Received: 11 October 2021 | Revised: 15 February 2022 | Accepted: 17 May 2022
DOI: 10.1002/we.2760

RESEARCH ARTICLE

WILEY

Predictive analytics beyond time series: Predicting series of events extracted from time series data

Sambeet Mishra[1,2] | Chiara Bordin[3] | Kota Taharaguchi[1] | Adri Purkayastha[4]

[1]Tallinn University of Technology, Tallinn, Estonia

[2]Danish Technical University, Kongens Lyngby, Denmark

[3]UiT, The Arctic University of Norway, Tromsø, Norway

[4]BNP Paribas, London, UK

Correspondence
Chiara Bordin, UiT, The Arctic University of Norway, Tromsø, Norway.
Email: chiara.bordin@uit.no

Funding information
Estonian Research Council, Grant/Award Number: PUTJD915

Abstract

Realizing carbon neutral energy generation creates the challenge of accurately predicting time-series generation data for long-term capacity planning and for short-term operational decisions. The key challenges for adopting data-driven decision-making, specifically predictive analytics, can be attributed to data volume and velocity. Data volume poses challenges for data storage and retrieval. Data velocity poses challenges for processing the data near real time for operational decisions or for capacity building. This manuscript proposes a novel prediction method to tackle the above two challenges by using an event-based prediction in place of traditional time series prediction methods. The central concept is to extract meaningful information, denoted by events, from time-series data and use these events for predictive analysis. These extracted events retain the information required for predictive analytics while significantly reducing the volume of the velocity of data; consequently, a series of events present the information at a glance, effectively enabling data-driven decision-making. This method is applied to a data set consisting of six years of historical wind power capacity factor and temperature measurements. Deploying five deep learning models, a comparison is drawn between classical time-series predictions and series of events predictions based on computational time and several error metrics. The computational analysis results are presented in graphical format and a comparative discussion is drawn on the prediction results. The results indicate that the proposed method obtains the same or better prediction accuracy while significantly reducing computational time and data volume.

KEYWORDS
features extraction, green computing, machine learning, multivariate time-series, predictive analytics, renewable energy, time-series forecasting, virtual power plants, wind energy

1 | INTRODUCTION

The total electricity production from wind resources increased by 170 TWh in 2020 with 11% rate of growth from 2019, as reported by International Energy Agency (IEA).[1] This pattern of growth can be observed for both onshore and offshore installations with off-shore installations leading at 29% rate of growth. A considerable growth of large-scale offshore wind farms installation is noticeable in Europe at 18.5 GW, and it is projected that the installed capacity will continue to increase, as among others, the capacity factor of the off-shore is higher than onshore wind turbines. Then, European Union (EU) aims to achieve about 100 GW of offshore wind capacity installation by 2030.[2]

Wind being an uncertain and variable resource, precise and accurate projection of wind power production is crucial for planning to build new capacity, operations and maintenance scheduling. Several studies have been conducted exploring analytical tools and improvements to existing ones in the literature. Following texts lists review of several of such tools to provide an impression of the methods. Predictive analytics of wind power data is used for operational control of energy storage units.[3] A strategy for offshore wind park maintenance is presented in Zhou and Yin.[4] In Mujeeb et al,[5] deep learning models are used for predictive analytics of wind power production, and in Jamil et al,[6] a predictive analytics model is proposed to facilitate peer to peer energy trading. From the above examples, it is clear that predictive analytics is used at all the various stages of wind park planning, operations, and maintenance.

As discussed in Lerner et al,[7] accurate forecasting of wind energy has important implications, as it reduces imbalance charges and penalties, it provides a competitive knowledge advantage in real-time "balancing" and in "day ahead" energy markets tradings, and it allows for more efficient project construction, operations, and maintenance planning. Renewable energy forecasting in general, and wind energy forecasting in particular, is traditionally performed in the literature in terms of "time-series predictions."

In the last decade, there is a shift from traditional time-series tools such as auto-regressive moving average (ARIMA) for statistical inference of patterns in the historical data to machine learning tools as it becomes more efficient and widely available.[8] A comprehensive review of intelligent predictors in the field of wind energy forecasting is provided in Liu et al,[9] where the authors discuss works that discuss four types of shallow predictors (artificial neural network, extreme learning machine, support vector machine, and fuzzy logic model) and four types of deep learning-based predictors (autoencoder, restricted Boltzmann machine, convolutional neural network, and recurrent neural network). Traditional methodologies implemented in literature can be classified as deterministic forecasting and probabilistic forecasting. While deterministic forecasting gives a certain predicted value at a specified time, the probabilistic forecasting provides a prediction interval or expected value and a probability density function. The latter is therefore applicable to predictions with implicit uncertainty and risk control; however, unreasonable assumptions about the distribution may introduce errors.[10] A review of deterministic wind energy forecasting is proposed in Liu et al,[9] while a review of probabilistic forecasting for wind power generation can be found in Zhang et al.[11] A very recent work reviewing models, methods and future research of both deterministic and probabilistic wind power forecasting is proposed in Bazionis and Georgilakis.[12]

Based on the reviews proposed in the listed literature, it is evident that deterministic forecasting models have been widely used over the last few decades for wind power predictions. Among these, the most popular are conventional time series based models, such as autoregressive moving average,[13] autoregressive integrated moving average,[14] and Grey method,[15] as well as artificial intelligence-based models,[16] like artificial neural networks[17] and support vector machine.[18] Continuous technological development has also led to new methodologies being considered for wind power forecasting, like deep learning for example.[19] The challenges related to the uncertainty of predictions in wind power forecasting have gained great interest in the recent literature. For this purpose, various models have been proposed, like kernel density estimation,[20] quantile regression,[21] lower upper bound estimation,[22] bootstrap,[23] and ensemble-based methods.[24] As shown in the literature, probabilistic forecasting models are more difficult to evaluate as compared to deterministic models; however, specific metrics (namely, reliability, sharpness, resolution, and the overall skill score) have been developed in order to improve the implementation and validation of probabilistic forecasting approaches.

1.1 | Main methodological challenges in wind power forecasting

Based on the available literature, two main challenges can be identified when it comes to wind power predictions: accuracy and computational resources.

In terms of accuracy, one common conclusion in the literature is that wind power generation is a non-linear and non-stable process, and therefore more research should be done to handle the wind power uncertainty within forecasting methodologies. Li et al[25] propose a short-term wind power forecasting method where the prediction accuracy approaches 90%; however, the authors conclude that while the proposed model can improve the accuracy of prediction, there are still some limitations, especially when it comes to long-term forecasting. A novel decomposition approach, which takes the chaotic nature of wind power time series into account and improves the prediction accuracy, is proposed in Safari et al.[26] Here, the chaotic components become more predictable by eliminating high-frequency variations that are small in amplitude. With this approach the model accuracy is improved when compared with other state of the art models. A review of hybrid empirical mode decomposition models for wind speed and wind power prediction is proposed in Bokde et al.[27] According to this review, wind speed and wind power prediction techniques can be classified into three different groups: the physical, the statistical, and the intelligent approach. It should be noted, however, that many researchers hypothesize hybrid methods constitute a fourth approach and have observed that these hybrid methods demonstrates a better performance when it comes to predictions accuracy. It is nonetheless difficult to quantify the overall improvement in prediction accuracy with hybridized methods, but the review confirms the popularity and penetration of hybrid neural network and support vector machine based models in wind speed and power predictions.

In essence, improving the accuracy of existing deterministic forecasting models as well as developing more advanced probabilistic forecasting models are understood as the most promising research directions for the scientific community. Then, new forecasting products are required to more effectively handle probabilistic information.[28]

In terms of computational resources, traditional wind power predictions based on time-series data-set are characterized by high volume (many years of historical data with low resolution), velocity (factoring in the real-time component), variety (associated with highly variable weather and highly unstructured measurements gathered through the sensors), and veracity (although not all the points in the time-series may be meaningful and there is a need for pre-processing before predictions). These characteristics place these predictions within the recent domain of big energy data.[29]

The challenge of execution time due to data volume and real-time machine learning applications has gained attention in literature, especially in recent years, due to an increased availability of historical and real time data. Execution time improvements have been achieved, for example, in Naik et al[30] by using hybrid empirical mode decomposition and kernel ridge regression, in Naik et al[30] by using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network, or in Saroha and Aggarwal[31] by using wavelet transforms and neural networks with tapped delay.

The literature implies that the nature of big data that rests behind the wind power predictions raise computational challenges. Indeed, when predictions are based on time-series data, a high availability of historical data is key for improving accuracy, albeit the computational resources required increase accordingly. Balancing prediction accuracy and computational resources is certainly a key challenge in current energy research.[32] Moreover, as the amount of data increases over time, the available methods tend to predict energy consumption with lower accuracy and/or with much higher completion times.[33]

1.2 | Virtual power plants, a challenging application for wind forecasting

Virtual power plants (VPP), which form a local energy markets, contain a very high share of variable and uncertain renewable energy resources (such as wind) as well as responsive energy demand. Real-time data processing and accessibility would provide information that could empower VPP owners, stakeholders, and decision makers alike.[34] The existing processing methods currently in use are computationally resource intensive and take longer processing time. In addition to problems with energy data, the communication infrastructure is operating at over capacity, and so it is lagging behind in the ability to enable services such as demand response.[35,36] The decision-makers are keen on identifying trends and patterns (information) from historical time-series data in order to perform predictive analytics of future scenarios (insight). For instance, identifying the timing and duration of peak demand is of significance to the decision-maker who is looking at a time-series data and must make decisions in a short time window, thereby making near-real time or real time decision a key issue. The models also influence policy making in terms of frequency, level of details, and target setting.[37] Models can also enable and support ambitious targets such as CO_2 reduction. In addition, transparency in operation and a simplification of the results would lead to inclusive policy design. Simplifying the whole time series into a set of events might also lead to clarity in decision-making. Therefore, data volume and velocity are among the key bottle necks for VPP applications.

1.3 | Overcoming forecasting challenges: extracting more out of the data

One key pathway for improving wind power forecasting has been identified in "extracting more out of the data."[38] The ability to identify key features intrinsic to the data set can indeed lead to improved and more accurate predictions. Features extracted from wind ramp events from a virtual wind park have been discussed in Mishra et al[39] where the authors present a novel ramping behavior analysis (RBA) method that identifies and quantifies variations in a time-varying data set. Two types of events are identified: significant and stationary events. Significant events refer to significant variations beyond a set threshold range, and stationary events refer to variations that mostly happen within the threshold limits. The authors associate the following features with each event: start time, end time, peak value, change in magnitude, the persistence of an event, the angle at which the event took place, and frequency of occurrences of the features. Once significant events are identified within a time-series data set, the next step is to utilize them to perform predictions.

While there are many works in the literature performing predictions based on traditional time-series data (which we will refer to as "traditional time series prediction"), to the authors' knowledge, there are currently no works in literature that perform predictions based on significant events extracted from time-series data. Such types of predictions would be beneficial to generate a better dataset that can be used as input to mathematical optimization models for smart energy and power systems.[40] Indeed, mathematical optimization has been widely used for decision-making within both electrical and thermal energy systems.[41] However, the traditional approach is to input a dataset in the form of time series, with hourly or finer resolution, that negatively affects the solution time. In addition, most of these models involve a high number of binary variables, as well as uncertainty. Many models are technology-oriented and they include details of specific technologies, often leading to non linear formulations (see, for instance, Ibrahim et al.[42]). When combined with multihorizon decision making such as Bordin et al[43] and Bordin and

Tomasgard,[44] they can easily become intractable. For these types of optimization models, the computational time of solving big instances is usually a challenge. An event-based dataset, that extracts only relevant events from the full dataset, would contribute to a faster solution time, compared to the traditional time-series dataset that is normally utilized.

The importance and the value of coupling predictions with decision support systems tools based on mathematical optimization have been discussed in Bordin et al.[45] It is the fundamental concept of the modern applications of prescriptive analytics, that is at the top of the analytics value escalator. However, predictions must be performed in ways suitable for inclusion within mathematical optimization models such that the generated dataset does not negatively impact the solution time. This is where predictive analytics based on events extracted from time-series data can greatly contribute to the next more advanced level of prescriptive analytics where the final decisions are taken utilizing the core subject of mathematical optimization.

1.4 | Hypothesis and contributions

This research work hypothesizes that an event based prediction method can reattain or be more effective than the time-series based methods while using deep learning models. Systematic computational experiments are conducted to test and validate the hypothesis. The primary contribution of this work is to introduce an event based prediction method using events extracted from time-series data. The secondary contribution of this paper is to present a comparative investigation between predictions using time-series and events from time-series as inputs. Then discuss the proposed novel predictive analytics product that is being developed as an extension of the research works in Mishra et al.[39,46] The first research work introduces RBA method with an application to wind power production that identifies and extracts events (significant and stationary) from the time-series data (wind power generation). The second proposes a comparison of a selection of established deep learning models for short and long term predictions of time series wind power generation and temperature data-set. Essentially, this paper investigates how events using RBA method can be utilized and measures the efficiency on time-series data (wind power production).

The main reasons why it is worthwhile to predict events instead of time-series can be summarized as follows. Not all the data points of a time-series data set may be relevant for the specific purpose of a certain predictions. Indeed, new forecast products should include probabilistic information, but should deliver it in a way that is tailored to the end user and their specific decision making problems.[28] From this point of view, identifying and predicting significant events is more meaningful for delivering predictions that matter for the specific needs of the final users. The accuracy of predictions is therefore higher when the focus is on significant events (which matter for a specific end user application), rather than on a full time-series data set that contains many data points that might be not relevant for the end user application. Moreover, the size of a data set containing significant events is notably smaller than the size of a full time-series data set, having immediate positive implications for reduced complexity and faster computations.

Descriptive statistics explains the underlying distribution of a data set; however, descriptive statistics is restricted by a set number of features that describe the statistical properties of a specific time-series. The proposed method extracts events from a time-series that depends on a threshold and cut-off frequency.[39] For instance, if the threshold is changed, the same data set is then represented by an alternative set of events. This implies that the user can define what counts as a significant and stationary event for the specific application. Thereby the proposed RBA method is a fast and flexible measure for extracting information from time-series data, and it is especially useful for application to a chaotic time-series, as this proposed method can reduce the complexity in events.

A highlight of key contributions are summarized as follows:

- Propose a novel predictive analytics method that is based on predicting events extracted from time series data.
- Perform a comparative analysis between predicting series of events extracted from time-series versus raw time-series data while using deep learning models.
- Demonstrate that predicting series of events maintain and at times excel in accuracy and precision matrices in comparison with predicting time-series.
- Demonstrate that predicting series of events is computationally faster and less resource intensive (memory requirements and computations), therefore more efficient than predicting time-series (faster computation due to reduction in size and complexity).
- Quantify the amount of energy consumed in predicting series of events and time-series data.
- Contribute to the domains of green computing[47] and green artificial intelligence[48] by proposing more efficient forecasting method and tools that enhance the sustainability within the domain of computer science.

Based on the contributions outlined above, the five deep learning models developed in Mishra et al[46] are applied to predict both conventional time-series and more innovative series of events. The five models used in this investigation are deep feed forward (DFF),[49,50] deep convolutional network (DCN),[51,52] recurrent neural network (RNN),[53,54] attention mechanism (Attention),[55,56] and long short-term memory networks

(LSTM).[57,58] These models are classified as deep learning models which are in turn a part of the machine learning field and therefore tightly connected to artificial Intelligence based forecasting. For comparison purposes, four established error metrics are used.

2 | PREDICTION OF EVENTS

2.1 | Extraction of events from time series wind power production

Wind ramp events are extracted using the RBA_θ methodology proposed in Mishra et al.[39] A ramp event can be stated as a swing across the reference axis. RBA_θ classifies the ramp events into significant and stationary events. Significant events are those with sudden and substantial swings, while stationary events have relatively smaller magnitude swings and are persistent over time. This study focuses on the significant events that exist in time series wind power production data. The features describing an event, classified as a significant event, are graphically drawn in Figure 1. Each significant event is described through five features. $w_s(t)$ measures the peak value of a ramp event, which is the peak value of capacity factor at the time of vertex. t denotes the time when the ramp event occurred. Δw_s measures the magnitude of a ramp event, which is value change between the set threshold and current value of wind capacity factor. Δt measures the length of time a ramp event has persisted. $\theta^{\Delta W_s}$ is the angle from the first departure point in contact with the threshold to the peak point $w_s(t)$.

The RBA_θ algorithm is applied to the wind power production data with hourly resolution for 5 years (2011–2016) to extract a series of events. The wind power production data is localized to Estonia and comes from the previous studies.[59] The statistical properties of the features explaining the events are depicted in Table 1.

There are 6131 significant events extracted from 52,500 time-series data points. The peak capacity factor is 0.97, and there is an 88% reduction in the amount of data points. The last event occurred at the 52,500 h or on the 6th year. The highest shift in wind capacity factor was 0.83, and the longest event lasted for 56 h (2.3 days). The highest angle was 45.62. Note that all the events are significant events, as stationary events are not considered for this analysis. More details related to the two types of events are presented here.[39]

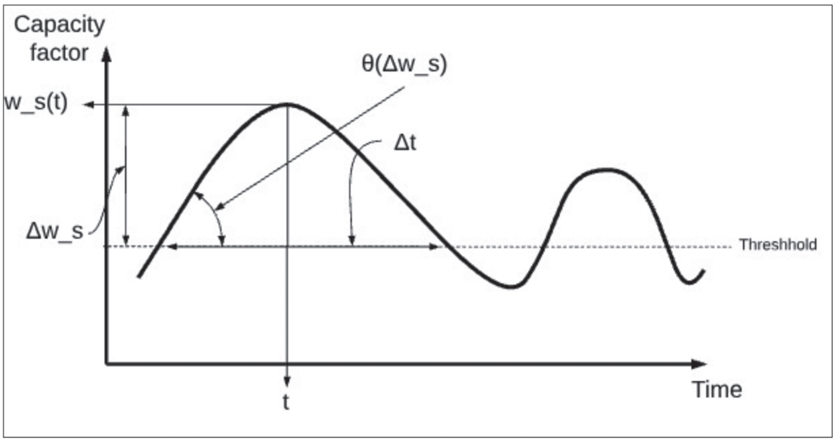


FIGURE 1 Concept diagram of RBA_θ features describing a significant event

TABLE 1 Statistical attributes of the significant ramp event features

Statistical properties	w_s	t	Δw_s	Δt	$\theta^{\Delta W_s}$
count	6131	6131	6131	6131	6131
mean	0.25	26326.4	0	9.53	0.37
std	0.2	15242.5	0.22	6.6	10.72
min	0	0	-0.95	1	-36.37
0.25	0.09	13124	-0.1	5	-7.36
0.5	0.2	26469	-0.02	8	-1.58
0.75	0.37	39550	0.11	13	8.05
max	0.97	52500	0.83	56	45.62

2.2 | Statistical analysis of the event features

Hourly time series data about wind power capacity factors and temperature in year 2016 is predicted using historical data from 2011 to 2015. Then the predictions are tested by matching with the statistical properties of the original records.

The wind power and temperature input data are localized[59,60] for Tallinn, Estonia. The data contains global re-analysis models and satellite observations. The wind power capacity factor (WF) and temperature (TEMP) data are retrieved for five years from 2011 to 2016. The data are segregated into training set 2011–2015 and testing set 2016, for model validations. Table 2 presents the statistical information of the wind power capacity factor from years 2011 to 2016. The table contains annual- number of data count, mean, standard deviation (std), minimum (min), 0.25 percentile, 0.50 percentile, and 0.75 percentile values from the numeric data.

Figure 2 shows a heat-map on the left and 3-d plot on right for the mean value of persistence Δt. It can be observed that the Δt of the Ramp behavior is higher in December–January during every year, similar to the capacity factors. The recorded maximum value is 12.265 in 2014. Conversely, the value is generally lower in June and August. From the 3-d plot, Δt, the persistence of a ramp event increases in proportion to the increase of the wind power capacity factor from November to December. Note that the value of Δt has generally decreased from 2011 to 2016.

The mean value of the peak events (W_s) are extracted and plotted in Figure 3. On the left hand, there is a heat map and on the right there is a 3-d plot. The trends indicate that the maximum value in each year is recorded during December. It can also be noted that the value has risen notably from September to October except in 2012 and in 2015. In the 3-d plot, W_s has risen in January–February 2015–2016. The value of W_s has been decreasing in October–November period since 2013.

Figure 4 shows the heat map and 3-d plot of the amplitude of ramp events (ΔW_s). There was not a significant difference in each year and month. The values for December are lower for all years. From this point, ΔW_s becomes smaller in December when wind power production rises.

Figure 5 shows the heat map and 3-d plot of angle at which an event occurs ($\theta^{\Delta W_s}$). The significant angle changes mostly occur in winter seasons over the years.

TABLE 2 Statistical properties of wind power capacity factor and temperature data

Year	2011		2012		2013		2014		2015	
Parameter	WF	TEMP	WF	TEMP	WF	TEMP	WF	TEMP	WF	TEMP
count	8760	8760	8760	8760	8760	8760	8760	8760	8760	8760
mean	0.26	6.27	0.248	4.57	0.228	6.31	0.266	6.87	0.232	6.09
std	0.21	10.19	0.183	10.47	0.196	9.48	0.216	7.64	0.192	9.3
min	0	−29.69	0	−30.11	0.001	−19.1	0.001	−16.83	0	−18.08
0.25	0.09	−0.22	0.101	−1.8	0.072	−0.38	0.098	0.93	0.079	−0.47
0.5	0.2	6.06	0.207	5.43	0.165	5.5	0.202	5.77	0.179	4.36
0.75	0.39	14.51	0.358	13.09	0.34	13.72	0.384	13.2	0.335	14.42
max	0.97	28.2	0.892	27.97	0.971	29.47	0.962	27.27	0.908	26.73

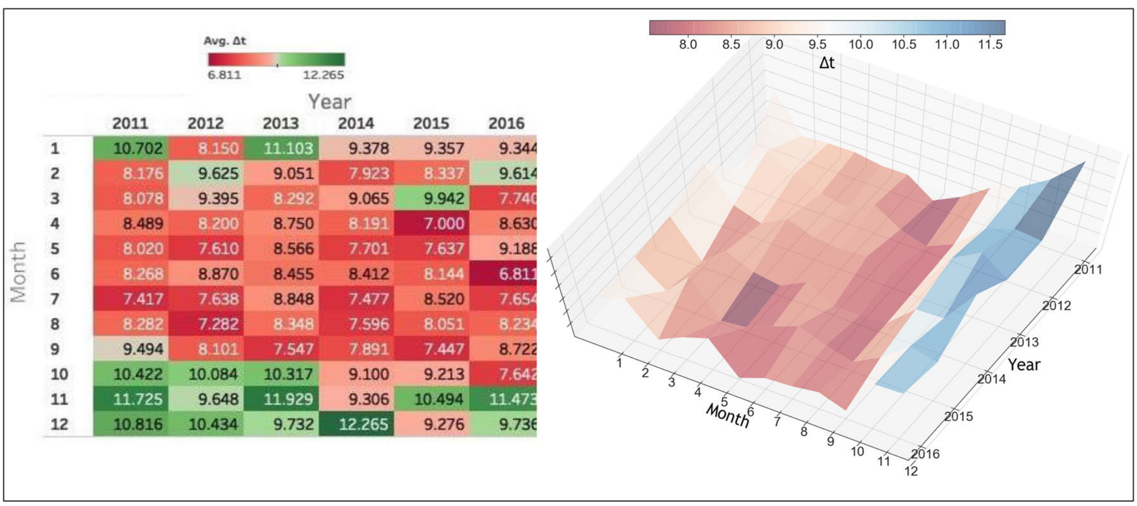


FIGURE 2 Heat map and surface plot of Δt

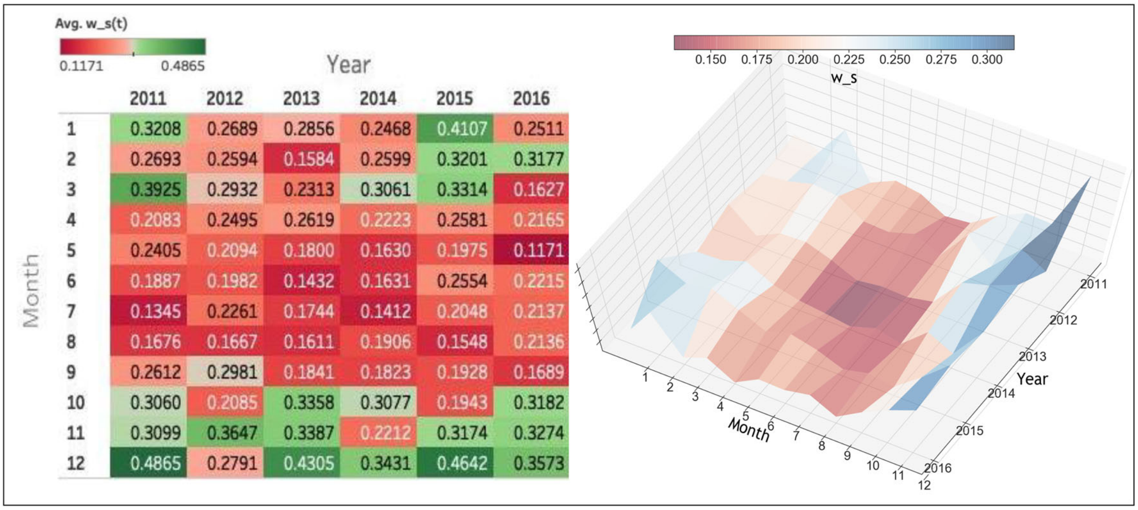


FIGURE 3 Heat map and surface plot of W_s

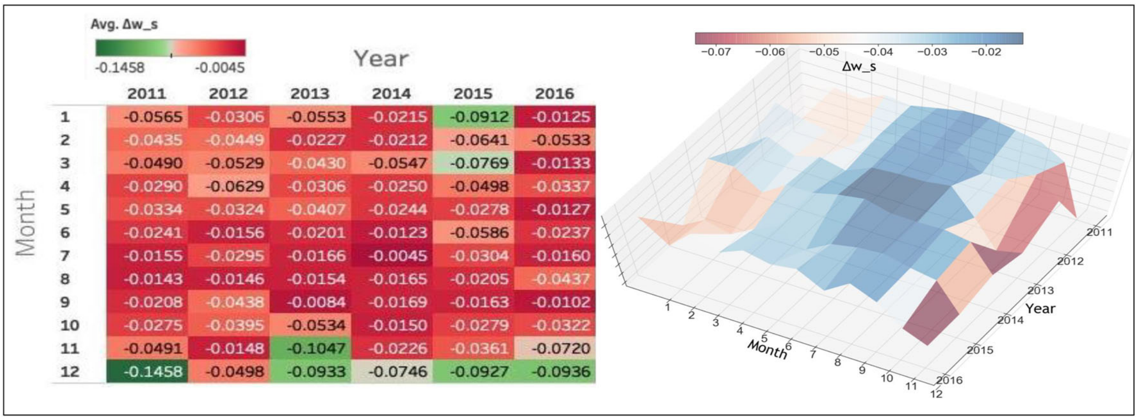


FIGURE 4 Heat map and surface plot of ΔW_s

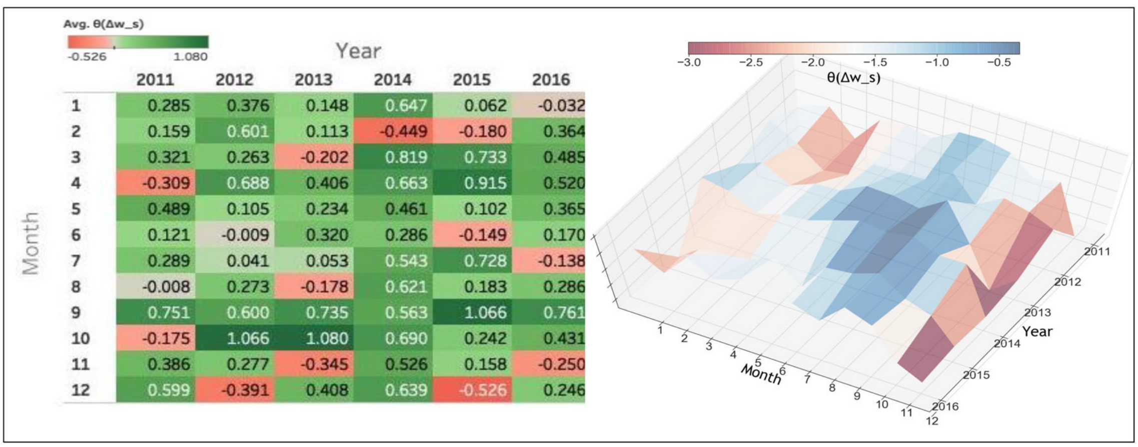


FIGURE 5 Heat map and surface plot of $\theta^{\Delta W_s}$

In this section the statistical properties of the extracted features of the significant wind ramp events are discussed. The features' persistence (Δt) andpeak event (W_s) are decreasing over the year in the data set for wind capacity factors, while, amplitude (ΔW_s) and angle of an event ($\theta^{\Delta W_s}$) has trivial changes over the six years time period. However there is a seasonal effect in the angle of an event wherein the changes intensify during the winter. The original time series data and the extracted events are then used for prediction using five deep learning models as in the following section.

3 | COMPUTATIONAL EXPERIMENTS

3.1 | Experimental setup

This investigation builds upon and utilizes the comparative analysis drawn between five deep learning models: Attention mechanism, deep convolutional neural network (DCN), deep feed forward (DFF), long short-term memory (LSTM), and independently recurrent neural network (indRNN).[46] These deep learning techniques are implemented using Tensorflow 2.0 package from Google[61] in Python programming language (3.6.4) in windows environment. In the previous work,[46] time series data are predicted over four time horizons, ranging from a month to a year long period, with hourly resolution. In this paper, significant events are extracted from 6 years of wind power production data. The features describing a significant event include persistence (Δt), the magnitude of power swing (Δw_s), time at which an event took place (t), time at which the power swing took place ($w_s(t)$) and the angle of the swing ($\theta^{\Delta w_s}$). The input data are processed in three formats: original, fast Fourier transformation (FFT), and discrete wavelet transformation (WT). Through the five deep learning techniques mentioned above, the features are predicted for a time period of 1 year. The model parameters are kept the same as in the previous work on time-series prediction.

3.2 | Results—Predictions of event features

Figures 6 and 7 depict the accuracy of the prediction. The models with accuracies better than the average gap of the attention model are highlighted in red. Among the five models, the Attention mechanism performs the best in all round tests.

Regarding prediction of Δt, the most accurate result is registered using the original data with Attention model. Pre-processing the data through time to signal domain using FFT has deteriorated the prediction through DFF, indRNN, and LSTM models. Meanwhile, the wavelet transformation for input data has declined the prediction accuracy using Attention mechanism and DCN.

In terms of computational times of execution for event prediction, the DFF with original event data is the fastest, but not the most accurate, while wavelet pre-processing with LSTM has the longest execution time. The same tendency as Δt was observed also for Δw_s, with the reason being that each feature, together with other features, describe a single event. In terms of model accuracy, the Attention model performs the best with original data. For prediction of t, the model accuracy using FFT transformation is the lowest. The Attention model so far has been most accurate for time series data and event series, however for prediction of t DCN it also performs better. For $w_s(t)$ prediction accuracy, the same tendency as $\Delta t, \Delta w_s$ was observed. The difference is that in case of the indRNN, the result of the original data was erroneous, and the result of the WT was accurate. In case of $\theta^{\Delta w_s}$, the DCN had larger gaps in all the input data compared with the Attention model; however, wavelet filters in other models have improved the accuracy by about 5–15% with the Attention models. In brief, the attention mechanism with original data performs the best for all five features, while for individual features the specific model and particular pre-processing performs better than the Attention mechanism.

For prediction of Δt, the Attention model is most accurate in terms of performance.

Comparing the short and long time horizons, short-term prediction was more accurate with the original data, while for other data types, the long time horizon prediction was more accurate. IndRNN is second to the attention model in terms of model accuracy. In this model, prediction period did not make a difference in the results for any data type. At Δw_s, the DCN results were improved. The values were almost the same as with the original data using the Attention model, and the short-period prediction was relatively better. Looking at indRNN and LSTM, with original data, long period prediction had a smaller gap. For t, it is clear that the preprocessing using the FFT filter did not improve the short and long period predictions. As for the tendency, similar to $\Delta t, \Delta w_s, t$ described above, the results of short-term period prediction were most accurate in the case of Attention model. The results for long period prediction were better for DFF, indRNN, LSTM. For $w_s(t)$ and for $\theta_{\Delta w_s}$, the same tendency as Δt and Δw_s was observed. The difference is that in case of indRNN, the result of the original data was erroneous, and the result for the WT is better.

3.3 | Results—Predictions of time series data

In Figure 8, the prediction accuracy of deep learning models with time-series input data is presented. The root mean square error (RMSE) is utilized as the metric to compare the models. The Attention mechanism performs best in comparison with the other four models for long-term prediction; therefore, the prediction errors in the Attention mechanism was selected as the the base case highlighted in red. As the errors increase, the red color is retained and if the error decreases, the color is changed to black. For instance, the error in predicting original WF using DFF is

Avg Gap

0.0071 0.6806

Data	Type	Attention	DCN	DFF	indRNN	LSTM
Δt	Original	0.0071	0.0228	0.0272	0.0182	0.0702
	WT	0.1043	0.1119	0.1002	0.1004	0.0877
	FFT	0.0527	0.0534	0.1473	0.3671	0.1345
Δw_s	Original	0.0145	0.0156	0.0210	0.0418	0.0643
	WT	0.1066	0.1145	0.1063	0.1056	0.0957
	FFT	0.0670	0.0542	0.1866	0.1160	0.1818
t	Original	0.0215	0.0114	0.0535	0.0267	0.0205
	WT	0.0367	0.0217	0.0416	0.0128	0.0160
	FFT	0.6790	0.6806	0.6531	0.3644	0.6624
w_s(t)	Original	0.0191	0.0215	0.0289	0.2476	0.0635
	WT	0.1583	0.1689	0.1696	0.1667	0.1394
	FFT	0.0457	0.0368	0.1896	0.1711	0.2321
θ(Δw_s)	Original	0.0131	0.0219	0.0326	0.0350	0.0821
	WT	0.1303	0.1365	0.1228	0.1251	0.1135
	FFT	0.0498	0.0555	0.1714	0.1312	0.1408

		Attention	DCN	DFF	indRNN	LSTM
Δt	Original	1.00	3.20	3.81	2.56	9.86
	WT	1.00	1.07	0.96	0.96	0.84
	FFT	1.00	1.01	2.79	6.97	2.55
Δw_s	Original	1.00	1.08	1.44	2.88	4.44
	WT	1.00	1.07	1.00	0.99	0.90
	FFT	1.00	0.81	2.78	1.73	2.71
t	Original	1.00	0.53	2.48	1.24	0.95
	WT	1.00	0.59	1.13	0.35	0.44
	FFT	1.00	1.00	0.96	0.54	0.98
w_s(t)	Original	1.00	1.13	1.51	12.96	3.33
	WT	1.00	1.07	1.07	1.05	0.88
	FFT	1.00	0.80	4.15	3.74	5.08
θ(Δw_s)	Original	1.00	1.67	2.49	2.67	6.28
	WT	1.00	1.05	0.94	0.96	0.87
	FFT	1.00	1.11	3.44	2.64	2.83

FIGURE 6 Comparison of model accuracy for prediction of event features

0.21 in black in comparison with the Attention mechanism. The reader is invited to refer to the work in Mishra et al[46] for a more comprehensive discussion on the results and for the algorithms' performance for time series predictions.

3.4 | Discussion—Time series predictions versus event predictions

A comparison of computational times between prediction of time series and prediction of series of events is performed in Figure 9. It is evident that prediction of events is significantly faster than predicting the original time-series data. The intuitive reasoning is that the absolute number of data points are reduced in several folds when events are extracted from data. For the proposed hypothesis to hold, prediction of events should also fare well in model accuracy. Thereby it can be established that prediction of events is accurate and computationally faster than predicting time-series data with and without preprocessing. This is because the the number of data points is reduced when events are extracted from the original time series.

A comparison of energy consumption between time series predictions and predicting a series of events is performed in Figure 10. A typical laptop consumes 100 W/h while in use. Taking it as a base case, the total energy consumption can be calculated through the following relationship: energy consumed (in Watt) $=$ (100 W / 3600 s) $*$ computing time (in seconds). The results of the comparison between the two types of

Data	Type	Attention				DCN				DFF				indRNN				LSTM			
Δt	Original	0.0069	0.0073	0.0069	0.0074	0.0170	0.0195	0.0235	0.0311	0.0299	0.0278	0.0247	0.0263	0.0197	0.0183	0.0169	0.0180	0.0768	0.0715	0.0648	0.0678
	WT	0.1134	0.1079	0.0968	0.0990	0.1203	0.1165	0.1041	0.1065	0.1065	0.1032	0.0932	0.0979	0.1058	0.1034	0.0931	0.0991	0.0943	0.0907	0.0810	0.0846
	FFT	0.0230	0.0323	0.0542	0.1013	0.0248	0.0338	0.0536	0.1013	0.1340	0.1427	0.1414	0.1709	0.3456	0.3597	0.3725	0.3907	0.1085	0.1088	0.1175	0.2030
Δw_s	Original	0.0140	0.0140	0.0139	0.0161	0.0142	0.0143	0.0156	0.0183	0.0223	0.0185	0.0184	0.0246	0.0618	0.0386	0.0320	0.0347	0.0747	0.0618	0.0585	0.0623
	WT	0.1050	0.1079	0.1053	0.1083	0.1111	0.1145	0.1139	0.1185	0.1060	0.1075	0.1037	0.1080	0.1014	0.1079	0.1045	0.1086	0.0941	0.0967	0.0930	0.0989
	FFT	0.0651	0.0655	0.0693	0.0682	0.0554	0.0526	0.0545	0.0542	0.1780	0.1911	0.1876	0.1898	0.1089	0.1187	0.1157	0.1205	0.1916	0.1854	0.1771	0.1731
t	Original	0.0198	0.0203	0.0216	0.0244	0.0108	0.0105	0.0105	0.0136	0.0481	0.0516	0.0551	0.0591	0.0298	0.0283	0.0237	0.0249	0.0331	0.0186	0.0156	0.0148
	WT	0.0345	0.0348	0.0366	0.0407	0.0209	0.0200	0.0200	0.0258	0.0375	0.0393	0.0426	0.0469	0.0173	0.0111	0.0103	0.0125	0.0275	0.0144	0.0113	0.0107
	FFT	0.6708	0.6818	0.6871	0.6763	0.6737	0.6829	0.6889	0.6769	0.6475	0.6564	0.6593	0.6493	0.3446	0.3555	0.3692	0.3881	0.6803	0.6899	0.6763	0.6029
w_s(t)	Original	0.0177	0.0190	0.0195	0.0202	0.0148	0.0185	0.0231	0.0296	0.0319	0.0267	0.0267	0.0303	0.2805	0.2456	0.2186	0.2456	0.0709	0.0609	0.0587	0.0636
	WT	0.1592	0.1597	0.1525	0.1618	0.1731	0.1685	0.1611	0.1727	0.1778	0.1619	0.1596	0.1789	0.1701	0.1576	0.1599	0.1791	0.1486	0.1363	0.1310	0.1417
	FFT	0.0459	0.0458	0.0444	0.0467	0.0365	0.0370	0.0365	0.0371	0.2130	0.1818	0.1733	0.1903	0.1883	0.1644	0.1586	0.1729	0.2452	0.2289	0.2226	0.2316
θ(Δw_s)	Original	0.0110	0.0119	0.0131	0.0163	0.0192	0.0217	0.0234	0.0232	0.0301	0.0319	0.0336	0.0348	0.0317	0.0332	0.0369	0.0380	0.0791	0.0795	0.0850	0.0849
	WT	0.1186	0.1295	0.1377	0.1355	0.1268	0.1324	0.1440	0.1429	0.1144	0.1235	0.1269	0.1262	0.1133	0.1254	0.1301	0.1314	0.1035	0.1137	0.1174	0.1195
	FFT	0.0506	0.0507	0.0476	0.0502	0.0546	0.0571	0.0551	0.0551	0.1625	0.1718	0.1737	0.1776	0.1242	0.1307	0.1337	0.1361	0.1445	0.1404	0.1400	0.1382

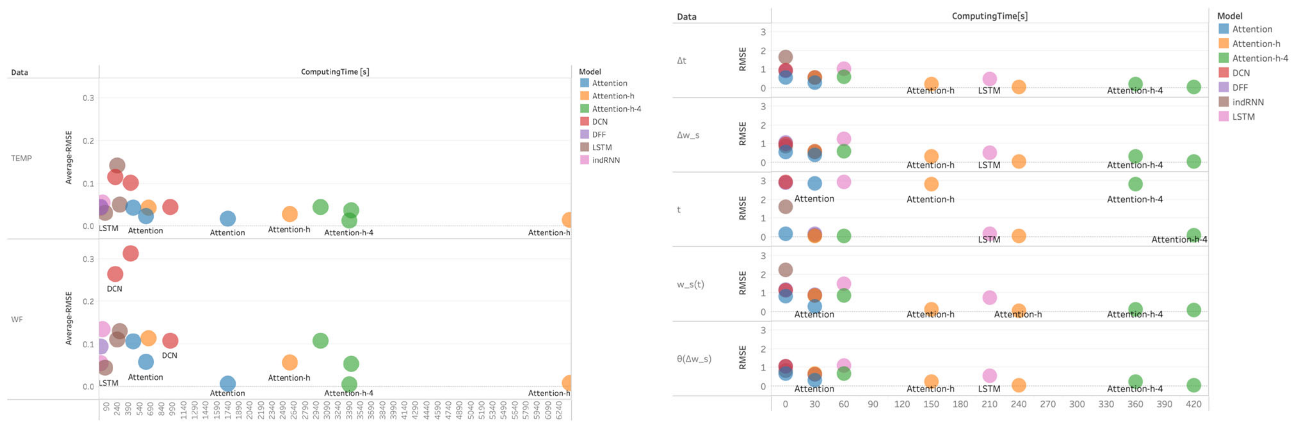
		Attention				DCN				DFF				indRNN				LSTM			
Data	**Type**	1	3	6	12	1	3	6	12	1	3	6	12	1	3	6	12	1	3	6	12
Δt	Original	1.00	1.06	1.00	1.07	1.00	1.15	1.38	1.83	1.00	0.93	0.83	0.88	1.00	0.93	0.86	0.91	1.00	0.93	0.84	0.88
	WT	1.00	0.95	0.85	0.87	1.00	0.97	0.87	0.89	1.00	0.97	0.88	0.92	1.00	0.98	0.88	0.94	1.00	0.96	0.86	0.90
	FFT	1.00	1.40	2.36	4.40	1.00	1.36	2.16	4.08	1.00	1.06	1.06	1.28	1.00	1.04	1.08	1.13	1.00	1.00	1.08	1.87
Δw_s	Original	1.00	1.00	0.99	1.15	1.00	1.01	1.10	1.29	1.00	0.83	0.83	1.10	1.00	0.62	0.52	0.56	1.00	0.83	0.78	0.83
	WT	1.00	1.03	1.00	1.03	1.00	1.03	1.03	1.07	1.00	1.01	0.98	1.02	1.00	1.06	1.03	1.07	1.00	1.03	0.99	1.05
	FFT	1.00	1.01	1.06	1.05	1.00	0.95	0.98	0.98	1.00	1.07	1.05	1.07	1.00	1.09	1.06	1.11	1.00	0.97	0.92	0.90
t	Original	1.00	1.03	1.09	1.23	1.00	0.97	0.97	1.26	1.00	1.07	1.15	1.23	1.00	0.95	0.80	0.84	1.00	0.56	0.47	0.45
	WT	1.00	1.01	1.06	1.18	1.00	0.96	0.96	1.23	1.00	1.05	1.14	1.25	1.00	0.64	0.60	0.72	1.00	0.52	0.41	0.39
	FFT	1.00	1.02	1.02	1.01	1.00	1.01	1.02	1.00	1.00	1.01	1.02	1.00	1.00	1.03	1.07	1.13	1.00	1.01	0.99	0.89
w_s(t)	Original	1.00	1.07	1.10	1.14	1.00	1.25	1.56	2.00	1.00	0.84	0.84	0.95	1.00	0.88	0.78	0.88	1.00	0.86	0.83	0.90
	WT	1.00	1.00	0.96	1.02	1.00	0.97	0.93	1.00	1.00	0.91	0.90	1.01	1.00	0.93	0.94	1.05	1.00	0.92	0.88	0.95
	FFT	1.00	1.00	0.97	1.02	1.00	1.01	1.00	1.02	1.00	0.85	0.81	0.89	1.00	0.87	0.84	0.92	1.00	0.93	0.91	0.94
θ(Δw_s)	Original	1.00	1.08	1.19	1.48	1.00	1.13	1.22	1.21	1.00	1.06	1.12	1.16	1.00	1.05	1.16	1.20	1.00	1.01	1.07	1.07
	WT	1.00	1.09	1.16	1.14	1.00	1.04	1.14	1.13	1.00	1.08	1.11	1.10	1.00	1.11	1.15	1.16	1.00	1.10	1.13	1.15
	FFT	1.00	1.00	0.94	0.99	1.00	1.05	1.01	1.01	1.00	1.06	1.07	1.09	1.00	1.05	1.08	1.10	1.00	0.97	0.97	0.96

FIGURE 7 Comparison of model predictions over the prediction period

			Data	
Type	Model	ComputingTime[sec]	TEMP	WF
Original	DFF	4.6	0.056	0.053
	indRNN	7.5	0.066	0.037
	LSTM	62.7	0.084	0.127
	DCN	221.5	0.357	0.844
	Attention	469	0.126	0.332
	Attention-h	670	0.122	0.352
	Attention-h-4	3000	0.127	0.335
WT	DFF	29.1	0.195	0.490
	indRNN	39.4	0.164	0.423
	LSTM	289.1	0.149	0.409
	DCN	427.4	0.311	1.000
	Attention	657.3	0.063	0.172
	Attention-h	2596	0.075	0.166
	Attention-h-4	3435	0.106	0.161
FFT	DFF	8.1	0.132	0.335
	indRNN	28.4	0.186	0.289
	LSTM	267.2	0.446	0.346
	DCN	964.9	0.126	0.336
	Attention	1764.4	0.039	0.006
	Attention-h-4	3390	0.026	0.000
	Attention-h	6399	0.030	0.011

Normalized-RMSE
0.000 1.000

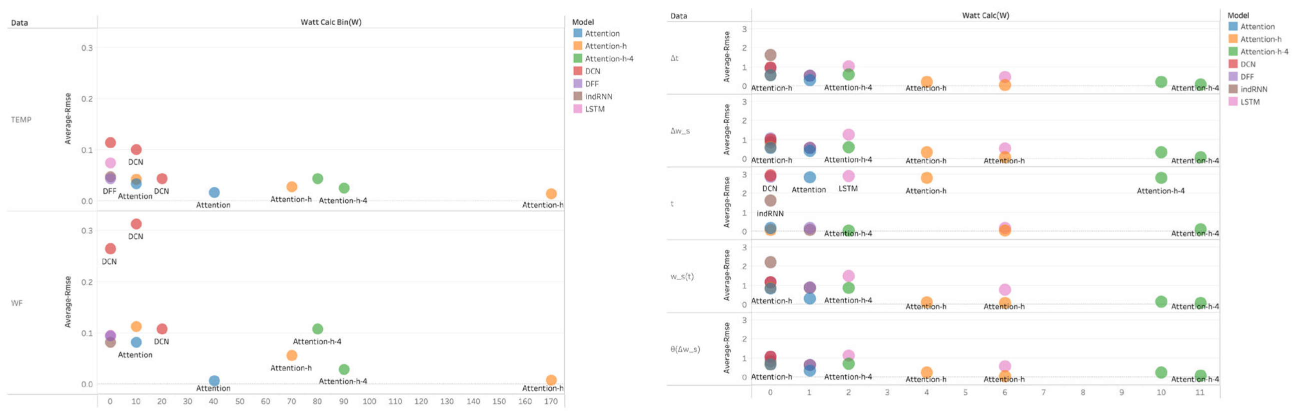
FIGURE 8 Time series prediction performance



(A) Computational time for time-series predictions

(B) Computational time for events prediction

FIGURE 9 Comparison of computational times between time series and a series of events



(A) Energy consumption for time-series predictions

(B) Energy consumption for events predictions

FIGURE 10 Comparison of energy consumption between time series and a series of events

predictions using the deep learning models are presented in Figure 10. It is evident that prediction of events is significantly more efficient in terms of energy consumption than that of predicting the original time-series data.

A comparison is drawn between prediction performance using series of events in Figures 6 and 7 and using time series in Figure 8. Taking a closer look at the model performances for the event features and four different time horizons (1, 3, 6, and 12 months), the model error in the LSTM, for the most part, is better than the base case (the Attention model). The indRNN stands in close second place in performance, where it fares well across all features except $\theta(\Delta W_s)$. This is followed by DFF closely and by DCN loosely in terms of performance. It was also observed that, for most part, the original data reduced the model prediction errors. Summing up the results of the four models, LSTM, indRNN, DFF, and DCN, have performed better than the base case of the Attention model in predicting the series of events. While the same pattern can be observed in model performances with time-series input, the degree of accuracy is higher in cases where a series of events are used as the input. Evidently, since the the total number of events extracted from the time-series is significantly less than the number of data points in the time-series, the input data size and the computational resources are significantly reduced. In summary, predicting a series of events is clearly computationally efficient, reduces the data volume and partially improves the model performance over that of time-series as an input. This means event based prediction might be better suited for real-time real-world machine learning applications for prediction. Having said that, more investigation is required to investigate how the performance variation in predicting event features can be improved. Expanding on the previous research, if LSTM performs well within the error threshold in predicting four out of five features, then perhaps the fifth feature can be adjusted. Finally, a pertinent question to be investigated in the future was introduced in Mishra et al,[39] and this is to research reconstruction of the original or near original time-series from the extracted events.

4 | CONCLUSIONS AND IMPLICATIONS

This paper explores a comparative investigation on multi-variate forecasting using five deep learning models for four time horizons. Predicting large time-series data consumes a lot of memory and is computationally expensive, thereby resulting in slower operational time. This paper proposes an alternative to address these bottlenecks, specifically predicting a series of events where each event is described by four features. The hypothesis is that if the trends and patterns are the information expected from a time-series, then the events carrying the relevant information could be extracted from the time series and a series of such events can encapsulate the actionable information in the time series. If predicting a series of events yields equal or improved prediction accuracy then the hypothesis is proved. A selection of five deep learning models, namely, Attention mechanism, deep convolutional neural network (DCN), deep feed foward (DFF), long short-term memory (LSTM), and independently recurrent neural network (indRNN) are used for the prediction. Two data pre-processing techniques, fast Fourier transformation (FFT) and discrete wavelet, are used. For comparison purposes, all the hyper parameters of the selected models are kept the same for all the tests. For testing and validation, 6 years of historical data of wind power production and temperature localized to Estonia are selected.

From the tests conducted, it is evident that event-based prediction reduces the total data volume by nearly 50% with the specific dataset and parameters selection, requires less computational resources, and performs at a similar or higher degree of efficiency than that of the prediction using time-series data as input. The investigation indicates that Attention models performs well in all the tests with DNN as a close second. Pre-processing while increases the computational time not necessarily improve the efficiency of predictions. In addition to that, the energy consumption in computation has decreased in one to two-folds when using event based prediction.

The proposed method is especially useful for the large volume and velocity of time-series data. In domains such as predictive analytics, it could be specially advantageous. Event-based prediction could accelerate real-time applications in various domains. The reduction in memory size and computational resources could result in cost saving. The domains of green computing and embedded systems require energy-efficient predictive analytics approaches. Embedded systems are widespread as they power the modern technology economy. There are over 28 billion microcontrollers in use each year. Embedded systems typically operate in resource-constrained environments such that they need to run on battery; hence, energy efficiency, limited memory, or extremely slow clock rates are key challenges. The proposed RBA method that focuses on predicting events extracted from time-series data could be relevant to program embedded systems. As sensor data arrives at a high frequency-potentially many thousands of times per second, an embedded device running event-based predictions can operate in a limited duty cycle comprising of the stages of data collection, processing, and providing inference on the results.

The event identification algorithm used in the work is generalized because the algorithm identifies variations in any data set; however, threshold and cut-off frequency parameters are required to execute the algorithm. These parameters can regulate the number and type of events extracted from a single data set. The threshold sets a limit, while cut-off frequency sets a range within which the variations are captured. This is location specific, therefore different thresholds and cut-off frequencies must be set when extracting events from different locations. Further research on the effect of threshold, cutt-off frequency to find an optimal values is a good direction for future investigation. Other directions for investigations include determining whether the features should be predicted separately and reconstructed at a later stage as a single event. Then, research will need to be done on how to reduce the losses in reconstructing a time-series from the extracted events. Also more investigation could be performed for addressing uncertainty that comes from assigning probability to features rather than to a single time-series data point. For instance, it might be known when t occurs, but not the magnitude ΔW_s of change or the rate of change $\theta(\Delta W_s)$. In this case a hyper parameter tuning can also be conducted to improve the model results. Finally, since the accuracy of the models in this work is compared through the RMSE error metrics, a more in-depth analysis can be performed focused on the error metrics.

ACKNOWLEDGEMENT

This work is supported by the Estonian Research Council grant PUTJD915.

DATA AVAILABILITY STATEMENT

The research data used are openly available and accessible.

ORCID

Sambeet Mishra https://orcid.org/0000-0002-0529-0620

REFERENCES

1. Piotr Bojek HB. Wind Power – Analysis - IEA. https://www.iea.org/reports/wind-power, [Online; accessed 14. Feb. 2022]; 2022.
2. Soares-Ramos EPP, de Oliveira-Assis L, Sarrias-Mena R, Fernández-Ramírez LM. Current status and future trends of offshore wind power in Europe. *Energy*. 2020;202:117787.
3. Haghi HV, Lotfifard S, Qu Z. Multivariate predictive analytics of wind power data for robust control of energy storage. *IEEE Trans Ind Inf*. 2016;12(4): 1350-1360.

4. Zhou P, Yin PT. An opportunistic condition-based maintenance strategy for offshore wind farm based on predictive analytics. *Renew Sustain Energy Rev*. 2019;109:1-9.
5. Mujeeb S, Alghamdi TA, Ullah S, Fatima A, Javaid N, Saba T. Exploiting deep learning for wind power forecasting based on big data analytics. *Appl Sci*. 2019;9(20):4417.
6. Jamil F, Iqbal N, Ahmad S, et al. Peer-to-peer energy trading mechanism based on blockchain and machine learning for sustainable electrical power supply in smart grid. *IEEE Access*. 2021;9:39193-39217.
7. Lerner J, Grundmeyer M, Garvert M. The importance of wind forecasting. *Renew Energy Focus*. 2009;10(2):64-66.
8. Meenal R, Binu D, Ramya KC, et al. Weather forecasting for renewable energy system: a review. *Arch Comput Methods Eng*. 2022;2022:1-17.
9. Liu H, Chen C, Lv X, Wu X, Liu M. Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods. *Energy Convers Manag*. 2019;195:328-345.
10. Liu H, Li Y, Duan Z, Chen C. A review on multi-objective optimization framework in wind energy forecasting techniques and applications. *Energy Convers Manag*. 2020;224:113324.
11. Zhang Y, Wang J, Wang X. Review on probabilistic forecasting of wind power generation. *Renew Sustain Energy Rev*. 2014;32:255-270.
12. Bazionis IK, Georgilakis PS. Review of deterministic and probabilistic wind power forecasting: Models, methods, and future research. *Electricity*. 2021;2(1):13-47.
13. Wang J, Zhou Q, Zhang X. Wind power forecasting based on time series ARMA model. In: IOP conference series: Earth and environmental science, Vol. 199 IOP Publishing; 2018:022015.
14. Yatiyana E, Rajakaruna S, Ghosh A. Wind speed and direction forecasting for wind power generation using ARIMA model. In: 2017 Australasian Universities power engineering conference (AUPEC) IEEE; 2017:1-6.
15. Meng X, Zhao W, Wang H, Gao Y, Zhong H, Li C. Application of grey theory and neural network in medium term wind power forecasting. In: 2017 Chinese automation congress (CAC) IEEE; 2017:5328-5331.
16. Kosovic B, Haupt SE, Adriaansen D, et al. A comprehensive wind power forecasting system integrating artificial intelligence and numerical weather prediction. *Energies*. 2020;13(6):1372.
17. Nazaré G, Castro R, Gabriel Filho LRA. Wind power forecast using neural networks: tuning with optimization techniques and error analysis. *Wind Energy*. 2020;23(3):810-824.
18. Pawar A, Jape VS, Mathew S. Wind power forecasting using support vector machine model in rstudio. *Cognitive informatics and soft computing*: Springer; 2019:289-298.
19. Wang H, Lei Z, Zhang X, Zhou B, Peng J. A review of deep learning for renewable energy forecasting. *Energy Convers Manag*. 2019;198:111799.
20. Yang X, Ma X, Kang N, Maihemuti M. Probability interval prediction of wind power based on KDE method with rough sets and weighted Markov chain. *IEEE Access*. 2018;6:51556-51565.
21. Hu J, Tang J, Lin Y. A novel wind power probabilistic forecasting approach based on joint quantile regression and multi-objective optimization. *Renew Energy*. 2020;149:141-164.
22. Wu Y-K, Su P-E, Wu T-Y, Hong J-S, Hassan MY. Probabilistic wind-power forecasting using weather ensemble models. *IEEE Trans Ind Appl*. 2018;54(6):5609-5620.
23. Xiyun Y, Xue M, Guo F, Huang Z, Jianhua Z. Wind power probability interval prediction based on bootstrap quantile regression method. In: 2017 chinese automation congress (CAC) IEEE; 2017:1504-1509.
24. Kim D, Hur J. Short-term probabilistic forecasting of wind energy resources using the enhanced ensemble method. *Energy*. 2018;157:211-226.
25. Li C, Lin S, Xu F, Liu D, Liu J. Short-term wind power prediction based on data mining technology and improved support vector machine method: a case study in northwest China. *J Cleaner Product*. 2018;205:909-922.
26. Safari N, Chung CY, Price GCD. Novel multi-step short-term wind power prediction framework based on chaotic time series analysis and singular spectrum analysis. *IEEE Trans Power Syst*. 2017;33(1):590-601.
27. Bokde N, Feijóo A, Villanueva D, Kulat K. A review on hybrid empirical mode decomposition models for wind speed and wind power prediction. *Energies*. 2019;12(2):254.
28. Sweeney C, Bessa RJ, Browell J, Pinson P. The future of forecasting for renewable energy. *Wiley Interdiscip Rev: Energy Environ*. 2020;9(2):e365.
29. Potdar V, Chandan A, Batool S, Patel N. Big energy data management for smart grids issues, challenges and recent developments. *Smart cities*: Springer; 2018:177-205.
30. Naik J, Dash S, Dash PK, Bisoi R. Short term wind power forecasting using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network. *Renew Energy*. 2018;118:180-212.
31. Saroha S, Aggarwal SK. Wind power forecasting using wavelet transforms and neural networks with tapped delay. *CSEE J Power Energy Syst*. 2018;4(2):197-209.
32. Grolinger K, Capretz MAM, Seewald L. Energy consumption prediction with big data: balancing prediction accuracy and computational resources. In: 2016 ieee international congress on big data (bigdata congress) IEEE; 2016:157-164.
33. Parhizkar T, Rafieipour E, Parhizkar A. Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction. *J Cleaner Product*. 2021;279:123866.
34. Kenner S, Thaler R, Kucera M, Volbert K, Waas T. Comparison of smart grid architectures for monitoring and analyzing power grid data via modbus and rest. *EURASIP J Embedded Syst*. 2017;2017(1):1-13.
35. Yan Y, Qian Y, Sharif H, Tipper D. A survey on smart grid communication infrastructures: motivations, requirements and challenges. *IEEE Commun Surv Tutor*. 2012;15(1):5-20.
36. Gao J, Xiao Y, Liu J, Liang W, Chen CLP. A survey of communication/networking in smart grids. *Future Gener Comput Syst*. 2012;28(2):391-404.
37. Süsser D, Ceglarz A, Gaschnig H, et al. Model-based policymaking or policy-based modelling? How energy models and energy policy interact. *Energy Res Soc Sci*. 2021;75:101984.
38. Pinson P. Wind energy: forecasting challenges for its operational management. *Stat Sci*. 2013;28(4):564-585.
39. Mishra S, Ören E, Bordin C, Wen F, Palu I. Features extraction of wind ramp events from a virtual wind park. *Energy Rep*. 2020;6:237-249.
40. Bordin C, Håkansson A, Mishra S. Smart energy and power systems modelling: an IoT and cyber-physical systems perspective, in the context of energy informatics. *Procedia Comput Sci*. 2020;176:2254-2263.

41. Bordin C. Mathematical optimization applied to thermal and electrical energy systems. *Ph.D. Thesis*: Alma Mater Studiorum; 2015.
42. Ibrahim AA, Kazemtabrizi B, Bordin C, Dent CJ, McTigue JD, White AJ. Pumped thermal electricity storage for active distribution network applications. In: 2017 ieee manchester powertech IEEE; 2017:1-6.
43. Bordin C, Mishra S, Palu I. A multihorizon approach for the reliability oriented network restructuring problem, considering learning effects, construction time, and cables maintenance costs. *Renew Energy*. 2021;168:878-895.
44. Bordin C, Tomasgard A. Smacs model, a stochastic multihorizon approach for charging sites management, operations, design, and expansion under limited capacity conditions. *J Energy Storage*. 2019;26:100824.
45. Bordin C, Skjelbred HI, Kong J, Yang Z. Machine learning for hydropower scheduling: state of the art and future research directions. *Procedia Comput Sci*. 2020;176:1659-1668.
46. Mishra S, Bordin C, Taharaguchi K, Palu I. Comparison of deep learning models for multivariate prediction of time series wind power generation and temperature. *Energy Rep*. 2020;6:273-286.
47. Muniswamaiah M, Agerwala T, Tappert CC. Green computing for internet of things. In: 2020 7th ieee international conference on cyber security and cloud computing (cscloud)/2020 6th IEEE international conference on edge computing and scalable cloud (edgecom) IEEE; 2020:182-185.
48. Schwartz R, Dodge J, Smith NA, Etzioni O. Green ai. arXiv preprint arXiv:190710597; 2019.
49. Telgarsky M. Representation benefits of deep feedforward networks. arXiv preprint arXiv:150908101; 2015.
50. Luo XJ, Oyedele LO, Ajayi AO, et al. Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings. *Energy and AI*. 2020;2:100015.
51. Aloysius N, Geetha M. A review on deep convolutional neural networks. In: 2017 international conference on communication and signal processing (ICCSP) IEEE; 2017:0588-0592.
52. Houidi S, Fourer D, Auger F. On the use of concentrated time–frequency representations as input to a deep convolutional neural network: application to non intrusive load monitoring. *Entropy*. 2020;22(9):911.
53. Medsker LR, Jain LC. Recurrent neural networks. Design and Applications, 5; 2001.
54. Kisvari A, Lin Z, Liu X. Wind power forecasting—a data-driven method along with gated recurrent neural network. *Renew Energy*. 2021;163:1895-1909.
55. Heidari A, Khovalyg D. Short-term energy use prediction of solar-assisted water heating system: Application case of combined attention-based LSTM and time-series decomposition. *Solar Energy*. 2020;207:626-639.
56. Xiang L, Wang P, Yang X, Hu A, Su H. Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism. *Measurement*. 2021;175:109094.
57. Hong T, Pinson P, Wang Y, Weron R, Yang D, Zareipour H. Energy forecasting: a review and outlook. *IEEE Open Access J Power Energy*. 2020.
58. Tan M, Yuan S, Li S, Su Y, Li H, He F. Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning. *IEEE Trans Power Syst*. 2019;35(4):2937-2948.
59. Pfenninger S, Staffell I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*. 2016;114:1251-1265.
60. Staffell I, Pfenninger S. Using bias-corrected reanalysis to simulate current and future wind power output. *Energy*. 2016;114:1224-1239.
61. Abadi M, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:160304467; 2016.

How to cite this article: Mishra S, Bordin C, Taharaguchi K, Purkayastha A. Predictive analytics beyond time series: Predicting series of events extracted from time series data. *Wind Energy*. 2022;1-14. doi:10.1002/we.2760