Engineering Applications of Artificial Intelligence 69 (2018) 157–167

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting

O. García-Hinde [a,*], G. Terrén-Serrano [b], M.Á. Hombrados-Herrera [b], V. Gómez-Verdejo [a], S. Jiménez-Fernández [e], C. Casanova-Mateo [c,d], J. Sanz-Justo [d], M. Martínez-Ramón [b], S. Salcedo-Sanz [e]

[a] *Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Leganés, Madrid, Spain*
[b] *Department of Electrical and Computing Engineering, The University of New Mexico, NM, USA*
[c] *Department of Civil Engineering: Construction, Infrastructure and Transport, UPM, Madrid, Spain*
[d] *LATUV, Universidad de Valladolid, Valladolid, Spain*
[e] *Department of Signal Processing and Communications, Universidad de Alcalá, Madrid, Spain*

ARTICLE INFO

Keywords:
Dimensionality reduction
Interpretability
Solar radiation forecast
Weather research and forecasting model
Support vector regression
Restricted Boltzmann machine

ABSTRACT

The interest in solar radiation prediction has increased greatly in recent times among the scientific community. In this context, Machine Learning techniques have shown their ability to learn accurate prediction models. The aim of this paper is to go one step further and automatically achieve interpretability during the learning process by performing dimensionality reduction on the input variables. To this end, three non standard multivariate feature selection approaches are applied, based on the adaptation of strong learning algorithms to the feature selection task, as well as a battery of classic dimensionality reduction models. The goal is to obtain robust sets of features that not only improve prediction accuracy but also provide more interpretable and consistent results. Real data from the Weather Research and Forecasting model, which produces a very large number of variables, is used as the input. As is to be expected, the results prove that dimensionality reduction in general is a useful tool for improving performance, as well as easing the interpretability of the results. In fact, the proposed non standard methods offer important accuracy improvements and one of them provides with an intuitive and reduced selection of features and mesoscale nodes (around 10% of the initial variables centered on three specific nodes).

1. Introduction

According to the December 13th, 2016 report of Solar Energy Industries Association (SEIA),[1] the total installed solar power capacity in the United States of America reached 35.8 GW in the third quarter of 2016, representing over 60% of the total installed electric capacity. This report states that there are more than 1 million residential solar installations across the country, and their industry growth nearly doubles every year. Regarding European countries, emphasis on solar power is decreasing. A report from Solar Power Europe[2] showed that a total of 1.56 GW in solar capacity were installed from June to September, which was 10% less than in the previous quarter. Nevertheless, during the first quarter of 2016, Europe reached 100GW of installed solar capacity. On the other hand, and according to SEIA, China and Japan lead the solar power market with 50% of new installed capacity. A fundamental tool for this active and growing market is, obviously, solar forecast (Voyant et al., 2017).

Atmospheric behavior makes solar power highly stochastic. Therefore, an efficient use of solar energy requires intelligent systems, specifically ones that are able to forecast the energy to be produced at different time-horizon scales, ranging from minutes to days. The amount of generated solar power primarily depends on cloud coverage, but also on other factors such as the presence of light absorbing particles in the air (Khatib et al., 2012). Many cloud coverage prediction methods

* Corresponding author.
E-mail addresses: oghinde@tsc.uc3m.es (O. García-Hinde), guillermoterren@unm.edu (G. Terrén-Serrano), mangelhombradosherre@unm.edu (M.Á. Hombrados-Herrera), vanessa@tsc.uc3m.es (V. Gómez-Verdejo), silvia.jimenez@uah.es (S. Jiménez-Fernández), carlos.casanova@upm.es (C. Casanova-Mateo), julia@latuv.uva.es (J. Sanz-Justo), manel@unm.edu (M. Martínez-Ramón), sancho.salcedo@uah.es (S. Salcedo-Sanz).

1 Source: www.seia.org.
2 Source: www.solarpowereurope.org.

https://doi.org/10.1016/j.engappai.2017.12.003
Received 22 May 2017; Received in revised form 11 October 2017; Accepted 5 December 2017
0952-1976/

rely on direct measurements, including ground or satellite observations of the clouds. The most widely used approach to address the task of solar radiation prediction consists in the physical modeling of the deterministic part of solar radiation, by computing the relative position of the sun with respect to the facility in order to obtain a clear sky model, and then adding the atmospheric conditions, including rain, wind speed and other variables (Bilgili and Ozoren, 2011). However, some causes of radiation attenuation are not easily predicted by direct observation. For this reason, numerical methods to construct weather models, such as the Weather Research and Forecasting (WRF) mesoscale model (Skamarock et al., 2005), have also been applied, as they provide significant atmospheric information in the surroundings of the location under study, thus improving predictions (Dasarathy, 2011).

In this work we present a study on feature selection and extraction methods for solar radiation forecast from a WRF model. The WRF model provides forecast of atmospheric variables at different heights for a given area. This model has the drawback of presenting a large number of dimensions (prediction variables). Indeed, the WRF model used in this paper produces around 10^4 prediction variables per time instant, while the number of samples available for the algorithm's training is much lower. This situation makes the use of powerful dimensionality reduction strategies mandatory. We propose and evaluate a series of novel methods that automatically select the most powerful features by adapting strong regression algorithms, such as SVMs or Deep Neural Networks (DNN), to the task of feature selection. We will compare the performance of these methods to other classic selection and extraction strategies and see their effect on interpretability. Our experiments show how our methods can maintain high prediction accuracies, while increasing interpretability by finding relationships and patterns within the data that are opaque to expert human knowledge.

2. Related work

There is a significant amount of work devoted to the prediction of solar radiation. Most approaches tackle the problem from a computational intelligence perspective. These strategies make use of different data sources. Some use meteorological data as inputs for *Machine Learning* (ML) predictors in order to improve their forecasting performance. Indeed, ML models have been used for example in Benghanem and Mellit (2010), where a Radial Basis Function was used in solar radiation prediction in a power plant using weather data. A comparison of prediction techniques can be found in Paoli et al. (2010) where the authors take a time series prediction approach where the input data consists of historical solar radiation data. There, Multi-Layer Perceptron (MLP) neural networks, Markov chains, Bayesian inference and ARIMA models are compared. Support Vector Machine models (SVM), and specifically Support Vector Regressors (SVR) (Burges, 1998), have also been widely used in energy production forecast (see e.g. Salcedo-Sanz et al., 2014b). For example in Chen et al. (2011) SVRs are used to predict monthly solar radiation from meteorological data, and the same authors use them in Chen et al. (2015) to estimate solar radiation from air temperature. Extreme learning machines (ELM) have also been applied to solar prediction using meteorological variables in Salcedo-Sanz et al. (2014a); and in Shamshirband et al. (2015), a Kernel Extreme Learning Machine (KELM) has been compared to a kernel SVR. Other works introduce neuro-fuzzy approaches (Olatomiwa et al., 2015), or hybrid models combining ARMA and artificial neural networks (Voyant et al., 2013), to cite some.

In other cases, researchers make use of observations of cloud evolution from satellite data. For example, in Jang et al. (2016) SVRs and Multilayer Perceptrons are used to predict the evolution of the clouds seen from satellite images. The data is pre-processed to generate variables related to the size, motion and other factors of the clouds.

In Kato et al. (2016), a numerical Weather Forecast System is combined with satellite infrared images to predict several hours ahead. In Cros et al. (2014), a model that estimates cloud motion is used to predict the future positions of the clouds from satellite imagery. In Linares-Rodriguez et al. (2013), a MLP combined with a genetic algorithm is used to perform solar radiation forecast. The authors use satellite images to perform radiation prediction in large areas of Spain. Other works, such as Perez et al. (2010), Schillings et al. (2004), Hammer et al. (1999) and Ineichen and Perez (1999), also make use of satellite data in radiation prediction. Forecasting with shorter horizons can be implemented using ground images of the clouds. In Li et al. (2015) a cloud identification model is constructed from RGB images. Cloud monitoring is introduced in Tapakis and Charalambides (2013) for prediction. A short term solar radiance prediction scenario using observations of the whole sky was presented in Chow et al. (2011). Other related works are Huang et al. (2013), Cervantes et al. (2016), Marquez and Coimbra (2013), Cheng and Yu (2016) and Sun et al. (2014). While the previous works use RGB images, in Mammoli et al. (2013) a prediction model is developed using a LAPART neural network (Caudell and Healy, 1996) and infrared images.

A portion of the strategies described above employ ML models for prediction using low or moderate dimensionality databases. However, many of the aforementioned scenarios use sources that produce data of very high dimensionality (Bolon-Canedo et al., 2015), i.e. satellite images or WRF mesoscale models. In these cases, it is of great importance to use dimensionality reduction methods to manage the structural complexity of the learning algorithm. To this end, there are two basic approaches: feature selection and feature extraction (Guyon and Elisseeff, 2003; Guyon et al., 2008; Fodor, 2002; Martin and Maes, 1979). Regarding specific application of dimensionality reduction methods to solar energy, in Yadav and Chandel (2014) a first general review of some works dealing with relevant parameters selection in solar energy prediction problems is offered. In Fu and Cheng (2013) a system for solar irradiance very short-term prediction (minutes time-horizon) is proposed in which a correlation filter is applied to select relevant features. In Yadav et al. (2014) a study of the main influencing input parameters for solar radiation prediction with neural networks is carried out in different locations of India, by using a Decision Tree variable selection method. In Rana et al. (2016) the problem of forecasting the electricity power generation by a solar photo-voltaic system is tackled, comparing multivariate and univariate correlation measurements to select useful features. In Mohammadi et al. (2016) an adaptive neuro-fuzzy inference system (ANFIS) has been applied to select the most influential variables in a daily horizontal diffuse solar radiation prediction problem. In Will et al. (2013) two applications of hybrid niching genetic algorithms are presented to solve the problem of variable selection for the estimation of Solar Radiation.

3. Solar radiance problem formulation

The solar radiation prediction problem formulation can be stated in the following way: let $\mathcal{R}_t$ be the global solar radiation registered at a given time t in a location $\mathcal{L}$ of the Earth's surface, and let $\hat{\mathcal{R}}_t$ be the prediction of the global solar radiation under the same considerations. In order to predict $\hat{\mathcal{R}}_t$, let us consider a set of N atmospheric variables $\mathcal{V}$ (outputs of a mesoscale model), some of them referred to different pressure levels (ranging from the pressure level corresponding to the ground level to 50 hPa). Considering each variable at each pressure level as a different predictor, the set of variables can be expressed as $\mathcal{V} = (v_{11}, \dots, v_{1N}, v_{21}, \dots, v_{2N}, \dots, v_{M1}, \dots, v_{MN})$, where M stands for the total number of grid points where the variables were obtained. Fig. 1 shows an outline of a generic grid ($m = 1 \dots M$) and the set of variables ($n = 1 \dots N$) considered at each point.

3.1. Model $\mathcal{M}$: the weather research and forecasting model

In this work, the Weather Research and Forecasting (WRF) mesoscale model (Skamarock et al., 2005) has been considered to obtain the set of atmospheric variables used as predictive variables. In this study, WRF