness when using a spectral domain Cirrus OCT or a Heidelberg Spectralis are due to hardware differences, or if they are caused by the segmentation algorithms.

Methods
Thirty-seven healthy eyes were examined within the same session with a Cirrus OCT and a Spectralis OCT, the latter using averaged B-scans. Scans from similar positions and passing the fovea were analyzed by custom-made software. Thickness was analyzed at the fovea, the central 1-mm line and the 6-mm line.

Results
When Cirrus and Spectralis scans were analyzed with the same software, the retinal thickness at the foveal center was 225.92 μm (SD 17.0) using the Cirrus and 228.70 μm (SD 18.4) using the Spectralis; the difference of 2.78 μm was not significant (p=0.055). For the central 1 mm, the difference was 1.78 μm (p=0.0414), and for all points out to 6 mm, the Spectralis retinal thickness was also significantly larger than the Cirrus thickness (p=0.0052), though the mean difference was only 1.85 μm. Also for the RPE_OS complex, Spectralis measured a greater thickness than did Cirrus, with a mean of 3.32 μm (p<0.0001) for all points.

Conclusion
The retinal thicknesses from the Cirrus and from the Spectralis differed by 14 μm with the standard software of the instruments, and by less than 3 μm when analyzed with the same custom-made software, indicating that the major differences between the two SD-OCT systems are due to differences in their built-in software algorithms.
Effects of school meals based on the New Nordic Diet on intake of signature foods: a randomised controlled trial. The OPUS School Meal Study

A New Nordic Diet (NND) was developed in the context of the Danish OPUS Study (Optimal well-being, development and health for Danish children through a healthy New Nordic Diet). Health, gastronomic potential, sustainability and Nordic identity were crucial principles of the NND. The aim of the present study was to investigate the effects of serving NND school meals compared with the usual packed lunches on the dietary intake of NND signature foods. For two 3-month periods, 834 Danish children aged 8-11 years received NND school meals or their usual packed lunches brought from home (control) in random order. The entire diet was recorded over 7 consecutive days using a validated Web-based Dietary Assessment Software for Children. The NND resulted in higher intakes during the entire week (% increase) of root vegetables (116 (95 % CI 1.93, 2.42)), cabbage (26 (95 % CI 1.08, 1.47)), legumes (22 (95 % CI 1.06, 1.40)), herbs (175 (95 % CI 2.36, 3.20)), fresh berries (48 (95 % CI 1.13, 1.94)), nuts and seeds (18 (95 % CI 1.02, 1.38)), lean fish and fish products (47 (95 % CI 1.31, 1.66)), fat fish and fish products (18 (95 % CI 1.02, 1.37)) and potatoes (129 (95 % CI 2.05, 2.56)). Furthermore, there was a decrease in the number of children with zero intakes when their habitual packed lunches were replaced by NND school meals. In conclusion, this study showed that the children increased their intake of NND signature foods, and, furthermore, there was a decrease in the number of children with zero intakes of NND signature foods when their habitual packed lunches were replaced by school meals following the NND principles.

General information
Publication status: Published
Organisations: Division of Nutrition, National Food Institute, Department of Informatics and Mathematical Modeling, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Management Engineering, University of Copenhagen
Number of pages: 8
Pages: 772-779
Publication date: 2015
Peer-reviewed: Yes
Efficient Estimating Functions for Stochastic Differential Equations

The overall topic of this thesis is approximate martingale estimating function-based estimation for solutions of stochastic differential equations, sampled at high frequency. Focus lies on the asymptotic properties of the estimators. The first part of the thesis deals with diffusions observed over a fixed time interval. Rate optimal and efficient estimators are obtained for a one-dimensional diffusion parameter. Stable convergence in distribution is used to achieve a practically applicable Gaussian limit distribution for suitably normalised estimators. In a simulation example, the limit distributions of an efficient and an inefficient estimator are compared graphically. The second part of the thesis concerns diffusions with finite-activity jumps, observed over an increasing interval with terminal sampling time going to infinity. Asymptotic distribution results are derived for consistent estimators of an ageneral multidimensional parameter. Conditions for rate optimality and efficiency of estimators of drift-jump and diffusion parameters are given in some special cases. These conditions are found to extend the pre-existing conditions applicable to continuous diffusions, and impose much stronger requirements on the estimating functions in the presence of jumps. Certain implications of these conditions are discussed, as is a heuristic notion of how efficient estimating functions might be constructed, thus setting the stage for further research.

Estimation of on-farm interventions to control Campylobacter

Before making risk management decisions to control Campylobacter prevalence in broiler flocks, it is useful to identify effective interventions. Given risk factors may seem to have a large effect, but in practice interventions related to this risk factor may have only limited effect due to a relative small proportion of the farms that can actually be intervened for the given risk factors. We present a novel tool for risk assessors to obtain such estimates of the effect of interventions before it is implemented at the farms. A statistical method was developed in order to estimate the flock prevalence if an intervention was to be implemented in a given population of broiler farms. The method is anchored in the ideas behind standardized population estimations. In order to obtain a country wise population estimate the predicted prevalence values are multiplied with elements from a reference population. In the present study risk factor estimates from a European study was used and the reference population consisted of data from the risk factor study plus extra data from a large questionnaire survey to improve the representativeness of the reference population. The results showed that some individual interventions gave only a limited reduction in prevalence if the biosecurity was not accounted for. Furthermore, the effect of the interventions differed between countries, depending on current farm management practices and Campylobacter prevalence. The most effective interventions were “building new houses with strict biosecurity for all houses older than 15 years” and “apply drinkers with nipples without cups”.

Evaluation of acceptance sampling related to sealing plug and pipe stub.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Thyregod, C., Ersbøll, B. K.
Number of pages: 20
Publication date: 2015

Evaluation of back scatter interferometry, a method for detecting protein binding in solution
Back Scatter Interferometry (BSI) has been proposed to be a highly sensitive and versatile refractive index sensor usable for analytical detection of biomarker and protein interactions in solution. However the existing literature on BSI lacks a physical explanation of why protein interactions in general should contribute to the BSI signal. We have established a BSI system to investigate this subject in further detail. We contribute with a thorough analysis of the robustness of the sensor including unwanted contributions to the interferometric signal caused by temperature variation and dissolved gases. We report a limit of the effective minimum detectability of refractive index at the $10^{-7}$ level. Long term stability was examined by simultaneously monitoring the temperature inside the capillary revealing an average drift of $2.0 \times 10^{-7}$ per hour. Finally we show that measurements on protein A incubated with immunoglobulin G do not result in a signal that can be attributed to binding affinities as otherwise claimed in literature.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Photonics Engineering, Aalborg University Hospital
Contributors: Jepsen, S. T., Jørgensen, T. M., Zong, W., Trydal, T., Kristensen, S. R., Sørensen, H. S.
Number of pages: 7
Pages: 895-901
Publication date: 2015
Peer-reviewed: Yes
Experiments for Multi-Stage Processes

Multi-stage processes are very common in both process and manufacturing industries. In this article we present a methodology for designing experiments for multi-stage processes. Typically in these situations the design is expected to involve many factors from different stages. To minimize the required number of experimental runs, we suggest using mirror image pairs of experiments at each stage following the first. As the design criterion, we consider their projectivity and mainly focus on projectivity 3 designs. We provide the methodology for generating these designs for processes with any number of stages and also show how to identify and estimate the effects. Both regular and non-regular designs are considered as base designs in generating the overall design.

Farm specific risk factors for Campylobacter colonization of broilers in six European countries

What:
This study, part of the EU financed CamCon project, presents the results from a multi-national risk factor study, where farm data were collected through a standardised questionnaire survey carried out in six EU countries: Denmark, Norway, the Netherlands, Poland, Spain and the UK.

Why:
To identify common and country-specific, on-farm risk factors that can be used to identify the most efficient on-farm measures for preventing broiler flocks from becoming colonized with Campylobacter (See presentation by H.M. Sommer, Wednesday morning).

How:
By applying a variance model that handles the explanatory variables in the generalized linear model using backward elimination and forward selection.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Norwegian University of Science and Technology, Wageningen University & Research, CReSA - Centre for Animal Health Research, National Veterinary Research Institute
Farm specific risk factors for Campylobacter colonization of broilers in six European countries

This study was part of the EU financed project CamCon. The objective of the study was to identify on-farm risk factors for Campylobacter colonization of broiler flocks based on comparable data from six European countries. The data included explanatory variables from a large questionnaire concerning production, farm management procedures and conditions, climate data on mean temperature, sunshine hours, precipitation, as well as data on Campylobacter status of broiler flocks. Overall, the study comprised data from more than 6000 flocks. The data were analyzed using a generalized linear model using backwards elimination and forward selection. Due to the structure of the data, several models were explored, by applying different strategies for categorizing explanatory variables and for selection and elimination of variables in the model. The risk of broiler flocks becoming colonized with Campylobacter was clearly affected by country. In descending order; broiler flocks were more likely to be colonized in Poland, the UK, Spain, the Netherlands, Denmark and Norway due to country specific factors that could not be explained by the management and climate variables in the explored models. The seasonality in the prevalence of Campylobacter was described nicely by temperature, i.e. the number of positive flock increased with increasing temperatures. The age of broiler houses, presence of anterooms and barriers in all houses, designated tools for each house as well as length of downtime and the type of drinker systems were found to affect the risk of the broiler flocks becoming colonized by Campylobacter.

Fitting a distribution to microbial counts: Making sense of zeroes

The accurate estimation of true prevalence and concentration of microorganisms in foods is an important element of quantitative microbiological risk assessment (QMRA). This estimation is often based on microbial detection and enumeration data. Among such data are artificial zero counts, that originated by chance from contaminated food products. When these products are not differentiated from uncontaminated products that originate true zero counts, the estimates of true prevalence and concentration may be inaccurate. This inaccuracy is especially relevant in situations where highly pathogenic bacteria are involved and where growth can occur along the food pathway. Our aim was to develop a method that provides accurate estimates of concentration parameters and differentiates between artificial and true zeroes, thus also accurately estimating true prevalence.

We first show the disadvantages of using a limit of quantification (LOQ) threshold for the analysis of microbial enumeration data. We show that, depending on the original distribution of concentrations and the LOQ value, it may be incorrect to treat artificial zeroes as censored below a quantification threshold.

Next, a method is developed that estimates the true prevalence of contamination within a food lot and the parameters
characterizing the within-lot distribution of concentrations, without assuming a LOQ, and using raw plate count data as an input. Counts resulting both from contaminated and uncontaminated sample units are analysed together. This procedure allows the estimation of the proportion of artificial zeroes among the total of zero counts, and therefore the estimation of true prevalence from enumeration results.

We observe that this method yields best estimates of mean, standard deviation and prevalence at low true prevalence levels and low expected standard deviation. Furthermore, we conclude that the estimation of prevalence and the estimation of the distribution of concentrations are interrelated and therefore should be estimated simultaneously. We also conclude that one of the keys to an accurate characterization of the overall microbial contamination is the correct identification and separation of true and artificial zeroes.

Our method for the analysis of quantitative microbial data shows a good performance in the estimation of true prevalence and the parameters of the distribution of concentrations, which indicates that it is a useful data analysis tool in the field of QMRA.

**General information**
Publication status: Published
Organisations: National Food Institute, Division of Epidemiology and Microbial Genomics, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Ribeiro Duarte, A. S., Stockmarr, A., Nauta, M.
Pages: 40-50
Publication date: 2015
Peer-reviewed: Yes

**Publication information**
Journal: International Journal of Food Microbiology
Volume: 196
ISSN (Print): 0168-1605
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 4.02 SJR 1.64 SNIP 1.626
Web of Science (2015): Indexed yes
Original language: English
DOIs: 10.1016/j.ijfoodmicro.2014.11.023
Source: PublicationPreSubmission
Source-ID: 103606159
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

**Fiere overvægtige danske kvinder**

**General information**
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Matthiessen, J., Stockmarr, A.
Number of pages: 10
Publication date: 2015
Peer-reviewed: No

**Publication information**
Journal: E-artikel fra DTU Fødevareinstitutet
Volume: 2015
Issue number: 2
ISSN (Print): 1904-5581
Original language: Danish
Electronic versions: Flere_overvgtige_danske_kvinder_011015.pdf
Source: PublicationPreSubmission
Source-ID: 116759850
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research
Generalized requirements and decompositions for the design of test parts for micro additive manufacturing research

The design of experimental test parts to characterize micro additive manufacturing (AM) processes is challenging due to the influence of the manufacturing and metrology processes. This work builds on the lessons learned from a case study in the literature to derive generalized requirements and high level decompositions for the design of test parts and the design of experiments to characterize micro additive manufacturing processes. While the test parts and the experiments described are still work in progress, the generic requirements derived from them can serve as a starting point for the design of other micro additive manufacturing related studies and their decompositions can help structure future work.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Department of Mechanical Engineering, Manufacturing Engineering, Statistics and Data Analysis
Contributors: Thompson, M. K., Clemmensen, L. K. H.
Pages: 229-235
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Procedia CIRP
Volume: 34
ISSN (Print): 2212-8271
Ratings:
Scopus rating (2015): SJR 0.572 SNIP 0.878
Original language: English
Keywords: Additive manufacturing, micro manufacturing, metrology, design of experiments, requirements, decomposition
Electronic versions:
Thompson_and_Clemmensen_2015.pdf
DOIs:
10.1016/j.procir.2015.07.075

Bibliographical note
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Source: PublicationPreSubmission
Source-ID: 115098587
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

IC3 and IC4 Trains Under Risk of Blocking their Wheels

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Stockmarr, A.
Publication date: 2015
Media of output: PowerPoint

Event information
Event: Big Data, Data Warehousing, and Data Analytics
Location: Technical University of Denmark, Lyngby, Denmark
Electronic versions:
Big_Data_Data_Warehousing_and_Data_Analytics_ANST_2608_2015.pdf
Research output: Non-textual form › Sound/Visual production (digital) – Annual report year: 2015 › Research

iCFD: Interpreted Computational Fluid Dynamics – Degeneration of CFD to one-dimensional advection-dispersion models using statistical experimental design – The secondary clarifier

The present study aims at using statistically designed computational fluid dynamics (CFD) simulations as numerical experiments for the identification of one-dimensional (1-D) advection-dispersion models – computationally light tools, used e.g., as sub-models in systems analysis. The objective is to develop a new 1-D framework, referred to as interpreted CFD (iCFD) models, in which statistical meta-models are used to calculate the pseudo-dispersion coefficient (D) as a function of design and flow boundary conditions. The method – presented in a straightforward and transparent way – is illustrated using the example of a circular secondary settling tank (SST). First, the significant design and flow factors are screened out by applying the statistical method of two-level fractional factorial design of experiments. Second, based on the number of significant factors identified through the factor screening study and system understanding, 50 different sets of design and flow conditions are selected using Latin Hypercube Sampling (LHS). The boundary condition sets are imposed on a 2-D axi-symmetrical CFD simulation model of the SST. In the framework, to degenerate the 2-D model structure, CFD model outputs are approximated by the 1-D model through the calibration of three different model structures for D. Correlation
ICFD modeling of final settlers - developing consistent and effective simulation model structures

Summary of key findings The concept of interpreted computational fluid dynamic (ICFD) modelling and the development methodology are presented (Fig. 1). The 1-D advection-dispersion model along with the statistically generated, meta-model for pseudo-dispersion constitutes the newly developed iCFD concept. The case of secondary settling tanks (SSTs) is used to demonstrate the methodological steps using the validated CFD model with the hindered-transient compression settling velocity model by (10). Factor screening and latin hypercube sampling (LSH) are used to degenerate a 2-D axi-symmetrical CFD model structure (10) into a one-dimensional (1-D) advection-dispersion model structure. The boundary condition sets, obtained in the LHS, are imposed on the 2-D CFD simulation model of the SST. In the framework, to degenerate the 2-D model structure, CFD model outputs are approximated by the 1-D model through the calibration of three different model structures for D, the pseudo-dispersion coefficient. Correlation equations for the D parameter (meta-models) are then identified as a function of the selected design and flow boundary conditions, and their accuracy is evaluated against the D values estimated in each numerical experiment. The evaluation and validation of the iCFD model structure is carried out using scenario simulation results obtained with parameters sampled from the corners of the LHS experimental region. For the studied SST, additional iCFD model development was carried out in terms of (i) assessing different density current sub-models; (ii) implementation of a combined flocculation, hindered, transient and compression settling velocity function; and (iii) assessment of modelling the onset of transient and compression settling. Furthermore, the optimal level of model discretization both in 2-D and 1-D was undertaken. Results suggest that the iCFD model developed for the SST through the proposed methodology is able to predict solid distribution with high accuracy -- taking a reasonable computational effort -- when compared to multi-dimensional numerical experiments, under a wide range of flow and design conditions. iCFD tools could play a crucial role in reliably predicting systems’ performance under normal and shock events.

General information
Publication status: Published
Organisations: Department of Environmental Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Urban Water Engineering, Technical University of Denmark
Contributors: Guyonvarch, E., Ramin, E., Kulahci, M., Plosz, B. G.
Pages: 396-411
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Water Research
Volume: 83
ISSN (Print): 0043-1354
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 6.63 SJR 2.772 SNIP 2.397
Web of Science (2015): Impact factor 5.991
Web of Science (2015): Indexed yes
Original language: English
Keywords: Computational Fluid Dynamics (CFD), One-dimensional advection dispersion model, Interpreted CFD model (iCFD), Statistical factor screening, Degeneration of model structural complexity, Secondary settling tank
DOIs: 10.1016/j.watres.2015.06.012
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review
designed moderate (under-loaded) and extreme (under-, critical and overloaded) operational boundary conditions on the 2-D CFD SST model (8). Results obtained in the statistical analysis of the CFD outputs in the extreme scenario suggest that the loading conditions characterised with feed solid concentration (Xin), SST overflow rate (Qov), recycle ratio (R), and, to a minor extent, the inlet height (Hin), are the four significant factors, impacting the SST performance, in terms of sludge blanket height (SBH), solids concentration in recycle of activated sludge (XRAS) and solids concentration in SST effluent (Xeff). Statistical results obtained in the moderate scenario indicate five factors significantly influencing the SST performance, i.e. Xin, Qov, R, sidewater depth (SWD) and Hin. It should be noted that the baffles installed in the SST inlet and outlet (characterised with 4 parameters), are found to have negligible effect compared to the aforementioned five factors. Internal baffling probably has a positive effect, especially on effluent quality albeit at a lower order of magnitude not visible in our results. LHS. Based on the screening study, the five significant factors used in the subsequent LHS are Xin, Qov, R, SWD and Hin. The LHS yielded 50 CFD simulations, and the outputs, in terms of SBH, XRAS, Xeff, total amount of solids in SST (Mtot) values, are extracted from the steady-state solution, and then used to perform the iCFD model identification, iCFD model identification. We assess the 1-D model structure in terms of setting the feed location (layer) and transient/compression threshold concentration, XTC. Using the CFD outputs from LHS, we assessed nine different model structures based on literature (1; 3; 2; 10; 9) and on more recent considerations (Fig. 2a). Validation tests were done using the CFD outputs from extreme scenarios. The most effective model structure (relatively low the sum of square of relative errors, SSRE, and computational time) obtained is that in which the XTC is set at the concentration of the layer just below the feed-layer. The feed-layer location is set to the highest location where X>Xin (solids concentration in SST influent). An effective discretization level (computational time/numerical error) is assessed by approximating the LHS outputs with an iCFD model discretised using 10-200 layers. Solutions obtained show convergence, and the discretisation scheme with 60 layers seems an effective trade-off. Identification and validation of the D-model. To identify a formulation for the pseudo-dispersion coefficient in the iCFD model, we tested three structural scenarios by defining (i) one single pseudo-dispersion coefficient (D0) for all the layers; (ii) one pseudo-dispersion coefficient (D1) above and another pseudodispersion coefficient (D2) below the feed-layer; (iii) one pseudo-dispersion coefficient (Df) just around the feed-layer. These scenarios were inspired by literature (1; 2; 9). As for the D0--iCFD model, values of SSRE obtained are below 1 with an average SSRE=0.206. The simulation model thus can predict the solids distribution inside the tank with a satisfactory accuracy. Averaged relative errors of 8.1 %, 3.1 %, 18.1 % and 11.8 % are observed for SBH, XRAS, Xeff and Mtot, respectively. A statistical analysis of the calibrated D0 compared to the five input factors is performed. In addition to the elementary factors, four interactions are found significant: Xin2, Xin*Hin, R*Hin and Xin*Qov. A correlation is obtained between the loading and design factors and D0 with an R² of 0.853 (adjusted R²=0.819), i.e. Eq. 1. D0 values predicted with Eq. 1 are compared to the values estimated using 1-D model approximation of the 50 CFD outputs, and results suggest a relatively effective correlation as well (Fig. 2b). The three iCFD models, employing different formulations for D, calibrated using the meta-models (Eq. 1 for D0), are used to approximate the CFD outputs obtained in the extreme scenarios. Using the D0--iCFD model, the solid distribution at the corners of the LHS experimental region can be predicted with the highest accuracy (average SSRE=0.71), thereby validating the simulation model. For D1.2--iCFD, the meta-model is limited in calculating D, and the model fails the validation test using the extreme scenario (SSRE=386). Using the Df--iCFD model, the predictive accuracy obtained is comparable with that obtained with the D0--iCFD (compare 1.73 to 0.71). The computational time required through simulation with Df--iCFD, however, is significantly higher than that obtained with the D0--iCFD (on average 3.5 times longer). Therefore, this study concludes that, considering the capability and usability - in terms of complexity and computational time - the D0--iCFD model is preferred.

General information
Publication status: Published
Organisations: Department of Environmental Engineering, Urban Water Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Technical University of Denmark
Contributors: Plósz, B. G., Guyonvarch, E., Ramin, E., Kulahci, M.
Number of pages: 2
Publication date: 2015
Peer-reviewed: Yes
Event: Abstract from 9th IWA Symposium on Systems Analysis and Integrated Assessment (Watermatex 2015), Gold Coast, Queensland, Australia.
Source: PublicationPreSubmission
Source-ID: 118851291
Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2015 › Research › peer-review

Impact of Autocorrelation on Principal Components and Their Use in Statistical Process Control
A basic assumption when using principal component analysis (PCA) for inferential purposes, such as in statistical process control (SPC), is that the data are independent in time. In many industrial processes, frequent sampling and process dynamics make this assumption unrealistic rendering sampled data autocorrelated (serially dependent). PCA can be used to reduce data dimensionality and to simplify multivariate SPC. Although there have been some attempts in the literature to deal with autocorrelated data in PCA, we argue that the impact of autocorrelation on PCA and PCA-based SPC is neither well understood nor properly documented. This article illustrates through simulations the impact of autocorrelation on the descriptive ability of PCA and on the monitoring performance using PCA-based SPC when autocorrelation is ignored. In the simulations, cross-correlated and autocorrelated data are generated using a stationary first-order vector autoregressive model. The results show that the descriptive ability of PCA may be seriously affected by autocorrelation causing a need to incorporate additional principal components to maintain the model's explanatory ability. When all
variables have equal coefficients in a diagonal autoregressive coefficient matrix, the descriptive ability is intact, while a significant impact occurs when the variables have different degrees of autocorrelation. We also illustrate that autocorrelation may impact PCA-based SPC and cause lower false alarm rates and delayed shift detection, especially for negative autocorrelation. However, for larger shifts, the impact of autocorrelation seems rather small.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Luleå University of Technology
Contributors: Vanhatalo, E., Kulahci, M.
Publication date: 2015
Peer-reviewed: Yes

**Publication information**
Journal: Quality and Reliability Engineering International
ISSN (Print): 0748-8017
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 1.71 SJR 1.3 SNIP 1.524
Web of Science (2015): Impact factor 1.457
Web of Science (2015): Indexed yes
Original language: English
Keywords: Safety, Risk, Reliability and Quality, Management Science and Operations Research, Multivariate data, Principal component analysis, Serial dependence, Vector autoregressive model, Autocorrelation, Process control, Statistical process control, Auto-correlated data, Autoregressive coefficient, Monitoring performance, Negative autocorrelation, Statistical process controls (SPC)
Electronic versions:
qre1858.pdf
DOIs:
10.1002/qre.1858
Source: FindIt
Source-ID: 2280951912
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

Individualized Learning Through Non-Linear use of Learning Objects: With Examples From Math and Stat
Our aim is to ensure individualized learning that is fun, inspiring and innovative. We believe that when you enjoy, your brain will open up and learning will be easier and more effective. The methods use a non-linear learning environment based on self-contained learning objects which are pieced together by a Hyperbolic Graph, or by the students themselves. This learning system makes it easy for students to find a path through the course material which suits his/her personal learning style and which makes learning more motivating, and efficient, and which leads to better learning. The methods have been tested in two case studies. One was a continuing education course in statistics for a global medical company, and the other was a “big” - both in terms of the number of students and in the number of ECTS points - introductory course in mathematics at a major technical university. The continuing education course made it possible for the company’s employees, from many different parts of the company and from all around the world, to learn from the same course. For the university course we started with a pilot project where our methods were used during only one course week. The pilot was a success and we then used the experiences from it to reshape the entire course. This course has now been running for 5 years and consistently receives very good evaluations, both from students and teachers. A clear finding from the test cases is that our learning method creates more motivation and makes the students use more time on the course and prepare better for the lectures. An important discussing point is how much “free choice” is best for the learners. We believe that the possibility to follow your own learning style by choosing between different types of material is important and ensures better learning and more motivation for all students, and that for the best students it also gives courage to go beyond the curriculum.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Rootzén, H.
Pages: 500-506
Publication date: 2015

**Host publication information**
Title of host publication: Proceedings of the 14th European Conference on e-Learning (ECEL 2015)
Keywords: Individual learning, Learning objects, Non-linear learning, Hyperbolic graph, Continuing education course, University mathematics course
In silico modelling of permeation enhancement potency in Caco-2 monolayers based on molecular descriptors and random forest

Structural traits of permeation enhancers are important determinants of their capacity to promote enhanced drug absorption. Therefore, in order to obtain a better understanding of structure–activity relationships for permeation enhancers, a Quantitative Structural Activity Relationship (QSAR) model has been developed. The random forest-QSAR model was based upon Caco-2 data for 41 surfactant-like permeation enhancers from Whitehead et al. (2008) and molecular descriptors calculated from their structure. The QSAR model was validated by two test-sets: (i) an eleven compound experimental set with Caco-2 data and (ii) nine compounds with Caco-2 data from literature. Feature contributions, a recent developed diagnostic tool, was applied to elucidate the contribution of individual molecular descriptors to the predicted potency. Feature contributions provided easy interpretable suggestions of important structural properties for potent permeation enhancers such as segregation of hydrophilic and lipophilic domains. Focusing on surfactant-like properties, it is possible to model the potency of the complex pharmaceutical excipients, permeation enhancers. For the first time, a QSAR model has been developed for permeation enhancement. The model is a valuable in silico approach for both screening of new permeation enhancers and physicochemical optimisation of surfactant enhancer systems.
Kage årsag til iltmangel og indlæggelse på hospital
Ambulance bragte en 58-årig forretningsrejsende til akutmodtagelsen med åndenød, svimmelhed og kvalme. Denne og andre cases viser, at man skal have respekt for kuldioxid.

Kraftig eksplosion efter sammenblanding af salpetersyre og 2-propanol

Layered Surface Detection in Micro-CT Tetra Pak Data
Learning and Motivational Processes When Students Design Curriculum-Based Digital Learning Games

This design-based research (DBR) project has developed an overall gamified learning design (big Game) to facilitate the learning process for adult students by inviting them to be their own learning designers through designing digital learning games (small games) in cross-disciplinary subject matters. The DBR project has investigated and experimented with which elements, methods, and processes are important when aiming at creating a cognitive complex (Anderson and Krathwohl, 2001) and motivating learning process within a reusable game-based learning design. This project took place in a co-design process with teachers and students. The learning approach was founded in problem-based learning (PBL) and constructionist pedagogical methodology, building on the thesis that there is a strong connection between designing and learning. The belief is that activities that involve making, building, or programming provide a rich context for learning, since the construction of artefacts, in this case learning games, enables reflection and new ways of thinking. The students learned from reflection and interaction with the tools alone as well as in collaboration with peers. After analysing the students' learning trajectories within this method of learning, this study describes seven areas of the iterative learning and game design process. The analysis also shows that the current learning design is constructed as a hierarchy supported through different roles as learning designers contained within one another. The study found that the students benefitted from this way of learning as a valid variation to more conventional teaching approaches, and teachers found that the students learned at least the same amount or more compared to traditional teaching processes. The students were able to think outside the box and experienced hard fun (Papert, 2002) - the phenomena that everyone likes challenging things to do, as long as they are the right things matched to the individual. They were motivated by hands-on work and succeeded in developing four very different and meaningful learning games and game concepts, which contributed to achieving their learning goals.

Linear mixed models in sensometrics

Today's companies and researchers gather large amounts of data of different kind. In consumer studies the objective is the collection of the data to better understand consumer acceptance of products. In such studies a number of persons (generally not trained) are selected in order to score products in terms of preferences. In sensory studies the aim is the collection of the data to better describe products and differences of the products according to a number of sensory attributes. Here trained persons, so-called assessors, score the products in terms of different characteristics such as smell, taste, texture, sound - depending on the aim of a study. It is a common approach in both studies to consider persons coming from a larger population, which, from the statistical perspective, leads to the use of mixed effects models.
Mixed effects models have been used extensively in analysis of both consumer and sensory studies. However, too simplistic models are considered, important effects are not accounted for and as a consequence important information is not gained or analysis leads to improper conclusions. The focus of this project is to propose a methodology for analyzing more complex models together with tools facilitating the methodology. This was accomplished by contributing to the mixed effects ANOVA modelling in general and specifically applied to sensory and consumer studies through a series of papers and software tools facilitating the developed methodologies. The primary advantage of the ANOVA approach is that it gives confidence intervals and significance tests for the various effects including the background variables used in the model and consequently a fast and reliable assessment and ranking of the importance of different factors.

There exists today very little easily available methodology and software which supports consumer studies with both sensory properties and background information related to health benefits, environment and user-friendliness. In close collaboration with the industrial partners an open-source software tool ConsumerCheck was developed in this project and now is available for everyone. will represent a major step forward when concerns this important problem in modern consumer driven product development. Standard statistical software packages can be used for some of the purposes, but for the specific problems considered here and for the typical users in industry, these programs are far from satisfactory. Therefore, the ConsumerCheck software represents a novel source of information for all quality-oriented industries. The effect is improved procedures for product development and hence improved quality of decision making in Danish as well as international food companies and other companies using the same methods.

The two open-source R packages lmerTest and SensMixed implement and support the methodological developments in the research papers as well as the ANOVA modelling part of the ConsumerCheck software. The SensMixed package is a package for semi-automated analysis of sensory and consumer studies within linear mixed effects framework. The lmerTest package supports tests for linear mixed effects models fitted with the lmer function of the lme4 package (Bates et al., 2013). While SensMixed is closely connected with sensometrics field, the lmerTest package has developed into a generic statistical package.

Reference manuals accompany these R packages.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Kuznetsova, A.
Number of pages: 76
Publication date: 2015

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
(DTU Compute PHD-2015; No. 374).
Electronic versions:
phd374_Kuznetsova_A.pdf

**Location and social context does matter when conducting consumer studies!**

**General information**
Publication status: Published
Organisations: National Food Institute, Division of Industrial Food Research, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Research group for Bioactives – Analysis and Application, TINE R&D
Contributors: Andersen, B. V., Kraggerud, H., Brockhoff, P. B., Hyldig, G.
Number of pages: 1
Publication date: 2015
Peer-reviewed: Yes
Electronic versions:
PANG2015_0270_paper.pdf
Research output: Contribution to conference › Poster – Annual report year: 2016 › Research › peer-review

**Model for TampImp-DeltaSigmaH Interaction.**

**General information**
Publication status: Published
Multivariate Analysis Techniques for Optimal Vision System Design

The present thesis considers optimization of the spectral vision systems used for quality inspection of food items. The relationship between food quality, vision based techniques and spectral signature are described. The vision instruments for food analysis as well as datasets of the food items used in this thesis are described. The methodological strategies are outlined including sparse regression and pre-processing based on feature selection and extraction methods, supervised versus unsupervised analysis and linear versus non-linear approaches.

One supervised feature selection algorithm based on the existing sparse regression methods (EN and lasso) and one unsupervised feature selection strategy based on the local maxima of the spectral 1D/2D signals of food items are proposed. In addition, two novel feature extraction and selection strategies are introduced; sparse supervised PCA (SSPCA) and DCT based characterization of the spectral diffused reflectance images for wavelength selection and discrimination.

These methods together with some other state-of-the-art statistical and mathematical analysis techniques are applied on datasets of different food items; meat, diaries, fruits and vegetables. These datasets are acquired using three different vision systems; a spectral imaging device called VideometerLab, spectroscopy, and diffused reflectance imaging systems called Static Light Scattering (SLS).

These analyses result in significant reduction in the number of required wavelengths and simplification of the design of practical vision systems.

Non-negative Matrix Factorization for Binary Data

We propose the Logistic Non-negative Matrix Factorization for decomposition of binary data. Binary data are frequently generated in e.g. text analysis, sensory data, market basket data etc. A common method for analysing non-negative data is the Non-negative Matrix Factorization, though this is in theory not appropriate for binary data, and thus we propose a novel Non-negative Matrix Factorization based on the logistic link function. Furthermore we generalize the method to handle missing data. The formulation of the method is compared to a previously proposed method (Tome et al., 2015). We compare the performance of the Logistic Non-negative Matrix Factorization to Least Squares Non-negative Matrix Factorization and Kullback-Leibler (KL) Non-negative Matrix Factorization on sets of binary data: a synthetic dataset, a set of student comments on their professors collected in a binary term-document matrix and a sensory dataset. We find that choosing the number of components is an essential part in the modelling and interpretation, that is still unresolved.
Optimization of Overflow Policies in Call Centers

We examine how overflow policies in a multi-skill call center should be designed to accommodate performance measures that depend on waiting time percentiles such as service level. This is done using a discrete Markovian approximation of the waiting time of the first customers waiting in line. A Markov decision chain is used to determine the optimal policy. This policy outperforms considerably the ones used most often in practice, which use a fixed threshold. The present method can be used also for other call-center models and other situations where performance is based on actual waiting times and customers are treated in a FCFS order.

General information

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, VU University Medical Centre
Contributors: Koole, G., Nielsen, B., Nielsen, T.
Pages: 461-471
Publication date: 2015
Peer-reviewed: Yes

Publication information

Journal: Probability in the Engineering and Informational Sciences
Volume: 29
Issue number: 3
ISSN (Print): 0269-9648
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 0.48 SJR 0.611 SNIP 0.622
Web of Science (2015): Impact factor 0.39
Web of Science (2015): Indexed yes
Original language: English
DOIs: 10.1017/S0269964815000091
Source: FindIt
Source-ID: 275155137
Research output: Contribution to journal → Journal article – Annual report year: 2015 → Research → peer-review

Pig herd monitoring and undesirable tripping and stepping prevention

Humane handling and slaughter of livestock are of major concern in modern societies. Monitoring animal wellbeing in slaughterhouses is critical in preventing unnecessary stress and physical damage to livestock, which can also affect the meat quality. The goal of this study is to monitor pig herds at the slaughterhouse and identify undesirable events such as pigs tripping or stepping on each other. In this paper, we monitor pig behavior in color videos recorded during unloading from transportation trucks. We monitor the movement of a pig herd where the pigs enter and leave a surveyed area. The method is based on optical flow, which is not well explored for monitoring all types of animals, but is the method of choice for human crowd monitoring. We recommend using modified angular histograms to summarize the optical flow vectors. We show that the classification rate based on support vector machines is 93% of all frames. The sensitivity of the model is 93.5% with 90% specificity and 6.5% false alarm rate. The radial lens distortion and camera position required for convenient surveillance make the recordings highly distorted. Therefore, we also propose a new approach to correct lens and foreshortening distortions by using moving reference points. The method can be applied real-time during the actual unloading operations of pigs. In addition, we present a method for identification of the causes leading to undesirable events, which currently only runs off-line. The comparative analysis of three drivers, which performed the unloading of the pigs from the trucks in the available datasets, indicates that the drivers perform significantly differently. Driver 1 has 2.95 times higher odds to have pigs tripping and stepping on each other than the two others, and Driver 2 has 1.11 times higher odds than Driver 3. (C) 2015 Elsevier B.V. All rights reserved.

General information
Plate waste and intake of school lunch based on the New Nordic Diet and on packed lunches: A randomised controlled trial in 8- to 11-year-old Danish children

The aim of the present study was to compare total food intake, total and relative edible plate waste and self-reported food likenings between school lunch based on the new Nordic diet (NND) and packed lunch from home. In two 3-month periods in a cluster-randomised controlled unblinded cross-over study 3rd- and 4th-grade children (n 187) from two municipal schools received lunch meals based on NND principles and their usual packed lunch (control). Food intake and plate waste (n 1558) were calculated after weighing lunch plates before and after the meal for five consecutive days and self-reported likenings (n 905) assessed by a web-based questionnaire. Average food intake was 6 % higher for the NND period compared with the packed lunch period. The quantity of NND intake varied with the menu (P < 0·0001) and was positively associated with self-reported likenings. The edible plate waste was 88 (sd 80) g for the NND period and 43 (sd 60) g for the packed lunch period whereas the relative edible plate waste was no different between periods for meals having waste (n 1050). Edible plate waste differed between menus (P < 0·0001), with more waste on soup days (36 %) and vegetarian days (23 %) compared with the packed lunch period. Self-reported likenings were negatively associated with percentage plate waste (P < 0·0001). The study suggests that portion sizes need to be considered in new school meal programmes. New strategies with focus on reduction of plate waste, children's likenings and nutritious school meals are crucial from both a nutritional, economic and environmental point of view.

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Research group for Risk Benefit, University of Copenhagen
Number of pages: 9
Publication date: 2015
Peer-reviewed: Yes
Recent Advances and Future Directions for Quality Engineering
The origins of quality engineering are in manufacturing, where quality engineers apply basic statistical methodologies to improve the quality and productivity of products and processes. In the past decade, people have discovered that these methodologies are effective for improving almost any type of system or process, such as financial, health care, and supply chains. This paper begins with a review of key advances and trends within quality engineering over the past decade. The second part uses the first part as a foundation to outline new application areas for the field. It also discusses how quality engineering needs to evolve in order to make significant contributions to these new areas. © 2015 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Virginia Polytechnic Institute and State University, Technical University of Denmark
Contributors: Vining, G., Kulahci, M., Pedersen, S.
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Quality and Reliability Engineering International
ISSN (Print): 0748-8017
Ratings:
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 1.71 SJR 1.3 SNIP 1.524
Web of Science (2015): Impact factor 1.457
Web of Science (2015): Indexed yes
Original language: English
Electronic versions:
Vining_et_al_2016_Quality_and_Reliability_Engineering_International.pdf
DOIs:
10.1002/qre.1797
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

Reduction in pedometer-determined physical activity in the adult Danish population from 2007 to 2012
Aims: To examine the development in pedometer-determined physical activity from 2007–2008 to 2011–2012 in the adult Danish population. Methods: The study population comprised two random samples of 18–75-year-old individuals who took part in cross-sectional studies in 2007–2008 (n=224) and 2011–2012 (n=1515). Pedometer data (sealed Yamax SW 200) were obtained for seven consecutive days. Data for 1624 participants (48.2% men) were included in the analysis. An overall step-defined activity level was examined based on a graduated step index (sedentary, low active, somewhat active, active, highly active). The pedometer-determined outcomes were analysed using regression models. Results: A borderline significant decline (p=0.077) from 8788 to 8341 steps/day (−446 (95% confidence intervals −50, 943)) was found between 2007–2008 and 2011–2012. Furthermore, a 23.7% (95% confidence intervals −41.7%, −0.1%) lower overall step-defined activity level was observed in 2011–2012 compared to 2007–2008. These changes were primarily due to a reduced level of activity among women. The proportion of individuals taking 10,000 steps/day decreased non-significantly from 34.8% to 29.3%, whereas the proportion taking

General information
Publication status: Published
Organisations: National Food Institute, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Research group for Risk Benefit, Linnaeus University
Contributors: Matthiessen, J., Andersen, E. W., Raustorp, A., Knudsen, V. K., Serensen, M. R.
Pages: 525-533
Publication date: 2015
Peer-reviewed: Yes

Publication information
Regularized generalized eigen-decomposition with applications to sparse supervised feature extraction and sparse discriminant analysis

We propose a general technique for obtaining sparse solutions to generalized eigenvalue problems, and call it Regularized Generalized Eigen-Decomposition (RGED). For decades, Fisher's discriminant criterion has been applied in supervised feature extraction and discriminant analysis, and it is formulated as a generalized eigenvalue problem. Thus RGED can be applied to effectively extract sparse features and calculate sparse discriminant directions for all variants of Fisher discriminant criterion based models. Particularly, RGED can be applied to matrix-based and even tensor-based discriminant techniques, for instance, 2D-Linear Discriminant Analysis (2D-LDA). Furthermore, an iterative algorithm based on the alternating direction method of multipliers is developed. The algorithm approximately solves RGED with monotonically decreasing convergence and at an acceptable speed for results of modest accuracy. Numerical experiments based on four data sets of different types of images show that RGED has competitive classification performance with existing multidimensional and sparse techniques of discriminant analysis.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Hong Kong
Contributors: Han, X., Clemmensen, L. K. H.
Pages: 43-54
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Pattern Recognition
Volume: 49
ISSN (Print): 0031-3203
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 4.83 SJR 2.051 SNIP 3.166
Web of Science (2015): Impact factor 3.999
Web of Science (2015): Indexed yes
Original language: English
Keywords: Regularized generalized eigen-decomposition, Sparse 2D-LDA, Sparse 3D-LDA, Sparse discriminant analysis, Sparse supervised feature extraction
DOIs:
10.1016/j.patcog.2015.07.008
Source: FindIt
Source-ID: 276241928
Research output: Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

Revising an OECD Test Guideline on toxicity in aquatic environments

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Ersbøll, B. K.
Risk factors for Campylobacter infection of broiler flocks

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Division of Risk Assessment and Nutrition, University of Barcelona, Veterinary Institute of Poland, Norwegian Veterinary Institute, Utrecht University, University of Liverpool
Number of pages: 1
Publication date: 2015

Segmentation and characterization of fibers

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Dahl, V. A.
Publication date: 2015
Media of output: PowerPoint

SensMixed R package: Easy-to-use application with graphical user interface for analyzing sensory and consumer data within a mixed effects model framework

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University, Universidade Federal de Lavras, Bang & Olufsen A/S
Short communication: Artificial ultraviolet B light exposure increases vitamin D levels in cow plasma and milk

The number of dairy cows without access to pasture or sunlight is increasing; therefore, the content of vitamin D in dairy products is decreasing. Ultimately, declining vitamin D levels in dairy products will mean that dairy products are a negligible source of natural vitamin D for humans. We tested the ability of a specially designed UVB lamp to enhance the vitamin D-3 content in milk from dairy cows housed indoors. This study included 16 cows divided into 4 groups. Each group was exposed daily to artificial UVB light simulating 1, 2, 3, or 4 h of summer sun at 56 degrees N for 24 d, and the group with simulated exposure to 2 h of summer sun daily continued to be monitored for 73 d. We found a significant increase in 25-hydroxyvitamin D-3 (25OHD(3)) levels in plasma as well as vitamin D-3 and 25OHD(3) levels in milk after daily exposure for 24 d in all treatment groups. Extending daily exposure to artificial UVB light to 73 d did not lead to an increase of vitamin D-3 or 25OHD(3) level in the milk. In conclusion, the change in production facilities for dairy cows providing cows with no access to pasture and sunlight causes a decrease of vitamin D levels in dairy products. This decrease may be prevented by exposing cows to artificial UVB light in the stable.
Significant decrease in yield under future climate conditions: Stability and production of 138 spring barley accessions

The response in production parameters to projected future levels of temperature, atmospheric carbon dioxide ([CO₂]), and ozone ([O₃]) was investigated in 138 spring barley accessions. The comprehensive set of landraces, cultivars, and breeder-lines, were during their entire life cycle exposed to a two-factor treatment of combined elevated temperature (+5°C day/night) and [CO₂] (700 ppm), as well as single-factor treatments of elevated temperature (+5°C day/night), [CO₂] (700 ppm), and [O₃] (100–150 ppb). The control treatment was equivalent to present average South Scandinavian climate (temperature: 19/12°C day/night, [CO₂]: 385 ppm). Overall grain yield was found to decrease 29% in the two-factor treatment with concurrent elevation of [CO₂] and temperature, and this response could not be predicted from the results of treatments with elevated [CO₂] and temperature as single factors, where grain yield increased 16% and decreased 56%, respectively. Elevated [O₃] was found to decrease grain yield by 15%. Substantial variation in response to the applied climate treatments was found between the accessions. The results revealed landraces, cultivars, and breeder-lines with phenotypes applicable for breeding towards stable and high yield under future climate conditions. Further, we suggest identifying resources for breeding under multifactor climate conditions, as single-factor treatments did not accurately forecast the response, when factors were combined.
Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range

1. Predicting and explaining the distribution and density of species is one of the oldest concerns in ecology. Species distributions can be estimated using geostatistical methods, which estimate a latent spatial variable explaining observed variation in densities, but geostatistical methods may be imprecise for species with low densities or few observations. Additionally, simple geostatistical methods fail to account for correlations in distribution among species and generally estimate such cross-correlations as a post hoc exercise.

2. We therefore present spatial factor analysis (SFA), a spatial model for estimating a low-rank approximation to multivariate data, and use it to jointly estimate the distribution of multiple species simultaneously. We also derive an analytic estimate of cross-correlations among species from SFA parameters.

3. As a first example, we show that distributions for 10 bird species in the breeding bird survey in 2012 can be parsimoniously represented using only five spatial factors. As a second case study, we show that forward prediction of catches for 20 rockfishes (Sebastes spp.) off the U.S. West Coast is more accurate using SFA than analysing each species individually. Finally, we show that single-species models give a different picture of cross-correlations than joint estimation using SFA.

4. Spatial factor analysis complements a growing list of tools for jointly modelling the distribution of multiple species and provides a parsimonious summary of cross-correlation without requiring explicit declaration of habitat variables. We conclude by proposing future research that would model species cross-correlations using dissimilarity of species' traits, and the development of spatial dynamic factor analysis for a low-rank approximation to spatial time-series data.
Statistical Data Mining for Efficient Quality Control in Manufacturing

Extensive use of machines, flexible/re-configurable manufacturing and transition towards the fully automated factories call for intelligent use of information recorded during the manufacturing process. Modern manufacturing processes produce Terabytes of information during different stages of the process e.g. sensor measurements, machine readings etc, and the major contributor of these big data sets are different quality control processes. In this article we will present methodology to extract valuable insight from manufacturing data. The proposed methodology is based on comparison of probabilities and extension of likelihood principles in statistics as a performance function for Genetic Algorithm.

Student employment and study effort for engineering students

The aim of this paper is to examine which factors effect student employment and study effort in a setting where engineering students are financially supported, such that their education is free of cost and that they receive financial support for living costs while studying. In addition, we wish to answer if the full-time student is under demise in these settings as opposed to settings without financial support [1, 2]. The research consisted of a web-based survey amongst all students at the Technical University of Denmark (DTU). The students in this survey had fewer employment hours and studied more than those in studies from e.g. UK and US [3, 4, 5]. A similar trend was seen in a study from Norway [6]. Government financial support seems to limit the amount of hours spent on paid work but not the percentage of students who take on paid work. Thus, full-time studies with benefits of increased capabilities and experience gained through employment could be aided by proper policies. Additionally, one of the highest impacts on study activity was the perceived study environment. As the engineering students have four hours per week of interaction with an instructor for each five ECTS, it is to be expected that the students generally spend a majority of their studying hours at the university. This study is to our knowledge the first to study student employment and study effort for Danish engineering students.
Surface Detection and Segmentation

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Dahl, V. A.
Publication date: 2015
Media of output: PowerPoint

Event information
Event: MEK/ Compute seminar
Location: Kgs. Lyngby, Denmark
Electronic versions:
VAND_MEK.pdf
Research output: Non-textual form › Sound/Visual production (digital) – Annual report year: 2015 › Research

Taking individual scaling differences into account by analyzing profile data with the Mixed Assessor Model

Scale range differences between individual assessors will often constitute a non-trivial part of the assessor-by-product interaction in sensory profile data (Brockhoff, 2003, 1998; Brockhoff and Skovgaard, 1994). We suggest a new mixed model ANOVA analysis approach, the Mixed Assessor Model (MAM) that properly takes this into account by a simple inclusion of the product averages as a covariate in the modeling and allowing the covariate regression coefficients to depend on the assessor. This gives a more powerful analysis by removing the scaling difference from the error term and proper confidence limits are deduced that include scaling difference in the error term to the proper extent. A meta study of 8619 sensory attributes from 369 sensory profile data sets from SensoBase (www.sensobase.fr) is conducted. In 45.3% of all attributes scaling heterogeneity is present (P-value <0.05). For the 33.9% of the attributes having a product difference P-value in an intermediate range by the traditional approach, the new approach resulted in a clearly more significant result for 42.3% of these cases. Overall, the new approach claimed significant product difference (P-value <0.05) for 66.1% of the attributes compared to the 60.3% of traditional approach. Still, the new, and non-symmetrical, confidence limits are more often wider than narrower compared to the classical ones: in 72.6% of all cases.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Institute for Agronomic Research, University of Copenhagen
Contributors: Brockhoff, P. B., Schlich, P., Skovgaard, I.
Pages: 156-166
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Food Quality and Preference
Volume: 39
ISSN (Print): 0950-3293
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 3.92 SJR 1.042 SNIP 1.781
The Effect of Autocorrelation on the Hotelling T-2 Control Chart

One of the basic assumptions for traditional univariate and multivariate control charts is that the data are independent in time. For the latter, in many cases, the data are serially dependent (autocorrelated) and cross-correlated because of, for example, frequent sampling and process dynamics. It is well known that the autocorrelation affects the false alarm rate and the shift-detection ability of the traditional univariate control charts. However, how the false alarm rate and the shift-detection ability of the Hotelling T-2 control chart are affected by various autocorrelation and cross-correlation structures for different magnitudes of shifts in the process mean is not fully explored in the literature. In this article, the performance of the Hotelling T-2 control chart for different shift sizes and various autocorrelation and cross-correlation structures are compared based on the average run length using simulated data. Three different approaches in constructing the Hotelling T-2 chart are studied for two different estimates of the covariance matrix: (i) ignoring the autocorrelation and using the raw data with theoretical upper control limits; (ii) ignoring the autocorrelation and using the raw data with adjusted control limits calculated through Monte Carlo simulations; and (iii) constructing the control chart for the residuals from a multivariate time series model fitted to the raw data. To limit the complexity, we use a first-order vector autoregressive process and focus mainly on bivariate data. (c) 2014 The Authors. Quality and Reliability Engineering International published by John Wiley & Sons Ltd.

The effect of skatole and androstenone on consumer response towards streaky bacon and pork belly roll

Consumer liking was assessed for streaky bacon and pork belly roll from entire male pigs with an androstenone (AND) content of up to 9.4ppm and a skatole (SKA) content of up to 0.92ppm in the back fat and castrates. No clear effect of either AND or SKA was seen in consumer liking, although an insignificant tendency was seen for SKA. A sensory profile analysis showed that AND increased the boar taint of bacon, while both AND and SKA increased the boar taint of the pork belly roll. Consumer sensitivity towards AND and SKA did not affect liking of the meat products. The lack of effect of AND and SKA on consumer liking could be due to a masking effect of the spices and smoke. Three consecutive weeks'
exposure to bacon did not change the liking score, irrespective of the AND and SKA content. This indicates that the consumers did not become more sensitive towards boar taint.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Danish Technological Institute
Contributors: Aaslyng, M. D., De Lichtenberg Broge, E. H., Brockhoff, P. B., Christensen, R. H. B.
Pages: 52-61
Publication date: 2015
Peer-reviewed: Yes

**Publication information**
Journal: Meat Science
Volume: 110
ISSN (Print): 0309-1740
Ratings:
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 3.04 SJR 1.963 SNIP 1.662
Web of Science (2015): Impact factor 2.801
Web of Science (2015): Indexed yes
Original language: English
Keywords: Entire male, Boar taint, Consumer study, Tenderness, Central location test, Home use test, Sensory profile, Bacon, Pork belly roll

**DOIs:**
10.1016/j.meatsci.2015.07.001
Source: FindIt
Source-ID: 2279691237

**Research output:** Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

**The likelihood principle and its proof – a never-ending story...**

An ongoing controversy in philosophy of statistics is the so-called "likelihood principle" essentially stating that all evidence which is obtained from an experiment about an unknown quantity \( \theta \) is contained in the likelihood function of \( \theta \). Common classical statistical methodology, such as the use of significance tests, and confidence intervals, depends on the experimental procedure and unrealized events and thus violates the likelihood principle. The likelihood principle was identified by that name and proved in a famous paper by Allan Birnbaum in 1962. However, ever since both the principle itself as well as the proof has been highly debated. This presentation will illustrate the debate of both the principle and its proof, from 1962 and up to today. An often-used experiment to illustrate the controversy between classical interpretation and evidential confirmation based on the likelihood principle will also be presented and discussed.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Jørgensen, T. M.
Number of pages: 8
Publication date: 2015

**Host publication information**
Title of host publication: Symposium i anvendt statistik 2015.
Publisher: Danmarks Statistik
ISBN (Print): 978-87-501-2171-8
Source: PublicationPreSubmission
Source-ID: 113807767
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2015 › Research › peer-review

**Two segmentation methods**

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Dahl, V. A.
Publication date: 2015
Media of output: PowerPoint
VirtualTable: a projection augmented reality game

VirtualTable is a projection augmented reality installation where users are engaged in an interactive tower defense game. The installation runs continuously and is designed to attract people to a table, which the game is projected onto. Any number of players can join the game for an optional period of time. The goal is to prevent the virtual stylized soot balls, spawning on one side of the table, from reaching the cheese. To stop them, the players can place any kind of object on the table, that then will become part of the game. Depending on the object, it will become either a wall, an obstacle for the soot balls, or a tower, that eliminates them within a physical range. The number of enemies is dependent on the number of objects in the field, forcing the players to use strategy and collaboration and not the sheer number of objects to win the game.

General information
Publication status: Published
Organisations: IT Service, Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Mathematics, Statistics and Data Analysis
Number of pages: 1
Publication date: 2015

Host publication information
Title of host publication: Proceedings of SIGGRAPH Asia 2015 Posters
Publisher: Association for Computing Machinery
Article number: 40
ISBN (Print): 978-1-4503-3926-1
Electronic versions:
VirtualTable_abstract.pdf
DOIs:
10.1145/2820926.2820950
Source: PublicationPreSubmission
Source-ID: 117830919
Research output: Chapter in Book/Report/Conference proceeding » Conference abstract in proceedings – Annual report year: 2015 » Research » peer-review

Vi skal sørge for at kvinder også passer ind

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Rootzén, H.
Pages: 22-23
Publication date: 2015
Peer-reviewed: Unknown

Publication information
Journal: Alt for Damerne
Volume: 2015
Issue number: 28
ISSN (Print): 0002-6506
Original language: Danish
Source: PublicationPreSubmission
Source-ID: 118086588
Research output: Contribution to journal » Journal article – Annual report year: 2015 » Communication

Vitamin D intake-status relationship among Danes aged 4-60 years during winter

General information
What do Danish children eat, and does the diet meet the recommendations? Baseline data from the OPUS School Meal Study

A child's diet is an important determinant for later health, growth and development. In Denmark, most children in primary school bring their own packed lunch from home and attend an after-school care institution. The aim of the present study was to evaluate the food, energy and nutrient intake of Danish school children in relation to dietary guidelines and nutrient recommendations, and to assess the food intake during and outside school hours. In total, 834 children from nine public schools located in the eastern part of Denmark were included in this cross-sectional study and 798 children (95·7 %) completed the dietary assessment sufficiently (August-November 2011). The whole diet was recorded during seven consecutive days using the Web-based Dietary Assessment Software for Children (WebDASC). Compared with the food-based dietary guidelines and nutrient recommendations, 85 % of the children consumed excess amounts of red meat, 89 % consumed too much saturated fat, and 56 % consumed too much added sugar. Additionally 35 or 91 % of the children (depending on age group) consumed insufficient amounts of fruits and vegetables, 85 % consumed insufficient amounts of fish, 86 % consumed insufficient amounts of dietary fibre, 60 or 84 % had an insufficient Fe intake (depending on age group), and 96 % had an insufficient vitamin D intake. The study also showed that there is a higher intake of fruits and bread during school hours than outside school hours; this is not the case with, for example, fish and vegetables, and future studies should investigate strategies to increase fish and vegetable intake during school hours.

General information

Publication status: Published
Organisations: National Food Institute, Research group for Risk Benefit, Division of Risk Assessment and Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Number of pages: 9
Publication date: 2015
Peer-reviewed: Yes

Publication information
Journal: Journal of Nutritional Science
Volume: 4
Article number: e29
Ratings:
Scopus rating (2015): SJR 0.381 SNIP 0.376
Web of Science (2015): Indexed yes
Original language: English
Which sensory perception is primarily considered, in consumers' hedonic evaluation of foods?
An analysis of the primary hedonic drivers of liking and sensory satisfaction will provide valuable information to product developers on which sensory properties to emphasise the most. The aims of the present study were: a) to study if liking of the sensory properties: appearance, odour, taste and texture were considered equally, when consumers rated overall liking and sensory satisfaction b) to study if the relation depended on, whether liking of sensory properties were related to overall liking or sensory satisfaction, and c) to study individual differences in which sensory properties the consumers primarily paid attention to when rating overall liking and sensory satisfaction, respectively. Four apple-cherry fruit drinks were used, varying in: type of sweetener, and addition of aroma and fibre. The fruit drinks were used in a in a cross-over consumer study on 67 subjects together with sensory profiling. For data analysis mixed three-way analysis of variance and principal component analysis was applied to study and visualise sensory differences. The relative importance of liking of sensory properties; appearance, odour, taste and texture was analysed using slopes, when consumers rated overall liking and sensory satisfaction, respectively. Sensory profiling confirmed that fruit drinks varied in sensory attributes related to appearance, odour, taste and texture. From the consumer study a significant main product effect was found for the six response variables; overall liking, liking of appearance, odour, taste and texture respectively, and sensory satisfaction. Further, liking of sensory properties differed in relation to overall liking and sensory satisfaction, respectively. Consumers primarily paid attention to liking of taste, when evaluating overall liking and sensory satisfaction, respectively. However, individual differences were found.

Les éléphants en danger

Analysis of micro-structure in raw and heat treated meat emulsions from multimodal X-ray microtomography
This study presents a novel non-destructive X-ray technique for analyzing meat emulsions before and after heat treatment. The method is based on X-ray grating-interferometry where three complementary imaging modalities are obtained simultaneously measuring the absorption, refraction and scattering properties of the sample. Enhanced contrast capabilities of this X-ray technique makes studies on materials with similar attenuation properties possible. The emulsion samples were imaged both in a raw and cooked state. Additionally, different fat types were used in the emulsions in order to compare microstructural differences when either pork fat or sunflower oil was added. From the reconstructed tomograms the different constituents in the emulsions were segmented using a multivariate segmentation method. From
this, a quantitative analysis was performed between the different samples, determining properties such as percent object volumes, porosity, average structure thickness and cooking loss. The grating-based X-ray technique and multivariate segmentation made the analysis of the microstructure possible which further gives insight to how both heat treatment, and the use of different lipid types, affect the final protein network quality.

Industrial relevance: Meat emulsions have previously been thoroughly studied, and the use of various fat substitutes and protein stabilizers has been investigated. The grating-based multimodal X-ray tomography method presented here is a feasible method to investigate the microstructural changes induced by heat treatment. It provides high-resolution three dimensional spatial information and in contrast to 2D imaging methods, quantitative parameters can be extracted by image analysis for the entire sample volume. Additionally, the non-destructive method allows for imaging the same sample before and after cooking.

A Comparison between Two Simulation Models for Spread of Foot-and-Mouth Disease

Two widely used simulation models of foot-and-mouth disease (FMD) were used in order to compare the models’ predictions in term of disease spread, consequence, and the ranking of the applied control strategies, and to discuss the effect of the way disease spread is modeled on the predicted outcomes of each model. The DTU-DADS (version 0.100), and ISP (version 2.001.11) were used to simulate a hypothetical spread of FMD in Denmark. Actual herd type, movements, and location data in the period 1st October 2006 and 30th September 2007 was used. The models simulated the spread of FMD using 3 different control scenarios: 1) A basic scenario representing EU and Danish control strategies, 2) pre-emptive depopulation of susceptible herds within a 500 meters radius around the detected herds, and 3) suppressive vaccination of susceptible herds within a 1,000 meters radius around the detected herds. Depopulation and vaccination started 14 days following the detection of the first infected herd. Five thousand index herds were selected randomly, of which there were 1,000 cattle herds located in high density cattle areas and 1,000 in low density cattle areas, 1,000 swine herds located in high density swine areas and 1,000 in low density swine areas, and 1,000 sheep herds. Generally, DTU-DADS predicted larger, longer duration and costlier epidemics than ISP, except when epidemics started in cattle herds located in high density cattle areas. ISP supported suppressive vaccination rather than pre-emptive depopulation, while DTU-DADS was indifferent to the alternative control strategies. Nonetheless, the absolute differences between control strategies were small making the choice of control strategy during an outbreak to be most likely based on practical reasons.
A national drug related problems database: evaluation of use in practice, reliability and reproducibility

Background A drug related problems database (DRP-database) was developed on request by clinical pharmacists. The information from the DRP-database has only been used locally e.g. to identify focus areas and to communicate identified DRPs to the hospital wards. Hence the quality of the data at the national level is unknown, which may compromise national analyses for benchmarking and identification of national focus areas. Objective The aim of the study was to evaluate the use in practice, reliability and reproducibility of the DRPs documented in the Danish drug related problems database. Setting Danish hospital pharmacies. Methods Practice use of the DRP-database was explored by an electronic questionnaire distributed to hospital pharmacies, and consisted of questions regarding current and previous use of the DRP-database. The reliability was evaluated by comparing the categorization of 24 cases by clinical pharmacists with categorization performed by the project group. Reproducibility was explored by re-categorization of a sample of existing records in the DRP-database by two project group members individually. Main outcome measures Observed proportion of agreement and Fleiss’ kappa as measures of inter-rater reliability and reproducibility. Results The practice use study of 12 hospital pharmacy locations revealed that when implementing the DRP-database, the majority of identified DRPs are documented in the DRP-database, however, some variations throughout the country exist. The interrater reliability study of 34 clinical pharmacists showed high inter-rater reliability with the project group (Fleiss’ kappa = 0.79 with 95 % CI (0.70; 0.88)), and the reproducibility study also documented high inter-rater reliability of a sample of 379 records from the DRP-database re-categorized by two project group members (Fleiss’ kappa = 0.81 with 95 % CI (0.78; 0.85)). Conclusion The study showed high reliability and reproducibility of the DRP-database, however, some local variation in the use of the DRP-database throughout the country existed affecting the overall quality. These findings indicate that data in the DRP-database may be pooled, and national analyses may be conducted to explore development areas for common interest.
A Thurstonian comparison of the Tetrad and Degree of Difference tests
The recurring need to assess product reformulations has kept difference testing at the forefront of sensory science. Within the realm of difference testing, the Tetrad test has risen in popularity recently as its superiority over the Triangle test has been demonstrated both in theory and in practice. But it remains to compare the Tetrad test in detail with other commonly used testing methods such as the Degree of Difference (DOD) test. In this paper, we provide such a comparison by considering, from a theoretical perspective, the differences between both power and precision for the Tetrad and DOD tests. In particular we show that, theoretically and for the range of sensory effect sizes likely to be of interest in consumer research, the Tetrad test is more powerful and more precise than the DOD test. Even so, if there is substantially more perceptual noise in the Tetrad test from the two additional stimuli, it is possible that performance of the DOD could surpass the performance of the Tetrad test in practice. To investigate this last statement, we quantify the additional noise required to negate the theoretical advantage of the Tetrad test.
Automatic sleep classification using a data-driven topic model reveals latent sleep states

Background: The golden standard for sleep classification uses manual scoring of polysomnography despite points of criticism such as oversimplification, low inter-rater reliability and the standard being designed on young and healthy subjects. New method: To meet the criticism and reveal the latent sleep states, this study developed a general and automatic sleep classifier using a data-driven approach. Spectral EEG and EOG measures and eye correlation in 1 s windows were calculated and each sleep epoch was expressed as a mixture of probabilities of latent sleep states by using the topic model Latent Dirichlet Allocation. Model application was tested on control subjects and patients with periodic leg movements (PLM) representing a non-neurodegenerative group, and patients with idiopathic REM sleep behavior disorder (iRBD) and Parkinson's Disease (PD) representing a neurodegenerative group. The model was optimized using 50 subjects and validated on 76 subjects. Results: The optimized sleep model used six topics, and the topic probabilities changed smoothly during transitions. According to the manual scorings, the model scored an overall subject-specific accuracy of 68.3 +/- 7.44 (% mu +/-sigma) and group specific accuracies of 69.0 +/- 4.62 (control), 70.1 +/- 5.10 (PLM), 67.2 +/- 8.30 (iRBD) and 67.7 +/- 9.07 (PD). Comparison with existing method: Statistics of the latent sleep state content showed accordances to the sleep stages defined in the golden standard. However, this study indicates that sleep contains six diverse latent sleep states and that state transitions are continuous processes. Conclusions: The model is generally applicable and may contribute to the research in neurodegenerative diseases and sleep disorders. (C) 2014 Elsevier B.V. All rights reserved.

General information
Publication status: Published
Organisations: Department of Electrical Engineering, Biomedical Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Technical University of Denmark, University of Copenhagen, H. Lundbeck A/S
Contributors: Koch, H., Christensen, J. A. E., Frandsen, R., Zoetmulder, M., Arvastson, L. J., Christensen, S. R., Jennum, P., Sørensen, H. B. D.
Pages: 130-137
Publication date: 2014
Peer-reviewed: Yes

Publication information
Journal: Journal of Neuroscience Methods
Volume: 235
ISSN (Print): 0165-0270
Ratings:
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 2.25 SJR 1.035 SNIP 0.85
Web of Science (2014): Impact factor 2.025
Web of Science (2014): Indexed yes
Original language: English
Keywords: BIOCHEMICAL, NEUROSCIENCES, Automatic sleep classification, Topic modeling, Electroencephalography (EEG), Electrooculography (EOG), Sleep state switching, Neurodegenerative diseases, latent sleep state, idiopathic REM sleep behavior disorder nervous system disease, Parkinson's disease nervous system disease, periodic leg movement nervous system disease, Primates Mammalia Vertebrata Chordata Animalia (Animals, Chordates, Humans, Mammals, Primates, Vertebrates) - Hominidae [86215] human common middle age, aged female, male, 04500, Mathematical biology and statistical methods, 10515, Biophysics - Biocybernetics, 20004, Sense organs - Physiology and biochemistry, 20506, Nervous system - Pathology, 24500, Gerontology, Computational Biology, Human Medicine, Medical Sciences, eye sensory system, EEG electroencephalography clinical techniques, diagnostic techniques, electrooculography clinical techniques, diagnostic techniques, optimized sleep model mathematical and computer techniques, topic model Latent Dirichlet Allocation mathematical and computer techniques, Models and Simulations, Neurology

Can We Find Organic Materials in Food Using X-rays?

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis

10.1016/j.foodqual.2014.05.004
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review
Common variants in CYP2R1 and GC genes are both determinants of serum 25-hydroxyvitamin D concentrations after UVB irradiation and after consumption of vitamin D3-fortified bread and milk during winter in Denmark

Background: Little is known about how the genetic variation in vitamin D modulating genes influences ultraviolet (UV)B–induced 25-hydroxyvitamin D [25(OH)D] concentrations. In the Food with vitamin D (VitmaD) study, we showed that common genetic variants rs10741657 and rs10766197 in 25-hydroxylase (CYP2R1) and rs842999 and rs4588 in vitamin D binding protein (GC) predict 25(OH)D concentrations at late summer and after 6-mo consumption of cholecalciferol (vitamin D3)–fortified bread and milk.

Objectives: In the current study, called the Vitamin D in genes (VitDgen) study, we analyzed associations between the increase in 25(OH)D concentrations after a given dose of artificial UVB irradiation and 25 single nucleotide polymorphisms located in or near genes involved in vitamin D synthesis, transport, activation, or degradation as previously described for the VitmaD study. Second, we aimed to determine whether the genetic variations in CYP2R1 and GC have similar effects on 25(OH)D concentrations after artificial UVB irradiation and supplementation by vitamin D3–fortified bread and milk.

Design: The VitDgen study includes 92 healthy Danes who received 4 whole-body UVB treatments with a total dose of 6 or 7.5 standard erythema doses during a 10-d period in winter. The VitmaD study included 201 healthy Danish families who were given vitamin D3–fortified bread and milk or placebo for 6 mo during the winter.

Results: After UVB treatments, rs10741657 in CYP2R1 and rs4588 in GC predicted UVB-induced 25(OH)D concentrations as previously shown in the VitmaD study. Compared with noncarriers, carriers of 4 risk alleles of rs10741657 and rs4588 had lowest concentrations and smallest increases in 25(OH)D concentrations after 4 UVB treatments and largest decreases in 25(OH)D concentrations after 6-mo consumption of vitamin D3–fortified bread and milk.

Conclusion: Common genetic variants in the CYP2R1 and GC genes modify 25(OH)D concentrations in the same manner after artificial UVB-induced vitamin D and consumption of vitamin D3–fortified bread and milk. The VitDgen study was registered at clinicaltrials.gov as NCT01741233. The VitmaD study was registered at clinicaltrials.gov as NCT01184716.
Common Variants in CYP2R1 and GC Genes Predict Vitamin D Concentrations in Healthy Danish Children and Adults

Environmental factors such as diet, intake of vitamin D supplements and exposure to sunlight are known to influence serum vitamin D concentrations. Genetic epidemiology of vitamin D is in its infancy and a better understanding on how genetic variation influences vitamin D concentration is needed. We aimed to analyse previously reported vitamin D-related polymorphisms in relation to serum 25(OH)D concentrations in 201 healthy Danish families with dependent children in late summer in Denmark. Serum 25(OH)D concentrations and a total of 25 SNPs in GC, VDR, CYP2R1, CYP24A1, CYP27B1, C10or88 and DHCR7/NADSYN1 genes were analysed in 758 participants. Genotype distributions were in Hardy-Weinberg equilibrium for the adult population for all the studied polymorphisms. Four SNPs in CYP2R1 (rs1562902, rs7116978, rs10741657 and rs10766197) and six SNPs in GC (rs4588, rs842999, rs2282679, rs12512631, rs16846876 and rs17468725) were statistically significantly associated with serum 25(OH)D concentrations in children, adults and all combined. Several of the SNPs were in strong linkage disequilibrium, and the associations were driven by CYP2R1-rs10741657 and rs10766197, and by GC-rs4588 and rs842999. Genetic risk score analysis showed that carriers with no risk alleles of CYP2R1-rs10741657 and rs10766197, and/or GC rs4588 and rs842999 had significantly higher serum 25(OH)D concentrations compared to carriers of all risk alleles. To conclude, our results provide supporting evidence that common polymorphisms in GC and CYP2R1 are associated with serum 25(OH)D concentrations in the Caucasian population and that certain haplotypes may predispose to lower 25(OH)D concentrations in late summer in Denmark.
Consideration of sample heterogeneity and in-depth analysis of individual differences in sensory analysis

In descriptive sensory analysis, large variations may be observed between scores. Individual differences between assessors have been identified as one cause for these variations. Much work has been done on modeling these differences and accounting for them through analysis of variance (ANOVA). When the products studied are prone to biological heterogeneity (e.g. fruits, vegetables, cheeses, etc.), variations in the data may be due to assessor differences and/or product heterogeneity. The present paper proposes an approach for quantifying these two sources of variation. For individual differences, an extended version of the assessor model approach is applied. The data set used in the paper is based on sensory evaluations of three apple samples scored by a panel of 19 assessors using seven descriptors in four replicates. The application of the extended assessor model approach to unbalanced data provides more insight into assessor differences and a better test for product differences. These results demonstrate the importance of choosing the right model and taking all potential sources of variation into account.

Danish children born to parents with lower levels of education are more likely to become overweight

AIM:
Little is known about whether the socio-economic status of parents is linked to their children becoming overweight. This study examined the association between parents’ educational level and overweight Danish children in a nationally representative sample.

METHODS:
Body mass index was calculated for a random sample of 512 children aged from four to 14 from the Danish National Survey of Diet and Physical Activity 2005-2008. Their parents provided weight and height data during an interview, together with details of their own educational level. Children were classified as overweight/obese in accordance with the International Obesity Task Force. Frequency estimates of prevalence and logistic regression models were used to correlate childhood overweight/obesity with the mothers’ and fathers’ educational levels as the main outcome measures.

RESULTS:
Danish mothers tended to be more highly educated than fathers and their educational level was inversely associated with their child being overweight, especially if it was a boy. However, the highest educational level of the parents was the only significant educational variable, suggesting that education was associated with overweight children irrespective of the gender of the parent.
CONCLUSION:
Public health initiatives should target parents with low educational levels to prevent, and reduce, social inequality in overweight children.

Data-driven modeling of sleep EEG and EOG reveals characteristics indicative of pre-Parkinson's and Parkinson's disease

Background: Manual scoring of sleep relies on identifying certain characteristics in polysomnograph (PSG) signals. However, these characteristics are disrupted in patients with neurodegenerative diseases. New method: This study evaluates sleep using a topic modeling and unsupervised learning approach to identify sleep topics directly from electroencephalography (EEG) and electrooculography (EOG). PSG data from control subjects were used to develop an EEG and an EOG topic model. The models were applied to PSG data from 23 control subjects, 25 patients with periodic leg movements (PLMs), 31 patients with idiopathic REM sleep behavior disorder (iRBD) and 36 patients with Parkinson's disease (PD). The data were divided into training and validation datasets and features reflecting EEG and EOG characteristics based on topics were computed. The most discriminative feature subset for separating iRBD/PD and PLM/controls was estimated using a Lasso-regularized regression model. Results: The features with highest discriminability were the number and stability of EEG topics linked to REM and N3, respectively. Validation of the model indicated a sensitivity of 91.4% and a specificity of 68.8% when classifying iRBD/PD patients. Comparison with existing method: The topics showed visual accordance with the manually scored sleep stages, and the features revealed sleep characteristics containing information indicative of neurodegeneration. Conclusions: This study suggests that the amount of N3 and the ability to maintain NREM and REM sleep have potential as early PD biomarkers. Data-driven analysis of sleep may contribute to the evaluation of neurodegenerative patients. (C) 2014 Elsevier B.V. All rights reserved.
Demystifying back scatter interferometry: a sensitive refractive index detector.

BACKGROUND: Back Scatter Interferometry (BSI) is a sensitive method for detecting changes of the refractive index (RI) in small capillaries. The method was originally developed as an off-axial column detector for use in Liquid Chromatography or Capillary Electrophoresis systems, but it has been proposed that this method can also be used to detect molecular binding in a label-free manner. Recent work proposes BSI to be a unique sensor for detecting protein binding with various ligands and other protein interactions in order to obtain relevant binding kinetics. We hypothesize that BSI is actually acting like a common-path interferometer. METHODS: A HeNe laser is directed at a glass capillary with inner diameter of 1.4 mm and reflected light from air/glass and liquid/glass interfaces interfere to form an RI dependent intensity fringe pattern at a CCD detector. The fringe shift relative to the change of RI of the sample; i.e. the sensitivity; is controlled by the physical interaction length between the interferometric sample beam and the sample itself. Using optical ray-tracing we calculate the sensitivity. We validate these theoretical findings by determining the RI increment (dn/dc) from a set of NaCl standard solutions. RESULTS: Ray-tracing show that the basic interference pattern recorded with BSI can be fully described by two beams, one reflected from the surface of the capillary and a beam reflected from the back of the capillary wall. In accordance we find that the interferometric interaction length is given by twice the diameter of the capillary. Experimentally we find a sensitivity of 4700 rad/(g/ml) and estimate dn/dc for NaCl to be 0.169 ml/g, which is in accordance with literature. Furthermore we report a minimum detectability of 7x10^-7 RI Units. CONCLUSIONS: BSI works like a common-path interferometer. The sensitivity of the BSI system is given by twice the inner diameter of the capillary times the wavenumber of the light source. Our results suggest that Back Scatter Interferometry does not provide a unique measurement principle for sensing biochemical bindings compared to what should be possible using many commercial available refractometers.
Detecting correlation between allele frequencies and environmental variables as a signature of selection. A fast computational approach for genome-wide studies

Genomic regions (or loci) displaying outstanding correlation with some environmental variables are likely to be under selection and this is the rationale of recent methods of identifying selected loci and retrieving functional information about them. To be efficient, such methods need to be able to disentangle the potential effect of environmental variables from the confounding effect of population history. For the routine analysis of genome-wide datasets, one also needs fast inference and model selection algorithms. We propose a method based on an explicit spatial model which is an instance of spatial generalized linear mixed model (SGLMM). For inference, we make use of the INLA–SPDE theoretical and computational framework developed by Rue et al. (2009) and Lindgren et al. (2011). The method we propose allows one to quantify the correlation between genotypes and environmental variables. It works for the most common types of genetic markers, obtained either at the individual or at the population level. Analyzing the simulated data produced under a geostatistical model then under an explicit model of selection, we show that the method is efficient. We also re-analyze a dataset relative to nineteen pine weevils (Hylobius abietis) populations across Europe. The method proposed appears also as a statistically sound alternative to the Mantel tests for testing the association between the genetic and environmental variables.

Dictionary Snakes

Visual cues like texture, color and context make objects appear distinct from the surroundings, even without gradients between regions. Texture-rich objects are often difficult to segment because algorithms need advanced features which are unique for the image. In this paper we suggest a method for image segmentation that operates without training data. Our method is based on a probabilistic dictionary of image patches coupled with a deformable model inspired by snakes and active contours without edges. We separate the image into two classes based on the information provided by the evolving curve, which moves according to the probabilistic information obtained from the dictionary. Initially, the image patches are assigned to the nearest dictionary element, where the image is sampled at each pixel such that patches overlap. The curve divides the image into an inside and an outside region allowing us to estimate the pixel-wise probability of the dictionary elements. In each iteration we evolve the curve and update the probabilities, which merges similar texture...
patterns and pulls dissimilar patterns apart. We experimentally evaluate our approach, and show how textured objects are precisely segmented without any prior assumptions about image features. In addition, a texture probability image is obtained.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Dahl, A. B., Dahl, V. A.
Pages: 142-147
Publication date: 2014

**Host publication information**
Title of host publication: Proceedings of the 22nd International Conference on Pattern Recognition (ICPR 2014)
Publisher: IEEE
Editor: O’Connor, L.
ISBN (Print): 978-1-4799-5208-3

**Electronic versions:**
DicSnake
DOIs:
10.1109/ICPR.2014.34

**Research output:** Chapter in Book/Report/Conference proceeding» Article in proceedings – Annual report year: 2014» Research» peer-review

---

**Dietary effects of introducing school meals based on the New Nordic Diet: a randomised controlled trial in Danish children.**

**The OPUS School Meal Study**
The OPUS (Optimal Well-being, development and health for Danish children through a healthy New Nordic Diet (NND)) School Meal Study investigated the effects on the intake of foods and nutrients of introducing school meals based on the principles of the NND covering lunch and all snacks during the school day in a cluster-randomised cross-over design. For two 3-month periods, 834 Danish children aged 8-11 years from forty-six school classes at nine schools received NND school meals or their usual packed lunches brought from home (control) in random order. The whole diet of the children was recorded over seven consecutive days using a validated Web-based Dietary Assessment Software for Children. The NND resulted in higher intakes of potatoes (130 %, 95 % CI 2·07, 2·58), fish (48 %, 95 % CI 1·33, 1·65), cheese (25 %, 95 % CI 1·15, 1·36), vegetables (16 %, 95 % CI 1·10, 1·21), eggs (10 %, 95 % CI 1·01, 1·19) and beverages (6 %, 95 % CI 1·02, 1·09), and lower intakes of bread (13 %, 95 % CI 0·84, 0·89) and fats (6 %, 95 % CI 0·90, 0·98) were found among the children during the NND period than in the control period (all, P< 0·05). No difference was found in mean energy intake (P= 0·4), but on average children reported 0·9 % less energy intake from fat and 0·9 % higher energy intake from protein during the NND period than in the control period. For micronutrient intakes, the largest differences were found for vitamin D (42 %, 95 % CI 1·32, 1·53) and iodine (11 %, 95 % CI 1·08, 1·15) due to the higher fish intake. In conclusion, the present study showed that the overall dietary intake at the food and nutrient levels was improved among children aged 8-11 years when their habitual packed lunches were replaced by school meals following the principles of the NND.

**General information**
Publication status: Published
Organisations: National Food Institute, Division of Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Pages: 1967–1976
Publication date: 2014
Peer-reviewed: Yes

**Publication information**
Journal: British Journal of Nutrition
Volume: 111
Issue number: 11
ISSN (Print): 0007-1145
Ratings:
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 3.18 SJR 1.364 SNIP 1.149
Web of Science (2014): Indexed yes
Original language: English
Discussion of *The Statistical Evaluation of Categorical Measurements: Simple Scales, but Treacherous Complexity Underneath*" 

The authors start their review with the discussion on the kappa statistic commonly used in agreement studies. The practical appeal of single summary statistics to summarize an entire statistical analysis such as the ubiquitous R² cannot be denied. However, one should always be quite cautious in concluding solely based on these statistics. As indicated by the authors, the kappa statistic is no different. Often defined as a sample statistic, potentially lead to inconsistent results. In his early work in this area, Professor de Mast explored the paradoxical behavior of this statistic. His more recent work on introducing more sophisticated methods in alleviating some of the issues discussed in this article is certainly worth praise. However, considering the widespread use of the kappa statistic, it is also worth looking into some of the authors’ comments on the behavior of this statistic.

Disparities in dietary habits and physical activity in Denmark and trends from 1995 to 2008. 

In all analyses, both diet quality and physical inactivity differed systematically with educational group and with healthier habits for the group with long education. Only for alcohol habits was an opposite social gradient seen. Disparities persisted over all time periods. The analyses using "The Concentration Index" did not change any of the conclusions. Gender also influenced diet quality. Men had dietary and alcohol habits furthest away from the guidelines. A trend of healthier habits was also found over the time period. CONCLUSIONS SOCIAL DISPARITIES IN DIET QUALITY AND LEISURE-TIME PHYSICAL ACTIVITY WERE SEEN FOR ALL EXAMINED TIME PERIODS AND WITH NO SIGNIFICANT CHANGE THEREFORE, THERE IS NO SUPPORT TO THE HYPOTHESIS OF WIDENING SOCIAL DISPARITIES IN ALL EDUCATIONAL GROUPS SOME IMPROVEMENT OF DIETARY HABITS WAS FOUND, ESPECIALLY FOR THOSE FOODS WHERE COMPREHENSIVE INITIATIVES HAD TAKEN PLACE HOWEVER, SOCIAL DISPARITIES STILL EXIST THERE IS AN ONGOING NEED TO ADDRESS THESE DISPARITIES IN ORDER TO REDUCE SOCIAL INEQUALITY IN HEALTH:
Effectiveness of offering healthy labelled meals in improving the nutritional quality of lunch meals eaten in a worksite canteen

Healthier meal selections at restaurants and cafeterias are often limited and not actively promoted. In this Danish study the effectiveness of a healthy labelling certification program in improving dietary intake and influencing edible plate waste was evaluated in a quasi-experimental study design. Employees from an intervention worksite canteen and a matched control canteen were included in the study at baseline (February 2012), after completing the certification process (end-point) and six month from end-point (follow-up) (total n=270). In order to estimate nutrient composition of the consumed lunch meals and plate waste a validated digital photographic method was used combining estimation of food intake with food nutrient composition data. Food satisfaction was rated by participants using a questionnaire. Several significant positive nutritional effects were observed at the intervention canteen including a mean decrease in energy density in the consumed meals from 561kJ/100g at baseline to 368 and 407kJ/100g at end-point and follow-up, respectively (P<0.001). No significant changes were seen with regard to food satisfaction and plate waste. In the control canteen no positive nutritional effects were observed. The results of the study highlight the potential of using healthy labelling certification programs as a possible driver for increasing both the availability and awareness of healthy meal choices, thereby improving dietary intake when eating out.
Enabling optimization in LCA: from "ad hoc" to "structural" LCA approach—based on a biodiesel well-to-wheel case study

Purpose Applied life cycle assessment (LCA) studies often lead to a comparison of rather few alternatives; we call this the “ad hoc LCA approach.” This can seem surprising since applied LCAs normally cover countless options for variations and derived potentials for improvements in a product life cycle. In this paper, we will suggest an alternative approach to the ad hoc approach, which more systematically addresses the many possible variations to identify the most promising. We call it the “structural LCA approach.” The goals of this paper are (1) to provide basic guidelines for the structural approach, including an easy expansion of the LCA space; (2) to show that the structural LCA approach can be used for different types of optimization in LCA; and (3) to improve the transparency of the LCA work.

Methods The structural approach is based on the methodology “design of experiments” (Montgomery 2005). Through a biodiesel well-to-wheel study, we demonstrate a generic approach of applying explanatory variables and corresponding impact categories within the LCA methodology. Explanatory variables are product system variables that can influence the environmental impacts from the system. Furthermore, using the structural approach enables two different possibilities for optimization: (1) single-objective optimization (SO) based on response surface methodology (Montgomery 2005) and (2) multiobjective optimization (MO) by the hypervolume estimation taboo search (HETS) method. HETS enables MO for more than two or three objectives.

Results and discussion Using SO, the explanatory variable “use of residual straw from fields” is, by far, the explanatory variable that can contribute with the highest decrease of climate change potential. For the respiratory inorganics impact category, the most influencing explanatory variable is found to be the use of different alcohol types (bioethanol or petrochemical methanol) in biodiesel production. Using MO, we found the Pareto front based on 5 different life cycle pathways which are nondominated solutions out of 66 different analyzed solutions. Given that there is a fixed amount of resources available for the LCA practitioner, it becomes a prioritizing problem whether to apply the structural LCA approach or not. If the decision maker only has power to change a single explanatory variable, it might not be beneficial to apply the structural LCA approach. However, if the decision maker (such as decision makers at the societal level) has power to change more explanatory variables, then the structural LCA approach seems beneficial for quantifying and comparing the potentials for environmental improvement between the different explanatory variables in an LCA system and identifying the overall most promising product system configurations among the chosen PWs.

Conclusions The implementation of the structural LCA approach and the derived use of SO and MO have been successfully achieved and demonstrated in the present paper. In addition, it is demonstrated that the structural LCA approach can lead to more transparent LCAs since the potentially most important explanatory variables which are used to model the LCAs are explicitly presented through the structural LCA approach. The suggested structural approach is a new approach to LCA and it seems to be a promising approach for searching or screening product systems for environmental optimization potentials. In the presented case, the design has been a rather simple full factorial design. More complicated problems or designs, such as fractional designs, nested designs, split plot designs, and/or unbalanced data, in the context of LCA could be investigated further using the structural approach.
Estimating spatio-temporal dynamics of size-structured populations
Spatial distributions of structured populations are usually estimated by fitting abundance surfaces for each stage and at each point of time separately, ignoring correlations that emerge from growth of individuals. Here, we present a statistical model that combines spatio-temporal correlations with simple stock dynamics, to estimate simultaneously how size distributions and spatial distributions develop in time. We demonstrate the method for a cod population sampled by trawl surveys. Particular attention is paid to correlation between size classes within each trawl haul due to clustering of individuals with similar size. The model estimates growth, mortality and reproduction, after which any aspect of size-structure, spatio-temporal population dynamics, as well as the sampling process can be probed. This is illustrated by two applications: 1) tracking the spatial movements of a single cohort through time, 2) predicting the risk of by-catch of undersize individuals. The method demonstrates that it is possible to combine stock assessment and spatio-temporal dynamics, however at a high computational cost. The model can be extended by increasing its ecological fidelity, although computational feasibility eventually becomes limiting.

Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models
Indices of abundance from fishery-independent trawl surveys constitute an important source of information for many fish stock assessments. Indices are often calculated using area stratified sample means on age-disaggregated data, and finally treated in stock assessment models as independent observations. We evaluate a series of alternative methods for calculating indices of abundance from trawl survey data (delta-lognormal, delta-gamma, and Tweedie using Generalized Additive Models) as well as different error structures for these indices when used as input in an age-based stock assessment model (time-constant vs time-varying variance, and independent versus correlated age groups within years). The methods are applied to data on North Sea herring (Clupea harengus), sprat (Sprattus sprattus), and whiting (Merlangius merlangus), and the full stock assessments are carried out to evaluate the different indices produced. The stratified mean method is found much more imprecise than the alternatives based on GAMs, which are found to be similar. Having time-varying index variances is found to be of minor importance, whereas the independence assumption is not only violated but has significant impact on the assessments.
Factors influencing observed and self-reported functional ability in women with chronic widespread pain: A cross-sectional study

Objective: To evaluate the relationships between key outcome variables, classified according to the International Classification of Functioning, Disability and Health (ICF), and observed and self-reported functional ability in patients with chronic widespread pain.

Design: Cross-sectional with systematic data collection in a clinical setting. Subjects: A total of 257 consecutively enrolled women with chronic widespread pain.

Methods: Multidimensional assessment using self-report and observation-based assessment tools identified to cover ICF categories included in the brief ICF Core Set for chronic widespread pain.

Results: Relationships between ICF variables and observed functional ability measured with the Assessment of Motor and Process Skills (AMPS) were few. Out of 36 relationships analysed, only 4 ICF variables showed a moderate correlation with the AMPS motor ability measure. A moderate to strong correlation between numerous ICF variables and self-reported functioning was noted. Multivariate regression modelling supported significant contributions from pain and psychosocial variables to the variability in self-reported functional ability, but not to the variability in AMPS ability measures.

Conclusion: Observation-based assessment of functional ability in patients with chronic widespread pain is less influenced by pain and psychosocial factors than are self-reported evaluations. Valid observation-based assessment tools, such as the AMPS, should be included in clinical evaluation and future research addressing functional outcomes in this patient population.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Southern Denmark, Copenhagen University Hospital
Contributors: Amris, K., Wæhrens, E. E., Stockmarr, A., Bliddal, H., Danneskiold-Samsøe, B.
Pages: 1014-1021
Publication date: 2014
Peer-reviewed: Yes

Fluorescence in situ hybridization investigation of potentially pathogenic bacteria involved in neonatal porcine diarrhea
Background
Neonatal diarrhea is a multifactorial condition commonly present on pig farms and leads to economic losses due to
increased morbidity and mortality of piglets. Immature immune system and lack of fully established microbiota at birth predispose neonatal piglets to infection with enteric pathogens. The microorganisms that for decades have been associated with enteritis and diarrhea in suckling piglets are: rotavirus A, coronavirus, enterotoxigenic Escherichia coli (ETEC), Clostridium perfringens type C, Cryptosporidium spp., Giardia spp., Cystoisospora suis and Strongyloides ransomi. However, in recent years, the pig industry has experienced an increased number of neonatal diarrhea cases in which the above mentioned pathogens are no longer detected. Potentially pathogenic bacteria have recently received focus in the research on the possible etiology of neonatal diarrhea not caused by common pathogens. The primary aim of this study was to investigate the role of E. coli, Enterococcus spp., C. perfringens and C. difficile in the pathogenesis of neonatal porcine diarrhea with no established casual agents. Fluorescence in situ hybridization with oligonucleotide probes was applied on the fixed intestinal tissue samples from 51 diarrheic and 50 non-diarrheic piglets collected from four Danish farms during outbreaks of neonatal diarrhea not caused by well-known enteric pathogens. Furthermore, an association between the presence of these bacteria and histological lesions was evaluated.

Results
The prevalence of fluorescence signals specific for E. coli, C. perfringens and C. difficile was similar in both groups of piglets. However, Enterococcus spp. was primarily detected in the diarrheic piglets. Furthermore, adherent bacteria were detected in 37 % diarrheic and 14 % non-diarrheic piglets. These bacteria were identified as E. coli and Enterococcus spp. and their presence in the intestinal mucosa was associated with histopathological changes.

Conclusions
The results of this study showed that simultaneous colonization of the intestinal mucosa by adherent non-ETEC E. coli and Enterococcus spp. can be involved in the pathogenesis of neonatal porcine diarrhea. These bacteria should be considered in diagnosis of diarrhea in piglets, when detection of common, well-known enteric agents is unsuccessful.
How do student evaluations of courses and of instructors relate?

Course evaluations are widely used by educational institutions to assess the quality of teaching. At the course evaluations, students are usually asked to rate different aspects of the course and of the teaching. We propose to apply canonical correlation analysis (CCA) in order to investigate the degree of association between how students evaluate the course and how students evaluate the teacher. Additionally it is possible to reveal the structure of this association. Student evaluations data is characterized by high correlations between the variables within each set of variables, therefore two modifications of the CCA method; regularized CCA and sparse CCA, together with classical CCA were applied to find the most interpretable model. Both methods give results with increased interpretability over traditional CCA on the present student evaluation data. The method shows robustness when evaluations over several years are examined.

Hyperspectral imaging based on diffused laser light for prediction of astaxanthin coating concentration

We present a study on predicting the concentration level of synthetic astaxanthin in fish feed pellet coating using multi- and hyperspectral image analysis. This was done in parallel using two different vision systems. A new instrument for hyperspectral imaging, the SuperK setup, using a super-continuum laser as the light source was introduced. Furthermore, a parallel study with the commercially available multispectral VideometerLab imaging system was performed. The SuperK setup used 113 spectral bands (455–1,015 nm), and the VideometerLab used 20 spectral bands (385–1,050 nm). To predict the astaxanthin concentration from the spectral image data, the synthetic astaxanthin content in the pellets was measured with the established standard technique; high-pressure liquid chromatography (HPLC). Regression analysis was done using partial least squares regression (PLSR) and the sparse regression method elastic net (EN). The ratio of standard error of prediction (RPD) is the ratio between the standard deviation of the reference values and the prediction error, and for both PLSR and EN both devices gave RPD values between 4 and 24, and with mean prediction error of 1.4–8.0 parts per million of astaxanthin concentration. The results show that it is possible to predict the synthetic
astaxanthin concentration in the coating well enough for quality control using both multi- and hyperspectral image analysis, while the SuperK setup performs with higher accuracy than the Videometer Lab device for this particular problem. The spectral resolution made it possible to identify the most significant spectral regions for detection of astaxanthin. The results also imply that the presented methods can be used in general for quality inspection of various coating substances using similar coating methods.

**General information**

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, National Food Institute, Division of Industrial Food Research, Division of Toxicology and Risk Assessment, Statistics and Data Analysis
Contributors: Ljungqvist, M. G., Nielsen, O. H. A., Frosch, S., Nielsen, M. E., Clemmensen, L. K. H., Ersbøll, B. K.
Number of pages: 17
Pages: 327-343
Publication date: 2014
Peer-reviewed: Yes

**Publication information**

Journal: Machine Vision & Applications
Volume: 25
Issue number: 2
ISSN (Print): 0932-8092
Ratings:
- BFI (2014): BFI-level 1
- Scopus rating (2014): CiteScore 1.84 SJR 0.625 SNIP 2.177
- Web of Science (2014): Impact factor 1.351
- Web of Science (2014): Indexed yes
Original language: English
Keywords: Hyperspectral imaging, Multispectral imaging, Spectral imaging, Image analysis, Diffused laser, Partial least squares, Elastic net, Astaxanthin, Pellet coating
DOI: 10.1007/s00138-013-0512-2
Source: dtu
Source-ID: u::7689
Research output: Contribution to journal > Journal article – Annual report year: 2013 > Research > peer-review

**Improved overall survival after implementation of targeted therapy for patients with metastatic renal cell carcinoma:**

**Results from the Danish Renal Cancer Group (DARENCA) study-2**

Abstract

Aim To evaluate the implementation of targeted therapy on overall survival (OS) in a complete national cohort of patients with metastatic renal cell carcinoma (mRCC). Methods All Danish patients with mRCC referred for first line treatment with immunotherapy, TKIs or mTOR-inhibitors between 2006 and 2010 were included. Baseline and outcome data were collected retrospectively. Prognostics factors were identified using log-rank tests and Cox proportional hazard model. Differences in distributions were tested with the Chi-square test. Results 1049 patients were referred; 744 patients received first line treatment. From 2006 to 2010 we observed a significant increase in the number of referred patients; a significant increase in treated patients (64% versus 75%, P = 0.0188); a significant increase in first line targeted therapy (22% versus 75%, P < 0.0001); a significant increase in second line treatment (20% versus 40%, P = 0.0104), a significant increased median OS (11.5 versus 17.2 months, P = 0.0435) whereas survival for untreated patients remained unchanged. Multivariate analysis validated known prognostic factors. Moreover, treatment start years 2008 (HR 0.74, 95% CI, 0.55–0.99; P = 0.0415), 2009 (HR 0.72, 95% CI, 0.54–0.96; P = 0.0277) and 2010 (HR 0.63, 95% CI, 0.47–0.86; P = 0.0035) compared to 2006, and more than two treatment lines received for patients with performance status 0–1 (HR 0.76, 95% CI, 0.58–0.99; P = 0.0397) and performance status 2–3 (HR 0.19, 95% CI, 0.06–0.60; P = 0.0051) were significantly associated with longer OS. Conclusion This retrospective study documents that the implementation of targeted therapy has resulted in significantly improved treatment rates and overall survival in a complete national cohort of treated mRCC patients.

**General information**

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University, Odense University Hospital, Aalborg University Hospital, Pfizer Oncology, Copenhagen University Hospital
Pages: 553–562
Publication date: 2014
Peer-reviewed: Yes
Influence of consumer's background on product perception

General information
Publication status: Published
Organisations: Department of Management Engineering, Technology and Innovation Management, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Perez Mata, M., Ahmed-Kristensen, S., Brockhoff, P. B.
Number of pages: 10
Pages: 2125-2134
Publication date: 2014

Host publication information
Title of host publication: 13th International Design Conference - Design 2014
Publisher: Design Society
Keywords: Creativity, Bioinspired design, Analogies, Design cognition
Electronic versions:
Influence_of_consumers_background.pdf
Research output: Chapter in Book/Report/Conference proceeding > Article in proceedings – Annual report year: 2014 > Research > peer-review

In vivo Comet assay – statistical analysis and power calculations of mice testicular cells
The in vivo Comet assay is a sensitive method for evaluating DNA damage. A recurrent concern is how to analyze the data appropriately and efficiently. A popular approach is to summarize the raw data into a summary statistic prior to the statistical analysis. However, consensus on which summary statistic to use has yet to be reached. Another important consideration concerns the assessment of proper sample sizes in the design of Comet assay studies. This study aims to identify a statistic suitably summarizing the % tail DNA of mice testicular samples in Comet assay studies. A second aim is to provide curves for this statistic outlining the number of animals and gels to use. The current study was based on 11 compounds administered via oral gavage in three doses to male mice: CAS no. 110-26-9, CAS no. 512-56-1, CAS no. 111873-33-7, CAS no. 79-94-7, CAS no. 115-96-8, CAS no. 598-55-0, CAS no. 636-97-5, CAS no. 85-28-9, CAS no. 13674-87-8, CAS no. 43100-38-5 and CAS no. 60965-26-6. Testicular cells were examined using the alkaline version of the Comet assay and the DNA damage was quantified as % tail DNA using a fully automatic scoring system. From the raw data 23 summary statistics were examined. A linear mixed-effects model was fitted to the summarized data and the estimated variance components were used to generate power curves as a function of sample size. The statistic that most appropriately summarized the within-sample distributions was the median of the log-transformed data, as it most consistently conformed to the assumptions of the statistical model. Power curves for 1.5-, 2-, and 2.5-fold changes of the highest dose group compared to the control group when 50 and 100 cells were scored per gel are provided to aid in the design of future Comet assay studies on testicular cells.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, National Food Institute, Division of Toxicology and Risk Assessment, Statistics and Data Analysis
Contributors: Hansen, M. K., Sharma, A. K., Dybdahl, M., Boberg, J., Kulahci, M.
IT-pedagogical Think Tank for Teacher-teams in Global Classroom, A Model for Continuous Competence Development with a Focus on Reflection, Innovation, Motivation and Engagement

The IT pedagogical think tank for teacher-teams (ITP4T) is the first version of a “continuous competence development model”. The purpose of this model is to provide teachers and educational organizations with a reflective tool that enables them to create pedagogical innovation in an on-going and structured way. The teachers continuously create new results through this new practice in regard to the learning design and the use of educational technology based on issues, wishes and visions of the teachers as well as the organization. The model is a response to the needs and challenges the teachers and organization at VUC’s Global Classroom are met by when having to be pedagogical innovators, being able to take new educational technology into use and change learning designs accordingly (Collins & Halverson 2010). The reflective tool should also create qualified and motivational learning for the students (Hutters et al. 2013). ITP4T answers the need for teachers to have the possibility to access sustainable competence development in their busy lives, taking outset in their daily problems and with team-support in their teaching environment (Dede et al. 2009). The findings were, that it was possible to establish an agile on-going practice, enabling the teacher-team to reflect, innovate and create new solutions for the constantly occurring It-pedagogical issues on a theoretical and practical level in a way that empowers, engages and motivates the teachers in their daily working life.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Weitze, C. L.
Number of pages: 6
Publication date: 2014

Host publication information
Title of host publication: Designs for learning 2014, 4th International Conference, Expanding the Field, 6-9 May 2014, Stockholm, Sweden: Conference Proceedings
Electronic versions:
CL.Weitze_IT_pedagogical_Think_Tank_for_Teacher_teams_in_Global_Classroom_Designs_for_learning.pdf

Bibliographical note
Source: PublicationPreSubmission
Source-ID: 147733571
Research output: Chapter in Book/Report/Conference proceeding → Article in proceedings – Annual report year: 2014 → Research → peer-review

Maximizing entropy over Markov processes
The channel capacity of a deterministic system with confidential data is an upper bound on the amount of bits of data an attacker can learn from the system. We encode all possible attacks to a system using a probabilistic specification, an
Interval Markov Chain. Then the channel capacity computation reduces to finding a model of a specification with highest entropy.

Entropy maximization for probabilistic process specifications has not been studied before, even though it is well known in Bayesian inference for discrete distributions. We give a characterization of global entropy of a process as a reward function, a polynomial algorithm to verify the existence of a system maximizing entropy among those respecting a specification, a procedure for the maximization of reward functions over Interval Markov Chains and its application to synthesize an implementation maximizing entropy.

We show how to use Interval Markov Chains to model abstractions of deterministic systems with confidential data, and use the above results to compute their channel capacity. These results are a foundation for ongoing work on computing channel capacity for abstractions of programs derived from code.

© 2014 Elsevier Inc. All rights reserved.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, French National Institute for Computer Science and Applied Mathematics, IT University of Copenhagen
Contributors: Biondi, F., Legay, A., Nielsen, B. F., Wasowski, A.
Pages: 384-399
Publication date: 2014
Peer-reviewed: Yes

Method for detection of an abnormal sleep pattern in a person
The present disclosure relates to a method for detection of an abnormal sleep pattern based on a dataset of Electrooculography (EOG) signals obtained from a sleeping subject over a time interval, the method comprising the steps of dividing the time interval into a plurality of subintervals, each subinterval preferably corresponding to a sleep epoch, classifying each subinterval in terms of sleep stages, thereby obtaining a temporal sleep stage pattern, wherein a subject having an uncharacteristic temporal distribution of sleep stages is characterized as having an abnormal sleep pattern.

General information
Publication status: Published
Organisations: Department of Electrical Engineering, Biomedical Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Sørensen, H. B. D., Christensen, J. A. E., Jennum, P. J., Rahn Christensen, S., Arvastson, L. J.
Publication date: 2014

Multiphase Image Segmentation Using the Deformable Simplicial Complex Method
The deformable simplicial complex method is a generic method for tracking deformable interfaces. It provides explicit interface representation, topological adaptivity, and multiphase support. As such, the deformable simplicial complex method can readily be used for representing active contours in image segmentation based on deformable models. We show the benefits of using the deformable simplicial complex method for image segmentation by segmenting an image into a known number of segments characterized by distinct mean pixel intensities.

General information
Publication status: Published
Multivariate phase type distributions - Applications and parameter estimation

The best known univariate probability distribution is the normal distribution. It is used throughout the literature in a broad field of applications. In cases where it is not sensible to use the normal distribution alternative distributions are at hand and well understood, many of these belonging to the class of phase type distributions. Phase type distributions have several advantages. They are versatile in the sense that they can be used to approximate any given probability distribution on the positive reals. There exist general probabilistic results for the entire class of phase type distributions, allowing for different estimation methods for the whole class or subclasses of phase type distributions. These attributes make this class of distributions an interesting alternative to the normal distribution.

When facing multivariate problems, the only general distribution that allows for estimation and statistical inference, is the multivariate normal distribution. Unfortunately only little is known about the general class of multivariate phase type distribution. Considering the results concerning parameter estimation and inference theory of univariate phase type distributions, the class of multivariate phase type distributions shows potential for similar great results.

My PhD studies were part of the the work package 3 of the UNITE project. The overall goal of the UNITE project is to improve the decision support prior to deciding on a project by reducing systematic model bias and by quantifying and reducing model uncertainties.

Research has shown that the errors on cost estimates for infrastructure projects clearly do not follow a normal distribution but is skewed towards cost overruns. This skewness can be described using phase type distributions. Cost benefit analysis assesses potential future projects and depend on reliable cost estimates. The Successive Principle is a group analysis method primarily used for analyzing medium to large projects in relation to cost or duration. We believe that the mathematical modeling used in the Successive Principle can be improved. We suggested a novel approach for modeling the total duration of a project using a univariate phase type distribution. The model is then extended to catch the correlation between duration and cost estimates using a bivariate phase type distribution. The use of our model can improve estimates for duration and costs and therefore help project management to make the optimal decisions.

The work conducted during my PhD studies aimed at shedding light on the class of multivariate phase type distributions. This thesis contains analytical and numerical results for parameter estimations and inference theory for a family of multivariate phase type distributions. The results can be used as a stepping stone towards understanding multivariate phase type distributions better. However, we are far from uncovering the full potential of general multivariate phase type distributions. Deeper understanding of multivariate phase type distributions will open up a broad field of research areas they can be applied to.

This thesis consists of a summary report and two research papers. The work was carried out in the period 2010 - 2014.
Novel X-ray phase-contrast tomography method for quantitative studies of heat induced structural changes in meat

The objective of this study was to evaluate the use of X-ray phase-contrast tomography combined with 3D image segmentation to investigate the heat induced structural changes in meat. The measurements were performed at the Swiss synchrotron radiation light source using a grating interferometric setup. The non-destructive method allowed the same sample to be measured before and after cooking. Heat denaturation resulted in a 36% decrease in the volume of the muscle fibers, while solubilization of the connective tissues increased the volume from 8.4% to 24.9%. The cooking loss was quantified and separated into a water phase and a gel phase formed by the sarcoplasmic proteins in the exudate. The results show that X-ray phase contrast tomography offers unique possibilities in studies both the meat structure and the different meat component such as water, fat, connective tissue and myofibrils in a qualitative and quantitative manner without prior sample preparation as isolation of single muscle components, calibration or histology.
On the use of functional calculus for phase-type and related distributions

The area of phase-type distributions is renowned for its ability to obtain closed form formulas or algorithmically exact solutions to many complex stochastic models. The method of functional calculus will provide an additional tool along these lines for establishing results in terms of functions of matrices. Functional calculus, which is a branch of operator theory frequently associated with complex analysis, can be applied to phase-type and matrix-exponential distributions in a rather straightforward way. In this paper we provide a number of examples on how to execute the formal arguments.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Bladt, M., Campillo Navarro, A., Nielsen, B. F.
Number of pages: 25
Publication date: 2014

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
(DTU Compute-Technical Report-2014; No. 18).
Electronic versions:
tr14_18_Nielsen_BF.pdf

On Weighted Support Vector Regression

We propose a new type of weighted support vector regression (SVR), motivated by modeling local dependencies in time and space in prediction of house prices. The classic weights of the weighted SVR are added to the slack variables in the objective function (OF-weights). This procedure directly shrinks the coefficient of each observation in the estimated functions; thus, it is widely used for minimizing influence of outliers. We propose to additionally add weights to the slack variables in the constraints (CF-weights) and call the combination of weights the doubly weighted SVR. We illustrate the differences and similarities of the two types of weights by demonstrating the connection between the Least Absolute Shrinkage and Selection Operator (LASSO) and the SVR. We show that an SVR problem can be transformed to a LASSO problem plus a linear constraint and a box constraint. We demonstrate the capabilities of the doubly weighted approach through an example of prediction of house prices. The weight functions in the house pricing model depend on the geographical distance to the house of interest and the difference in time of sale (CF-weights) as well as the differences lying in variables (OF-weights), such as house size and number of floors. The results illustrate that the combination of the two types of weights describes the relative importance of observations very well and lowers the influence of possible outliers. Therefore, it enables the SVR models to have good performance. Copyright © 2014 John Wiley & Sons, Ltd.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Technical University of Denmark
Contributors: Han, X., Clemmensen, L. K. H.
Pages: 891-903
Publication date: 2014
Peer-reviewed: Yes

Publication information
Journal: Quality and Reliability Engineering International
Volume: 30
Issue number: 6
ISSN (Print): 0748-8017
Ratings:
BFI (2014): BFI-level 1
Optimering af billetkontrol.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Thyregod, C., Ersbøll, B. K.
Number of pages: 38
Publication date: 2014

Publication information
Publisher: DTU Compute
Original language: English

Bibliographical note
Confidential report, not accessible to the public.

Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

Paired preference data with a no-preference option – Statistical tests for comparison with placebo data
It is well-established that when respondents are presented with identical samples in a preference test with a no preference option, a sizable proportion of respondents will report a preference. In a recent paper (Ennis, D. M., & Ennis, J. M. (2012a). Accounting for no difference/preference responses or ties in choice experiments. Food Quality and Preference, 23, 13–17) noted that this proportion can depend on the product category, have proposed that the expected proportion of preference responses within a given category be called an identically norm, and have argued that knowledge of such norms is valuable for more complete interpretation of 2-Alternative Choice (2-AC) data. For instance, these norms can be used to indicate consumer segmentation even with non-replicated data. In this paper, we show that the statistical test suggested by Ennis and Ennis (2012a) behaves poorly and has too high a type I error rate if the identically norm is not estimated from a very large sample size. We then compare five $\chi^2$ tests of paired preference data with a no preference option in terms of type I error and power in a series of scenarios. In particular, we identify two tests that are well behaved for sample sizes typical of recent research and have high statistical power. One of these tests has the advantage that it can be decomposed for more insightful analyses in a fashion similar to that of ANOVA F-tests. The benefits are important because they enable more informed business decisions, particularly when ingredient changes are considered for cost-reduction or health initiative purposes.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, The Institute for Perception
Contributors: Christensen, R. H. B., Ennis, J. M., Ennis, D. M., Brockhoff, P. B.
Pages: 48-55
Publication date: 2014
Peer-reviewed: Yes

Publication information
Journal: Food Quality and Preference
Volume: 32
ISSN (Print): 0950-3293
Ratings:
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 3.18 SJR 0.997 SNIP 1.676
Web of Science (2014): Impact factor 2.779
Web of Science (2014): Indexed yes
Quantitative assessment of the risk of introduction of bovine viral diarrhea virus in Danish dairy herds

A quantitative risk assessment was carried out to estimate the likelihood of introducing bovine viral diarrhea virus (BVDV) in Danish dairy herds per year and per trimester, respectively. The present study gives important information on the impact of risk mitigation measures and sources of uncertainty due to lack of data. As suggested in the Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement), the OIE Terrestrial Animal Health Code was followed for a transparent science-based risk assessment. Data from 2010 on imports of live cattle, semen, and embryos, exports of live cattle, as well as use of vaccines were analyzed. Information regarding the application of biosecurity measures, by veterinarians and hoof trimmers practicing in Denmark and in other countries, was obtained by contacting several stakeholders, public institutions, and experts. Stochastic scenario trees were made to evaluate the importance of the various BVDV introduction routes. With the current surveillance system, the risk of BVDV introduction was estimated to one or more introductions within a median of nine years (3–59). However, if all imported animals were tested and hoof trimmers always disinfected the tools used abroad, the risk could be reduced to one or more introductions within 33 years (8–200). Results of this study can be used to improve measures of BVD surveillance and prophylaxis in Danish dairy herds.

Real-life use of vitamin D₃-fortified bread and milk during a winter season: the effects of CYP2R1 and GC genes on 25-hydroxyvitamin D concentrations in Danish families, the VitmaD study.

Common genetic variants rs10741657 and rs10766197 in CYP2R1 and rs4588 and rs842999 in GC and a combined genetic risk score (GRS) of these four variants influence late summer 25-hydroxyvitamin D (25(OH)D) concentrations. The objectives were to identify those who are most at risk of developing low vitamin D status during winter and to assess whether vitamin D₃-fortified bread and milk will increase 25(OH)D concentrations in those with genetically determined low 25(OH)D concentrations at late summer. We used data from the VitmaD study. Participants were allocated to either vitamin D₃-fortified bread and milk or non-fortified bread and milk during winter. In the fortification group, CYP2R1 (rs10741657) and GC (rs4588 and rs842999) were statistically significantly associated with winter 25(OH)D concentrations and CYP2R1 (rs10766197) was borderline significant. There was a negative linear trend between 25(OH)D concentrations and carriage of 0–8 risk alleles (p < 0.0001). No association was found for the control group (p = 0.1428). There was a significant positive linear relationship between different quintiles of total vitamin D intake and the increase in 25(OH)D concentrations among carriers of 0–2 (p = 0.0012), 3 (p = 0.0001), 4 (p = 0.0118) or 5 (p = 0.0029) risk alleles, but not among carriers of 6–8 risk alleles (p = 0.1051). Carriers of a high GRS were more prone to be vitamin D deficient during winter.
compared to carriers of a low GRS. Furthermore, rs4588-AA carriers have a low but very stable 25(OH)D concentration, and interestingly, also low PTH level.

**General information**
Publication status: Published
Organisations: National Food Institute, Division of Nutrition, Division of Toxicology and Risk Assessment, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Aarhus University
Number of pages: 15
Publication date: 2014
Peer-reviewed: Yes

**Publication information**
Journal: Genes & Nutrition
Volume: 9
Issue number: 4
Article number: 413
ISSN (Print): 1555-8932
Ratings:
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 3.4 SJR 1.219 SNIP 0.931
Web of Science (2014): Impact factor 2.794
Web of Science (2014): Indexed yes
Original language: English
Electronic versions:
Genes_Nutrition_1.pdf
DOIs:
10.1007/s12263-014-0413-7
Source: FindIt
Source-ID: 268448325
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

Segmentation of Connective Tissue in Meat from Microtomography Using a Grating Interferometer
It has been demonstrated that phase contrast imaging provides superior contrast of soft tissues in biological material over typical absorption tomography [1-2]. In meat science, this imaging modality can provide valuable information of the effects of heat treatment on muscle tissue. Although microtomography provides high resolution, the thin structures of the connective tissues are difficult to segment. This is mainly due to partial object voxels, image noise and artifacts. The segmentation of connective tissue is important for quantitative analysis purposes. Factors such as the surface area, relative volume and the statistics of the electron density of the connective tissue could prove useful for understanding the structural changes occurring in the meat sample due to heat treatment.

In this study a two step segmentation algorithm was implemented in order to segment connective tissue from phase contrast microtomograms obtained by a grating interferometer. This segmentation has previously been demonstrated for the segmentation of the optic nerve head from microscopic images of stained slices [3]. The first step is to model the data as a mixture of Gaussians using an expectation-maximization (EM) algorithm [4]. This iterative process finds the maximum likelihood of parameters where the model depends on unobserved latent variables. The spatial information of the data is next incorporated into the segmentation process by modeling the data as a Markov random field (MRF) [5]. It models the a priori probability of neighborhood dependencies, and the field can either be isotropic or anisotropic. For the segmentation of connective tissue, the local information of the structure orientation and coherence is extracted to steer the smoothing (anisotropy) of the final segmentation.

The results show that the segmentation provides a superior classification of connective tissue over conventional threshold segmentation. Additionally modeling the data as a mixture of Gaussians made it possible to segment the connective tissue into two separate classes. The segmentation results provide the means for further analysis of the structural changes in the meat due to heat treatment.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Einarsdottir, H., Ersbøll, B. K., Larsen, R.
Number of pages: 1
Publication date: 2014
Segmentation Toolbox for Tomographic Image Data
Motivation: Image acquisition has vastly improved over the past years, introducing techniques such as X-ray computed tomography (CT). CT images provide the means to probe a sample non-invasively to investigate its inner structure. Given the wide usage of this technique and massive data amounts, techniques to automatically analyze such data becomes ever more important. Most segmentation methods for large datasets, such as CT images, deal with simple thresholding techniques, where intensity values cut offs are predetermined and hard coded. For data where the intensity difference is not sufficient, and partial volume voxels occur frequently, thresholding methods do not suffice and more advanced methods are required.

Contribution: To meet these requirements a toolbox has been developed, combining well known methods within the image analysis field. The toolbox includes cluster-based methods to automatically determine parameters of the different classes present in the data, and edge weighted smoothing of the final segmentation based on Markov Random Fields (MRF). The toolbox is developed for Matlab users and requires only minimal background knowledge of Matlab.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Einarsdottir, H.
Number of pages: 1
Publication date: 2014
Peer-reviewed: No
Electronic versions:
Pages_from_FAIMIII_1_poster.pdf
URLs:

Bibliographical note
Published in 'Farm Animal Imaging' a report of the 3rd annual conference of the COST Action FA1102 (FAIM III)
Source: PublicationPreSubmission
Source-ID: 120651287
Research output: Contribution to conference › Poster – Annual report year: 2014 › Research

Selecting groups of covariates in the elastic net
This paper introduces a novel method to select groups of variables in sparse regression and classification settings. The groups are formed based on the correlations between covariates and ensure that for example spatial or spectral relations are preserved without explicitly coding for these. The preservation of relations gives increased interpretability. The method is based on the elastic net and adaptively selects highly correlated groups of variables and does therefore not waste time in grouping irrelevant variables for the problem at hand. The method is illustrated on a simulated data set and on regression of moisture content in multispectral images of sand. In both cases, the predictions were better or similar to existing regression and classification algorithms and the interpretation was enhanced using the grouping method. On top of that, the grouping method more consistently selects the important variables.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Clemmensen, L. K. H.
Number of pages: 14
Publication date: 2014
**Short mandible - a possible risk factor for cleft palate with/without a cleft lip**

Structured Abstract

Objectives To estimate the influence of a short mandible on the risk of developing a cleft palate with/without a cleft lip (CP). Setting and sample population The retrospective sample consisted of 115 2-month-old Danish infants with CP, and 70 control infants with unilateral incomplete cleft lip (UICL). Material and Methods Cephalometric X-rays were obtained. Mandibular length (L-m) was measured and corrected for body length (L-b) to remove influence of varying body length in the sample. Logistic regression was applied to the corrected mandibular length (L-mc) to calculate the risk of having a cleft palate. Results The mean mandibular length in the group with CP was about 4mm shorter than in the control group. Odds ratio (OR) was calculated to be 0.58 (95% confidence interval 0.48-0.68), implying that an individual's risk of cleft palate with/without a cleft lip increases about 50% per mm decrease in mandibular length. Conclusions A special facial type including a short mandible is a possible risk factor for cleft palate, and it was found that the risk of cleft palate increases 58% per mm decreases in mandibular length.

**Sparse and shrunken estimates of MRI networks in the brain and their influence on network properties**

Estimation of morphometric relationships between cortical regions is a widely used approach to identify and characterize structural connectivity. The elevated number of regions that can be considered in a whole-brain correlation analysis might lead to overfitted models. However, the overfitting can be avoided by using regularization methods. We found that, as expected, non-regularized correlations had low variability when a scarce number of variables were considered. However, a slight increase of variables led to an increase of variance of several magnitude orders. On the other hand, the regularized approaches showed more stable results with a relative low variance at the expense of a little bias. Interestingly, topological properties as local and global efficiency estimated in networks constructed from traditional non-regularized correlations also showed higher variability when compared to those from regularized networks. Our findings suggest that a population-based connectivity study can achieve a more robust description of cortical topology through regularization of the correlation estimates. Four regularization methods were examined: Two with shrinkage (Ridge and Schäfer's shrinkage), one with sparsity (Lasso) and one with both shrinkage and sparsity (Elastic net). Furthermore, the different regularizations resulted in different correlation estimates as well as network properties. The shrunken estimates resulted in lower variance of the estimates than the sparse estimates.
Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass

Mixture models are introduced as a new and stronger methodology for statistical prediction of biomethane potentials (BMP) from lignocellulosic biomass compared to the linear regression models previously used. A large dataset from literature combined with our own data were analysed using canonical linear and quadratic mixture models. The full model to predict BMP ($R^2>0.96$), including the four biomass components cellulose ($x_C$), hemicellulose ($x_H$), lignin ($x_L$) and residuals ($x_R=1-x_C-x_H-x_L$) had highly significant regression coefficients. It was possible to reduce the model without substantially affecting the quality of the prediction, as the regression coefficients for $x_C$, $x_H$ and $x_R$ were not significantly different based on the dataset. The model was extended with an effect of different methods of analysing the biomass constituents content (DA) which had a significant impact. In conclusion, the best prediction of BMP is $p_{BMP}=347x_C+x_H+x_R-438x_L+63DA$. 

Status for NEXIM New X-ray Imaging Modalities for safe and high quality food

The main objectives of the NEXIM project are to develop the novel X-ray grating interferometry technique (Weitkamp et al. 2005; Pfeiffer et al. 2008) specifically towards food application and to identify the areas within the Danish food industry with the highest technological and commercial impact. The main focuses are determined to be threefold:
1) Improving the detectability of low density foreign bodies incidentally present in food products.
2) Development of new modalities for assessment of quality traits in food production, for instance connective tissue and fatty acid composition.
3) Develop a proof-of-principle of a conveyor belt solution that can form the basis for real product development.

In the past year the NEXIM project has focused on these three objectives, studying the applicability of GBI to meat quality assessment and foreign object detection. Some efforts have been put to developing laboratory-based setups further towards an in-line scanning system. Additionally, close co-operation with industrial partners has further emphasized the need for new techniques for quality control, product development and foreign object detection.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Contributors: Einarsdottir, H., Ersbøll, B. K., Larsen, R.
Pages: 121-124
Publication date: 2014

**Host publication information**
Title of host publication: Farm Animal Imaging
Editors: Maltin, C., Craigie, C., Bünger, L.
ISBN (Print): 978-0-9931063-0-9
Electronic versions:
Pages_from_FAIMIII_1.pdf
URLs:
Source: PublicationPreSubmission
Source-ID: 120651373

**Stochastic simulation modeling to determine time to detect Bovine Viral Diarrhea antibodies in bulk tank milk**

A stochastic simulation model was developed to estimate the time from introduction of Bovine Viral Diarrhea Virus (BVDV) in a herd to detection of antibodies in bulk tank milk (BTM) samples using three ELISAs. We assumed that antibodies could be detected, after a fixed threshold prevalence of seroconverted milking cows was reached in the herd. Different thresholds were set for each ELISA, according to previous studies. For each test, antibody detection was simulated in small (70 cows), medium (150 cows) and large (320 cows) herds. The assays included were: (1) the Danish blocking ELISA, (2) the SVANOVIR®BVDV-Ab ELISA, and (3) the ELISA BVD/MD p80 Institute Pourquier. The validation of the model was mainly carried out by comparing the predicted incidence of persistently infected (PI) calves and the predicted detection time, with records from a BVD infected herd. Results showed that the SVANOVIR, which was the most efficient ELISA, could detect antibodies in the BTM of a large herd 280 days (95% prediction interval: 218; 568) after a transiently infected (TI) milking cow has been introduced into the herd. The estimated time to detection after introduction of one PI calf was 111 days (44; 605). With SVANOVIR ELISA the incidence of PIs and dead born calves could be limited and the impact of the disease on the animal welfare and income of farmers (before detection) could be minimized. The results from the simulation modeling can be used to improve the current Danish BVD surveillance program in detecting early infected herds.

**General information**
Publication status: Published
Organisations: National Veterinary Institute, Section for Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Seges Knowledge Centre for Agriculture
Contributors: Foddai, A., Eneee, C., Krogh, K., Stockmarr, A., Hisham Beshara Halasa, T.
Pages: 149-159
Publication date: 2014
Peer-reviewed: Yes

**Publication information**
Journal: Preventive Veterinary Medicine
Volume: 117
Issue number: 1
ISSN (Print): 0167-5877
Ratings:
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 2.37 SJR 1.27 SNIP 1.415
Web of Science (2014): Impact factor 2.167
Superior performance of constant-saltier-reference DTF and DTFM to same-different tests by consumers for discriminating products varying sodium contents

Reducing sodium content in foods and beverages has become very important, and great efforts are being made to achieve this while maintaining overall taste/acceptance of food. This requires more robust sensory discrimination test methods in terms of operational power because discrimination tests using food/beverages with high sodium contents might be more prone to adaptation and memory bias which might reduce the operational power of the test methods. The operational test power of six versions of the duo-trio test method (two different versions of replicated tests under three reference modes – constantly stronger-reference (SR) vs. balanced-reference vs. constantly weaker-reference (WR)) were investigated using a pair of tomato juices with different sodium content in comparison to the same-different tests using a balanced-design. The two duo-trio versions were the traditional duo-trio with a reference presented first (DTF) and the duo-trio with the reference presented both at the first and in the middle between the two alternative test samples (DTFM). An examination of the d’ estimates indicated that discrimination significantly improved across all duo-trio tests and the same-different tests when using naïve consumer subjects and when the sample in the first position in a test was a stronger (saltier) sample. This observation granted operational superiority to the DTF and DTFM in comparison to the same-different test, using the stronger (saltier) product as a constant-reference for discriminating products with high sodium content.
Supervised feature selection for linear and non-linear regression of Lab color from multispectral images of meat

In food quality monitoring, color is an important indicator factor of quality. The CIELab (Lab) color space as a device independent color space is an appropriate means in this case. The commonly used colorimeter instruments can neither measure the Lab color in a wide area over the target surface nor in a contact-less mode. However, developing algorithms for conversion of food items images into Lab color space can solve both of these issues. This paper addresses the problem of Lab color prediction from multispectral images of different types of raw meat. The efficiency of using multispectral images instead of the standard RGB is investigated. In addition, it is demonstrated that due to the fiber structure and transparency of raw meat, the prediction models built on the standard color patches do not work for raw meat test samples. As a result, multispectral images of different types of meat samples (430–970 nm) were used for training and testing of the Lab prediction models. Finding a sparse solution or the use of a minimum number of bands is of particular interest to make an industrial vision set-up simpler and cost effective. In this paper, a wide range of linear, non-linear, kernel-based regression and sparse regression methods are compared. In order to improve the prediction results of these models, we propose a supervised feature selection strategy which is compared with the Principal component analysis (PCA) as a pre-processing step. The results showed that the proposed feature selection method outperforms the PCA for both linear and non-linear methods. The highest performance was obtained by linear ridge regression applied on the selected features from the proposed Elastic net (EN) -based feature selection strategy. All the best models use a reduced number of wavelengths for each of the Lab components.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Danish Meat Research Institute
Contributors: Sharifzadeh, S., Clemmensen, L. K. H., Borggaard, C., Støier, S., Ersbøll, B. K.
Number of pages: 17
Pages: 211-227
Publication date: 2014
Peer-reviewed: Yes

Publication information
Journal: Engineering Applications of Artificial Intelligence
Volume: 27
ISSN (Print): 0952-1976
Ratings:
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 3.56 SJR 1.437 SNIP 2.741
Web of Science (2014): Impact factor 2.207
Web of Science (2014): Indexed yes
Original language: English
Keywords: L a b color space, Multispectral imaging, Sparse regression, Artificial neural networks, Support vector machine, Supervised feature selection
DOIs:
10.1016/j.engappai.2013.09.004
Source: dtu
Source-ID: u::9232
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

Surface Detection using Round Cut
We propose an iterative method for detecting closed surfaces in a volumetric data, where an optimal search is performed in a graph build upon a triangular mesh. Our approach is based on previous techniques for detecting an optimal terrain-like or tubular surface employing a regular grid. Unlike similar adaptations for triangle meshes, our method is capable of capturing complex geometries by iteratively refining the surface, where we obtain a high level of robustness by applying explicit mesh processing to intermediate results. Our method uses on-surface data support, but it also exploits data information about the region inside and outside the surface. This provides additional robustness to the algorithm. We demonstrate the capabilities of the approach by detecting surfaces of CT scanned objects.

General information
Publication status: Published
The importance of spatial models for estimating the strength of density dependence

Identifying the existence and magnitude of density dependence is one of the oldest concerns in ecology. Ecologists have aimed to estimate density dependence in population and community data by fitting a simple autoregressive (Gompertz) model for density dependence to time series of abundance for an entire population. However, it is increasingly recognized that spatial heterogeneity in population densities has implications for population and community dynamics. We therefore adapt the Gompertz model to approximate local densities over continuous space instead of population-wide abundance, and to allow productivity to vary spatially. Using simulated data generated from a spatial model, we show that the conventional (nonspatial) Gompertz model will result in biased estimates of density dependence, e.g., identifying oscillatory dynamics when not present. By contrast, the spatial Gompertz model provides accurate and precise estimates of density dependence for a variety of simulation scenarios and data availabilities. These results are corroborated when comparing spatial and nonspatial models for data from 10 years and ~100 sampling stations for three long-lived rockfishes (Sebastes spp.) off the California Coast. In this case, the nonspatial model estimates implausible oscillatory dynamics on an annual time scale, while the spatial model estimates strong autocorrelation and is supported by model selection tools. We conclude by discussing the importance of improved data archiving techniques, so that spatial models can be used to re-examine classic questions regarding the presence and strength of density dependence in wild populations.

Read More: http://www.esajournals.org/doi/abs/10.1890/14-0739.1
The MAM-CAP Table: A New Tool for Monitoring Panel Performances

Assessor performances in sensory analysis are usually represented by three indicators: repeatability, discrimination and agreement. However, assessors can also differ on the range of their scores, the so-called “scaling effect”. Brockhoff, Schlich, and Skovgaard (2013) proposed the mixed assessor model (MAM) which, as the original assessor model (Brockhoff & Skovgaard, 1994), takes this effect into account, but also allows for the product effect to be tested against a new interaction free of the scaling effect. The present paper proposes a unified system for monitoring assessor and panel performances based on the MAM. In addition to the product effect (tested at panel and individual levels), scaling and repeatability panel heterogeneities are tested and, when significant, assessors who were responsible for this heterogeneity are identified. Further, the pure disagreement is decomposed over assessors allowing for the individuals significantly responsible for it to be highlighted. Those numerous statistics are summarized into a few diagnostics presented in a synthetic single table called the MAM CAP table.

General Information

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Centre des Sciences du Goût et de l’Alimentation
Contributors: Peltier, C., Brockhoff, P. B., Visalli, M., Schlich, P.
Pages: 24-27
Publication date: 2014
Peer-reviewed: Yes

Publication Information

Journal: Food Quality and Preference
Volume: 32
Issue number: Part A
ISSN (Print): 0950-3293
Ratings:
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 3.18 SJR 0.997 SNIP 1.676
Web of Science (2014): Impact factor 2.779
Web of Science (2014): Indexed yes
Original language: English
Keywords: Panel performances, Mixed assessor model, Scaling
DOIs:
10.1016/j.foodqual.2013.07.004
Source: dtu
Source-ID: n::oai:DTIC-ART:elsevier/409662303::34550
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

The Relationship between Mechanical Hyperalgesia Assessed by Manual Tender Point Examination and Disease Severity in Patients with Chronic Widespread Pain: A Cross-Sectional Study

The clinical utility of tender point (TP) examination in patients reporting chronic widespread pain (CWP) is the subject of contemporary debate. The objective of this study was to assess the relationship between mechanical hyperalgesia assessed by manual TP examination and clinical disease severity. 271 women with CWP were recruited from a clinical setting. Data collection included patient-reported symptoms, health-related quality of life variables, and observation-based measures of functional ability, muscle strength, 6-minute walk, and pressure pain thresholds measured by cuff algometry. TP examination was conducted according to ACR-guidelines. Relationships between disease variables and TP count (TPC) were analyzed with logistic regression in a continuum model, allowing the TPC to depend on the included disease variables and two regression models carried out for a TPC threshold level, varying between 1 and 17. The threshold analyses indicated a TPC threshold at 8, above which a large number of disease variables became consistently significant explanatory factors, whereas none of the disease variables reached a significance level in the continuum model. These results support the premise that the presence of mechanical hyperalgesia influences symptomatology in CWP and that the severity of clinical expression is related to a threshold of TPs, rather than being part of a continuum.

General Information

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Oregon Health and Science University, Copenhagen University Hospital
Contributors: Amris, K., Wæhrens, E. E., Jespersen, A., Stockmarr, A., Bennett, R., Bliddal, H., Danneskiold-Samsoe, B.
Number of pages: 13
Publication date: 2014
Peer-reviewed: Yes
The type I error rate for in vivo Comet assay data when the hierarchical structure is disregarded

The Comet assay is a sensitive technique for detection of DNA strand breaks. The experimental design of in vivo Comet assay studies are often hierarchically structured, which should be reflected in the statistical analysis. However, the hierarchical structure sometimes seems to be disregarded, and this imposes considerable impact on the type I error rate. This study aims to demonstrate the implications that result from disregarding the hierarchical structure. Different combinations of the factor levels as they appear in a literature study give type I error rates up to 0.51 and for all combinations the type I error rate is greater than the nominal \( \alpha \) at 0.05. Closed-form expressions based on scaled F-distributions using the Welch-Satterthwaite approximation are provided to show how the type I error rate is affected. With this study we hope to motivate researchers to be more precise regarding the exposition of the statistical methodology and to suitably account for the hierarchical structure of Comet assay data whenever present.

Trends in overweight and obesity in Danish children and adolescents: 2000-2008 – exploring changes according to parental education

Aims: To examine the hypotheses that an overall levelling off in the prevalence of overweight and obesity during the period 2000-2008 has occurred, and that increasing social inequality in overweight and obesity exists in a nationally representative sample of Danish children and adolescents.

Methods: The population comprised a random sample of 1849 children aged 4-14 years who participated in the Danish National Survey of Diet and Physical Activity in 2000-2002, 2003-2004 and 2005-2008. Parental education was chosen as an indicator of children's socioeconomic status. Body mass index (BMI) was calculated from parent-reported weight and height. Subjects were classified as overweight and obese according to the International Obesity Task Force age- and gender-specific BMI cut-off values. Crude prevalence estimates and logistic regression models were used to analyse trends in overweight and obesity as the main outcome measures.

Results: An increase was found in the crude prevalence of overweight (including obesity) in boys (12.8-21.7%, \( p = 0.0006 \), but not in girls (17.6-15.9%, \( p = 0.56 \)), between 2000-2002 and 2005-2008. The prevalence of overweight increased significantly in boys of parents with low educational level only. A strong inverse social gradient in overweight and obesity was documented for boys and girls during the whole survey period.

Conclusions: The present study showed an increase in the prevalence of overweight in Danish boys, but not in girls. This increase was due to increasing social inequality in overweight among boys. Public health initiatives aimed at preventing and reducing overweight and obesity should consider gender difference and especially target boys with parents of low
educational level.

**General information**

Publication status: Published  
Organisations: National Food Institute, Division of Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis  
Pages: 385-392  
Publication date: 2014  
Peer-reviewed: Yes

**Publication information**

Volume: 42  
Issue number: 4  
ISSN (Print): 1403-4948  
Ratings:  
BFI (2014): BFI-level 1  
Scopus rating (2014): CiteScore 3.47 SJR 2.275 SNIP 1.675  
Web of Science (2014): Impact factor 1.832  
Web of Science (2014): Indexed yes  
Original language: English  
Keywords: Denmark, Development, Obesity, Scandinavia, Social disparity, Socioeconomic status  
DOIs: 10.1177/1403494813520356

Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

**Trends in pedometer-measured steps per day in Danish adults: 2007 to 2012**

**General information**

Publication status: Published  
Organisations: National Food Institute, Division of Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Linnaeus University  
Contributors: Matthiessen, J., Andersen, E. W., Raustorp, A., Sørensen, M. R.  
Number of pages: 1  
Publication date: 2014  
Peer-reviewed: Yes  

Electronic versions:  
Source: PublicationPreSubmission  
Source-ID: 97170551

Research output: Contribution to conference › Poster – Annual report year: 2014 › Research › peer-review

**Understanding traffic crash under-reporting: linking police and medical records to individual and crash characteristics**

**General information**

Publication status: Published  
Organisations: Department of Transport, Traffic modelling and planning, National Food Institute, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Odense University Hospital  
Contributors: Janstrup, K. H., Hels, T., Kaplan, S., Sommer, H. M., Lauritsen, J.  
Number of pages: 10  
Publication date: 2014  
Peer-reviewed: Yes  

Electronic versions:  
TRA_2014_Janstrup_Hels_Kaplan_Sommer_Lauritsen.pdf

Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2014 › Research › peer-review
Vitamin D status and its determinants in children and adults among families in late summer in Denmark.

The impact of the familial relationship on vitamin D status has not been investigated previously. The objective of the present cross-sectional study was to assess serum 25-hydroxyvitamin D (25(OH)D) concentration and its determinants in children and adults among families in late summer in Denmark (56°N). Data obtained from 755 apparently healthy children (4-17 years) and adults (18-60 years) recruited as families (n 200) in the VitmaD study were analysed. Blood samples were collected in September-October, and serum 25(OH)D concentration was measured by liquid chromatography-tandem MS. Information on potential determinants was obtained using questionnaires. The geometric mean serum 25(OH)D concentration was 72·1 (interquartile range 61·5-86·7) nmol/l (range 9-162 nmol/l), with 9 % of the subjects having 25(OH)D concentrations <50 nmol/l. The intra-family correlation was 0·27 in all subjects, 0·24 in the adults and 0·42 in the children. Serum 25(OH)D concentration was negatively associated with BMI (P<0·001) and positively associated with dietary vitamin D intake (P= 0·008), multivitamin use (P= 0·019), solarium use (P= 0·006), outdoor stay (P= 0·001), sun preference (P= 0·002) and sun vacation (P<0·001), but was not associated with lifestyle-related factors in the adults when these were assessed together with the other determinants. In conclusion, the majority of children and adults among the families had serum 25(OH)D concentrations >50 nmol/l in late summer in Denmark. Both dietary and sun-related factors were determinants of vitamin D status and the familial component was stronger for the children than for the adults.
Analysis of factors important for the occurrence of Campylobacter in Danish broiler flocks

For more than a decade human campylobacteriosis has been the leading zoonosis in many developed countries. Consumption of poultry or poultry products has been identified as the primary source of infection in humans. This study was conducted to identify risk factors for the occurrence of Campylobacter in Danish broiler flocks. The study was based on a large data set consisting of Campylobacter positive status for nearly 6000 broiler flocks and 43 explanatory variables. Data were obtained from the Danish Campylobacter surveillance programme in poultry and from the responses to a standardized questionnaire answered via interviews with broiler farm owners. Two hundred and forty broiler farms, comprising 539 broiler houses, were included in the study and their Campylobacter status was followed over a 2-year period (1999–2000). The large number of observations made it possible to carry out a multivariate analysis including all 43 variables. A multivariate analysis was conducted using a generalized linear model, and the correlations between the houses from the same farms were accounted for by adding a variance structure to the model. The procedures for analyses included backward elimination, forward selection and expanding of the number of observations used in the variance analysis along with the reduction of the number of parameters in the model. The unit of analysis was ‘broiler house’, meaning that all results from a broiler house were aggregated into one prevalence figure (number of positive flocks/total number of flocks delivered over the 2-year period). The following factors were found to be significantly associated with the occurrence of Campylobacter in the broiler flocks: old broiler houses, late introduction of whole wheat in the feed, relatively high broiler age at slaughter, improper rodent control, large number of chimneys on the broiler house, farm located in an area with a high density of cattle farms, having more than one broiler house on the farm, and improper storage of wheat. This large-scale study confirms several risk factors identified in previous studies. The results concerning chimneys may be explained by the easier access that flies have to the broiler houses, which seems in agreement with recent Danish studies on the significance of fly-screens to reduce Campylobacter in broiler flocks. The results of this study may be used in identification of effective interventions aimed at controlling Campylobacter in Danish broiler flocks.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Division of Epidemiology and Microbial Genomics, National Veterinary Institute
Contributors: Sommer, H. M., Heuer, O. E., Sørensen, A. I. V., Madsen, M.
Pages: 100-111
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Preventive Veterinary Medicine
Volume: 111
Issue number: 1-2
ISSN (Print): 0167-5877
Ratings:
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 2.49 SJR 1.219 SNIP 1.54
Web of Science (2013): Impact factor 2.506
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Keywords: Campylobacter, Broiler flocks, Risk factors, Multivariate, Generalized linear model
DOI:
10.1016/j.prevetmed.2013.04.004
Source: dtu
Source-ID: n::oai:DTIC-ART:elsevier/387810504::29098
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

Analysis of sensory ratings data with cumulative link models
Examples of categorical rating scales include discrete preference, liking and hedonic rating scales. Data obtained on these scales are often analyzed with normal linear regression methods or with omnibus Pearson chi² tests. In this paper we propose to use cumulative link models that allow for regression methods similar to linear models while respecting the categorical nature of the observations. We describe how cumulative link models are related to the omnibus chi² tests and how they can lead to more powerful tests in the non-replicated setting. For replicated categorical ratings data we present a quasi-likelihood approach and a mixed effects approach both being extensions of cumulative link models. We contrast population-average and subject-specific interpretations based on these models and discuss how different approaches lead to different tests. In replicated settings, naive tests that ignore replications are often expected to be too liberal because of over-dispersion. We describe how this depends on whether the experimental design is fully randomized or blocked. For the latter situation we describe how naive tests can be stronger than over-dispersion adjusting approaches, and that mixed effects models can provide even stronger tests than naive tests. Examples will be given throughout the paper and the methodology is implemented in the authors’ free R-package ordinal.
A new method for image segmentation based on Fuzzy C-means algorithm on pixonal images formed by bilateral filtering

In this paper, a new pixon-based method is presented for image segmentation. In the proposed algorithm, bilateral filtering is used as a kernel function to form a pixonal image. Using this filter reduces the noise and smoothes the image slightly. By using this pixon-based method, the image over segmentation could be avoided. Indeed, the bilateral filtering, as a preprocessing step, eliminates the unnecessary details of the image and results in a few numbers of pixons, faster performance and more robustness against unwanted environmental noises. Then, the obtained pixonal image is segmented using the hierarchical clustering method (Fuzzy C-means algorithm). The experimental results show that the proposed pixon-based approach has a reduced computational load and a better accuracy compared to the other existing pixon-based image segmentation techniques.

A New Paradigm in Mortgage Loan Advice

The Danish mortgage market has undergone considerable changes during the last 15 years. New and more complex variations of loan products have been introduced. Nevertheless, mortgage loan advice has remained, by large,
unchanged. This paper addresses a study where a number of new refinancing rules are constructed with the help of a stochastic optimization model and the data mining method, CART.

**General information**
Publication status: Published
Organisations: Department of Management Engineering, Management Science, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Otterstedt, M. S., Rasmussen, K. M., Kulahci, M.
Number of pages: 9
Publication date: 2013

**Host publication information**
Title of host publication: Proceedings of ORS 2013
Publisher: Social Science Electronic Publishing, Inc

**Application of X-ray phase-contrast tomography in quantitative studies of heat induced structural changes in meat**
X-ray computed tomography is increasingly used in the studies of food structure. This paper describes the perspectives of use of phase contrast computed tomography in studies of heat induced structural changes in meat. From the data it was possible to obtain reconstructed images of the sample structure for visualization and qualitative studies of the sample structure. Further data segmentation allowed structural changes to be quantified.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen
Contributors: Miklos, R., Nielsen, M. S., Einarsdottir, H., Lametsch, R.
Number of pages: 4
Publication date: 2013

**Host publication information**
Title of host publication: Proceedings of the 59th International Congress of Meat Science and Technology (ICoMST 2013)
Keywords: Connective tissue, Myofibrils, Cooking, Quantitative analysis

**A sampling approach for predicting the eating quality of apples using visible–near infrared spectroscopy**
BACKGROUND
Visible–near infrared spectroscopy remains a method of increasing interest as a fast alternative for the evaluation of fruit quality. The success of the method is assumed to be achieved by using large sets of samples to produce robust calibration models. In this study we used representative samples of an early and a late season apple cultivar to evaluate model robustness (in terms of prediction ability and error) on the soluble solids content (SSC) and acidity prediction, in the wavelength range 400–1100 nm.

RESULTS
A total of 196 middle–early season and 219 late season apples (Malus domestica Borkh.) cvs ‘Aroma’ and ‘Holsteiner Cox’ samples were used to construct spectral models for SSC and acidity. Partial least squares (PLS), ridge regression (RR) and elastic net (EN) models were used to build prediction models. Furthermore, we compared three sub-sample arrangements for forming training and test sets (‘smooth fractionator’, by date of measurement after harvest and random). Using the ‘smooth fractionator’ sampling method, fewer spectral bands (26) and elastic net resulted in improved performance for SSC models of ‘Aroma’ apples, with a coefficient of variation CVSSC = 13%. The model showed consistently low errors and bias (PLS/EN: R2cal = 0.60/0.60; SEC = 0.88/0.88°Brix; Biascal = 0.00/0.00; R2val = 0.33/0.44; SEP = 1.14/1.03; Biasval = 0.04/0.03). However, the prediction acidity and for SSC (CV = 5%) of the late cultivar ‘Holsteiner Cox’ produced inferior results as compared with ‘Aroma’.
CONCLUSION

It was possible to construct local SSC and acidity calibration models for early season apple cultivars with CVs of SSC and acidity around 10%. The overall model performance of these data sets also depend on the proper selection of training and test sets. The ‘smooth fractionator’ protocol provided an objective method for obtaining training and test sets that capture the existing variability of the fruit samples for construction of visible–NIR prediction models. The implication is that by using such ‘efficient’ sampling methods for obtaining an initial sample of fruit that represents the variability of the population and for sub-sampling to form training and test sets it should be possible to use relatively small sample sizes to develop spectral predictions of fruit quality. Using feature selection and elastic net appears to improve the SSC model performance in terms of R², RMSECV and RMSEP for ‘Aroma’ apples. © 2013 Society of Chemical Industry

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Dayenú, University of Copenhagen
Contributors: Vega, M. V. M., Sharifzadeh, S., Wulfsohn, D., Skov, T., Clemmensen, L. K. H., Toldam-Andersen, T. B.
Pages: 3710–3719
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Journal of the Science of Food and Agriculture
Volume: 93
Issue number: 15
ISSN (Print): 0022-5142
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 2.22 SJR 0.841 SNIP 1.217
Web of Science (2013): Impact factor 1.879
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Keywords: Malus domestica, SSC, Representative sample, Training set formation, Variability
DOIs:
10.1002/jsfa.6207
Source: dtu
Source-ID: u::7660
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

Automatic scatter detection in fluorescence landscapes by means of spherical principal component analysis

In this paper, we introduce a new method, based on spherical principal component analysis (S-PCA), for the identification of Rayleigh and Raman scatters in fluorescence excitation–emission data. These scatters should be found and eliminated as a prestep before fitting parallel factor analysis models to the data, in order to avoid model degeneracies. The work is inspired and based on a previous research, where scatter removal was automatic (based on a robust version of PCA called ROBPCA) and required no visual data inspection but appeared to be computationally intensive. To overcome this drawback, we implement the fast S-PCA in the scatter identification routine. Moreover, an additional pattern interpolation step that complements the method, based on robust regression, will be applied. In this way, substantial time savings are gained, and the user's engagement is restricted to a minimum, which might be beneficial for certain applications. We conclude that the subsequent parallel factor analysis models fitted to excitation–emission data after scatter identification based on either ROBPCA or S-PCA are comparable; however, the modified method based on S-PCA clearly outperforms the original approach in relation to computational time.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Division of Industrial Food Research
Contributors: Kotwa, E. K., Jørgensen, B. M., Brockhoff, P. B., Frosch, S.
Pages: 3-11
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Journal of Chemometrics
Volume: 27
Cell-Mediated and Humoral Immune Responses after Immunization of Calves with a Recombinant Multiantigenic Mycobacterium avium subsp. paratuberculosis Subunit Vaccine at Different Ages

Neonates and juvenile ruminants are very susceptible to paratuberculosis infection. This is likely due to a high degree of exposure from their dams and an immature immune system. To test the influence of age on vaccine-induced responses, a cocktail of recombinant Mycobacterium avium subsp. paratuberculosis proteins (MAP0217, MAP1508, MAP3701c, MAP3783, and MAP1609c/Ag85B) was formulated in a cationic liposome adjuvant (CAF01) and used to vaccinate animals of different ages. Male jersey calves were divided into three groups that were vaccinated at 2, 8, or 16 weeks of age and boosted twice at weeks 4 and 12 relative to the first vaccination. Vaccine-induced immune responses, the gamma interferon (IFN-γ) cytokine secretion and antibody responses, were followed for 20 weeks. In general, the specific responses were significantly elevated in all three vaccination groups after the first booster vaccination with no or only a minor effect from the second booster. However, significant differences were observed in the immunogenicity levels of the different proteins, and it appears that the older age group produced a more consistent IFN-γ response. In contrast, the humoral immune response is seemingly independent of vaccination age as we found no difference in the IgG1 responses when we compared the three vaccination groups. Combined, our results suggest that an appropriate age of vaccination should be considered in vaccination protocols and that there is a possible interference of vaccine-induced immune responses with weaning (week 8).

General information
Publication status: Published
Organisations: National Veterinary Institute, Section for Immunology and Vaccinology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Statens Serum Institut
Contributors: Thakur, A., Aagaard, C., Stockmarr, A., Andersen, P., Jungersen, G.
Pages: 551-558
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Clinical and Vaccine Immunology (Online)
Volume: 20
Issue number: 4
ISSN (Print): 1556-679X
Ratings:
Scopus rating (2013): CiteScore 2.69
ISI indexed (2013): ISI indexed no
Web of Science (2013): Indexed yes
Original language: English
DOIs:
10.1128/CVI.05574-11
Source: dtu
Source-ID: n:oai:DTIC-ART:highwire/384397893::27610
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review
Code Generation for Protocols from CPN models Annotated with Pragmatics

Model-driven engineering (MDE) provides a foundation for automatically generating software based on models. Models allow software designs to be specified focusing on the problem domain and abstracting from the details of underlying implementation platforms. When applied in the context of formal modelling languages, MDE further has the advantage that models are amenable to model checking which allows key behavioural properties of the software design to be verified. The combination of formally verified models and automated code generation contributes to a high degree of assurance that the resulting software implementation satisfies the properties verified for the model. Coloured Petri Nets (CPNs) have been widely used to model and verify protocol software, but limited work exists on using CPN models of protocol software as a basis for automated code generation. In this report, we present an approach for generating protocol software from a restricted class of CPN models. The class of CPN models considered aims at being descriptive in that the models are intended to be helpful in understanding and conveying the operation of the protocol. At the same time, a descriptive model is close to a verifiable version of the same model and sufficiently detailed to serve as a basis for automated code generation when annotated with code generation pragmatics. Pragmatics are syntactical annotations designed to make the CPN models descriptive and to address the problem that models with enough details for generating code from them tend to be verbose and cluttered. Our code generation approach consists of three main steps, starting from a CPN model that the modeller has annotated with a set of pragmatics that make the protocol structure and the control-flow explicit. The first step is to compute for the CPN model, a set of derived pragmatics that identify control-flow structures and operations, e.g., for sending and receiving packets, and for manipulating the state. In the second step, an abstract template tree (ATT) is constructed providing an association between pragmatics and code generation templates. The ATT then directs the code generation in the third step by invoking the code templates associated with each node of the ATT in order to generate code. We illustrate our approach using an example of a unidirectional data framing protocol.

Comparison of estimated energy intake in children using a Web-based Dietary Assessment Software with accelerometer-estimated energy expenditure in children

Background
The OPUS (Optimal well-being, development and health for Danish children through a healthy New Nordic Diet) project carried out a school meal study to assess the impact of a New Nordic Diet (NND). The random controlled trial involved 834 children aged 8–11 in nine local authority schools in Denmark. Dietary assessment was carried out using a program known as WebDASC (Web-based Dietary Assessment Software for Children) to collect data from the children.

Objective
To compare the energy intake (EI) of schoolchildren aged 8–11 estimated using the WebDASC system against the total energy expenditure (TEE) as derived from accelerometers worn by the children during the same period. A second objective was to evaluate the WebDASC’s usability.

Design
Eighty-one schoolchildren took part in what was the pilot study for the OPUS project, and they recorded their total diet
using WebDASC and wore an accelerometer for two periods of seven consecutive days: at baseline, when they ate their usual packed lunches and at intervention when they were served the NND. EI was estimated using WebDASC, and TEE was calculated from accelerometer-derived activity energy expenditure, basal metabolic rate, and diet-induced thermogenesis. WebDASC's usability was assessed using a questionnaire. Parents could help their children record their diet and answer the questionnaire.

Results
Evaluated against TEE as derived from the accelerometers worn at the same time, the WebDASC performed just as well as other traditional methods of collecting dietary data and proved both effective and acceptable with children aged 8–11, even with perhaps less familiar foods of the NND.

Conclusions
WebDASC is a useful method that provided a reasonably accurate measure of EI at group level when compared to TEE derived from accelerometer-determined physical activity in children. WebDASC will benefit future research in this area.

General information
Publication status: Published
Organisations: National Food Institute, Division of Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Biltoft-Jensen, A. P., Hjort, M. F., Trolle, E., Christensen, T., Brockhoff, P. B., Andersen, L. F., Tetens, I., Matthiessen, J.
Number of pages: 12
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Food & Nutrition Research
Volume: 57
ISSN (Print): 1654-6628
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 1.82 SJR 0.693 SNIP 0.798
Web of Science (2013): Impact factor 1.785
ISI indexed (2013): ISI indexed no
Web of Science (2013): Indexed yes
Original language: English
Electronic versions:
FNR_57_21434.pdf
DOI:
10.3402/fnr.v57i0.21434
URLs:
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

Contextual Multivariate Segmentation of Pork Tissue from Grating-Based Multimodal X-Ray Tomography
X-ray computed tomography is increasingly used as a nondestructive method for studying three dimensional food structures. For meat products, studies have focused mainly on fat and protein content due to limited contrast capabilities of absorption based techniques. Recent advances in X-ray imaging have made novel X-ray image modalities available, where the refraction and scattering of X-rays is obtained simultaneously with the absorption properties, providing enhanced contrast for soft biological tissues. This paper demonstrates how data obtained from grating-based imaging can be segmented by means of multivariate and contextual methods to improve the classification of soft tissues in meat products. The results show that the presented segmentation method provides improved classification over univariate segmentation.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Technische Universität München, Danish Technological Institute, University of Copenhagen
Contributors: Einarsdóttir, H., Nielsen, M. S., Ersbøll, B. K., Larsen, R., Bech, M., Pfeier, F., Christensen, L. B., Feidenhans'l, R.
Pages: 37-42
Publication date: 2013

Host publication information
Title of host publication: Workshop on Farm Animal and Food Quality Imaging 2013 : Espoo, Finland, June 17, 2013, Proceedings
Correlation of iris biometrics and DNA

The presented work concerns prediction of complex human phenotypes from genotypes. We were interested in correlating iris color and texture with DNA. Our data consist of 212 eye images along with DNA: 32 single-nucleotide polymorphisms (SNPs). We used two types of biometrics to describe the eye images: One for iris color and one for iris texture. Both biometrics were high dimensional and a sparse principle component analysis (SPCA) reduced the dimensions and resulted in a representation of data with good interpretability. The correlations between the sparse principal components (SPCs) and the 32 SNPs were found using a canonical correlation analysis (CCA). The result was a single significant canonical correlation (CC) for both biometrics. Each CC comprised two correlated canonical variables, consisting of a linear combination of SPCs and a linear combination of SNPs, respectively. The significant canonical variables for color and texture were primarily explained by the first SPC (SPC1). Therefore, we made a visual inspection of the first SPCs. The color based SPC1 explained a blue to brown variation in iris color and the texture based SPC1 gave a general explanation of iris texture. The SNPs (rs12896399, rs3733542, rs6475555, rs12913832) and (rs12896399, rs3733542, rs12913832) had the highest correlation to the canonical variable for color and texture, respectively. Three of the most contributing SNPs were the same for both biometrics, revealing a covariance between iris color and texture.

Data Mining—A Special Issue of Quality and Reliability Engineering International (QREI)

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Arizona State University
Contributors: Li, J. (ed.), Kulahci, M. (ed.)
Number of pages: 1
Pages: 437
Publication date: 2013
Peer-reviewed: Yes
DCT-Based Characterization of Milk Products Using Diffuse Reflectance Images
We propose to use the two-dimensional Discrete Cosine Transform (DCT) for decomposition of diffuse reflectance images of laser illumination on milk products in different wavelengths. Based on the prior knowledge about the characteristics of the images, the initial feature vectors are formed at each wavelength. The low order DCT coefficients are used to quantify the optical properties. In addition, the entropy information of the higher order DCT coefficients is used to include the illumination interference effects near the incident point. The discrimination powers of the features are computed and used to do wavelength and feature selection. Using the selected features of just one band, we could characterize and discriminate eight different milk products. Comparing this result with the current characterization method based of a fitted log-log linear model, shows that the proposed method can discriminate milk from yogurt products better.

Dynamic changes in antibody levels as an early warning of Salmonella Dublin in bovine dairy herds
Salmonella Dublin is a bacterium that causes disease and production losses in cattle herds. In Denmark, a surveillance and control program was initiated in 2002 to monitor and reduce the prevalence of Salmonella Dublin. In dairy herds, the surveillance includes herd classification based on bulk tank milk measurements of antibodies directed against Salmonella Dublin at 3-mo intervals. In this study, an "alarm herd" concept, based on the dynamic progression of these repeated measurements, was formulated such that it contains predictive power for Salmonella Dublin herd classification change from "likely free of infection" to "likely infected" in the following quarter of the year, thus warning the farmer 3mo earlier than the present system. The alarm herd concept was defined through aberrations from a stable development over time of antibody levels. For suitable parameter choices, alarm herd status was a positive predictor for Salmonella Dublin status change in dairy herds, in that alarm herds had a higher risk of changing status in the following quarter compared with nonalarm herds. This was despite the fact that both alarm and nonalarm herds had antibody levels that did not indicate the herds being "likely infected" according to the existing classification system in the present quarter. The alarm herd concept can be used as a new early warning element in the existing surveillance program. Additionally, to improve accuracy of herd classification, the alarm herd concept could be incorporated into a model including other known risk factors for change in herd classification. Furthermore, the model could be extended to other diseases monitored in similar ways.
Effectiveness of offering keyhole labelled meals in improving the nutritional quality of lunch meals eaten in worksite canteens

General information
Publication status: Published
Organisations: National Food Institute, Division of Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Technical University of Denmark, Copenhagen University Hospital
Contributors: Lassen, A. D., Beck, A. M., Leedo, E., Andersen, E. W., Christensen, T., Mejborn, H., Thorsen, A. V., Tetens, I.
Pages: 954-954
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Annals of Nutrition and Metabolism
Volume: 63
Issue number: SI
ISSN (Print): 0250-6807
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 2.46 SJR 0.932 SNIP 1.008
Web of Science (2013): Impact factor 2.747
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Research output: Contribution to journal › Conference abstract in journal – Annual report year: 2013 › Research › peer-review

Effect of fat type and heat treatment on the microstructure of meat emulsions

In comminuted meat products the gel-forming abilities of the myofibrillar proteins are of major importance. In meat emulsions fat will be present in globules which are stabilized by a membrane coating made of salt-soluble proteins. These discontinuous fat particles act as fillers or co-polymers and stabilize the protein network. Differences in the physicochemical properties of saturated and unsaturated lipids affect the distribution of fat and thereby the functionality and quality of the final product. The objectives were to study the effects of lipid type and heat treatment on changes in microstructure of meat emulsions by use of a novel quantitative application of absorption- and phase-contrast tomography. The non-invasive technique offered the possibility to study the same sample in both raw and cooked condition. The samples were raw and heat treated meat emulsions (10% protein, 25% fat, 60% moisture) prepared with either pork fat or sunflower oil. The tomograms were obtained at a synchrotron facility using a grating interferometer which measured three different properties in the sample simultaneously: The attenuation length, the electron density and the diffusion length. Phase contrast imaging of the tomograms was used to analyse the impact of lipid type on spatial fat distribution, microstructure of the protein network and structural changes caused by heat treatment. The tomograms showed that the fat distribution in the meat emulsions depended on the physicochemical properties of the added fat. Use of vegetable oil resulted in homogeneous emulsions with smaller fat globules compared to the use of pork fat. This has previously been shown by the use of light micrographs. However, with the use of phase contrast imaging it was, from the same image,
possible to resolve the protein phase to obtain information about the quality of the protein network and of the changes in microstructure caused by heat treatment. Further it was possible to compare the amounts of cooking loss from the emulsions. In conclusion phase contrast imaging with its high spectral resolution offers a unique possibility for studies of microstructure and is superior to histology since the information is obtained for the full volume.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen
Contributors: Miklos, R., Lametsch, R., Nielsen, M. S., Lauridsen, T., Einarsdottir, H.
Number of pages: 5
Publication date: 2013

**Host publication information**
Title of host publication: InsideFood Symposium : Book of Proceedings
Electronic versions:
URLs:
http://www.insidefood.eu/proceedings.awp

**Bibliographical note**
Poster presentation.
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2013 › Research › peer-review

**Effects of intraspecific variation in reproductive traits, pectoral fin use and burst swimming on metabolic rates and swimming performance in the Trinidadian guppy (Poecilia reticulata)**

There is considerable intraspecific variation in metabolic rates and locomotor performance in aquatic ectothermic vertebrates; however, the mechanistic basis remains poorly understood. Using pregnant Trinidadian guppies (Poecilia reticulata), a livebearing teleost, we examined the effects of reproductive traits, pectoral fin use and burst-assisted swimming on swimming metabolic rate, standard metabolic rate (MO2std) and prolonged swimming performance (Ucrit). Reproductive traits included reproductive allocation and pregnancy stage, the former defined as the mass of the reproductive tissues divided by the total body mass. Results showed that the metabolic rate increased curvilinearly with swimming speed. The slope of the relationship was used as an index of swimming cost. There was no evidence that reproductive traits correlated with swimming cost, MO2std or Ucrit. In contrast, data revealed strong effects of pectoral fin use on swimming cost and Ucrit. Poecilia reticulata employed body-caudal fin (BCF) swimming at all tested swimming speeds; however, fish with a high simultaneous use of the pectoral fins exhibited increased swimming cost and decreased Ucrit. These data indicated that combining BCF swimming and pectoral fin movement over a wide speed range, presumably to support swimming stability and control, is an inefficient swimming behaviour. Finally, transition to burst-assisted swimming was associated with an increase in aerobic metabolic rate. Our study highlights factors other than swimming speed that affect swimming cost and suggests that intraspecific diversity in biomechanical performance, such as pectoral fin use, is an important source of variation in both locomotor cost and maximal performance.

**General information**
Publication status: Published
Organisations: National Institute of Aquatic Resources, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Section for Freshwater Fisheries Ecology, University of California, University of Copenhagen
Contributors: Svendsen, J. C., Banet, A. I., Christensen, R. H. B., Steffensen, J. F., Aarestrup, K.
Pages: 3564-3574
Publication date: 2013
Peer-reviewed: Yes

**Publication information**
Journal: Journal of Experimental Biology
Volume: 216
ISSN (Print): 0022-0949
Ratings:
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 2.75 SJR 1.715 SNIP 1.328
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Effects of mid-term student evaluations of teaching as measured by end-of-term evaluations: An empirical study of course evaluations

Universities have varying policies on how and when to perform student evaluations of courses and teachers. More empirical evidence of the consequences of such policies on quality enhancement of teaching and learning is needed. A study (35 courses at the Technical University of Denmark) was performed to illustrate the effects caused by different handling of mid-term course evaluations on student's satisfaction as measured by end-of-term evaluations. Midterm and end-of-term course evaluations were carried out in all courses. Half of the courses were allowed access to the midterm results. The evaluations generally showed positive improvements over the semester for courses with access, and negative improvements for those without access. Improvements related to: Student learning, student satisfaction, teaching activities, and communication showed statistically significant average differences of 0.1-0.2 points between the two groups. These differences are relatively large compared to the standard deviation of the scores when student effect is removed (approximately 0.7). We conclude that university policies on course evaluations seem to have an impact on the development of the teaching and learning quality as perceived by the students and discuss the findings.

Estimating passenger numbers in trains using existing weighing capabilities

Knowing passenger numbers is important for the planning and operation of the urban rail systems. Manual and electronic counting systems (typically infrared or video) are expensive and therefore entail small sample sizes. They usually count boarding and alighting passengers, which means that errors in estimates of total numbers of passengers propagate along train runs. Counting errors in manual and electronic counting systems are typically flow-dependent, making uncertainty a function of volume. This paper presents a new counting technique that exploits the weighing systems installed in most modern trains to control braking. This technique makes passenger counting cheaper and ensures a complete sample. The paper compares numbers estimated by this technique with manual counts and counts from an infrared system in trains in urban Copenhagen. It shows that the weighing system provides more accurate passenger counts than the infrared equipment. The method has been validated on a large data set and is now in full operation in the urban Copenhagen rail system.
Favorable results after total wrist arthroplasty: 65 wrists in 60 patients followed for 5–9 years

Background and purpose

During the past 40 years, several attempts have been made with total wrist arthroplasty to avoid fusion in severely destroyed wrists. The results have often been disappointing. There is only modest clinical documentation due to the small number of patients (especially non-rheumatoid cases) and short follow-up times. Here we report a multicenter series using a third-generation implant with a minimum follow-up time of 5 years.

Methods

In 2012, data were retrieved from a registry of consecutive wrist operations at 7 centers with units specialized in hand surgery, between 2003 and 2007. The wrists had been reviewed annually and analysis was done on the latest follow-up data.

Results

60 patients had been operated (5 bilaterally), 5 wrists had been revised, and 52 were available for follow-up (with the revised cases excluded). The pain scores, QuickDASH scores, ulnar flexion, and supination for the whole group were statistically significantly better at follow-up. There were no statistically significant differences between the rheumatoid and the non-rheumatoid patients except for motion, which was better in the non-rheumatoid group. The motion obtained depended on the preoperative motion. Implant survival was 0.9 at 5–9 years.

Interpretation

The clinical results in terms of pain, motion, strength, and function were similar to those in previous reports. The implant survival was 0.9 at 9 years, both in rheumatoid and non-rheumatoid cases, which is an important improvement compared to the earlier generations of total wrist arthroplasty.
Impact of changes in metabolic control on progression to photocoagulation for clinically significant macular oedema: a 20 year study of type 1 diabetes

Aims/hypothesis
Although increasing hyperglycaemia, arterial hypertension and longer duration of diabetes raise the risk of progression of diabetic retinopathy, short-term benefits in terms of improved metabolic control and lowered blood pressure have not been demonstrated. We therefore examined the effect of changes in glycaemia and arterial blood pressure on the incidence of clinically significant macular oedema in a population of diabetic patients.

Methods
We performed a retrospective review of all patients with type 1 diabetes who attended the retinopathy screening clinic at the Steno Diabetes Center from 1988 to 2008, using the endpoint referral to first photocoagulation treatment for clinically significant diabetic macular oedema. The analysis included 1,878 patients (median observation, 8 years). Changes were defined as the inter-visit change; in the case of an event the last event-free interval before referral, where the median screening interval was 6 months.

Results
Risk of progression to photocoagulation for macular oedema increased with duration of diabetes (p < 0.001), current HbA1c (p < 0.0001) and with the magnitude of changes in HbA1c (p = 0.0002) and systolic blood pressure (p < 0.0001) in a multiple regression model. A recent decrease of ≥0.5 percentage points or an increase in HbA1c of >0.5 percentage points per 6 months was associated with HRs of 3.04 and 1.28, respectively, compared with lesser changes in HbA1c.

Conclusions/interpretation
In this study, large recent changes in metabolic control and systolic blood pressure, irrespective of direction, were independent risk factors for progression to photocoagulation for diabetic macular oedema. The effects of metabolic and haemodynamic stability on diabetic retinopathy should be examined in prospective studies.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Copenhagen University Hospital
Contributors: Sander, B., Larsen, M., Andersen, E. W., Lund-Andersen, H.
Pages: 2359-2366
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Diabetologia
Volume: 56
Issue number: 11
ISSN (Print): 0012-186X
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 6 SJR 3.188 SNIP 2.031
Web of Science (2013): Impact factor 6.88
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Keywords: Blood pressure, Clinically significant macular oedema, HbA1c, Risk factors, Type 1 diabetes
Electronic versions:
art_3A10.1007_2Fs00125_013_3027_5.pdf
DOIs: 10.1007/s00125-013-3027-5
Source: dtu
Source-ID: u::8568
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

Impact of product information and repeated exposure on consumer liking, sensory perception and concept associations of local apple juice
The impact of product information and repeated exposure of local apple juice on consumer liking, sensory perception and concept associations was studied. Findings showed that consumers had high expectations towards the studied local apples juices. Consequently, the liking for the local apple juices was higher when the consumers were informed about the product versus blind tasting. This shift was explained by an assimilation effect i.e., the consumers rated the liking in the informed product condition closer to their expectations. However, the assimilation effect was incomplete as the product
information did not fully determine consumer liking. Thus the sensory characteristics were also found to be of importance. Besides liking, the concept associations were also affected by product information. All local apple juices were associated as being more exclusive under informed tasting conditions. No effects were found for liking of the apple juices over repeated exposure. Also, the consumers did not change their sensory perception over repeated consumption. However, the local apple juices were conceptualised as more familiar but remained exclusive after repeated consumption. When local producers have to market their products, they are recommended to focus on product information as a selling point as this information strongly influences liking, concept associations and sensory perception of the product. However, in order to ensure successful products, the sensory characteristics of the product must not be compromised as these influence liking in a manner increasing over repeated consumption.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Contributors: Stolzenbach, S., Bredie, W. L. P., Christensen, R. H. B., Byrne, D. V.
Pages: 91-98
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Food Research International
Volume: 52
Issue number: 1
ISSN (Print): 0963-9969
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.68 SJR 1.514 SNIP 1.818
Web of Science (2013): Impact factor 3.05
ISI indexed (2013): ISI indexed yes
Original language: English
Keywords: Local foods, Expectations, Assimilation, Repeated exposure, Danish apple juice
DOIs: 10.1016/j.foodres.2013.02.018
Source: dtu
Source-ID: n:oai:DTIC-ART:compendex/385331860::27972
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

Importance of fruit variability in the assessment of apple quality by sensory evaluation
The assessment of produce quality is a major aspect of applied postharvest biology. Horticultural researchers working on organoleptic quality of fruit need objective methods for the evaluation of sensory properties. The development of sensory methodologies specifically for apples highlighted the problem of handling variation due to fruit variability and assessor differences. The aim of this study was to investigate the weight of within-batch variability in sensory evaluation of apples and to propose a methodology that accounts for this variability. Prior to sensory analysis, for three apple cultivars, apples were sorted into homogenous acoustic firmness categories within each cultivar. The discrimination ability of the trained panel was observed not only between cultivars but also within each cultivar for crunchiness, firmness, juiciness and acidity. Following these results, a mixed hierarchical model for the analysis of the sensory data was proposed to measure the contribution of fruit variability to the variability of sensory scores. The results showed the efficiency of the model in quantifying within-batch variability. Fruit sampling and presentation methods as well as data handling procedures are suggested for obtaining reliable sensory results in the assessment of apple quality.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Université Nantes Angers Le Mans
Contributors: Bavay, C., Symoneaux, R., Maître, I., Kuznetsova, A., Brockhoff, P. B., Mehinagic, E.
Pages: 67-74
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Postharvest Biology and Technology
Volume: 77
ISSN (Print): 0925-5214
Localisation of nursery areas based on comparative analyses of the horizontal and vertical distribution patterns of juvenile Baltic cod (Gadus morhua)

Knowledge of the spatial distribution of juvenile cod is essential for obtaining precise recruitment data to conduct sustainable management of the eastern and western Baltic cod stocks. In this study, the horizontal and vertical distribution and density patterns of settled juvenile 0- and 1-group Baltic cod are determined, and their nursery areas are localised according to the environmental factors affecting them. Comparative statistical analyses of biological, hydrographic and hydroacoustic data are carried out based on standard ICES demersal trawl surveys and special integrated trawl and acoustic research surveys. Horizontal distribution maps for the 2001–2010 cohorts of juvenile cod are further generated by applying a statistical log-Gaussian Cox process model to the standard trawl survey data. The analyses indicate size-dependent horizontal and distinct vertical and diurnal distribution patterns related to the seabed topography, water layer depth, and the presence of hydrographic frontal zones (pycnoclines) as well as intraspecific patterns in relation to the presence of adult cod. The extent of the nursery areas also depends on the cod year class strength. Juvenile cod (≥3 cm) are present in all areas of the central Baltic Sea (CBS), showing broad dispersal. However, their highest density in the Baltic Basins is found at localities with a 40–70 m bottom depth in waters with oxygen concentrations above 2 ml O2.l−1 and temperatures above 5°C. The smallest juveniles are also found in deep sea localities down to a 100 m depth and at oxygen concentrations between 2–4 ml O2.l−1. The vertical, diurnally stratified and repeated trawling and hydroacoustic target strength-depth distributions obtained from the special surveys show juvenile cod concentrations in frontal zone water layers (pycnocline). However, the analyses indicate that in the CBS, juvenile cod of all sizes do not appear to aggregate in dense schooling patterns, which differs from what has been reported from the North Sea.
Making sense of zeros: impact on human health risk estimates

General information
Publication status: Published
Organisations: National Food Institute, Division of Epidemiology and Microbial Genomics, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Ribeiro Duarte, A. S., Stockmarr, A., Nauta, M.
Pages: 55
Publication date: 2013

Host publication information
Place of publication: Lyngby
Electronic versions:
MVN2013ConfAbstractBook_v1.0.pdf
URLs:

Bibliographical note
RR07
Research output: Chapter in Book/Report/Conference proceeding › Conference abstract in proceedings – Annual report year: 2013 › Research › peer-review

Matrix Analytic Methods in Applied Probability with a View towards Engineering Applications
Queueing permeates most of man's commercial behaviour. People queue in stores, at banks, and at restaurants. Cars, ships and air-planes queue at roads, ports and runways, while electronic messages queue at transmission lines and in servers, waiting to be processed. The mathematical modelling of queueing phenomena has developed at ever increasing speed since the birth of queueing theory in the early 20th century, usually ascribed to A. K. Erlang, who worked as a mathematician for the Copenhagen Telephone Company (KTAS – Københavns Telefons Aktie Selskab). Erlang is perhaps the foremost representative among many Danish and Scandinavian scientists who have contributed profoundly to queueing theory and the closely related field of risk theory, with insurance being its main application area. When modelling queueing systems a model is needed for the description of the random arrival stream of demands, in terms of customers in the various forms of people, cars, ships, or electronic messages. Point process theory arose from the field of applied probability to address this need.

The Markovian Arrival Process (MAP) is one of the main concrete manifestations of point process theory. The MAP is an essential building block within matrix analytic methods in queueing theory pioneered by Neuts and coauthors. The theory of matrix analytic methods is appealing from a practical point of view as many systems can be analytically and numerically evaluated using this approach.

In this thesis we present contributions to the theoretical development of the field of matrix analytic methods including an extension to a multivariate setting. We further demonstrate the applicability of the theory, giving examples from telecommunications engineering and computer science.

The thesis is based on a number of original contributions and a summary introductory paper. The outline of the summary is as follows.

The class of MAPs and the related class of Phase Type (PH) distributions belong to the slightly larger classes of what
have been termed Rational Arrival Processes (RAP) and Matrix Exponential (ME) distributions, respectively. In Chapter 2 we present the basic constructions of phase-type and matrix-exponential distributions along with the Markovian and rational arrival processes. We briefly mention some well-known properties of these constructions while describing our own contributions in more detail. Chapter 3 is devoted to discussion of parameter estimation in the models described in Chapter 2. We give a very brief review of current estimation methods while focusing on our own contributions.

Chapters 4 and 5 contain different aspects of applications. The MAP is a versatile tool in sensitivity analyses of stochastic systems since point process descriptors of a MAP can be evaluated numerically. Sensitivity analyses based on the MAP are described in Chapter 4. That chapter is somewhat more generic in nature than Chapter 5 in which some concrete examples of engineering applications are presented.

In Chapter 6 we present two different ways of proving how the matrix analytic results related to the classical models of phase-type distributions and Markovian arrival processes extend to the case of matrix-exponential distributions and rational arrival processes.

In Chapter 7 we introduce the classes multivariate matrix-exponential and bilateral multivariate matrix-exponential distributions. The chapter starts with a small review of previous work on multivariate phase-type distributions while the rest of the chapter contains recent results of our own research.

The main contributions of the thesis are described in Chapters 4, 6, and 7. Chapter 4 is important from an engineering perspective. The approach described in reference [8] was somewhat controversial at the time. Measurements in packet based communication networks made some researchers call for a paradigm shift in queueing theory, where models based on Markovian assumptions would be, if not superfluous, then at least of minor importance. The contribution of [8] was to show that the Markovian arrival process could indeed remain a useful tool in modelling modern communication systems. The paper and its preliminary version reference [7] have been widely cited. Also reference [4] is important as this paper exemplifies how sensitivity analyses of queueing systems can be carried out using the Markovian arrival process, frequently leading to conclusions of general validity.

The two final chapters, 6 and 7, contain substantial theoretical contributions. The importance of Chapter 6 is at present primarily the mathematical content. It has been satisfying to finally settle the common anticipation that results for PH distributions and MAPs carry over verbatim to the case of ME distributions and RAPs. The method of proof has to rely on new ideas, as the standard probabilistic line of reasoning breaks down in the case of matrix-exponential distributions and rational arrival processes. Two different proof techniques were applied. In reference [17] a continuous time analysis based on a last exit time approach was applied. The approach taken in reference [18] was that of an embedded Markov chain with a general state space.

Finally Chapter 7 describes the contributions of the references [22, 24, 25, 28] containing the definition of the important class of bilateral multivariate matrix-exponential distributions together with examples of their use and various related results. These distributions provide a very flexible tool for modelling multivariate phenomena. The definition seems to be the natural multivariate generalisation of matrix-exponential distributions. The main result is a characterisation theorem similar to the main characterisation theorem of the multivariate normal distribution. Finally we demonstrate how the MVME distribution class unifies a number of previously published models in a way quite similar to the way PH and ME distributions unified a number of seemingly loosely connected models and results. The work described in Chapter 7 opens several non-trivial mathematical and theoretical questions. If just some of these problems can be solved satisfactorily it will pave the way for a huge application potential, and it is very likely that the distributions can and will be useful in statistical analysis too. The research on multivariate distributions lead to reference [27] describing a closure property of matrix-exponential and phase-type distributions.

In general, results from our own research will be stated as definitions, lemmas, corollaries, and theorems, while other results will be part of the text flow. The notation used in the papers is generally similar to that of the summary, and it is my hope that the slight differences will not reduce the accessibility of the papers.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Nielsen, B. F.
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
ISBN (Print): 978-87-643-1291-1
Original language: English
Electronic versions:
Thesis summary
Research output: Book/Report › Doctoral thesis – Annual report year: 2013 › Research
Maximizing Entropy over Markov Processes
The channel capacity of a deterministic system with confidential data is an upper bound on the amount of bits of data an attacker can learn from the system. We encode all possible attacks to a system using a probabilistic specification, an Interval Markov Chain. Then the channel capacity computation reduces to finding a model of a specification with highest entropy. Entropy maximization for probabilistic process specifications has not been studied before, even though it is well known in Bayesian inference for discrete distributions. We give a characterization of global entropy of a process as a reward function, a polynomial algorithm to verify the existence of an system maximizing entropy among those respecting a specification, a procedure for the maximization of reward functions over Interval Markov Chains and its application to synthesize an implementation maximizing entropy. We show how to use Interval Markov Chains to model abstractions of deterministic systems with confidential data, and use the above results to compute their channel capacity. These results are a foundation for ongoing work on computing channel capacity for abstractions of programs derived from code.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, IT University of Copenhagen, French National Institute for Computer Science and Applied Mathematics
Contributors: Biondi, F., Legay, A., Nielsen, B. F., Wąsowski, A.
Pages: 128-140
Publication date: 2013

Monitoring Motion of Pigs in Thermal Videos
We propose a new approach for monitoring animal movement in thermal videos. The method distinguishes movements as walking in the expected direction from walking in the opposite direction, stopping or lying down. The method utilizes blob detection combined with optical flow to segment the pigs and extract features which characterize a pig’s movement (direction and speed). Subsequently a multiway principal component analysis is used to analyze the movement features and monitor their development over time. Results are presented in the form of quality control charts of the principal components. The method works on-line with pre-training.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Gronskyte, R., Kulahci, M., Clemmensen, L. K. H.
Pages: 31-36
Publication date: 2013
Multispectral Image Analysis for Robust Prediction of Astaxanthin Coating

The aim of this study was to investigate the possibility of predicting the type and concentration level of astaxanthin coating of aquaculture feed pellets using multispectral image analysis. We used both natural and synthetic astaxanthin, and we used several different concentration levels of synthetic astaxanthin in combination with four different recipes of feed pellets. We used a VideometerLab with 20 spectral bands in the range of 385-1050 nm. We used linear discriminant analysis and sparse linear discriminant analysis for classification and variable selection. We used partial least squares regression (PLSR) for prediction of the concentration level. The results show that it is possible to predict the level of synthetic astaxanthin coating using PLSR on either the same recipe, or when calibrating on all recipes. The concentration prediction is adequate for screening for all recipes. Moreover, it shows that it is possible to predict the type of astaxanthin used in the coating using only ten spectral bands. Finally, the most selected spectral bands for astaxanthin prediction are in the visible range of the spectrum.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, National Food Institute, Division of Industrial Food Research, Division of Toxicology and Risk Assessment, Statistics and Data Analysis
Contributors: Ljungqvist, M. G., Frosch, S., Nielsen, M. E., Erskøll, B. K.
Pages: 738-746
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Applied Spectroscopy
Volume: 67
Issue number: 7
ISSN (Print): 0003-7028
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 2.08 SJR 0.641 SNIP 1.183
Web of Science (2013): Impact factor 2.014
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Keywords: Multispectral, Image analysis, Spectral imaging, NIR, Astaxanthin, Fish feed, Coating
Electronic versions:
Multispectral Image Analysis for Robust Prediction of Astaxanthin Coating AS 2013.pdf
DOIs: 10.1366/12-06823

Bibliographical note
This paper was published in Applied Spectroscopy and is made available as an electronic reprint with the permission of OSA. The paper can be found at the following URL on the OSA website:
http://www.opticsinfobase.org/as/abstract.cfm?URI=as-67-7-738. Systematic or multiple reproduction or distribution to multiple locations via electronic or other means is prohibited and is subject to penalties under law.
Source: dtu
Source-ID: u::7980
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

Multivariate Statistical Process Control
As sensor and computer technology continues to improve, it becomes a normal occurrence that we confront with high dimensional data sets. As in many areas of industrial statistics, this brings forth various challenges in statistical process control (SPC) and monitoring for which the aim is to identify “out-of-control” state of a process using control charts in order to reduce the excessive variation caused by so-called assignable causes. In practice, the most common method of monitoring multivariate data is through a statistic akin to the Hotelling’s T2. For high dimensional data with excessive amount of cross correlation, practitioners are often recommended to use latent structures methods such as Principal Component Analysis to summarize the data in only a few linear combinations of the original variables that capture most of the variation in the data. Applications of these control charts in conjunction with image data are plagued with various challenges beyond the usual ones encountered in current applications. In this presentation we will introduce the basic ideas of SPC and the multivariate control charts commonly used in industry. We will further discuss the challenges the practitioners are facing with in the implementation of these charts.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
On discriminant analysis techniques and correlation structures in high dimensions

This paper compares several recently proposed techniques for performing discriminant analysis in high dimensions, and illustrates that the various sparse methods differ in prediction abilities depending on their underlying assumptions about the correlation structures in the data. The techniques generally focus on two things: Obtaining sparsity (variable selection) and regularizing the estimate of the within-class covariance matrix. For high-dimensional data, this gives rise to increased interpretability and generalization ability over standard linear discriminant analysis. Here, we group the methods in two: Those who assume independence between the variables and thus use a diagonal estimate of the within-class covariance matrix, and those who assume dependence between the variables and thus use an estimate of the within-class covariance matrix, which also estimates the correlations between variables. The two groups of methods are compared and the pros and cons are exemplified using different cases of simulated data. The results illustrate that the estimate of the covariance matrix is an important factor with respect to choice of method, and the choice of method should thus be driven by the nature of the problem at hand.

Optimal vision system design for characterization of apples using US/VIS/NIR spectroscopy data

Quality monitoring of the food items by spectroscopy provides information in a large number of wavelengths including highly correlated and redundant information. Although increasing the information, the increase in the number of wavelengths causes the vision set-up to be more complex and expensive. In this paper, three sparse regression methods; lasso, elastic-net and fused lasso are employed for estimation of the chemical and physical characteristics of one apple cultivar using their high dimensional spectroscopic measurements. The use of sparse regression reduces the number of required wavelengths for prediction and thus, simplifies the required vision set-up. It is shown that, considering a tradeoff between the number of selected bands and the corresponding validation performance during the training step can result in a significant reduction in the number of bands at a small price in the test performance. Furthermore, appropriate regression methods for different number of bands and spectrophotometer design are determined.
Persistent of low-pathogenic H5N7 and H7N1 avian influenza subtypes in filtered natural waters

Wild aquatic birds are the natural reservoir of avian influenza virus (AIV), and the virus is transmitted among birds through a fecal-oral route. Infected birds excrete significant amounts of AIV into the environment, and thereby sustain the circulation of AIV in the bird populations. Improved knowledge on the influence of environmental factors on the persistence of AIV in natural habitats would be valuable for risk assessments. The presented work investigated the persistence of two low-pathogenic AIV subtypes in natural water samples. The study included two AIVs formerly isolated from wild ducks, which were suspended in filtered natural fresh, brackish or sea water with salinity of 0, 8000 and 20,000 parts per million (ppm), respectively. Also sterilized brackish and sea waters were included in order to examine the influence of microbial flora on virus persistence. All water samples were incubated at temperatures representative for seasonal variation of ambient temperatures in Northern Europe (4, 17 and 25°C). The results showed a clear correlation between persistence of viral infectivity and temperature, salinity and presence of microbial flora. While independent of virus subtype, the persistence of infectivity was negatively affected by increased temperature, salinity as well as presence of natural microbial flora. The study provides insight on impact of essential physical, chemical and biological parameters on persistence of AIV in aquatic environments. Studies determining the importance of additional environmental parameters and the detailed mechanisms of microbial inactivation of AIV should be encouraged.

General information
Publication status: Published
Organisations: National Veterinary Institute, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Nielsen, A. A., Jensen, T. H., Stockmarr, A., Jørgensen, P. H.
Pages: 419-428
Publication date: 2013
Peer-reviewed: Yes

PMWS development in pigs from affected farms in Spain and Denmark

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Stockmarr, A.
PMWS development in pigs from affected farms in Spain and Denmark

Postweaning multisystemic wasting syndrome (PMWS) is a worldwide spread condition that affects pigs in nursery and/or fattening units, and is considered to have a severe economic impact on swine production. The main clinical sign of PMWS is wasting, but can also include pallor of the skin, icterus, respiratory distress and diarrhoea. The main essential infectious agent for PMWS development is porcinecircovirus type 2 (PCV2), but the exact cause of PMWS is still unclear. PCV2 is present in most pig herds, but the occurrence of PMWS is more sporadic, and it is been difficult to reproduce PMWS by inoculating PCV2 alone. However, studies where co-infections have been applied have been more successful. Based on this, we modeled PMWS development based on longitudinal data on antibodies and PMWS status from herds in Denmark and Spain, where presence of a range of pathogens were considered as explanatory variables in the form of maternal immunity and the occurrence of seroconversion against the considered pathogens. However, maternal immunity could not be measured from mother animals due to cross fostering, no time points for seroconversion was available, and no case/control status could be assigned as PMWS do not have an ‘infectious period’ after which animals may be assigned control status. The talk will concentrate on the framework in which this was handled, which may be translated to similar settings for similar studies. We found that seroconversion towards PCV2 and Lawsonia intracellularis had a significant impact on PMWS in the Danish data, but it appears that the effect is positive, in the sense that seroconverted animals were less likely to develop PMWS. A number of maternal immunities also significantly affected PMWS development. Furthermore it was uncovered that most of these effects would not have been detected if pathogens were considered by themselves and not simultaneously.
Population dynamics of species-rich ecosystems: the mixture of matrix population models approach

Matrix population models are widely used to predict population dynamics, but when applied to species-rich ecosystems with many rare species, the small population sample sizes hinder a good fit of species-specific models. This issue can be overcome by assigning species to groups to increase the size of the calibration data sets. However, the species classification is often disconnected from the matrix modelling and from the estimation of matrix parameters, thus bringing species groups that may not be optimal with respect to the predicted community dynamics.

We proposed here a method that jointly classified species into groups and fit the matrix models in an integrated way. The model was a special case of mixture with unknown number of components and was cast in a Bayesian framework. An MCMC algorithm was developed to infer the unknown parameters: the number of groups, the group of each species and the dynamics parameters.

We applied the method to simulated data and showed that the algorithm efficiently recovered the model parameters.

We applied the method to a data set from a tropical rain forest in French Guiana. The mixture matrix model classified tree species into well-differentiated groups with clear ecological interpretations. It also accurately predicted the forest dynamics over the 16-year observation period.

Our model and algorithm can straightforwardly be adapted to any type of matrix model, using the life cycle diagram. It can be used as an unsupervised classification technique to group species with similar population dynamics.

General information

Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Centre de cooperation Internationale en Recherche Agronomique pour le Développement
Contributors: Mortier, F., Rossi, V., Guillot, G., Gourlet-Fleury, S., Picard, N.
Pages: 316–326
Publication date: 2013
Peer-reviewed: Yes

Publication information

Journal: Methods in Ecology and Evolution
Volume: 4
Issue number: 4
ISSN (Print): 2041-210X
Ratings: BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 5.34 SJR 2.998 SNIP 2.422
Web of Science (2013): Impact factor 5.322
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Electronic versions:
Preprint
DOIs:
10.1111/2041-210x.12019
Source: dtu
Source-ID: u::5399
Research output: Contribution to journal › Journal article – Annual report year: 2012 › Research › peer-review

PorkCAD: Case study of the design of a pork product prototyper

With the help of industry experts we developed porkCAD, an application intended to aid in the communication process between producer and retailer when developing new meat products for a constantly evolving market. The application interface allows the user to make planar cuts to a virtual pig formed from CT-scans of a real-world pig carcass. We present a case study of the design process from conceptualization to intended introduction into the work flow of a meat production company. We discuss critical design decisions during development and present perspectives for future development.

To determine the usability of porkCAD, we tested it with personnel from the pork industry, using two different controller interfaces, one being a traditional mouse and keyboard input, and the other a six degrees of freedom haptic feedback device. The accurate depiction of pig anatomy guided trained professionals to re-create standardized pig products using porkCAD. The quantitative results of the usability test with sales personnel did not lean significantly in favor of either interface.
Since one interface was extremely well known and the other highly unfamiliar, the fact that users did not express a clear preference for the known input modality is deemed important. We report on the observed user experience regarding the two interfaces.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Cognitive Systems, Statistics and Data Analysis, University of Tokyo, Danish Technological Institute
Contributors: Laursen, L. F., Bærentzen, J. A., Igarashi, T., Petersen, M. K., Clemmensen, L. K. H., Ersbøll, B. K., Christensen, L. B.
Number of pages: 12
Pages: 1134-1
Publication date: 2013

Host publication information
Title of host publication: Proceedings of IASDR 2013
URLs:
Source: dtu
Source-ID: u::9245
Research output: Chapter in Book/Report/Conference proceeding Article in proceedings – Annual report year: 2013 peer-review

Quantifying Dispersal of European Culicoides (Diptera: Ceratopogonidae) Vectors between Farms Using a Novel Mark-Release-Recapture Technique

Studying the dispersal of small flying insects such as Culicoides constitutes a great challenge due to huge population sizes and a lack of a method to efficiently mark and objectively detect many specimens at a time. We here describe a novel mark-release-recapture method for Culicoides in the field using fluorescein isothiocyanate (FITC) as marking agent without anaesthesia. Using a plate scanner, this detection technique can be used to analyse thousands of individual Culicoides specimens per day at a reasonable cost. We marked and released an estimated 853 specimens of the Pulicaris group and 607 specimens of the Obsoletus group on a cattle farm in Denmark. An estimated 9,090 (8,918–9,260) Obsoletus group specimens and 14,272 (14,194–14,448) Pulicaris group specimens were captured in the surroundings and subsequently analysed. Two (0.3%) Obsoletus group specimens and 28 (4.6%) Pulicaris group specimens were recaptured. The two recaptured Obsoletus group specimens were caught at the release point on the night following release. Eight (29%) of the recaptured Pulicaris group specimens were caught at a pig farm 1,750 m upwind from the release point. Five of these were recaptured on the night following release and the three other were recaptured on the second night after release. This is the first time that movement of Culicoides vectors between farms in Europe has been directly quantified. The findings suggest an extensive and rapid exchange of disease vectors between farms. Rapid movement of vectors between neighboring farms may explain the the high rate of spatial spread of Schmallenberg and bluetongue virus (BTV) in northern Europe.

General information
Publication status: Published
Organisations: National Veterinary Institute, Section for Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Section for Immunology and Vaccinology
Contributors: Kirkeby, C., Bødker, R., Stockmarr, A., Lind, P., Heegaard, P. M. H.
Number of pages: 7
Pages: e61269
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: P L o S One
Volume: 8
Issue number: 4
ISSN (Print): 1932-6203
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.94 SJR 1.722 SNIP 1.134
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Electronic versions:
journal.pone.0061269 (1).pdf
Quantitative Analysis of Micro-Structure in Meat Emulsions from Grating-Based Multimodal X-Ray Tomography

Using novel X-ray techniques, based on grating-interferometry, new imaging modalities can be obtained simultaneously with absorption computed tomography (CT). These modalities, called phase contrast and dark field imaging, measure the electron density and the diffusion length of the sample. Enhanced contrast capabilities of this X-ray technique makes studies on materials with similar attenuation properties possible. In this paper the focus is set on processing grating-based X-ray tomograms of meat emulsions to quantitatively measure micro-structural changes due to heat treatment. The emulsion samples were imaged both in a raw and cooked state. Additionally, different fat types were used in the emulsions in order to compare micro-structural differences when either pork fat or sunflower oil was used. From the reconstructed tomograms the different ingredients in the emulsions were segmented using a multivariate segmentation method. From this, a quantitative analysis was performed between the different samples, determining properties such as percentage object volumes and cooking loss. Additionally, the porosity, degree of anisotropy and average structure thickness of the protein networks were determined. Analyzing the multivariate dataset instead of the single univariate absorption modality gave superior segmentation results. The quantitative analysis of the micro-structure gives insight to how both heat treatment, and the use of different lipid types, affect the final protein network.

Quantitative assessment of course evaluations

Student evaluation of teaching has been used in educational institutions around the world as a means of providing feedback on the quality of teaching. Nowadays, it is one of the most widespread tools used to inform teachers and administration about the instruction given in an institution.

The goal of the thesis is to develop efficient tools to analyze the data from student evaluations of teaching and courses at the Technical University of Denmark.

The thesis explores both classical and modern methods of multivariate statistical data analysis to address different issues of student evaluation of teaching (SET). In particular, the thesis includes results on the investigation of the association between the student evaluations of the course and the student evaluations of the teachers, the investigation of the effects of the mid-term evaluation on the end-of-term evaluations and the investigation of the student non-response on SETs. In order to utilize information from open-ended qualitative student answers, text-mining methods were applied in order to extract points of students praise and complaints.

The methods proposed contribute to the knowledge about student evaluation at the Technical University of Denmark. The
results provided some new information that will help teachers and university managers to better understand results of course evaluations.

Mid-term course evaluation was found to be able to capture both types of course issues: issues that can be addressed during the semester and also issues that can only be addressed at the next semester. Therefore, it seems to be preferable to conduct general mid-term evaluations instead of end-of-term evaluation, so the current course students can benefit. Additionally, it might be beneficial to conduct a short end-of-term evaluation with very limited number of questions that focus on general course issues after the final exams in order to obtain student feedback on the entire teaching and learning process, including the alignment of assessment of students’ learning with course objectives and teaching activities.

Student-specific and course-specific characteristics was found to be related with whether students participate in SETs and with how students evaluate courses and teachers. The DTU administrations should be aware that high achievers are more likely to participate in course evaluation survey and are more likely to give higher scores to courses. Students diversity on the course should be taken into account while making comparisons of evaluation results between courses.

In the student written feedback was found be able to provide additional knowledge of student point of satisfaction or dissatisfaction. However, in order to build an automated tool that can help to extract patterns from student comments higher quality of the collected data is needed.

Randomized controlled trial of the effects of vitamin D–fortified milk and bread on serum 25-hydroxyvitamin D concentrations in families in Denmark during winter: the VitmaD study

Background: Vitamin D intakes are lower than dietary recommendations in most populations, and thus, a low vitamin D status is widespread, especially during winter.

Objective: We investigated the effects of increasing vitamin D intake to the recommended amount by fortification of milk and bread on serum 25-hydroxyvitamin D [25(OH)D] concentrations in families during winter in Denmark.

Design: The study was a randomized controlled trial in 782 children and adults (4–60 y old) recruited as 201 families. Families were randomly assigned to vitamin D–fortified or nonfortified milk and bread for 6 mo starting in September. The milk and bread replaced the participants’ usual consumptions of products.

Results: Median (IQR) vitamin D intakes (habitual diet plus fortified products) were 9.4 μg/d (6.5, 12.3 μg/d) and 2.2 μg/d (1.5, 3.0 μg/d) in fortification and control groups, respectively. Geometric mean (IQR) serum 25(OH)D concentrations decreased from 73.1 nmol/L (61.9, 88.5 nmol/L) to 67.6 nmol/L (56.2, 79.4 nmol/L) in the fortification group and from 71.1 nmol/L (61.2, 85.9 nmol/L) to 41.7 nmol/L (29.5, 58.9 nmol/L) in the control group (both P < 0.001). The final 25(OH)D concentration was significantly higher in the fortification group than in the control group (P < 0.001). By the end of the study, <1% of subjects in the fortification group and 25% of subjects in the control group had 25(OH)D concentrations <30 nmol/L and 16% and 65% of subjects, respectively, had 25(OH)D concentrations <50 nmol/L.

Conclusion: Vitamin D fortification of milk and bread reduces the decrease in serum 25(OH)D concentrations during winter and ensures 25(OH)D concentrations >50 nmol/L in children and adults in Denmark. This trial was registered at clinicaltrials.gov as NCT01184716.
Spatial abundance and clustering of Culicoides (Diptera: Ceratopogonidae) on a local scale

Background
Biting midges, Culicoides, of the Obsoletus group and the Pulicaris group have been involved in recent outbreaks of bluetongue virus and the former was also involved in the Schmallenberg virus outbreak in northern Europe.

Methods
For the first time, here we investigate the local abundance pattern of these two species groups in the field by intensive sampling with a grid of light traps on 16 catch nights. Neighboring trap catches can be spatially dependent on each other, hence we developed a conditional autoregressive (CAR) model framework to test a number of spatial and non-spatial covariates expected to affect Culicoides abundance.

Results
The distance to sheep penned in the corner of the study field significantly increased the abundance level up to 200 meters away from the sheep. Spatial clustering was found to be significant but could not be explained by any known factors, and cluster locations shifted between catch nights. No significant temporal autocorrelation was detected. CAR models for both species groups identified a significant positive impact of humidity and significant negative impacts of precipitation and wind turbulence. Temperature was also found to be significant with a peak at just below 16 degrees Celsius. Surprisingly, there was a significant positive impact of wind speed. The CAR model for the Pulicaris group also identified a significant attraction to the smaller groups of sheep placed in the field. Furthermore, a large number of spatial covariates which were incorrectly found to be significant in ordinary regression models were not significant in the CAR models. The 95% C.I. on the prediction estimates ranged from 20.4% to 304.8%, underlining the difficulties of predicting the abundance of Culicoides.

Conclusions
We found that significant spatial clusters of Culicoides moved around in a dynamic pattern varying between catch nights. This conforms with the modeling but was not explained by any of the tested covariates. The mean abundance within these clusters was up to 11 times higher for the Obsoletus group and 4 times higher for the Pulicaris group compared to the rest of the field.
Spatio-temporal abundance of Culicoides on a local scale

General information
Publication status: Published
Organisations: National Veterinary Institute, Section for Epidemiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Kirkeby, C., Bødker, R., Stockmarr, A., Lind, P.
Number of pages: 15
Publication date: 2013
Peer-reviewed: Yes
Event: Abstract from EDENext Annual Meeting, Barcelona, Spain.
Electronic versions:
Spatio_temporal_abundance.pdf

Bibliographical note
The presentation won 3rd prize (300 Euro) among 27 other contributions at the conference.

Spatio-temporal optimization of sampling for bluetongue vectors (Culicoides) near grazing livestock

BACKGROUND: Estimating the abundance of Culicoides using light traps is influenced by a large variation in abundance in time and place. This study investigates the optimal trapping strategy to estimate the abundance or presence/absence of Culicoides on a field with grazing animals. We used 45 light traps to sample specimens from the Culicoides obsoletus species complex on a 14 hectare field during 16 nights in 2009.

FINDINGS: The large number of traps and catch nights enabled us to simulate a series of samples consisting of different numbers of traps (1-15) on each night. We also varied the number of catch nights when simulating the sampling, and sampled with increasing minimum distances between traps. We used resampling to generate a distribution of different mean and median abundance in each sample. Finally, we used the hypergeometric distribution to estimate the probability of falsely detecting absence of vectors on the field. The variation in the estimated abundance decreased steeply when using up to six traps, and was less pronounced when using more traps, although no clear cutoff was found.

CONCLUSIONS: Despite spatial clustering in vector abundance, we found no effect of increasing the distance between traps. We found that 18 traps were generally required to reach 90% probability of a true positive catch when sampling just one night. But when sampling over two nights the same probability level was obtained with just three traps per night. The results are useful for the design of vector monitoring programmes on fields with grazing animals.

General information
Statistical Outlier Detection for Jury Based Grading Systems

This paper presents an algorithm that was developed to identify statistical outliers from the scores of grading jury members in a large project-based first year design course. The background and requirements for the outlier detection system are presented. The outlier detection algorithm and the follow-up procedures for score validation and appeals are described in detail. Finally, the impact of various elements of the outlier detection algorithm, their interactions, and the sensitivity of their numerical values are investigated. It is shown that the difference in the mean score produced by a grading jury before and after a suspected outlier is removed from the mean is the single most effective criterion for identifying potential outliers but that all of the criteria included in the algorithm have an effect on the outlier detection process.

General information
Publication status: Published
Organisations: Department of Mechanical Engineering, Manufacturing Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Valparaíso
Contributors: Thompson, M. K., Clemmensen, L. K. H., Rosas, H.
Number of pages: 13
Publication date: 2013

Host publication information
Title of host publication: 120th ASEE Annual Conference and Exhibition
Publisher: American Society for Engineering Education
Electronic versions:
Statistical_Outlier_Detection.pdf
URLs:
http://www.asee.org/public/conferences/20/papers/6423/download
Statistical Quality Assessment of Pre-fried Carrots Using Multispectral Imaging

Multispectral imaging is increasingly being used for quality assessment of food items due to its non-invasive benefits. In this paper, we investigate the use of multispectral images of pre-fried carrots, to detect changes over a period of 14 days. The idea is to distinguish changes in quality from spectral images of visible and NIR bands. High dimensional feature vectors were formed from all possible ratios of spectral bands in 9 different percentiles per piece of carrot. We propose to use a multiple hypothesis testing technique based on the Benjamini-Hachberg (BH) method to distinguish possible significant changes in features during the inspection days. Discrimination by the SVM classifier supported these results. Additionally, 2-sided t-tests on the predictions of the elastic-net regressions were carried out to compare our results with previous studies on fried carrots. The experimental results showed that the most significant changes occurred in day 2 and day 14.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Division of Industrial Food Research
Contributors: Sharifzadeh, S., Clemmensen, L. K. H., Løje, H., Ersbøll, B. K.
Pages: 620-629
Publication date: 2013

Host publication information
Title of host publication: Image Analysis : 18th Scandinavian Conference, SCIA 2013, Espoo, Finland, June 17-20, 2013.
Proceedings
Publisher: Springer
ISBN (Print): 978-3-642-38885-9
ISBN (Electronic): 978-3-642-38886-6
(Lecture Notes in Computer Science, Vol. 7944).
Keywords: Multispectral imaging, Multiple hypothesis testing, Segmentation, Food quality assessment, SVM classification, Elastic-net regression
DOIs: 10.1007/978-3-642-38886-6_58
Source: dtu
Source-ID: u::7661
The dietary effect of serving school meals based on the new Nordic diet – A randomised controlled trial in Danish children

Background and objectives:
The OPUS study is a school-based intervention study testing selected health effects of New Nordic Diet (NND). Children are served lunch and snacks based on NND. The hypothesis is that Danish school children eat a healthier diet when receiving NND school meals as compared with packed lunch brought from home. To investigate the effects on intake of selected macronutrients in Danish school children when served school meals based on NND compared with packed lunch.

Methods:
In a cluster-randomized controlled unblinded cross-over study children received school meals based on NND for 3 months and their usual packed lunch for 3 months. The daily intake of food and beverages was recorded 3 times during 7 consecutive days using a validated self-administered web-based dietary assessment software tool for children. Statistical analysis was performed by hierarchical mixed models.

Results:
834 children from 9 schools were included and 96%, 89% and 80% filled out the first, second and third dietary assessment sufficiently (4-7 days), respectively. The preliminary results showed that the effect of serving NND resulted in a reduction in fat E% (P<0.0001), total fat (P=0.0007) and saturated fat (P<0.0001) intake for the NND compared to packed lunch; and an increase in protein E% (P<0.0001), and a borderline significant increase in dietary fiber intake (P=0.0471). There was no effect for energy intake, carbohydrate E% and added sugar E% (P>0.05). Effects are adjusted for BMI, season and household education.

Conclusions:
Danish school children's dietary intake of total and saturated fat decreased, fat E% decreased and protein E% increased when eating NND lunch and snacks compared to packed lunch brought from home. The OPUS project (Optimal well-being, development and health for Danish children through a healthy New Nordic Diet) is supported by the Nordea Foundation.

General information
Publication status: Published
Organisations: National Food Institute, Division of Nutrition, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Copenhagen
Pages: 197-198
Publication date: 2013
Peer-reviewed: Yes

The Effect of Rubric Rating Scale on the Evaluation of Engineering Design Projects

This paper explores the impact of the rubric rating scale on the evaluation of projects from a first year engineering design course. A small experiment was conducted in which twenty-one experienced graders scored five technical posters using one of four rating scales. All rating scales tested produced excellent results in terms of inter-rater reliability and validity. However, there were significant differences in the performance of each of the scales. Based on the experiment’s results and past experience, we conclude that increasing the opportunities for raters to deduct points results in greater point deductions and lower overall scores. Increasing the granularity of the scale can reduce this effect. Rating scales that use letter grades are less reliable than other types of scale. Assigning weights to individual criteria can lead to problems with validity if the weights are improperly balanced. Thus, heavily weighted rubrics should be avoided if viable alternatives exist. Placing more responsibility for the final score on the grader instead of the rubric seems to increase the validity at the cost of rater satisfaction. Finally, rater discomfort can lead to intentional misuse of a rating scale. This, in turn, increases the need to perform outlier detection on the final scores. Based on these findings, we recommend rating scale rubrics that
use simple 3 or 4-point ordinal rating scales (augmented checks) for individual criteria and that assign numerical scores to groups of criteria.

**General information**

Publication status: Published
Organisations: Department of Mechanical Engineering, Manufacturing Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Thompson, M. K., Clemmensen, L. K. H., Ahn, B.
Pages: 1490–1502
Publication date: 2013
Peer-reviewed: Yes

**Publication information**

Journal: International Journal of Engineering Education
Volume: 29
Issue number: 6
ISSN (Print): 0949-149X
Ratings:
- BFI (2013): BFI-level 2
- Scopus rating (2013): CiteScore 0.67 SJR 1.323 SNIP 1.134
- Web of Science (2013): Impact factor 0.36
- ISI indexed (2013): ISI indexed yes
- Web of Science (2013): Indexed yes
Original language: English
Keywords: Engineering design, Evaluation, Rating scale, Rubric, Psychometrics
Electronic versions:
- Effect_of_Rubric_Rating_Scale.pdf
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

**The Gut Microbiotassay: a high-throughput qPCR approach combinable with next generation sequencing to study gut microbial diversity**

**Background**
The intestinal microbiota is a complex and diverse ecosystem that plays a significant role in maintaining the health and well-being of the mammalian host. During the last decade focus has increased on the importance of intestinal bacteria. Several molecular methods can be applied to describe the composition of the microbiota. This study used a new approach, the Gut Microbiotassay: an assembly of 24 primer sets targeting the main phyla and taxonomically related subgroups of the intestinal microbiota, to be used with the high-throughput qPCR chip ‘Access Array 48.48’, AA48.48, (Fluidigm®) followed by next generation sequencing. Primers were designed if necessary and all primer sets were screened against DNA extracted from pure cultures of 15 representative bacterial species. Subsequently the setup was tested on DNA extracted from small and large intestinal content from piglets with and without diarrhoea. The PCR amplicons from the 2304 reaction chambers were harvested from the AA48.48, purified, and sequenced using 454-technology.

**Results**
The Gut Microbiotassay was able to detect significant differences in the quantity and composition of the microbiota according to gut sections and diarrhoeic status. 454-sequencing confirmed the specificity of the primer sets. Diarrhoea was associated with a reduced number of members from the genus Streptococcus, and in particular S. alactolyticus.

**Conclusion**
The Gut Microbiotassay provides fast and affordable high-throughput quantification of the bacterial composition in many samples and enables further descriptive taxonomic information if combined with 454-sequencing.

**General information**

Publication status: Published
Organisations: National Veterinary Institute, Section for Bacteriology, Pathology and Parasitology, Section for Immunology and Vaccinology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Danish Genome Institute
Number of pages: 14
Publication date: 2013
Peer-reviewed: Yes

**Publication information**

Journal: B M C Genomics
The range of attraction for light traps catching Culicoides biting midges

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Cryptology, Dynamical Systems, National Veterinary Institute, Section for Epidemiology, Statistics and Data Analysis
Contributors: Græsbøll, K., Kirkeby, C., Bødker, R., Stockmarr, A., Christiansen, L. E., Lind, P.
Number of pages: 1
Publication date: 2013
Peer-reviewed: Yes
Event: Poster session presented at EDENext Annual Meeting, Barcelona, Spain.

Electronic versions:
prod11363846960013.Poster_EDENext_3C2.pdf
Source: dtu
Source-ID: u::7293
Research output: Contribution to conference › Poster – Annual report year: 2013 › Research › peer-review

The range of attraction for light traps catching Culicoides biting midges

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Cryptology, Dynamical Systems, National Veterinary Institute, Section for Epidemiology, Statistics and Data Analysis
Contributors: Græsbøll, K., Kirkeby, C., Bødker, R., Stockmarr, A., Christiansen, L. E., Lind, P.
Number of pages: 1
Publication date: 2013
Peer-reviewed: Yes

Electronic versions:
prod21377679734408.GEOVET13.pdf
Source: dtu
Source-ID: u::8524
Research output: Contribution to conference › Poster – Annual report year: 2013 › Research › peer-review

The range of attraction for light traps catching Culicoides biting midges (Diptera: Ceratopogonidae)

Background
Culicoides are vectors of e.g. bluetongue virus and Schmallenberg virus in northern Europe. Light trapping is an important tool for detecting the presence and quantifying the abundance of vectors in the field. Until now, few studies have
investigated the range of attraction of light traps.

Methods
Here we test a previously described mathematical model (Model I) and two novel models for the attraction of vectors to light traps (Model II and III). In Model I, Culicoides fly to the nearest trap from within a fixed range of attraction. In Model II Culicoides fly towards areas with greater light intensity, and in Model III Culicoides evaluate light sources in the field of view and fly towards the strongest. Model II and III incorporated the directionally dependent light field created around light traps with fluorescent light tubes. All three models were fitted to light trap collections obtained from two novel experimental setups in the field where traps were placed in different configurations.

Results
Results showed that overlapping ranges of attraction of neighboring traps extended the shared range of attraction. Model I did not fit data from any of the experimental setups. Model II could only fit data from one of the setups, while Model III fitted data from both experimental setups.

Conclusions
The model with the best fit, Model III, indicates that Culicoides continuously evaluate the light source direction and intensity. The maximum range of attraction of a single 4W CDC light trap was estimated to be approximately 15.25 meters. The attraction towards light traps is different from the attraction to host animals and thus light trap catches may not represent the vector species and numbers attracted to hosts.

General information
Publication status: Published
Organisations: National Veterinary Institute, Section for Epidemiology, Department of Applied Mathematics and Computer Science, Cryptology, Dynamical Systems, Statistics and Data Analysis
Contributors: Kirkeby, C., Græsbøll, K., Stockmarr, A., Christiansen, L. E., Bødker, R.
Number of pages: 19
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Parasites & Vectors
Volume: 6
Issue number: 1
Article number: 67
ISSN (Print): 1756-3305
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.52 SJR 1.496 SNIP 1.453
Web of Science (2013): Impact factor 3.251
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Keywords: Culicoides, Range of attraction, Vector abundance, Light traps, Vector monitoring
Electronic versions:
1756_3305_6_67.pdf
DOIs: 10.1186/1756-3305-6-67

Bibliographical note
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Source: dtu
Source-ID: u::7314
Research output: Contribution to journal › Journal article – Annual report year: 2013 › Research › peer-review

The Statistical Value Chain - a Benchmarking Checklist for Decision Makers to Evaluate Decision Support Seen from a Statistical Point-Of-View
When decisions are made, by decision makers (DMs) in private and public organizations the DMs are supported by analysts (ANs) who provide decision support to the DM. Therefore, the quality of decision support provided by the AN directly affects the quality of a DM’s decision. At present, many quantitative methods exist for evaluating uncertainty—for example, Monte Carlo simulation—and such methods work very well when the AN is in full control of the data collection and model-building processes. In many cases, however, the AN is not in control of these processes. In this article we develop a simple method that a DM can employ in order to evaluate the process of decision support from a statistical
point-of-view. We call this approach the "Statistical Value Chain" (SVC): a consecutive benchmarking checklist with eight steps that can be used to evaluate decision support seen from a statistical point-of-view.

**General information**
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, Energy Systems Analysis, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, University of Southampton, University of Copenhagen
Pages: 71-83
Publication date: 2013
Peer-reviewed: Yes

**Publication information**
Journal: International Journal of Decision Sciences
Volume: 4
Issue number: 2
ISSN (Print): 2229-5879
Original language: English
Keywords: Statistical Value Chain, Decision Theory, Benchmarking Checklist, Decision Makers, Evaluate, Decision Support, Statistics, Data Analysis, Uncertainty, Quality
Electronic versions:
THE_STATISTICAL_VALUE_CHAIN.pdf
Source: PublicationPreSubmission
Source-ID: 101105312
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

**Using multilevel systematic sampling to study apple fruit (Malus domestica Borkh.) quality and its variability at the orchard scale**
We report on the performance of a novel sampling method for determining fruit quality variability and yield from an orchard, which focus on its applicability for the fruit industry. We used the 'fractionator' tree sampling method to investigate the quality variability of a small, representative sample of ‘Granny Smith’ (Malus x domestica cv. ‘Granny Smith’) apples obtained from a 17 ha orchard based on a final sample of 74 fruit. Estimates of fruit marketable yield and fruit size distribution agreed well with packing house records. The estimated marketable yield was 356.6 ± 89.2 t compared to 374.9 t of fruit packed for export. Distributions of starch (S), soluble solids content (SSC) and flesh firmness (F) were also estimated from the sample. The distribution of starch (S) and fruit mass (M) showed high variability (CVS = SD/mean = 0.32 and CVM = 0.23), whereas SSC and flesh firmness showed moderate variability (CVSSC = 0.11 and CVF = 0.10). The average within-tree variabilities were estimated as CVM = 0.04, CVSSC = 0.10, CVS = 0.15 and CVF = 0.07. Between-tree variabilities were similar to the within-tree variabilities, except for starch (CVM = 0.04, CVSSC = 0.13, CVS = 0.29 and CVF = 0.09). From the quality characteristics studied only fruit mass could be significantly related to position of the fruit in the canopy, represented by height of the fruit above ground, the fruit position along the branch and position relative to the tree row orientation in the orchard. Variations in starch, SSC and flesh firmness could not be explained by position of the fruit in the canopy. The methods used in this paper are proposed as tools for studies aimed at understanding sources of quality variability as well as for management purposes. Further research is needed to determine recommended sample sizes to accurately describe the distribution of various quality variables of apples at the orchard scale.

**General information**
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Dayenú, University of Copenhagen
Contributors: Martínez Vega, M. V., Wulfsohn, D., Clemmensen, L. K. H., Toldam-Andersen, T. B.
Pages: 58-64
Publication date: 2013
Peer-reviewed: Yes

**Publication information**
Journal: Scientia Horticulturae
Volume: 161
ISSN (Print): 0304-4238
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 2 SJR 0.767 SNIP 1.447
Web of Science (2013): Impact factor 1.504
Using Multispectral Imaging for Spoilage Detection of Pork Meat

The quality of stored minced pork meat was monitored using a rapid multispectral imaging device to quantify the degree of spoilage. Bacterial counts of a total of 155 meat samples stored for up to 580 h have been measured using conventional laboratory methods. Meat samples were maintained under two different storage conditions: aerobic and modified atmosphere packages as well as under different temperatures. Besides bacterial counts, a sensory panel has judged the spoilage degree of all meat samples into one of three classes. Results showed that the multispectral imaging device was able to classify 76.13% of the meat samples correctly according to the defined sensory scale. Furthermore, the multispectral camera device was able to predict total viable counts with a standard error of prediction of 7.47%. It is concluded that there is a good possibility that a setup like the one investigated will be successful for the detection of spoilage degree in minced pork meat.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Image Analysis & Computer Graphics, Agricultural University of Athens, Institute of Technology of Agricultural Products
Contributors: Dissing, B. S., Papadopoulou, O. S., Tassou, C., Ersbøll, B. K., Carstensen, J. M., Panagou, E., Nychas, G.
Pages: 2268-2279
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Food and Bioprocess Technology
Volume: 6
Issue number: 9
ISSN (Print): 1935-5130
Ratings:
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 2.97 SJR 1.205 SNIP 1.679
Web of Science (2013): Impact factor 3.126
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
Original language: English
Keywords: Multispectral imaging, Meat spoilage, Chemometrics, Computational biology, Meat quality, Non-invasive methods, Converging technologies, Predictive modelling
DOIs:
10.1007/s11947-012-0886-6
Source: dtu
Source-ID: u::7981
Research output: Contribution to journal › Journal article – Annual report year: 2014 › Research › peer-review

Using X-ray imaging to study thermal-induced changes in food

The food quality in many food processes relies greatly on the structural changes that take place during heating or freezing of the food product. So far, it has only been possible to study these changes indirectly but recent new X-ray imaging modalities allow for direct visualization. We present preliminary results of structural changes by heating of bovine meat and freezing of berries inspected with X-ray phase-contrast and dark-field imaging.

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, University of Copenhagen
Contributors: Nielsen, M. S., Miklos, R., Lametsch, R., Einarsdottir, H., Feidenhans'l, R.
Number of pages: 1
Vitamin D status among families in Denmark: Baseline data from the vitmad study

Background and objectives:
The beneficial effect of vitamin D in bone health is acknowledged and the vitamin has also been associated with several chronic diseases. It is therefore relevant to determine the prevalence of vitamin D insufficiency in different groups, and vitamin D statuses within families have not been studied previously. The objective of the present study was to evaluate serum 25-hydroxyvitamin D (25(OH)D) concentrations among families in Denmark (56°N) after seasonal UVB peak and to ascertain determining factors.

Methods:
Cross-sectional study with 755 children and adults (4-60 y) recruited as families in the VitmaD study. Blood samples were collected in September-October 2010, and vitamin D status was measured as serum 25(OH)D concentration by LC-MS/MS. Vitamin D intake and life style factors were assessed in self-administered questionnaires. Determinants of vitamin D status were identified in a linear mixed model with family as a random variable.

Results:
Mean (±SD) serum 25(OH)D concentration was 75 ± 20 nmol/l (range 9-162 nmol/l) and only 10 % had 25(OH) D <50 nmol/l. Determinants of serum 25(OH)D were age (p=0.036), BMI class (p=0.001), multi vitamin use (p=0.033), sun behaviour (p=0.005), outdoor stay (p=0.033), sun vacation (p<0.001), and physical activity (p=0.040). Gender (p=0.692) and vitamin D intake (p=0.238) were not associated to serum 25(OH)D.

Conclusions:
The prevalence of vitamin D insufficiency among families in Denmark was low after seasonal UVB peak. Sun vacation was the strongest determinant for vitamin D status at this time of the year.
consumers always rely on their own judgment in such matters, or delegate the final judgment to experts of some sort? The present experimental study addressed these issues in combination by testing the limits for consumers’ acceptance of three different name–product combinations when exposed to taste samples alone (sensory product attributes), taste samples in combination with ingredients lists and nutrition facts (adding factual information), and both, in combination with authoritative definitions (adding experts’ final judgments). The examples were modelled around authentic cases from the Danish food market which have been subject to vast legal as well as public concern. The results provide new insights into the socio-cognitive dynamics behind consumers’ acceptance or rejection of specific name–product combinations and new leads for supporting the fairness of food naming practices with a view also to the product type, the stage it has reached in its life-cycle, and its degree of familiarity on the market.

Dismantling the Mantel tests
1. The simple and partial Mantel tests are routinely used in many areas of evolutionary biology to assess the significance of the association between two or more matrices of distances relative to the same pairs of individuals or demes. Partial Mantel tests rather than simple Mantel tests are widely used to assess the relationship between two variables displaying some form of structure.

2. We show that contrary to a widely shared belief, partial Mantel tests are not valid in this case, and their bias remains
close to that of the simple Mantel test.

3. We confirm that strong biases are expected under a sampling design and spatial correlation parameter drawn from an actual study.

4. The Mantel tests should not be used in case autocorrelation is suspected in both variables compared under the null hypothesis. We outline alternative strategies. The R code used for our computer simulations is distributed as supporting material.

**General information**
**Publication status:** Published
**Organisations:** Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Université de Montpellier
**Contributors:** Guillot, G., Rousset, F.
**Pages:** 336–344
**Publication date:** 2012
**Peer-reviewed:** Yes

**Publication information**
**Journal:** Methods in Ecology and Evolution
**Volume:** 4
**Issue number:** 4
**ISSN (Print):** 2041-210X
**Ratings:**
- Scopus rating (2012): CiteScore 3.56 SJR 2.119 SNIP 1.651
- Web of Science (2012): Impact factor 5.924
- ISI indexed (2012): ISI indexed no
- Web of Science (2012): Indexed yes
**Original language:** English
**Keywords:** Landscape ecology, Landscape genetics, Phylogeography, Geographic epidemiology, Spatial structure, Isolation by distance, Isolation by resistance, Autocorrelation, Type I error, Loa loa
**Electronic versions:** 2F250d01.pdf
**DOIs:** 10.1111/2041-210x.12018
**Source:** dtu
**Source-ID:** n::oai:DTIC-ART:arxiv/372849633::20949
**Research output:** Contribution to journal › Journal article – Annual report year: 2012 › Research › peer-review

**Fitting a distribution to microbial counts: making sense of zeros**
Non-detects or left-censored results are inherent to the traditional methods of microbial enumeration in foods. Typically, a low concentration of microorganisms in a food unit goes undetected in plate counts or most probable number (MPN) counts, and produces “artificial zeros”. However, these “artificial zeros” are only a share of the total number of zero counts resulting from a sample, as their number adds up to the number of “true zeros” resulting from uncontaminated units. In the process of fitting a probability distribution to microbial counts, “artificial” and “true” zeros are usually undifferentiated. This practice may lead to errors in the estimation of the parameters for the distribution of microbial concentrations, most specifically to the underestimation of the mean and overestimation of the variance. Distributions of microbial counts are often used as input in quantitative microbial risk assessment; therefore it is possible that errors related to these distributions have an impact in terms of food safety, if an influence on the estimated risk is observed.

In this study, we developed a method to estimate both the parameters of a lognormal distribution of microbial concentrations (mean and standard deviation) and the prevalence of contaminated food units (one minus the proportion of “true zeros”) from a set of microbial counts.

By running the model with in silico generated concentration and count data, we could evaluate the performance of this method in terms of estimation of the three different parameters. In principle, the higher the proportion of zeros in a dataset, the higher the error in the estimation will be, and a lower prevalence contributes to a higher proportion of “true zeros” in microbial counts. Therefore, we also investigated the effect of the prevalence on the estimation of the distribution parameters mean and standard deviation by running the same model for different prevalence scenarios.

**General information**
**Publication status:** Published
**Organisations:** National Food Institute, Division of Epidemiology and Microbial Genomics, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
**Contributors:** Ribeiro Duarte, A. S., Stockmarr, A., Nauta, M.
**Number of pages:** 1
"Remotion" Total Wrist Arthroplasty: Preliminary Results of a Prospective International Multicenter Study of 215 Cases

This study reports the current results of an international multicenter study of one last generation total wrist arthroplasty (TWA) ("Remotion," Small Bone Innovation, Morristown, PA). The two first authors (G.H. and M.B.) built a Web-based prospective database including clinical and radiological preoperative and postoperative reports of "Remotion" TWA at regular intervals. The cases of 7 centers with more than 15 inclusions were considered for this article. A total of 215 wrists were included. In the rheumatoid arthritis (RA; 129 wrists) and nonrheumatoid arthritis (non-RA; 86 wrists) groups, there were respectively 5 and 6% complications requiring implant revision with a survival rate of 96 and 92%, respectively, at an average follow-up of 4 years. Within the whole series, only one dislocation was observed in one non-RA wrist. A total of 112 wrists (75 rheumatoid and 37 nonrheumatoid) had more than 2 years of follow-up (minimum: 2 years, maximum: 8 years). In rheumatoid and non-RA group, visual analog scale (VAS) pain score improved by 48 and 54 points, respectively, and QuickDASH score improved by 20 and 21 points, respectively, with no statistical differences. Average postoperative arc of wrist flexion-extension was 58 degrees in rheumatoid wrists (loss of 1 degree) compared with 63 degrees in non-RA wrists (loss of 9 degrees) with no statistical differences. Grip strength improved respectively by 40 and 19% in rheumatoid and non-RA groups (p = 0.033). Implant loosening was observed in 4% of the rheumatoid wrists and 3% of the non-RA wrists with no statistical differences. A Web-based TWA international registry was presented. Our results suggest that the use of the "Remotion" TWA is feasible in the midterm both for rheumatoid and non-RA patients. This is a significant improvement compared with the previous generation TWA. The level of evidence for this study is IV.
Susse Winther interviews Frank Hedlund in: [Når katastrofen rammer] /When disaster strikes: Hurricane Katrina
Striking a balance between, on the one hand, the need for centralization of the disaster emergency response, and on the other hand, the ability to achieve a flexible and improvised response, that uses scarce resources in an optimal manner – a flexibility which by nature requires decentralization.

The crustal uplift determined at the Jakobshavn glacier (West Greenland) using ATM and GPS data
The Greenland ice sheet has experienced record melting in recent years. In order to estimate the ice loss we can make use of the earth's natural elasticity to weigh the ice. Ice bends down the bedrock so when the ice melts away, the bedrock rises measurably in response. Throughout this abstract we present both a predicted and observed crustal uplift for the Jakobshavn glacier using ATM data (Airborne Topographic Mapper) from NASA ATM flights during 1997, 2005 and 2010 supplemented with data provided from continuous Global Positioning System (GPS), measurements made on bedrock between 2005-2010.

In order to compute the crustal uplift in response to the ice mass loss of the Jakobshavn area from the GPS stations, the convolution of the gridded thinning rates has been computed with the vertical-displacement Green's function as described in [1].

Several manipulations of data were required in order to achieve a good prediction of the crustal uplift. In this sense the programs Matlab and Geogrid-Gravsoft were used along with some Fortran executable files. Furthermore, the GPS data which presents the difference in uplift is provided processed as a difference of data from the permanent GPS stations KAGA, ILUL and QEQE relative to the AASI station (Figure 1). Also, in order to compare the predicted uplift from ATM data with the observed uplift from GPS data the post-glacial rebound (PGR) rates have been subtracted.

The results obtained for the predicted crustal uplift for KAAS is 11.62 mm/yr while the observed value was 16.32 mm/yr, for ILAS 1.74 mm/yr and 1.53 mm/yr, for QEQE-0.189 mm/yr and 1.15 mm/yr. That being an difference of 4.701 mm/yr is found for KAAS, 0.21 mm/yr for ILAS and 1.339 mm/yr for QEQE. The uncertainties associated both with the ATM and GPS results are 0.8 mm/yr for ATM and 0.5 mm/yr for GPS. The total ice mass loss in km³ of water predicted from the ATM data concern the Jakobshavn area is -88.815 between 2010 and 2005 and -83.599 between 2005 and 1997.

It seems fair to state that this differences, between the predicted and observed rates, may also be due to the fact that not all the errors have been taken into account when computing the observed results and also due to the fact that, perhaps, ice is melting in Greenland much faster than predicted.
Deliverable 3.1: Sampling protocols for inorganic ENP from at least 3 matrices (meat, soup, olive oil)

General information
Publication status: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Food Institute, Research group for Nano-Bio Science
Contributors: Ersbøll, B. K., Dahl, V. A., Löschner, K., Larsen, E. H.
Number of pages: 25
Publication date: 2011

Publication information
Publisher: NanoSafety Cluster
Original language: English
Electronic versions:
NanoLyse_D3_1.pdf
Source: PublicationPreSubmission
Source-ID: 149478710
Research output: Book/Report › Report – Annual report year: 2011 › Commissioned

Optimal interventions to control campylobacter in broilers in Denmark
In a multi disciplinary project we have evaluated interventions against Campylobacter in the broiler production chain. Taking into account risk reduction, costs, practicability and public acceptance of decontamination, it was concluded that at present the optimal control measure for the Danish situation is screening broiler houses with fly nets.

General information
Publication status: Published
Organisations: National Food Institute, Division of Epidemiology and Microbial Genomics, Division of Microbiology and Risk Assessment, Statistics and Data Analysis, Technical University of Denmark, University of Copenhagen
Number of pages: 1
Publication date: 2011
Peer-reviewed: Yes
Event: Poster session presented at 16th International Workshop on Campylobacter, Helicobacter, and Related Organisms, Vancouver, Canada.
Source: dtu
Source-ID: u::8257
Research output: Contribution to conference › Poster – Annual report year: 2011 › Research › peer-review

Persistence of Low-Pathogenic Avian Influenza H5N7 and H7N1 Subtypes in House Flies (Diptera: Muscidae)
Avian influenza caused by avian influenza virus (AIV) has a negative impact on poultry production. Low-pathogenic AIV (LPAIV) is naturally present in wild birds, and the introduction of the virus into domestic poultry is assumed to occur through contact with wild birds and by human activity, including the movement of live and dead poultry, and fomites such as clothing and vehicles. At present, the possible role of insects in the spread of AIV is dubious. The objective of the present work was to investigate the potential transmission of LPAIV by persistence of the virus in the alimentary tract of house flies, Musca domestica L. (Diptera: Muscidae). Flies were fed three virus concentrations of two AIV strains and then incubated at different temperatures for up to 24 h. The persistence of the two virus strains in the flies declined with increasing incubation temperatures and incubation periods. Similarly, increased virus uptake by the flies increased the persistence of virus. Persistence of infective AIV in flies differed significantly between the two virus strains. The laboratory experiments of the present study indicate that the house fly can be a potential carrier of AIV.

General information
Publication status: Published
Organisations: National Food Institute, Division of Food Microbiology, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, National Veterinary Institute, Aarhus University Hospital, Aarhus University
Contributors: Nielsen, A. A., Skovgård, H., Stockmarr, A., Handberg, K., Jørgensen, P. H.
Pages: 608-614
Publication date: 2011
Peer-reviewed: Yes

Publication information
Journal: Journal of Medical Entomology
Volume: 48
A Generalization of the Alias Matrix

The investigation of aliases or biases is important for the interpretation of the results from factorial experiments. For two-level fractional factorials this can be facilitated through their group structure. For more general arrays the alias matrix can be used. This tool is traditionally based on the assumption that the error structure is that associated with ordinary least squares. For situations where that is not the case, we provide in this article a generalization of the alias matrix applicable under the generalized least squares assumptions. We also show that for the special case of split plot error structure, the generalized alias matrix simplifies to the ordinary alias matrix.
Split-plot fractional designs: Is minimum aberration enough?
Split-plot experiments are commonly used in industry for product and process improvement. Recent articles on designing split-plot experiments concentrate on minimum aberration as the design criterion. Minimum aberration has been criticized as a design criterion for completely randomized fractional factorial design and alternative criteria, such as the maximum number of clear two-factor interactions, are suggested (Wu and Hamada (2000)). The need for alternatives to minimum aberration is even more acute for split-plot designs. In a standard split-plot design, there are several types of two-factor interactions, not all of them equally interesting. However, minimum aberration is not designed to distinguish among the different types of two-factor interactions. It should be noted that this criticism is valid not only for the minimum aberration but also for any other design criteria originally proposed for completely randomized designs. Consequently, we provide a modified version of the maximum number of clear two-factor interactions design criterion to be used for split-plot designs.

Correlating phospholipid fatty acids (PLFA) in a landfill leachate polluted aquifer with biogeochemical factors by multivariate statistical methods
Different multivariate statistical analyses were applied to phospholipid fatty acids representing the biomass composition and to different biogeochemical parameters measured in 37 samples from a landfill contaminated aquifer at Grindsted Landfill (Denmark). Principal component analysis and correspondence analysis were used to identify groups of samples showing similar patterns with respect to biogeochemical variables and phospholipid fatty acid composition. The principal component analysis revealed that for the biogeochemical parameters the first principal component was linked to the pollution effect and to redox processes and the second principal component described the geological and geochemical features of the samples. Dependent on the data transformation of the phospholipid fatty acid profiles in either absolute concentrations (logarithm transformed) or in mol% of total phospholipid fatty acids, different groups of samples and outliers were revealed by the principal component analysis. The principal component analysis on data in absolute concentrations revealed that many phospholipid fatty acids reflected the pollution effect on the biomass composition. In contrast, the phospholipid fatty acids in mol% divided the samples into one group of the more polluted samples and another with the nearly unpolluted samples. The important phospholipid fatty acids for this grouping were mainly a few of the normal saturated phospholipid fatty acids (10:0, 16:0 and 18:0). Discriminant analysis was used to allocate samples of phospholipid fatty acids into predefined classes. A large percentages of samples were classified correctly when discriminating samples into groups of dissolved organic carbon and specific conductivity, indicating that the biomass is highly influenced by the pollution. In contrast, the discriminant analysis revealed that on the basis of the profiles of phospholipid fatty acids no good discrimination between samples showing dominant sulfate reduction and dominant iron reduction could be made, nor between samples showing dominant nitrate reduction and aerobic respiration. Partial least square analysis related the phospholipid fatty acids data to the biogeochemical parameters assuming linear relationships. After selection of the optimal phospholipid fatty acid combination by genetic algorithms, good partial least squares models with low prediction errors were gained primarily between the biogeochemical parameters describing total contents of carbon, pH and chloride. The models predicting specific activity in terms of, e.g., sulfate reduction activity in a sample had
relatively higher prediction errors and low correlation coefficients. This indicates that the phospholipid fatty acid profiles from complex habitats have limited value for identifying more specific microbial populations.

General information
Publication status: Published
Organisations: Department of Environmental Science and Engineering, Department of Environmental Engineering, Urban Water Engineering, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Contributors: Ludvigsen, L., Albrechtsen, H., Rootzén, H., Christensen, T. H.
Pages: 447-460
Publication date: 1997
Peer-reviewed: Yes

Publication information
Journal: F E M S Microbiology Reviews
Volume: 20
Issue number: 3-4
ISSN (Print): 0168-6445
Original language: English
DOIs: 10.1111/j.1574-6976.1997.tb00329.x
Source: orbit
Source-ID: 169668
Research output: Contribution to journal › Journal article – Annual report year: 1997 › Research › peer-review