ENHANCING RENEWABLE ENERGY FORECASTING FOR HYBRID POWER STATIONS: A DATA-DRIVEN APPROACH

Francesco Superchi[1]*, Antonis Moustakis[2], George Pechivanoglou[2], Alessandro Bianchini[1]

[1]Department of Industrial Engineering, University of Florence, Florence, Italy

[2]Eunice Energy Group, Athens, Greece

*Corresponding Author: francesco.superchi@unifi.it

ABSTRACT

The present study investigates the correction of bias errors in solar photovoltaic production forecasting by looking for a correlation between commercial predictions and the history of the actual operation. Using a whole year of real production data from an actual hybrid power station and forecasts from a commercial provider, a comprehensive comparison of various Machine Learning (ML) algorithms, from a simple Linear Regression to more complex methods such as Neural Networks, is carried out. By mimicking the actual application, ML algorithms were used to improve the forecasts of 24 hours of operation at a time, after being trained on several previous days. This rolling horizon training approach allows models to capture seasonal variations in Renewable Energy Sources production and other possible phenomena. The effectiveness of methods is measured by means of error metrics such as the Normalized Mean Absolute Error (NMAE). This work also dives into the optimization of key parameters for the application of the data-driven model: the optimal number of days used for the model training phase and their hyperparameters (i.e., parameters used to control the learning process). The investigation indeed reveals an optimal data window for the training phase. Moreover, the optimization of the hyperparameters that tune the training phase of algorithms is key to tailoring the methods to the specific characteristics of the application. In addition to the reduction of the bias error, results show that the application of optimized methods can reduce the NMAE of solar production from almost 9% to less than 4%.

1 INTRODUCTION

The intrinsic intermittence of solar radiation and wind speed makes the power generation from energy systems making use of them inevitably variable and non-constant. On the other hand, those sources are highly available all over the world and their adoption is crucial for the decarbonization of our society. Power generation based on Renewable Energy Sources (RES) is spreading all over the world, even where outdated electrical grids may struggle to host a high share of intermittent supply. To allow for the integration of this kind of production, reliable methods must be developed to control the output of intermittent generators. Grid operators must pose limits and constraints to producers to manage the distribution of electricity without violating technical constraints. Moreover, to obtain the optimal power output schedule from every kind of generator, it is key to know in advance its production capabilities. In this context, a reliable forecasting of renewable power production is crucial.

Hybrid Power Stations (HPSs), by combining RES generation with storage systems, help to improve the flexibility of the green power output. With an optimized control of storage and accurate scheduling, these systems may behave similarly to dispatchable generation units. However, to produce the required scheduling in advance, it is necessary to know the time and level of the power generation from renewables. Current forecasting methods of RES production are based on predictions made by the closest weather station. Solar and wind generation depends indeed on weather conditions, whose forecasts have always been affected by a certain degree of uncertainty. Moreover, stations can be situated at a considerable distance from power generators, and this may lead to additional bias errors.

https://doi.org/10.52202/077185-0210

Several methods have been developed and tested to obtain reliable forecasts for solar production. Indeed, accurate PV power forecasting is crucial for transmission and distribution system operators to improve the management of energy flows (Pierro et al., 2022).

Yin et al. (2020) studied the error distribution characteristics of PV power prediction and proposed a correction method to improve the forecasting. Their method succeeded in reducing the Mean Absolute Error (MAE) by about 2.6%. Ma et al. (2021) analyzed the performance of PV plants and showed that the PV power forecasting error mainly comes from numerical weather prediction and forecasting processes. They proposed a short-term PV power forecasting method based on irradiance correction and error forecasting. Bright et al. (2018) proposed a methodology to improve satellite-derived PV power prediction by combining two data sources.

Data-driven approaches are promising in improving renewable power forecasting, especially when involving artificial intelligence (AI). In their review, Sobri et al. (2018) stated that AI approaches outperform traditional methods due to their capability to solve the non-linear and complex structure of data. Sharifzadeh et al. (2019) compared three predictive methods including NNs, SVR, and Gaussian Process Regression (GPR) to predict the behavior of renewable energy from wind and solar resources and electricity demand. They found that all tested models were able to predict wind and solar power.

Even if a simple LR can lead to good results, more advanced models can improve consistently the prediction. Markovics and Mayer (2022) compared 24 machine learning models and stated that the best one outperformed LR by 13.9% in Root Mean Squared Error (RMSE) terms. They also stated that hyperparameter tuning is essential to exploit the full potential of the models.

Das et al. (2017) applied an SVR-based model to forecast PV power generation and, for their application, this method was more accurate than NNs. Also De Leone et al. (2015) succeeded in obtaining accurate PV production forecasts using SVR. On the other hand, AlShafeey and Csáki (2021) compared regression techniques with NNs to perform 24-hour ahead horizon PV forecasts. Their results show that NNs perform better than regressions regardless of the input method used.

Commercial providers often deliver generic forecasts that may fail to capture site-specific conditions such as variable shadowing and conversion losses. Böök and Lindfors (2020) proposed a method to adjust a generic PV output forecast to any specific site. Their model succeeded in reducing the MAE from 8 to 6%. Differently from the other studies cited before, the approach proposed in this study is similar only to the latter work but proposes a more direct approach.

The aim of the present study is the development of a reliable method to correct generic renewable power production forecasts using a data-driven approach. Three methods have been employed and compared: Linear Regression (LR), Support Vector Regression (SVR), and Neural Networks (NNs).

The time series of the production from an actual PV field has been used to train a machine learning (ML) algorithm that aims to reduce the error between the forecasts from a commercial provider and the actual production. The solar generator considered in this study is part of the HPS of the Greek island of Tilos, which comprises solar and wind power generation and an energy storage system. This HPS must produce every day an output schedule based on RES production forecasts to optimize the use of the high-temperature battery that transforms the intermittent production into dispatchable power to sell to the grid operator (Superchi et al., 2024). Those methods could also be used to schedule the activation of different components of more complex hybrid energy systems based on renewables, as one proposed in previous studies by the same authors (Superchi et al., 2023). In this sense, an improved forecast has the potential to enhance the performance of HPSs, since the accuracy of the prediction directly affects the effectiveness of energy storage scheduling, key to maximizing renewable source utilization and revenues.

The present work may contribute to the body of literature by applying ML methods to enhance a commercial forecast and testing its effectiveness on the production history of an actual HPS. A data-driven approach is promising for this application because it can correlate a generic prediction with the real performance of the specific plant. The actual production may be different from the predicted one for several reasons: the degradation of components, the specific location, environmental effects, and many other factors that may influence the performance of devices. Data-driven models have the potential to capture all those effects by learning from historical datasets. A correct model fit can post-process generic predictions to introduce the effect of all the phenomena that are influencing the operation of the plant.

https://doi.org/10.52202/077185-0210

The paper is structured as follows. Section 2 presents the research approach, i.e., which ML methods have been applied and how. Section 3 presents how the algorithms have been tuned and which error metrics have been selected to evaluate the effectiveness of three proposed methodologies. Finally, Section 4 presents the results of the sensitivity analysis and highlights which method is the most promising one to reduce the forecasting error of solar production.

2 RESEARCH APPROACH

2.1 Training and application of ML algorithms

The method proposed herein makes use of a moving window approach to train the ML algorithm on several days of operation and then applies it to improve the forecast related to the day after the training. This mimics the actual operation of the algorithm: this method can be used to improve the forecasts related to the next day, before running the algorithm that produces the optimal energy scheduling of the HPS. Three machine learning methods have been tested and compared to assess which one would be better suited for this specific application. The first was a simple LR, that, given the repetitive nature of the PV production, could produce good results with low computational efforts. Then, a more complex model was tested: the SVR. This more advanced method allows for parameter tuning that may enhance the effectiveness of its application for this specific task. The third and last tested methods are NNs, the most advanced algorithm tested in this work. NNs may capture non-linear correlations between the time series used for training and lead to better results.

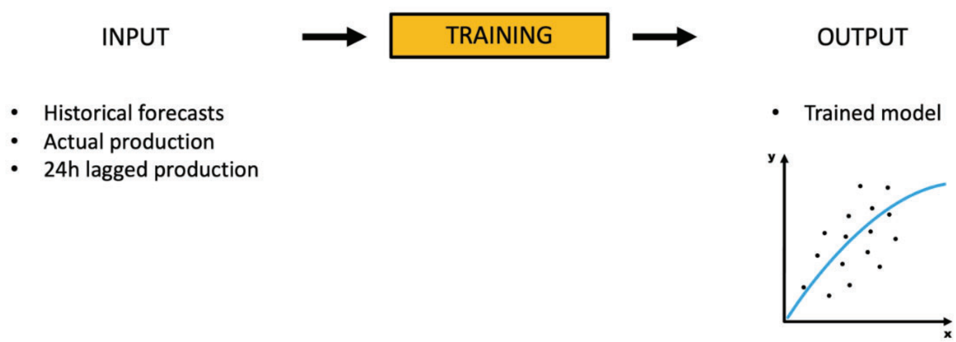

Figure 1: Training Phase.

During the training phase (Figure 1), the algorithm takes as input the historical production from the PV field and the historical forecast of the production that was delivered from the commercial provider. The algorithm also exploits the autocorrelation of solar production, an additional information to improve the correction on the forecast further. The time series is lagged by 24 hours and, during the phase of the forecast correction, the algorithm also considers what was the PV production the day before. The output of the training phase is a tuned model that links the original forecast from the provider with the actual production.

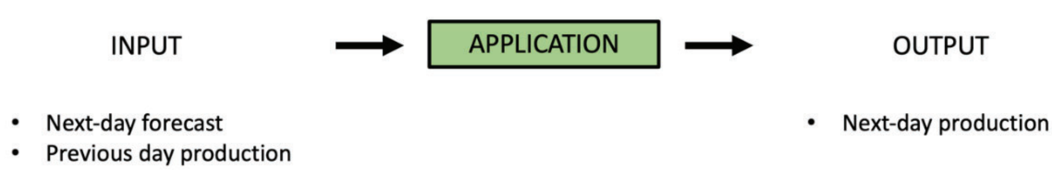

Figure 2: Application Phase.

https://doi.org/10.52202/077185-0210

During the application phase (Figure 2), the algorithm takes as input the solar production forecast for the next day of operation and the actual operation during the 24 hours before the application. The output is, ideally, the actual solar production of the next day.

2.2 Moving window approach

This analysis simulates this kind of approach to improve the following-day renewables forecasting over one year of operation. The algorithm is trained on the first N days of operation and then applied to improve the forecast of the first "application day". Then, the training window shifts forward to test the algorithm on again N days before the second "application day", and so on (Figure 3). Results are presented as the average error that would characterize a forecast improved using a specific method when applied during this entire year of operation.

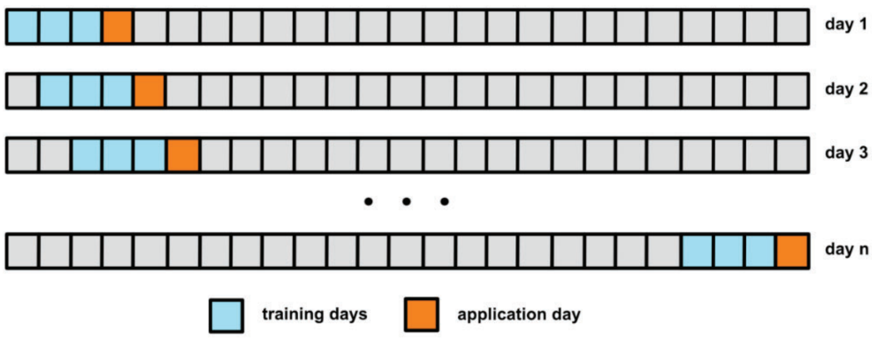

Figure 3: Moving window approach to test the actual performance of the algorithm.

To capture all the effects that create the mismatch between the generic forecast with the actual production from the PV field, it is key to train the algorithm on an operating window in proximity to the application day. In this way, all the phenomena that are affecting the real operation are considered. On the other hand, training the algorithm with data far in time can induce bias errors in the prediction.
The results section presents several sensitivity analyses to optimize the method for this specific application. The size of the training window is varied to understand what the optimal number of previous days is to train the algorithm. Moreover, machine learning models have been tuned to fit the specific requirements of this problem.

3 METHODS

3.1 Photovoltaic production data and power production forecasts

The HPS of Tilos comprises wind and solar energy production. The Energy Management System (EMS) of the HPS receives forecasts from the commercial provider and real-time measurements of the performance of each generator. Both the prediction and the production are stored in datasets, available for data analyses. This case study considers specifically the solar production of the HPS.
The PV field installed in Tilos is characterized by a peak power of 160 kWp and composed of 592 solar panels of 270 Wp each, mounted with a tilt angle of 30° and oriented towards the south. The time series of historical solar production has been kindly provided by Eunice Energy Group (EEG), owner and manager of the HPS of Tilos. The production dataset is characterized by a 1-hour time resolution and provides the average power production from the PV field in that time step. The historical forecasts come from a commercial provider, with the same hourly resolution and covering the same time span as the historical production dataset.
The training window runs through thirteen months of operation of an actual PV field, from September 1st, 2022, to October 1st, 2023. The application day is instead selected on a year of operation, from October 1st, 2022, to October 1st, 2023.

https://doi.org/10.52202/077185-0210

3.2 Machine learning methods and hyperparameters optimization

The methods proposed here can be tuned by modifying the parameters of the algorithms and the time series that they take as input for the training phase. The parameters that specify the details of the learning process of machine learning algorithms are called "hyperparameters". Those hyperparameters are related to the specific machine-learning algorithm and must be tuned according to the specific application. Moreover, to understand the optimal number of days to train the algorithm, train training window was varied from just three previous days to an entire month before the "application day".
As stated in Section 2, the algorithms proposed in this work are LR, SVR, and NNs. All methods are described in detail by Superchi et al. (2021).

3.3 Linear Regression (LR)

LR is a linear model that provides the equation of a line describing the relationship between a predictor (original forecast and lagged production) and an outcome (actual prediction). The training phase tunes the coefficients of the model to minimize the distance between the model outcome (predicted power) and the target (historical power production).
LR method was implemented through the *LinearRegression* function from the *scikit-learn* library (Pedregosa et al., 2011), and it was optimized by varying the number of days that the algorithm can use for the training phase.

3.4 Support Vector Regression (SVR)

In SVR is possible to define a degree of acceptance for the error, hence to tailor the algorithm to a specific task (Maciejowska et al., 2021). In addition to the optimization of the number of training days, two hyperparameters of the SVR were tuned:

- ε: maximum error that the model accepts during the training phase. This parameter specifies the range in which no penalty is associated to predicted points that fall within a distance equal to or smaller than ε from the actual value. Its initial trial value is 0.0 and is optimized between 0.0 and 1.0.
- C: hyperparameter that tunes the tolerance of errors greater than ε. A higher value of C makes the tolerance for points outside the maximum error epsilon increase. When C is null, there is no tolerance for values outside the range. Its first trial value is 1.0 and it is optimized in a range from 1.0 to 20.0.

SVR was implemented through the *LinearSVR* function from the *scikit-learn* library.

3.5 Neural networks (NNs)

As discussed in section 2, NNs have been proposed to understand if a more complex method may lead to better results, maybe capturing non-linear correlations between the time series used as input.
A prediction algorithm based on NNs consists of an iterative procedure by which a collection of units (neurons) transmits and processes signals to correlate the predictor to the outcome. Each neuron is characterized by a weight and a bias that modifies the signal when it passes through it. In the training process, the weights and biases are adjusted at each iteration to minimize the error between the outcome and of NNs the target. The process is iterated until maximum iterations, or a convergence criterion, are reached (Catalão et al., 2007).
NNs were implemented using the *MLPRegressor* function from the *scikit-learn* library, using the solver "adam". An initial learning rate of 0.0001 was used to contain the step size in updating the weights.
The number of training days was then optimized also for the NNs method. Moreover, two hyperparameters were tuned to enhance its performance:

- Topology: structure of the network as the number of hidden layers and number of neurons inside each of them.
- Maximum iterations: the maximum number of iterations that the solver can perform. The default value is 200.

https://doi.org/10.52202/077185-0210

3.6 Normalized Mean Absolute Error (NMAE)

The error that characterizes the forecast when trying to predict the power production of the PV field in a given instant can be normalized and expressed as a Normalized Error (NE):

$$NE = \frac{P_{forecast} - P_{actual}}{P_{rated}} \cdot 100 \tag{1}$$

The error between the forecasted power $P_{forecast}$ and the actual power production P_{actual} is normalized to the rated power P_{rated} of the PV field i.e., 160 kWp.
To assess the overall accuracy of the forecast, it was applied the NAME metric, computed as:

$$NMAE = \frac{1}{n}\sum_{i=1}^{n} \frac{\left|P_{forecast} - P_{actual}\right|}{P_{rated}} \cdot 100 \tag{2}$$

The NMAE represents the arithmetic average of the absolute value of the normalized errors that characterize the forecast in each timestep n considered in the analysis. In this case, n is equal to the number of hours in a year (8760).
The original forecast from the commercial provider is characterized by an average NMAE for the selected year of 8.6%. The models will be tested, and their effectiveness will be evaluated by computing this same metric upon the adjusted prediction that they will produce.

3.7 Normalized Root Mean Squared Error (NRMSE)

The RMSE is the quadratic mean of the differences between the predicted and the actual values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (P_{forecast} - P_{actual})^2} \tag{3}$$

The RMSE is more sensitive to outliers than the NMAE, since larger errors have a disproportionally large effect on it. It is useful to evaluate both of these error metrics since they highlight different aspects of forecasting accuracy.
As the NMAE, it was normalized to the nominal power of the PV field and expressed in percentage terms:

$$NRMSE = \frac{RMSE}{P_{rated}} \cdot 100 \tag{4}$$

The NRMSE of the forecast from the commercial provider is 15.95%. Methods were optimized to minimize the NMAE, but the NRMSE of final results is evaluated to understand if, from the point of view of this metric, the forecasts were improved too and to which extent.

https://doi.org/10.52202/077185-0210

4 RESULTS

4.1 Linear Regression (LR) sensitivity analysis on the number of training days

LR was tested by training the algorithm on an increasing number of previous days of operation, starting from 3 and up to 30.

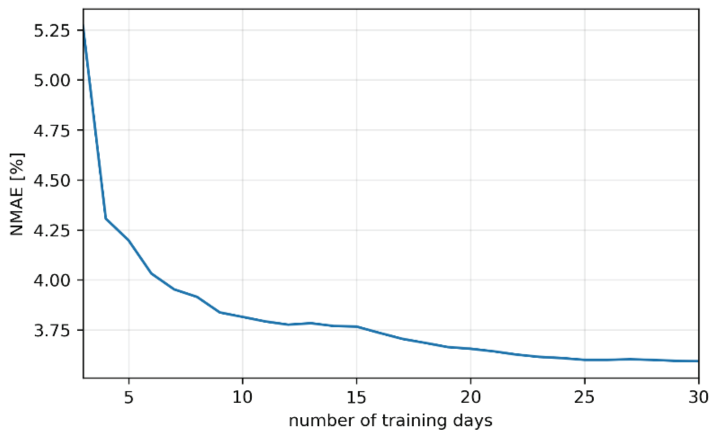

Figure 4: Sensitivity analysis NAME of forecast corrected by LR, varying the number of days used for the training phase of the algorithm.

Figure 4 shows that the average NMAE of the prediction adjusted by the LR method decreases with the number of previous days used for the training phase of the algorithm. In this case, it seems that the more days the algorithm can analyze, the more accurate the improved prediction will be. Trained on 30 days of previous operation, LR produces a forecast characterized by an NMAE of 3.59%. and a NRMSE of 7.68%.

4.2 Support Vector Regression (SVR) sensitivity analysis on C hyperparameter and number of training days

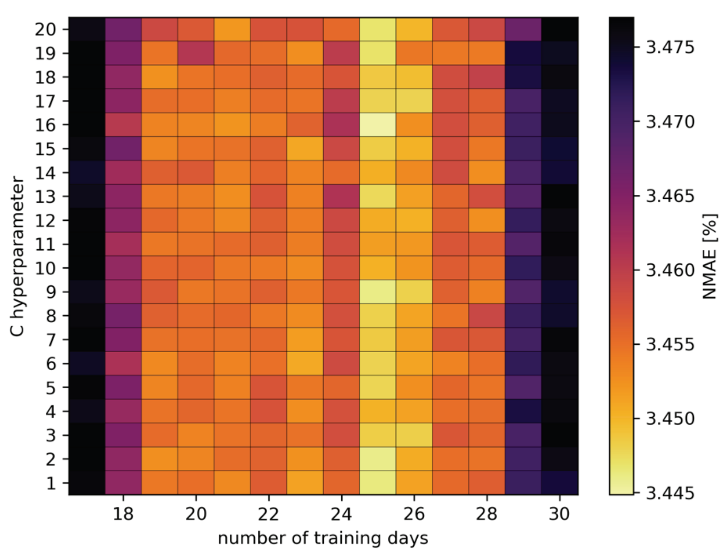

Figure 5: Sensitivity analysis on the NMAE of the forecast corrected by SVR, varying the C parameter and the number of training days used for the training phase of the algorithm.

https://doi.org/10.52202/077185-0210

The SVR method was tested with the hyperparameter ε varying from zero to one, with a resolution of 0.1. Several tests have shown that the best results in terms of minimizing the NMAE of the prediction come from algorithms with ε equal to zero. For this specific application, it seems better to always assign penalties to predicted points far from the target.
The contour plot in Figure 5 shows instead the results of the sensitivity analysis made by varying the value of the hyperparameter C and the number of days used for the training phase. The value of C varies from 1 to 20, with a resolution of 1. The number of training days again varies from 3 to 30, but the figure displays only the area around the optimal value (i.e., 17 to 30 days).
Results show that the best combination to minimize the forecasting error is given by 25 training days and a C equal to 16. With this configuration, the NMAE is equal to 3.44%.
Even if the optimal null value of epsilon means that it is better to always penalize errors during the training phase, a high C means that is also better to increase their tolerance.
In this case, the optimal number of days to train the algorithm is different than the maximum. Differently from LR, the correlation that SVR can find may be biased by values too far away in time.
The optimal SVR performs slightly worse than LR in NRMSE terms, lowering the metric down to 7.8%.

4.3 Neural Networks (NNs) sensitivity analysis on topology and number of training days

NNs were tested by varying the topology of the network, the maximum number of iterations that the algorithm can perform during the training phase, and the number of days used for the training.
Several tests have shown that the default maximum number of iterations (200) was never sufficient to reach convergence during the training phase for this kind of application. A reasonable limit that most of the time allows the training phase to converge, without extending excessively computational times, is ten thousand iterations.

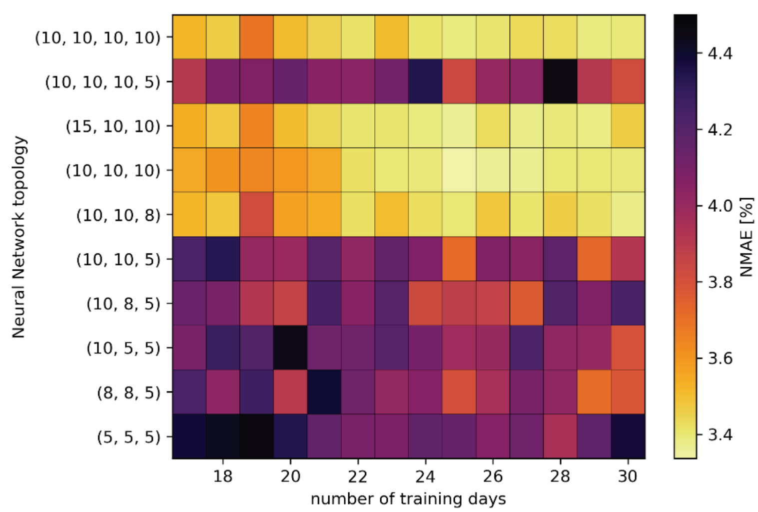

Figure 6: Sensitivity analysis on NMAE varying the topology and the number of training days for NNs.

The contour plot in Figure 6 shows the results of the sensitivity analysis varying the number of days to train the algorithm and the topology of the neural network. The NN topology is described by a sequence of numbers in round brackets indicating how many neurons each layer of the NN contains, from the input to the output. Also in this case, the figure only shows the results from the number of training days around the optimal one (i.e., 17 to 30), even if the sensitivity analysis was performed on the same range as the other ones.
Results show that the best configuration of the NNs method involves a training of 25 days and a network topology characterized by three hidden layers, each containing ten neurons (10, 10, 10). In this case, the NMAE of the improved prediction is reduced to 3.34%. The optimal number of days to train the model is the same as in the case of SVR (i.e., 25), confirming that too many training days may bias the algorithm.

https://doi.org/10.52202/077185-0210

Several possible network topologies were tested to find the best-performing one. Structures composed of a number of neurons lower than 28 were the worst performing. Due to the complexity of the forecasting problem, a higher number of neurons is required to find patterns that a simpler structure may not see. However, results show that the introduction of a fourth layer or the addiction of more neurons was also not beneficial. In this case, it can be a problem of overfitting i.e., the network memorizes training data instead of generalizing the solution. This test case uses only two time series for the training phase, meaning that a too complex structure may be unable to capture any other additional trend. A trade-off between the two extremes seems to be represented by a structure composed of three layers, each containing ten neurons.
NNs achieve very good results also in NRMSE terms, and the optimal topology trained on 25 days of operation lowers the metric down to 7.62%.

4.4 Critical comparison

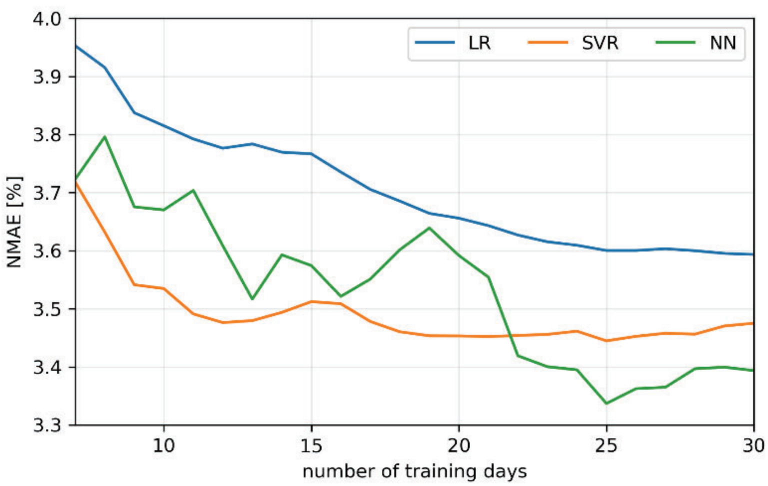

Figure 7: Comparison among the NMAE between LR (blue line), SVR (orange line), and NNs (green line) varying the number of days that the algorithm can use for the training phase.

Figure 7 shows a comparison among the NMAE of the improved forecast resulting from LR (blue line), SVR (orange line), and NNs (green line), varying the number of days that the algorithm can use for the training phase. Results from SVR and NNs come from the best hyperparameters tuning when working with that specific number of training days.
LR performs best if the algorithm is allowed to be trained on the maximum possible number of days. On the other hand, both NNs and SVR perform best when they are trained on 25 days before the "application day". Moreover, LR is always able to improve the original forecast, but the other two tested algorithms always produce better results than LR, no matter how many days are used to train them.
The SVR represents the best solution when the algorithm must be trained on a number of days lower than 22. The SVR well captures the linear correlation between the following day's production and its trend during the previous days. Above this threshold of 22 days, NNs are more precise. When there are enough days to train the algorithm, NNs are able to capture trends that the linear models cannot see and, in this way, they can produce a better final prediction.
Figure 8 shows an example of forecast correction during two weeks of operation, using the best method tested in this work: NNs trained on 25 days of operation. While results express the average accuracy of forecasts during the whole considered year of operation, this graph allows to visualize a portion of the comparison.
Figure 8 (a) shows the time comparison between the original forecast, the improved forecast, and the actual production. The improved forecast is always closer to the actual production since the model can reduce the initial bias error. The figures below show instead the NE trend between the actual production and the original forecast (b) and the improved forecast (c). Figure 8 (b) shows that the original NE was

https://doi.org/10.52202/077185-0210

always negative, meaning that the forecast was always higher than the actual production (bias error). Figure 8 (c) shows instead that the NE of the new forecast is reduced and evenly distributed. The NNs method can remove the bias error and improve the quality of the forecast.

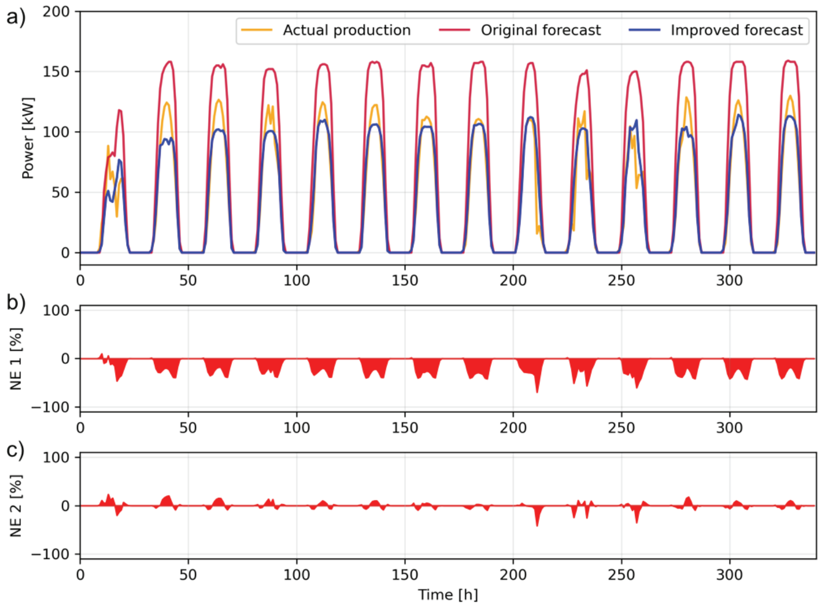

Figure 8: Example of forecast correction during two weeks of operation.
a) comparison between the original forecast, the improved forecast, and the actual production
b) normalized error trend between the original forecast and the actual production
c) normalized error trend between the improved forecast and the actual production.

Table 1: Comparison of best results from different algorithms.

Metric	Original	LR	SVR	NNs
Optimal number of training days	-	30	25	25
Optimal Hyperparameters	-	-	ε = 0 C = 16	Max iter. = 10k Topology = 10,10,10
Final NRMSE	15.95%	7.68%	7.8 %	7.62 %
Final NMAE	8.6 %	3.59 %	3.44 %	3.34 %

Table 1 summarizes the results from the sensitivity analysis made on the three tested methods and their best possible outcome. From the original NMAE of 8.6%, LR trained on 30 days of previous operation can reduce the error down to 3.59%. LR performs best when it is trained on the highest possible number of days prior to its application. This was not the case for SVR and NNs, which performed better when trained on a high number of previous days, but not too many.
SVR tuned with ε equal to zero and C equal to 16 can reduce the NMAE even more when trained on 25 days of operation, bringing the metric to 3.44%. Results of the sensitivity analysis on SVR show that it is better to always assign a penalty to errors during the training phase of the method, but better forecasts come from methods that mitigate this penalty.
NNs lead to the best forecast and lower the error down to 3.34% when they are trained on 25 days prior to their application. They must be properly set, with a sufficient number of neurons that allow them to capture the complexity of patterns, but not excessively complicated to avoid overfitting problems. To

https://doi.org/10.52202/077185-0210

allow the method to converge during the training process, the maximum number of admissible iterations must be sufficiently high (10k).

From the NRMSE point of view, all the tested methods were also beneficial. From 15.59% of the original forecasts, the LR can reduce the metric to 7.68%. SVR performed slightly worse in this sense, bringing the NRMSE to 7.8%. Once again, the method that brings the highest benefit is NNs, decreasing the metric down to 7.62%.

5 CONCLUSIONS

The present study compared different data-driven methods to improve the forecast of production from renewables in a Mediterranean case study. The HPS paradigm is spreading in areas with outdated electrical grids, where the intermittent production from RES must respect strict technical constraints.

Every method tested in this work succeeded in improving the original forecast of the production capabilities of a PV field that feeds an actual HPS. Results show that a data-driven model has the potential to correct a generic forecast delivered by a real commercial provider. Such an application is very promising since a more accurate prediction can have highly beneficial effects in the scheduling phase of an HPS which uses batteries to manage the intermittent nature of renewables.

Despite its simplicity, LR leads to good results. SVR was the best performing approach when the available number of days for the algorithm training was limited. Overall, the method that performed best comprises the NNs algorithm. Every tested method was able to improve the commercial forecast both in terms of NMAE and NRMSE.

NOMENCLATURE

AI	Artificial Intelligence
HPS	Hybrid Power Station
LR	Linear Regression
MAE	Mean Absolute Error
ML	Machine Learning
NE	Normalized Error
NMAE	Normalized Mean Absolute Error
NRMSE	Normalized Root Mean Squared Error
NNs	Neural Networks
P	Power
RMSE	Root Mean Squared Error
SVR	Support Vector Regression

REFERENCES

AlShafeey, M., Csáki, C., 2021. Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods. Energy Reports 7, 7601–7614. https://doi.org/10.1016/j.egyr.2021.10.125

Böök, H., Lindfors, A.V., 2020. Site-specific adjustment of a NWP-based photovoltaic production forecast. Solar Energy 211, 779–788. https://doi.org/10.1016/j.solener.2020.10.024

Bright, J.M., Killinger, S., Lingfors, D., Engerer, N.A., 2018. Improved satellite-derived PV power nowcasting using real-time power data from reference PV systems. Solar Energy, Advances in Solar Resource Assessment and Forecasting 168, 118–139. https://doi.org/10.1016/j.solener.2017.10.091

Catalão, J.P.S., Mariano, S.J.P.S., Mendes, V.M.F., Ferreira, L.A.F.M., 2007. Short-term electricity prices forecasting in a competitive market: A neural network approach. Electric Power Systems Research 77, 1297–1304. https://doi.org/10.1016/j.epsr.2006.09.022

Das, U.K., Tey, K.S., Seyedmahmoudian, M., Idna Idris, M.Y., Mekhilef, S., Horan, B., Stojcevski, A., 2017. SVR-Based Model to Forecast PV Power Generation under Different Weather Conditions. Energies 10, 876. https://doi.org/10.3390/en10070876

https://doi.org/10.52202/077185-0210

De Leone, R., Pietrini, M., Giovannelli, A., 2015. Photovoltaic energy production forecast using support vector regression. Neural Comput & Applic 26, 1955–1962. https://doi.org/10.1007/s00521-015-1842-y

Ma, Y., Lv, Q., Zhang, R., Zhang, Y., Zhu, H., Yin, W., 2021. Short-term photovoltaic power forecasting method based on irradiance correction and error forecasting. Energy Reports 7, 5495–5509. https://doi.org/10.1016/j.egyr.2021.08.167

Maciejowska, K., Nitka, W., Weron, T., 2021. Enhancing load, wind and solar generation for day-ahead forecasting of electricity prices. Energy Economics 99, 105273. https://doi.org/10.1016/j.eneco.2021.105273

Markovics, D., Mayer, M.J., 2022. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. Renewable and Sustainable Energy Reviews 161, 112364. https://doi.org/10.1016/j.rser.2022.112364

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Pierro, M., Gentili, D., Liolli, F.R., Cornaro, C., Moser, D., Betti, A., Moschella, M., Collino, E., Ronzio, D., van der Meer, D., 2022. Progress in regional PV power forecasting: A sensitivity analysis on the Italian case study. Renewable Energy 189, 983–996. https://doi.org/10.1016/j.renene.2022.03.041

Sharifzadeh, M., Sikinioti-Lock, A., Shah, N., 2019. Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression. Renewable and Sustainable Energy Reviews 108, 513–538. https://doi.org/10.1016/j.rser.2019.03.040

Sobri, S., Koohi-Kamali, S., Rahim, N.Abd., 2018. Solar photovoltaic generation forecasting methods: A review. Energy Conversion and Management 156, 459–497. https://doi.org/10.1016/j.enconman.2017.11.019

Superchi, F., Giovannini, N., Moustakis, A., Pechlivanoglou, G., Bianchini, A., 2024. Optimization of the power output scheduling of a renewables-based hybrid power station using MILP approach: The case of Tilos island. Renewable Energy 220, 119685. https://doi.org/10.1016/j.renene.2023.119685

Superchi, F., Mariuzzo, I., Repetto, M., 2021. Machine learning strategies assessment for energy commodities forecasting. Politecnico di Torino.

Superchi, F., Schepers, S., Moustakis, A., Pechivanoglou, G., Bianchini, A., 2023. Towards the Introduction of Green Hydrogen in the Energy Mix of Mediterranean Islands Through the Integration of Wind and Solar Power: A Techno-Economic Case Study. Presented at the ECOS2023, p. 3361. https://doi.org/10.52202/069564-0301

Yin, W., Han, Y., Zhou, H., Ma, M., Li, L., Zhu, H., 2020. A novel non-iterative correction method for short-term photovoltaic power forecasting. Renewable Energy 159, 23–32. https://doi.org/10.1016/j.renene.2020.05.134

ACKNOWLEDGEMENTS

Thanks are due to Eunice Energy Group for providing access to data regarding the Tilos Hybrid Power Station and for the co-financing of the PhD project of Dr. Francesco Superchi.
The authors would like also to thank Prof. Giovanni Ferrara from the University of Florence for the coordination of the PhD project of Dr. Superchi.

https://doi.org/10.52202/077185-0210