# PVNet: A LRCN Architecture for Spatio-Temporal Photovoltaic Power Forecasting from Numerical Weather Prediction

**Johan Mathe** [1] **Nina Miolane** **Nicolas Sebastien** [2] **Jeremie Lequeux** [1]

## Abstract

Photovoltaic (PV) power generation has emerged as one of the leading renewable energy sources. Yet, its production is characterized by high uncertainty, being dependent on weather conditions like solar irradiance and temperature. Predicting PV production, even in the 24-hour forecast, remains a challenge and leads energy providers to left idling - often carbon-emitting - plants. In this paper, we introduce a Long-Term Recurrent Convolutional Network using Numerical Weather Predictions (NWP) to predict, in turn, PV production in the 24-hour and 48-hour forecast horizons. This network architecture fully leverages both temporal and spatial weather data, sampled over the whole geographical area of interest. We train our model on a prediction dataset from the National Oceanic and Atmospheric Administration (NOAA) to predict spatially aggregated PV production in Germany. We compare its performance to the persistence model and state-of-the-art methods.

## 1. Introduction

There is an increased commitment worldwide to mitigate greenhouse gas emissions and countries are taking actions to better integrate renewable clean energies into their grids. As an example, many scientific studies tackle the challenge of photovoltaic (PV) power forecast. They rely on statistical time-series methods, physical methods, or ensemble methods which combine different models to enhance accuracy (Sobri et al., 2018). Widely used time-series models can be separated into two groups: linear and non-linear models, where the latter often shows better prediction accuracy when enough data are available. Among non-linear time-series models, Artificial Neural Networks (ANN) have become increasingly popular for PV forecast (Ding et al., 2011; Kardakos et al., 2013; Dolara et al., 2015) and among them Recurrent Neural Networks (Malvoni et al., 2013) and Long-Short Term Memory Networks (Abdel-Nasser & Mahmoud, 2017; Gensler et al., 2017).

Meanwhile, flexible non-linear predictions models taking into account the spatiotemporal structure of the data, like Long-Term Recurrent Convolutional Network (LRCN) (Donahue et al., 2015), 2D LSTM or Convolutional LSTM architectures (ConvLSTM) (Shi et al., 2015), have been successfully applied to a variety of problems. LRCN models have been used in activity recognition, image captioning and visual question answering (Donahue et al., 2015). A 2D LSTM model has been applied to traffic forecasting (Zhao et al., 2017) while a ConvLSTM has shown promising results on a precipitation forecast that predicts rainfall intensity in a local region on a short time horizon (Shi et al., 2015) (also known as nowcasting), a weather-related prediction problem that shares similarities with PV forecast.

**Contributions** This paper presents PVNet, a PV forecasting model that introduces an LRCN architecture to integrate past PV power 1D time-series with dense spatiotemporal NWP and physical models' inputs. Our model extends the state-of-the-art 1D time-series models by learning spatial features with CNN modules, whose channels encode the NWP and physical models variables. We focus our analysis on the day-ahead PV forecast, whose accuracy is known to be highly dependent on NWP integration, and which is particularly interesting as energy market bids are placed one day in advance. Our experimental results show that our architecture manages to leverage these additional spatial data for prediction accuracy. We train a non-linear prediction model that reaches a high accuracy on our validation dataset.

## 2. Data Representation

In what follows, $t$ is the time, measured in hours from the current time and $(x, y)$ is the longitude-latitude coordinates of a point in the region of interest.

**Numerical Weather Predictions** We incorporate numerical weather prediction (NWP) data and their non-linear relationships in our model. Rather than just incorporating

[1]Froglabs, San Francisco [2]Reuniwatt. Correspondence to: Johan Mathe <johan@froglabs.ai>, Nina Miolane <nina.miolane@polytechnnique.fr>.

the irradiance and temperature at the specific location of the power plant(s), our model aims to leverage the full spatial scalar fields of the meteorological parameters over the whole area of interest. These forecasts can be obtained from global forecast systems, like the ECMWF HRES model (temporal resolution of 1 hour and spatial resolution of 0.1 degree) [1] or the NOAA GFS [2] (temporal resolution of 3 hours and spatial resolution of 0.5 degrees). Each of these models provides a set of weather-related variables:

$$X_{NWP,t}(x,y) = \begin{pmatrix} NWP_1(x,y) \\ \vdots \\ NWP_K(x,y) \end{pmatrix} \tag{1}$$

where $K$ is the number meteorological factors of interests.

**Standard Models Predictions** We also incorporate predictions of standard forecast models into our feature vector. We consider first the persistence model, a simple - yet popular - forecast model. The persistence model (PSS) predicts that future PV output is likely to be similar to the current value at the same time of the day (Cornaro et al., 2015):

$$X_{\text{PSS}}(x,y,t) = P_{pv}(t - T_p) \tag{2}$$

where time $T_p$ is measured in hours. Here we take two specific values of persistence. One will be used for comparing the performance of the algorithm at $T_{24}$ and the one we use as a feature of our algorithm is at $T_{48}$. Indeed when used in the context of energy forecast, there is a 24 hours delay.

Second, we consider the clear sky model (CSM), which estimates irradiance $X$ under the assumption that there are no clouds in the sky (Reno et al.):

$$X_{\text{CSM},t}(x,y) = \text{CSM}(t,x,y,\alpha(t,x,y)) \tag{3}$$

where $\alpha$ is a set of external variables used for the Clear Sky Model, for example various atmospheric conditions.

**Feature vector** We form a feature vector $X_t$ by aggregating the weather forecast variables with the predictions of the standard models:

$$X_t(x,y) = \begin{pmatrix} X_{\text{NWP},t}(x,y) \\ X_{\text{PSS},t} \\ X_{\text{CSM},t}(x,y) \end{pmatrix} \in \mathbb{R}^{K+2} \tag{4}$$

Our feature vector incorporates both spatial weather forecasts as well as some predictions from existing models.

## 3. PVNet Model

The goal of PV power output forecasting is to use numerical weather predictions, to forecast the future PV output in a region, in our case: at the day ahead forecast at the scale of a country. We consider $t = 0$ the time at which we make the prediction, and thus $t_{\text{forecast}} = 24h$ is the forecast time. Precisely, we consider a sliding window of size $T$, and $(X_{t_1}, ..., X_{t_T})$ where $X_t$ is the feature maps at time $t$, as defined in Equation 4 and $t_T = t_{\text{forecast}}$. Our goal is to predict $\widehat{P_{pv}}(t_T)$,, an estimate of the PV production at the forecast horizon. Our training set is thus a set of $\left(\left(X_{t_0+t_1}, ..., X_{t_0+t_T}\right), P_{pv}(t_0 + t_T)\right)_{t_0}$ for different values of $t_0$.

Our model presents a long term recurrent convolutional network which combines a convolutional network (CNN) and a bi-directional long short term memory (LSTM) (Schuster & Paliwal, 1997) network, as shown in Figure 1. In practice, we use a time window of $T = 8$ values, spread at 3 hours time interval, so as to represent a window that is a full day of data history.

**Architecture** First, we use the CNN module to encode the spatial data into a new feature vector of lower dimension as running the LSTM directly on the whole input tensor would be too computationally expensive. Rather than hand-crafting spatially aggregated features, we let the CNN learn spatial features that are the most relevant for the following LSTM-based PV power predictions. Then, we use Bidirectional RNNs (Schuster & Paliwal, 1997), taking into account past and future data, as this has shown better efficiency over standard LSTM where any hidden layers only have access to past data (Arisoy et al., 2015). The BiLSTM module designs features, at each time $t$, that take into account the temporal structure of the input signal.

The convolutional network extracts spatial features and then forwards them to the LSTM. We then have a fully connected layer taking all the LSTM hidden outputs and producing a single scalar value that is the estimated value of aggregated PV output prediction for the whole country. We highlight that this model differs from the ConvLSTM model from (Shi et al., 2015), in the sense that the ConvLSTM model integrates convolutional structures directly in the LSTM cells. Both CNN and LSTM modules are trained simultaneously, to minimize a mean squared error loss function.

**Implementation Details** We implement PVNet using the Tensorflow (Abadi et al.) framework. Our convolutional networks are built alternating 2D convolution layers, PReLu activations (He et al.), dropout and max pooling. There is a 20% dropout after each convolutional layer. We added a fully connected layer between the CNN layers and the LSTM, with a 30% dropout in between. We use the hard sigmoid for the recurrent activation and the hyperbolic tangent for the activation itself.

## 4. Experiments

**Dataset** We consider an aggregated production of PV power across Germany, from 2014 to 2018, sampled at a time resolution of 15 minutes. We use NWP data from NOAA, more

[1] https://www.ecmwf.int/en/forecasts/datasets/set-i
[2] https://www.emc.ncep.noaa.gov/GFS/doc.php

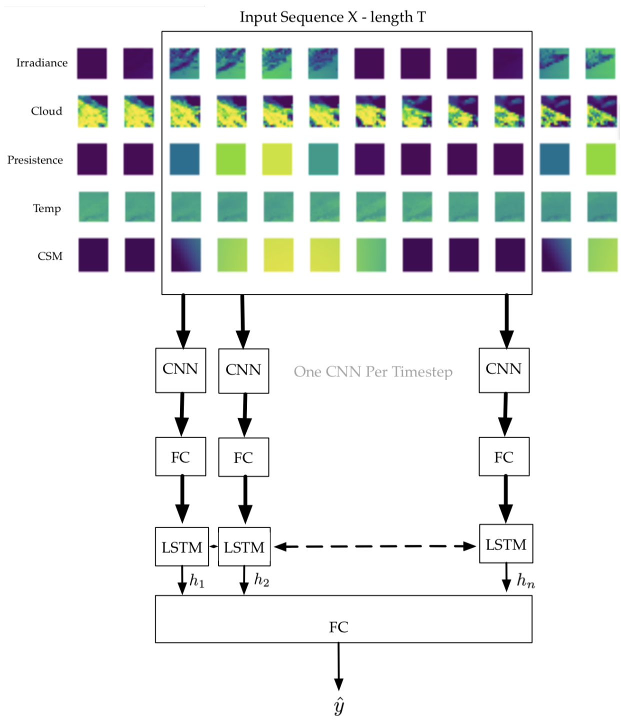


*Figure 1.* PVNet Architecture Details

specifically the Global Forecast System (GFS). The time resolution of the PV power is downsampled at 3h, to match the time resolution of the NWP. We consider $K = 3$ meteorological variables: Downward short wave radiation flux (DSWRF), Cloud Cover (EACC) and Temperature (TMP). We incorporate data from the persistence model (PSS) and the Clear Sky Model (CSS). In a production setting, one only has access to the data from the day before. Thus, we use the persistence data from only 48 hours before the point estimate.

**Training** We split the training and validation dataset as follows: first 3 years for training and the last year for validation. This approach guarantees that our validation score is not affected by seasonal patterns.

We train our model using the Adam optimizer (Kingma & Ba, 2014), with a learning rate of 0.0015 and a batch size of 32. We run our training for a total of 500 epochs.

**Metrics** We use two metrics to evaluate the quality of PVNet: the root mean squared error (RMSE) and mean absolute error (MAE) of the prediction. We compute these metrics only when the measured and predicted power is higher than 0 ($P_{PV}(t) > 0$). We do not take nights into accounts because these values tend to increase the performance since the estimator of the PV night is $\hat{P_{PV}}(t) = 0$ The normalized RMSE and MAE (nRMSE and nMAE) are normalized with the country-wide PV capacity, which is 41.2 GW in our case.

**Results** Table 1 presents our results for PVNet compared to the persistence model and the current state of the art for day ahead country wide PV prediction (Lorenz et al., 2011). We compute the persistence model RMSE by taking the value of the country-aggregated prediction 24 hours before the current value. The authors of (Lorenz et al., 2011) used data from 2009-2010 and a different formulation of the nRMSE. They normalized the production on a per plant basis while we normalize it at the scale of the whole country. We re-implement their method in order to compute the metrics MAE, nMAE, RMSE, and nRMSE for the same times as the ones used in our validation dataset. We refer to this method as "L" in our table.

*Table 1.* PVNet Experimental Performance.

| | NRMSE | NMAE | RMSE | MAE |
|---|---|---|---|---|
| PSS | 22.04% | 15.28% | 8816 MW | 6297 MW |
| L | 6.11% | 4.37% | 2518 MW | 1798 MW |
| **PVNET** | **4.73%** | **3.63%** | **1949 MW** | **1499 MW** |

We first observe an RMSE improvement of 17.31 percents compared to the persistence model, which is expected since the persistence model is relatively naive. We do also use the persistence data in our model as one of the inputs - even though the timing is quite different. We use 24 hours persistence for verification and 48 hours persistence as the PVNet input layer. We also observe a 1.38% decrease in accuracy error in nRMSE and a 0.74% decrease in accuracy error in nMAE compared to state of the art in the country-wide day ahead aggregation (Lorenz et al., 2011).

## 5. Conclusion and Future Work

Accurate PV forecast remains challenging because of its correlation with highly fluctuating weather variables. This paper introduced PVNet, a country-wide PV forecast model that fully integrates 1D time-series of past PV power with dense spatiotemporal exogenous inputs. The model enjoys good prediction performances, with a decrease in nRMSE of 1.38% compared to the state-of-the-art model for country-aggregated PV output prediction. The model also demonstrates inference capability, e. g. learning the geographic impact of different meteorological factors on the PV power prediction and the surface density of PV power production for a given area.

Our future work will involve the addition of interpolation of NWP inputs, in order to avoid temporal down-sampling and data loss, and adding new spatial information, like satellite imagery or fish-eye data imagery. Our objective is to enable a secure and economic integration of PV power into countries' smart energy grids.

## References


Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., and Research, G. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Technical report. URL www.tensorflow.org.

Abdel-Nasser, M. and Mahmoud, K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Computing and Applications*, pp. 1–14, 2017. ISSN 09410643. doi: 10.1007/s00521-017-3225-z.

Arisoy, E., Sethy, A., Ramabhadran, B., and Chen, S. Bidirectional recurrent neural network language models for automatic speech recognition. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5421–5425. IEEE, 4 2015. ISBN 978-1-4673-6997-8. doi: 10.1109/ICASSP.2015.7179007. URL http://ieeexplore.ieee.org/document/7179007/.

Cornaro, C., Pierro, M., and Bucci, F. Master optimization process based on neural networks ensemble for 24-h solar irradiance forecast. *Solar Energy*, 111:297–312, 2015. ISSN 0038092X. doi: 10.1016/j.solener.2014.10.036. URL http://dx.doi.org/10.1016/j.solener.2014.10.036.

Ding, M., Wang, L., and Bi, R. An ANN-based Approach for Forecasting the Power Output of Photovoltaic System. 2011. doi: 10.1016/j.proenv.2011.12.196. URL www.elsevier.com/locate/procedia.

Dolara, A., Grimaccia, F., Leva, S., Mussetta, M., and Ogliari, E. A physical hybrid artificial neural network for short term forecasting of PV plant power output. *Energies*, 8(2):1138–1153, 2015. ISSN 19961073. doi: 10.3390/en8021138.

Donahue, J., Hendricks, L. A., Guadarrama, S., Rohrbach, M., Venugopalan, S., Darrell, T., and Saenko, K. Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015. ISBN 9781467369640. doi: 10.1109/CVPR.2015.7298878.

Gensler, A., Henze, J., Sick, B., and Raabe, N. Deep Learning for solar power forecasting - An approach using AutoEncoder and LSTM Neural Networks. *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, pp. 2858–2865, 2017. ISSN 1879-0852. doi: 10.1109/SMC.2016.7844673.

He, K., Zhang, X., Ren, S., and Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. Technical report. URL https://arxiv.org/pdf/1502.01852.pdf.

Kardakos, E. G., Alexiadis, M. C., Vagropoulos, S. I., Simoglou, C. K., Biskas, P. N., and Bakirtzis, A. G. Application of time series and artificial neural network models in short-term forecasting of PV power generation. *Proceedings of the Universities Power Engineering Conference*, pp. 1–6, 2013. ISSN 1944-9925. doi: 10.1109/UPEC.2013.6714975.

Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization. 12 2014. URL http://arxiv.org/abs/1412.6980.

Lorenz, E., Heinemann, D., and Christian, K. Local and regional photovoltaic power prediction for large scale grid integration: Assessment of a new algorithm for snow detection. *International Journal of ChemTech Research*, 9(11):261–270, 2011. ISSN 09744290. doi: 10.1002/pip.

Malvoni, M., De Giorgi, M. G., and Congedo, P. M. Photovoltaic power forecasting using statistical methods: impact of weather data. *IET Science, Measurement & Technology*, 8(3):90–97, 2013. ISSN 1751-8822. doi: 10.1049/iet-smt.2013.0135. URL http://digital-library.theiet.org/content/journals/10.1049/iet-smt.2013.0135.

Reno, M. J., Hansen, C. W., and Stein, J. S. SANDIA REPORT Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis. Technical report. URL http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online.

Schuster, M. and Paliwal, K. K. Bidirectional Recurrent Neural Networks. Technical Report 11, 1997. URL https://pdfs.semanticscholar.org/4b80/89bc9b49f84de43acc2eb8900035f7d492b2.pdf.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k., and Woo, W.-c. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. pp. 1–12, 2015. ISSN 10495258. doi: 10.1074/jbc.M200827200. URL http://arxiv.org/abs/1506.04214.

Sobri, S., Koohi-Kamali, S., and Rahim, N. A. Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management*, 156(May 2017):459–497, 2018. ISSN 01968904. doi: 10.1016/j.enconman.2017.11.019. URL https://doi.org/10.1016/j.enconman.2017.11.019.

Zhao, Z., Chen, W., Wu, X., Chen, P. C. Y., and Liu, J. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2):68–75, 2017. ISSN 1751-956X. doi: 10.1049/iet-its.2016.0208. URL http://digital-library.theiet.org/content/journals/10.1049/iet-its.2016.0208.