In geosciences, as well as in astrophysics, direct observations of a studied physical system and objects are not always accessible. Instead, indirect observations have to be used in order to obtain information about the unknown system, which leads to an inverse problem. Such geoscientific inverse problems face the challenge of determining a set of unknown parameters based on a set of indirect observations of the subsurface. In a traditional least-squares formulation of the solution to an inverse problem, a subjectively chosen regularization parameter is used to obtain a unique solution to this problem, which leads to a smooth solution with no geological realism. Moreover, such an optimization-based framework does not allow introducing realistic geological prior information (due to a vectorial normed space structure). This thesis focuses on a more sophisticated approach based on a probabilistic formulation of the solution to the inverse problem. This formulation, different sources of information about the subsurface can be weighted with regard to their relative quality and reliability (i.e., uncertainties) using probability distributions and subsequently integrated into a posterior probability distribution over the model parameters. The different sources of information are provided in form of a set of observed data, uncertainties related to the data, and geological prior information, which is established from, e.g., expert knowledge and old data sets. The prior information, when being informative and realistic, has a regulating effect on the solution to the inverse problem as geological and geophysical information are orthogonal in some ways, which allows reducing the underdetermination of the inverse problem. At the same time, such prior information also reduces the effective dimension of the inverse problem, which may considerably reduce the computationally cost related to such problems. Moreover, the probabilistic formulation of the inverse problem allows the use of geologically more realistic prior information that leads to solutions to the inverse problem with a higher degree of geological realism. Finally, the probabilistic formulation provides the means of analyzing uncertainties and potential approximations. Approximations can be used for risk assessments in relation to, e.g., reservoir characterization and forecasting. Prior models rely on information from old data sets or expert knowledge in form of, e.g., training images that express structural, lithological, or textural features. Statistics obtained from these types of observations will be referred to as sample models. Geostatistical sampling algorithms use a sample model as input and produce multiple realizations of the model parameters that, to some degree, honor this information. Such algorithms can be used to define the prior information for probabilistic inverse problems. In this way, very informative and geologically more realistic prior information can be provided. This thesis provides an overview of the scientific developments within the field of probabilistic inverse problem theory and geostatistics, with emphasis on the combination of these scientific disciplines. In particular, the focus will be on consistent probabilistic formulations of this problem, which means that a correct weighting of the different sources of information is obeyed such that no unknown assumptions and biases influence the solution to the inverse problem. This involves a definition of the probabilistically formulated inverse problem, a discussion about how prior models can be established based on statistical information from sample models, and an analysis of geostatistical algorithms in order to understand the implicit assumptions made by such “black box” algorithms. A description of the posterior distribution can be obtained by drawing a representative sample from this distribution. Methodologies to be used for this purpose are presented. An example of sampling the posterior probability distribution of a computationally hard full waveform inverse problem using prior information based on multiple-point statistics, obtained from a training image, is demonstrated. For some computationally challenging inverse problems, a sample from the posterior distribution might still be too laborious to be obtained. Instead, a set of model parameters with (near) maximum posterior probability can be obtained. In order to do this, a closed form mathematical formulation of the prior probability distribution has to be established, such that the posterior probability distribution can be evaluated. Different solutions to this problem are presented and discussed. The prior probability distribution that is sampled by geostatistical sampling algorithms is typically unknown or sometime only a part of or an approximation to the distribution is known. This thesis provides an analysis and a discussion of how these prior probability distributions can be established, such that it is consistent with information provided by a known sample model. It is described how assumptions about the distribution, in addition to the information provided by the sample model, have to be made in order to end up with a unique solution to this problem. It is shown that these sampling algorithms typically provide samples from a prior probability distribution that is not consistent with the sample model. However, examples of consistent algorithms are also provided. A likelihood function is part of the probabilistic formulation of the inverse problem. This function is based on an uncertainty model that describes the uncertainties related to the observed data. In a similar way, a formulation of the prior probability distribution that takes into account uncertainties related to the sample model statistics is formulated. Prior models that are consistent with the statistics from a training image do not necessarily reproduce realizations with the same spatial patterns as seen in the training image because the local Markov properties that is satisfied in this way does not lead to a global reproduction of the pattern distribution from the training image. A prior probability distribution, with realizations that resemble the patterns as seen in the training image, is described and an efficient sampling algorithm that samples this distribution is provided. Moreover, an example of using this prior model for an inverse problem is demonstrated. The theoretical forward problem that describes the relation between data and model parameters is often associated with some degree of approximation. This approximation may have a great impact on the solution to the inverse problem because such approximate calculations of the data have an impact similar to observation uncertainties. We refer to the effect of these approximations as modeling errors. Examples that show how the modeling error is estimated are provided. Moreover, it is shown how these effects can be taken into account in the formulation of the posterior probability distribution. Common to the methods and strategies presented in this thesis is that they strive for a solution to the inverse problem that is consistent with the available information and to a less degree based on unconscious or subjective choices and implicit assumptions. Future studies related to theoretical developments of these strategies have to be provided. Moreover, applications of these strategies will reveal the practical implications of these consistent formulations. This will in particular be of great importance when it comes to assessments related cases of high risk such as human health or resources of high
economical potentials.

**General information**
State: Published
Organisations: Center for Energy Resources Engineering, National Space Institute, Mathematical and Computational Geoscience
Contributors: Cordua, K. S., Mosegaard, K.
Number of pages: 306
Publication date: 2013

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: DTU Space
Original language: English
Electronic versions:
Cordua 2013 - PhD thesis 17x24 Compressed.pdf
Source: dtu
Source-ID: u::10783
Research output: Research - peer-review → Book – Annual report year: 2014