Data Article

Title: Spatio-temporal forecasting of global horizontal irradiance using Bayesian inference

Authors: Caston Sigauke, Edina Chandiwana and Alphonce Bere

Affiliations: Department of Mathematical and Computational Sciences, University of Venda, Private Bag X5050, Thohoyandou, 0950, South Africa.

Contact email: caston.sigauke@univen.ac.za

Abstract
The article presents a description of the global horizontal irradiance data with the important drivers (covariates) used in developing spatio-temporal forecasting models. The data is from the Southern African Universities Radiometric Network (SAURAN) https://sauran.ac.za/. The following radiometric stations are used in this study: CSIR - CSIR Energy Centre (1), CUT-Central University of technology (2), UFH-University of Fort Hare (3), UNV- University of Venda (4), UNZ-University of Zululand (5), UPR-University of Pretoria(6), MIN-CRSES Mintek (7) and NUST-Namibian University of Science and Technology (8).

The data are stored in an Excel file.

Specifications Table

Subject area	*Spatial Predictive Modelling*
More specific subject area	*Spatio-temporal forecasting of renewable energies*
Type of data	*Excel file*
How data was acquired	*Provided and from the internet* https://sauran.ac.za/
Data format	*Filtered and analysed.*
Experimental factors	*N/A*
Experimental features	*N/A*
Data source location	*South Africa*
Data accessibility	*Data is hosted on GitHub* https://github.com/csigauke
Related research article	*The relevant research article is: Spatio-temporal forecasting of global horizontal irradiance using Bayesian inference.*

Value of the Data
- The data allows researchers to develop spatio-temporal predictive models for short-term forecasting of global horizontal irradiance. This is important for the large-scale integration of solar power onto the grid.

- The data can help researchers identify important drivers of global horizontal irradiance, such as nonlinear trend variables, which will result in improved forecasts.

Data

The data are daily averaged global horizontal irradiance from SAURAN. The data are stored in an Excel file. While cleaning data, it is noted that some stations have too many missing values. These are removed, and simple imputation techniques are implemented using the R package "mice" for the remaining ones. Adaptive nonlinear trend variables are used in the study. The trend variables are determined by decomposing the response variables for the eight radiometric stations used in the study using penalised cubic splines. The trend values are then extracted for the variable 'noltrend'. The other covariates are elevation (Elev), temperature (Temp), relative humidity (RH), wind speed (WS), wind direction (WD), wind direction standard deviation (WDStdDev), barometric pressure (BP), month, day and the dependent variable GHI.

Experimental Design, Materials, and Methods

Data used in the study is from SAURAN.

References

Southern African Universities Radiometric Network (SAURAN) https://sauran.ac.za/ (Accessed on 13 August 2022).