García-Cuesta, E., Aler, R., Pózo-Vázquez, D.d. et al.
A combination of supervised dimensionality reduction
and learning methods to forecast solar radiation. *Appl
Intell* (2022).

DOI: 10.1007/s10489-022-04175-y

A combination of supervised dimensionality reduction and learning methods to forecast solar radiation

Esteban García-Cuesta[1] · Ricardo Aler[2] · David del Pózo-Vázquez[3] · Inés M. Galván[2]

Abstract

Machine learning is routinely used to forecast solar radiation from inputs, which are forecasts of meteorological variables provided by numerical weather prediction (NWP) models, on a spatially distributed grid. However, the number of features resulting from these grids is usually large, especially if several vertical levels are included. Principal Components Analysis (PCA) is one of the simplest and most widely-used methods to extract features and reduce dimensionality in renewable energy forecasting, although this method has some limitations. First, it performs a global linear analysis, and second it is an unsupervised method. Locality Preserving Projection (LPP) overcomes the locality problem, and recently the Linear Optimal Low-Rank (LOL) method has extended Linear Discriminant Analysis (LDA) to be applicable when the number of features is larger than the number of samples. Supervised Nonnegative Matrix Factorization (SNMF) also achieves this goal extending the Nonnegative Matrix Factorization (NMF) framework to integrate the logistic regression loss function. In this article we try to overcome all these issues together by proposing a Supervised Local Maximum Variance Preserving (SLMVP) method, a supervised non-linear method for feature extraction and dimensionality reduction. PCA, LPP, LOL, SNMF and SLMVP have been compared on Global Horizontal Irradiance (GHI) and Direct Normal Irradiance (DNI) radiation data at two different Iberian locations: Seville and Lisbon. Results show that for both kinds of radiation (GHI and DNI) and the two locations, SLMVP produces smaller MAE errors than PCA, LPP, LOL, and SNMF, around 4.92% better for Seville and 3.12% for Lisbon. It has also been shown that, although SLMVP, PCA, and LPP benefit from using a non-linear regression method (Gradient Boosting in this work), this benefit is larger for PCA and LPP because SMLVP is able to perform non-linear transformations of inputs.

Keywords Dimensionality reduction · Hybrid learning · Solar radiation forecast · Data mining

1 Introduction

Considerable efforts have been made in the past decades to make solar energy a real alternative to the conventional energy generation system. There are two main technologies, solar thermal electricity (STE) and solar photovoltaic (PV) energy, and many countries have already reached a notable solar share in their energy mixes. Moreover, important growth is expected in the near future (International Energy Agency, 2018).

Contrary to conventional generation, solar electricity generation is conditioned by weather, and thus it is highly intermittent. Transient clouds and aerosol intermittency lead to considerable variability in the solar power plants yield on a wide range of temporal scales, particularly in minutes to hours time scales. This presents serious issues regarding solar power plant management and their yield integration into the electricity grid [1]. Currently, in addition to expensive storage-based solutions, the use of solar radiation forecasts is the only plausible way to mitigate the intermittency. Therefore, the development of accurate solar radiation forecasting methods has become an essential research topic [2].

Solar forecasting methods can be classified depending on the forecasting horizon. Nowcasting methods are mostly related to one-hour ahead forecasts, short-term forecasting with up to 6 hours ahead forecasts and forecasting methods are aimed at producing days ahead forecasts. The techniques associated with these methods are essentially different [3–5]. In recent years, these has been increasing interest,

✉ Esteban García-Cuesta
esteban.garcia@fi.upm.es

Extended author information available on the last page of the article.

particularly, in short-term forecasting, fostered by the expected massive deployment of solar PV energy. Accurate short-term solar forecasts are important to ensure the quality of the PV power delivered to the electricity network and, thus, to reduce the ancillary costs [6, 7]. Short-term forecasting has also been successfully used for the management of STE plants [8, 9] and for the participation of PV and STE plants in the energy market [8, 10].

Short-term forecasts can be derived either from satellite imagery [11, 12] or from Numerical Weather Prediction (NWP) models [13–15]. As solar radiation measured datasets have become progressively available, the use of data-driven methods have become increasingly popular [16]. In [15, 17, 18] a comparison of the performance of different methods is assessed.

The use of NWP models for short-term solar forecasting has some important advantages, such as the global and easy availability of the forecasts. Because of that, this approach was extensively evaluated during the past decade [14, 15, 19]. Nevertheless, the reliability is far from optimal and machine-learning methods play an important role in providing enhanced solar forecasts derived from NWPs models [20, 21]. In this context, the inputs for machine learning techniques are forecasts of several meteorological variables provided by numerical weather prediction (NWP) physical models such as the European Center for Medium Weather Forecasts (ECMWF) and the Global Ensemble Forecast System (GEFS). Meteorological variables are forecast for the points of a grid over the area of interest. However, the number of features resulting from these grids is usually large, especially if several vertical levels are included in the grid. This may result in models that do not generalize well, and techniques to reduce the dimensionality of data are required.

Dimensionality reduction techniques can be divided into feature selection and feature extraction. Feature selection methods select the most relevant variables in the grid, while feature extraction summarizes information from the whole grid into fewer features. Both approaches have been used in the context of renewable energy forecasting with machine learning [22, 23]. Feature selection techniques have been used in [24] where methods such as Linear Correlation, ReliefF, and Local Information Analysis have been explored to study the influence on forecast accuracy of the number of NWP grid nodes used as input for the solar forecasting model.

In [25], feature extraction (PCA) is compared with feature selection (a minimal redundancy and maximal-relevance method) to reduce the dimensionality of variables in a grid for wind power forecasting in the east of China. The authors conclude that PCA is a good choice to simplify the feature set, while obtaining competitive results. PCA has also been used in [26] together with domain knowledge

to extract features from a NWP grid to improve renewable energy forecasting. Advanced machine learning methods, such as convolutional neural networks, have also been used as a feature extraction scheme for wind power prediction using NWPs, showing competitive results compared to a PCA baseline [27]. García-Hinde et al. [28] presents a study on feature selection and extraction methods for solar radiation forecasting. The study includes classical methods, such as PCA or variance and correlation filters, and novel methods based on the adaptation of the support vector machines and deep Boltzmann machines for the task of feature selection. Results show that one of the novel methods (the adaptation of support vector machine) and PCA select high relevance features. Verbois et al. [29] combine feature extraction (PCA) and stepwise feature selection of NWP variables for solar irradiance forecasting, comparing favorably with other benchmark methods. In [30] a hybrid approach that combines PCA and deep learning is presented to forecast wind power from hours to years, showing a good performance. A recent study on solar irradiance forecasting has compared many methods on different datasets, where PCA has been used as the main method for feature extraction and dimensionality reduction [31]. In general, it is observed that PCA, even in recent works, is one of the most widely-used methods to extract features in renewable energy forecasting.

PCA is a multivariate statistical analysis that transforms a number of correlated variables into a smaller group of uncorrelated variables called principal components [32]. PCA has two main limitations. First, it performs a global linear analysis by an axis transformation that best represents the mean and variance of the given data, but lacks the ability to give local information representation. Second, PCA is an unsupervised method, that is, the target output is not used to extract the new features and this may be a drawback to finding the best low dimensional representation whenever labels are available.

In this article we propose Supervised Local Maximum Variance Preserving (SLMVP), a kernel method for supervised feature extraction and dimensionality reduction. The method considers both characteristics: it preserves the maximum local variance and distribution of the data, but also considers the distribution of the data by the response variable to find an embedding that best represents the given data structure. This method can be applied to multiclass and regression problems when the sample size m is small and the dimensionality p is relatively large or very large as opposed to Fisher's Linear Discriminant Analysis (LDA) [33], one of the foundational and most important approaches to classification. In summary, SLMVP uses the full or partially labeled dataset to extract new features that maximize the variance of the embedding that best represents the common local distances [34] and computationally is

based on weighted graphs [35]. Additionally, this method is able to perform a linear and non-linear transformation of the original space by using different kernels as the similarity metric.

To validate the SLMVP method, it has been tested to extract features in order to improve solar radiation forecasting (both Global Horizon Irradiance (GHI) and Direct Normal Irradiance (DNI)) for a 3-hour forecasting horizon, and compared to PCA (the most popular workhorse in the area), but also to other state-of-the-art methods that have not been previously used in the context of solar radiation forecasting. These methods are (1) Locality Preserving Projection (LPP, an unsupervised local dimensionality reduction method) that finds linear projective maps that arise by solving a variational problem that optimally preserves the neighborhood structure of the dataset [36]; (2) Linear Optimal Low-Rank (LOL, a supervised dimensionality reduction method) that learns a lower-dimensional representation in a high-dimensional low sample size setting extending PCA by incorporating class-conditional moment estimates into the low-dimensional projection [37], and (3) Supervised Non-negative Matrix Factorization (SNMF) that extends Negative Matrix Factorization (NMF) to be supervised [38, 39]. SNMF integrates the logistic regression loss function into the NMF framework and solves it with an alternating optimization procedure. All of these methods are able to solve the "large p, small m" problem as opposed to many classical statistical approaches that were designed with a "small p, large m" situation in mind (e.g. LDA). Features have been extracted from meteorological forecasts (obtained from the GEFS) in points of a grid around two locations in the Iberian peninsula: Seville and Lisbon. Two grid sizes have been tested, small and large. The performance of SLMVP has been compared with PCA, LPP, LOL, and SNMF using two different regressors, a linear one (standard Linear Regression (LR)) and a non-linear technique (Gradient Boosting). Thus the main contributions of this work are:

- A new local and supervised dimensionality reduction method capable of solving the "large p, small m" problem.
- The application of SLMVP to reduce the dimensionality of the NWP variables in a grid for the solar radiation forecasting problem.
- The comparison with PCA, one of the most widely-used methods in the context of renewable energy for feature extraction, LPP, and two state-of-the-art recent supervised methods, LOL and SNMF, showing the usefulness of the proposed method.

The structure of the article is as follows. Section 2 explains the SLMVP method, which is tested using the data described in Section 3 and the experimental design included in Section 4. The Conclusions section summarizes the main results.

2 Supervised dimensionality reduction method: Kernel-SLMVP

As has been mentioned in Section 1, PCA is an unsupervised method that performs a global analysis of the whole dataset. As opposed to the global-based data projection techniques like PCA, other methods based on local structure preservation i.e. ISOMAP [40], LLP [36], Laplacian Eigenmaps [41], and Locally Linear Embedding [42] have been proposed in order to overcome the characteristic of being global. Although these techniques use linear optimization solutions, they are also able to represent nonlinear geometric features by local linear modeling representation that lies in a low dimensional manifold [43]. Note that these non-linear methods still do not consider labeled data, that is, they are unsupervised methods. Recently, Linear Optimal Low-Rank (LOL) projection has been proposed incorporating class-conditional means. The key intuition behind LOL is that it can jointly use the means and variances from each class (like LDA), but without requiring more dimensions than samples [37]. Another recent method is Supervised Non-negative Matrix Factorization (SNMF) that extends Negative Matrix Factorization (NMF) to be supervised [38]. SNMF integrates the logistic regression loss function into the NMF framework and solves it with an alternating optimization procedure. For both methods, regression can be done by projecting the data onto a lower-dimensional subspace followed by the application of linear or non-linear regression techniques. This mitigates the curse of high-dimensions.

The Supervised Local Maximum Variance Preserving (SLMVP) dimensionality reduction method solves the problem of LPP to work on problems with "large p small m" and the global approach of LOL and SNMF, despite being supervised. Therefore, SLMVP preserves the maximum local variance of the data being able to represent non-linear properties, but also considers the output information (in a supervised mode) to preserve the local patterns between inputs and outputs. In summary, it uses the full or partially labeled dataset to extract new features that best represent the local maximum joint variance.

SLMVP is based on a graph representation for a given set of inputs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \Re^p$, and a set of outputs $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m \in \Re^l$. With m being the sample data points and, p and l the number of input and output features, in our case the dimensionality of $l = 1$ and $p = 342$

for the small grid, and $m = 12274$ for the large grid. The application of any similarity function $\mathscr{S}$ on the inputs $\mathscr{S}_x(\mathbf{X}) : \mathbf{X} \in \Re^{m \times p}$ and $\mathscr{S}_y(\mathbf{Y}) : \mathbf{Y} \in \Re^{m \times l}$ defines an input weighted graph $\{H, U\}$ and an output weighted graph $\{I, V\}$ with H and I being the nodes, and U and V the vertex, respectively. The graphs are not constrained and can be fully connected, or some weights can have a zero value meaning that the connection between those points disappears. The weight of the links represents the similarity between two data points. These characteristics allow the method with the capability of being local. Following [41] and [35] a graph embedding viewpoint can be used to reduce the dimensionality, mapping a weighted connected graph $G = (V, E)$ to a line so that the connected points stay as close together as possible.

The unsupervised dimensionality reduction problem aims to choose the mapping $\mathbf{y}'_i = \mathbf{A}^T \mathbf{x}_i : \mathbf{y}'_i \in \Re^k$ and $k \ll p$, which minimizes the distance with its neighbors in multidimensional data and can be expressed by the next cost function:

$$J_{ns} = \frac{1}{2} \sum_{ij} \| \mathbf{y}'_i - \mathbf{y}'_j \|^2 w_{ij} \tag{1}$$

where $\mathbf{W} \in \Re^{m \times m}$ is the similarity matrix $\mathscr{S}_x(\mathbf{X})$.

Following this graph embedding approach, SLMVP solves the supervised version and the wish to choose the mapping $\mathbf{y}'_i = \mathbf{A}^T \mathbf{x}_i : \mathbf{y}'_i \in \Re^k$ and $k \ll p$, which minimizes the distance with its neighbors in multidimensional data but preserves only those distances that are shared in the input and output spaces, given the similarity functions for each of them $\mathscr{S}_x$ and $\mathscr{S}_y$. The cost function is then expressed by:

$$J_s = \frac{1}{2} \sum_{ij} \| \mathbf{y}'_i - \mathbf{y}'_j \|^2 z_{ij} \tag{2}$$

where $\mathbf{Z} \in \Re^{m \times m}$ represents the joint similarity matrix between input $\mathscr{S}_x(\mathbf{X})$ and output $\mathscr{S}_y(\mathbf{Y})$ similarity matrices, being $z_{ij} = \sum_{k=1}^{m} u_{ik} v_{kj}$. Note the difference between (1) that is non supervised and (2) that defines a supervised manifold learning problem using the similarity matrix between inputs and outputs.

The minimization of the cost function (2) can be expressed in its kernelization form (Kernel-SLMVP) after some transformations as the following maximization problem:

$$\max \ tr(\mathbf{Y}^T \mathbf{K}_x \mathbf{K}_y \mathbf{Y}) \tag{3}$$

where $\mathbf{K}_x = \mathscr{S}_x(\mathbf{X})$ and $\mathbf{K}_y = \mathscr{S}_y(\mathbf{Y})$ are the input and output similarity graphs expressed as kernel functions (i.e.

polynomial $K(a, b) = (1 + a \cdot b)^p$ or Gaussian $K(a, b) = e^{\frac{-|a-b|^2}{2\sigma^2}}$). Finally, the (3) can be solved as an eigenvector problem on $\mathbf{B}$ as follows:

$$\mathbf{X} \mathbf{K}_x \mathbf{K}_y \mathbf{X}^T \mathbf{B} = \lambda \mathbf{B} \tag{4}$$

where B is the learned latent space. The projection of the input space data $\mathbf{X}$ on this space $\mathbf{P} = \mathbf{B}^T \mathbf{X}$ are the new extracted features to be used by the machine learning model. The Python code of SLMVP has been released publicly at [44].

3 Data description

The dataset used in this study concerns GHI and DNI measurements at two radiometric solar stations in the Iberian Peninsula: Seville and Lisbon. GHI and DNI have been acquired with a Kipp & Zonen CMP6 pyranometer, with a 15-minute resolution.

The set of inputs is a collection of forecasted meteorological variables obtained from GEFS at different levels of the atmosphere and at different latitudes and longitudes. More specifically, 9 meteorological variables at different levels are used (see Table 1), making a total of 38 attributes at each latitude-longitude pair. Latitudes go from 32 to 51 and longitudes go from -18 to 6 with a resolution of 0.5 degrees. In this work, two grids of different sizes have been used: a small grid with $3 \times 3 = 9$ points around the solar station (Seville and Lisbon) and a larger one with $17 \times 19 = 323$ points. For Seville, the larger grid covers the Iberian Peninsula (latitudes: 36 to 44, longitudes: 350 to 359.5, both with a resolution of 0.5 degrees). In the case of the Lisbon solar station, the larger grid has been shifted to cover part of the Atlantic Ocean (latitudes also go from 36 to 44 and longitudes go from 346 to 355.5). Figure 1 shows both the wide and narrow grids, centered around Seville and Lisbon (in blue). Since each point in the grid contains 38 attributes, the small grid results in $3 \times 3 \times 38 = 342$ input variables, and the larger one in $17 \times 19 \times 38 = 12274$ inputs.

GEFS provides predictions of meteorological variables for a 3-hour forecasting horizon every 6 hours each day (00:00am, 06:00am, 12:00pm, and 18:00pm). The corresponding GHI and DNI measurements are also used. To select the relevant hours of the day for GHI and DNI, samples with a zenithal angle larger than 75 degrees have been removed. Given this restriction, data times range from 9:15am to 6:00pm. The total input–output data covers from March 2015 to March 2017.

In this study, GHI and DNI are normalized by the irradiance of clear sky according to (5).

$$I_{kt}(t) = I(t) / I_{cs}(t) \tag{5}$$

Table 1 Meteorological Variables

Variable	Description	Levels
CLWMR	Cloud mixing ratio	300, 350, 400, 450, 500, 550 mb
		600, 650, 700, 750, 800, 850 mb
		900, 925, 950, 975, 1000 mb
HGT	Geopotential Height	500, 850, 925, 1000
RH	Relative humidity	500, 850, 925
UGRD	U component of wind	500, 850, 925 mg
VGRD	V component of wind	500, 850, 925
SOILW	Soil Temperature	0.0-0.1 m
TMP	2-meter temperature	2 m
CAPE	Convective available potential energy	surface, 255
PRMSL	Pressure reduced to MSL	surface

where $I(t)$ stands for GHI or DNI at time t and $I_{cs}(t)$ is the irradiance of clear sky at a particular at time t.

4 Experimental validation

In order to study the performance of the SLMVP algorithm, it has to be combined with a regression method to predict normalized GHI and DNI for a 3-hour forecasting horizon and compared with the other above mentioned methods PCA, LPP, LOL, and SNMF. The regression technique uses as inputs the attributes/features from the input-space transformation obtained by the SLMVP, PCA, LPP, LOL, and SNMF methods. As suggested in [37], to learn the projection matrix for the LOL method, we partition the data into K partitions (we select $K = 10$) equally separated between the target variable range $[0 - 1]$ to obtain a K-class classification problem. In this work, linear and non-linear regression methods have been tested. As a non-linear method, a state-of-the-art machine learning technique has been used: Gradient Boosting Regression (GBR) [45, 46]. This technique has shown considerable success in predictive accuracy in recent years (see for instance [24, 47–49]).

In this Section, first the methodology employed is described. Then, the results comparing SLMVP with PCA, LPP, LOL, and SNMF for different GEFS grid sizes will be presented.

4.1 Methodology

Cross-validation (CV) has been applied to study the performance of SLMVP, PCA, LPP, LOL, and SNMF. In standard CV, instances are distributed randomly into CV partitions. But our study involves time series of data, and therefore there are temporal dependencies between consecutive samples (in other words, consecutive samples can be highly correlated). Hence, in this study, group 4-fold CV has been used, as explained next. Data has been split into 4 groups, one for each week of every month. Fold 1 thus contains the first week of each month (January, February, ...). Fold 2, the second week of every month and so on. This guarantees that, at least training and testing partitions will never contain instances belonging to the same week, which allows a more realistic analysis of the performance of the methods. Since in this work the optimal number of features must be selected, a validation set strategy has been

Fig. 1 17×19 (black) and 3×3 (red) grids

(a) Seville.   (b) Lisbon.
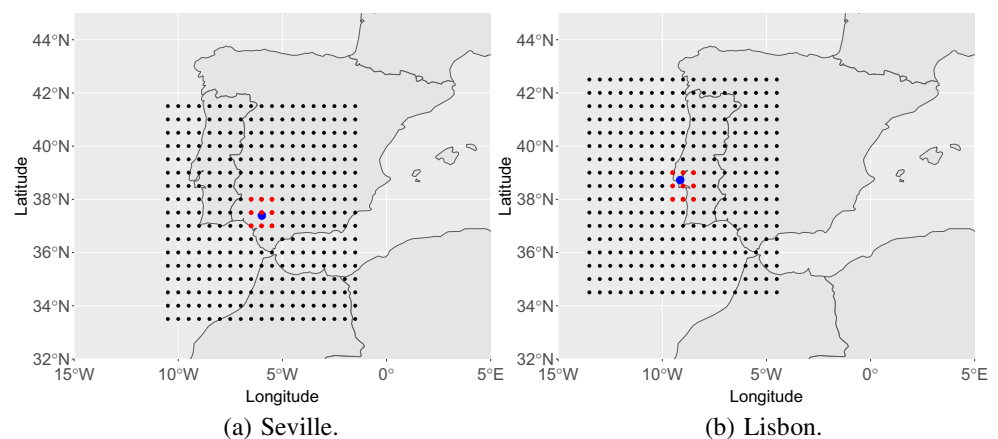
used. For this purpose, each training partition (that contains 3 folds) is again divided into training and validation sets. The validation set contains a week of each month out of the three weeks of data available in the training partition. The remaining two weeks (the ones not used for validation) are used for training.

Mean Absolute Error (MAE) has been used as the performance measure (6). Given that a 4-fold CV has been employed, results are the CV-average of MAE.

$$MAE = \frac{\sum_{i=1}^{i=N} |y_i - o_i|}{N} \qquad (6)$$

where N is the number of samples and y_i and o_i are the actual value and the output of the model, respectively. Note that the number of samples for training are 480, which is smaller than the number of dimensions for the large grid $480 << 12274$ and within the same scale factor for the small grid $342 <\approx 480$.

The performance of the methods are evaluated as follows. Recall that the number of the selected projected features is very relevant and the obtained features for the different methods are also ordered by their importance. Then, in order to analyze the optimal number of dimensions, the performance of both linear and GBR regression methods is evaluated for 5, 10, 20, 50, 100 and 150 projected features.

Given that 4-fold CV is used for performance evaluation, in each of the 4 CV iterations there is a training, validation, and testing partition. For each iteration, the regression models are trained with the training partition and then, the validation and test errors are obtained. The averages of the 4 iterations are obtained for the three errors (train, validation and test). The validation error is used to select the optimal number of features.

SLMVP, SNMF, and GBR have some hyper-parameters that require tuning in order to improve results. Five hyper-parameters were fitted: gamma parameter $\gamma = \frac{1}{2\sigma}^2$ that defines the Gaussian kernel function of the SLMVP method, α, β, and θ that defines the weight of each term of the SNMF method, and number of estimators and tree depth (which belongs to GBR). The following range of values for each hyper-parameter were tested:

- γ (SLMVP): from 0 to 2 in steps of 0.1
- α, β, θ (SNMF): from 0, 0.1, 0.01, 0.001
- Number of estimators (GBR): from 10 to 200 in steps of 10
- Tree depth (GBR): from 1 to 10 in steps of 1

In order to tune the hyper-parameters, a systematic procedure known as grid-search was used. This method tries all possible combinations of hyper-parameter values.

Table 2 Average test MAE and number of selected components for different methods and for GHI at Seville and Lisbon locations

Method	Small-grid		Large-grid	
	MAE	Components	MAE	Components
Seville				
SLMVP - LR	0.1673	20	0.1845	50
PCA - LR	0.8417	20	0.7929	20
LPP - LR	0.3655	3	>10	20
LOL - LR	0.1660	50	0.1949	50
SNMF - LR	0.1699	100	0.1890	50
SLMVP - GBR	**0.1562**	**20**	0.1653	50
PCA - GBR	0.1688	20	0.1808	50
LPP - GBR	0.2008	150	0.2605	20
LOL - GBR	0.1875	150	0.1813	100
SNMF - GBR	0.1653	50	0.1885	50
Lisbon				
SLMVP - LR	0.2035	50	0.2217	50
PCA - LR	0.7734	100	0.7706	10
LPP - LR	>5	100	>10	10
LOL - LR	0.2029	50	0.2233	100
SNMF - LR	0.2055	50	0.2209	50
SLMVP - GBR	**0.1974**	**20**	0.2084	100
PCA - GBR	0.2008	10	0.2167	50
LPP - GBR	0.2269	150	0.2548	5
LOL -GBR	0.2272	100	0.2254	150
SNMF - GBR	0.2023	20	0.2278	100

The bold entries are the best model for each case (Seville GHI and Lisbon GHI) independently of the grid size (small or large)

Models for each hyper-parameter combination are trained with the training partition and evaluated with the validation partition. The best combination on the validation set is selected.

4.2 Results

Table 2 shows the average GHI MAE for the best number of components for different methods and grid sizes. Table 3 displays the same information for DNI. The best number of components has been selected using the MAE for the validation set. In all cases, it is observed that the use of the nonlinear regression technique (GBR) improves considerably the errors for PCA and LPP, in some cases for LOL and SNMF, and always minor improvements for SLMVP. For instance, in the case of the small-grid for GHI in Seville (Table 2 top left), the use of GBR with PCA improves the MAE considerably (from 0.6467 with LR to 0.0126 with GBR accountable for a 6.8% improvement). Similar improvements for PCA, LPP, and LOL MAE can be observed for the large-grid, from 0.6084 to 0.0155 accountable for a 8.51% improvement (Table 2 top right). Lisbon GHI (Table 2 bottom) behaves in a similar way. For SNMF the differences are almost nonexistent. In the case of

SLMVP, although GBR obtains better errors than LR, the difference between linear and non-linear is smaller than for PCA and LPP cases. For instance, observing the GHI results for Seville (top of Table 2), it can be seen that for the small grid (top left), the difference between GBR and LR (when using SLMVP) is only 0.1562 vs. 0.1673, and for the large-grid (top right), is 0.1653 vs. 0.1845. Similar differences can be observed for GHI at Lisbon (bottom of Table 2). This is reasonable because SLMVP uses a non-linear kernel, so even when using LR, some of the non-linearity of the problem has been included by SLMVP feature extraction process. Conclusions for DNI (Table 3) follow a similar trend: PCA and LPP benefit more from using a non-linear method (GBR) than SLMVP and SNMF, but LOL benefits more using a regularized linear regressor. LOL includes linear class prior information about which is beneficial for LR.

Analyzing the results depending on the size of grid (small vs. large), it is observed that the use of a large grid does not result in better MAE values. The best errors are always obtained with the small grid in all cases of Tables 2 (GHI) and 3 (DNI). When the large grid is used, more components are used for SLMVP but, as already mentioned, this does not improve the results.

Table 3 Average test MAE and number of selected components for different methods and for DNI at Seville and Lisbon locations

Method	Small-grid		Large-grid	
	MAE	Components	MAE	Components
	Seville			
SLMVP - LR	0.2580	50	0.2787	50
PCA - LR	0.6628	5	0.6531	20
LPP - LR	0.9360	10	>10	10
LOL - LR	0.2534	50	0.2785	50
SNMF - LR	0.2598	100	0.2888	50
SLMVP - GBR	**0.2446**	**20**	0.2600	50
PCA - GBR	0.2536	20	0.2788	10
LPP - GBR	0.3017	150	0.3586	5
LOL -GBR	0.2900	100	0.2640	150
SNMF - GBR	0.2704	50	0.3021	20
	Lisbon			
SLMVP - LR	0.2845	50	0.3076	100
PCA - LR	0.6048	100	0.6034	10
LPP - LR	3.3840	100	>10	10
LOL - LR	0.2873	20	0.3020	50
SNMF - LR	0.2896	100	0.3090	50
SLMVP - GBR	**0.2732**	**50**	0.2874	100
PCA - GBR	0.2855	10	0.3082	50
LPP - GBR	0.3214	100	0.3884	150
LOL -GBR	0.3278	100	0.3140	100
SNMF - GBR	0.2809	20	0.3228	150

The bold entries are the best model for each case (Seville DNI and Lisbon DNI) independently of the grid size (small or large)

Table 4 Percentage improvement of SLMVP relative to PCA, LPP, LOL, and SNMF

Method	GHI		DNI		
	Small-grid	Large-grid	Small-grid	Large-grid	Avg.
	Seville				
PCA	8.07 %	9.34 %	3.68 %	7.20 %	6.68 %
LPP	28.50 %	57.57 %	23.36 %	37.90 %	34.72 %
LOL	6.24 %	7.14 %	3.60 %	1.52 %	3.95 %
SNMF	5.82 %	13.98%	6.23%	11.07%	8.65%
	Lisbon				
PCA	1.73 %	3.98 %	4.50 %	7.23%	3.87 %
LPP	14.96 %	22.27 %	17.62 %	35.14 %	21.31 %
LOL	2.80 %	7.18 %	5.17 %	5.09 %	4.80 %
SNMF	2.50 %	6.03%	2.81%	7.53%	4.23%

Summarizing the results so far, for both irradiances, GHI and DNI, and both locations (Seville and Lisbon), the best performance is always obtained with the SLMVP method and the non-linear regression method (GBR). In order to quantify this improvement better, Table 4 shows the percentage improvement of SLMVP relative to PCA, LPP, LOL, and SNMF for the best models (SLMVP+GBR, PCA+GBR, LPP+GBR, LOL+LR, and SNMF-GBR/LR). In summary, it can be said that SLMVP offers results 4.92% better than LOL for Seville and around 3.99% than LOL for Lisbon, 5.88% better than PCA for Seville and around 3.12% than PCA for Lisbon, 25.93% better than LPP for Seville and around 16.29% than LPP for Lisbon, 6.21% better than SNMF for Seville and around 2.82% for Lisbon.

In order to visualize the relation between the number of components and error, Fig. 2 shows the GHI validation and test MAE for the different number of components. This is done for SLMVP, PCA, LPP, LOL and SNMF using GBR as regressor, for Seville and Lisbon (top/botton, respectively), and for small and large grids (left/right, respectively). The same information is displayed in Fig. 3 for DNI. It is observed that the best number of PCA components is usually smaller than for other methods and that LPP and LOL usually benefit slightly with larger number of components.

SNMF and SLMVP have similar behavior with the number of components with the optimal number being slightly smaller for SLMVP. In contrast to PCA and LPP,
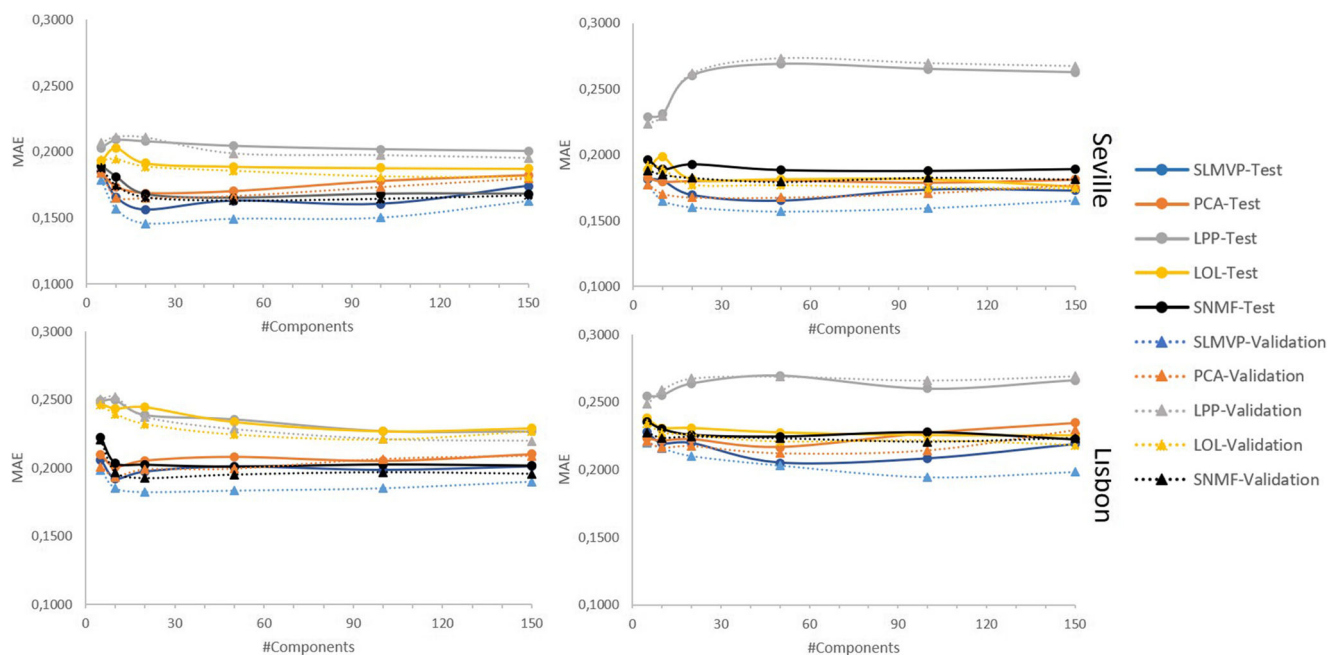
Fig. 2 Average MAE for GHI of SLMVP-GBR, PCA-GBR, LPP-GBR, LOL-GBR, and SNMF-GBR along the number of components (x-axis) for the small and large grids in Seville and Lisbon
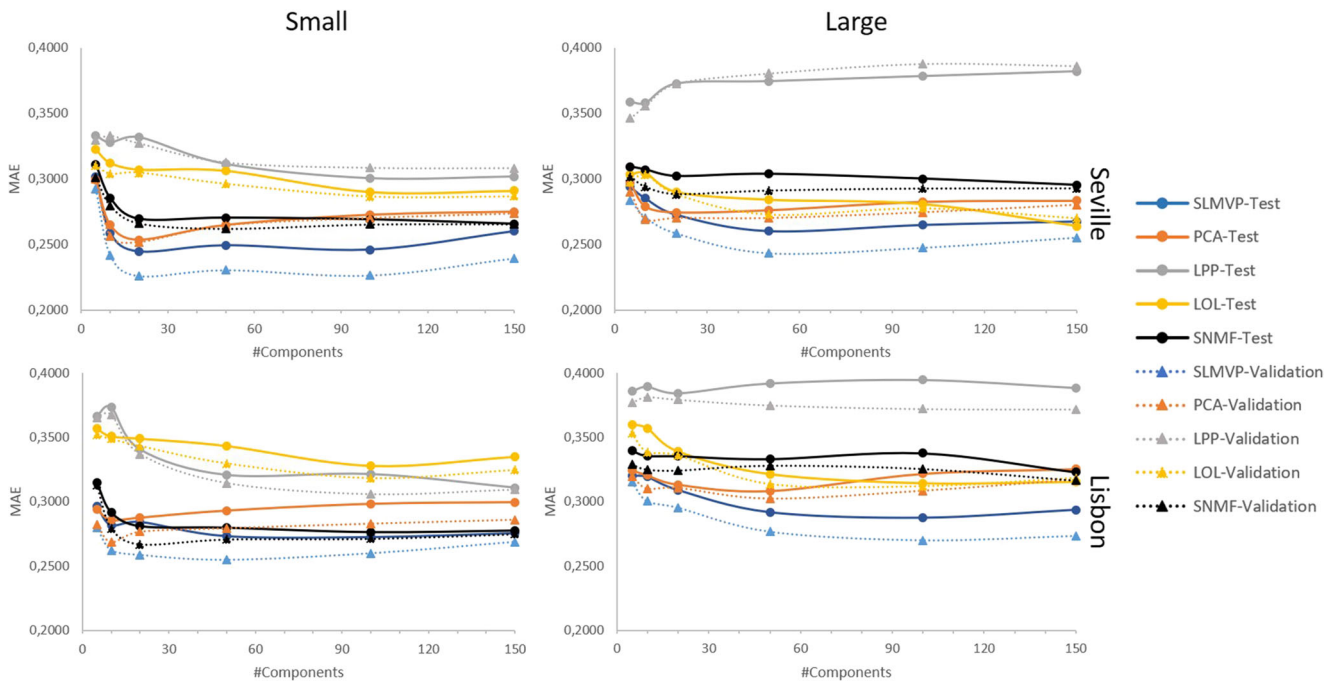
Fig. 3 Average MAE for DNI of SLMVP-GBR, PCA-GBR, LPP-GBR, LOL-GBR, and SNMF-GBR along the number of components (x-axis) for the small and large grids in Seville and Lisbon

the information and components found by SLMVP benefits the performance of the regression method. In Figs. 2 and 3 it is observed that up to 20 components, the errors decrease in all study cases. With a small grid, 20 components is the best solution for all datasets except Lisbon DNI, which reached the best solution with 50 components (see left part in Fig. 2 for GHI and for DNI left part of Fig. 3). When a large grid is used, more than 20 components are generally beneficial, with 50 or 100 components being selected as the best options (50 components for Seville and 100 components for Lisbon, although 50 and 100 components perform similarly for both locations). In those figures, it is also observed that although validation and test errors follow a similar trend, it is not always the case that the best error in validation corresponds to the best error in test. This should be expected because validation error is only an estimation obtained with

a finite independent sample. But at least it can be seen that in all cases, using the validation error to determine the best number of components is a reliable way of achieving a reasonable test error.

Figures 2 and 3 also show that the performance of SLMVP is always better than PCA, LPP, LOL, and SNMF for every number of components (but for a few PCA exceptions and one for SNMF). In order to quantify these improvements, Tables 5, 6, 7 and 8 show the percentage of improvements of SLMVP over PCA, LPP, LOL and SNMF using the best results regression model for the different number of components used, respectively. The superiority of SLMVP is clearly observed, but it is interesting to note that when 5 components are used (and some cases with 10 components), either PCA is better or the improvement of SLMVP is smaller. This suggests that PCA is able to find

Table 5 Improvements in percentage (%) of SLMVP over PCA for the different number of components

Location	Small-grid						Large-grid					
	5	10	20	50	100	150	5	10	20	50	100	150
	GHI											
Seville	-0.47	4.37	6.80	17.49	9.41	4.43	0.58	-0.37	5.87	8.51	2.93	4.65
Lisbon	1.76	4.05	3.83	3.36	3.15	4.29	-1.99	0.76	1.11	4.89	8.11	6.68
	DNI											
Seville	-0.79	2.02	2.98	5.14	8.81	4.93	0.99	-2.14	0.63	5.46	6.02	5.39
Lisbon	-0.84	1.32	1.08	6.64	8.67	8.10	1.35	0.34	1.28	5.14	10.70	9.89

Table 6 Improvements in percentage (%) of SLMVP over LPP for the different number of components

Location	Small-grid						Large-grid					
	5	10	20	50	100	150	5	10	20	50	100	150
	GHI											
Seville	9.66	23.53	28.06	22.46	22.34	14.28	26.14	28.11	50.14	57.16	50.47	49.52
Lisbon	20.91	27.89	20.13	16.80	13.87	12.51	11.75	15.65	19.28	28.31	22.69	20.68
	DNI											
Seville	10.28	22.83	28.82	20.51	18.01	13.77	22.01	24.79	34.08	38.95	38.65	39.01
Lisbon	16.98	24.39	17.96	14.05	12.95	8.86	20.45	21.85	23.48	31.28	33.48	29.59

relevant information when only very few components are allowed. In any case, it is clear from Figs. 2 and 3 that more than 5 components are required in order to obtain the best results.

Finally, to also verify that the SLMVP technique is superior to the current use of PCA, LPP, LOL, and SNMF not only for the optimal number of dimensions but independently of the number of dimensions selected, we have used a two-sample t-test for equal means to test the hypothesis that the obtained average error improvement for the different number of dimensions for each dataset is not due to chance. The obtained significance is shown in Table 9. We applied this test under the null hypothesis that the means are equal and the observations have different standard deviations. We used as observations the 6 test error data results (5, 10, 20, 50, 100, and 150 extracted components) obtained for each dataset . We conclude that the improvement obtained for the analysis SLMVP vs. LPP is significant, rejecting the hypothesis of equal means with a p-value always below < 0.001. Vs. PCA this p-value is below < 0.05 in 4 out of 8 cases (3 for Lisbon and 1 for Seville), vs. LOL the p-value is below < 0.05 also in 4 out of 8 cases (3 for Lisbon and 1 for Seville), and vs. SNMF the p-value is below < 0.05 also in 4 out of 8 cases (2 for Lisbon and 2 for Seville) rejecting the hypothesis of equal means.

In summary, we observed that for Lisbon, the null hypothesis is rejected for 12 out of 16 cases (and the other

two cases have a p-value close to the 5% threshold being the p-value=0.08 and 0.1) and 8 out of 16 for Seville. These insights suggest that the source data for both locations have different properties and Lisbon may contain more noisy data and therefore our method obtains larger improvements because of its noise tolerant characteristics introduced by the use of locality.

5 Conclusions

Using Machine Learning methods to forecast GHI or DNI radiation, based on features that use NWP grids, typically results in a large number of attributes. In this article, a supervised method for feature transformation and reduction (SLMVP) has been proposed to extract the most relevant features solving the limitations of PCA technique to represent locality, non-linear patterns, and use labeled data. The PCA method is one of the most widely used methods to extract features and reduce dimensionality in renewable energy. Three other state-of-the-art dimensionality methods that include locality (LPP), and supervision (LOL and SNMF) have been also compared with.

The five methods have been tested and compared on radiation data at two different Iberian locations: Seville and Lisbon. Both linear and non-linear (GBR) regression

Table 7 Improvements in percentage (%) of SLMVP over LOL for the different number of components

Location	Small-grid						Large-grid					
	5	10	20	50	100	150	5	10	20	50	100	150
	GHI											
Seville	4.32	20.01	19.03	14.00	14.70	7.12	3.98	10.01	6.50	8.86	4.25	1.52
Lisbon	20.04	24.92	22.91	15.91	13.89	13.67	4.45	4.98	4.77	9.83	7.55	2.74
	DNI											
Seville	6.76	17.72	20.71	18.85	14.59	10.21	3.28	6.67	5.95	8.21	5.43	1.18
Lisbon	13.62	16.69	20.80	21.68	15.12	17.03	12.42	11.74	9.30	9.19	8.30	6.70

Table 8 Improvements in percentage (%) of SLMVP over SNMF for the different number of components

	Small-grid						Large-grid					
Location	5	10	20	50	100	150	5	10	20	50	100	150
	GHI											
Seville	2.23	8.44	6.35	1.22	4.13	3.23	7.72	11.30	6.10	9.86	10.74	8.58
Lisbon	7.86	5.36	2.40	-0.10	1.91	0.25	3.37	4.60	2.56	8.53	8.53	1.46
	DNI											
Seville	0.79	2.02	2.98	5.14	8.81	4.93	5.24	7.50	10.03	14.87	11.99	9.44
Lisbon	0.84	1.32	1.08	6.64	8.67	8.10	60.2	5.07	8.20	12.91	15.67	9.12

methods have been used on the components extracted from SLMVP, PCA, LPP, LOL, and SNMF.

Results show that for both types of radiation (GHI and DNI) and both locations, SLMVP offers smaller MAE errors than the other methods. In order to assess the influence of the size of the NWP grid, two sizes have been tested, small and large. SLMVP results in better radiation estimates, but the small size grids display slightly better errors. It has also been shown that PCA tends to underestimate the number of features required to obtain the best results. LPP obtains the worst results and this is noticeable for large grids. SNMF has also shown a degradation in its performance for the large grid compared with SLMVP. In summary, it can be said that the small grid works better and the improvement of SLMVP over the other methods is about 6.24% at Seville GHI, 3.60% at Seville

DNI, 1.73% at Lisbon GHI, and around 4.50% at Lisbon DNI.

Finally, although both SLMVP, PCA, and LPP benefit from using a non-linear regression method (GBR), this benefit is larger for PCA and LPP because they are not able to perform non-linear transformations. LOL does not benefit from non-linear regression and for some cases obtained better results using the regularized linear regressor. SNMF benefits slightly from non-linear regression for all but one of the small grids, but not for large ones. Because SMLVP is able to use non-linear transformations, the difference between using the linear and non-linear regression method is smaller as expected (but still present).

We can conclude that SLMVP is a competitive method for dimensionality reduction in the context of solar radiation forecast using NWP variables beating PCA, which is

Table 9 Dimensionality two-sample t-test analysis for equal means and 5, 10, 20, 50, and 150 dimensions

		GHI		DNI	
		Small-grid	Large-grid	Small-grid	Large-grid
		Seville			
PCA	t-value	1.74	2.59	1.07	1.29
	p-value	0.11	<0.03	0.31	0.23
LPP	t-value	8.19	10.20	5.38	14.36
	p-value	<0.001	<0.001	<0.001	<0.001
LOL	t-value	2.25	2.56	1.55	1.7
	p-value	0.07	<0.003	0.18	0.12
SNMF	t-value	1.01	5.8	1.63	5.08
	p-value	0.34	<0.001	0.13	<0.001
		Lisbon			
PCA	t-value	3.04	1.8	2.83	2.34
	p-value	<0.02	0.1	< 0.02	<0.05
LPP	t-value	8.52	10.65	5.17	13.92
	p-value	<0.001	<0.001	<0.001	<0.001
LOL	t-value	2.55	3.30	10.99	2.06
	p-value	<0.03	<0.02	<0.001	0.08
SNMF	t-value	1.57	2.8	0.88	4.74
	p-value	0.15	<0.02	0.4	<0.001

currently the most widely used, and LOL and SNMF which are two recent supervised dimensionality reduction state-of-the-art methods. Overall SLMVP also obtains better results independently of the number of dimensions used, showing its robustness.

We envision that different machine learning methods would benefit by their combination with SLMVP, and thus it will be of interest to verify it, using this and other domain datasets.

References

1. Yang D, Wang W, Gueymard CA, Hong T, Kleissl J, Huang J, Perez MJ, Perez R, Bright JM, Xia X et al (2022) A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality. Renew Sust Energ Rev 161:112348

2. Haupt SE (2018) Short-range forecasting for energy. Springer, Berlin, pp 97–107. https://doi.org/10.1007/978-3-319-68418-5_7

3. Sobri S, Koohi-Kamali S, Rahim NA (2018) Solar photovoltaic generation forecasting methods: A review. Energy Convers Manag 156:459–497

4. Yang D, Kleissl J, Gueymard CA, Pedro HTC, Coimbra CFM (2018) History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining. Sol Energy 168:60–101. https://doi.org/10.1016/j.solener.2017.11.023

5. Singla P, Duhan M, Saroha S (2021) A comprehensive review and analysis of solar forecasting techniques. Frontiers in Energy, pp 1–37

6. Litjens GBMA, Worrell E, van Sark WGJHM (2018) Assessment of forecasting methods on performance of photovoltaic-battery systems. Appl Energy 221:358–373. https://doi.org/10.1016/j.apenergy.2018.03.154

7. Agüera-Pérez A, Palomares-Salas JC, González de la Rosa JJ, Florencias-Oliveros O (2018) Weather forecasts for microgrid energy management: Review, discussion and recommendations. Appl Energy 228(C):265–278. https://doi.org/10.1016/j.apenergy.2018.0

8. Dersch J, Schroedter-Homscheidt M, Gairaa K, Hanrieder N, Landelius T, Lindskog M, Müller SC, Ramirez Santigosa L, Sirch T, Wilbert S (2019) Impact of dni nowcasting on annual revenues of csp plants for a time of delivery based feed in tariff. Meteorol Z 28(3):235–253. https://doi.org/10.1127/metz/2019/0925

9. Alonso-Montesinos J, Polo J, Ballestrín J, Batlles FJ, Portillo C (2019) Impact of DNI forecasting on CSP tower plant power production. Renew Energy 138(C):368–377. https://doi.org/10.1016/j.renene.2019.01

10. Antonanzas J, Pozo-Vázquez D, Fernand ez-Jimenez LA, Martinez-de-Pison FJ (2017) The value of day-ahead forecasting for photovoltaics in the Spanish electricity market. Sol Energy 158:140–146. https://doi.org/10.1016/j.solener.2017.09.043

11. Blanc P, Remund J, Vallance L (2017) Short-term solar power forecasting based on satellite images, pp 179–198. https://doi.org/10.1016/B978-0-08-100504-0.00006-8

12. Bright JM, Killinger S, Lingfors D, Engerer NA (2018) Improved satellite-derived pv power nowcasting using real-time power data from reference pv systems. Sol Energy 168:118–139

13. Arbizu-Barrena C, Ruiz-Arias JA, Rodríguez-Benítez FJ, Pozo-Vázquez D, Tovar-Pescador J (2017) Short-term solar radiation forecasting by advecting and diffusing msg cloud index. Sol Energy 155:1092–1103. https://doi.org/10.1016/j.solener.2017.

14. Lopes FM, Silva HG, Salgado R, Cavaco A, Canhoto P, Collares-Pereira M (2018) Short-term forecasts of ghi and dni for solar energy systems operation: assessment of the ecmwf integrated forecasting system in southern portugal. Sol Energy 170:14–30

15. Rodríguez-Benítez FJ, Arbizu-Barrena C, Huertas-Tato J, Aler-Mur R, Galván-León I, Pozo-Vázquez D (2020) A short-term solar radiation forecasting system for the iberian peninsula. Part 1: Models description and performance assessment. Sol Energy 195:396–412. https://doi.org/10.1016/j.solener.2019.11.028

16. McCandless TC, Haupt SE, Young GS (2016) A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. Renew Energy 89(C):351–359. https://doi.org/10.1016/j.renene.2015.12

17. Lee JA, Haupt SE, Jiménez PA, Rogers MA, Miller SD, McCandless TC (2017) Solar irradiance Nowcasting case studies near sacramento. J Appl Meteorol Climatol 56(1):85–108. https://doi.org/10.1175/JAMC-D-16-0183.1

18. Ahmed R, Sreeram V, Mishra Y, Arif M (2020) A review and evaluation of the state-of-the-art in pv solar power forecasting: Techniques and optimization. Renew Sust Energ Rev 124: 109792

19. Yang D, Wang W, Bright JM, Voyant C, Notton G, Zhang G, Lyu C (2022) Verifying operational intra-day solar forecasts from ecmwf and noaa. Sol Energy 236:743–755

20. Mellit A, Massi Pavan A, Ogliari E, Leva S, Lughi V (2020) Advanced methods for photovoltaic output power forecasting: A review. Appl Sci 10(2):487

21. Markovics D, Mayer MJ (2022) Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. Renew Sust Energ Rev 161:112364

22. Salcedo-Sanz S, Cornejo-Bueno L, Prieto L, Paredes D, García-Herrera R (2018) Feature selection in machine learning prediction systems for renewable energy applications. Renew Sust Energ Rev 90:728–741

23. Liu H, Chen C (2019) Data processing strategies in wind energy forecasting models and applications: A comprehensive review. Appl Energy 249:392–408

24. Martin R, Aler R, Valls JM, Galván IM (2016) Machine learning techniques for daily solar energy prediction and interpolation using numerical weather models. Concurr Comput Pract Exp 28(4):1261–1274

25. Wang Z, Wang W, Wang B (2017) Regional wind power forecasting model with nwp grid data optimized. Front Energy 11(2):175–183

26. Andrade JR, Bessa RJ (2017) Improving renewable energy forecasting with a grid of numerical weather predictions. IEEE Trans Sustain Energy 8(4):1571–1580

27. Higashiyama K, Fujimoto Y, Hayashi Y (2017) Feature extraction of numerical weather prediction results toward reliable wind power prediction. In: 2017 IEEE PES Innovative smart grid technologies conference europe (ISGT-Europe), pp 1–6. IEEE

28. García-Hinde O, Terrén-Serrano G, Hombrados-Herrera M, Gómez-Verdejo V, Jiménez-Fernández S, Casanova-Mateo C, Sanz-Justo J, Martínez-Ramón M, Salcedo-Sanz S (2018) Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting. Eng Appl Artif Intell 69:157–167

29. Verbois H, Huva R, Rusydi A, Walsh W (2018) Solar irradiance forecasting in the tropics using numerical weather prediction and statistical learning. Sol Energy 162:265–277

30. Khan M, Liu T, Ullah F (2019) A new hybrid approach to forecast wind power for large scale wind turbine data using deep learning with tensorflow framework and principal component analysis. Energies 12(12):2229

31. Verbois H, Saint-Drenan Y-M, Thiery A, Blanc P (2022) Statistical learning for nwp post-processing: a benchmark for solar irradiance forecasting. Sol Energy 238:132–149

32. Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24(6):417–441. https://doi.org/10.1037/h0071325

33. Fisher RA (1925) Theory of statistical estimation. Math Proc Philos Soc 22:700–725

34. García-Cuesta E, Iglesias JA (2012) User modeling: through statistical analysis and subspace learning. Expert Syst Appl 39(5):5243–5250

35. McInnes L, Healy J, Saul N, Großberger L (2018) Umap: Uniform manifold approximation and projection. J Open Source Softw 3(861)

36. He X, Niyogi P (2003) Locality preserving projections. Advances in neural information processing systems, p 16

37. Vogelstein JT, Bridgeford EW, Tang M et al (2021) Supervised dimensionality reduction for big data. Nat Commun 12(2872). https://doi.org/10.1038/s41467-021-23102-2

38. Chao G, Mao C, Wang F, Zhao Y, Luo Y (2018) Supervised nonnegative matrix factorization to predict icu mortality risk. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp 1189–1194. IEEE

39. Chao G, Luo Y, Ding W (2019) Recent advances in supervised dimension reduction: a survey. Mach Learn Knowl Extract 1(1):341–358

40. Tenenbaum JB, Silva VD, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290(5500):2319–2323. https://doi.org/10.1126/science.290.5500.2319

41. Belkin M, Niyogi P (2003) Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput 15(6):1373–1396. https://doi.org/10.1162/089976603321780317

42. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500):2323–2326

43. Weinberger KQ, Sha F, Saul LK (2004) Learning a kernel matrix for nonlinear dimensionality reduction. In: Proceedings of the Twenty-first international conference on machine learning, p 106

44. García-Cuesta E (2022) Supervised Local Maximum Variance Preserving (SLMVP) Dimensionality Reduction Method (1.0). https://doi.org/10.5281/zenodo.6856079, Online; Accessed 18 July 2022

45. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Annals of statistics, pp 1189–1232

46. Friedman JH (2002) Stochastic gradient boosting. Comput Stat Data Anal 38(4):367–378

47. Aler R, Galván IM, Ruiz-Arias JA, Gueymard CA (2017) Improving the separation of direct and diffuse solar radiation components using machine learning by gradient boosting. Sol Energy 150:558–569

48. Wu J, Zhou T, Li T (2020) Detecting epileptic seizures in eeg signals with complementary ensemble empirical mode decomposition and extreme gradient boosting. Entropy 22(2):140

49. Asante-Okyere S, Shen C, Ziggah YY, Rulegeya MM, Zhu X (2019) A novel hybrid technique of integrating gradient-boosted machine and clustering algorithms for lithology classification. Natural Resources Research, pp 1–17

Esteban García is assistant professor with the Artificial Intelligence department at Universidad Politécnica de Madrid since 2001. He received his PhD in Computer Science from Universidad Carlos III de Madrid (Spain) in 2010 and his MSc in Computer Science from Universidad Carlos III (2005). He has worked in European and Spanish research projects, related to machine learning and artificial intelligence on several application domains, including remote sensing, renewable energy forecasting, affective computing, and medicine. His current research interests include machine learning and explainable AI. He has been visiting scientist at Carnegie Mellon (Pittsburgh, US) and collaborates actively with industry initiatives.

Ricardo Aler is associate professor with the Computer Science Department at Universidad Carlos III de Madrid since 2001. He received his PhD in Computer Science from Universidad Politécnica de Madrid (Spain) in 1999 and his MSc in Decision Support Systems from Sunderland University (UK) (1993). He has worked in European and Spanish projects, related to evolutionary computation and machine learning on several application domains, including telecommunications, robosoccer, brain-computer interfaces, and renewable energy forecasting. His current research interests include Machine Learning, Evolutionary Optimization, and the Energy Forecasting field.

David Pozo-Vázquez holds a Ph. D in Atmospheric Science (2000) from University of Granada (Spain) and B.S. in Applied Physics from the same University (1994). Since 1998 he is on the Faculty of the Department of Physics of the University of Jaen, where he is responsible for courses on meteorology and renewable energy resources and leads the Solar Radiation and Atmosphere Modeling research group. He obtained a permanent appointment (Associated Professor) in 2003 and is full Professor since September 2018. In the last decade, his research focused on the solar and wind energy resources assessment, including their spatial and temporal balancing at different spatial and temporal scales. In addition, he conducted research aimed at the improvement of solar radiation forecasting techniques at different spatial scales (plant, utility scale) and forecasting lead times (from minutes to hours and days). He has been visiting scientist at the University of East Anglia, the European Center for Medium Range Weather Forecasting (both in the U.K.) and the National Center for Atmospheric Research (Boulder, CO, USA).

Inés M. Galván is full professor at the Computer Science and Engineering Department at Carlos III University of Madrid since 2020. She received her PhD degree in Computer Science at Universidad Politécnica de Madrid (Spain), in 1998. She has worked in several research European and Spanish projects related with control of chemical reactors, optimization and evolutionary computation, and solar radiation forecasting. Her current research focuses on Machine Learning Techniques, Evolutionary Computation and Multi-objective algorithms. Her research interests also cover different applications fields, as control of dynamic process, times series prediction, probabilistic forecasting, and renewable energy.

Affiliations

Esteban García-Cuesta[1] (iD) · Ricardo Aler[2] · David del Pózo-Vázquez[3] · Inés M. Galván[2]

Ricardo Aler
aler@inf.uc3m.es

David del Pózo-Vázquez
dpozo@ujaen.es

Inés M. Galván
igalvan@inf.uc3m.es

[1] Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, E.T.S.I.I Campus de Montegancedo s/n, Boadilla del Monte, 28660, Madrid, Spain

[2] Departamento de Informática, Universidad Carlos III de Madrid, Av. de la Universidad, 30, Leganés, 28911, Madrid, Spain

[3] Departamento de Física, Universidad de Jaen, Campus Las Lagunillas, s/n, Jaen, 23071, Jaen, Spain