RESEARCH ARTICLE

Incorporating Spatial and Temporal Correlations to Improve Aggregation of Decentralized Day-Ahead Wind Power Forecasts

NDAMULELO MARARAKANYE[iD]**, AMARIS DALTON**[iD]**, (Member, IEEE), AND BERNARD BEKKER, (Member, IEEE)**
Department of Electrical and Electronic Engineering, Stellenbosch University, Matieland, Stellenbosch 7602, South Africa

Corresponding author: Ndamulelo Mararakanye (ndamulelo@sun.ac.za)

ABSTRACT In some electricity markets, individual wind farms are obliged to provide point forecasts to the power purchaser or system operator. These decentralized forecasts are usually based on on-site meteorological forecasts and measurements, and thus optimized for local conditions. Simply adding decentralized forecasts may not capture some of the spatial and temporal correlations of wind power, thereby lowering the potential accuracy of the aggregated forecast. This paper proposes the explanatory variables that are used to train the kernel density estimator and conditional kernel density estimator models to derive day-ahead aggregated point and probabilistic wind power forecasts from decentralized point forecasts of geographically distributed wind farms. The proposed explanatory variables include (a) decentralized point forecasts clustered using the clustering large applications algorithm to reduce the high-dimensional matrices, (b) hour of day and month of year to account for diurnal and seasonal cycles, respectively, and (c) atmospheric states derived from self-organizing maps to represent large-scale synoptic circulation climatology for a study area. The proposed methodology is tested using the day-ahead point forecast data obtained from 29 wind farms in South Africa. The results from the proposed methodology show a significant improvement as compared to simply adding the decentralized point forecasts. The derived predictive densities are shown to be non-Gaussian and time-varying, as expected given the time-varying nature of wind uncertainty. The proposed methodology provides system operators with a method of not only producing more accurate aggregated forecasts from decentralized forecasts, but also improving operational decisions such as dynamic operating reserve allocation and stochastic unit commitment.

INDEX TERMS Aggregated wind power forecasting, diurnality, large-scale atmospheric circulations, probabilistic, seasonality.

I. INTRODUCTION

The electricity generation from wind energy is increasing worldwide, as different regions continue with the transition towards a decarbonized future. Wind power forecasting remains one of the most effective methods of reducing the impacts of wind power intermittency on power system operations. Traditionally, wind power forecasts consist of one value

(also called point or deterministic forecasts) for each wind farm [1], [2], [3]. These are usually based on on-site meteorological forecasts and measurements, and thus optimized for local conditions [4]. However, as wind capacity continues to increase in a region, producing accurate aggregated wind power forecasts becomes a concern for system operators. One of the key considerations in aggregated wind power forecasting is spatial-temporal correlations between wind farms that are geographically distributed. The power from geographically distributed wind farms exhibits spatial-temporal

The associate editor coordinating the review of this manuscript and approving it for publication was Ton Duc Do[iD].

correlations, the degree of which can vary based on numerous factors such as separation distance and direction, timescale, diurnal weather variations and movement of synoptic weather systems [1], [2], [5]. Recent literature has shown the potential benefit of incorporating spatial-temporal correlations in wind forecasting, especially in terms of improving forecasting accuracy [1], [2], [3], [6], [7], [8]. Another factor to consider is that uncertainty in wind power forecasting is unavoidable due to time-varying meteorological conditions, weather-to-power conversion process, and dynamic behavior of wind turbines [7]. Based on this realization, utilities are moving away from only focusing on increasing the accuracy of wind power forecasts towards also quantifying the uncertainty inherent in the forecast and incorporating this into their decisions [1], [2], [7], [9], [10], i.e. attaching a risk metric to the forecast. This is also referred to as probabilistic forecasting.

Probabilistic wind power forecasting approaches that incorporate spatial-temporal correlations can be divided into two main categories: physical and statistical [1], [7], [11], [12], [13]. Physical approaches typically use numerical weather prediction (NWP) models and current weather conditions to predict wind speed [7], [8], [14]. NWP models formulate the problem of wind speed prediction as a set of mathematical equations describing the atmosphere and oceans. In [7], a recursively backtracking framework based on the particle filter was used to estimate the atmospheric state (with near-surface measurements) and forecast samples of aggregated wind power. The samples were used to derive predictive densities using a kernel density estimator (KDE). In [3], weighted Euclidean distance was proposed to search for similar wind characteristics in historical NWP data, and used corresponding aggregated wind power measurements to construct probabilistic forecasts based on distance-weighted KDE. Weather ensemble predictions based on atmospheric models and time series were used together with Gaussian KDE [15], normalized prediction risk index [16], and generalized autoregressive conditional heteroscedasticity [17] to derive wind power forecasts and associated uncertainty. In [18], Gaussian processes combined with NWP were used to derive day-ahead wind power forecasts. In [19], the poor man's ensemble was used to estimate forecast errors for one wind farm while in [20] used input from 16 different European meteorological services for Previento to derive probabilistic forecasts for Germany. More studies based on physical models can be found in literature reviews conducted in [11], [12], [21], [22], and [23]. Statistical approaches, on the other hand, take historical wind power data and/or NWP as inputs and use machine learning algorithms and/or other statistical models to generate aggregated wind forecasts [7], [8], [14]. These approaches assume that historical data can be used to infer spatial-temporal correlations among wind farms. In [8], machine learning algorithms were used to generate point forecasts of wind farms, the copula method to build spatial-temporal correlated aggregated wind power forecasts, and Bayesian theory to derive predictive densities.

In [2], Bayesian hierarchical models were used for obtaining spatial-temporal correlated probabilistic wind power forecasts. In [1], the alternating direction method of multipliers was proposed to capture spatial-temporal correlations of geographically distributed wind farms and used multiple quantile regression to derive predictive densities. In [24] and [25], the resampling approach is used to estimate the confidence intervals for wind power forecasts. In [26], local quantile regression is compared with the local Gaussian model and the Nadaraya-Watson Estimator. In [27], historical forecast error distributions were used to obtain scenarios for stochastic wind power generation. In [10], time adaptive conditional KDE was proposed for probabilistic wind power forecasting. In [28], [29], and [30], the beta distribution was used to estimate forecast errors at different wind power forecast bins while in [31] the gamma-like distributions were used to achieve the same. In [32], the logit transformation approach was used to estimate the confidence intervals of forecast errors. More studies based on statistical models can be found in literature reviews conducted in [11], [12], [21], [22], and [23].

The first observation made from the literature is that the majority of proposed forecasting methodologies assume that the problem of aggregated wind power forecasting is solved in a centralized manner, i.e., forecasts of all wind farms in a region are derived centrally by one forecasting company (facilitated by the system operator). A centralized forecaster will often use a consistent forecasting approach for all wind farms, leading to more consistent results [33]. In addition, a forecaster will have access to measurements from all wind farms, making it easier to incorporate spatial-temporal correlations between wind farms into their forecasting methodologies [1]. However, in some markets (e.g. South Africa), individual wind farms are obliged to provide point forecasts to the power purchaser or system operator [34]. These decentralized point forecasts are optimized for local conditions and therefore simply adding these decentralized point forecasts may not capture some of the common spatial and temporal correlations of wind power, thereby lowering the potential accuracy of the aggregated wind power forecast. It is well known that the correlation of wind power from geographically distributed wind farms depends on the proximity of wind farms [2]. Wind farms in close proximity are highly correlated, whereas wind farms that are further apart are not. In addition, wind power profiles exhibit a high degree of statistical regularity along diurnal and seasonal timescales in the literature [35], [36], [37], [38], [39], [40], [41]. These spatial and temporal correlations form the basis of understanding wind power variability and ultimately improving wind power forecasting. Therefore, there is a need for a model that aggregates decentralized point forecasts while considering these correlations.

The second observation made from the literature is that most forecasting methodologies are based on the microscale and/or mesoscale NWP models. However, it was illustrated in [42], [43], and [44] that the probabilistic properties of wind
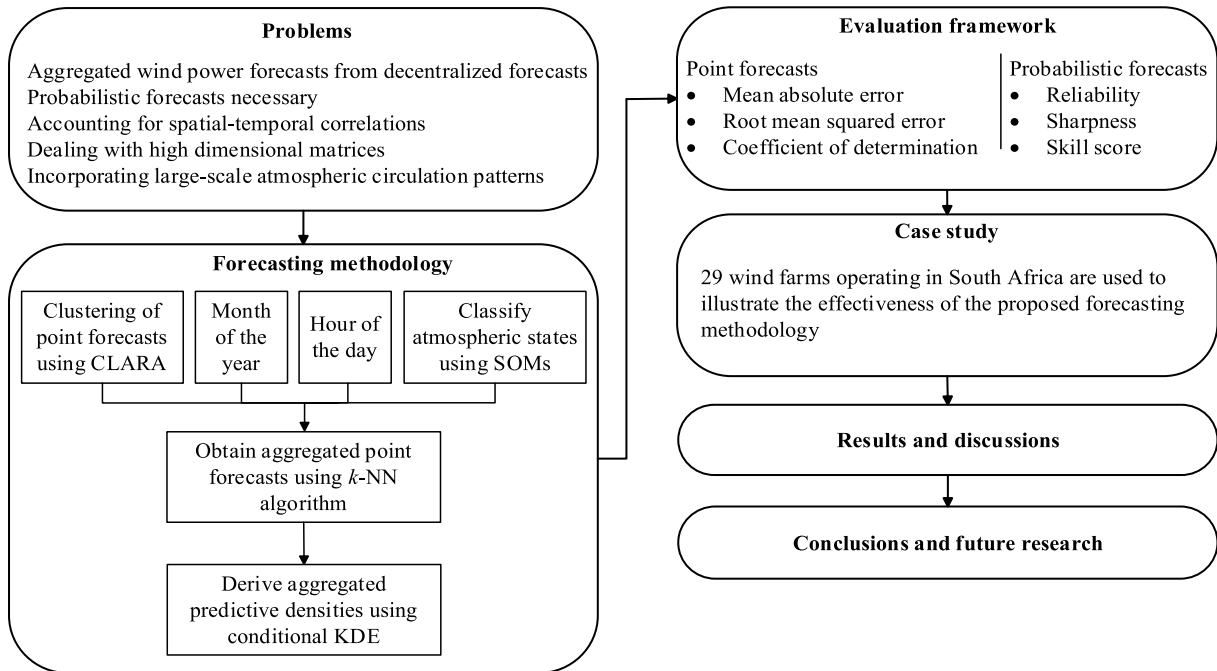
power generators' output, along with the level of correlation between wind generators' output, are dependent on the dominant large-scale atmospheric circulation archetypes. Thus, the information contained in large-scale atmospheric circulations can be useful in improving wind power forecasting. Large-scale atmospheric circulations have been incorporated in medium- to long-term wind forecasting [15], [45], [46], but have only been alluded to in short-term wind forecasting [7], [15], [47]. There is a need for more applied research on the potential benefits of incorporating large-scale atmospheric circulations in short-term wind power forecasting.

In light of the two aforementioned needs, the primary hypothesis of this paper is that: the accuracy of aggregated forecasts can be improved by training machine learning models with features that account for some of the common spatial and temporal correlations of wind power, including those correlations caused by large-scale atmospheric circulations.

To test this hypothesis, a methodology is proposed that trains machine learning models using the explanatory variables listed below to derive aggregated point and probabilistic wind power forecasts:

(a) Decentralized point forecasts – To eliminate duplicate features and reduce the high dimension matrices required to model a high number of wind farms in a region (without losing important spatial information), the correlated point forecasts are first clustered into k clusters using the clustering large applications algorithm (CLARA).

(b) The hour of day and month of year – These are included to model the well-known statistical regularity of wind profiles along diurnal and seasonal timescales.

(c) Atmospheric states – These are derived from self-organizing maps (SOMs) to represent large-scale synoptic circulation climatology for a study area.

The machine learning models proposed to test the hypothesis are: (a) the k-nearest neighbor (k-NN) to derive the aggregated wind power point forecast and (b) conditional KDE to derive aggregated wind power predictive densities. It should be noted here that there are numerous machine learning models that have been applied for wind power forecasting as shown in the literature review conducted above. However, the k-NN- and KDE-based approaches are proposed to test this hypothesis because they are common and easy to implement. In addition, the conditional KDE is a nonparametric approach for predicting wind power densities and thus can account for the time-varying and non-Gaussian nature of wind power uncertainty. The proposed methodology is demonstrated using the day-ahead point forecast data obtained from 29 wind farms in South Africa.

An overview of this paper is shown in Fig. 1, with the paper organized as follows: Section II provides a theoretical framework of the methodology for aggregated wind power forecasting. Section III defines the forecasting evaluation framework used in this paper. Section IV introduces the case study used for illustrating the proposed methodology and presents the results and discussions. Section V gives conclusions and identifies further research that may arise from this work.

The most significant contribution made by this paper is in proposing a simple approach to improve the accuracy of aggregated point and probabilistic wind power forecasts that can be derived from decentralized point forecasts. This is particularly important in regions where individual wind

farms generate their own forecasts and do not necessarily have access to the measurements from other wind farms. The proposed approach provides system operators with a way of aggregating these forecasts while taking into account spatial and temporal correlations of wind power. In addition, the derived predictive densities can be used to improve operational decisions such as dynamic operating reserve allocation and stochastic unit commitment.

An additional contribution of the paper is contained in the proposed approach towards atmospheric states, derived from SOMs, which demonstrates another way in which large-scale atmospheric circulation patterns can be incorporated into short-term wind power forecasting.

II. FORECASTING METHODOLOGY
A. FEATURE SELECTION
1) CLUSTERS OF WIND POWER POINT FORECASTS
There can be hundreds of wind farms within a region and thus the dataset containing individual wind power point forecasts can be high dimensional. In most machine learning algorithms, as the number of features grows, the amount of data required to generalize accurately grows exponentially (also known as the 'curse of dimensionality'). In addition, wind farms that are in close proximity are highly correlated as (mentioned in Section I) and therefore, the machine learning model might not learn anything insightful by considering these wind farms independently. To mitigate against this phenomenon, this paper clusters wind farms with correlated point forecasts together. Some of the most common clustering algorithms that can be used to achieve this include partitioning, include hierarchical, k-means, partitioning around medoids (PAM) and the clustering large applications algorithm (CLARA). Recent studies have shown that the CLARA algorithm produces better clustering results for large datasets [48], [49], and hence it is used in this paper. The CLARA algorithm can be summarised in the following steps:

- Creating random subsets with fixed size from original dataset.
- Choosing the number of clusters k and corresponding k medoids for each subset.
- Calculating the dissimilarity matrix and assigning each observation of the dataset to the closest medoid.
- Calculating the mean of the dissimilarities of the observations to their closest medoid.
- Repeating the process while retaining the sub-dataset for which the mean is minimal.

The silhouette coefficient is used in this paper to find the optimal number of clusters k, while the distance metric used is the Euclidean distance (as also recommended for wind resource clustering in [48]).

2) ATMOSPHERIC STATES
A useful way to classify atmospheric circulation is by using SOMs, which are a class of self-learning artificial neural networks [50]. The classification output from the SOM procedure is a two-dimensional array of nodes, spaced on a lattice topology, which may be interpreted as maps showing typical patterns within a dataset. SOMs have often been used in meteorology and climatology [51], and may indeed be preferential to other commonly employed classification procedures (e.g. k-means clustering or principal component analysis (PCA)), notably principal components analysis, as it does not discretise data and does not force orthogonality [52].

The SOM training process is based on a competitive learning algorithm which successively measures the Euclidean distance between a predefined set of SOM-nodes (or reference vectors) and the input feature vectors. For each iteration of the training process, the best matching unit's (BMU) weight, along with the weights of nodes located in the BMUs proximity on the SOM lattice, is updated towards that of the feature vector. Reference nodes on the SOM lattice thereby develop towards a generalized configuration of the training dataset. Once the training process has been completed, each feature vector in the classification time-series may be clustered based on the weighted Euclidean to each node on the SOM map. In other words, each time-step in the input-data is retroactively clustered by being assigned the node number (or atmospheric state in this instance), to which it is most similar.

3) HOUR OF DAY AND MONTH OF YEAR
The cyclical nature of the processes responsible for diurnal and seasonal variability – i.e., the rotation of the earth around its axis and around the sun – however does imbue these processes with a measure of statistical regularity. This statistical regularity increases the wind power predictability associated with such cyclical diurnal and seasonal processes. The value of modelling these variations in wind power forecasting has been shown in the recent literature [53]. This paper captures these variations by including the attributes 'hour of day' (for diurnal variations) and 'month of year' (for seasonal variations), which takes the values $0, 1, \ldots 23$ and $0, 1, \ldots 11$, respectively.

B. K-NN ALGORITHM
The k-NN algorithm is a non-parametric method that averages the k closest training examples in feature space to predict the new data point (also known as the query point). This method resembles the similar-day approach that is still used by many system operators for short term load demand forecasting [53]. The similar-day approach predicts the load demand using historical days with similar weather conditions and day types to the day of forecast. In the same way, the k-NN algorithm is used in this paper to predict the aggregated wind power using historical examples with similar month of year, hour of day, point forecasts (of wind farm clusters) and atmospheric states. The k-NN algorithm can be summarised in three steps:

- Calculating the distance between the query point and each training point.

- Selecting the k nearest neighbors from the training set with k smallest distances.
- Predicting the aggregated wind power output based on a weighted average of k nearest neighbors.

If $X_1, X_2, \ldots, X_K$ are the k nearest neighbors and $P_1, P_2, \ldots, P_K$ are their corresponding aggregated wind power observations, then the aggregated wind power prediction $\hat{p}(x)$ can be derived using the estimator given by:

$$\hat{p}(x) = \frac{\sum_{k=1}^{K} K(x, X_k) . P_k}{\sum_{k=1}^{K} K(x, X_k)} \quad (1)$$

where $K(x, X_k)$ is a kernel function that assigns weights on aggregated wind power observations P_k based on the distance from the neighbor X_k to the query point x. The distance metric used in this paper is the Manhattan distance. The aggregated wind power observations of the neighbor with the smallest distance from the query point has more influence on the final point prediction $\hat{p}(x)$, as compared to the other observations. For a Gaussian kernel function (which is used in this paper) with a smoothing bandwidth h, the function $K(x, X_k)$ can be written as:

$$K(x, X_k) = e^{-\frac{(x - X_k)^2}{2h^2}} \quad (2)$$

C. CONDITIONAL KERNEL DENSITY ESTIMATOR

Once the aggregated wind power point forecasts have been determined using (5), the KDE-based approach is used to derive the conditional aggregated wind power predictive density. This involves estimating the conditional probability density function of aggregated wind power P, given that the explanatory variable X is equal to x. The main advantage of KDE is that it is a non-parametric and data driven approach that can estimate the density of a random variable without distribution hypothesis. As a result, KDE-based approaches are becoming popular in recent developments of probabilistic wind power forecasting [3], [10], [53], [54], [55], [56]. The main drawback of KDE-based approaches is the difficulty in selecting good bandwidths, especially in the presence of large datasets and high dimensionality. For this reason, this paper considers different combinations of explanatory variables (explained in 2.1) to avoid using all variables at once (and thus reducing the dimensionality of the dataset).

The standard conditional KDE (also known as the Nadaraya-Watson Conditional Estimator) can be written as:

$$\hat{f}(p \mid X = x) = \frac{1}{h_p} \sum_{i=1}^{N} K\left(\frac{p - P_i}{h_p}\right) . w_i(x) \quad (3)$$

Having

$$w_i(x) = \frac{K\left(\frac{x - X_i}{h_x}\right)}{\sum_{i=1}^{N} K\left(\frac{x - X_i}{h_x}\right)} \quad (4)$$

where h_p and h_x are bandwidths controlling the smoothness of each conditional density in p and x directions, respectively,

X_i is a point in a training set and P_i is the corresponding aggregated wind power observation, and K is a kernel function.

The choice of a kernel function K and bandwidths h_p and h_x has a significant influence on estimated conditional densities. For most applications, Gaussian kernel function is the popular choice. However, it is well known that wind power output follows a non-Gaussian distribution [10], [53]. In general, the mean squared error of the Epanechnikov kernel function is optimal [55], [57], and hence it is used in this paper. The Epanechnikov kernel function can be written as:

$$K(\partial) = \begin{cases} \frac{3}{4}(1 - \partial^2), & \partial \in [-1, 1] \\ 0, & otherwise \end{cases} \quad (5)$$

where $\partial = \frac{p - P_i}{h_p}$ and $\partial = \frac{x - X_i}{h_x}$ in p and x directions, respectively.

To select the bandwidths h_p and h_x, this paper uses the least-squares cross-validation (LSCV) method. The method is based on selecting h_p and h_x that minimises the integrated squared error (ISE) given by:

$$
\begin{aligned}
ISE(h_p, h_x) &= \int \left(\hat{f}(p \mid x) - f(p \mid x)\right)^2 dp \\
&= \int (\hat{f}(p \mid x))^2 dp - 2 \int \hat{f}(p \mid x) f(p \mid x) dp \\
&\quad + \int (f(p \mid x))^2 dp
\end{aligned} \quad (6)
$$

Therefore, minimising ISE is equivalent to minimising the first two terms, $\int (\hat{f}(p \mid x))^2 dp - 2 \int \hat{f}(p \mid x) f(p \mid x) dp$, since the last term does not involve h_p and h_x. Thus, a cross-validated estimate of ISE is given by:

$$LSCV(h_p, h_x) = \frac{1}{n} \sum_{i=1}^{n} \int (\hat{f}_{-i}(p \mid X_i))^2 dp - \frac{2}{n} \sum_{i=1}^{n} \hat{f}_{-i}(P_i \mid X_i) \quad (7)$$

where, $\hat{f}_{-i}$ is $\hat{f}$ evaluated with (X_i, P_i) left out.

D. SUMMARY OF THE PROPOSED METHODOLOGY

In summary, a methodology is proposed that trains k-NN and conditional KDE models to derive aggregated point and probabilistic wind power forecasts, respectively. The explanatory variables considered in training these models include: decentralized point forecasts (clustered using CLARA), atmospheric states (derived using SOMs), month of year and hour of day. Fig. 2 shows the flowchart summarising the proposed methodology for aggregated wind power forecasting.

III. EVALUATION FRAMEWORK
A. POINT FORECASTS
There are many metrics that can be used to evaluate the accuracy of wind power point forecasts. The most frequently used metrics are mean absolute error (MAE), root mean squared error (RMSE) and coefficient of determination (R^2) [8], [13], [14], [55]. The same metrics are used in this paper to evaluate
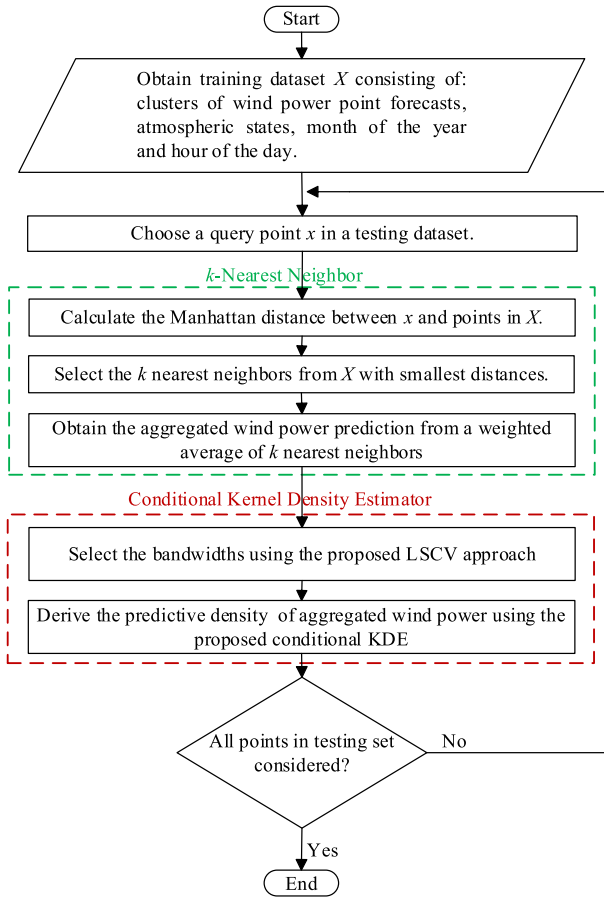
FIGURE 2. Flowchart summarizing the proposed methodology.

the accuracy of aggregated wind power point forecasts, and are calculated as:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_i - \hat{p}_i}{C_t} \right| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{P_i - \hat{p}_i}{C_t} \right)^2} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (P_i - \hat{p}_i)^2}{\sum_{i=1}^{N} (P_i - \bar{P})^2} \tag{10}$$

where P_i and $\hat{p}_i$ are the i-th actual and forecasted values of aggregated wind power, respectively, N is the total number of forecasting samples and C_t is the total wind capacity in a region. Note that the total wind capacity C_t is used as a denominator for (8) and (9) instead of actual wind power P_i to avoid the effects of aggregated wind power that is close to zero. In general, smaller MAE and RMSE, and R^2 value that is close to 1, indicate better performance of the forecasting model.

B. PROBABILISTIC FORECASTS

To evaluate the aggregated wind power predictive densities, this paper adopts the framework defined in [58].

The framework is based on three metrics: reliability or calibration, sharpness and skill score. Evaluating probabilistic forecasts based on these metrics requires the evaluation set consisting of quantile forecasts (of various nominal proportions) and observations.

1) RELIABILITY OR CALIBRATION

This metric measures the statistical consistency between quantile forecasts and observations. For example, a quantile forecast with a nominal level of 0.5 should contain 50% of the observed values lower or equal to its value. For a given quantile forecast $\hat{q}_t^{(\alpha)}$ at time t with nominal level α, and the corresponding observation p_t, the indicator variable $\xi_t^{(\alpha)}$ is given by:

$$\xi_t^{(\alpha)} = \begin{cases} 1, & if \ p_t < \hat{q}_t^{(\alpha)} \\ 0, & otherwise \end{cases} \tag{11}$$

The empirical level $a_k^{(\alpha)}$ is obtained by calculating the mean of indicator variable $\xi_t^{(\alpha)}$ over a set of T quantile forecasts.

$$a_k^{(\alpha)} = \frac{1}{T} \sum_{t=1}^{T} \xi_t^{(\alpha)} \tag{12}$$

In a 'perfect' calibration, the empirical levels match the nominal proportions.

2) SHARPNESS

This metric measures how tight or concentrated the predictive densities are. Let $\delta_t^{(\beta)} = \hat{q}_t^{(1-\alpha/2)} - \hat{q}_t^{\alpha/2}$ be the width of a given prediction interval estimated at time t. The sharpness is obtained by calculating the average width $\bar{\delta}_t^{(\beta)}$ of the prediction interval over a set of T quantile forecasts. Mathematically, this is given by:

$$\bar{\delta}_t^{(\beta)} = \frac{1}{T} \sum_{t=1}^{T} \delta_t^{(\beta)} \tag{13}$$

In general, narrow prediction intervals (subject to calibration) are preferred because they suggest that predictions have more information content.

3) SKILL SCORE

This metric assesses probabilistic forecasts by a single score, just like MAE and RMSE for the point forecasts. This score incorporates a variety of aspects of probabilistic forecast evaluation and hence allows for easier comparison between probabilistic forecasting approaches. In this paper, the skill score S_c for a set of M quantiles is calculated as:

$$S_c = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{M} (\xi_t^{\alpha_i} - \alpha_i).(p_t - \hat{q}_t^{\alpha_i}) \tag{14}$$

This score cannot be decomposed to provide contributions of calibration, sharpness and other aspects of probabilistic forecast evaluation. In general, a higher score means higher skill of the probabilistic forecast approach and the score is zero for a 'perfect' probabilistic forecast.
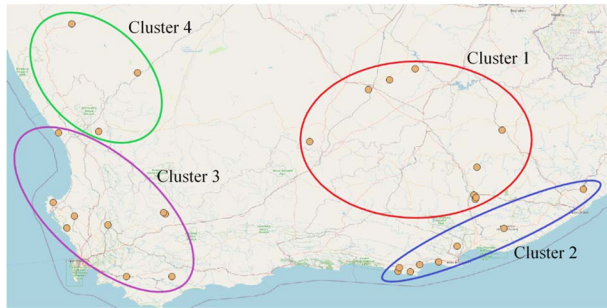
FIGURE 3. Geographical locations of wind farms used for testing the methodology.

IV. CASE STUDY

A. DESCRIPTION OF CASE STUDY

The proposed methodology for aggregated wind power forecasting is tested using the data obtained from 29 wind farms in South Africa from 01 January 2018 to 31 March 2021. The average installed capacity of these wind farms is 92 MW (ranging between 13 MW and 140 MW) and their geographical locations are shown in Fig. 3. Since the individual wind power point forecast data is sensitive and confidential, Eskom (the utility company in South Africa) was only able to send the day-ahead wind power point forecasts in groups of three nearest wind farms (summed together). The challenge here is that some of the wind farms may not fall in the same cluster as its two nearest neighbors if CLARA was applied to the original dataset, instead of applying it to already clustered dataset (as received from the utility). This implies a drop in resolution of the data, which may have impacted the performance of the proposed methodology in this particular case study.

The SOM Toolbox developed for Matlab by Helsinki University of Technology was used [59] for the classification of atmospheric states. As classification parameter, the hourly ERA5 850hPa geopotential height reanalysis dataset [60] was selected, as geopotential heights provide a good representation of large scale atmospheric circulation, and is a wind speed predictor e.g. [47] and [61]. It was further deemed that the 850 hPa pressure level provides sufficient elevation to capture circulation above the South African escarpment. The classification area was bounded between 22-40°S and 5-40°E, which includes the South African subcontinent along with significant areas of the Indian and Atlantic oceans so as to adequately capture the large-scale circulation impacting the study area. To illustrate the proposed aggregation methodology, a 4×5 SOM-node topology was selected, which was deemed to provide sufficient detail. It should however be noted that the number of SOM nodes chosen represents what is often a subjective trade-off between level of generalization or detail required and quantization error.

The SOM was trained in two phases with a batch training algorithm, using a rectangular lattice as recommended for batch training [62]. The first (rough) training phase consisted of 1000 iterations with the neighborhood decreasing

TABLE 1. Overall point prediction results of the proposed methodology.

Method	Sum of point forecasts	Proposed methodology
MAE	7.12	3.18
RMSE	9.11	4.63
R^2	0.77	0.94

from 5 to 1, which was followed by a fine-tuning phase of 5000 iterations where the neighborhood function decreased from 2 to 1, thereafter remaining fixed after which only the BMU was updated. A number of studies have showed the starting point of the neighborhood function to have little impact on patterns produced [63], however in continuing the training once the radius parameter equals 1, the SOM approach becomes the equivalent of the k-means clustering procedure, which has been shown to be advantages in terms of average classification error [64]. The Epanechikov neighborhood function was used, as recommended by [62] when training small SOM maps.

The full dataset (consisting of hour of day, month of year, actual wind power generation, clusters of wind power point forecasts, and atmospheric states) was split into a training set (by randomly picking 90% rows from the full dataset) and testing set (remaining 10% from the full dataset). The seven-fold cross validation was used to determine the optimal value of k in k-NN algorithm. The aggregated wind power point forecast results obtained by training the k-NN algorithm with the proposed explanatory variables were compared to the results obtained by simply summing up the point forecasts. In addition, the k-NN algorithm was trained using different combinations of explanatory variables to evaluate the impact of each variable on the accuracy of the resulting aggregated forecasts. Furthermore, in order to assess the proposed model under various conditions, its performance evaluated during different hours of the day, months of the year, and atmospheric states.

This paper trained the conditional KDE using various combinations of proposed explanatory variables, and the results were compared to those obtained by training the conditional KDE with just the sum of point forecasts. This was done because it was difficult to train the conditional KDE using all proposed explanatory variables at once (due to bandwidth selection as explained in Section II-C).

B. EVALUATION OF POINT FORECASTS

In Table 1, the overall performance of training the k-NN model with proposed explanatory variables is compared to simply adding the point forecasts by individual wind farms. The proposed methodology performs better than the sum of point forecasts within the context of the proposed evaluation metrics. The MAE and RMSE of the proposed methodology are 55% and 47% less than those achieved by adding the point forecasts, respectively. Likewise, the R^2 values show that 94% of the observed data fit the proposed model, while 77% of the data fit the sum of point forecasts model. Therefore, taking into account the spatial and temporal

TABLE 2. Scenarios for explanatory variables to show how explanatory variables impact forecasting accuracy.

Explanatory variables	Clusters	Clusters + time of day	Clusters + month of year	Clusters + atmospheric states	All
MAE	6.25	5.81	4.62	4.92	3.18
RMSE	8.45	8.03	6.70	7.08	4.63
R^2	0.80	0.82	0.87	0.86	0.94

TABLE 3. Point prediction results for different hours of the day.

Hour	0	1	2	3	4	5	6	7	8	9	10	11
MAE	5.22	4.83	3.18	3.19	2.65	1.97	2.43	2.46	3.14	2.61	2.92	2.75
RMSE	6.96	6.82	4.43	5.04	3.98	3.12	3.56	3.84	4.74	3.38	4.30	3.70
R^2	0.86	0.83	0.93	0.91	0.94	0.97	0.95	0.95	0.91	0.96	0.92	0.95

Hour	12	13	14	15	16	17	18	19	20	21	22	23
MAE	2.53	3.21	2.94	3.66	3.11	3.57	3.74	3.43	3.37	3.79	3.50	3.27
RMSE	3.28	5.05	4.5	5.19	4.6	4.96	5.43	5.29	4.38	5.88	4.85	4.56
R^2	0.97	0.94	0.96	0.93	0.95	0.95	0.93	0.92	0.95	0.88	0.93	0.92

TABLE 4. Point prediction results for different months of the year.

Month	1	2	3	4	5	6	7	8	9	10	11	12
MAE	3.77	3.17	2.87	2.68	2.63	2.87	2.82	3.06	2.74	3.78	3.89	4.46
RMSE	5.22	4.65	3.89	3.82	4.10	4.71	4.09	4.63	3.80	5.81	5.37	6.23
R^2	0.91	0.91	0.95	0.96	0.95	0.94	0.96	0.94	0.96	0.92	0.90	0.87

TABLE 5. Point prediction results for different atmospheric states.

Atmospheric state	1,1	1,2	1,3	1,4	1,5	2,1	2,2	2,3	2,4	2,5
MAE	3.77	4.04	4.09	3.67	3.07	3.37	3.34	4.2	2.58	2.85
RMSE	5.37	5.84	5.80	5.48	4.27	4.47	4.65	6.24	3.61	4.22
R^2	0.89	0.88	0.90	0.92	0.93	0.94	0.92	0.92	0.96	0.93

Atmospheric state	3,1	3,2	3,3	3,4	3,5	4,1	4,2	4,3	4,4	4,5
MAE	2.59	3.61	2.61	2.88	2.74	3.54	2.62	2.80	2.80	3.03
RMSE	3.57	5.60	3.87	4.25	4.21	5.07	4.08	3.81	3.96	4.56
R^2	0.96	0.94	0.95	0.92	0.94	0.95	0.93	0.95	0.93	0.94

correlations through the proposed explanatory variables improves the accuracy of aggregating decentralized point forecasts.

Table 2 displays the results of the various scenarios for explanatory variables to show how the proposed explanatory variables impact forecasting accuracy. As shown in Table 2, taking into account the spatial information found within various clusters result in better forecasting accuracy compared to the sum of point forecasts. The accuracy improves further when taking into account diurnal and seasonal cycles as well as large-scale atmospheric circulations. The seasonal cycles show superior impact followed by the large-scale atmospheric circulations.

Tables 3, 4 and 5 shows the performance of the proposed model during different hours of the day, months of the year and atmospheric states, respectively. The proposed model seems to perform better in the hours between 4h and 14h and in the months between March and July (which is the autumn and winter seasons). Fig. 4 shows the hourly average wind power generation for January (summer), April (autumn), July (winter) and October (spring). The average wind generation

profile shows daily ramp-up in wind power, especially during summer, autumn, and spring seasons, from approximately 9h to 18h, which is in turn followed by an equivalent ramp-down during the evening. This ramp-up is likely associated with daily surface heating due to the diurnal cycle. A possible reason that prediction errors are comparatively small during a period associated with significant variability in the wind power profile (4h-14h), is due to the cyclical nature of said variability, which is a function diurnal surface heating and cooling. Notwithstanding diurnal variability, Fig. 4 shows that the autumn and winter wind power profiles are comparatively flat throughout the day, which is in turn a possible reason for the better model performance during these seasons.

The model also performs better during the atmospheric states (2,4), (2,5), (3,1), (3,3), (3,4), (3,5), (4,2), (4.3) and (4.4). Fig. 5 shows the 20 (4 × 5) node SOM representing the 20 classified atmospheric states together with the frequency of SOM node occurrence as a percentage above each node. It is evident that the nodes where the proposed model performs better tends to represent atmospheric circulation
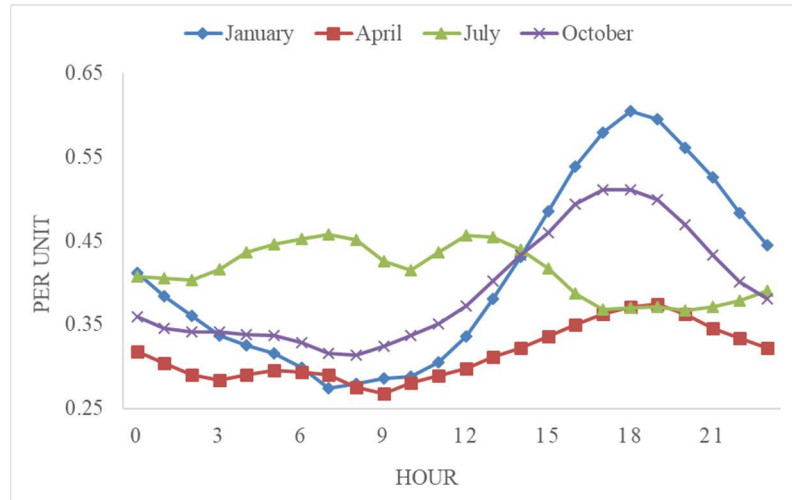
FIGURE 4. Hourly average wind generation for different months representing different seasons.
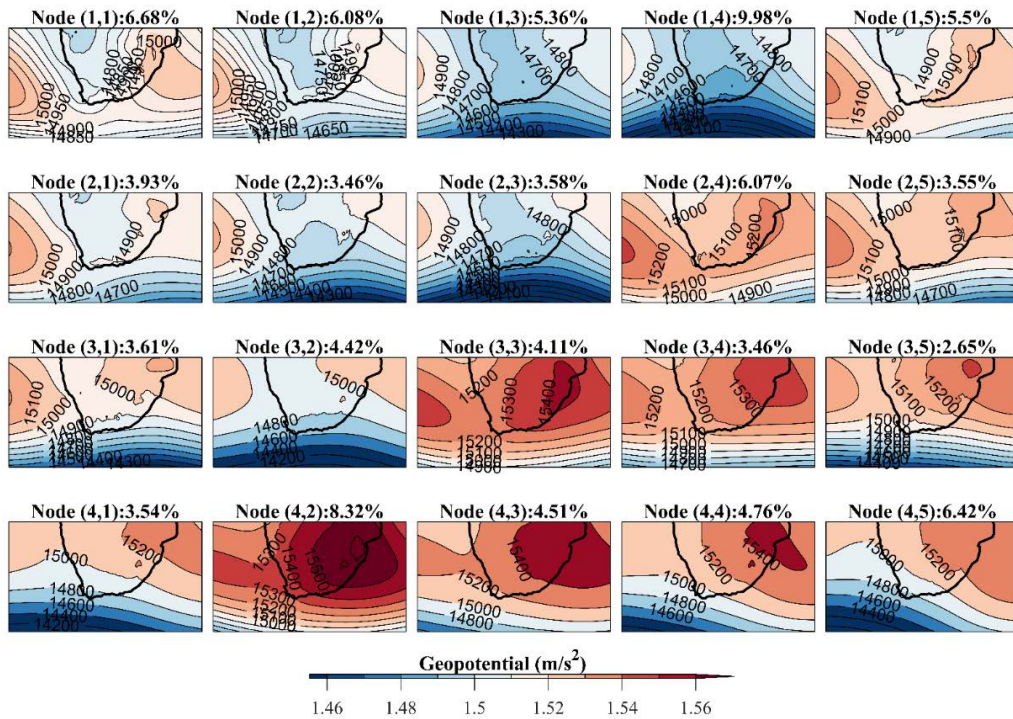

FIGURE 5. SOM geopotential heights together with frequency of SOM node occurrence over Southern Africa.

dominated by high pressure conditions. Such conditions are associated with clear skies and more stable weather conditions, and thus less variability over different timescales which increases predictive skill.

C. EVALUATION OF PROBABILISTIC FORECASTS

Fig. 6 shows the day-ahead prediction intervals (10-90%) together with the actual observations of aggregated wind power generation on 30 March 2021 (which is reflective

of a 'typical' autumn day in South Africa). Note that for Fig. 6 only clusters of wind power point forecasts were used as the explanatory variables for illustration purposes. It can be seen that only two observations (at 7h and 21h) falls outside the 90% prediction intervals. The width of the prediction intervals is time-varying with the actual observations as expected. To get a different view on these results, Fig. 7 shows the aggregated wind power generation predictive densities at 1h, 8h, 14h and 20h on 30 March 2021. It can be seen from Fig. 7 that the predictive densities are not only
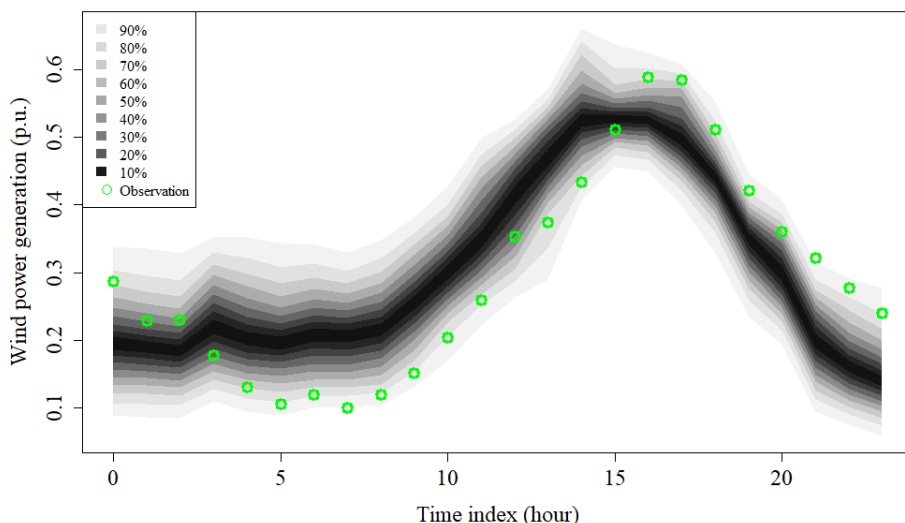
FIGURE 6. Day-ahead prediction intervals of aggregated wind power and actual observations on 30 March 2021 using clusters of point forecasts as explanatory variables.
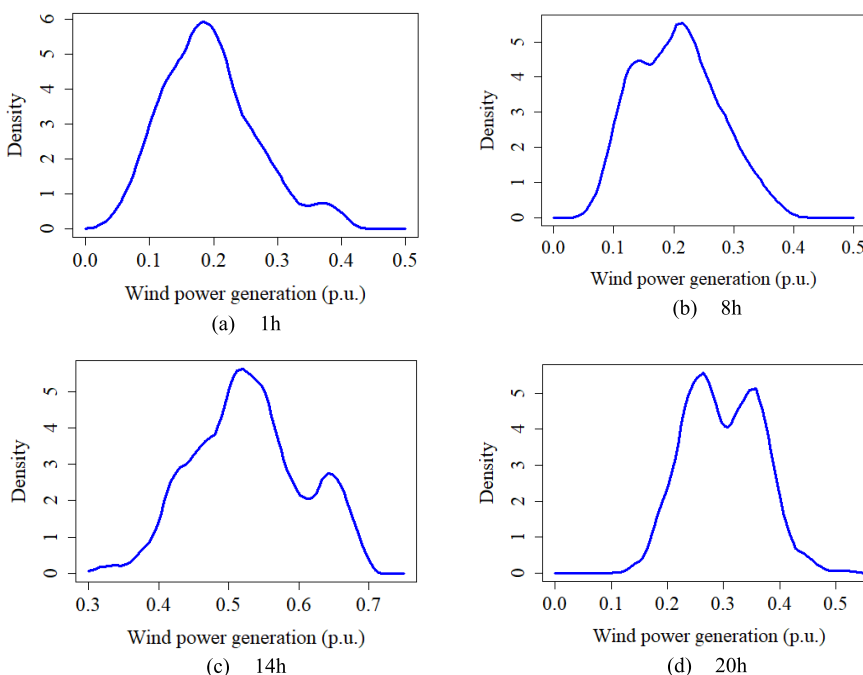

FIGURE 7. Day-ahead predictive densities of the aggregated wind power generation at 1h, 8h, 14h and 20h on 30 March 2021 using clusters of point forecasts as explanatory variables.

time-varying, but also multimodal and asymmetric (and thus non-Gaussian).

Fig. 8 shows the reliability diagrams of the probabilistic forecasts using the proposed model for different explanatory variables scenarios. In Fig. 8 (a), the empirical levels are plotted against the nominal proportions while also showing the desired line (where the empirical levels are equal to the nominal proportions). It can be seen from Fig. 8 (a) that results achieved from all considered scenarios aligns well with the desired line (indicating reliable

probabilistic forecasts). To conduct a comparative analysis between considered scenarios, Fig. 8 (b) plots the errors between empirical levels and desired outcomes, against the nominal proportions. A general conclusion from Fig. 8 (b) is that the 'sum' scenario produces better reliability results than all other scenarios. In addition, the reliability of forecasts seems to drop with the increase in explanatory variables considered in a scenario. One possible explanation for this is the bandwidths selection that becomes more complex for high number of explanatory variables. As seen through negative
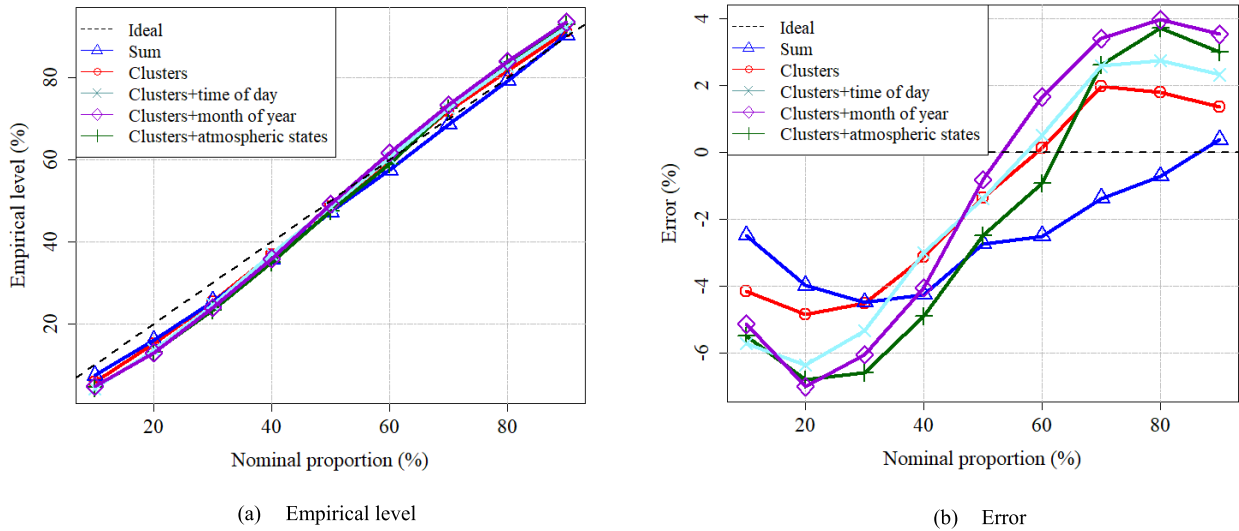
(a) Empirical level

(b) Error

FIGURE 8. Reliability diagrams of prediction intervals provided by the proposed model for different combinations of explanatory variables.

TABLE 6. Skill score results for different explanatory variables scenarios.

Explanatory variables	Sum	Clusters	Clusters + time of day	Clusters + month of year	Clusters + atmospheric states
Skill score	-0.283	-0.174	-0.157	-0.148	-0.153

reliability results, the proposed model tends to underestimate the aggregated wind power generation for the 'sum' scenario. For the rest of scenarios, the model tends to underestimate the aggregated wind power generation when nominal proportion is below 60% and overestimate for larger nominal proportions.

Fig. 9 shows the sharpness results of prediction intervals for different explanatory variables scenarios. It can be seen that the 'clusters + month of year' scenario achieves the best sharpness results (ranging between 2% and 20%) while the 'sum' scenario achieves the worst sharpness results (ranging between 2% and 28%). The results show a significant improvement from the 'sum' scenario to the 'clusters' scenario, followed by noticeable improvements from the 'clusters' scenario to the rest of scenarios. This highlights the importance of considering spatial correlations, larger atmospheric circulations, inter-hour and seasonal variations in aggregated wind power forecasting. However, good sharpness results may lead to underestimation and overestimation of aggregated wind power uncertainty as deduced from the reliability diagram in Fig. 8. Therefore, there is a need for a trade-off between reliability and sharpness depending on the specific requirements of the grid and its customers.

Table 6 shows the skill score results for different combinations of explanatory variables. It can be seen that the 'clusters + month of year' scenario achieves the best skill score (−0.148) while, the lowest skill score (−0.283) is achieved on
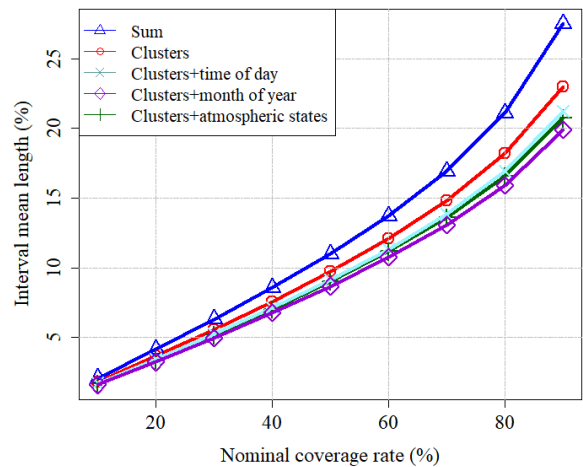

FIGURE 9. Sharpness results with different nominal coverage rates for different combinations of explanatory variables.

'sum' scenario. The results show a significant improvement from the 'sum' scenario to the 'clusters' scenario, followed by smaller improvements from the 'clusters' scenario to the rest of scenarios.

The results above show that taking into account the spatial and temporal correlations through the proposed explanatory variables also improves the accuracy of aggregated probabilistic forecasts that can be derived from decentralized forecasts.

V. CONCLUSION

This paper proposed explanatory variables that were used to train the k-NN and conditional KDE models to derive day-ahead aggregated point and probabilistic wind power forecasts from decentralized point forecasts of geographically distributed wind farms. The proposed explanatory variables include clusters of point forecasts (to account for spatial correlations between wind farms), hour of day (to account for diurnal cycles), month of year (to account for seasonal cycles) and, atmospheric states (to account for correlations due to large-scale atmospheric circulations). The main findings of this paper can be summarized as follows:

- The results from the proposed approach showed a significant improvement (47% less RMSE) as compared to simply adding the decentralized point forecasts. Therefore, the proposed well-known spatial and temporal correlations as well as large-scale atmospheric circulations can be effective in capturing some spatial and temporal correlations that are not accounted for by simply adding the decentralized forecasts.
- In the case study, the proposed approach performed better in hours between 4h and 14h. One possible reason for this can be that the variability of the wind power profile between these hours is a function of diurnal surface heating and cooling, and this was incorporated by including the hour of day feature in the proposed approach. In addition, the autumn and winter wind power profiles are comparatively flat throughout the day, which can be a possible reason for the better performance of the approach during these seasons. Furthermore, the proposed approach performed better in SOM nodes that tends to represent atmospheric circulation dominated by high pressure conditions.
- The predictive densities obtained from the proposed approach were shown to be non-Gaussian and time-varying as expected given the time-varying nature of wind uncertainty. The probabilistic forecasts from the considered scenarios were relatively reliable (less than 8% in error between empirical levels and desired outcomes). However, the reliability of forecasts appears to drop with the increase in explanatory variables considered in a scenario. This can be due to bandwidths selection that becomes more complex in the proposed KDE approach for a high number of explanatory variables.
- The sharpness and skill score results showed a significant improvement when the proposed explanatory variables were considered as compared to simply adding decentralized forecasts. In addition, the month of year feature produce slightly better sharpness and skill scores compared to the hour of day and atmospheric state features, which highlights the importance of incorporating effects of seasonality in aggregated forecasting.

The most significant contribution made by this paper is in proposing an approach for incorporating some spatial and temporal correlations to improve accuracy of aggregated forecasts derived from decentralized point forecasts. In addition, knowing the hours, months, and atmospheric states under which the proposed approach performs better together with the derived predictive densities can be used as inputs to operational processes such as stochastic unit commitment and dynamic operating reserve allocation. This paper also demonstrated a way in which large-scale atmospheric circulation patterns can be incorporated into short-term wind power forecasting – which has been lacking in the literature. In future, one can test the proposed methodology using ungrouped wind data to see if there can be further improvements on the results. In addition, more work needs to be done in terms of bandwidths selection of high dimensional dataset in KDE-based approaches.

REFERENCES

[1] Y. Zhang and J. Wang, "A distributed approach for wind power probabilistic forecasting considering spatiooral correlation without direct access to off-site information," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5714–5726, Sep. 2018.

[2] A. Lenzi, I. Steinsland, and P. Pinson, "Benefits of spatiotemporal modeling for short-term wind power forecasting at both individual and aggregated levels," *Environmetrics*, vol. 29, no. 3, pp. 1–17, 2018.

[3] Z. Wang, W. Wang, C. Liu, B. Wang, and S. Feng, "Probabilistic forecast for aggregated wind power outputs based on regional NWP data," *J. Eng.*, vol. 2017, no. 13, pp. 1528–1532, Jan. 2017.

[4] J. Tastu, P. Pinson, E. Kotwa, H. Madsen, and H. A. Nielsen, "Spatio-temporal analysis and modeling of short-term wind power forecast errors," *Wind Energy*, vol. 14, no. 1, pp. 43–60, 2010.

[5] A. Malvaldi, S. Weiss, D. Infield, J. Browell, P. Leahy, and A. M. Foley, "A spatial and temporal correlation analysis of aggregate wind power in an ideally interconnected Europe," *Wind Energy*, vol. 20, no. 8, pp. 1315–1329, Aug. 2017.

[6] M. Sun, C. Feng, and J. Zhang, "Aggregated probabilistic wind power forecasting based on spatio-temporal correlation," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2019, pp. 1–5.

[7] P. Li, X. Guan, and J. Wu, "Aggregated wind power generation probabilistic forecasting based on particle filter," *Energy Convers. Manag.*, vol. 96, pp. 579–587, May 2015.

[8] M. Sun, C. Feng, and J. Zhang, "Conditional aggregated probabilistic wind power forecasting based on spatio-temporal correlation," *Appl. Energy*, vol. 256, Dec. 2019, Art. no. 113842.

[9] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.

[10] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660–669, Oct. 2012.

[11] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renew. Sustain. Energy Rev.*, vol. 13, no. 4, pp. 915–920, 2009.

[12] G. Giebel, G. N. Kariniotakis, and C. Draxl, "The state of the art in short-term prediction of wind power—A literature review," ANEMOS, Oldenburg, Germany, ANEMOS.plus Project Deliverable Rep. D1.2, 2011.

[13] P. Mandal, H. Zareipour, and W. D. Rosehart, "Forecasting aggregated wind power production of multiple wind farms using hybrid wavelet-PSO-NNs," *Int. J. Energy Res.*, vol. 38, no. 13, pp. 1654–1666, 2014.

[14] Q. Zhu, J. Chen, L. Zhu, X. Duan, and Y. Liu, "Wind speed prediction with spatio-temporal correlation: A deep learning approach," *Energies*, vol. 11, no. 4, pp. 1–18, 2018.

[15] B. Alonzo, P. Tankov, P. Drobinski, and R. Plougonven, "Probabilistic wind forecasting up to three months ahead using ensemble predictions for geopotential height," *Int. J. Forecasting*, vol. 36, no. 2, pp. 515–530, Apr. 2020.

[16] P. Pinson, H. A. Nielsen, H. Madsen, and G. Kariniotakis, "Skill forecasting from ensemble predictions of wind power," *Appl. Energy*, vol. 86, nos. 7–8, pp. 1326–1334, Jul. 2009.

[17] J. W. Taylor, P. E. McSharry, and R. Buizza, "Wind power density forecasting using ensemble predictions and time series models," *IEEE Trans. Sustain. Energy*, vol. 24, no. 3, pp. 775–782, Sep. 2009.

[18] N. Chen, Z. Qian, I. T. Nabney, and X. Meng, "Wind power forecasts using Gaussian processes and numerical weather prediction," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 656–665, Mar. 2014.

[19] L. Landberg, G. Giebel, L. Myllerup, J. Badger, H. Madsen, and T. S. Nielsen, "Poor-man's ensemble forecasting for error estimation," in *Proc. Amer. Wind Energy Assoc. Conf.*, 2002, pp. 1–8.

[20] M. Lange, U. Focken, R. Meyer, M. Denhardt, B. Ernst, and F. Berster, "Optimal combination of different numerical weather models for improved wind power predictions," in *Proc. 6th Int. Workshop Large-Scale Integr. Wind Power Transmiss. Networks Offshore Wind Farms*, 2006, pp. 273–276.

[21] M. Negnevitsky, P. Mandal, and A. K. Srivastava, "An overview of forecasting problems and techniques in power systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2009, pp. 1–4.

[22] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renew. Sustain. Energy Rev.*, vol. 32, pp. 255–270, Apr. 2014.

[23] W.-Y. Chang, "A literature review of wind forecasting methods," *J. Power Energy Eng.*, vol. 2, no. 4, pp. 161–168, 2014.

[24] P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, no. 2, pp. 119–132, 2004.

[25] P. Pinson and G. N. Kariniotakis, "On-line adaption of confidence intervals based on weather stability for wind power forecsting," in *Proc. Global Windpower Conf.*, 2004, pp. 1–12.

[26] J. B. Bremnes, "A comparison of a few statistical models for making quantile wind power forecasts," *Wind Energy*, vol. 9, nos. 1–2, pp. 3–11, 2006.

[27] C. Yildiz, M. Tekin, A. Gani, Ö. Keçecioglu, H. Açıkgö, and M. Sekkeli, "A day-ahead wind power scenario generation, reduction, and quality test tool," *Sustainability*, vol. 9, no. 5, p. 864, May 2017.

[28] H. Bludszuweit, J. A. Dominguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 983–991, Aug. 2008.

[29] A. Fabbri, T. G. S. Roman, J. R. Abbad, and V. H. M. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1440–1446, Aug. 2005.

[30] S. Bofinger, A. Luig, and H. G. Beyer, "Qualification of wind power forecasts," in *Proc. Global Wind Power Conf.*, 2002, pp. 2–5.

[31] N. Menemenlis, M. Huneault, and A. Robitaille, "Computation of dynamic operating balancing reserve for wind power integration for the time-horizon 1–48 hours," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 692–702, Oct. 2012.

[32] B. Mauch, J. Apt, P. M. S. Carvalho, and M. J. Small, "An effective method for modeling wind power forecast uncertainty," *Energy Syst.*, vol. 4, no. 4, pp. 393–417, Dec. 2013.

[33] K. Porter, J. Rogers, K. Porter, and J. Rogers, "Status of centralized wind power forecasting in North America: May 2009-may 2010," Nat. Renew. Energy Lab. (NREL), Golden, CO, USA, Tech. Rep. NREL/SR-550-47853, 2010.

[34] *Implementation Agreement Between the Seller and the Department of Energy Pursuant to the Renewable Energy Independent Power Producer Procurement Programme*, Ministry of Energy, Pretoria, South Africa, 2012.

[35] M. Fripp and R. H. Wiser, "Effects of temporal wind patterns on the value of wind-generated electricity in California and the Northwest," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 477–485, May 2008.

[36] S. Rehman, "Wind energy resources assessment for Yanbo, Saudi Arabia," *Energy Convers. Manag.*, vol. 45, nos. 13–14, pp. 2019–2032, Aug. 2004.

[37] B. Karki and R. Billinton, "Effects of seasonality and locality on the operating capacity benefits of wind power," in *Proc. IEEE Electr. Power Energy Conf. (EPEC)*, Oct. 2009, pp. 1–6.

[38] K. Knorr, B. Zimmermann, S. Bofinger, A. Gerlach, T. Bischof-Niemz, and C. Mushwana, "Wind and solar PV resource aggregation study for South Africa," Council Sci. Ind. Res., Pretoria, South Africa, Tech. Rep. 542-23-02-2015, 2016.

[39] F. M. Mulder, "Implications of diurnal and seasonal variations in renewable energy generation for large scale energy storage," *J. Renew. Sustain. Energy*, vol. 6, no. 3, May 2014, Art. no. 033105.

[40] R. Carapellucci and L. Giordano, "The effect of diurnal profile and seasonal wind regime on sizing grid-connected and off-grid wind power plants," *Appl. Energy*, vol. 107, pp. 364–376, Jul. 2013.

[41] L. Zhou, Y. Tian, S. B. Roy, Y. Dai, and H. Chen, "Diurnal and seasonal variations of wind farm impacts on land surface temperature over Western Texas," *Climate Dyn.*, vol. 41, no. 2, pp. 307–326, Jul. 2013.

[42] A. Dalton, B. Bekker, and M. J. Koivisto, "Atmospheric circulation archetypes as clustering criteria for wind power inputs into probabilistic power flow analysis," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Aug. 2020, pp. 1–6.

[43] A. Dalton, B. Bekker, and M. J. Koivisto, "Simulation and detection of wind power ramps and identification of their causative atmospheric circulation patterns," *Electric Power Syst. Res.*, vol. 192, Mar. 2021, Art. no. 106936.

[44] A. Dalton, B. Bekker, and M. J. Koivisto, "Classified atmospheric states as operating scenarios in probabilistic power flow analysis for networks with high levels of wind power," *Energy Rep.*, vol. 7, pp. 3775–3784, Nov. 2021.

[45] D. J. Brayshaw, A. Troccoli, R. Fordham, and J. Methven, "The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: A case study over the U.K.," *Renew. Energy*, vol. 36, no. 8, pp. 2087–2096, Aug. 2011.

[46] H. C. Bloomfield, D. J. Brayshaw, P. L. M. Gonzalez, and A. Charlton-Perez, "Sub-seasonal forecasts of demand and wind power and solar power generation for 28 European countries," *Earth Syst. Sci. Data*, vol. 13, no. 5, pp. 2259–2274, May 2021.

[47] R. J. Davy, M. J. Woods, C. J. Russell, and P. A. Coppin, "Statistical downscaling of wind variability from meteorological fields," *Boundary-Layer Meteorol.*, vol. 135, no. 1, pp. 161–175, Apr. 2010.

[48] C. Y. J. van Vuuren and H. J. Vermeulen, "Clustering of wind resource data for the south African renewable energy development zones," *J. Energy Southern Afr.*, vol. 30, no. 2, pp. 126–143, Jun. 2019.

[49] T. Gupta and S. Panda, "A comparison of K-means clustering algorithm and CLARA clustering algorithm on Iris dataset," *Int. J. Eng. Technol.*, vol. 7, no. 4, pp. 4766–4768, 2018.

[50] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990.

[51] R. Huth, C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul, M. Cahynová, J. Kyselý, and O. E. Tveito, "Classifications of atmospheric circulation patterns," *Ann. New York Acad. Sci.*, vol. 1146, no. 1, pp. 105–152, Dec. 2008.

[52] C. Lennard and G. Hegerl, "Relating changes in synoptic circulation to the surface rainfall response using self-organising maps," *Climate Dyn.*, vol. 44, nos. 3–4, pp. 861–879, Feb. 2015.

[53] Y. Zhang and J. Wang, "K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting," *Int. J. Forecasting*, vol. 32, no. 3, pp. 1074–1080, Jul. 2016.

[54] Y. He, R. Liu, H. Li, S. Wang, and X. Lu, "Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory," *Appl. Energy*, vol. 185, pp. 254–266, Jan. 2017.

[55] Y. He and H. Li, "Probability density forecasting of wind power using quantile regression neural network and kernel density estimation," *Energy Convers. Manag.*, vol. 164, pp. 374–384, May 2018.

[56] X. Yang, X. Ma, N. Kang, and M. Maihemuti, "Probability interval prediction of wind power based on KDE method with rough sets and weighted Markov chain," *IEEE Access*, vol. 6, pp. 51556–51565, 2018.

[57] A. Araveeporn, "The comparison of bandwidth selection methods," *KMITL Sci. Tech. J.*, vol. 11, no. 2, pp. 64–78, 2011.

[58] P. Pinson, H. A. Nielsen, J. K. Moller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: Required properties and evaluation," *Wind Energy*, vol. 10, no. 6, pp. 497–516, Nov./Dec. 2007.

[59] E. Alhoniemi, J. Himberg, J. Parhankangas, and J. Vesanto, "SOM Toolbox," Lab. Inf. Comput. Sci., Helsinki Univ. Technol., Helsinki, Tech. Rep. A57, 2002.

[60] Copernicus, "ERA5: Fifth generation of ECMWF 940 atmospheric reanalyses of the global climate," Copernicus Climate Change Service Climate Data Store (CDS), Eur. Union, Brussels, Belgium, 2017.

[61] J. Browell, D. R. Drew, and K. Philippopoulos, "Improved very short-term spatio-temporal wind forecasting using atmospheric regimes," *Wind Energy*, vol. 21, no. 11, pp. 968–979, Nov. 2018.

[62] Y. Liu, R. H. Weisberg, and C. N. K. Mooers, "Performance evaluation of the self-organizing map for feature extraction," *J. Geophys. Res.*, vol. 111, no. 5, pp. 1–14, 2006.

[63] P. B. Gibson, S. E. Perkins-Kirkpatrick, P. Uotila, A. S. Pepler, and L. V. Alexander, "On the use of self-organizing maps for studying climate extremes," *J. Geophys. Res., Atmos.*, vol. 122, no. 7, pp. 3891–3903, Apr. 2017.

[64] N. Jiang, K. Luo, P. J. Beggs, K. Cheung, and Y. Scorgie, "Insights into the implementation of synoptic weather-type classification using self-organizing maps: An Australian case study," *Int. J. Climatol.*, vol. 35, no. 12, pp. 3471–3485, Oct. 2015.

AMARIS DALTON (Member, IEEE) received the B.Sc. degree (Hons.) in meteorology, in 2014, the M.Sc. degree in environmental management from the University of Pretoria, in 2017, and the Ph.D. degree in electrical engineering from Stellenbosch University, in 2020, worked under Dr. Bernard Bekker's supervision. From 2015 to 2018, he worked at Eskom as an Environmentalist. Since 2021, he has been working as a Postdoctoral Research Fellow at the Centre for Renewable and Sustainable Energy Studies, Stellenbosch University. His research interests include energy meteorology, wind power prediction, and power systems modeling.

NDAMULELO MARARAKANYE received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Cape Town, South Africa, in 2013 and 2017, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with Stellenbosch University. He works as a Senior Engineer at the Centre for Renewable and Sustainable Energy Studies. His research interests include variable renewable energy grid integration studies, forecasting, characterizing uncertainty, and power system operations.

BERNARD BEKKER (Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from Stellenbosch University, South Africa, in 1996 and 2004, respectively, and the Ph.D. degree in electrical engineering from the University of Cape Town, South Africa, in 2010. He is currently the Eskom Research Chair in power system simulation with Stellenbosch University, where he is also an Associate Director of the Centre for Renewable and Sustainable Energy Studies. His research interests include the grid integration of distributed energy resources, probabilistic modeling, and variable renewable energy forecasting.

• • •