Short-Term Forecasting of Wind Power Plant Generation for System Stability and Provision of Ancillary Services

Dr Harley Mackenzie
Managing Director
HARD software
Melbourne, Australia
hjm@hardsoftware.com

Jonathon Dyson
Director
Greenview Strategic Consulting
Melbourne, Australia
jonathon@greenviewstrategicconsulting.com.au

Abstract—**Renewable wind generation presents many challenges for accurate and timely generation forecasts due to site topology, turbine layout and their potential wake effects and local network constraints, particularly for generators operating in fast changing, low or high wind resource conditions. Internationally, short term generation forecasting for wind power plants has become a critical issue for market operators and transmission networks to adequately manage system reserve margins and allow for the delivery of ancillary services by wind generators.**

The authors have investigated and developed machine learning techniques in combination with robust site and plant condition heuristics to produce a short-term generation forecast of wind power plant generation for the 5 to 7 minute timeframes to meet new requirements in the Australian National Electricity Market (NEM) for wind power plant sourced unconstrained generation forecasts. Similarly, wind power plants will need reliable short term generation forecasts to be able to participate in contingency and regulation raise markets in the NEM to offset a significant increase in the generator costs for ancillary services over the last three years.

The forecasting system will be implemented as a standalone device designed to integrate directly with wind power plant SCADA systems and therefore can be deployed at many new and existing wind generators. Preliminary generation forecasting results are very encouraging and already demonstrate a substantial improvement in the accuracy of short term forecasts in comparison to simple generation state estimation and market operator generated forecasts for actual wind power plant case studies for a wide range of generator sites.

Accurate and timely short term forecasts will have a critical role to play in ensuring improved system stability for power systems with a large or increasing proportions of wind power plant generation. The proposed system has the potential to fulfill a missing vital technological component for improved wind generator revenue through improved dispatch and ancillary services income streams and cost reductions and lead to increased power system stability and reliability.

Keywords - wind; forecasting; short-term; generation;

INTRODUCTION

The Australian National Electricity Market (NEM), like many other international competitive market based electricity systems, must address system security and dispatch challenges associated with the introduction of increasing levels of intermittent generators. Both wind and solar power plants are being integrated into an existing market design that were based on the scheduled dispatch of hydro, coal and gas power plants. As countries respond to their agreed emission targets, intermittent generators are now forming a substantial proportion of the generation portfolio and the need to efficiently dispatch and manage system security is becoming critical in many international electricity markets.

As noted in [1], wind power plants in the NEM are currently dispatched as a semi-scheduled generator and can generate at any level unless constrained by network or pricing constraints. Dispatch solutions within the NEM are co-optimized between energy and Frequency Control Ancillary Services (FCAS) market bids with the NEM dispatch engine (NEMDE) optimizing the price paid by the market for both energy and ancillary services by analyzing the capability of all generating units, loads, and unit ramp rates.

For NEMDE to optimally dispatch a generator, the Australian Energy Market Operator (AEMO) requires accurate information on the prevailing conditions at the wind power plant. The Australian Wind Energy Forecasting System (AWEFS) uses a mixture of persistence forecasts and turbine availability to create the Unconstrained Intermittent Generation Forecast (UIGF) of wind power plants to produce a central forecast ranging from 5 minutes to 7 days.

Changes in the past 12 months [2] have increased the level of SCADA information sent to AEMO by wind power plants to refine the inputs to NEMDE, including an optional estimated power capability of the wind power plant: a SCADA signal from the SCADA system using vendor specific variants of a “capability right now” (also known as Possible Power, Capable Power, Park Potential Power or Active Power Available). The challenge with using a ‘now’ value is that the dispatch algorithm is working towards a power system security solution for 5 minutes into the future.

As NEM connected wind power plants progressively work towards implementing FCAS, the criticality of ensuring that the power system takes account of the variability in the wind forecasts coming from the wind power plants in the coming 5-7 minutes becomes even more important, especially for market and power system operators.

I. FORECASTING APPROACHES

There are many different wind forecasting techniques that have been used for predicting the output of wind power plants and the nature of the approaches used is determined by the forecasting horizons required and the time scales being studied. A detailed review of published wind power and speed forecasting techniques can be found in [3] and classifies the published forecasting approaches based on their intended time horizons and application.

TABLE I. TIME-SCALE CLASSIFICATION FOR FORECASTS[1]

Time horizon	***Range***	***Applications***
Very short-term	Few seconds to 30 minutes ahead	Electricity market clearing Regulation actions
Short-term	30 minutes to 6 hours ahead	Economic load dispatch planning Load increment/decrement decisions
Medium-term	6 hours to 1 day ahead	Generator online/offline decisions Operational security in day ahead markets
Long-term	1 day to 1 week or more ahead	Unit commitment decisions Reserve requirement decisions Maintenance scheduling

A more recent study [4] also includes a summary of the recent advances for wind forecasting and power prediction including global to local scales, ensemble forecasting and upscaling and downscaling processes.

Based upon the time scale of the forecasting required, the most common forecasting techniques developed and reported can be classified as follows.

A. Persistence Method

The "Naive Predictor" is used as the benchmark for many forecasting techniques and simply uses the wind speed or generation at time t as the forecast for a later time [5].

$$\hat{P}_P(t+k|t) = P(t) \quad (1)$$

The equation specifies for the persistence forecast that the power forecast for time $t + k$ made at the time origin t is the measured power at time t. Many papers have made the point that for very short to short time scales, the persistence forecast performs better than many forecasting techniques and is very difficult to improve upon [4][5].

B. Physical approach

Physical systems are based on modelling the physical atmosphere and environmental conditions of the wind power plant and often use weather services on a course grid adapted to the location and topology of the site. Complex mathematical models used by Numerical Weather Prediction (NWP) approaches are provided by weather services to produce forecasts based on temperature, pressure, surface conditions and roughness and are run a few times over a single day for timescales up to 7 days and are therefore only suited to long time-scale forecasting.

C. Statistical approach

Statistical wind speed or generation forecasts are based on the approach of training a model using measured data and using the difference between predicted and actual results to develop and tune the numerical model. Statistical forecasting methods have proven to be useful over many forecasting time scales and use a variety of machine learning techniques such as Auto-Regressive Moving Average (ARMA) models, Artificial Neural Networks (ANN), Bayesian Learning, Support Vector Machines and Stochastic Gradient Boosting (SGB) to produce wind speed and generation forecasts.

D. Hybrid approaches

Many of the statistical and NWP models can be combined or multiple NWP models based on spatial correlation can be consolidated to produce forecasts that improve upon the use of a single technique and have shown some promise for the medium and long term wind speed and power forecasts.

For the study described in this paper, wind power generation forecasts are required for time scales that are usually classified as “very short” time-scales by most review papers [4][5] but referred to in this paper as "short" time-scales. Therefore, the focus of the study has been on the use of statistical techniques to model the demanding requirement of producing accurate 5 minute ahead forecasts of wind power generation.

The forecasting of the short time-scale wind generation for the integration of wind power into the Australian NEM was described in the early stages of the market in a series of papers describing the modelling of Tasmanian wind power plants in advance of Tasmania being connected to the Australian mainland NEM market with the Basslink DC inter-connector using a combination approach of ANN models and fuzzy inference systems [6][7][8]. These papers correctly anticipated the requirement for the NEM to be able to accurately forecast wind power plants with the anticipated increased participation of renewable wind energy, the need for reliable dispatch of generators and the requirements for the provision of ancillary services.

E. Machine learning aproaches

Modern machine learning techniques and the availability of cheap fast computing resources have presented the opportunity to significantly improve the forecasting capabilities of wind power plants and provide reliable short term forecasts that are a substantial improvement of simple persistence forecasts that remain the benchmark for short time-scale generation forecasting techniques.

Many short-term modelling approaches have been used for the wind generation forecasting problem. Wake effects and physical configuration of offshore wind power plants was presented in [9] where possible power forecasts were made using SCADA plant measurements and numerical wake models that focused on the interaction effects between turbines that are common to many wind power plants locations.

Other machine learning approaches for short term power forecasting have included Support Vector Machines [10], neural networks and adaptive Bayesian learning [11] and k-means clustering with ANN models [12]. The modelling approach in this paper is based on an initial experimentation process where a wide variety of machine learning approaches were evaluated and an approach formulated that provided the best combination of predictive power and computational efficiency and performance to produce a robust and practical short time-scale wind power generation forecast for a wide variety of conditions.

A standardized performance evaluation of short term wind power prediction models has been proposed by [5] and shall be used for the operational framework for the presented case study and the evaluation and comparison of the model results.

Given that the focus of this research is to find a reliable and effective forecasting approach over a single time horizon of 5 minutes, many of the lead time measures proposed in [5] are not appropriate for this study.

F. Evaluation and comparison of forecasting models

The main measure that will be used for the forecast model development, evaluations and comparisons will be based on the Root Mean Squared Error (*RMSE*) value defined as:

$$RMSE(k) = \sqrt{\frac{\sum_{t=1}^{N}\left(e(t+k|t)\right)^2}{N-P}} \tag{2}$$

And for model evaluation, the Mean Absolute Error (*MAE*) shall also be used:

$$MAE(k) = \frac{1}{N}\sum_{t=1}^{N}|e(t+k|t)| \tag{3}$$

Statistically the *MAE(k)* is associated with the first moment of the prediction error that are directly related to the produced energy and the *RMSE(k)* values are associated with the second order moment relating to the variance of the prediction error. All error measures are calculated using the prediction error $e(t+k|t)$.

For comparisons between models to highlight any relative improvement of one technique or set of parameters, a relative gain will be used to measure the improvement with respect to the reference model that will be the persistence model for all studies:

$$Imp_{ref,EC}(k) = \frac{EC_{ref}(k) - EC(k)}{EC_{ref}(k)} \tag{4}$$

Here the EC is the Evaluation Criterion such as MAE or RMSE for the reference persistence model and for the model being considered.

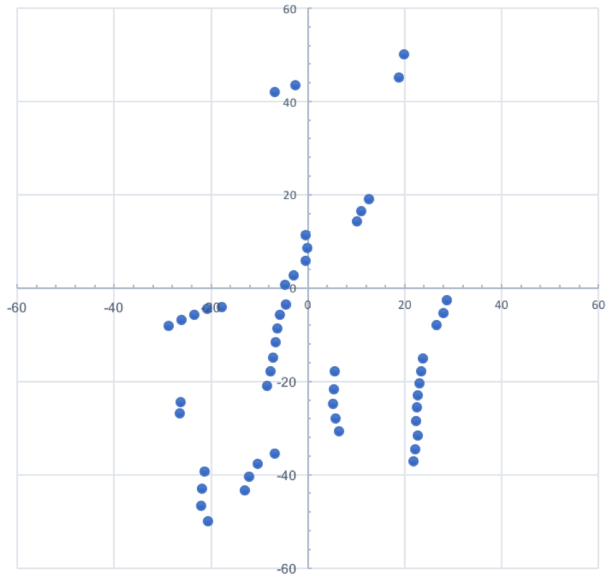

Figure 1 – Normalized turbine layout of studied wind power plant

II. Characteristics of the Studied Wind Plant

The studied wind power plant is a wind farm located in South East New South Wales, Australia with 51 turbines incorporating a mixture of Vestas V90 and V100 type 1.8 MW, 2 MW and 3 MW turbines. Total generation capacity for the wind power plant is 107 MW. The normalized turbine layout where the dominant axis has been mapped to a 100 unit nominal length is shown in Fig. 1.

The wind power plant provided very detailed data collected at a polling frequency of one minute, including a substantial range of turbine measurements and possible power estimate for the entire plant and was selected for the initial study to ensure a wide range of data could be considered.

The turbine locations for the wind power plant are a mix of hill-top, increasing gradient as well as flatter landscape profiles with a vertical distance between the lowest and highest wind turbines of nearly 90 m at approximately 960 m above sea level. The variation in site topology and turbine types presents challenging operating conditions and an even more challenging forecasting environment.

III. Forecasting Approach

A. Exploratory analysis

The initial investigation explored the nature of the very short-term forecasting problem from the perspective of using SCADA data that was easily available from common wind power plant control systems. Typically, the minimum data that is available from the control system are the meteorological measurements at key positions on the power plant site, individual turbine generation, turbine availability and total plant generation. Additionally, some power plants also measure wind speed and direction at the turbine nacelle, nacelle direction and calculate a possible power measurement for the current plant state.

The Weka workbench [13] was used extensively in the initial stages to explore the data and identify the best predictive data fields as well as exploring many different machine learning techniques to identify those machine learning techniques that best predicted the 5-minute generation forecast.

In the initial investigation phase, it became apparent that there was a very weak correlation between any of the individual available variables from the power plant control system to the forecast 5-minute generation value and that the very short term forecasting task is therefore a very demanding problem. There were no stand out predictive variables, although a useful conclusion from the initial analysis was that delta values of wind and generation, that calculated the difference of these variables between one minute samples, did add to the predictive capabilities of the models.

Simple regression models were evaluated by random sampling and comparing the results to independent validation datasets using linear regression, Support Vector Regression, k-Nearest Neighbors, Stochastic Gradient Boosting and ANN with the most promising results being produced by the ANN models and the SGB model with most techniques producing RMSE error values of around 4 MW that were greater than the persistence benchmark forecast RMSE at this stage of the investigation.

B. Single generator forecasting model

A single model was then developed for the entire wind power plant using an ANN model for the wind power plant using many different neural network topologies with wide and deep models being explored. Each of the turbine generations were included in the input variables, MET mast measurements and overall plant performance. A filtering step was introduced to remove input readings that showed signs of SCADA failure of any significant inputs, or were outside of the wind power plant minimum and maximum speeds.

The ANN results produced were at about the same or up to 2% to 3% improvement in RMSE and MAE values in comparison to the persistence model so some progress was made but no substantial improvements were made for any of the results produced after experimenting with a wide variety of ANN network topologies and parameters.

Another model using Stochastic Gradient Boosting [14] was implemented for the same filtered input data as the ANN model, as this machine learning technique also showed promise in the preliminary investigation. The advantage that SGB models have in comparison to ANN models is the reduced training time for developing and training the supervised learning model.

The results produced by the stochastic gradient boosting model were slightly better than the ANN models with significantly reduced model training times.

Neither the ANN or the stochastic gradient boosting models produced results that were substantially better than the persistence forecast and so on this basis it was decided to investigate the performance of a clustering design that was representative of the locations and topology of the turbines and therefore could capture more local generation conditions than a single wind power plant model.

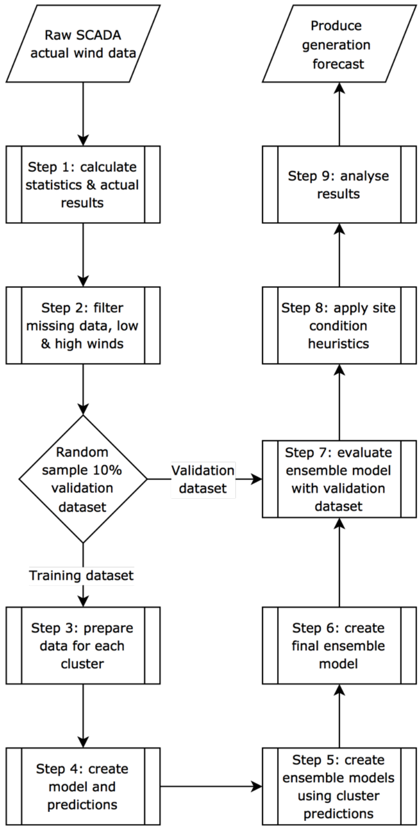

Figure 2 - Procedure for the short-term generation model.

C. Turbine cluster & ensemble generation forecast model

Previous studies that used a turbine clustering approach for the short-term forecasting of wind power plant generation used unsupervised statistical techniques such as k-means clustering that groups a set of observations around observed initial cluster centres [12] and supervised learning techniques, such as the k-nearest neighbors technique, that was considered in the initial exploratory investigation.

The approach used for this study clustered static locations of the turbines, the local topology and the dynamic environmental conditions. Each of the turbines are assigned to a static cluster for a set of conditions and then an overall ensemble model was used by aggregating the predicted generation cluster forecasts to then produce a forecast for the entire wind power plant.

For this study, stochastic gradient boosting (SGB) models were used for the individual cluster models and both stochastic gradient boosting and ANN models were used for the ensemble models, thus producing a forecast for the entire wind power plant over a very large range of differing input environmental conditions.

IV. Forecasting Results

Clustering the turbines into groups and running ensemble models to produce the power plant forecast had a very substantial effect on the model training times and computational resources required to train the model. For all the initial investigation and the single generator machine learning models it was feasible to train and run the

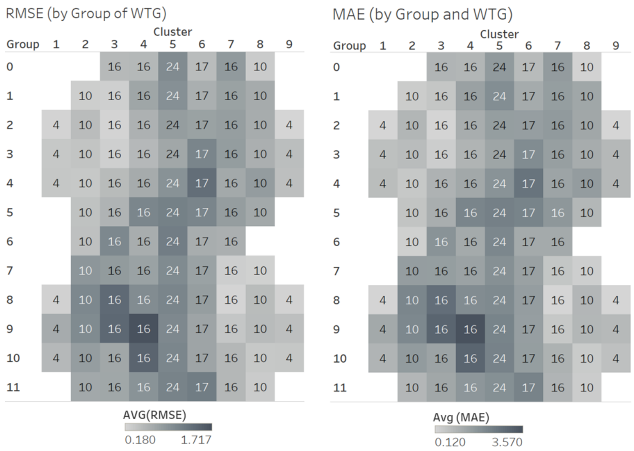

Figure 3 - RMSE and MAE error distribution

evaluations on a single MacBook Pro laptop but the learning times for the clustered model took too long to complete for this approach to be feasible for the clustered models.

A Compute Optimized 32 core 60 GB memory AWS C3.8xlarge instance running Ubuntu 16.04 was used to develop the clustering and ensemble turbine models for the studied wind power plant. Compute optimized instances have a higher ratio of vCPUs to memory than other families and are suitable for high performance science and engineering applications. Typical run times for each training and evaluation run for a turbine cluster model was about 24 hours on this optimized AWS instance.

Once the model has been trained and developed, evaluation is a very quick operation making both the stochastic gradient boost and ANN models ideally suited for production deployment.

For the cluster model, 314,357 observations were used with the input data being randomly split into 90% of the data forming the training set with 285,779 observations and an independent validation data set comprising 28578 records. No part of the validation set was used in the training of the turbine cluster or ensemble model.

A k-fold cross-validation approach was used to train the models with 10 splits on the training data and a negative mean squared loss function was used for the model training.

Various combinations of input fields were used for the stochastic gradient boosting model for both the individual turbine clusters and the ensemble model. The SGB model can display the feature importance of each of the input fields and so by determining the contribution each field makes to the forecast, the models can be improved.

There are many different parameters that are available to be tuned for the SGB model such as the learning rate, depth of trees, number of trees and the minimum child weight and the results were relatively insensitive to most parameter changes and continued to find consistent optimal values even when large ranges of parameters were tested. The number of trees parameter was the exception where large numbers of trees would give low training evaluation scores but perform worse on the independent validation dataset indicating that the model was over-fitted to the training data.

The turbine cluster and ensemble model using the SGB model for both the clusters and the ensemble produced a range of RMSE improvements in comparison to the persistence forecast for this preliminary model of **7.2%** to **10.7%** and for the MAE **9.1%** to **10.4%**. Considering the early stages of development of the forecasting model and the recent completion of the clustering implementation, the observed results are very encouraging and certainly much better than any of the single generation forecast wind power plant models.

The distribution of errors over the cluster groups for the SGB cluster and ensemble models can be seen in the heat map (Fig. 3) for both RMSE and MAE where the number of turbines in the cluster are indicated in each cell of the heat map. It can be seen from the graph that the internal clusters of the wind power plant have the largest errors and that work is required in the clustering procedure to try to improve the cluster modelling in these areas.

Lastly an ANN ensemble model was run on the ensemble data using the SGB model for the turbine clusters without any refinement or development just prior to this paper being completed with resulting improvements in RMSE compared to the persistence forecast of **8.8%** and MAE of **7.5%**. Given that no tuning or development of the ANN model has been possible, it may prove that ANN is a more effective means of producing the final ensemble forecast rather than the stochastic gradient model.

V. Production forecast implementation

The current study has been undertaken with the ultimate objective of implementing the developed and refined forecasting model into a production device that can be implemented within the wind power plant SCADA system.

The production forecast implementation is required to operate under a wide range of input conditions and must be robust and reliable. If the inputs to the forecast model do not meet the model validation requirements, the forecast model will revert to using the persistent forecast value rather than risk the production of an erroneous forecast value and produce poor outcomes (Fig. 4).

Individual wind power plant locations, site details and production data will be used to create a dedicated set of models for the generator. Once a suitably trained set of models has been developed for a specific site, the implementation of the new forecasting technique will be directed toward replacing the existing Australian NEM intermittent generator AWEFS dispatch forecast with the revised, trained estimation of the likely generation in the coming 5-7 minute dispatch period.

The implementation of the forecasting device will have the following effects for both the generator and the system operator:

- Provide a more accurate forecast for the wind power plant to reduce its deviation from the dispatch target calculated by AEMO and thereby reducing any ‘causer-pays’ penalties [15];
- allow the wind generator to potentially participate in raise FCAS contingency and regulation markets by providing accurate, unconstrained generation forecasts; and
- by contributing to the reduction in the uncertainty of individual wind generators, provide more accurate regulation requirement estimations for AEMO to

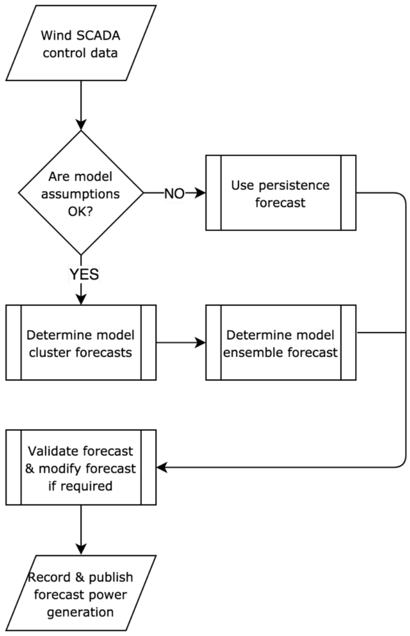

Figure 4 - Procedure for producing generation forecasts

ensure appropriate control of secondary-order primary frequency control.

The device is being designed as a dedicated server that will have a direct connection to the SCADA system through the modbus, DNP3 or possibly IEC 61400-25 protocols and can provide a web service interface, write directly to the SCADA and database for later validation and analysis. The device will be supported remotely and will be updated at regular intervals with model updates and refinements.

CONCLUSIONS AND FUTURE WORK

This paper presents the encouraging preliminary results of the development of a dedicated machine learning device for the 5-7 minute forecasting of wind power plants. Further development of the model is required to improve the clustering, tune the machine learning algorithms and demonstrate the capabilities and robustness of the forecasting approach over a sustained period of operation and for a larger set of diverse wind power plants.

ACKNOWLEDGMENT

The authors would like to thank Pacific Hydro Australia for permission to publish the forecasting results for their wind power plant. Also, this research would not be possible without the availability of the many excellent open source projects such Python, Pandas, Scikit-Learn, Keras, Theano, NumPy and Weka, that were used in this study.

REFERENCES

[1] Jennings, R., Dyson, J. and Summers, K., 2016, November. The maturing of wind integration in Australia: Improvements required in market operation for more consistent, economic and efficient dispatch outcomes. 15th Wind Integration Forum, Vienna.

[2] http://www.aemo.com.au/Stakeholder-Consultation/Consultations/AWEFS-and-ASEFS-Stakeholder-Consultation, accessed 27 July 2017.

[3] Soman, S.S., Zareipour, H., Malik, O. and Mandal, P., 2010, September. A review of wind power and wind speed forecasting methods with different time horizons. In North American Power Symposium (NAPS), 2010 (pp. 1-8). IEEE.

[4] Foley, A.M., Leahy, P.G., Marvuglia, A. and McKeogh, E.J., 2012. Current methods and advances in forecasting of wind power generation. *Renewable Energy*, *37*(1), pp.1-8.

[5] Madsen, H., Pinson, P., Kariniotakis, G., Nielsen, H.A. and Nielsen, T.S., 2005. Standardizing the performance evaluation of short-term wind power prediction models. *Wind Engineering*, *29*(6), pp.475-489.

[6] Potter, C., Ringrose, M. and Negnevitsky, M., 2004, September. Short-term wind forecasting techniques for power generation. In *Australasian Universities Power Engineering Conference (AUPEC 2004)* (pp. 26-29).

[7] Potter, C.W. and Negnevitsky, M., 2006. Very short-term wind forecasting for Tasmanian power generation. *IEEE Transactions on Power Systems*, 21(2), pp.965-972.

[8] Negnevitsky, M., Johnson, P.L. and Santoso, S., 2007. Short term wind power forecasting using hybrid intelligent systems.

[9] Gocmen, T., Giebel, G., Rethore, P.E. and Leon, J.P.M., 2016, November. Uncertainty quantification of the real-time reserves for offshore wind power plants. 15th Wind Integration Forum, Vienna.

[10] Zhou, J., Shi, J. and Li, G., 2011. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Conversion and Management*, *52*(4), pp.1990-1998.

[11] Blonbou, R., 2011. Very short-term wind power forecasting with neural networks and adaptive Bayesian learning. *Renewable Energy*, *36*(3), pp.1118-1124.

[12] Kusiak, A. and Li, W., 2010. Short-term prediction of wind power with a clustering approach. *Renewable Energy*, *35*(10), pp.2362-2369.

[13] Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J., 2017. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

[14] Friedman, J.H., 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), pp.367-378.

[15] Systems Capability, Causer Pays: Procedure for Determining Contribution Factors, AEMO, Tech. Rep., 2013.

BIOGRAPHICAL INFORMATION

Harley J. Mackenzie was awarded a PhD. in applied mathematics on the subject of numerical modelling of combustion processes from Monash University in Clayton, Australia in 1990 and has a Degree with Distinction in mechanical engineering from the Swinburne Institute of Technology in Hawthorn, Australia in 1984.

He has been the Managing Director for the last 20 years of HARD Software that develops technical tools and solutions for trading, generation and retail operations in the Australian NEM and International energy markets as well as providing consulting services in energy, risk management and aviation. His current passions are focussed on reducing the costs of FCAS regulation and contingency services for intermittent generators, developing effective wind and solar software tools and optimisations and developing short term wind and solar forecasting solutions. He is also a director of a joint venture with Jonathon Dyson of Dispatch Solutions, that was recently established to provide technical dispatch solutions for power system operators and generators.

Dr. Mackenzie is a member of Open Source Industry of Australia and is a contributor to the GLPK open source linear program application.

Jonathon Dyson holds a double degree in engineering and business from Monash University in Clayton, Australia in 1998.

He is a Director of a specialised market consulting firm with a specific focus in plant and trading operations and optimisation. With 18 years experience in the NEM, Jonathon has worked for numerous market participants, regulators and policy/observer organisations across all regions in Australia. He is also a director of a joint venture with Harley Mackenzie of Dispatch Solutions, that was recently established to provide technical dispatch solutions for power system operators and generators.

Mr. Dyson is an associate of the Australian Institute of Company Directors and is a member of IEEE, CIGRE and UV