Paper:
Monfared, M., Fazeli, M., Lewis, R. & Searle, J. (2019). Fuzzy Predictor With Additive Learning for Very Short-Term PV Power Generation. *IEEE Access, 7*, 91183-91192.
http://dx.doi.org/10.1109/ACCESS.2019.2927804

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 26, 2019, accepted July 2, 2019, date of publication July 10, 2019, date of current version July 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927804

Fuzzy Predictor With Additive Learning for Very Short-Term PV Power Generation

MOHAMMAD MONFARED[1,3], (Senior Member, IEEE), MEGHDAD FAZELI[1], RICHARD LEWIS[2], AND JUSTIN SEARLE[2]

[1]Department of Electrical and Electronics Engineering, College of Engineering, Swansea University, Swansea SA18EN, U.K.
[2]SPECIFIC, College of Engineering, Swansea University, Swansea SA18EN, U.K.
[3]Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad 9177948944, Iran

Corresponding author: Mohammad Monfared (mohammad.monfared@swansea.ac.uk; m.monfared@um.ac.ir)

This work was supported in part by the Engineering and Physical Science Research Council under Grant EP/N020863/1, in part by the Innovate U.K. under Grant 920036, and in part by the European Regional Development Fund under Grant c80892 through the Welsh Government.

ABSTRACT Photovoltaic (PV) power generation is highly intermittent in nature and any accurate very short-term prediction can decrease the impact of its uncertainties and operation costs and boost the reliable and efficient integration of PV systems into micro/smart grids. This work develops a new generalized technique for very short-term prediction of PV power generation from the lagged power generation data using fuzzy techniques. A preprocessor extracts relevant statistical features from the PV data which are fed to the fuzzy predictor. A modified version of Wang-Mendel training algorithm is employed to directly extract the fuzzy rules from the training data pairs. This methodology exploits the limited training data more efficiently. In addition, an online additive learning routine is proposed, which enables the predictor to learn from new data while running the predictions. So, the prediction accuracy increases over time and the predictor updates to account for long-term changing conditions of weather and PV system performance and its surroundings. Numerical results of the comparison of the proposed approach with simple fuzzy and traditional artificial neural network methods on a live PV system in the United Kingdom demonstrate its improved prediction accuracy, outperforming the benchmark approaches with a normalized mean absolute error (NMAE) of 3.6%.

INDEX TERMS Fuzzy theory, photovoltaic (PV), power generation forecasting, Wang-Mendel.

I. INTRODUCTION

With the global shift away from fossil fuels towards renewable energy sources, solar energy is expected to be one of the most profitable sectors. Estimates show that the global solar energy market share will jump to almost twice the wind share in 2025, while the residential end-users hold a dominant share in the global renewable energy market currently and in forecasts for 2025 [1]. At the end of September 2018, the overall UK solar PV capacity reached approximately 13 GW, an increase of 1.8% compared to September 2017 and more than half of this energy is derived from installations of less than 50 kW [2]. In Germany, which has the largest proportion of residential PV systems among European countries, the PV power feed-in tariff is less than the electricity price at the household level [3].

In the UK, the feed-in tariff for installations smaller than 50 kW is much less than the electricity price and the PV generation tariffs have now stopped for new applications since April 2019 [4]. Also, since the uncontrolled nature of PV export will increasingly impact the grid, many countries have already imposed some constraints to limit the export of PV energy to the public grid. For example, in Germany, the maximum PV power feed-in is limited to 50% of the peak power [3]. Indeed, in recent years, many countries encourage the self-consumption of the electricity generated from PV systems instead of feeding it into the grid [5]. This new approach necessitates using energy storage mechanisms to make the most of the installed PV system. For example, approximately 50% of the new PV installations in Germany in 2018 are with a battery system [6]. Due to this shift in PV system approach, a lot of research focus targeted at distributed small-scale PV systems with storage units [7].

Obviously, the energy management strategy of a PV system equipped with battery storage is different from that of a direct grid-connected PV system. In this condition, any kind of prediction of the very short-term intra-hour PV power variations, sometimes called ramping events, enables optimal

The associate editor coordinating the review of this manuscript and approving it for publication was Canbing Li.

operation of the combined PV and storage units and makes the most of the investment. It has been shown that short-term forecast-based storage management can significantly reduce the extent of PV power curtailment, such that the total electricity price is halved for a typical combined 5kWp PV and 5 kWh battery residential system in Germany [3]. As well as optimizing the operation, it is also demonstrated that a short-term forecast of generation disturbances, with around 15-minute forecast horizon, is necessary to prepare the backup system of a microgrid with high PV penetration in order to avoid instability or black out events [8]. Recently, many regulatory market environments allow very short-term electricity trading to reduce the required reserve capacity. For instance, the Energy Imbalance Market in the US or the Australian Energy Market Operator cover forecasts down to 5 minutes [9], [10]. In such markets, therefore, any improvement in very-short term PV forecasts directly leads to an increase in revenue. Also, it is already proposed that the distributed PV systems can be diagnosed by continuously comparing the short-term power production predictions and the actual values. Subsequently, a mismatch indicates a probable problem in the system, e.g., due to unusual shading, dust on panels or contacts corrosion [11], [12].

A. PV FORECAST BACKGROUND

Although the necessity of accurate and reliable short-term forecasts is becoming more apparent, the solution is still a real challenge. For example, ''improving short-term forecasting of generation/supply and demand'' is announced as the second priority of the UK National Grid system operator innovation strategy in 2018 [13]. Currently, no documented standards or recommended practices or even widely accepted definitions are available. However, the most common industry-requested very short-term predictions are for time horizons from 15 minutes to 2 hours ahead, with resolutions of 30 seconds to 5 minutes [14], [15]. Alternatively, the day ahead forecasts for up to two days ahead or beyond are crucial for participation in electricity markets of large-scale renewable farms and are widely in use by PV farm owners, portfolio managers, grid operators and trading companies to decrease the amount of procured reserves that are necessary with a high penetration of intermittent energy sources [15]–[17]. The forecasting techniques can be broadly classified as physical (or deterministic), statistical (or probabilistic) and hybrid (or ensembled) methods [17]–[25]. The early methods were purely physical and calculated the PV power from meteorological, topographical and geomorphological data as well as the PV characteristic curves. The numerical weather prediction (NWP) models are mainly used to forecast the solar data by evaluating extensive dynamic equations subject to vast input data and initial conditions. They are complex and the performance is highly dependent on the weather forecast accuracy. Statistical techniques use the historical PV power data to train forecasting models, which are then used to forecast PV generation from past inputs. The hybrid methods use the output of NWP models or sky and satellite imagery as extra inputs for the probabilistic techniques to improve the accuracy of predictions, especially for short-term estimations. Intra-hour and intra-day forecasts are not currently as economically valuable as day-ahead forecasts; however, with ever-increasing penetration of combined solar and storage systems that demands smart energy management strategies, significant market opportunities are anticipated. For very short-term forecasts applying to a single point installation rather than a wide geographic area, pure statistical techniques seem to be the most accurate solutions [23], [26]. The most successful statistical approaches reported in literature for very short-term intra-hour PV forecast are time series [27], regression [28] and Markov chain [29]. These techniques can be considered as ensemble models that combine the classical statistical techniques with the nonlinear correlation and classification analyses. While these techniques can be capable of extracting information from high-dimensional meteorological data, they could not prove efficacy with very short-term intra-hour time steps. Artificial Neural Networks (ANNs) are also employed for very-short term PV predictions [22], [30]–[33]. Chu *et al.* [30] has successfully integrated the ANN to three different baseline statistical prediction models to enhance their forecast skills for time horizons of 5, 10, and 15 minutes. Their method simply uses the predictions from the baseline forecast model and the measured previous generation values as inputs to the ANN. Rana *et al.* [31] has also presented a hybrid ANN and Support Vector Regression algorithm to generate the 5-minute to 60-minute predictions from the previous PV power and the meteorological data. Sharma *et al.* [32] has proposed a mixed wavelet and ANN for 15-minute solar irradiance forecasting, in which the hidden layer neurons of the ANN uses the wavelets as activation functions. Wang *et al.* [33] has employed the wavelet transform to decompose the PV data into different frequencies before feeding them to the ANN predictor. The output of the ANNs at different frequencies will then go through the inverse transform and then composed to construct the final prediction. Common among all, the parameters of ANN and the learning algorithm are chosen from trial and error for the best performance and the optimal architecture of the ANN varies for different weather conditions. Recently, Asrari *et al.* [22] has used the gradient-descent and the meta-heuristic optimization techniques to search for the parameters of the ANN to predict 15-minute solar powers. Besides the added complexity and computational burden their training technique is still vulnerable to the overfitting problem of ANN-based methods.

Recent literature shows that the nonlinear prediction algorithms based on the ANNs have found popularity and perform best for very short-term forecasts among other competitors. However, some of the challenges with ANNs, that cause severe barriers in industrial applications of them, are (1) ANNs have various parameters to be set, especially those related to the network topology, which is not a trivial task to do; also, training results are usually non-deterministic and depend on the initial parameters [12], [18], (2) a serious problem in the training process is the possibility of overfitting, i.e. if the training process is not stopped at the right point, the ANN begins to "memorize" the training dataset rather than achieve "generalization" from it [22] and (3) if one decides to add new input-output data pairs to an existing ANN, then the whole training process must be repeated; so it is not possible to learn from new data while running the forecasts.

B. FUZZY LOGIC FOR PREDICTION PURPOSES

Like ANNs, fuzzy logic is an artificial intelligence (AI) technique, that has been proven to inherit the universal nonlinear mapping of input-output data pairs using soft linguistic variables [34]. Some successful applications of fuzzy logic in weather forecasting can be found in [35]–[38]. The main advantage of fuzzy-based prediction techniques is that they can provide comparable prediction results to the ANNs without suffering from the three major problems mentioned above.

A fuzzy logic system is composed of the Knowledge Base (KB) and the Inference System (IS). The KB contains the available knowledge about the problem in the form of linguistic IF-THEN rules. The IS makes decisions on the system outputs from inputs using the information in the KB. A general method called Wang-Mendel is widely used to generate fuzzy rules from numerical input-output data pairs [34]. It is a five-step procedure for building the KB from numerical data pairs. As a recent and successful work, reference [38] has proposed a fuzzy model for the very short-term prediction of wind speed, using the Wang-Mendel technique to generate the KB of a Mamdani-type IS predictor. It has improved the accuracy of forecasts and reduced the computational requirements by extracting some statistical properties of the wind speed information, such as the standard deviation, the average and the slope, and uses them instead of the measured wind speeds, as inputs to the fuzzy predictor. Since the number of inputs is limited to three, the size of the KB is also considerably reduced. So, wind speed statistical properties can be summarized in a smaller KB whilst at the same time the accuracy of the predictions is improved. However, the choice of the statistical properties used in [38] is purely arbitrary, which can reduce the accuracy of the forecasts.

Taking inspiration from [38] this paper proposes a new fuzzy predictor with an online additive learning capability for very short-term PV power generation that just uses the historical PV power generation data. The main contributions of this work are, first, unlike the previous work in [38], it investigates different statistical properties of historical data based on a scientific justification and proposes the best three statistical features to be used as the outputs of the preprocessor to the fuzzy predictor. Second, the conflict resolution logic of the classical Wang-Mendel rule generation algorithm is modified such that the rules, which are already considered as conflicting with the classical algorithm (i.e. same IF parts) are no longer ignored. Instead, they can cooperatively contribute to the output formation, i.e. the same input can trigger different outputs simultaneously but with different weights. Third, an online additive learning algorithm, based on the modified Wang-Mendel algorithm, is proposed that enables acquiring additional knowledge about the process while the predictions are running and at the same time makes sure that no available information in the newly available data pairs are ignored or lost. The theoretical achievements are experimentally proven on a real PV system in the UK. The rest of this paper is organized as follows. Section II presents the proposed technique. A comparative performance evaluation is then reported in Section III and finally, Section IV concludes the paper.

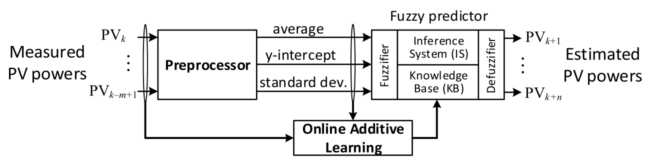


FIGURE 1. Proposed PV power predictor based on [38].

II. PROPOSED FUZZY PREDICTOR

The main structure of the proposed predictor is shown in Fig. 1. As it is illustrated in Fig.1, instead of directly feeding the measured PV power data (i.e. PV_k, PV_{k-1}, $PV_{k-2}, \ldots, PV_{k-m+1}$) to the predictor to estimate PV_{k+1}, $PV_{k+2}, \ldots, PV_{k+n}$, some statistical properties of time-series inputs are calculated and used as inputs to the fuzzy predictor. Although using statistical properties instead of the actual data was proposed in [38], that work does not provide any scientific justification on the selected statistical properties. While the previous work uses the average, the standard deviation and the slope of the measured data as the inputs, in this work, it became clear based on a Pearson correlation analysis (as will be demonstrated in the next subsection) that with the PV power data, the combination of the average, the standard deviation and the y-intercept offers much better prediction accuracy. The Wang-Mendel algorithm and the IS are also modified in this work. An online additive learning processor is also proposed that enriches and updates the KB while the prediction is running. More details are presented in section II.B.

A. SELECTED FEATURES

Different statistical features can be extracted from the almost non-stationary time series of PV power data. For the following analysis and the benchmark tests, the real PV power data measured at Swansea University in the UK (SPECIFIC Active Classroom) during 2018 are used. The one-minute measured data have been averaged over 5-minute intervals, which seems reasonable for the very unstable UK weather condition.

The number of outputs from the predictor is chosen to be three ($n = 3$) that means each run of the forecast algorithm provides PV power generation data for the 15-minute time ahead in three 5-minute consecutive steps. It has

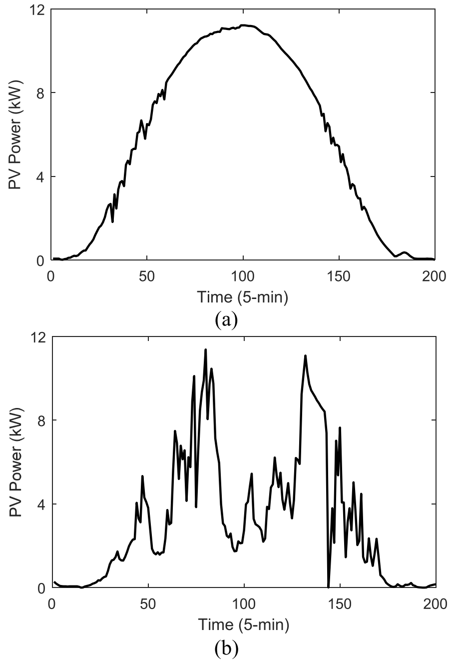


FIGURE 2. Sample days used for correlation analysis of different static characteristics: (a) sunny and (b) highly cloudy day.

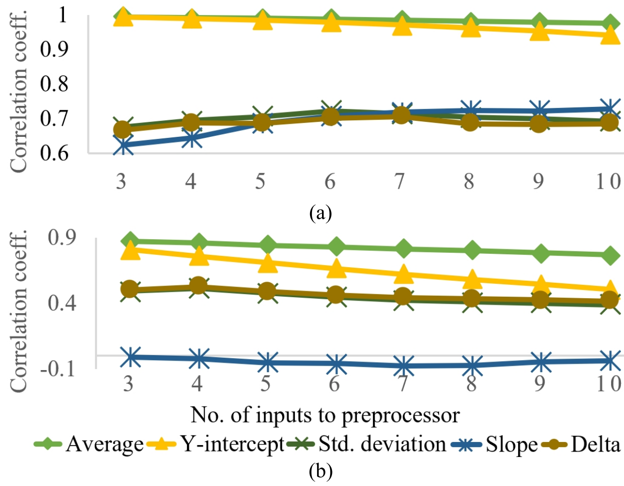


FIGURE 3. Correlation analysis of PV generation static characteristics (output to input data): (a) sunny and (b) highly cloudy day, $n = 3$ and m varies from 3 to 10.

previously been shown that the power generation forecasts for 15 minutes in advance are adequate to efficiently manage the backup system for a micro-grid [8], [28]. Also, the 5-minute resolution provides more accuracy in response to rapid weather changes, such as that experienced in the UK. Different statistical characteristics, including average, y-intercept, standard deviation, slope and delta are investigated as potential candidate features. The value of y-intercept and slope are calculated from the first order fitted line to the corresponding input data. The delta is defined as the difference between the maximum and the minimum of the input data. To find the best choices, a linear regression analysis using the well-known Pearson correlation coefficient was conducted between the stochastic values of the current and the two ahead data (output) and different numbers of lagged data points (input) for two different days of historical PV generation measurements. These two days, shown in Fig. 2, are selected from sunny and cloudy days. The other days have a power pattern between these two extreme samples. The results are plotted in Fig. 3. As well as selecting the best statistical characteristics to be used at the preprocessing stage, another important step in developing the forecasting algorithm is to determine the number of lagged (previous) data to be fed into the preprocessor, i.e. m in Fig. 1. For this purpose, the number of input samples in the correlation analysis was varied from 3 to 10. Clearly, for both days, the average and the y-intercept show a high correlation between the current/future and the lagged PV power data. The slope, which was previously proposed by Monfared *et al.* for the wind speed data in [38], is the worst among the studied characteristics for the PV data, especially under fast-changing conditions. The two other properties, i.e. the standard deviation and the delta, perform almost the same. Based on this analysis, one may choose the average and the y-intercept as the best characteristics for inclusion with the option of either standard deviation or delta or even both as extra characteristics. Indeed, by increasing the number of outputs from the preprocessor to the fuzzy logic, the size of KB increases, which demands higher computations. Also, with a higher number of characteristics to be extracted from the original data, a larger historical dataset is required in the early training stage to ensure that the KB has adequate information about the process to be forecast. Based on several test studies, the standard deviation, which shows slightly better performance in the correlation analysis with sunny days, is chosen in this work. Based on the correlation analysis, m is also chosen to be 5. It must be mentioned here that with the proposed prediction algorithm the KB continuously learns from new data while the forecasts are running based on the available information up to that instance of time. So, the size of KB increases gradually and therefore the quality of predictions improves from prediction to prediction.

B. MODIFIED WANG-MENDEL ALGORITHM

The details of the original Wang-Mendel technique can be found in [34]. This work proposes the following modifications to improve the performance of the predictor:

Firstly, the decision among the conflicting rules is now revisited in order to avoid any loss of information by omitting the rules that are already defined as conflicting in the original version of the algorithm. It is worth noting that the original Wang-Mendel algorithm interprets any two or more generated IF-THEN rules with the same IF parts as conflicting and only includes the one with the higher weight in the KB. The others will simply be ignored, which reduces the forecast accuracy. However, in the alternative algorithm, any new rule generated from new input-output data pairs that has the same IF part as another rule in the current KB is no longer classified as conflicting rule if the THEN parts are different. In doing this, it is accepted that the same input can trigger different outputs simultaneously but with different levels of effect (i.e. weight). Also, the IS is modified by using the sum aggregator instead of the maximum (one-winner) of the classical Mamdani-fuzzy version. Therefore, different rules with the same IF parts have the opportunity to contribute to the formation of the output according to their weights [39];

Secondly, to further prevent possible loss of information by ignoring the conflicting rules that have the same both IF and THEN parts, this paper proposes to modify the new conflicting rule such that it can be eventually added as a piece of new information to the KB. The proposed procedure can be summarized as below:

- find the input or the output variable with the minimum membership;
- for that variable, change the membership function (MF) to the adjacent one and update the membership value, which consequently changes the rule and its weight;
- if the new rule does not have the same both IF and THEN parts as another already available rule in the KB, then add it to the KB, otherwise, from the two conflicting rules keep the one with the higher weight.

The flowchart of the proposed fuzzy predictor with online additive learning is depicted in Fig. 4. The blocks in grey constitute the additive learning algorithm by the proposed conflict solving procedure. As it is illustrated, any new available datapoint is simultaneously employed to generate a possible new rule to be added as a piece of new information to the current KB. The proposed approach is especially valuable when the input-output pairs at the time of first training are not sufficient to successfully constitute a rich KB. Also, with the proposed online additive learning of the predictor, the maximum possible information is extracted from new examples. Moreover, this modification to the Wang-Mendel way of selecting the rules along with the sum aggregator offers inherent filtering of wild-data, i.e. the data-pairs that are generated as a result of noise or malfunction of equipment that can lead to an incorrect rule with a high weight. Thus, as the KB size increases, the effect of wild-data on the output formation decreases. Moreover, it can adaptively change the KB when new examples are available.

Evidently, after a while of running the prediction and at the same time learning (extracting new rules from new data), the KB size will increase sufficiently, and the accuracy will improve. While the shape and number of MFs are arbitrary, this work assumes 30 MFs as 50% overlapping isosceles triangles for all input and output variables. The more MF numbers mean the more prediction accuracy at the price of a larger KB dimension.

III. PERFORMANCE EVALUATION

In order to investigate the effectiveness of the proposed prediction strategy, many benchmark tests have been carried out using actual PV power data measured in Swansea University's Bay Campus in the UK for July 2018 and February 2019. These measured data have been averaged to obtain 5-minute interval data sets. For all studies, the measured data for the first week of July are used as the training data. Corresponding measurements for the next 13 days of July 2018 are used as the test data. Also, in order to verify the effectiveness of the proposed approach during different seasons, predictions for February 2019 are also provided. It is worth emphasizing again that the data of the first week of July 2018 are used for initial training of all predictors and for both prediction periods. In order to quantitatively examine the performance of different prediction methods, the root mean square error (RMSE), the mean absolute error (MAE), the normalized mean absolute error (NMAE), the standard deviation of error (STDE) and the coefficient of determination (COD), defined below [38], [40], are employed.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{PV}_i - \mathrm{PV}_{ip}\right)^2} \tag{1}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\mathrm{PV}_i - \mathrm{PV}_{ip}\right| \tag{2}$$

$$\mathrm{NMAE} = \frac{100}{N}\sum_{i=1}^{N}\frac{\left|\mathrm{PV}_i - \mathrm{PV}_{ip}\right|}{\mathrm{PV}_{inst}} \tag{3}$$

$$\mathrm{STDE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{PV}_i - \mathrm{PV}_m - (\mathrm{PV}_{ip} - \mathrm{PV}_{mp})\right)^2} \tag{4}$$

$$\mathrm{COD} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{PV}_i - \mathrm{PV}_{ip}\right)^2}{N-2} \Bigg/ \frac{\sum_{i=1}^{N}\left(\mathrm{PV}_i - \mathrm{PV}_m\right)^2}{N-1} \tag{5}$$

In the above equations, N is the total number of samples, PV_i and PV_{ip} are the actual and the predicted values of generated PV power data, PV_m and PV_{mp} are the mean of PV_is and PV_{ip}s, respectively and PV_{inst} is the maximum power capacity of the PV system. The ideal values for the RMSE (kW), MAE (kW), NMAE (%) and STDE (kW) are 0 and for the COD is 1, respectively. In order to better highlight the superior performance of the proposed method, three different fuzzy prediction strategies are implemented as follows:

- (Fuzzy-1) using the original PV data as inputs to the fuzzy predictor;
- Fuzzy-2) using the preprocessor with the proposed features, i.e. the average, the y-intercept and the standard deviation, as the inputs to the fuzzy system;
- (Fuzzy-3) using the proposed features, the same as the previous strategy, while simultaneously running the proposed online additive learning algorithm from the new data (test data).

The third strategy (i.e. Fuzzy-3) continuously learns from the new measured data, considered as new examples, as time goes on. In this way, after any 5 minutes, a new data pair is available to be amended, as a possible new rule, to the current KB. As a competitor, the classical multi-layer

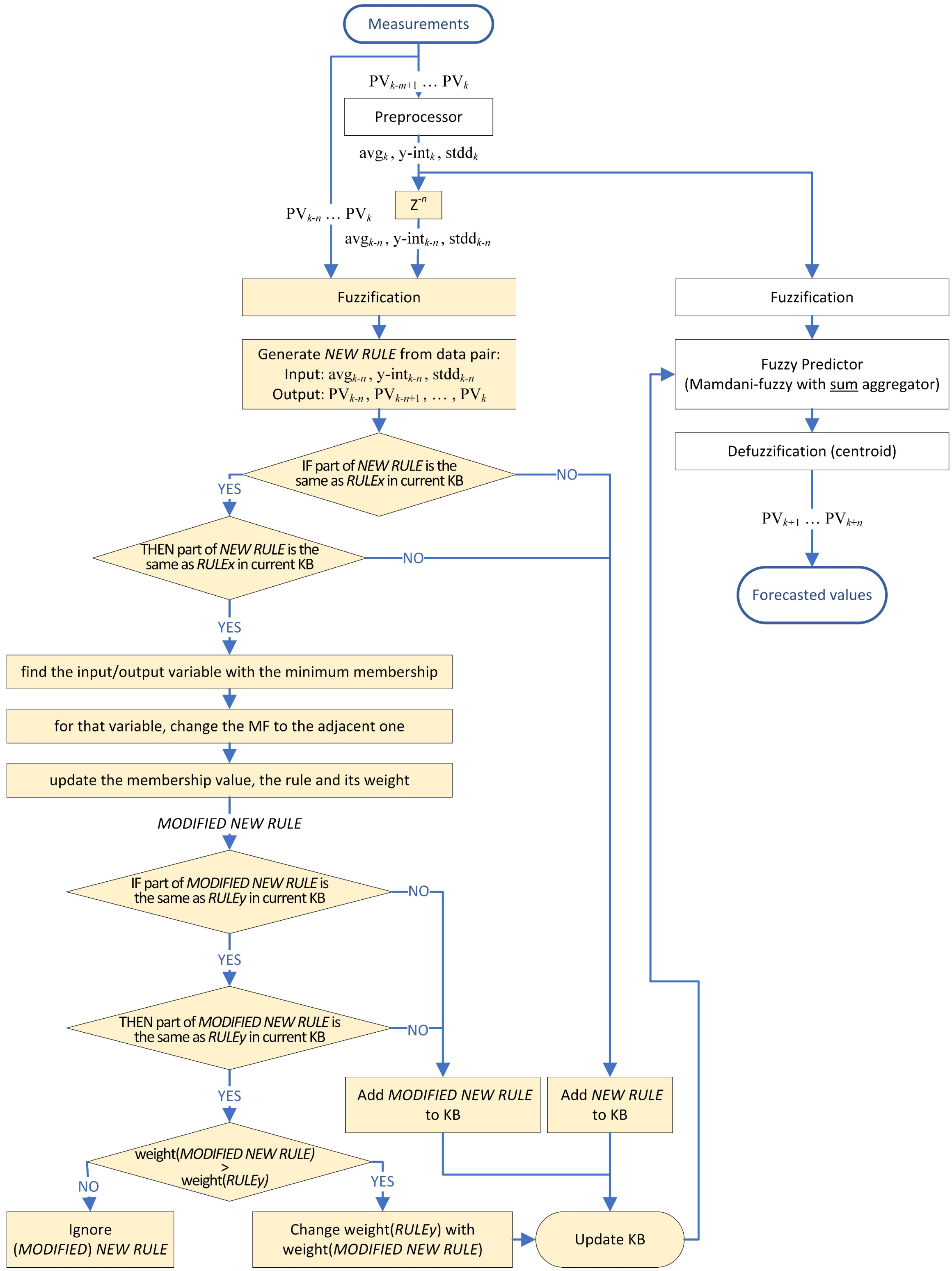


FIGURE 4. **Flowchart of the proposed fuzzy predictor with online additive learning ($n = 3$ and $m = 5$).**

feed-forward ANN, which is the most used in the literature for the PV prediction purposes, is implemented to be included in the benchmark study. The ANN predictor has a 30-node hidden layer and is trained by a standard back-propagation algorithm [41]. The training process of the ANN uses the actual PV power measurements as the input-output pairs. The numerical comparative results are summarized in Table 1. All metrices show that the prediction accuracy has increased with

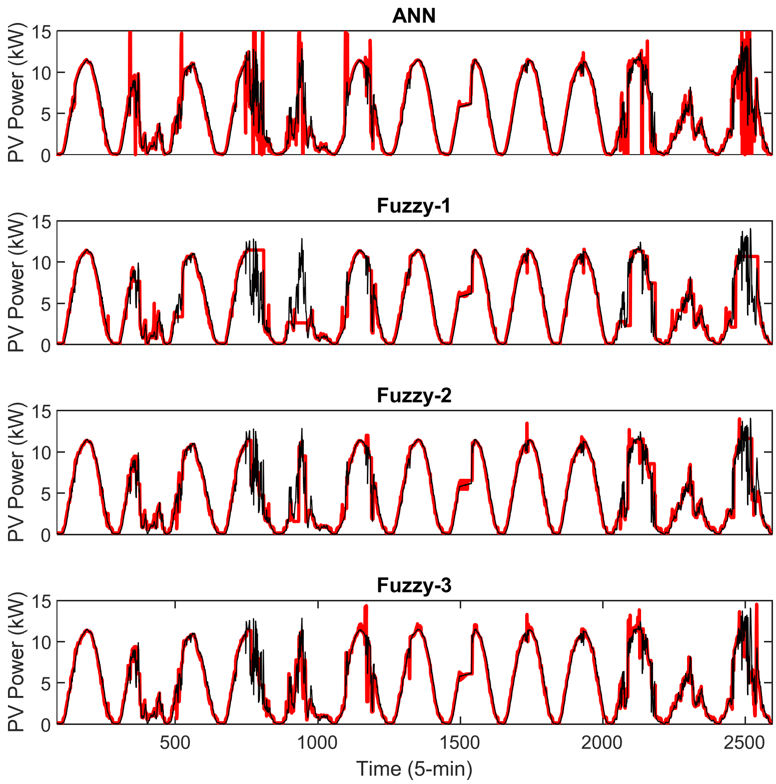


FIGURE 5. Predictions by different strategies for 8-20 July 2018: red (predicted) and black (actual).

TABLE 1. Statistical performance analysis.

Strategy	RMSE(kW)	MAE(kW)	NMAE(%)	STDE(kW)	COD
			July 2018		
ANN	2.05	0.71	3.92	2.05	0.784
Fuzzy-1	1.77	0.78	4.34	1.77	0.820
Fuzzy-2	1.43	0.70	3.88	1.43	0.879
Fuzzy-3	1.28	0.65	3.60	1.28	0.902
			February 2019		
ANN	1.64	0.78	4.34	1.63	0.583
Fuzzy-1	2.03	1.09	6.06	1.96	0.286
Fuzzy-2	1.52	0.86	4.78	1.51	0.537
Fuzzy-3	1.30	0.77	4.32	1.29	0.667

the preprocessor, i.e. strategies Fuzzy-2 and Fuzzy-3. The best results are obtained from the third fuzzy strategy, where the proposed online learning is also in progress during the predictions. While the effectiveness of the proposed additive learning technique is clear for the prediction period of July, it is better revealed for the prediction period of February. The predictors are all trained by using the July data and are now exposed to test with the February data. Clearly, the accuracy of predictions is decreased, because all strategies are providing predictions for a winter month whilst they have been trained with a summer month data. However, the Fuzzy-3, with additive learning, has the opportunity to learn from new data and consequently its prediction accuracy, in terms of all indices and for all cases in Table 1, is superior. Moreover, Fig. 5 compares the predicted PV power data with the actual values for the July days of test

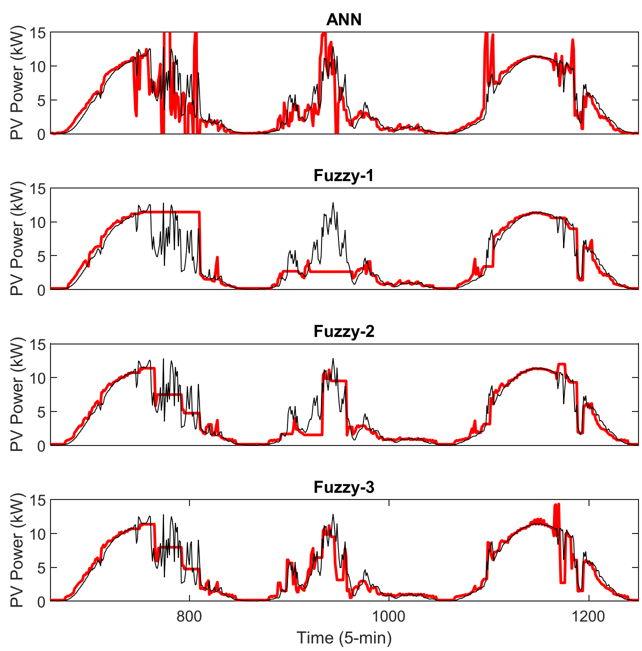


FIGURE 6. Zoomed view of predictions by different strategies for 11-13 July 2018: red (predicted) and black (actual).

and four different prediction techniques. For the sake of better presentation, the night-time data are removed. A zoomed view of days 4 to 6 and days 12 and 13 are also provided in Figs. 6 and 7, respectively. The ANN shows adequate

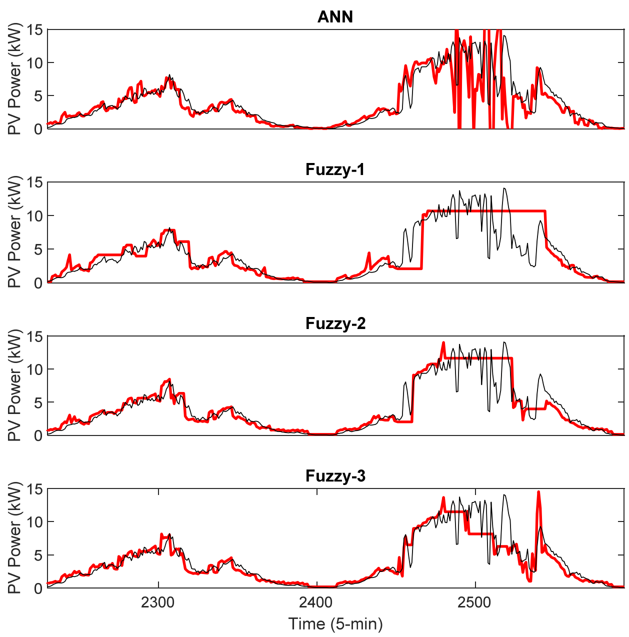


FIGURE 7. Zoomed view of predictions by different strategies for 19-20 July 2018: red (predicted) and black (actual).

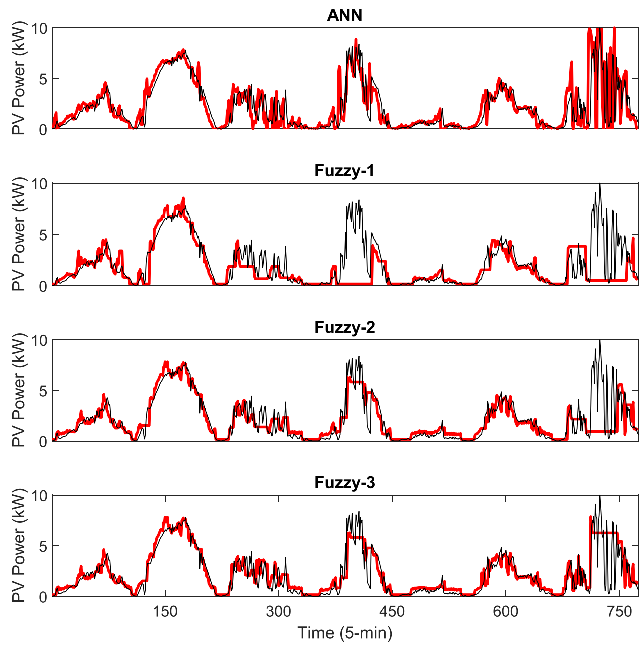


FIGURE 8. Predictions by different strategies for 1-7 February 2019: red (predicted) and black (actual).

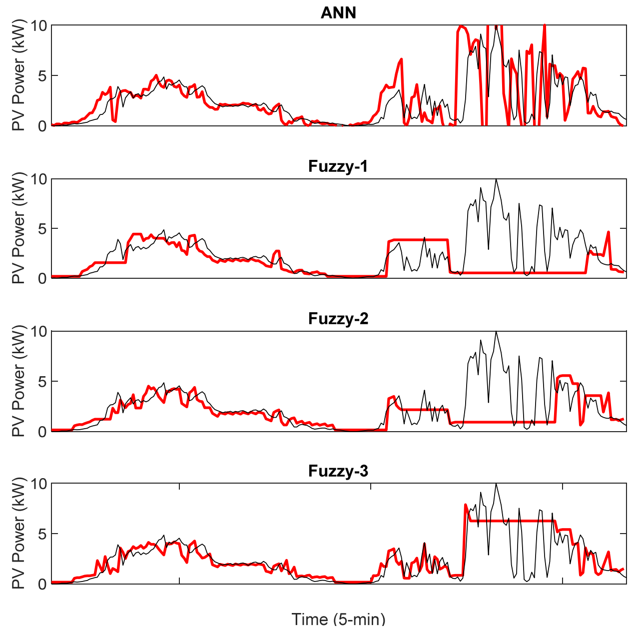


FIGURE 9. Zoomed view of predictions by different strategies for 6-7 February 2019: red (predicted) and black (actual).

prediction performance during the sunny days, but as can be readily realized from Figs. 6 and 7, it is unable to provide acceptable results under fast-changing conditions. Indeed, for these days, no historical information was available during the learning process and therefore the ANN is likely to generate incorrect outputs, even in excess of the dataset boundaries. The ANN just memorizes the relationship between the input and the output data without any interpretation capability and cannot deal with uncertainties. On the contrary, all fuzzy models can predict 5-minute PV power data adequately even with sudden falls and rises. The preprocessing enables the extraction of more knowledge about the process from the same amount of historical data available at the learning process. This directly translates to improved prediction capabilities, especially during fast-changing conditions. Also, the additive learning from new data continuously updates the KB at each prediction step with almost no effort and enables the enrichment of the available information about the process by using any available piece of recent measurement data. As can be seen in Fig. 5 (Fuzzy-3), the accuracy of predictions for the middle-day (day 5) has increased noticeably by learning from almost similar patterns of variations in the previous day. The efficiency of the additive learning is more evident in Fig. 7 (Fuzzy-3), where the predictions for the last test day (day 13) are considerably more accurate compared to the other strategies. Evidently, while the predictor goes on, more data becomes available about the current process to be included in the KB and consequently, the accuracy of predictions will increase gradually over time.

Figures 8 and 9 compare the predicted and actual PV power output during the first week of February 2019. An obvious difference to the training data is the lower level of PV power output. As already discussed, these figures again confirm the effectiveness of both preprocessing of PV data and active additive learning on providing acceptable predictions even under very different seasonal conditions in comparison to those conditions during the initial training process. Evident from Fig. 9, the Fuzzy-3 strategy can almost reach a very good prediction accuracy at the end of the 7-day February test period, which demonstrate the effectiveness of the proposed active learning strategy.

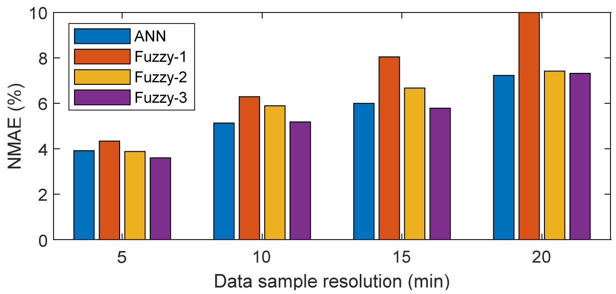


FIGURE 10. Performance comparison at higher prediction horizons.

As already mentioned in the introduction, the proposed algorithm (as well as the benchmarks) are pure statistical (probabilistic) techniques that only use the lagged PV power generation data as the input to the predictor. Keeping the same structure for all predictors, the interval of data samples is increased to 10, 15 and 20 minutes. The prediction errors, in terms of NMAE%, are calculated and compared in Fig. 10. As it can be seen, the increase of forecasting horizon results in higher forecasting errors. However, the proposed technique still reduces the error of the fuzzy predictor for all resolutions and provides the best accuracy among all benchmark techniques for intra-hour predictions.

IV. CONCLUSION

This paper proposed a fuzzy logic-based predictor for PV generation systems that uses the average, the standard deviation and the y-intercept of the actual data to define the KB. Moreover, an active additive learning feature is added to the conventional fuzzy-based method through modifying the Wang-Mendel algorithm. The main achievement and advantages of the proposed strategy can be listed as:

it offers a more accurate very short-term prediction, while the size of the KB and consequently the computational effort remains low;

it is a generalized technique and does not need any site-specific information;

it does not use any weather predictions; so, it is simpler to operate and immune to errors related to meteorological predictions;

it has the unique feature of an online additive self-learning process using new data; so, the prediction accuracy increases over time while it is running. Also, the KB will be adaptively updated to accommodate long-term changing conditions of weather or the PV system and its surroundings.

The experimental results of the prediction of 15-minute ahead PV power generation in the UK demonstrate the effectiveness of the proposed algorithm compared to the conventional fuzzy logic approach and the well-known ANN competitors. It presents a NRMSE of 3.6% and outperforms the benchmark scenarios in terms of different performance indices.

REFERENCES

[1] Allied Market Research. (2018). *Renewable Energy Market Size, Growth, Share and Analysis*. [Online]. Available: https://www.alliedmarketresearch.com/renewable-energy-market

[2] *Solar Photovoltaics Deployment—GOV.UK*. [Online]. Available: https://www.gov.uk/government/statistics/solar-photovoltaics-deployment

[3] E. Lorenz and G. Arne, "PV power forecasting: Applications and value," in *Proc. IEA PVPS Task Workshop Sol. Forecast Requirements Value Grid Appl.*, 2018.

[4] *Feed-in Tariff Scheme | Energy Saving Trust*. Accessed: Nov. 23, 2018. [Online]. Available: https://www.energysavingtrust.org.uk/renewable-energy/electricity/solar-panels/feed-tariffs

[5] V. Bertsch, J. Geldermann, and T. Lühn, "What drives the profitability of household PV investments, self-consumption and self-sufficiency?" *Appl. Energy*, vol. 204, pp. 1–15, Oct. 2017.

[6] J. Binder. (2018). PV Storage and Own-Consumption in the Residential Sector in Germany. Paris, [Online]. Available: https://www.energie-fr-de.eu

[7] *New Active Building Centre to Transform the UK Construction and Energy Sectors Through the Deployment of Solar-Powered Buildings—EPSRC Website*. Accessed: Feb. 11, 2019. [Online]. Available: https://epsrc.ukri.org/newsevents/news/activebuildingcentre/

[8] E. Rikos, S. Tselepis, C. Hoyer-Klick, and M. Schroedter-Homscheidt, "Stability and power quality issues in microgrids under weather disturbances," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 1, no. 3, pp. 170–179, Sep. 2008.

[9] A. Kaur, L. Nonnenmacher, H. T. C. Pedro, and C. F. M. Coimbra, "Benefits of solar forecasting for energy imbalance markets," *Renew. Energy*, vol. 86, pp. 819–830, Feb. 2016.

[10] Australian Energy Market Operator. *Semi-Scheduled Generation Dispatch Self-Forecast—Assessment Procedure*. Accessed: Feb. 11, 2019. [Online]. Available: https://www.aemo.com.au

[11] M. Gulin, T. Pavlović, and M. Vašak, "A one-day-ahead photovoltaic array power production prediction with combined static and dynamic on-line correction," *Sol. Energy*, vol. 142, pp. 49–60, Jan. 2017.

[12] L. Gigoni, A. Betti, E. Crisostomi, A. Franco, M. Tucci, F. Bizzarri, and D. Mucci, "Day-ahead hourly forecasting of power generation from photovoltaic plants," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 831–842, Oct. 2018.

[13] National Grid. (2018). *National Grid System Operator Innovation Strategy*. [Online]. Available: https://www.nationalgrideso.com/innovation

[14] V. Kostylev and A. Pavlovski, "Solar power forecasting performance–towards industry standards," in *Proc. 92nd Amer. Meteorolog. Soc. Annu. Meeting*, 2012, pp. 1–8.

[15] D. K. Chaturvedi and I. Isha, "Solar power forecasting: A review," *Int. J. Comput. Appl.*, vol. 145, no. 6, pp. 28–50, 2016.

[16] *Power Forecast Services—ConWX*. Accessed: Nov. 20, 2018. [Online]. Available: https://www.conwx.com/services/power-forecasts

[17] *Photovoltaic and Solar Forecasting*, document IEA PVPS T14-01, 2013.

[18] H.-T. Yang, C.-M. Huang, Y.-C. Huang, and Y.-S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 917–926, Jul. 2014.

[19] M. Yang and X. Huang, "Ultra-short-term prediction of photovoltaic power based on periodic extraction of PV energy and LSH algorithm," *IEEE Access*, vol. 6, pp. 51200–51205, 2018.

[20] X. G. Agoua, R. Girard, and G. Kariniotakis, "Probabilistic models for spatio-temporal photovoltaic power forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 780–789, Apr. 2019.

[21] P. Lin, Z. Peng, Y. Lai, S. Cheng, Z. Chen, and L. Wu, "Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets," *Energy Convers. Manage.*, vol. 177, pp. 704–717, Dec. 2018.

[22] A. Asrari, T. X. Wu, and B. Ramos, "A hybrid algorithm for short-term solar power prediction—Sunshine state case study," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 582–591, Apr. 2017.

[23] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy Convers. Manage.*, vol. 156, pp. 459–497, Jan. 2018.

[24] U. K. Das, K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. Van Deventer, B. Horan, and A. Stojcevski, "Forecasting of photovoltaic power generation and model optimization: A review," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 912–928, Jan. 2018.

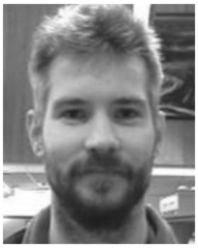
[25] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016.

[26] H. T. C. Pedro and C. F. M. Coimbra, "Assessment of forecasting techniques for solar power production with no exogenous inputs," *Sol. Energy*, vol. 86, no. 7, pp. 2017–2028, 2012.
[27] A. Tascikaraoglu, B. M. Sanandaji, G. Chicco, V. Cocina, F. Spertino, O. Erdinc, N. G. Paterakis, and J. P. S. Catalão, "Compressive spatio-temporal forecasting of meteorological quantities and photovoltaic power," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1295–1305, Jul. 2016.
[28] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, and S. Wang, "Short-term solar power forecasting based on weighted Gaussian process regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 300–308, Jan. 2018.
[29] M. J. Sanjari and H. B. Gooi, "Probabilistic forecast of PV power generation based on higher order Markov chain," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2942–2952, Jul. 2017.
[30] Y. Chu, B. Urquhart, S. M. I. Gohari, H. T. C. Pedro, J. Kleissl, and C. F. M. Coimbra, "Short-term reforecasting of power output from a 48 MWe solar PV plant," *Sol. Energy*, vol. 112, pp. 68–77, Feb. 2015.
[31] M. Rana, I. Koprinska, and V. G. Agelidis, "Univariate and multivariate methods for very short-term solar photovoltaic power forecasting," *Energy Convers. Manage.*, vol. 121, pp. 380–390, Aug. 2016.
[32] V. Sharma, D. Yang, W. Walsh, and T. Reindl, "Short term solar irradiance forecasting using a mixed wavelet neural network," *Renew. Energy*, vol. 90, pp. 481–492, May 2016.
[33] H. Wang, H. Yi, J. Peng, G. Wang, Y. Liu, H. Jiang, and W. Liu, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Convers. Manage.*, vol. 153, pp. 409–422, Dec. 2017.
[34] L.-X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 6, pp. 1414–1427, Nov. 1992.
[35] A. T. Eseye, J. Zhang, D. Zheng, H. Li, and G. Jingfu, "Short-term wind power forecasting using a double-stage hierarchical hybrid GA-ANFIS approach," in *Proc. IEEE 2nd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Chengdu, China, Apr. 2017, pp. 499–503.
[36] W. Zou, C. Li, and P. Chen, "An inter type-2 FCR algorithm based T-S fuzzy model for short-term wind power interval prediction," *IEEE Trans. Ind. Informat.*, to be published.
[37] L. Suganthi, S. Iniyan, and A. A. Samuel, "Applications of fuzzy logic in renewable energy systems—A review," *Renew. Sustain. Energy Rev.*, vol. 48, pp. 585–607, Aug. 2015.
[38] M. Monfared, H. Rastegar, and H. M. Kojabadi, "A new strategy for wind speed forecasting using artificial intelligent methods," *Renew. Energy*, vol. 34, no. 3, pp. 845–848, 2009.
[39] D. Alvarez-Estevez and V. Moret-Bonillo, "Revisiting the Wang–Mendel algorithm for fuzzy classification," *Expert Syst.*, vol. 35, no. 4, 2018, Art. no. e12268.
[40] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and meteorological information," *Renew. Energy*, vol. 118, pp. 357–367, Apr. 2018.
[41] G. Notton, C. Paoli, L. Ivanova, S. Vasileva, and M. L. Nivet, "Neural network approach to estimate 10-min solar global irradiation values on tilted planes," *Renew. Energy*, vol. 50, pp. 576–584, Feb. 2013.

MOHAMMAD MONFARED (S'07–M'10–SM'15) received the B.Sc. degree in electrical engineering from the Ferdowsi University of Mashhad, Iran, in 2004, and the M.Sc. and Ph.D. degrees (both with Hons.) in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 2006 and 2010, respectively. He is an Associate Professor with the Ferdowsi University of Mashhad, Iran, and currently a Visiting Researcher with the Department of Electrical and Electronics Engineering, Swansea University, Swansea, U.K. His research interests include power electronics, renewable energy systems, and power quality.

MEGHDAD FAZELI received the B.Sc. degree in electrical engineering from the Chamran University of Ahvaz, Ahvaz, Iran, in 2004, and the M.Sc. and Ph.D. degrees both from Nottingham University, Nottingham, U.K., in 2006 and 2010, respectively. Since 2011, he has been with Swansea University, Wales, Swansea, U.K., where he is a Senior Lecturer in electrical power engineering. His current research interests include virtual synchronous machines control, energy management, and forecasting algorithms.

RICHARD LEWIS received the B.Eng. and Ph.D. degrees from Swansea University, U.K., where he continues to hold a position as a Senior Technology Transfer Fellow within SPECIFIC, College of Engineering. His current research interests include real-time data acquisition from disparate systems such as energy generation and storage systems together with environmental monitoring and distributed computational infrastructure.



JUSTIN SEARLE received the bachelor's degree in analytical chemistry and the Ph.D. degree in engineering in photostability of PVC paints from Swansea University. On completion of his doctorate, he worked with Tata Colors, Shotton Works, Deeside. He is currently the Industrial Technology Director with SPECIFIC, Swansea University, U.K., leading the technical delivery teams for buildings, systems integration and technology demonstration, responsible for technology demonstration, and delivery at building scale.