

Technical report 18-018

# **A Generalized Model for Short-Term Forecasting of Solar Irradiance\***

J. Lago, K. De Brabandere, F. De Ridder, and B. De Schutter

*To cite this work, please refer to the published version:*

J. Lago, K. De Brabandere, F. De Ridder, and B. De Schutter, "A generalized model for short-term forecasting of solar irradiance," *Proceedings of the 57th IEEE Conference on Decision and Control*, Miami Beach, Florida, pp. 3165–3170, Dec. 2018. doi:[10.1109/CDC.2018.8618693](https://doi.org/10.1109/CDC.2018.8618693)

Delft Center for Systems and Control  
Delft University of Technology  
Mekelweg 2, 2628 CD Delft  
The Netherlands  
phone: +31-15-278.24.73 (secretary)  
URL: <https://www.dcsc.tudelft.nl>

---

\* This report can also be downloaded via <https://dpub.eu/18-018>

# A generalized model for short-term forecasting of solar irradiance

Jesus Lago<sup>1,2</sup>, Karel De Brabandere<sup>3</sup>, Fjo De Ridder<sup>2</sup>, and Bart De Schutter<sup>1</sup>

**Abstract**—In recent years, as the share of solar power in the electrical grid has been increasing, accurate methods for forecasting solar irradiance have become necessary to manage the electrical grid. More specifically, as solar generators are geographically dispersed, it is very important to have general models that can predict solar irradiance without the need of ground data. In this paper, we propose a novel technique that can accomplish that: using satellite images, the proposed model is able to forecast solar irradiance without the need of ground measurements. To illustrate the performance of the proposed model, we consider 15 locations in The Netherlands, and we show that the proposed model is as accurate as local models that are individually trained with ground data.

## I. INTRODUCTION

As the share of renewable energy sources in the electrical grid has been steadily increasing in recent years, forecasting their production has become key to safely manage the electricity grid. In particular, since the generation from renewable sources is intermittent and highly unpredictable, forecasting accurately this generation is critical to balance the electrical grid and to keep the system stable [1]. In this context, as the use of solar energy has dramatically increased, forecasting the *global horizontal irradiance (GHI)* has become a key element in safely operating the electrical grid.

The forecasting techniques in the literature for forecasting the GHI are typically divided into two categories according to the input data and the forecast horizon [1], [2]:

- 1) *Time series models*, which use sky/satellite images and/or ground measurements and are only suitable for short-term forecasts up to 4-6 hours. This category is further divided in three groups:
  - a) Classical statistical models, e.g. ARIMA models [3].
  - b) Machine learning models, e.g. neural networks [4].
  - c) Cloud-moving vector models that use satellite images [5].
- 2) *Numerical weather prediction (NWP)* models, which simulate weather conditions and outperform statistical models for longer forecast horizons, i.e. 4-6 hours onward [6].

<sup>1</sup>Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands, jlagogarcia@tudelft.nl.

<sup>2</sup>Algorithms, Modeling, and Optimization, VITO, Energyville, ThorPark, Genk, Belgium.

<sup>3</sup>3E, Brussels, Belgium.

This research has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 675318 (INCITE).

To the best of our knowledge, all the statistical and artificial intelligence proposed in the literature have considered ground measurements of the solar irradiance as input regressors. Considering the geographical dispersion of solar generators, this means that these methods need to gather local data across several geographical locations. Therefore, if these methods are to be used, the cost of forecasting solar irradiance can potentially become very expensive as the number of local sensors and data gathering points grows very large.

To address this issue and to obtain scalable solutions for solar irradiance forecasting, global models that can forecast the GHI without the need of ground data are needed. While current cloud-moving vectors could accomplish that, they are much more computationally intensive and they do not obtain more accurate predictions than local models that use ground data [2].

In this paper, we address the mentioned problem by proposing a novel forecasting technique that, using SEVIRI<sup>1</sup> satellite images, forecasts solar irradiance without the need of local data. The model is based on a *deep neural network (DNN)*, i.e. a neural network that uses more than one hidden layer and employs state of the art algorithms and functions from the field of deep learning. While the model considers satellite images just as cloud-moving vector models do, is easier to deploy as it is not computationally intensive.

To analyze and illustrate the performance of the proposed model, we consider 15 locations in The Netherlands: we employ 5 locations for estimating the model and the remaining 10 locations to evaluate its performance. In particular, we compare the performance of the model against the performance of literature models that are individual trained for each of the 10 sites using ground data. As we show, the proposed method is as accurate as the literature models that consider ground measurements.

The remainder of the paper is organized as follows: Section II introduces the background and preliminary concepts. Then, Section III presents the proposed model for irradiance forecasting. Next, Section IV introduces the case study and analyzes the performance of the model when compared with local models. Finally, Section V summarizes the results and concludes the paper.

## II. PRELIMINARIES

In this section we introduce the concepts and algorithms that are used in the paper.

<sup>1</sup>The SEVIRI (Spinning Enhanced Visible and InfraRed Imager) is a measurement instrument of the METEOSAT satellite.

### A. Deep Neural Networks

During the last decade, deep learning architectures [7] have become the state of the art learning techniques across several applications, e.g. image recognition [8] or speech recognition [9]. More recently, their benefits have also spread to several energy-related applications [10], [11], [12], [13]. For our application, we consider one of these deep architectures as a base model. While there exist different deep learning architectures, e.g. convolutional networks or recurrent networks, we consider a *deep neural network (DNN)*, i.e. a multilayer perceptron with more than one hidden layer. This selection is done because DNNs are less computationally intensive than the other DL architectures [7] and because DNNs have empirically outperformed the other DL architectures in similar energy-based forecasts [13].

Defining by  $\mathbf{X} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$  the input of the network, by  $\mathbf{Y} = [y_1, y_2, \dots, y_m]^T \in \mathbb{R}^m$  the output of the network, by  $n_i$  the number of neurons of the  $i^{\text{th}}$  hidden layer, and by  $\mathbf{z}_i = [z_{i1}, \dots, z_{in_i}]^T$  the state vector in the  $i^{\text{th}}$  hidden layer, a general DNN with two hidden layers can be represented as in Figure 1.

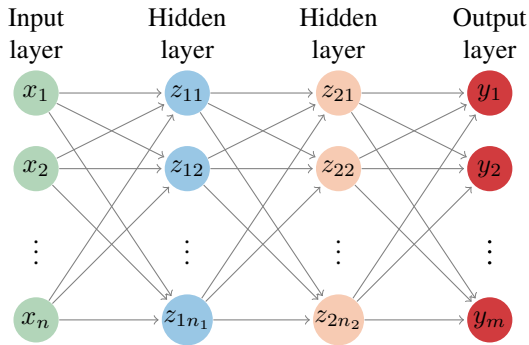


Fig. 1: Example of a DNN.

In this representation, the parameters of the model are represented by the set of weights  $\mathbf{W}$  that establish the mapping connections between the different neurons of the network. Moreover, in addition to the weights, the network has other parameters that need to be selected before the training process, e.g. the number of neurons of the hidden layers or the number of hidden layers. To distinguish them from the main parameters they are called hyperparameters.

### B. Hyperparameter Optimization and Feature Selection

In this paper, to perform the hyperparameter selection, we consider a Bayesian optimization algorithm that has been widely used in the machine learning community: the *tree-structured Parzen estimator (TPE)* [14], an optimization algorithm within the family of sequential model-based optimization methods.

In addition to optimizing the hyperparameters, the TPE algorithm is also employed for the selection of input features, i.e. for choosing the best subset of regressors to forecast the quantity of interest given the full set of available input data. In particular, using the feature selection method proposed in

[12], the TPE algorithm is employed to select the optimal historical data for forecasting the solar irradiance.

### C. Performance Metrics

In order to evaluate the accuracy of the proposed model, following the standards of the literature of solar irradiance forecasting, we consider the *relative root mean square error (rRMSE)*. Given a vector  $\mathbf{Y} = [y_1, \dots, y_N]^T$  of real outputs and a vector  $\hat{\mathbf{Y}} = [\hat{y}_1, \dots, \hat{y}_N]^T$  of predicted outputs, the rRMSE metric can be computed as:

$$\text{rRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}}{\frac{1}{N} \sum_{i=1}^N y_i} \cdot 100 \quad (1)$$

## III. PREDICTION MODEL

In this section, the proposed prediction model for solar irradiance forecasting is presented. In particular, the model structure, input features, model hyperparameters, and training algorithms are described.

### A. Model Structure

In order to forecast solar irradiance without ground measurements and across multiple locations, a model that can generalize across geographical areas is required. As DNNs are powerful models that can generalize across tasks [7], they are selected as the base model for the proposed forecaster. In particular, the proposed model is a DNN that consists of 6 output neurons representing the forecasted hourly irradiance over the next 6 hours. This horizon is the standard choice for short-term irradiance forecasting as after 6 hours NWP forecasts outperform time series models [2].

To have a model that can re-adapt to the geographical area where the forecasts are made, the proposed DNN is not subject to any specific depth nor number of neurons. Instead, depending on the geographical area under study, the number of hidden layers and neurons are optimized using hyperparameter optimization (see Section II-B).

### B. Model Inputs

The proposed model considers three types of possible input features: NWP forecasts of the solar irradiance, clear-sky irradiance values, and satellite images representing maps of past solar irradiance.

1) *Numerical weather prediction forecast*: NWP forecasts of the solar irradiance obtained from the *European center for medium-range weather forecasts (ECMWF)*. For the proposed model, the input data consists of the 6 forecasted values for the next 6 hours.

2) *Clear-sky irradiance*: The GHI under clear-sky conditions over the next 6 hours. These are 6 deterministic inputs that are obtained using the clear-sky model of [15].

3) *SEVIRI images*: Satellite data that represents the past irradiance values of a geographical area. In particular, the input data consists of images from the SEVIRI instrument of the METEOSAT satellite that are transformed to irradiance values using two different methods:

- 1) The *Surface insolation under clear and cloudy skies (SICSS)* algorithm [16] for data corresponding to solar elevation angles above  $12^\circ$ .
- 2) The interpolation method described in [17] for data corresponding to solar elevation angles below  $12^\circ$ , i.e. very early in the morning and late in the evening.

It is important to note that this selection of input features was done for a clear reason: as the model should be generalized across geographical locations, it is critical to consider input features that can be easily obtained for any given location. Considering this restriction, the NWP forecasts, the clear-sky irradiance, and the satellite images are excellent input features as they can be easily obtained for any area.

Another remark to be made regards the challenging nature of the problem we consider. In particular, the irradiance values obtained from the satellite images have a resolution of  $3 \times 3$  km and only represent the average irradiance in a  $9 \text{ km}^2$  area. As a result, when using these input features for forecasting the solar irradiance in a specific location, they can hardly be as accurate as ground measurements. Based on this premise, it becomes clear that building a global model as accurate as a local models is not a trivial task.

### C. Hyperparameter Optimization and Feature Selection

As indicated previously, the proposed model needs to be tuned for the specific geographical area where it is applied. To do so, three hyperparameter are optimized:

- i) Number of hidden layers.
- ii) Number of neurons per layer.
- iii) Dropout [18]: this is a regularization hyperparameter to reduce overfitting. The hyperparameter is defined by a real number between 0 and 1.

In combination with the hyperparameter optimization, the proposed model also involves a feature selection. More specifically, the feature selection method selects which and how many past historical irradiance values are needed to obtain the most accurate model.

### D. Training

The DNN is trained by minimizing the mean square error<sup>2</sup>. In particular, the optimization problem that is solved to train the neural network is:

$$\underset{\mathbf{W}}{\text{minimize}} \quad \sum_{i=1}^N \|\hat{\mathbf{I}}_i - F(\mathbf{X}_i, \mathbf{W})\|_2^2, \quad (2)$$

where  $\mathcal{S}_{\mathcal{T}} = \{(\mathbf{X}_i, \hat{\mathbf{I}}_i)\}_{i=1}^N$  is the training dataset,  $\hat{\mathbf{I}}_i$  the  $i^{\text{th}}$  irradiance value of the training dataset,  $\mathbf{X}_i$  the input vector

<sup>2</sup>Note that minimizing the mean square error is equivalent to minimizing the rRMSE metric used throughout the paper to evaluate and compare models.

used to predict  $\hat{\mathbf{I}}_i$ ,  $\mathbf{W}$  the set of network parameters, and  $F: \mathbb{R}^n \rightarrow \mathbb{R}^6$  the neural network map.

To solve (2), we use multi-start optimization and Adam [19], a version of the stochastic gradient descent method that uses adaptive learning rates for each model parameter. In addition, early stopping [20] is also considered to avoid overfitting.

### E. Generalizing across geographical sites

In order for the model to forecast without ground data, it is paramount that the model generalizes across geographical locations. To build this generalization capability, the training is performed across a small subset of sites so that the model learns to generalize across geographical areas.

## IV. CASE STUDY

To analyze the proposed general model, we consider 15 sites in the Netherlands. More specifically, to evaluate its performance, we compare its accuracy with that of local models from the literature that are individually trained for each specific location using local data.

### A. Data description

The dataset spans four years, i.e. from 01/01/2014 until 31/12/2017, and comprises four types of input data:

- i) Historical ground data measured on site.
- ii) Satellite-based irradiance values.
- iii) Daily ECMWF forecasts.
- iv) Deterministic clear-sky irradiance values.

In all the four cases, these data represent hourly average values between two consecutive hours, i.e. a variable given at a time step  $h$  represents the average variable between hours  $h$  and  $h + 1$ .

1) *Geographical Locations*: To select the geographical locations, we consider 15 meteorological stations in The Netherlands that are maintained by the *Royal Netherlands Meteorological Institute (KNMI)* [21]. In particular, we consider the following 15 locations: Arcen, Berkhout, Cabauw, De Kooy, Deelen, Eindhoven, Gilze-Rijen, Herwijnen, Lauwersoog, Lelystad, Maastricht, Schiphol, Twenthe, Westdorpe, and Wijk aan Zee.

2) *Data Sources*: The source used for the input features depends on the type of input:

- i) The irradiance values obtained from SEVIRI satellite images are obtained from the KNMI via their Cloud Physical Properties model [21] directly processed as irradiance values.
- ii) The ground measurements are obtained through pyranometer readings at the weather stations [22].
- iii) The ECMWF forecasts are directly obtained through the ECMWF website [23].
- iv) The clear-sky irradiance is obtained through the `python PVLIB` library [24] that implements the clear-sky model of [15].

3) *Data division*: To perform the study, the data is divided into three subsets:

- i) A 2-year training set (01/01/2014 to 31/12/2015) for estimating the various models.
- ii) A 1-year validation set (01/01/2016 to 31/12/2016) for selecting the optimal hyperparameters/input features and for monitoring early-stopping.
- iii) A 1-year test set (01/01/2017 to 31/12/2017) as the out-of-sample dataset to evaluate the proposed model and to compare its performance with that of local models.

Moreover, the data is also divided according to the location:

- i) Of the 15 sites, 5 are randomly selected to train the proposed model: Herwijnen, Wijk aan Zee, Schiphol, Twenthe, and Lelystad.
- ii) The remaining 10 act as out-of-sample data to evaluate the models.

4) *Data Preprocessing*: We disregard the data corresponding to the hours of the day for which during parts of the year the irradiance is zero, i.e. we limit the forecasts to predict solar irradiance from 8:00 to 19:00. As the prediction horizon is 6 hours ahead, this implies that we can evaluate the model on 6 forecast per day.

### B. Local models

To compare the proposed forecaster, we consider three types of local models from the literature: a persistence model [2], an *autoregressive model with exogenous inputs* (ARX) [25], and a local neural network [25].

1) *Persistence model*: A standard procedure in the literature of irradiance forecasting is to check whether new models provide better predictions than a trivial model [2]. Normally, the trivial model that is usually considered is the persistence model [2], a forecasting technique that assumes that the irradiance at the prediction time  $h+p$  is equal to the current irradiance but scaled by the irradiance diurnal cycle.

2) *Linear model*: Another standard benchmark in the literature of irradiance forecasting are linear autoregressive models [2], [25]. In this case, the exogenous inputs of this model are similar to the global model; however, instead of using the satellite irradiance maps  $\mathcal{I}_S$ , the model considers the historical irradiance ground measurements  $\mathcal{I}_G$ . As we evaluate the local models in 10 locations, we make 6 forecasts per day, and each forecast is made for 6 prediction times; so we estimate a total of  $10 \times 6 \times 6 = 360$  linear autoregressive models. For each of the 360 models, the number of historical  $\mathcal{I}_G$  values is optimally selected using the feature selection method that was also employed for the global DNN.

3) *Neural network*: A third benchmark standard in the literature of solar irradiance forecasting are neural networks [1], [17]. Thus, as a third local model, we consider a local DNN with a very similar structure as the proposed global DNN. The main difference w.r.t. to the proposed DNN is that the local DNN considers the local measurements of the irradiance  $\mathcal{I}_G$  instead of the satellite irradiance maps  $\mathcal{I}_S$ . As the model is local and we evaluate the proposed DNN in

10 sites, we estimate 10 different local DNNs. For each of these 10 local DNNs, the number of historical  $\mathcal{I}_G$  values and the mode hyperparameters are optimally selected using the optimization method that was also employed for the global DNN.

### C. Hyperparameter Optimization and Feature Selection

As defined in Section III, the hyperparameters and input features of the global DNN are optimally selected according to the geographical location. For our case study, the obtained optimal values are listed in Table I.

Hyperparameter	Value
Number of hidden layers	2
Neurons in 1 <sup>st</sup> layer	208
Neurons in 2 <sup>nd</sup> layer	63
Dropout	0.14

TABLE I: Optimal hyperparameters for the global DNN.

In terms of the historical satellite irradiance values  $\mathcal{I}_S$ , the optimal input features are defined by the irradiance values at lags 0, 1, 2, and 3 w.r.t. the current hour  $h$ ; i.e.  $\mathcal{I}_{S,h}, \dots, \mathcal{I}_{S,h-3}$ ; and at lag 24 w.r.t the 6 prediction hours  $h+1, \dots, h+6$ ; i.e.  $\mathcal{I}_{S,h-23}, \dots, \mathcal{I}_{S,h-18}$ .

Note that while the hyperparameters and input features for each local model are also optimized, this optimization is done for 360 linear models and 10 local DNNs. Therefore, displaying these hyperparameter and feature selection results cannot be done within the scope of this paper.

### D. Representation

Considering the obtained optimal hyperparameters and input features, the global DNN for this case study can be represented as in Figure 2. In this representation, the current hour is defined by  $h$ , the values of the ECMWF forecast by  $\hat{\mathcal{I}}_E$ , the clear-sky irradiance by  $\mathcal{I}_c$ , and the forecasted values of the proposed model by  $\hat{\mathcal{I}}$ .

### E. Results and Discussion

Now we compare the performance of the global DNN against the local models. As can be observed from Table II, which compares the average rRMSE across the 10 sites and the 6 prediction times, the proposed global model performs slightly better than the local linear model and the local DNN and significantly better than the persistent model.

Model	rRMSE [%]
Global DNN	31.66
Linear	32.31
Local DNN	32.56
Persistence	42.36

TABLE II: Comparison of the average prediction accuracy across sites and prediction times by means of rRMSE.

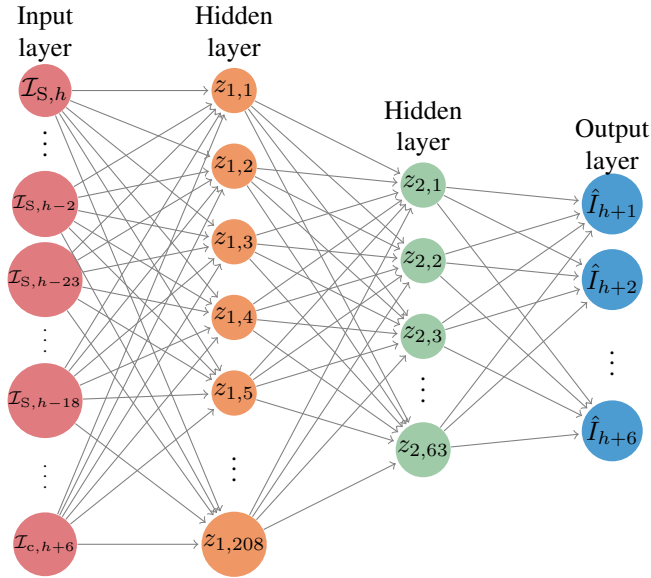


Fig. 2: Generalized DNN to forecast solar irradiance without ground data.

Considering these results we can observe that, in terms of overall performance, the proposed global model appears to be an excellent candidate to replace the local models. Nevertheless, before establishing this empirical observation, we need to ensure that the quality of its performance is kept across the 10 different sites. More specifically, it is paramount to check whether the global model can generalize across geographical areas without using ground data, i.e. whether it can keep the superior prediction accuracy at each of the 10 locations. Such a comparison across the individual locations is displayed in Table III and Figure 3. As can be observed, the global model does indeed obtain predictions that are equal or better than the local models across all 10 sites. Particularly, the global DNN is the best in 8 of the 10 locations and performs very similar to the best model in the remaining 2 locations.

As a consequence, since the global model performs better than or equally well as the local models across all individual locations, we can conclude that the proposed global model is an excellent replacement for the local models. In particular, as the global model does not require ground measurements, it has the potential to save the operational costs of local models without reducing their forecasting accuracy.

## V. CONCLUSION

We have proposed a general model for short-term forecasting of the solar irradiance. The end goal of the model was to provide a novel technique to forecast the solar irradiance that, unlike local models previously proposed in the literature, did not require ground measurements. The motivation behind building such a model was the fact that, as solar generators are geographical disperse, local models incur large operational costs as local sensors have to be installed and maintained at every generation site.

The proposed model successfully replaces ground measurements by using satellite-based irradiance values, numer-

ical weather forecasts, and an architecture based on a deep neural network. In particular, using 10 different geographical locations in The Netherlands, the accuracy of the proposed model was shown to be equal or better than that of local models proposed in the literature. Based on this result, we can conclude that the proposed model is an excellent replacement for the local models in order to save the operational costs of installing local sensors and gathering ground data.

In future work, the proposed model will be evaluated in other regions to analyze whether the model generalizes to larger geographical areas. Moreover, the comparison of the model will be extended to incorporate more local models from the literature.

## REFERENCES

- [1] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569–582, 2017.
- [2] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.
- [3] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," *Solar Energy*, vol. 83, no. 3, pp. 342–349, 2009.
- [4] A. Mellit and A. M. Pavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy," *Solar Energy*, vol. 84, no. 5, pp. 807–821, 2010.
- [5] E. Lorenz and D. Heinemann, "Prediction of solar irradiance and photovoltaic power," in *Comprehensive Renewable Energy*, A. Sayigh, Ed. Elsevier, 2012, pp. 239–292.
- [6] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, D. Renné, and T. E. Hoff, "Validation of short and medium term operational solar radiation forecasts in the US," *Solar Energy*, vol. 84, no. 12, pp. 2161–2172, 2010.
- [7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org/>.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, ser. NIPS'12. USA: Curran Associates Inc., 2012, pp. 1097–1105.

Model	Site									
	Arcen	Berkhout	Cabauw	De Kooy	Lauwersoog	Deelen	Maastricht	Eindhoven	Westdorpe	Gilze-Rijen
Global	32.55	30.39	30.90	29.63	30.47	34.71	30.97	32.26	32.22	32.52
Linear	33.18	31.20	31.16	30.01	31.31	35.64	31.89	32.44	33.27	33.05
DNN	33.59	33.14	31.42	31.34	31.09	35.93	31.63	32.19	32.08	33.20
Persistence	43.84	41.24	41.72	41.38	41.34	45.70	41.40	43.41	40.47	43.07

TABLE III: Comparison of the prediction accuracy of the various forecasters by means of rRMSE. The best and worst models are respectively marked with green and red cells.

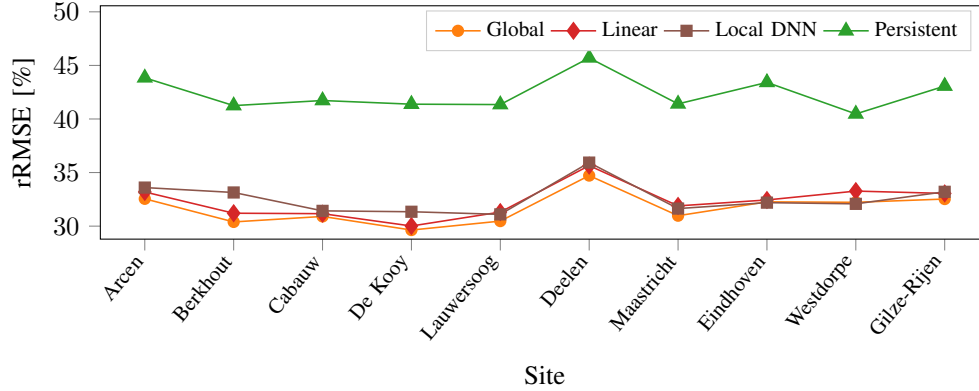


Fig. 3: Comparison of the prediction accuracy of the various forecasters across the 10 locations.

- [9] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.
- [10] I. Coelho, V. Coelho, E. Luz, L. Ochi, F. Guimarães, and E. Rios, "A GPU deep learning metaheuristic based model for time series forecasting," *Applied Energy*, vol. 201, pp. 412–418, 2017.
- [11] C. Fan, F. Xiao, and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Applied Energy*, vol. 195, pp. 222–233, 2017.
- [12] J. Lago, F. De Ridder, P. Vranx, and B. De Schutter, "Forecasting day-ahead electricity prices in Europe: The importance of considering market integration," *Applied Energy*, vol. 211, pp. 890–903, 2018.
- [13] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy (Accepted)*, 2018.
- [14] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in Neural Information Processing Systems*, 2011, pp. 2546–2554.
- [15] P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," *Solar Energy*, vol. 73, no. 3, pp. 151–157, 2002.
- [16] W. Greuell, J. Meirink, and P. Wang, "Retrieval and validation of global, direct, and diffuse irradiance derived from SEVIRI satellite observations," *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 5, pp. 2340–2361, 2013.
- [17] H. Deneke, A. Feijt, and R. Roebeling, "Estimating surface solar irradiance from METEOSAT SEVIRI-derived cloud properties," *Remote Sensing of Environment*, vol. 112, no. 6, pp. 3131–3141, 2008.
- [18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.
- [19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, arXiv eprint 1412.6980.
- [20] Y. Yao, L. Rosasco, and A. Caponnetto, "On early stopping in gradient descent learning," *Constructive Approximation*, vol. 26, no. 2, pp. 289–315, 2007.
- [21] "Royal Netherlands Meteorological Institute (KNMI)," <http://knmi.nl/>.
- [22] "Database of the weather stations of the KNMI," [http://www.sciamachy-validation.org/climatology/daily\\_data/selection.cgi](http://www.sciamachy-validation.org/climatology/daily_data/selection.cgi).
- [23] "European Centre for Medium-Range Weather Forecasts (ECMWF) website," <https://www.ecmwf.int/>.
- [24] R. W. Andrews, J. S. Stein, C. Hansen, and D. Riley, "Introduction to the open source PV LIB for python photovoltaic system modelling package," in *Photovoltaic Specialist Conference (PVSC), 2014 IEEE 40th*. IEEE, 2014, pp. 0170–0174.
- [25] P. Lauret, C. Voyant, T. Soubdhan, M. David, and P. Poggi, "A benchmarking of machine learning techniques for solar radiation forecasting in an insular context," *Solar Energy*, vol. 112, pp. 446–457, 2015.