ENHANCED TURBINE PERFORMANCE MONITORING

The authors of this article clearly demonstrate the benefits of applying data-mining techniques to the prediction of fault modes in wind turbines.

By Andrew Kusiak and Anopp Verma

Andrew Kusiak is faculty and Anopp Verma is a graduate student in the Department of Mechanical and Industrial Engineering at The University of Iowa. Kusiak can be reached at andrew-kusiak@uiowa.edu. Also visit www.uiowa.edu.

COMPONENTS OF WIND TURBINES are affected by asymmetric loads, variable wind speeds, and severe weather conditions which cause wind turbines to change their states. A typical wind turbine undergoes various states during its daily operations. The wind turbine states follow a certain pattern, such as: 1) turbine OK to run-up idling; 2) turbine online to maintenance/repair mode; 3) turbine weather conditions to external stop, and; 4) turbine OK to fault mode, and so on. A state change from normal turbine operations to a fault mode adversely impacts the performance of wind turbines and its components. Monitoring these states can greatly augment the maintenance operations.

Wind turbine monitoring can be done in two ways: condition-based, and performance-based. Condition monitoring requires installation of additional equipment/sensors to continuously monitor relevant parameters in real time. Performance monitoring utilizes historical data for prediction of turbine performance. Since performance monitoring relies on the existing data, there is no additional cost to the wind farm operators. Analyzing the historical data through data-mining techniques is a promising approach to

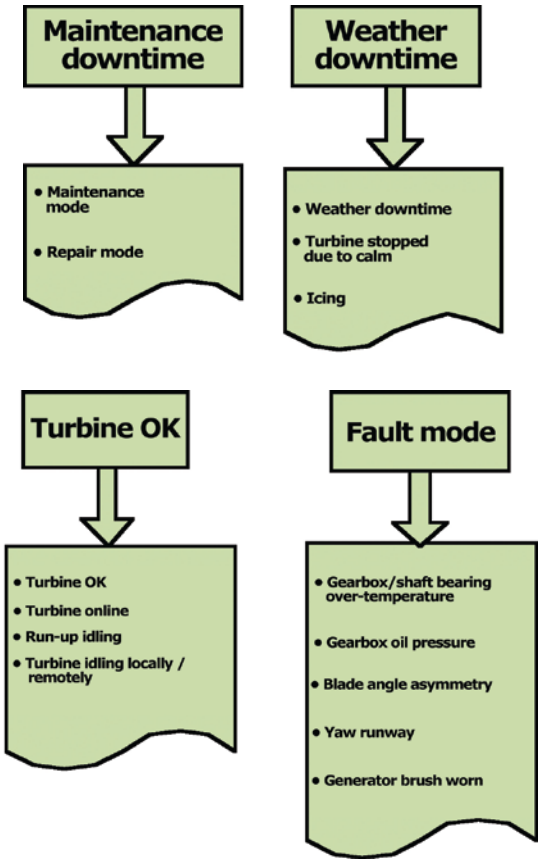


Fig. 1: Categories of states of a wind turbine.

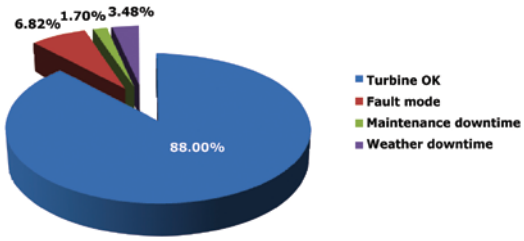


Fig. 2: Distribution of wind turbine states.

performance monitoring. The current SCADA systems already record wind turbine parameters. An improved SCADA system coupled with data-mining algorithms can be useful in identifying critical performance indicators of wind turbines. This article demonstrates application of data-mining techniques to prediction of fault modes of a wind turbine.

TURBINE STATES INFORMATION

In addition to wind turbine parameters, a SCADA system records the states of wind turbines. Consider 17 possible states in which a turbine can be found. The state "turbine in fault mode" (state No. 17 in Table 1) can usually be expressed in more than 400 ways. Components of a wind turbine such as wind turbine blades (e.g., blade angle asymmetry), turbine yaw (e.g., yaw runaway), wind turbine generator (e.g., generator brush worn), and wind turbine gearbox (e.g., gearbox over-temperature) can be affected. Using domain knowledge, the states listed in Table 1 can be broadly categorized into four main groups (fig. 1). At the top level, prediction of fault modes of wind turbine is crucial in identifying an actual fault.

Figure 2 displays the distribution of wind turbines states. The values are averaged for a wind farm over a year period. Almost 90 percent of the time the turbine is operating normally. However, faults make up 6.82 per-

cent of the total time. Any effort to minimize the frequency of faults would enhance wind turbine availability.

DEVELOPING PREDICTION MODELS

In this section, data-mining models are developed for predicting turbine states. The data collect by the SCADA system is sampled at the frequency of 0.1 Hz. The faults that have occurred at 17 wind turbines in a three-month period are plotted in fig. 3. Based on the frequency of fault modes, three wind turbines are identified (labeled as A, B, and C in fig. 3). Data from turbine A is used for training and testing data-mining algorithms, whereas turbine B and C are used for performance evaluation.

PARAMETERS FOR STATES PREDICTION

A SCADA system records various parameters, which can be categorized into: 1) non-controllable parameters such as wind speed and wind deviations; 2) performance parameters such as power output and rotor speed; 3) vibration parameters such as tower acceleration and drive train acceleration,

State Number	State Description	State Number	State Description
1	Turbine OK with no errors	10	Turbine stopped locally
2	Turbine running smoothly	11	Emergency stop
3	Turbine running up idling for cut in	12	Turbine stopped due to curtailment
4	Turbine in maintenance mode	13	Turbine stopped by customer
5	Turbine in repair mode	14	Turbine idling locally
6	Power failure/grid downtime	15	Turbine idling remotely
7	Weather downtime	16	Wind direction curtailment
8	Turbine stopped externally	17*	Turbine in fault mode
9	Turbine stopped locally		
			* Primary focus

Table 1: Turbines states information.

No.	WGS	WBFS	BTA
1	Nacelle revolution	Blade 3 pitch angle(actual)	Blade2 pitch angle (actual)
2	Blade 3 pitch angle (actual)	Current phase C	Blade3 pitch angle (actual)
3	Current Phase B	Temperature hub	Blade1 pitch angle (actual)
4	Nacelle Position	Temp. control box axis 1	Generator/gearbox speed
5	Generator/gearbox speed	Voltage phase C	Generator speed
6	Temperature, bearing B	Generator speed	Rotor speed
7	Temperature top box (°C)	Drive train acceleration	Blade2 pitch angle (set)
8	Power (Actual)	Temperature top box	Blade3 pitch angle (set)
9	Tower deflection	Nacelle revolution	Blade1pitch angle (set)
10	Wind deviation, 1 sec	Temperature bearing A	Drive train acceleration

Table 2: Parameters selected with data-mining algorithms.

Table 3: Prediction accuracy of turbine states (turbine A).

Time stamp [s]	Turbine state				Overall Accuracy [%]
	Turbine OK [%]	Fault mode [%]	Maintenance downtime [%]	Weather downtime [%]	
t	99.88	99.67	78.41	97.91	99.45
t + 10	99.56	99.00	77.22	95.04	98.39
t + 30	97.64	96.41	74.59	94.62	96.54
t + 60	95.70	95.64	71.67	92.55	94.43
t + 120	91.87	90.00	67.49	88.47	90.89
t + 180	88.58	87.34	64.94	84.43	86.82
t + 240	85.62	84.64	60.31	82.44	83.93
t + 300	83.05	82.76	59.67	80.39	81.76

and; 4) temperature parameters such as gearbox temperature and generator temperature. It is important to mention that not all the parameters are causing wind turbines to change its states. Therefore, we need to identify a set of parameters that impact the wind turbine states. This is the place where algorithms build using data-mining techniques are useful. Data-mining algorithms use statistical measure such as information gain and correlation coefficient to identify relevant set of parameters. Table 2 shows 10 parameters relevant to wind turbines states (data from turbine A), identified by three different data-mining algorithms. The techniques used are wrapper with genetic search (WGS), wrapper with best first search (WBFS), and boosting tree algorithms (BTA). For detailed description of these techniques refer to [1-3].

DATA-MINING MODELS

A combination of relevant input parameters (found in previous section) can be used to develop prediction models for state faults.

Actual state	Anticipated state	Correctly identified cases
Blade angle not plausible axis 2	Fault	76.92%
Gearbox oil pressure too low	Fault	0.00%
Maintenance downtime	Maintenance downtime	100%
Motor protection	Fault	100%
No activity CAN-Bus CCU	Fault	50.0%
Overproduction	Fault	100%
Pitch control deviation axis 3	Fault	100%
Safety chain	Fault	100%
Turbine OK	Turbine OK	99.45%
Weather downtime	Weather downtime	60.70%
Yaw runaway	Fault	87.50%

Table 4: Analysis on turbine B.

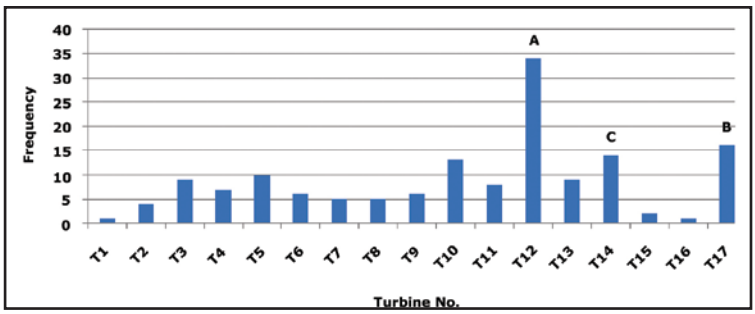


Fig. 3: Frequency of fault modes of wind turbines.

The evaluation of data-mining algorithms is based on the prediction accuracy of turbine states namely turbine OK (TO), weather downtime (WD), maintenance downtime (MD), and fault mode (FM) (fig. 4). The diagonal elements (e.g.

Actual \ Predicted	Turbine OK	Weather downtime	Maint. downtime	Fault mode
Turbine OK	TP_{TO}	$FP_{TO,WD}$	$FP_{TO,MD}$	$FP_{TO,FM}$
Weather downtime	$FP_{WD,TO}$	TP_{WD}	$FP_{WD,MD}$	$FP_{WD,FM}$
Maint. downtime	$FP_{MD,TO}$	$FP_{MD,WD}$	TP_{MD}	$FP_{MD,FM}$
Fault mode	$FP_{FM,TO}$	$FP_{FM,WD}$	$FP_{FM,MD}$	TP_{FM}

Fig. 4: Accuracy assessment of wind turbine states.

Actual state	Anticipated state	Correctly identified cases
Centrifugal switch	Fault	100%
Gearbox oil over-temperature	Fault	0.00%
Maintenance downtime	Maintenance downtime	100%
Pitch overrun 00	Fault	100%
Power failure	Weather downtime	60.00%
Turbine OK	Turbine OK	99.27%
Weather downtime	Weather downtime	95.54%
Yaw runaway	Fault	100%

Table 5: Analysis on turbine C.

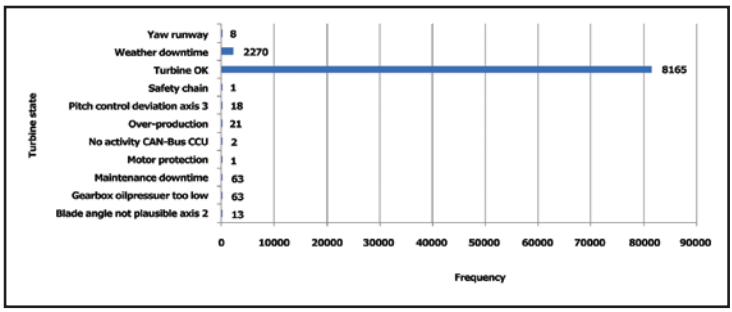
TPTO, TPWD) are correctly predicted turbine states, whereas non-diagonal elements are wrongly predicted (e.g., FPFM, MD: turbine in fault mode is predicted as turbine in weather downtime).

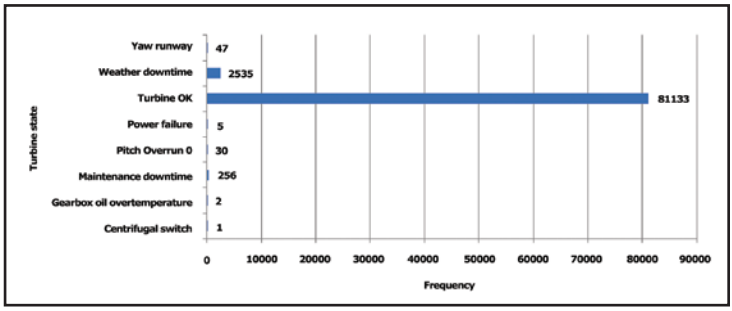
Data-mining algorithms namely random forest algorithm (RFA) [4] is used to build eight prediction models at various time stamps, with the maximum prediction length of 5 min. The selection of data-mining algorithm is based upon their performance on training data at time stamp t. The accuracy was found to be in the range 81-99 percent for all turbine states (Table 3).

ROBUSTNESS OF DATA-MINING MODELS

In order to check the robustness of developed models, various unseen faults were tested. The aim is to check the response of the model when unobserved faults are encountered. Figure 5 (a-b) displays the distribution states for turbines B and C plotted over a three-month period. The number of faults varies across turbines, however, turbines are found to be operational most of the time. The results shown in Tables 4 and 5 illustrate the response of data-mining algorithms on turbine A and B. Except of the gearbox-related faults, other fault modes of wind turbines are correctly identified.

Fig. 5: States distribution of wind turbines: (a) Turbine B, (b) Turbine C.





CONCLUSION

The parameters recorded by SCADA systems can be useful for monitoring purposes. Using prediction models derived by data-mining algorithms the states of wind turbines can be predicted ahead of the time, which can be helpful in maintenance planning. The models built using data-mining algorithms can be integrated with current SCADA system to enhance performance monitoring of wind turbines. Acknowledgement: The research reported in the paper has been supported by funding from the Iowa Energy Center, Grant 07-01.

REFERENCES:

1) Kohavi, R. and John, G.H., 1997, "Wrappers for feature subset selection," Artificial Intelligence, 97(1-2), pp. 273-324.
2) Sbihi, A., 2007, "A best first search exact algorithm for the Multiple-choice Multidimensional Knapsack Problem," Journal of Combinatorial Optimization, 13, pp. 337-351.
3) Kudo, T. and Matsumoto, Y., 2004, "A Boosting Algorithm for Classification of Semi-Structured Text," EMNLP.
4) Liwa, A. and Wiener, M., 2002, "Classification and regression by random forest," R News, 2(3), pp. 18-22.