Multi-turbine Associative Model for Wind Turbine Performance Monitoring

Onder Uluyol[1] and Girija Parthasarathy[2]

[1,2]*Honeywell International, Inc., Golden Valley, MN, 55422, USA*

onder.uluyol@honeywell.com
girija.parthasarathy@honeywell.com

ABSTRACT

Comparing the performance parameters of a set of wind turbines in a single region will provide insights that prevent raising unnecessary alarm, while confirming actual faults. Wind turbines operating in a wind farm experience similar operating and environmental conditions that could indicate either normality for that group or failures that manifest in those conditions. These norms and failures are an orthogonal set of data rich in information that can be utilized in performance monitoring algorithms to supply better prediction accuracy and low false positives. In this paper, we describe the use of an associative model (AM) for fault detection in a population. An associative model maps system parameters to an identical set of virtual parameters. The AM-based approach can be used to capture the underlying correlation of an observable system, such as performance parameters of a set of wind turbines in a wind farm. The residuals between the model output and the input can then be used to detect anomalies and isolate faults.

1. INTRODUCTION

A modern wind farm can hold hundreds of wind turbines at a site, located as close as several hundreds of meters from each other. They are usually situated in remote locations and operate under severe environments atop 60-90 m towers. It is no surprise that operations and maintenance costs of wind turbines run high. One of the biggest drivers of maintenance cost is unscheduled maintenance due to unexpected failures. Continuous performance monitoring of wind turbine health for automated failure detection can reduce maintenance costs by detecting failure pre-cursors before they reach a catastrophic stage and by keeping turbines operating at higher efficiencies.

Typical performance or health monitoring of a machine involves monitoring a few performance parameters over time to see changes with respect to its own history or manufacturer provided baseline. It has been successfully used in the process and aerospace industries [Bell & Foslien, 2005; Gorinevsky, Dittmar & Mylaraswamy, 2002; Kim & Mylaraswamy, 2006]. In the wind industry, fault detection methods include anomaly detection based on neural network models of normal operating modes [Zaher, McArthur, and Infield 2009]; power-curve-based performance monitoring analytic [Uluyol, Parthasarathy, Foslien & Kim, 2011]; data mining techniques based on wind speed and power output [Kusiak, 2011]; and classification methods of clustering and principal components analysis [Kim, Parthasarathy, Uluyol, Foslien, Sheng & Fleming, 2011].

An alternative, complementary practice would be to monitor performance parameters with respect to similar machines in similar operating conditions. A wind farm offers such an opportunity. This opportunity converts the inherent complexity of monitoring a large number of wind turbines individually into an asset that increases the robustness of fault and performance monitoring systems for all wind turbines in a wind farm.

A study by Ye et al. (2010) presented a method for detecting wind turbine anemometer failures based on the difference in measured wind speeds between a pair of closely situated wind turbines. However, in this study, the method is a comparison of only two wind turbines and operates on accumulated week's worth of data. The present work explores the use of an associative model (AM) to capture the relationship among several wind turbines on a farm to detect anomalous conditions. The AM is applied to high resolution time-series data from seven wind turbines.

2. FAULT DETECTION USING ASSOCIATIVE MODELS

An associative model maps system parameters to an identical set of virtual parameters. The AM-based approach can be used to capture the underlying dynamics of an observable system, such as performance parameters of a set of wind turbines in a wind farm. The residuals between the model output and the input can then be used to detect anomalies and isolate faults. See [Uluyol 2001 and Uluyol

2003] for previous applications of this approach for in-range sensor fault detection, isolation and recovery.

When applied to sensor data, AM captures analytical redundancy among sensors and maps the readings from a group of correlated sensors into an estimation set for an identical group. When an appreciable sensor fault is detected, the associated model estimate diverges from the actual sensor reading. Associative models in the form of an auto-associative neural network (AANN) [Kramer 1992] have also been applied for sensor validation in nuclear power plants, chemical process plants, and propulsion turbine engines.

Our application differs from conventional use of neural networks in that we employ a neural network as a model of the system that maintains dependencies among parameters of interest. A fault in our approach is a break in this overall correlation rather than a deviation in an individual parameter. The iterative associative model approach for fault isolation has been demonstrated with actual flight data collected from sensors installed on multiple types of turbine engines for fixed wing aircraft and helicopters.

3. WIND FARM APPLICATION

When applied to a wind farm, AM captures the performance correlation among the modeled wind turbines (see Figure 1). In this application of AM for wind farms, we explore the use of wind turbine power produced as the performance parameter. Many factors affect wind turbine performance. In applying an associative model, the analysis should include careful filtering of the data, both in training the model and in deploying it.

A wind turbine may exhibit anomalous behavior compared to its neighboring wind turbines for a number of reasons. It could be under repair or be subject to curtailment while others are operating normally. Or, the other turbines could be subject to curtailment while one is operating normally. These conditions can be obtained either through status parameters captured as part of SCADA (Supervisory Control And Data Acquisition) data or be detected using filters based on simple statistics on measured power.

The factors that affect the dynamics of wind turbine performance, and hence present an opportunity for the use of the AM approach, include location effects, park-wide control effects, and slowly-progressing fault in one of the monitored wind turbines. In a wind farm, wind turbines can be situated as far as 100 m apart from each other facing the wind direction and as far as 500 m apart along the wind direction. In order for all wind turbines to be under the same wind regime, the wind direction should be stable for a period of time that is greater than the wind turbine response time. While the wind direction is stable, the wind turbines don't necessarily produce exactly the same amount of energy; however, but we can correlate production to their location in the wind farm and capture this correlation in the associative model. Similarly, wind turbines may respond differently but consistently for their locations when the wind farm operator applies park-wide controls. These correlations can also be captured in the associative models.

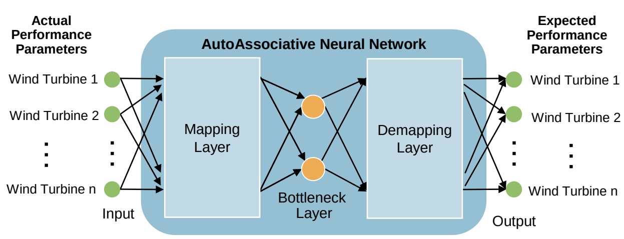

Figure 1. Wind turbine correlation mapping using AANN

After correlations among wind turbines for stable wind regimes and common control operation regimes are established, the associative model can be used to detect any deviation from the expected behavior.

3.1. SCADA Data Description and Access

We obtained archived wind turbine SCADA data from three wind parks. The large data set, maintained in a MatrikonOPC Desktop Historian, includes 16000+ tags, and spans a period of about 6 months.

Before developing the algorithms, we identified the initial parameters of interest. The flat set of 16000+ tags was organized into a metadata list so we could parse the data appropriately. We also established a procedure for accessing the data with MATLAB®, which is our development environment. We used MathWorks® OPC Toolbox™ for this purpose. The tags were classified, and tags belonging to one wind farm and their data were exported to MATLAB®. The tags in this farm looked most meaningful for inclusion in a multi-turbine diagnostics algorithm for tag consistency (i.e., the same tags are available for most wind turbines in the park) and relevancy (i.e., the tags are useful for performance analysis).

3.2. Data Pre-processing and Selection of Wind Turbines

Our objective was to select a representative set of turbines for developing and demonstrating the multiple wind turbine fault diagnoses. The wind park IDs in the tag names of the data were first fixed to eliminate discrepancies and redundant tags. Based on the wind turbines associated with one particular meteorological tower (MET) (502 marked with a red star in Figure 2) and their representative location in the geography around the tower, we chose seven wind turbines for further data analysis. These turbines are numbered T075, T081, T098, T104, T115, T118 and T127. There are about 30-50 wind turbines installed for each MET. The area where the turbines are installed forms a trapezoid shape behind each MET. There were 49 turbines

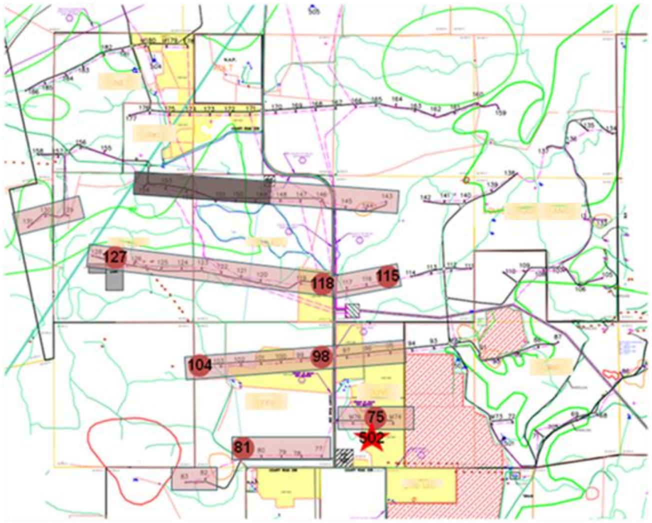

Figure 2. Location of selected wind turbines on the wind farm

installed in four rows within the trapezoid for MET #502. We selected two turbines from the first row, two from the second row, and three turbines from the third row for this analysis. The last row was omitted because of missing data.

The data from these turbines show that the tags have different sampling rates, and recorded parameters are not synchronized with each other. For example, wind turbine power, wind speed, and temperature are recorded at different timestamps and have differing timestamp intervals. The data statistics of wind turbine power produced, wind speed and ambient temperature are shown in Table 1. Notice that while the turbine power and speed may be recorded every second and the ambient temperature every 2 seconds, the median interval varies a lot from 5 seconds for power to 500 seconds for the ambient temperature. This results in having almost 300 times as many data points for power as the ambient temperature, and necessitates re-sampling for synchronizing data.

We reviewed several tags for filtering data to create a baseline data set. These tags included Operating, Faulted, and State_fault, among others. To obtain set interval time-synchronized data, each tag was re-sampled at a higher frequency using a proprietary technique (patent pending), and means were computed for the set interval (1-min, 5-min or 10-min) for periods common to the tags of interest.

3.3. Analysis Using Power Curve Analytic

To detect correlations and variations for the same periods, the synchronized data from the seven selected wind turbines were analyzed according to the power curve analytic (Uluyol, Parthasarathy, Foslien and Kim, 2011). The power curve for the wind turbines was obtained from the operator. For each of the wind turbines, we calculated residuals from the power curve, for all data, with a 30-day moving window

	Power	**Windspeed**	**Ambient temperature**
Min interval (sec)	1	1	2
Median interval (sec)	5	20	500
Mode (sec)	5	5	100
Max interval (hr)	4.33	7.99	8
No of samples	2083426	424568	7331

Table 1. Data Statistics

and 1-day interval or frequency of calculation. The residuals were further processed to obtain wind speed bin-based averages, standard deviations, skewness, and kurtosis. Our analysis reveals that a correlation exists amongst the turbines in terms of power production, and with a suitable multiple turbine algorithm, anomalous conditions and faults may be detected. The plots in Figure 3 and Figure 4 show the skewness and kurtosis statistics. The seven wind turbines are represented in rows, with the columns representing the dates 25 days apart.

We observe that, in the time periods ending on Jan 9 and Feb 3, skewness and kurtosis for each wind turbine have shapes deviating from zero values across the wind speed bins. Since all wind turbines show this behavior for these time periods, and since the skewness is measured with respect to the population mean (rather than the nominal power curve), we speculate that this behavior could be found in normally operated wind turbines at this park. It is also possible that the number of samples in these time periods may have skewed the statistics.

These plot sets in Figure 3 and Figure 4 are an important visualization tool for comparing wind turbines in terms of broad parameter values and spotting anomalies. They provide the means to explain any individual power curve analytic alarms that could have been raised in the vicinity of the dates in the first two columns, when there seems to have been controlled power curtailment.

This part of the analysis leads us to conclude that wind turbine performance is strongly correlated across the wind farm. Additional automated analytics with the associative models approach, which models the correlation at a much finer timescale, could be beneficial.

3.4. Wind Farm Data

A set of normal operating data under nominal conditions is extracted to train an associative model. Normal operating data are the data from periods where no known fault is present in the records. The nominal conditions are defined based on power magnitude and variation, and wind direction variation. Figure 5 shows the power profile of the seven wind turbines during the six-month period between September 2009 and February 2010. Notice that the bulk of the data are recorded between the red vertical lines indicating the 250 kW and 1250 kW levels. The start-up and

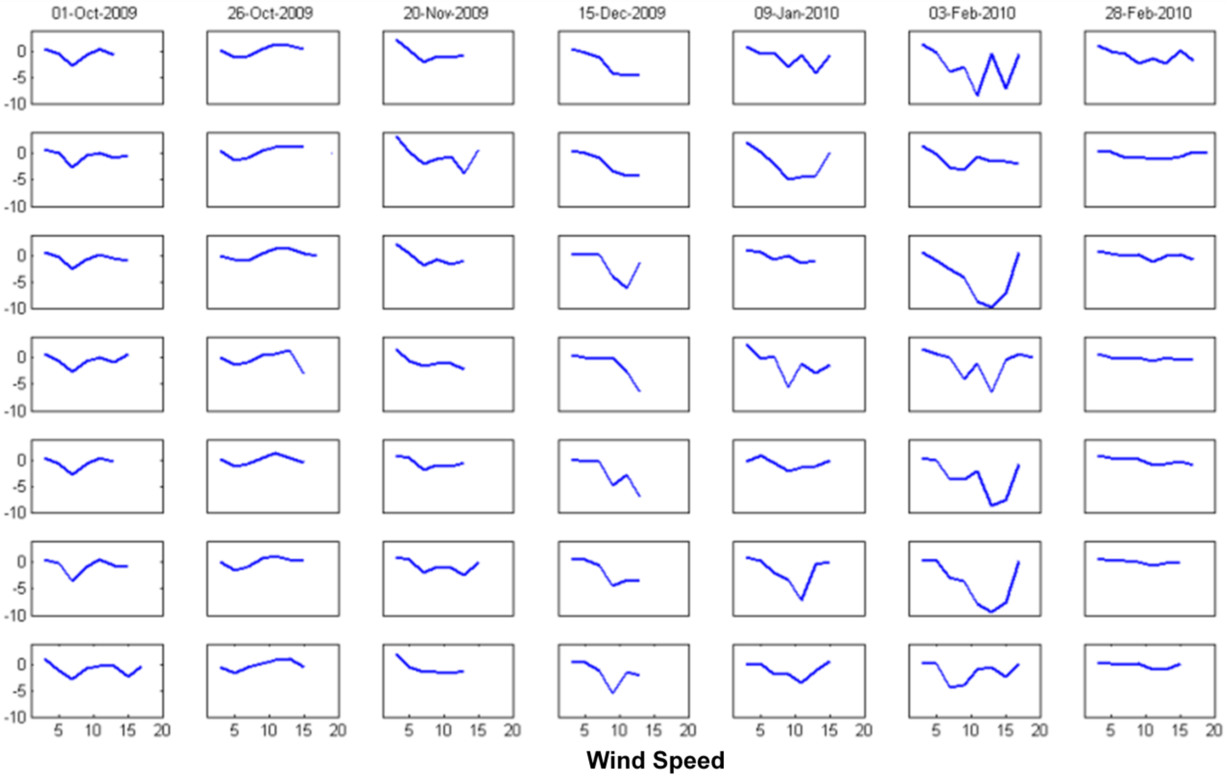

Figure 3: Skewness of power residuals in 30-day windows for 7 wind turbines (shown in rows)

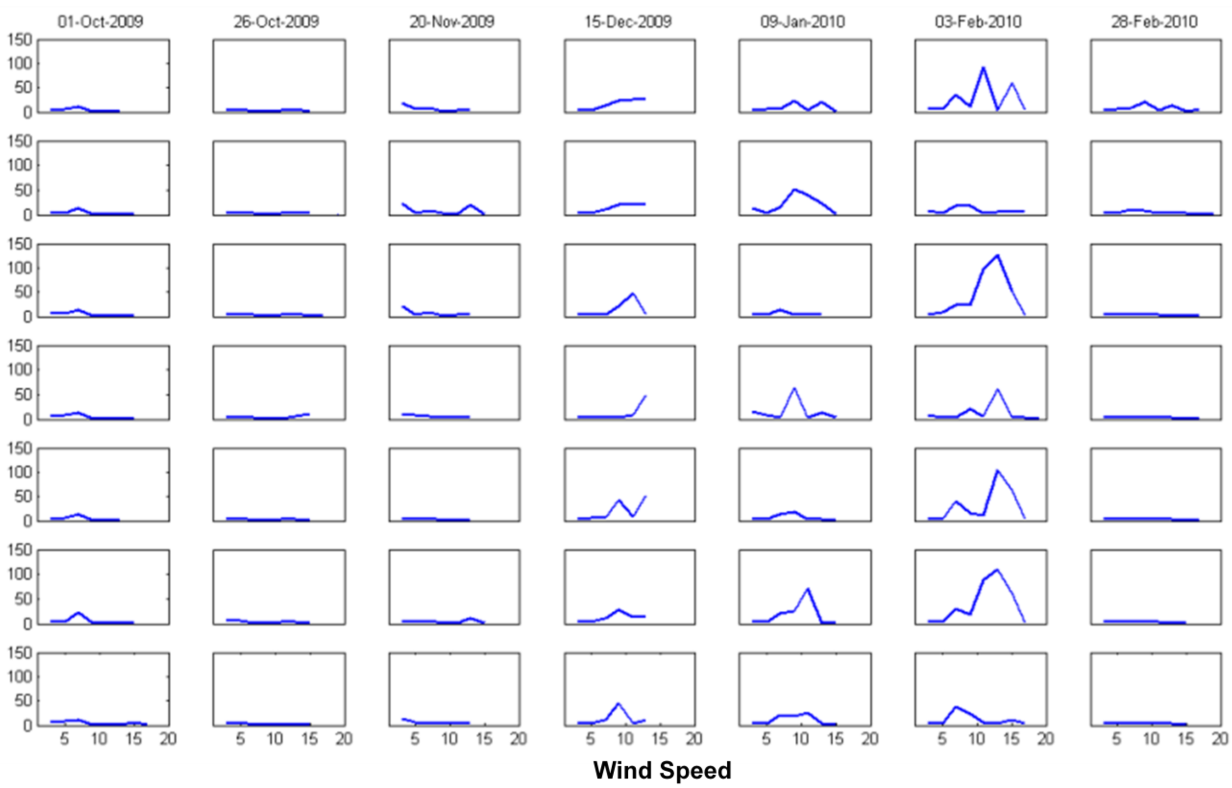

Figure 4: Kurtosis of power residuals in 30-day windows for 7 wind turbines (shown in rows)

the high power regions that fall outside of these limits contain highly non-linear operation periods and are excluded from the training set.

Figure 6 shows the variation in power in terms of the standard deviation calculated for each sampling interval. Notice that it shows a normal distribution superimposed with samples having very little or no variation. The samples that show very little variation in the wind turbine group correspond to shutdown, full saturated power, or curtailed power level conditions. For the baseline training, we exclude these points. The samples at the other end of the distribution mark very high variation conditions, and are similarly excluded from the baseline training. We chose the region between 50kW and 300 kW standard deviation as the representative region for a baseline nominal operation and used it for the AM training.

Figure 7 and Figure 8 show the wind direction and its variation. Wind direction is measured at the MET towers. Two sets of measurements were taken at different heights at each tower. Figure 7 shows Wind Direction 1 measurement at one of the towers (MET 2 in Park 2) for about a 2.5 week period as a function of time. Notice that while the wind direction is stable most of the time as indicated by small standard deviation values plotted in green in Figure 7, the wind direction varies greatly about 8 times during the period covered as indicated by the sharp rises in standard deviation. Figure 8 is a histogram of the standard deviation of the wind direction. For the training set, we excluded the high variation points and only included samples below 10 degrees, indicated by the vertical red line.

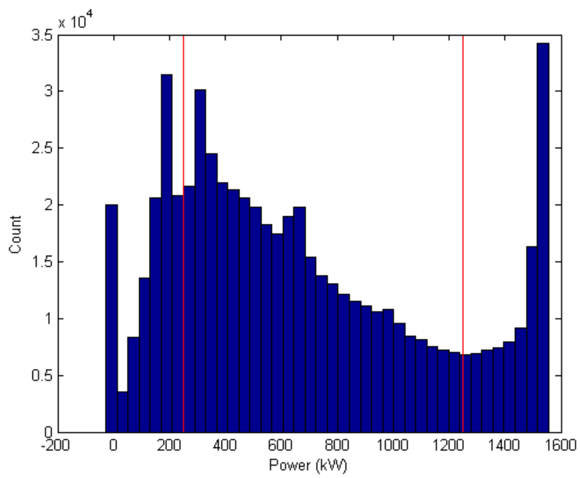

Figure 5: Power profile of 7 wind turbines during 6 months of operation

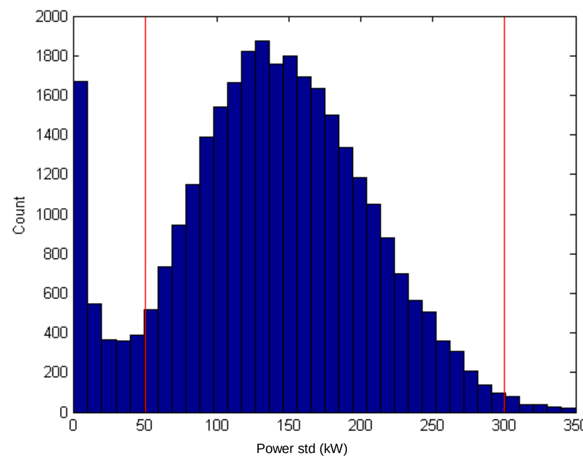

Figure 6: Standard deviation in power among 7 turbines for each sampling instance

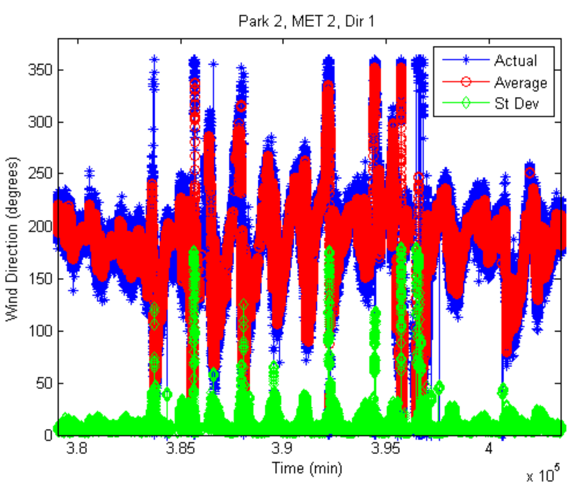

Figure 7. Wind direction at MET2

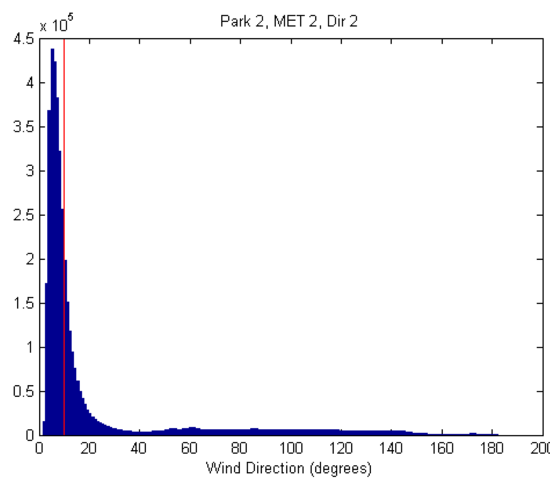

Figure 8. Variation in wind direction

4. RESULTS AND DISCUSSION

We first tested our approach using the complete data set to see how often the associative model approach would indicate an anomaly. This data set was filtered only for stable wind direction as measured by both Direction 1 and Direction 2 and included data from only those instances where the variation in wind direction was below a set threshold as shown in Figure 8. Although the ground truth was not available for each anomaly, an initial assessment for the efficacy of the approach can be made based on the distribution of the indicated anomalies.

Figure 9 and Figure 10 show the associative model results for two select months from the period of September, 2009 to February, 2010. The green, yellow, and red circles in the figures indicate healthy, low-confidence and unconfirmed anomaly, and high-confidence and confirmed anomaly conditions, respectively. The input is made up of data sampled at 1-min intervals, and each sample has a corresponding output indicated by one of the three colored circles.

The yellow indicator may appear in all turbines for each unconfirmed anomaly, while the red indicator is limited to only one turbine. The unconfirmed anomaly is triggered whenever a high residual occurs in any of the turbines, which indicates a mismatch between the associated model output and the measured input signal. However, for an anomaly to be confirmed, a certain amount of persistency in the fault is expected. This confirmation is achieved by a counter implemented as a leaky-integrator. Once the fault persists for a set number of samples, then a hypothesis testing procedure iteratively isolates the fault to one particular wind turbine.

As expected, the monthly data results show that turbines are operating without any indication of an anomaly most of the time. Moreover, when an unconfirmed fault is detected, it does not always persist long enough to trigger a confirmed fault. Notice that once the fault is confirmed, the isolation to a single turbine is sometimes achieved, as is the case on Sep 4 data in Figure 9; or not achieved, as in the case of Jan 5 data in Figure 10. The latter case shows that further tuning of the model may be necessary.

4.1. No Fault Data Set

Besides the faulty data cases, performance assessment of anomaly detection methods needs to include no-fault cases. Especially in a condition-based maintenance approach for remotely located wind turbine applications, the false alarm performance of anomaly detection algorithms is highly critical, as the wind farm operator would incur high cost for an unnecessary maintenance action.

Figure 11 shows the associative model results for a data set made up of nominal power and wind speed. The data set spans six months from September, 2009 to February, 2010. Notice that the model detected a mismatch between the actual and expected output only two times in this data set, as indicated by two sets of yellow circles at 55 and 145 days from September 1, 2009. Neither of these indications persisted, and the associative model output quickly reverted back to healthy. This shows that the approach is very robust with respect to inherent variations in the data and does not produce false alarms.

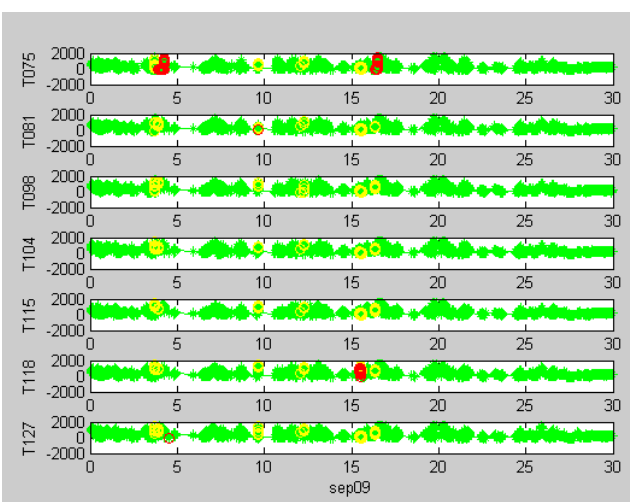

Figure 9. Associative model results for Sep 2009 data

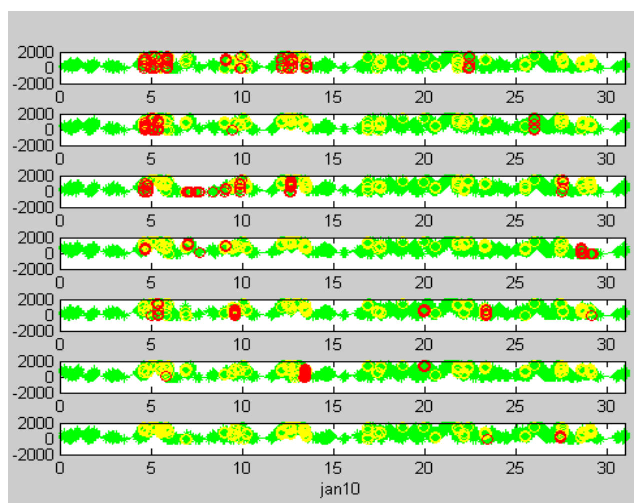

Figure 10. Associative model results for Jan 2010 data

4.2. Ice Condition Data

The SCADA data set includes an aggregate ice condition indicator. This parameter shown in Figure 12 indicates the sum of wind turbines in which an ice condition is detected at any given time. An ice condition denotes the possibility of icing based on low temperature and high humidity—not actual icing. The number of turbines is presented in terms of the percentage of total number of wind turbines in this particular wind park.

During the six-month period for which data is available, a high number of ice condition detections occurred on four occasions: Dec 7–8, 2009, Dec 29–30, 2009, Jan 7–8, 2010, and Jan 28–29, 2010. On these dates, more than 85 percent of the wind turbines out of 155 wind turbines in Park II experienced ice conditions.

Figure 13 and Figure 14 show the anomaly detection results around the four periods of ice conditions. The semi-transparent blue rectangles on the figures indicate the beginning and end of the periods of ice condition. The available SCADA data for the first three periods is very sparse; yet, the associative model was able detect many anomalies in all three cases. The Jan 28–29, 2010 case is preceded with continuous data and followed by no data for about a day. During this fourth ice condition period, all turbines are shown as producing power, and two of them (T104 and T115) are marked as anomalous.

4.3. "Faulted" Data Set

The SCADA data set includes a tag called "faulted," which is a binary parameter recorded for each wind turbine in February, 2010. The data is not available for the earlier periods. When this information is available, it is the closest we have to the ground truth about the state of health of each wind turbine.

One or more "faulted" conditions are recorded in the SCADA data for four of the seven wind turbines (T075, T081, T104, and T127) between Feb 13 and Feb 22, 2010. These faulted cases are indicated by the vertical red lines drawn on the subplots for the wind turbines in Figure 15. As before, the red circles indicate the anomaly isolated to a wind turbine by the associative model.

Notice that for T075 and T104, the red circles precede the red lines. The associative model is not only able to detect the developing anomaly, but it can do so up to two days in advance of the detection logic built into the SCADA system.

T081 seems to have an intermittent failure, as indicated by multiple vertical lines spanning the six days between February 12 and 18. The associative model has two clusters of indications for this wind turbine. The first one is produced on February 15 —about three days after the first faulted signal. The second one is produced on February 17—about one day before the last faulted signal. In other words, the associative model shows sensitivity to the intermitted failure and is able to detect it.

T127 contains one faulted flag during this period. This case is not detected by the associative model. It is interesting to note that this fault occurs at exactly the same time as the last fault on T081. Further information about the nature of this fault could help explain why it is missed by the associative model.

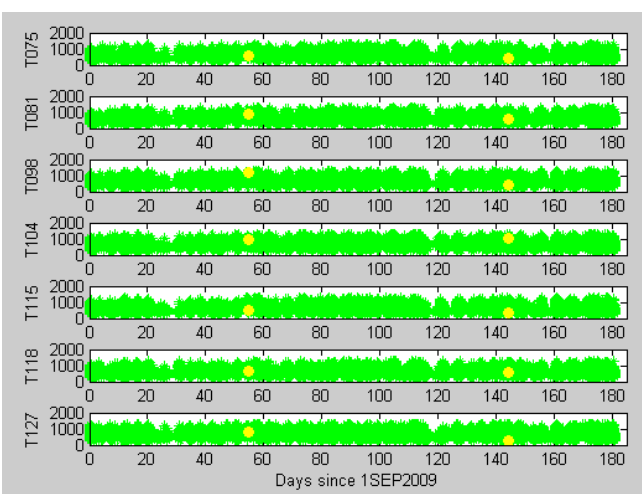

Figure 11. Associative model result for no fault data spanning 6 months

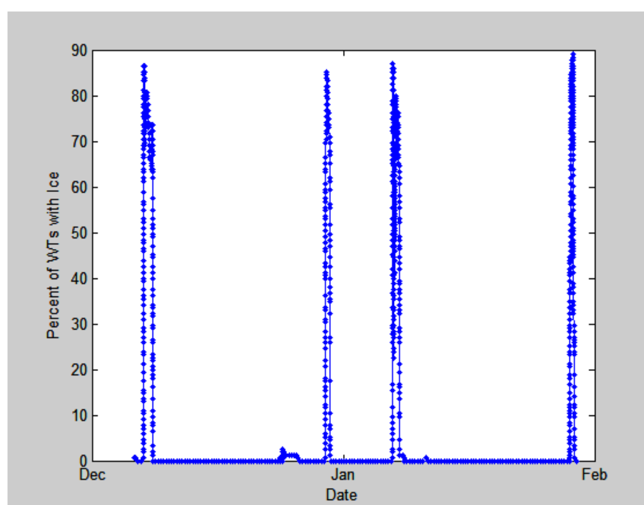

Figure 12. Ice condition times in Park II

Finally, the associative model indicates a fault on T115 on February 20 at about the same time as it indicates a fault on T104. In this case, the anomaly is correctly detected; however, the associative model could only narrow it down to T104 or T115, instead of firmly isolating it to T104.

Overall, the results in this test case show that 3 out of 4 anomalies were detected correctly; 2 out of 4 cases were detected in advance; 1 case was missed; and in one case the ambiguity set could only be reduced to 2 wind turbines.

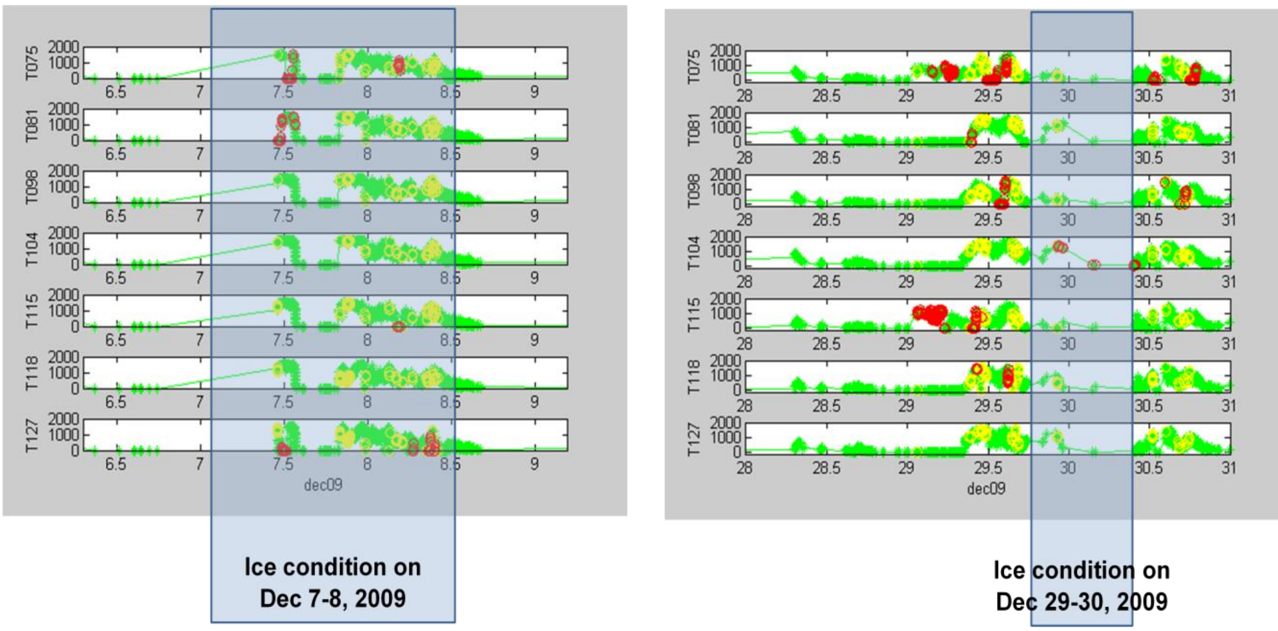

Figure 13. Associative model results for ice condition data in Dec 2009

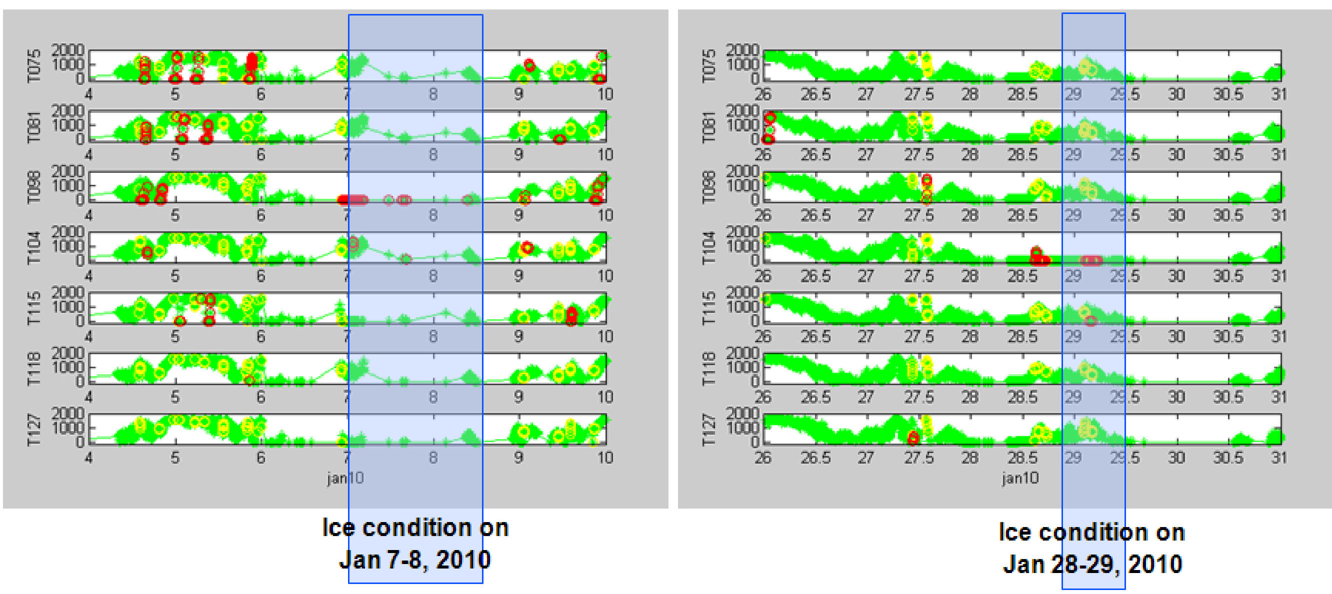

Figure 14. Associative model results for ice condition data in Jan 2010

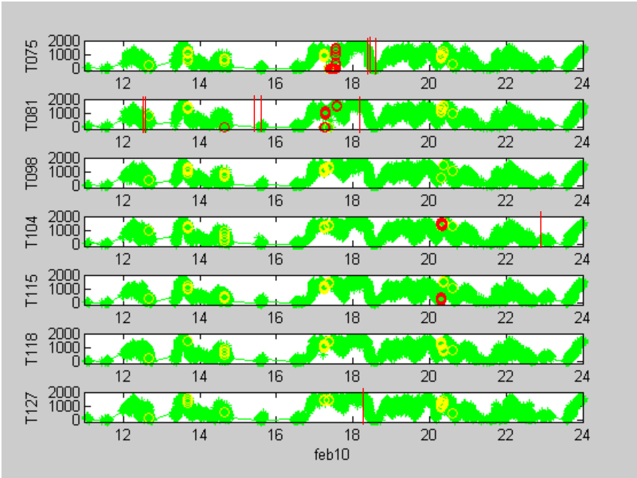

Figure 15. Associative model output for the "faulted" data

5. Conclusion

We analyzed data from several wind turbines on a wind farm and developed an automated fault detection approach using an associative model. The data analysis and visualization of the statistical features showed excellent cumulative correlation over large periods of time (several days) among different wind turbines. The associative model used correlations in a set of wind turbines on a much finer time scale, on the order of minutes, to detect anomalies. Several data filters were created to systematically segregate data for training and testing under different fault or no-fault conditions. The results show that the associative model is a promising approach for capturing correlated wind turbine behavior operating under similar conditions and for automated detection of anomalies when the correlation is broken.

Although we don't have access to the root causes behind anomaly indications, a rich set of configurable data filters provided us with testing data. Testing with the clean data showed no false positives. Testing with the data containing the 'Faulted' tag or large ice condition indicator (indicating possibility of wide-spread icing) showed that the AM is able to detect these conditions—sometimes in advance of the alarms. Further tuning with known fault data is needed to mature the approach.

Acknowledgment

We thank the U.S. Department of Energy for the support for this work under Award Number DE-EE0001368.

References

Bell, M. B. & Foslien, W. K. (2005). Early event detection-results from a prototype implementation. In 17th Annual Ethylene Producers' Conference, Session TA006-Ethylene Plant Process Control. Spring National Meeting (pp. 727-741), Apr. 10-14, Atlanta, GA.

Gorinevsky, D., Dittmar, K., Mylaraswamy, D. & Nwadiogbu, E. (2002). Model-based diagnostics for an aircraft auxiliary power unit. IEEE Conference on Control Applications. (pp 215-220), Sept 18–20, Glasgow, Scotland.

Kim, K. & Mylaraswamy, D. (2006). Fault Diagnosis and Prognosis of Gas Turbine Engines Based on Qualitative Modeling. In *ASME TurboExpo*. (881–889), May 8–11, Barcelona, Spain.

Kim, K., Parthasarathy, G., Uluyol, O., Foslien, W., Sheng, S., Fleming, P., (2011), Use of SCADA Data for Failure Detection in Wind Turbines, Proceedings of 2011 Energy Sustainability Conference and Fuel Cell Conference, August 7-10, 2011, Washington DC, USA.

Kramer, M. A., (1992) Autoassociative neural networks, *Computers Chem. Eng*., Vol. 16, No. 4, 313-328.

Kusiak, A., & Li, W. (2011). The prediction and diagnosis of wind turbine faults. *Renewable Energy*, 36, pp 16-23.

Uluyol, O., Buczak, A.L., Nwadiogbu, E., (2001), Neural Networks Based Sensor Validation and Recovery Methodology for Advanced Aircraft Engines, *Proceedings of AeroSense 2001*, *SPIE* Vol. 4389, pp. 102-109, Orlando, Florida..

Uluyol, O., Kim, K., Wrest, D., Nwadiogbu, E., (2003), Sensor Validation and Recovery through Iterative Association Models, *Proceedings of the IEEE Sensors 2003*, Oct. 22-24.

Uluyol, O., Parthasarathy, G., Foslien, W., Kim, K., (2011), Power curve analytic for wind turbine performance monitoring and prognostics, Annual Conference of the Prognostics and Health Management Society, Montreal, Canada.

Ye, X., Yan, Y. and Osadciw, L., (2010). Learning Decision Rules by Particle Swarm Optimization (PSO) for Wind Turbine Fault Diagnosis, Annual Conference of the Prognostics and Health Management Society, Oct 10-16, 2010, Portland, OR.

Zaher, A., McArthur, S.D.J. & Infield, D.G., (2009), Online wind turbine fault detection through automated SCADA data analysis, wind energy. Published online in *Wiley Interscience* http://www.interscience.wiley.com DOI: 10.1002/we.319.

Authors

Dr. Onder Uluyol is a senior scientist at Honeywell Aerospace Condition-Based Maintenance (CBM) Center of Excellence. He received his Ph.D. in Computational Science and Engineering from the University of Illinois, Urbana-Champaign, IL. Dr. Uluyol has extensive experience in system health management and published numerous technical articles on innovative modeling, measurement, diagnostics, and control approaches. Most recently, he has been focusing on wind and gas turbine performance monitoring using associative models and bearing diagnostics using oil debris and vibration monitoring. Dr. Uluyol holds seven patents on sensor validation, gas turbine light-off detection and start-up monitoring, substance detection using an electronic nose, and on various fault diagnostic and isolation technologies. He has eight additional patent disclosures filed and pending.

Dr. Girija Parthasarathy is a Senior Principal Research Scientist at Honeywell Automation and Control Solutions Labs. She received her Ph. D. in Mechanical Engineering from the University of Minnesota, Minneapolis. She has broad experience with increasing leadership responsibility in research and development programs. She has served roles of principal investigator and program manager in several government- and Honeywell- funded programs. Most recently, she served as PI for a DOE-funded program on Condition Based Maintenance for Wind Farms. Her technical experience spans wind turbine health monitoring, turbine engine diagnostics and prognostics, microgrid modeling and optimization, combined heat and power, combustion and renewable energy. Dr. Parthasarathy has authored several conference and journal papers and holds several patents.