Spatio-temporal precipitation climatology over complex terrain using a censored additive regression model

Flexible spatio-temporal models are widely used to create reliable and accurate estimates for precipitation climatologies. Most models are based on square root transformed monthly or annual means, where a normal distribution seems to be appropriate. This assumption becomes invalid on a daily time scale as the observations involve large fractions of zero observations and are limited to non-negative values. We develop a novel spatio-temporal model to estimate the full climatological distribution of precipitation on a daily time scale over complex terrain using a left-censored normal distribution. The results demonstrate that the new method is able to account for the non-normal distribution and the large fraction of zero observations. The new climatology provides the full climatological distribution on a very high spatial and temporal resolution, and is competitive with, or even outperforms existing methods, even for arbitrary locations.
Spatial ensemble post-processing with standardized anomalies

To post-process ensemble predictions for a particular location, statistical methods are often used, especially in complex terrain such as the Alps. When expanded to several stations, the post-processing has to be repeated at every station individually, thus losing information about spatial coherence and increasing computational cost. Therefore, the ensemble post-processing is modified and applied simultaneously at multiple locations. We transform observations and predictions to standardized anomalies. Seasonal and site-specific characteristics are eliminated by subtracting a climatological mean and dividing by the climatological standard deviation from both observations and numerical forecasts. This method allows us to forecast even at locations where no observations are available. The skill of these forecasts is comparable to forecasts post-processed individually at every station and is even better on average.
Ensemble postprocessing of daily precipitation sums over complex terrain using censored high-resolution standardized anomalies

Probabilistic forecasts provided by numerical ensemble prediction systems have systematic errors and are typically underdispersive. This is especially true over complex topography with extensive terrain-induced small-scale effects, which cannot be resolved by the ensemble system. To alleviate these errors, statistical postprocessing methods are often applied to calibrate the forecasts. This article presents a new full-distributional spatial postprocessing method for daily
precipitation sums based on the standardized anomaly model output statistics (SAMOS) approach. Observations and forecasts are transformed into standardized anomalies by subtracting the long-term climatological mean and dividing by the climatological standard deviation. This removes all site-specific characteristics from the data and makes it possible to fit one single regression model for all stations at once. As the model does not depend on the station locations, it directly allows the creation of probabilistic forecasts for any arbitrary location. SAMOS uses a left-censored power-transformed logistic response distribution to account for the large fraction of zero observations (dry days), the limitation to nonnegative values, and the positive skewness of the data. ECMWF reforecasts are used for model training and to correct the ECMWF ensemble forecasts with the big advantage that SAMOS does not require an extensive archive of past ensemble forecasts as only the most recent four reforecasts are needed, and it automatically adapts to changes in the ECMWF ensemble model. The application of the new method to the central Alps shows that the new method is able to depict the small-scale properties and returns accurate fully probabilistic spatial forecasts.
Nonhomogeneous boosting for predictor selection in ensemble postprocessing

Nonhomogeneous regression is often used to statistically postprocess ensemble forecasts. Usually only ensemble forecasts of the predictand variable are used as input, but other potentially useful information sources are ignored. Although it is straightforward to add further input variables, overfitting can easily deteriorate the forecast performance for increasing numbers of input variables. This paper proposes a boosting algorithm to estimate the regression coefficients, while automatically selecting the most relevant input variables by restricting the coefficients of less important variables to zero. A case study with ensemble forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) shows that this approach effectively selects important input variables to clearly improve minimum and maximum temperature predictions at five central European stations.
Simultaneous ensemble postprocessing for multiple lead times with standardized anomalies

Separate statistical models are typically fit for each forecasting lead time to postprocess numerical weather prediction (NWP) ensemble forecasts. Using standardized anomalies of both NWP values and observations eliminates most of the lead-time-specific characteristics so that several lead times can be forecast simultaneously. Standardized anomalies are formed by subtracting a climatological mean and dividing by the climatological standard deviation. Simultaneously postprocessing forecasts between +12 and +120 h increases forecast coherence between lead times, yields a temporal resolution as high as the observation interval (e.g., up to 10 min), and speeds up computation times while achieving a forecast skill comparable to the conventional method.

General information
State: Published
Organisations: University of Innsbruck
Authors: Dabernig, M. (Ekstern), Mayr, G. J. (Ekstern), Messner, J. (Intern), Zeileis, A. (Ekstern)
Pages: 2523-2531
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Monthly Weather Review
Volume: 145
Issue number: 7
ISSN (Print): 0027-0644
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.24 SJR 2.558 SNIP 1.322
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 3.158 SNIP 1.501 CiteScore 3.44
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 3.578 SNIP 1.697 CiteScore 3.46
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 3.934 SNIP 1.729 CiteScore 3.48
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 1
Heteroscedastic censored and truncated regression with crch

The crch package provides functions for maximum likelihood estimation of censored or truncated regression models with conditional heteroscedasticity along with suitable standard methods to summarize the fitted models and compute predictions, residuals, etc. The supported distributions include left- or right-censored or truncated Gaussian, logistic, or student-t distributions with potentially different sets of regressors for modeling the conditional location and scale. The models and their R implementation are introduced and illustrated by numerical weather prediction tasks using precipitation data for Innsbruck (Austria).

General information
State: Published
Organisations: University of Innsbruck
Authors: Messner, J. W. (Intern), Mayr, G. J. (Ekstern), Zeileis, A. (Ekstern)
Number of pages: 9
Pages: 173-181
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: R Journal
Volume: 8
Issue number: 1
Original language: English
Links:
http://www.scopus.com/inward/record.url?scp=85010282211&partnerID=8YFLogxK (Link to publication in Scopus)
Source: Scopus
Source-ID: 85010282211
Publication: Research - peer-review › Journal article – Annual report year: 2016
Predicting wind power with reforecasts
Energy traders and decision-makers need accurate wind power forecasts. For this purpose, numerical weather predictions (NWPs) are often statistically postprocessed to correct systematic errors. This requires a dataset of past forecasts and observations that is often limited by frequent NWP model enhancements that change the statistical model properties. Reforecasts that recompute past forecasts with a recent model provide considerably longer datasets but usually have weaker setups than operational models. This study tests the reforecasts from the National Oceanic and Atmospheric Administration (NOAA) and the European Centre for Medium-Range Weather Forecasts (ECMWF) for wind power predictions. The NOAA reforecast clearly performs worse than the ECMWF reforecast, the operational ECMWF deterministic and ensemble forecasts, and a limited-area model of the Austrian weather service [Zentralanstalt für Meteorologie und Geodynamik (ZAMG)]. On the contrary, the ECMWF reforecast has, of all tested models, the smallest squared errors and one of the highest financial values in an energy market.

General information
State: Published
Organisations: University of Innsbruck
Authors: Dabernig, M. (Ekstern), Mayr, G. J. (Ekstern), Messner, J. W. (Intern)
Number of pages: 8
Pages: 1655-1662
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Weather and Forecasting
Volume: 30
Issue number: 6
ISSN (Print): 0882-8156
Ratings:
Scopus rating (2016): SJR 1.48 SNIP 1.066
Scopus rating (2015): SJR 1.77 SNIP 1.246
Scopus rating (2014): SJR 1.907 SNIP 1.342
Scopus rating (2013): SJR 2.371 SNIP 1.316
Scopus rating (2012): SJR 2.013 SNIP 1.113
Scopus rating (2011): SJR 2.07 SNIP 1.235
Scopus rating (2010): SJR 2.018 SNIP 1.273
Scopus rating (2009): SJR 1.36 SNIP 1.328
Scopus rating (2008): SJR 1.533 SNIP 1.428
Scopus rating (2007): SJR 1.581 SNIP 1.405
Scopus rating (2006): SJR 2.006 SNIP 1.213
Scopus rating (2005): SJR 1.738 SNIP 1.208
Scopus rating (2004): SJR 1.622 SNIP 2.002
Scopus rating (2003): SJR 1.588 SNIP 1.048
Scopus rating (2002): SJR 1.287 SNIP 1.12
Scopus rating (2001): SJR 1.334 SNIP 0.817
Scopus rating (2000): SJR 1.517 SNIP 1.112
Scopus rating (1999): SJR 1.503 SNIP 0.91
Original language: English
DOIs:
10.1175/WAF-D-15-0095.1
Links:
http://www.scopus.com/inward/record.url?scp=84953406505&partnerID=8YFLogxK (Link to publication in Scopus)
Source: Scopus
Source-ID: 84953406505
Publication: Research - peer-review › Journal article – Annual report year: 2015

Probabilistic wind power forecasts with an inverse power curve transformation and censored regression
Forecasting wind power is an important part of a successful integration of wind power into the power grid. Forecasts with lead times longer than 6 h are generally made by using statistical methods to post-process forecasts from numerical weather prediction systems. Two major problems that complicate this approach are the non-linear relationship between wind speed and power production and the limited range of power production between zero and nominal power of the turbine. In practice, these problems are often tackled by using non-linear non-parametric regression models. However,
such an approach ignores valuable and readily available information: the power curve of the turbine’s manufacturer. Much of the non-linearity can be directly accounted for by transforming the observed power production into wind speed via the inverse power curve so that simpler linear regression models can be used. Furthermore, the fact that the transformed power production has a limited range can be taken care of by employing censored regression models. In this study, we evaluate quantile forecasts from a range of methods: (i) using parametric and non-parametric models, (ii) with and without the proposed inverse power curve transformation and (iii) with and without censoring. The results show that with our inverse (power-to-wind) transformation, simpler linear regression models with censoring perform equally or better than non-linear models with or without the frequently used wind-to-power transformation.
Heteroscedastic extended logistic regression for postprocessing of ensemble guidance

To achieve well-calibrated probabilistic forecasts, ensemble forecasts are often statistically postprocessed. One recent ensemble-calibration method is extended logistic regression, which extends the popular logistic regression to yield full probability distribution forecasts. Although the purpose of this method is to postprocess ensemble forecasts, usually only the ensemble mean is used as the predictor variable, whereas the ensemble spread is neglected because it does not improve the forecasts. In this study it is shown that when simply used as an ordinary predictor variable in extended logistic regression, the ensemble spread affects the location but not the variance of the predictive distribution. Uncertainty information contained in the ensemble spread is therefore not utilized appropriately. To solve this drawback a new approach is proposed where the ensemble spread is directly used to predict the dispersion of the predictive distribution. With wind speed data and ensemble forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) it is shown that by using this approach, the ensemble spread can be used effectively to improve forecasts from extended logistic regression. © 2014 American Meteorological Society.

General information
State: Published
Organisations: University of Innsbruck, Cornell University
Authors: Messner, J. W. (Intern), Mayr, G. J. (Ekstern), Zeileis, A. (Ekstern), Wilks, D. S. (Ekstern)
Number of pages: 9
Pages: 448-456
Publication date: Jan 2014
Main Research Area: Technical/natural sciences

Publication information
Journal: Monthly Weather Review
Volume: 142
Issue number: 1
ISSN (Print): 0027-0644
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.24 SJR 2.558 SNIP 1.322
Web of Science (2016): Indexed Yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 3.158 SNIP 1.501 CiteScore 3.44
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 3.578 SNIP 1.697 CiteScore 3.46
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 3.934 SNIP 1.729 CiteScore 3.48
Extending extended logistic regression: Extended versus separate versus ordered versus censored

Extended logistic regression is a recent ensemble calibration method that extends logistic regression to provide full continuous probability distribution forecasts. It assumes conditional logistic distributions for the (transformed) predictand and fits these using selected predictand category probabilities. In this study extended logistic regression is compared to the closely related ordered and censored logistic regression models. Ordered logistic regression avoids the logistic distribution assumption but does not yield full probability distribution forecasts, whereas censored regression directly fits the full conditional predictive distributions. The performance of these and other ensemble postprocessing methods is tested on wind speed and precipitation data from several European locations and ensemble forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF). Ordered logistic regression performed similarly to extended logistic regression for probability forecasts of discrete categories whereas full predictive distributions were better predicted by censored regression.
Brief communication Spatial and temporal variation of wind power at hub height over Europe

Wind power over Europe computed from two years of the new 100 m wind product from ECMWF at 16 km horizontal resolution is 20% of maximum capacity of an exemplary wind turbine power curve. This is five percent of maximum capacity less than extrapolated from 10 m winds using model roughness in the logarithmic law, but eight percent more than multiplying 10 m winds by a constant factor of 1.28 as in a previous study. The result from the new data set happens to be very close to the actual capacity factor of 21% for European wind turbines (Boccard, 2009). The capacity factor in high power regions between 50 and 58° N and most of northernmost Africa is almost 30%. The aggregation of wind power over Europe smooths onshore day-to-day fluctuations to at most 7 percentage points during 80% of the year. © 2013 Author(s).
Wind speeds at heights crucial for wind energy: Measurements and verification of forecasts

Wind speed measurements from one year from meteorological towers and wind turbines at heights between 20 and 250 m for various European sites are analyzed and are compared with operational short-term forecasts of the global ECMWF model. The measurement sites encompass a variety of terrain: offshore, coastal, flat, hilly, and mountainous regions, with
low and high vegetation and also urban influences. The strongly differing site characteristics modulate the relative contribution of synoptic-scale and smaller-scale forcing to local wind conditions and thus the performance of the NWP model. The goal of this study was to determine the best verifying model wind among various standard wind outputs and interpolation methods as well as to reveal its skill relative to the different site characteristics. Highest skill is reached by wind from a neighboring model level, as well as by linearly interpolated wind from neighboring model levels, whereas the frequently applied 10-m wind logarithmically extrapolated to higher elevations yields the largest errors. The logarithmically extrapolated 100-m model wind reaches the best compromise between availability and low cost for data even when the vertical resolution of the model changes. It is a good choice as input for further statistical postprocessing. The amplitude of measured, height-dependent diurnal variations is underestimated by the model. At low levels, the model wind speed is smaller than observed during the day and is higher during the night. At higher elevations, the opposite is the case. © 2012 American Meteorological Society.

General information
State: Published
Organisations: University of Innsbruck
Authors: Drechsel, S. (Ekstern), Mayr, G. J. (Ekstern), Messner, J. (Intern), Stauffer, R. (Ekstern)
Number of pages: 16
Pages: 1602-1617
Publication date: 1 Sep 2012
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of Applied Meteorology and Climatology
Volume: 51
Issue number: 9
ISSN (Print): 1558-8424
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 2.73 SJR 1.796 SNIP 1.183
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.87 SNIP 1.405 CiteScore 2.68
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.946 SNIP 1.428 CiteScore 2.62
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.962 SNIP 1.516 CiteScore 2.47
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 1.847 SNIP 1.351 CiteScore 2.03
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 1.672 SNIP 1.258 CiteScore 2.17
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 1.813 SNIP 1.197
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 1.883 SNIP 1.227
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 1.705 SNIP 1.418
Scopus rating (2007): SJR 1.895 SNIP 1.44
Scopus rating (2006): SJR 1.698 SNIP 1.425
Why Does It Always Rain on Me? A Spatio-Temporal Analysis of Precipitation in Austria

It is popular belief that the weather is “bad” more frequently on weekends than on other days of the week and this is often perceived to be associated with an increased chance of rain. In fact, the meteorological literature does report some evidence for such human-induced weekly cycles although these findings are not undisputed. To contribute to this discussion, a modern data-driven approach using structured additive regression models is applied to a newly available high-quality data set for Austria. The analysis investigates how an ordered response of rain intensities is influenced by a (potential) weekend effect while adjusting for spatio-temporal structure using spatially varying effects of overall level and seasonality patterns. The underlying data are taken from the HOMSTART project which provides daily precipitation quantities over a period of more than 60 years and a dense net of more than 50 meteorological stations all across Austria.

General information
State: Published
Organisations: University of Innsbruck
Authors: Umlauf, N. (Ekstern), Mayr, G. J. (Ekstern), Messner, J. (Intern), Zeileis, A. (Ekstern)
Pages: 81–92
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information
Journal: Austrian Journal of Statistics
Volume: 41
Issue number: 1
ISSN (Print): 1026-597X
Ratings:
BFI (2017): BFI-level 1
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 0.26 SJR 0.114 SNIP 0.346
BFI (2015): BFI-level 1
BFI (2014): BFI-level 1
BFI (2013): BFI-level 1
BFI (2012): BFI-level 1
BFI (2011): BFI-level 1
BFI (2010): BFI-level 1
BFI (2009): BFI-level 1
BFI (2008): BFI-level 1
Original language: English
Rainfall, Generalized Additive Model, Structured Additive Regression Model, Ordered Probit Model, HOMSTART, BayesX, R

Electronic versions:
121Zeileis.pdf
Publication: Research - peer-review › Journal article – Annual report year: 2012
Probabilistic forecasts using analogs in the idealized Lorenz96 setting
Three methods to make probabilistic weather forecasts by using analogs are presented and tested. The basic idea of these methods is that finding similar NWP model forecasts to the current one in an archive of past forecasts and taking the corresponding analyses as prediction should remove all systematic errors of the model. Furthermore, this statistical postprocessing can convert NWP forecasts to forecasts for point locations and easily turn deterministic forecasts into probabilistic ones. These methods are tested in the idealized Lorenz96 system and compared to a benchmark bracket formed by ensemble relative frequencies from direct model output and logistic regression. The analog methods excel at longer lead times. © 2011 American Meteorological Society.