

RESEARCH ARTICLE

Intelligent integrated maintenance for wind power generation

D. Pattison¹, M. Segovia Garcia¹, W. Xie¹, F. Quail¹, M. Revie², R. I. Whitfield³ and I. Irvine⁴

¹ Department of Electronic and Electrical Engineering, University of Strathclyde, 204 George Street, Glasgow, G11XW, UK

² Department of Management Science, University of Strathclyde, 40 George Street, Glasgow, G11QE, UK

³ Department of Design Manufacture and Engineering Management, University of Strathclyde, 75 Montrose Street, Glasgow, G11XJ, UK

⁴ SgurrEnergy, 225 Bath Street, Glasgow, G24GZ, UK

ABSTRACT

A novel architecture and system for the provision of Reliability Centred Maintenance (RCM) for offshore wind power generation is presented. The architecture was developed by conducting a bottom-up analysis of the data required to support RCM within this specific industry, combined with a top-down analysis of the required maintenance functionality. The architecture and system consists of three integrated modules for intelligent condition monitoring, reliability and maintenance modelling, and maintenance scheduling that provide a scalable solution for performing dynamic, efficient and cost-effective preventative maintenance management within this extremely demanding renewable energy generation sector. The system demonstrates for the first time the integration of state-of-the-art advanced mathematical techniques: Random Forests, dynamic Bayesian networks and memetic algorithms in the development of an intelligent autonomous solution. The results from the application of the intelligent integrated system illustrated the automated detection of faults within a wind farm consisting of over 100 turbines, the modelling and updating of the turbines' survivability and creation of a hierarchy of maintenance actions, and the optimizing of the maintenance schedule with a view to maximizing the availability and revenue generation of the turbines. © 2015 The Authors. *Wind Energy* published by John Wiley & Sons Ltd.

KEYWORDS

reliability centred maintenance; intelligent condition monitoring; reliability and maintenance modelling; maintenance scheduling

Correspondence

R. I. Whitfield, Department of Design Manufacture and Engineering Management, University of Strathclyde, 75 Montrose Street, Glasgow, G11XJ, UK.

E-mail: ian.whitfield@strath.ac.uk

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Received 16 May 2014; Revised 27 February 2015; Accepted 16 March 2015

1. INTRODUCTION

The wind power industry has grown considerably over the past 15 years, and with an expected installed capacity of 230 GW by 2020 and 400 GW by 2030, it is gearing up to become the main power technology in the EU.¹ To meet emission targets, the UK wind industry is planning very large offshore wind farms, some at considerable distance from shore and in deeper water, which pose challenges from the viewpoints of installation and maintenance.² As the number of wind turbines increases, the resources needed to keep these assets in optimal operating condition also increases. These resources include the infrastructure used to monitor the performance and health of the assets, the technology needed to analyse the data generated by the sensors to inform maintenance related decisions and the equipment needed to perform maintenance.

For an offshore wind farm, a failure can cause a significant downtime given the difficulties in accessing sites due to weather, sea state and availability of equipment, for example. In addition, a relatively large period of wind turbine downtime is required to conduct repairs, resulting in further loss of revenue.¹ Costs of operation and maintenance (O&M) actions in this case can be up to 25–30% of the cost of the energy² and typically estimated at five to 10 times of cost of onshore maintenance,³ and to this, a

major loss on the energy generating revenues must be added. For this reason, it is of paramount importance to prevent the failure of wind turbines and reduce unavailability. While traditional maintenance scheduling (MS) is often 'corrective', we recognize that offshore maintenance must be 'preventative' where possible, and that in order to achieve this, the solution must have direct online access to the necessary information with which it can optimize decisions relating to maintenance tasks.

1.1. Proposed modular architecture

This paper presents a novel architecture and system to support Reliability Centred Maintenance (RCM). The aim of the paper is to illustrate the integration of different mathematical methods within the development of a systemic solution to RCM, which is subsequently implemented within an offshore wind power generation context. Within the solution presented, decision making is informed by observations from the sensors installed on the turbines; condition monitoring provides the orientation of the data in the form of identifying anomalous behaviour; reliability and maintenance modelling (RMM) supports decision making by proposing different maintenance options, and finally MS generates a near-optimal decision with respect to the overall maintenance cost effectiveness of the wind farm. Figure 1 provides an illustration of how associated modules may be integrated from an information flow perspective to create the solution and consists of six high-level steps. The three modules of the solution are as follows:

- **Intelligent condition monitoring (ICM):** This module uses methods to analyse data from wind turbine supervisory control and data acquisition (SCADA) units and deployed sensors that measure features such as mechanical vibration⁴ to determine the state of components within the turbine and monitor deviations from normal behaviour (step 1 in Figure 1). If any such deviation/anomaly is found, the system identifies this state and informs the *Reliability and Maintenance Modelling* module (step 2).
- **RMM:** This module models a component's probability of survival across the duration of its intended service life. This module uses statistical analysis and input from the ICM module to adapt a generic lifetime model to the observed state of the component (step 3) and determine the best maintenance action for turbine survival, which resolves the fault (step 4).
- **MS:** This module uses the RMM lifetime estimations to produce a context-enhanced maintenance schedule (step 5). The overall goal is still to minimize cost and downtime, but by being able to better predict the state of a component, overall costs can be reduced through predictive maintenance. The implemented maintenance action will subsequently impact the turbine survivability within the RMM (step 6).

As illustration of the interaction of these modules, if a turbine is known to experience high wind shear, it is likely that the gearbox will be under higher stress, leading to a shorter service life. This can manifest itself in the ICM module detecting anomalous behaviour, which in turn informs the RMM module of additional wear on the component, which finally informs the MS module of the component's deteriorated state. If the deterioration is significant, the scheduler will determine that the component requires more than standard scheduled maintenance, resulting in the requirement of additional equipment for a more significant maintenance operation to be made available to the maintenance team. This maintenance intervention potentially extends the lifetime of the component by a greater amount than standard scheduled maintenance.

The maintenance problem is compounded by most large-scale wind farm arrays potentially having a range of wind turbine types, variants and entry into service dates, as well as the individual turbine wear being influenced by the 3D nature of

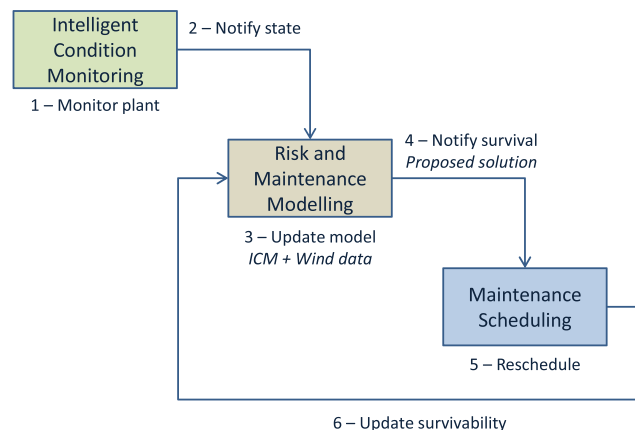


Figure 1. Proposed high-level information flow.

the inflow conditions: topography, wind shear and atmospheric stability assumptions. As such, each individual turbine on the wind farm requires individual treatment.

In order to consider the appropriateness of available approaches as potential candidates for an integrated solution, a more detailed understanding of the types and availability of information required for an integrated solution was developed. Figure 2 illustrates in more detail the type and flow of information required to enable intelligent integrated RCM. The data sources are distinguished as either ‘online’ or ‘offline’. The former indicates that the data are generated or computed at runtime by the system and will be passed between modules. The latter refers to anything that has been taken from an external data source, which includes information derived from expert knowledge, existing/fixed maintenance regimes and known safety limits.

The ICM module uses existing fault detection practices and component dependencies, to integrate with online turbine sensor SCADA data (including frequency, standard deviation and latency of the observations). The ICM module subsequently provides a diagnosis of each turbine state, identifies whether a component is failing and determines the overall system state.

The RMM module uses offline information relating to the mean time to failure (MTTF) for the turbine components as well as system diagrams illustrating the component dependencies and integrates this with the diagnosis output from the ICM module. SCADA data are also used within the RMM module to estimate the survivability (predicted repair state) for each turbine and subsequently generate a hierarchy of maintenance actions. In addition, the failure rate data for the components within the turbine are enhanced by incorporating these new observations to improve the accuracy of the MTTF.

The MS module uses information relating to existing scheduled maintenance for the turbines as well as constraints with respect to the availability of maintenance equipment and integrates this with the maintenance action hierarchy from the RMM module. Optimized maintenance schedules are created for each of the potential maintenance actions (integrated with other scheduled actions) with a view to optimizing the turbines’ availability. The RMM module is notified if a maintenance action cannot be scheduled, e.g. a generator replacement requiring lifting equipment, which is unavailable or unsafe to use given current meteo-oceanic conditions. In addition, the MS module notifies the RMM module of all scheduled maintenance actions in order that the survivability of each maintained turbine can be updated.

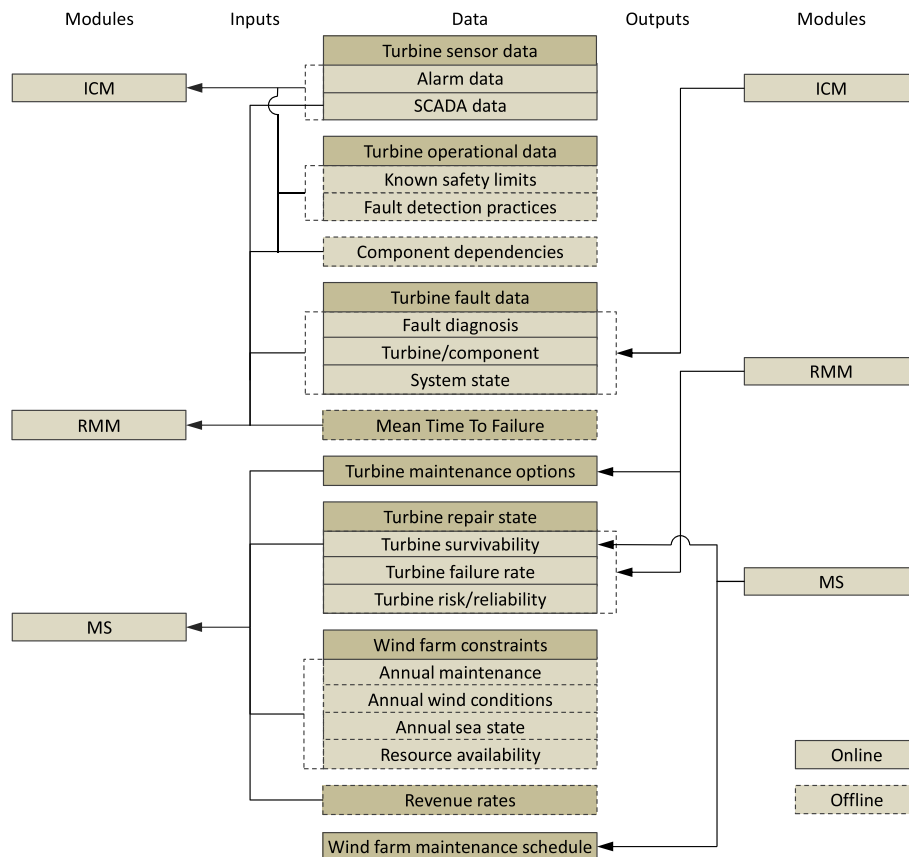


Figure 2. Input and output information flow.

1.2. Technology review

While ICM is a relatively new concept in the wind industry, it has been deployed with success in many related fields. Catterson *et al.* use a Gaussian mixture model (GMM) derived from transformer sensor data, in order to detect anomalous states within the plant,⁵ while Yu applies an adaptive version of a GMM to the problem of machine-tool degradation.⁶ Here, the use of ICM is in detecting anomalous readings with respect to other correlated variables, for which a hand-coded model would be intractable. Among the most commonly deployed ICM techniques has been artificial neural networks (ANNs), which can be used for detecting both anomalous sensor readings and diagnosing these as specific faults.^{7,8} Using only SCADA data, Zaher *et al.* successfully applied ANNs to detect anomalous temperature readings within a turbine's gearbox and cooling oil,⁹ while Kusiak and Li also applied ANNs for diagnosing the severity of a fault.¹⁰

Yan demonstrated that the use of Random Forests (RF)¹¹ for classifying faults outperformed conventional decision tree classifiers and support vector machines, as well as producing comparable performance to ANNs.¹² Kusiak and Verma compared a range of different data mining algorithms to the prediction of faults within wind turbines.¹³ The magnitude of the fault detection challenges within wind turbines was highlighted by Kusiak and Verma with 16 normal operating states, and one fault state, which could in turn have over 400 different reasons for the fault. SCADA data were used to train the algorithms, which consisted of over 100 parameters for 17 turbines, with two-thirds of the available data used for training. Using the geometric mean of the fault class as the algorithm performance metric, Kusiak and Verma established that the RF algorithm provided the most accurate results with an accuracy of 78–98%.

Complex system components exhibit dynamic behaviour, where not only the combination of failing components reflects the state of the system but also the sequence in which these components fail.¹⁴ While conventional techniques such as fault trees fail to capture this dynamic behaviour, Bayesian belief networks (BBNs) support reasoning under uncertainty and constitute a flexible and powerful probabilistic modelling framework that makes them suitable for applications in the field of reliability and maintenance.^{15,16} Recently, influence diagrams and BBNs have been used for modelling the reliability of civil structures. Models have been used to optimize the acquisition of condition data of structures¹⁷ and optimize the maintenance of individual structures.¹⁸ Typically, BBNs are static models that represent the joint probability distribution at a fixed point or interval of time. To account for temporal dependencies, an explicit representation of time in a BBN is needed. Dynamic Bayesian networks (DBNs), such as those described in the work of Straub and Kiureghian,¹⁸ extend BBNs to allow for reasoning in a dynamic world where changes occur over time. To date, these models have successfully captured the reliability of individual turbines. The research presented here expands on previous models by considering the impact of environmental characteristics gathered in real time and the challenges of accessing remote assets.

Maintenance scheduling problems can be generally classified as optimization problems where the aim is to find the optimal schedule(s) to satisfy the designed objective(s). Genetic algorithms (GAs) can be considered a suitable model owing to their ability to perform global optimization and intelligent parallel searching in non-linear solution space.¹⁹ To achieve optimal solutions, several strategies have been developed using a combination of GAs, Monte Carlo (MC) simulation, and simulated annealing (SA), for example. Dahal *et al.*²⁰ used a GA with a fuzzy evaluation function for generator MS. Later, the same author extended the study with GA/SA and GA/SA/heuristic hybrid approaches; the results of the investigation demonstrated that the GA/SA/heuristic hybrid approaches resulted with slightly better results than the GA/SA approach, while both approaches were better than the results generated from a simple GA.

Garcia *et al.* presented an integrated intelligent solution for predictive maintenance of a wind turbine gearbox.²¹ The SIMAP tool incorporated an ANN to detect anomalous behaviour within the gearbox as well as for health condition, a fuzzy expert system to establish the failure mode relating to the anomalous behaviour and a fuzzy GA to schedule the associated maintenance action. While the authors suggest that the tool may be applied to more complex components, as well as to more than one turbine, the approach was demonstrated for a single gearbox. The research undoubtedly lays the foundation for demonstrating the concept of intelligent integrated maintenance; however, one of the main challenges lies in the use of a fuzzy expert system for the diagnosis of the anomaly. The rules incorporated within the SIMAP fuzzy expert system are defined *a priori*, despite such a pre-defined approach not being able to detect previously unknown errors or patterns as is prevalent within this problem domain. In this paper, we develop an automated and intelligent approach to identify failure patterns or modes that does not require expert input and can be applied to numerous systems.

The framework developed in this paper is in contrast to the current suite of O&M models available,^{22,23} which focus on long-term planning of an offshore wind farm and strategic decisions relating to utilizing vessels, while the framework developed here addresses short-term decisions such as MS. This proposed modular structure for the solution provides a focus for the development of each of the modules (Sections 2) and the integration of the modules (Section 5). The modules are discussed within the context of gearbox anomalies; however, the modules are applicable for all wind turbine components. A case study is described within Section 6, which demonstrates the implementation of the integrated solution using over 100 turbines that included seven turbines that were known to possess anomalous behaviour.

2. INTELLIGENT CONDITION MONITORING MODULE

Like the system as a whole, the ICM module is designed to support generic turbine data in order that deployment time is minimal and that target turbine models are largely irrelevant. To achieve this, machine learning (ML) models form the core of the module, with both supervised and unsupervised paradigms supported. Such automated processing is a key requirement of continuous large-scale fleet monitoring where human observation is not cost feasible. Figure 3 illustrates the flow and processing of information within the ICM module.

It is rarely the case that operators will use a single manufacturer/model of turbine within their fleets, preferring instead to select these on a site-by-site basis. This leads to a mixture of assets throughout the fleet, making it intractable to design bespoke monitoring systems. However, most modern turbines are equipped with common sensors and logging abilities, generating data that can be interpreted by ML models.

Specifically, only SCADA and alarm data are required for fault detection and diagnosis models to be constructed. SCADA data refer to the industry standard of sensor readings, such as generator Revolutions Per Minute (RPM) yaw positions and wind speeds, which each have their mean, minimum, maximum and standard deviation recorded and averaged over a 10 min period.

It is assumed that the turbine will have an automated alarm system. Alarms are generated by a combination of the internal CM system and turbine controller and indicate warnings such as generator over-speed, low oil pressure and blade vibrations. For modern turbines, a typical internal CM system will have several thousand possible alarms. If available, we can also incorporate into the modelling data on safety limits, e.g. shutdown speeds, and data on current fault detection procedures.

Condition monitoring can largely be classed into three categories: detection, diagnosis and prognosis of faults. Of these, only detection and diagnosis are currently supported and are represented as an ML *classification* problem. Here, fault detection is simply detecting whether the turbine is in an anomalous state, where this means anything that is not the normal, intended behaviour. Fault diagnosis takes this further by determining the type of state the turbine is in (e.g. active fault). In both cases, this is implemented using an *RF*.¹¹

All training of ML models is currently performed offline. The model training system is provided with a historical SCADA dataset and associated alarm set for the same turbine, which then automatically derives the trained model. It is at this point that the input SCADA data must be annotated with a class indicating whether it is normal or anomalous. For detection, this is achieved by assuming that any data sample that has an alarm active at the associated time is anomalous, while if no alarm is active, it is normal. Classification is used in preference to anomaly detection methods as the ratio of samples during which an alarm is active is suitably high enough for classification to be valid. For fault diagnosis, the alarm code itself is used as the label, e.g. alarm code 10512 indicates a high temperature in the main bearing. Irrelevant alarms that may be active but do not indicate an anomalous turbine state are filtered, e.g. communication with the central database being offline. Additionally, if maintenance records are available for the turbine and time period considered (as used by the RMM module), any data samples associated during the work period are removed, as there is no way to know if the turbine is in a normal operating state.

In both of these cases, an RF is trained on historical SCADA data for each turbine, which is labelled as previously described. Of these data, 80% are used to train the RF with the remaining 20% used to evaluate the trained model. Each turbine-specific model is evaluated on approximately 5000 SCADA samples. Table I shows that the RF provides high accuracy in fault detection, with an overall accuracy of 98.83%. This percentage is calculated as the number of true positives (correctly classifying the turbine as being in a good state), plus the number of true negatives (correctly classifying the turbine as being in an anomalous state), divided by the total number of samples.

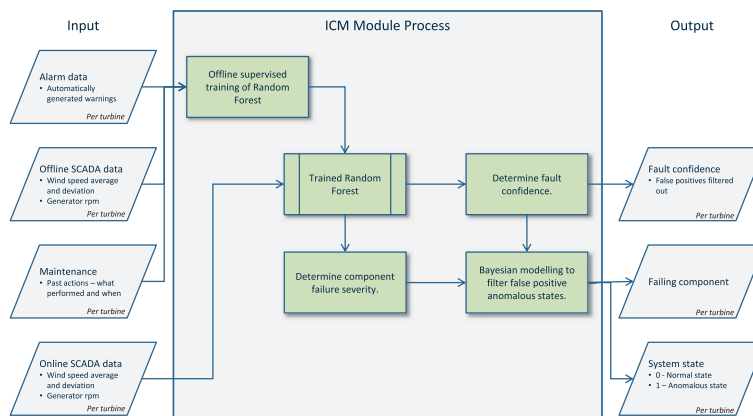


Figure 3. ICM module data and process flow.

Table I. Confusion matrix of the ICM module with regard to fault detection.

	Working (predicted) (%)	Failed (predicted) (%)
Working (actual)	99.7	0.30
Failed (actual)	9.3	90.7

The fault detection and diagnosis models offer high accuracy when trained on historical datasets, with detection being slightly higher. The reason for this discrepancy is that classes for fault diagnosis are based upon alarm codes, rather than the simple binary normal/anomalous behaviour of fault detection. Given that certain alarm codes will only appear once or twice in the training data, it is unlikely the model will be able to accurately classify any future data of this type without seeing many more samples during training. Similarly, a data-driven model cannot classify unknown alarms/faults.

The ICM module continuously monitors all turbines in real time. Regardless of whether a fault is encountered or not, the system notifies the RMM module of the turbine number and state. However, the ICM system is designed to detect anomalies in the turbine, including those that are transient, i.e. faults that only last for a single 10 min sample, such as an over-speed caused by high wind.

These can often be false positives, caused by erroneous sensor readings that do not truly represent the state of the system. Alternatively, these may be genuinely anomalous readings, which, while of interest to the ICM system, have negligible impact on the life of the turbine. As such, we choose to filter these anomalous classifications prior to them being transmitted to the RMM module. The accuracy of a fault detection model is determined by its overall accuracy, λ , the *true positive rate* (*TPR*) and *true negative rate* (*TNR*) as defined in equations (1)–(3). Here, *TP* refers to the true positives output by the classifier and the corresponding abbreviations refer to the true negative, false positive and false negative rate. In this case, a true positive is a correctly classified ‘normal’ SCADA sample, while a true negative is a correctly classified anomalous sample.

$$TPR = \frac{TP}{TP + FN} \quad (1)$$

$$TNR = \frac{TN}{TN + FP} \quad (2)$$

$$\lambda = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

The overall accuracy, λ , is used as an indicator of how strongly evidence should be weighted towards the current observation versus prior knowledge,²⁴ as shown in equations (4), (5), where S_A is an anomalous turbine state, S_N is a normal turbine state and C_{conf} is the confidence the classifier, C , has in the label produced for the respective SCADA sample. This is used to prevent poor classifiers continually providing false negatives/positives to the RMM module.

$$P(C|S_A) = \lambda(TNR \times C_{conf}) + \frac{1 - \lambda}{|S|} \quad (4)$$

$$P(C|S_N) = \lambda(TPR \times C_{conf}) + \frac{1 - \lambda}{|S|} \quad (5)$$

These values are used to derive a probability of the classification being true, given the classifier’s overall and individual accuracy on normal and anomalous data, using a simple Bayesian update. Unless the posterior probability of a fault occurring ($S=S_A$) exceeds a user-defined threshold, the RMM module will not be notified of a fault. Regardless of the filtered result, the RMM module uses the ICM output as an indicator of turbine state, in order to derive the remaining life of the asset.

3. RELIABILITY AND MAINTENANCE MODELLING MODULE

The output from the ICM provides an assessment of the current state of the system. However, it does not provide information on the future evolution of the system, particularly in relation to its current age. The purpose of the RMM module was to model the natural degradation of the system and to incorporate the output of the ICM into this calculation. A DBN was chosen as it met four key criteria. First, it has a strong theoretical foundation. Second, it is capable of capturing the dynamic nature of degradation over time and into the future. Third, it is capable of updating the reliability of the system based on the output of the ICM. Finally, it is capable of estimating the resulting state of the system under different types of maintenance actions.

Two workshops with maintenance engineers and senior management were conducted to qualitatively construct the model. The aims of these workshops were to understand degradation of a gearbox of an offshore wind turbine and identify factors influencing the deterioration of the system. The first workshop focused on understanding how gearboxes aged, the most appropriate metric for measuring the age of a turbine and the variables that influenced the rate of degradation. The second workshop focused on identifying data sources to quantify the model, and where necessary, for the experts to specify their subjective belief about the dependencies.

During the first workshop, engineers stated that calendar time was an inadequate measurement of the age of a turbine. The number of rotations, used as a proxy for use, was proposed but rejected by experts. They believed that turbulence intensity, defined as the ratio of the wind speed standard deviation and the mean wind speed determined from the same set of measured data samples of wind speed, and taken over a specified period of time was a key contributor to degradation.

From this, a new metric, the effective rotations of the generator (ERG) was defined. The ERG was dependent on the generator RPM, i.e. the number of rotations per time unit, which measures the usage of the generator, and on the turbulence intensity, which provides the conditions of usage, i.e. $ERG = NR * TI$, where NR is the number of rotations in a 10 min interval, and the TI is the turbulence intensity during that period, defined as high, medium or low. The latter variable corresponds to the external conditions accelerating the deterioration on the gearbox. For wind turbines exposed to high turbulence intensity, it is expected that the gearbox will deteriorate much faster than when this intensity is low. As such, during periods of high turbulence intensity, TI is greater than during periods of low turbulence intensity, therefore increasing the age of the system, effective generator rotations, faster. Experts used historical data and observation of onshore wind farms to define the boundaries between high, medium and low turbulence intensity, and to assess the impact that differing levels of turbulence intensity has on the failure rate.

The impact of maintenance actions was also considered. While the previous variables, i.e. generator RPM and turbulence intensity, increase the deterioration on the gearbox, maintenance actions, if well performed, will have an opposite impact on the deterioration. It is believed that the maintenance actions will reduce the deterioration of the system and will rejuvenate the system.

Finally, engineers believed that the survival probability would influence the output of any condition monitoring systems installed on the turbine. The output of CMSs was used to update the survival probability of the gearbox.

Considering these variables, the dynamic evolution of the system was represented by the DBN given by Figure 4. The arcs on the figure correspond to direct probabilistic dependencies between the different variables. Straight arrows indicate relationships within the same time slice, while circular arcs represent relationships from one time slice to another. The structure of the DBN was validated through interviews with additional experts.

Three different methods were used for quantifying the dependencies in the DBN. First, when large volumes of data were available, a Kalman filter (KF) was used to quantify the strength of dependency between two variables. Second, for dependencies where data were unavailable but the relationship was well understood by the operator or a proxy variable was created, structural equation modelling was used. An example of this is modelling the ERG. For that, the following equation was used:

$$ERG_{t+1} = (ERG_t + \gamma_{t+1}G_{t+1})(1 - \rho_{t+1}) \tag{6}$$

where ERG_{t+1} and ERG_t represent the ERG at time slices $t + 1$ and t , respectively, G_{t+1} quantifies the generator rotations between the time slices t and $t + 1$, γ_{t+1} is the turbulence intensity impact between time slices t and $t + 1$, and ρ_{t+1} represents the effectiveness of the maintenance action performed, if any. Finally, when data were unavailable and the relationship between two variables was unintuitive to the operator, copulas were used. Copulas are a flexible method for capturing many different dependency structures between two variables.

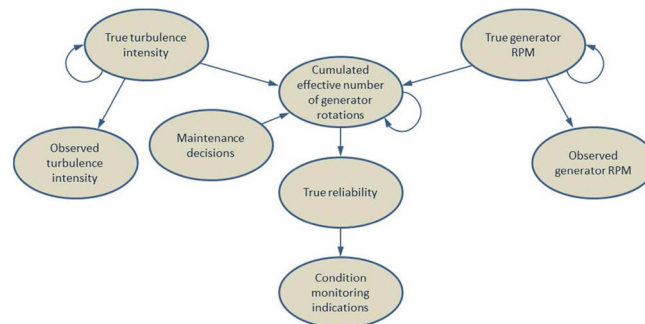


Figure 4. DBN for the deterioration of the gearbox.

Once the model was fully populated, it was used to forecast future deterioration beyond time step t . Using available data, the DBN calculated the probability of survival at any instant of time. The historical dataset of SCADA observations was used along with the KF to update the estimate about the turbulence intensity and generator RPM, i.e. to estimate the true value of these two variables. From this, the DBN infers ERG, which was then used to estimate the current probability of failure of the gearbox. The probability was then updated considering CM indications. Additionally, the KF was used to predict the turbulence intensity and the generator RPM over the next time steps as it can reproduce the turbulence intensity and generator RPM patterns observed in the previous time steps. In that way, it was possible to predict how the gearbox was going to deteriorate in the near future, e.g. the next 24 h.

The outputs of the RMM module were an estimation of survival probability of the gearbox at any instant of time, and an estimation of the impact of possible maintenance actions when necessary. The DBN also estimated the degradation in the near future by considering the short-term patterns of turbulence intensity and generator revolutions. This output was generated for each individual turbine and was used by the MS module to determine which action should be performed in each particular case.

Dependency between each gearbox is not explicitly modelled in the framework. However, as we assume each gearbox begins operation with the same life and each experiences similar wind conditions, the failure time of each gearbox will be implicitly dependent.

4. MAINTENANCE SCHEDULING MODULE

Offshore wind turbine maintenance tasks require the scheduling and management of different kinds of resources, such as skilled personnel, spare parts and special equipment such as vessels and ships. Maintenance teams are dispatched to farms and turbines in response to the maintenance schedule. A team can execute only one task at a time, and it is assumed that the team must finish the task before moving to the location of the next task. A domain-specific feature is the dependence of maintenance tasks on environmental conditions, such as wind speed or sea state. Certain maintenance tasks have maximum values of each weather parameter, depending on the type of work to be performed, as well as the safety regulations. Many maintenance tasks require access to appropriate spare parts, some of which may be immediately available and kept in stock, while others must be ordered, resulting with additional lead times before the maintenance tasks may be scheduled. More complicated tasks require special equipment such as jack-up barges, which must be hired from service suppliers and are subject to other demands, and subsequently have a service availability interval, which influences the maintenance window.

Periodic and preventive tasks are planned and released well before they become timely. Maintenance planners assign time windows to these tasks, in which the date of execution can be chosen according to the actual circumstances, e.g. the maintenance commitments of the technicians or the weather conditions.

In addition to failures, production loss is caused by the maintenance tasks themselves, since turbines may have to be stopped during maintenance. An interesting feature of this problem domain is that the maintenance of one turbine may stop other turbines as well; since several turbines are connected serially to the grid, the complete disconnection of one turbine stops its posterior ones as well. For each task, the set of affected turbines can be determined based on the states required by the given task.

Scheduling consists in determining the set of tasks that should be executed within the scheduling horizon and assigning a team and a start time to them, in order to minimize the total production loss of the turbines. The objective function contains both production loss due to failures and cost due to maintenance. Note that from the scheduling point of view, this optimization criterion belongs to the class of irregular criteria, which is an atypical and difficult-to-handle class. This means that it may be worth postponing certain tasks, e.g. from a period with high winds to a later period with low winds, even if all the resources are available to execute it earlier.

Maintenance scheduling for offshore wind farms is a typical multi-variable and multi-objective optimization problem. Figure 5 illustrates the flow of information within the MS module and with the other ICM and RMM modules. A memetic algorithm is used to perform the optimization, which consists of the following elements: input/output, parameter setting, genetic representation, population initialization, reproduction selection, genetic operations, local search, logistic optimizer, fitness measurement, generational selection, and stopping criteria.

A systematic data structure was constructed owing to the large amount of data required for the optimization of the maintenance schedule, which consisted of a 3D matrix consisting of turbine numbers, maintenance task codes and task dates as the dimensions. The gene is encoded as a combination of elements from these dimensions, which act as keys to locate and retrieve all other relevant information relating to the gene such as the number of maintenance personnel and transportation required.

The sequence of maintenance tasks within each possible schedule is encoded as a string, with each maintenance task represented as a gene in the string. The whole chromosome represents a complete schedule of the required maintenance tasks. The gene and chromosome encoding is illustrated in Figure 6. This coding scheme may be easily altered based upon

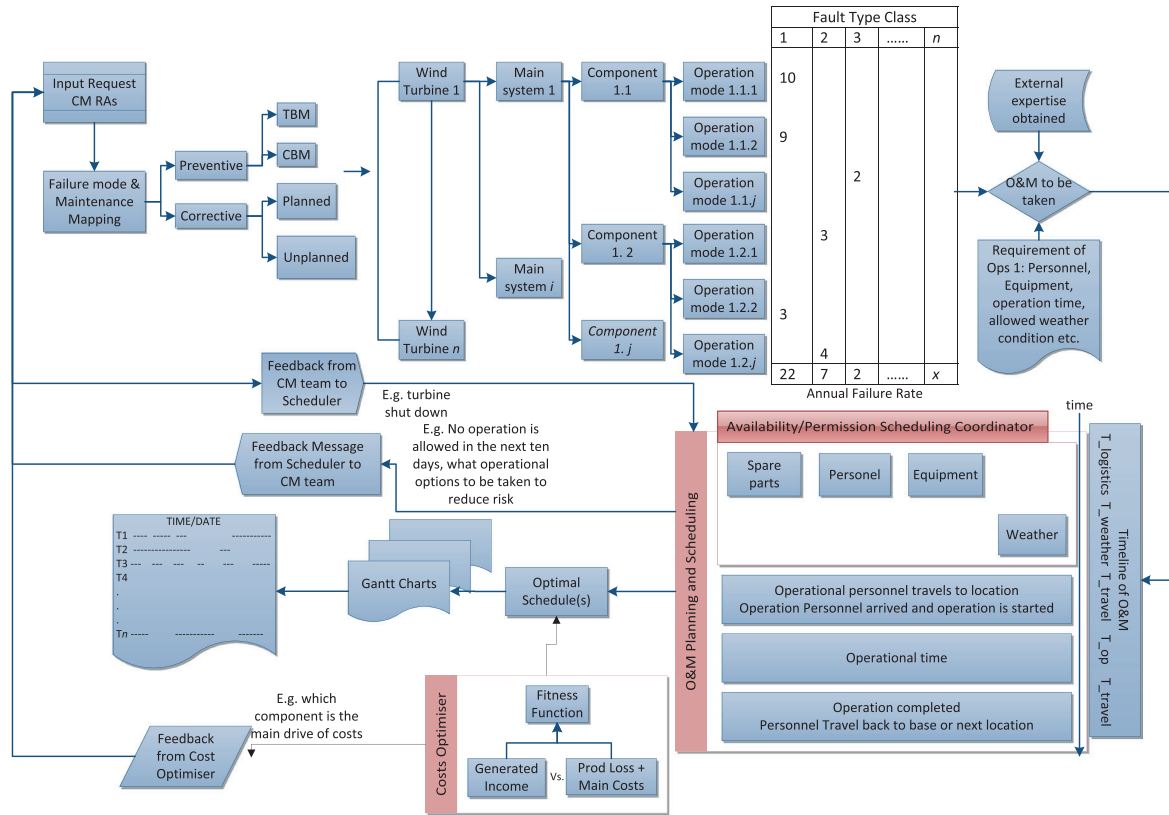


Figure 5. Maintenance scheduling module flow chart.

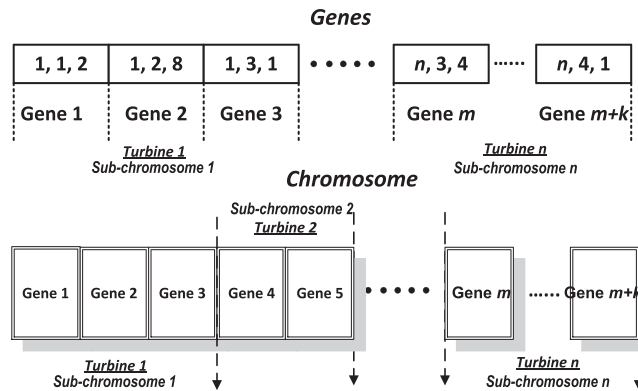


Figure 6. Gene and chromosome encoding.

the size of the scheduling problem and has been developed specifically to schedule maintenance tasks for hundreds of machines at any time.

A simple hill climbing local search is used to improve the search capability of the memetic algorithm. The local search chooses the direction randomly and takes one step each time by altering either the maintenance intervention category or the task date by 1 day. The local search procedure continues if a better solution is found and stops if a worse solution is found.

A logistic satisfier is used to ensure that the schedules generated by the MA do not violate any of the constraints at the time that the task is scheduled, which include meteo-oceanic, labour and transportation constraints. The meteo-oceanic conditions come from historical data to identify weather windows appropriate for the maintenance tasks. These data reflect daily averaged wave heights and wind speeds over a 1 year period reflective of the wind farm’s location.

In order to quantify the economic impact of individual maintenance tasks, the production loss due to the corresponding (present or future predicted) failure and its maintenance time is estimated. The turbine returns to an improved state after

maintenance is undertaken with an associated increase in power and revenue generated. However, the loss of the generated power cannot be modelled as constant over time: it largely depends on the wind speed, as the production of the turbine is proportional to this parameter. The wind speeds are averaged over a 24 h time interval using historical data, which allows an annual variation of power generation to be determined.

The fitness function captures the trade-off between the possible income revenue/power generated against the costs due to the downtime of the maintenance tasks and its associated direct costs such as transportation and component costs. The fitness function consists of the total sum of possible power generated before and after the scheduled maintenance minus the total sum of production lost due to the maintenance and the maintenance costs itself as seen in equation 7, where N represents the number of maintenance actions, i represents the maintenance action index, MC_i represents the maintenance cost for maintenance action i , J represents the number of turbines, j represents the turbine index, $PL_{j,t}$ represents the production loss on turbine j in time period t , P' and P'' represent the power generated before and after the maintenance action, p'_s and p''_s represent the survival probability before and after the maintenance action, and T and t represent the number and current time interval.

$$\text{Max} \left(\sum_{j=1}^J \sum_{t=0}^T (p'_s P' + p''_s P'') - \left(\sum_{i=1}^N MC_i + \sum_{j=1}^J \sum_{t=0}^T PL_{j,t} \right) \right) \quad (7)$$

A rolling 1 year maintenance window is typically used, which consists of 365 time intervals. While the maintenance window can be easily extended, it will impact the optimization since it will increase the search space for scheduling the maintenance actions. The overall maintenance cost includes transportation cost, access equipment cost, replacement component cost and labour cost. The transportation cost for the maintenance actions is not considered within the window of the maintenance action if using the same transportation, and the cost for hiring jack-up barges is considered on the first day of hire, and costs for all maintenance actions using this transportation during the duration of hire are not considered.

In order to quantitatively populate the model, both historical data and expert judgement were used. Examples of variables include the costs associated with transportation, labour, access and components. For wind turbines, a yearly scheduled maintenance is typically carried out, which provides a starting point for generating the revised schedule. The goal of the integrated solution was to automatically manage maintenance for up to 300 turbines at any time and plan the maintenance schedule over a 12 month window.

The MS module generates a project plan for annual and preventative maintenance. This was a list of dates and actions that should be carried out throughout the year. The output also included an expected cost for the upcoming year including generated income and costs associated with repairs. Once a maintenance task has been undertaken, the MS module updates the RMM module so that the reliability of the turbine can be updated.

5. INTELLIGENT INTEGRATED MAINTENANCE SYSTEM – AURA

Aura is a data-driven intelligent integrated maintenance system, with a number of inputs taken from heterogeneous sources. The modules discussed previously are controlled via a central engine, which monitors the input and output data from each module as well as from external resources. Figure 7 provides a simplified illustration of how the Aura system transfers data between each of the three modules.

For typical operation, the ICM module would continuously monitor the SCADA data of the remote assets in order to detect and diagnose anomalous behaviour. The ICM manages a number of models for each of the turbines within the wind farm. The data received from these assets include live data being streamed and data specific to condition monitoring units included on the turbine. Offline data were also used and could either be in the form of historical data collected from similar turbines on other wind farms or structured judgement elicited from experts. Once new data have been read, they are passed to the ICM model associated with the turbine that the data originated from. The ICM module analyses the data to detect whether they are anomalous and diagnoses the fault if it is. Regardless of whether the data are anomalous, the result is cached for future use by the ICM module.

While the ICM module operates continuously within the Aura system, both the RMM and MS modules are operated intermittently in response to the wind farms requirements for the maintenance schedule. This could, e.g. be performed continuously in real time, as a nightly operation or less frequently (weekly or monthly) and can be easily configured within the Aura system.

Once the ICM has identified one or more anomalies, the RMM is invoked. The cached ICM data are used to inform the calculation of the lifetime estimation of the turbine with the actual observed data. The RMM subsequently generates the lifetime pattern for 365 days from the current point in time. These lifetime estimations reflect the survival probability of the turbine if a maintenance operation is performed now. The RMM module considers four different maintenance actions: no maintenance, small repair, large repair and major replacement corresponding to the maintenance intervention categories

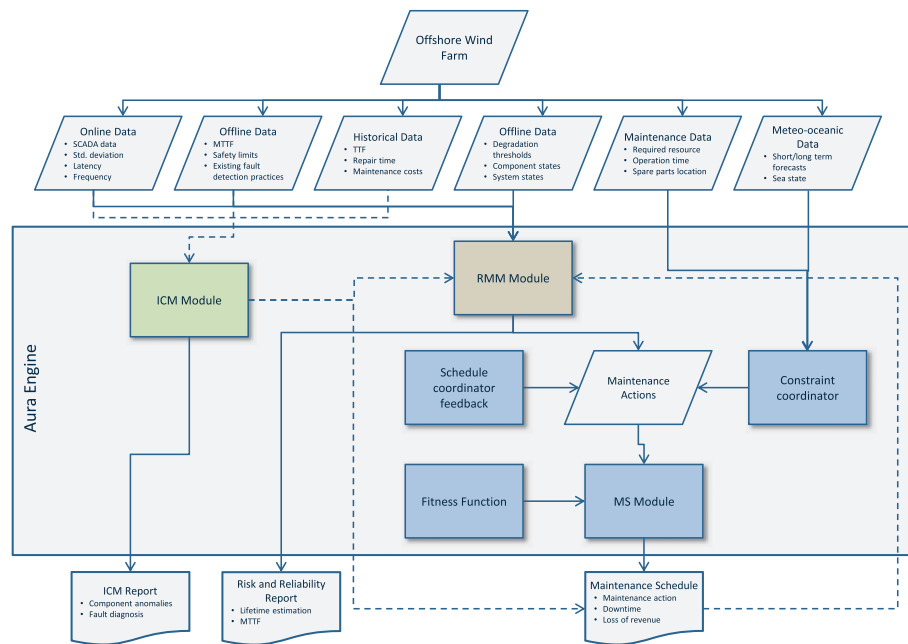


Figure 7. Data flow within the Aura system.

previously discussed. Each of these maintenance actions increases the lifetime of the turbine by varying degrees, with no maintenance having no effect, and major replacements assumed to return the turbine to near original condition.

Once these maintenance actions and survival probabilities are generated for each of the anomalous turbines, the MS module within Aura is invoked. The MS module uses these survival probabilities, information relating to existing planned maintenance, meteo-oceanic and resource constraints in order to generate the maintenance schedule. The MS module considers the different maintenance action options and optimizes the costs associated with the revenue generated for each anomalous turbine taking into account losses associated with equipment, access, resources and production lost. The MS module returns a decision in terms of which maintenance action provides the most economically viable choice. These maintenance actions for each of the anomalous turbines are added to the existing maintenance schedule, which is then updated within the Aura system.

The RMM is then re-invoked in order to provide more accurate survival estimations, given that the maintenance schedule indicates that maintenance actions for the anomalous turbines are scheduled for specific days. This step re-estimates the survival probability for these turbines knowing that a particular maintenance action will occur at some point in the future.

The final maintenance schedule and lifetime models can then be inspected by the wind-farm operator. Each module generates individual reports corresponding to the function that they provide. These reports include feedback inputs for the model, e.g. was the scheduled action actually executed, and what maintenance action was performed once on site. The document from the ICM module is used for evaluating the effectiveness of the additional condition monitoring equipment on the system. The RMM module report is used to assess the 'true' age of the system and to compare this to the chronological age. This is important for operators who were keen to manage wear-out failures.

6. CASE STUDY

As outlined in Section 1, the purpose of constructing a holistic asset management system was to enable better control of large numbers of assets. To demonstrate this behaviour, this section details a case study in which over 100 gearboxes are monitored for anomalies and subsequently have any maintenance appropriately scheduled. As data for an offshore site were not freely available at the time of publication, SCADA data have been taken from a large-scale onshore site and applied in the context of an offshore wind farm. This refers to the relevant logistical operations being translated to an offshore equivalent. All other datasets are reflective of offshore wind farms.

A typical wind farm consists of a number of turbines, switch gear and transformers (mostly located within the wind farm) and an onshore substation. All systems and components within the wind farm need to be maintained. Turbines are typically visited twice a year, and each visit has a duration of 3–5 days. In addition to turbine maintenance, regular inspections and maintenance are undertaken for the sub-structures, the scour protection, the cabling and the transformer station.

6.1. Test case setup

The raw data used in the case study are composed of 12 months of historical SCADA and alarm logs, which were used by both the ICM and RMM modules, where the RMM module used a small subset of SCADA data relevant to the generator. These data were taken from a fleet of over 100 onshore gearboxes, which have been operating for approximately 3 years. However, in the interest of demonstrating the behaviour of the system, some turbines were assumed to be newly commissioned, while others were assumed to be several years old and others largely degraded. These assumptions were reflected in the output of the RMM and ICM modules. The following sections shall expand upon these details where relevant.

Table II illustrates the different sources of data, including structured expert judgement that was used in AURA. This includes various maintenance intervention scenarios, covering resources required for specific maintenance tasks and their respective costs. The parameters for the Weibull distribution used by the RMM module, which represents the lifetime survival probability of a component, have also been elicited, to conform to the expected behaviour of an offshore wind gearbox.

For ease of presentation, this section shall only enumerate the results of specific gearboxes, which offer insight into holistic operation. These turbines are representative of the fleet as a whole, in that they reflect turbines of varying ages and are known to exhibit a variety of interesting faults, which can be categorized as trivial, minor or major.

As stated, the case study used historical SCADA data as input to the system. Naturally, this leads to difficulties in evaluating the performance of both the RMM and MS modules, as the observed SCADA data will not necessarily match the expected behaviour of a turbine once maintenance operations have been executed. That is, if the ICM module indicates that a gearbox is in an advanced state of degradation, which in turn causes a large-scale maintenance operation to be scheduled and then executed, the subsequent SCADA samples will still demonstrate the degraded behaviour of the original gearbox.

There are two possible solutions to this problem. The first was that the system can actually be deployed and utilised on a real-world wind farm, or can at least be accessed by all maintenance performed on the site. The second was that a model of turbine component behaviour could be developed, which can then be sampled, and could move the system towards model-based reasoning.

In the context of demonstrating the approach, accessing maintenance records of the wind farm was challenging, while the latter case was simply intractable—if such a model could be constructed, it would negate the need for the ICM module. Therefore, the results presented in this section are done so with the understanding that they cannot be empirically verified without further study.

The Aura system was initially provided with information to generate a representation of the existing state and maintenance plans of the wind farm. First, the ICM module was provided with a historical dataset containing relevant SCADA and alarm logs, which were used to construct the random forest associated with each turbine using the process described previously. The training of this model was performed offline, prior to system initialisation. The size and breadth of this dataset were sufficient to contain a wide variety of turbine conditions, such that the resulting model can produce accurate classifications. Therefore, at least 1 year of observations are recommended. Once trained, the individual classifiers in the RF were used to estimate the state of any SCADA data fed in. Figure 8 visualized this process for a simple RF containing three underlying classifiers.

A subset of this SCADA data was also used by the RMM module to train the KF. Figure 9 shows an example of the mapping of the KF to the observations of turbulence intensity and generator RPM. The solid green line corresponds to the KF, which illustrates the predictions for the turbulence intensity and generator RPM, once the KF has learnt the patterns. The age of each turbine within the wind farm was also provided to the RMM module, in order to derive the correct component survival distribution. In Figure 10, we can see the effects of different ranges of turbulence intensity on the survival probability curve. The survival probability decays much faster under high turbulence intensity. The RMM module initially generated a survival probability for each turbine within the wind farm, and using the turbine age, it was possible to predict the survival probability for each turbine.

The MS module was provided with the detail relating to existing scheduled maintenance, as shown in Table III, in which seven turbines have an annual inspection, which takes place midway through the year in June/July.

Table II. The source of the various inputs for each module.

Module	Raw data	Expert derived	Aura derived
ICM	SCADA, alarms	—	—
RMM	SCADA (partial)	Weibull parameters	ICM output, MS output
MS	Resource availability	Offshore logistics/maintenance operations and costs, annual meteo-oceanic conditions	RMM output

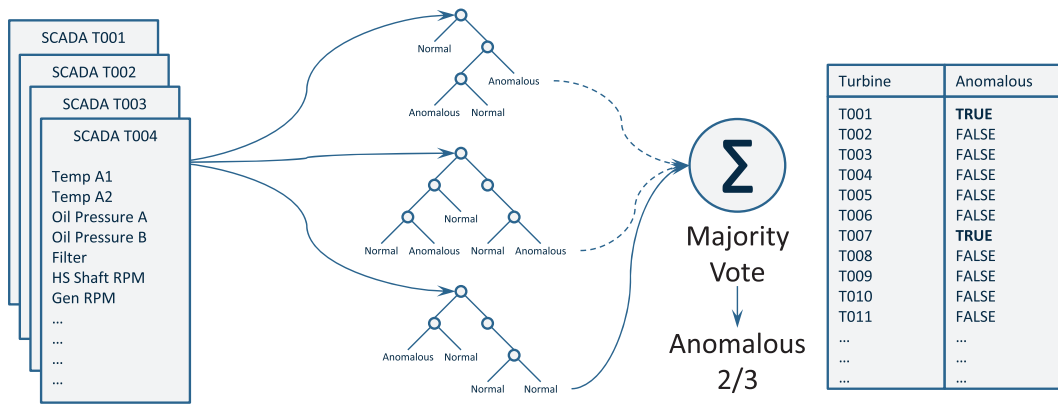


Figure 8. Fault detection in the ICM module, using the RF model associated with the turbine.

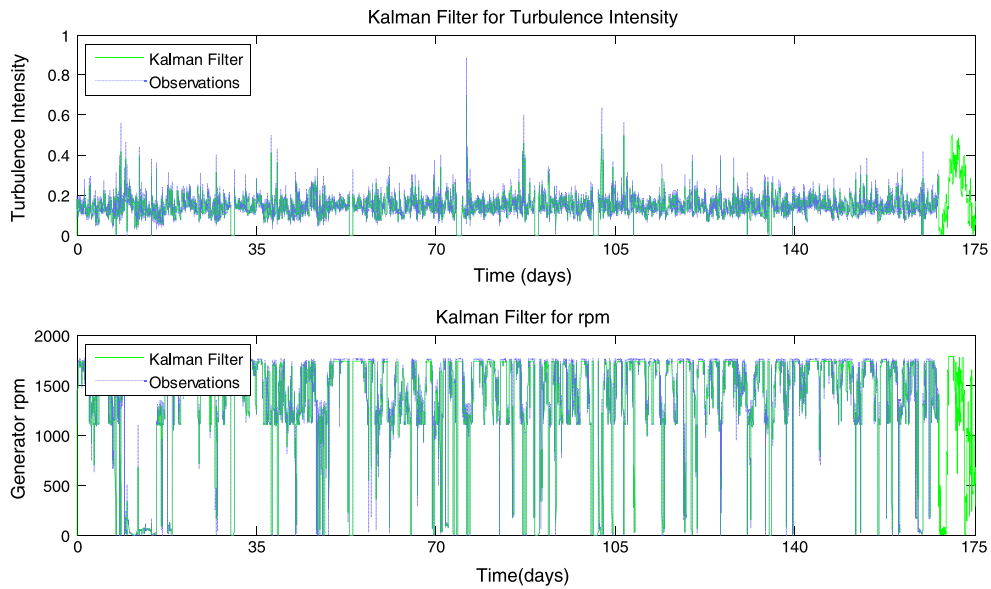


Figure 9. KF mapping to observations of turbulence intensity and generator RPM.

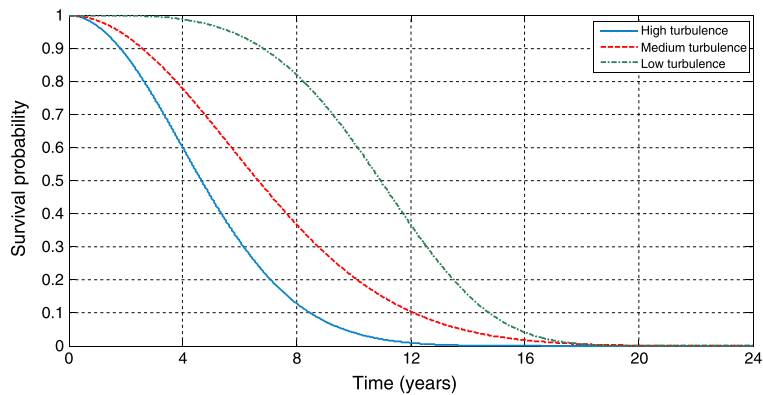


Figure 10. Turbulence intensity impact on the survival probability curve.

Table III. Annual maintenance.

Turbine	Maintenance	State	Component	Transportation	Schedule date	Duration (day)
6	Annual service	Degraded	N/A	Workboat/helicopter	Day 148	0.3
9	Annual service	New	N/A	Workboat/helicopter	Day 149	0.3
15	Annual service	Mid-life	N/A	Workboat/helicopter	Day 150	0.3
29	Annual service	Degraded	N/A	Workboat/helicopter	Day 151	0.3
56	Annual service	New	N/A	Workboat/helicopter	Day 152	0.3
85	Annual service and medium repair	Degraded	Gearbox	Workboat/helicopter	Day 153	1
101	Annual service	New	N/A	Workboat/helicopter	Day 154	0.3

Additional information for each of the maintenance interventions was provided to the MS module relating to the component cost, number of personnel required to perform the maintenance task, the transportation cost, the mean time to repair, and the wind and wave limits for the maintenance intervention. This information formed the basis for the MS module to consider current availability of personnel and transportation in scheduling additional maintenance.

6.2. System flow

Once this initial data were supplied, the system was launched, and the process of streaming SCADA data from remote sources begun. In the case study, streamed data were assumed to arrive every 10 min but was not required to do so—any gaps were ignored. Figure 11 shows a high-level overview of the fault detection and resolution process as data flows between modules.

As data entered the Aura system, these were logged to a database for historical access, before being passed onto the ICM module. Once inside this module, they were delegated to the appropriate model (e.g., data from Turbine 29 were delegated to the bespoke fault detection model for Turbine 29). The output of this module was a binary fault classification, indicating whether the SCADA sample was normal or anomalous. If it was anomalous, the sample was further analysed by the ICM module to attempt diagnosis of the fault. Regardless of this classification, the resulting labelled SCADA sample was saved to a temporary turbine-specific buffer until the RMM module was triggered.

The RMM module used the observations taken from the SCADA data and ICM output and adjusted the lifetime model of the turbine based upon this concrete evidence. This enabled the operator to view the expected life remaining on a turbine, based upon observed data rather than theory alone.

With the production of the lifetime estimation curves complete, the RMM module then computed how the probability of survival was affected by performing a maintenance intervention at the current time. That is, for each possible maintenance operation, a further lifetime estimation was produced based upon the assumption that the maintenance task was executed. Additionally, the original lifetime estimation was also included as an example of performing no maintenance (no improvement in component state).

The final part of the workflow was for the MS module to schedule appropriate maintenance for the turbine, given the context provided by the RMM module estimations. Table IV shows the new maintenance schedule produced at the end of month 3. The table illustrates the maintenance action to be performed (which may be no action required), the component on which the action may be performed, the transportation required, and the date and duration for the maintenance for each turbine. In addition, Table IV indicates the impact of undertaking no maintenance for each turbine, which was calculated on the basis of the revenue generated within the operating window, and was used as a basis to determine the economic impact of performing maintenance. The expected profit accounts for the revenue generated as a result of the specified scheduled maintenance.

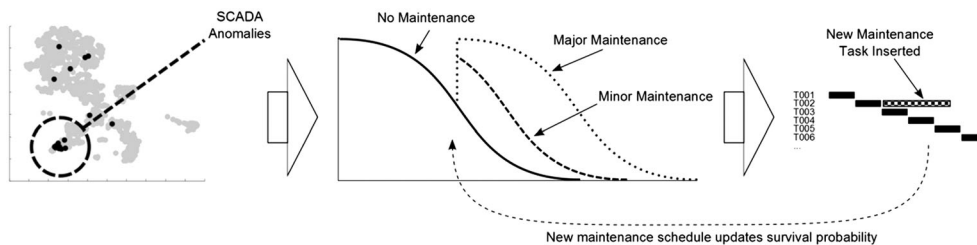


Figure 11. Fault detection and resolution in Aura.

Table IV. Scheduled maintenance.

Turbine	Maintenance	Component	Transportation	Schedule date	Duration (day)	No action	Expected profit
6	Major replacement	Gearbox	Jack-up vessel	Day 165	5	£715k	£1.46M
9	No action	N/A	N/A	N/A	N/A	£1.81M	£1.81M
15	Small repair	Gearbox oil pump	Workboat	Day 258	1	£1.08M	£1.18M
29	Medium replacement	Gearbox bearing	Workboat	Day 165	1	£951k	£1.14M
56	No action	N/A	N/A	N/A	N/A	£1.78M	£1.78M
85	Major replacement	Gearbox	Jack-up vessel	Day 160	5	£664k	£1.10M
101	No action	N/A	N/A	N/A	N/A	£1.81M	£1.81M

The schedule was produced using the techniques outlined in Section 4, whereby tasks were assigned based upon their cost, and revenue lost from associated downtime. In general, this means that maintenance tasks were not assigned to be executed during winter months when wind speeds and revenues were at their highest, and meteo-oceanic conditions were at their worst.

Table IV indicates that turbines 6, 29 and 85, having an advanced state of degradation, have benefitted from major and medium replacements with significant increases in revenue as a result of the maintenance action. Turbine 15 has a lesser degree of degradation but benefits from a small repair. Turbines 9, 56 and 101 have no maintenance action planned other than the annual inspection owing to being new relatively turbines. The benefits are measured here with respect to the expected additional income from the turbine. An alternative approach would be to use the Levelised Cost of Energy, which would require the management of additional information relating to investment, infrastructure and all operational expenditures.

With a schedule constructed and tasks assigned, control returned to the RMM module, which again updated the estimated lifetime for each turbine based upon any newly assigned future maintenance tasks.

Once the MS module has created a schedule and the RMM module has completed the final update process and modified the turbine lifetime estimations, Aura passed the control back to the ICM module. Aura returned to monitoring the incoming SCADA data and buffering data in preparation for the next invocation of the RMM and MS modules, with any maintenance tasks, which execute during the buffer period similarly cached in order to update the RMM models with observed maintenance (as opposed to only observed SCADA and ICM data). This ensured that the component lifetime was indeed extended by the respective task, even if the RMM module was not immediately invoked.

7. CONCLUSION

A novel architecture and system for the provision of RCM within the offshore wind power generation sector is presented. The architecture was created through a consideration of the data that are available from the various online and offline sources, and how this could be used to support the development of an intelligent, autonomous system. The online sources specifically related to the SCADA and alarm data gathered from the wind farm, as well as the data to be generated within the system. These data were augmented with offline data relating to existing fault detection practices, MTTF, component dependencies, scheduled maintenance and meteo-oceanic conditions, for example. The functionality required for the provision of an intelligent integrated maintenance system related to fault detection and diagnosis, health and survivability monitoring, and MS.

The system demonstrates for the first time the use and integration of state-of-the-art artificial intelligence techniques: RF for condition monitoring, DBNs for RMM and a mimetic algorithm for MS. The Aura system was constructed to respond to the demands of maximizing the revenue created from an offshore wind, operating between 13 and 195 km from shore, by ensuring the availability of the associated assets through a preventative maintenance regime. The system was developed as three modules, corresponding to the required functionality, and integrated and controlled within the Aura engine.

A case study using 12 months of SCADA and alarm data from a wind farm consisting of over 100 turbines was used to demonstrate the system. Specific focus was placed upon seven turbines, which were known to exhibit a range of different anomalous behaviours. The case study demonstrated the following: the automatic detection of faults within these turbines; the impacts these faults had on the survivability of the turbines; the assessment of different maintenance actions with the objective of maximizing availability; and the scheduling of maintenance and the updating of turbine survivability in response to the maintenance action.

The purposes of the case study were to illustrate the integration of the three methods and to demonstrate the potential of the framework. Further research is required to validate and refine parts of the modelling. For example, the data that were used in the case study only covered a 1 year period in the turbines' lifetime. A fuller study is required to evaluate the choice of the Random Forrest and the Weibull distribution over a longer period of data, and to assess to what extent ERG captures the degradation of a wind turbine.

ACKNOWLEDGEMENT

The authors would like to acknowledge the funding received that enabled this research to be undertaken. This project was funded by EPSRC KTA (project reference EP/H50009X/1) and SgurrEnergy. The opinions are those of the authors and should not be construed as representative of the views of any of the organizations that contributed towards the research.

REFERENCES

1. Spinato F, Tavner PJ, Van Bussel GJW, Koutoulakos E. Reliability of wind turbine subassemblies. *IET Renewable Power Generation* 2009; **3**: 387–401.
2. Nielsen JJ, Sørensen JD. Bayesian networks as a decision tool for O&M of offshore wind turbines. In *Fifth International ASRANet Conference, Integrating Structural Analysis, Risk & Reliability*, Edinburgh UK, 2010.
3. Van Bussel GJW, Zaaijer MB. Reliability, availability and maintenance aspects of large-scale offshore wind farms, a concepts study. In *MAREC 2001 Marine Renewable Energies Conference*, 2001.
4. Tavner PJ. *Offshore Wind Turbines: Reliability, Availability & Maintenance*. The Institution of Engineering and Technology: Stevenage, 2012.
5. Catterson VM, McArthur SDJ, Moss G. Online conditional anomaly detection in multivariate data for transformer monitoring. *IEEE Transactions on Power Delivery* 2010; **25**: 2556–2564.
6. Yu J. Machine tool condition monitoring based on an adaptive Gaussian mixture model. *Journal of Manufacturing Science and Engineering* 2012; **134**: 13.
7. Yang DM, AF Stronach, P MacConnell, J Penman. Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks. *Mechanical Systems and Signal Processing* 2002; **16**: 391–411.
8. Samanta B, Al-Balushi KR. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical Systems and Signal Processing* 2003; **17**: 317–328.
9. Zaher ASAE, McArthur SDJ, Infield DG, Patel Y. Online wind turbine fault detection through automated SCADA data analysis. *Wind Energy* 2009; **12**: 574–593.
10. Kusiak A, Li W. The prediction and diagnosis of wind turbine faults. *Renewable Energy* 2011; **36**: 16–23.
11. Breiman L. Random forests. *Machine Learning* 2001; **45**: 5–32.
12. Yan W. Application of random forest to aircraft engine fault diagnosis. In *IMACS Multiconference on Computational Engineering in Systems Applications*, IEEE, 2006; 468–475.
13. Kusiak A, Verma A. A data-mining approach to monitoring wind turbines. *IEEE Transactions on Sustainable Energy* 2012; **3**: 150–157.
14. Boudali H, Dugan JB. A continuous-time Bayesian network reliability modeling, and analysis framework. *IEEE Transactions on Reliability* 2006; **55**: 86–97.
15. Sigurdsson JH, Walls LA, Quigley JL. Bayesian belief nets for managing expert judgement and modelling reliability. *Quality and Reliability Engineering International* 2001; **17**: 181–190.
16. Langseth H, Portinale L. Bayesian networks in reliability. *Reliability Engineering & System Safety* 2007; **92**: 92–108.
17. Straub D. Value of information analysis with structural reliability methods. *Structural Safety* 2014; **49**: 75–85.
18. Straub D, Der Kiureghian A. Bayesian network enhanced with structural reliability methods: methodology. *Journal of Engineering Mechanics* 2010; **136**: 1248–1258.
19. Back T, Hammel U, Schwefel HP. Evolutionary computation: comments on the history and current state. *IEEE Transactions on Evolutionary Computation* 1997; **1**: 3–17.
20. Dahal KP, Aldridge CJ, McDonald JR. Generator maintenance scheduling using a genetic algorithm with a fuzzy evaluation function. *Fuzzy Sets and Systems* 1999; **102**: 21–29.
21. Garcia MC, Sanz-Bobi MA, del Pico J. SIMAP: intelligent system for predictive maintenance: application to the health condition monitoring of a windturbine gearbox. *Computers in Industry* 2006; **57**: 552–568.
22. Hofmann M, Sperstad IB. NOWIcob—a tool for reducing the maintenance costs of offshore wind farms. *Energy Procedia* 2013; **35**: 177–186.
23. Rademakers LWMM, Braam H, Obdam TS, Pieterman RPvd. Operation and maintenance cost estimator (OMCE) to estimate the future O&M costs of offshore wind farms. In *European Offshore Wind Conference*, Stockholm, Sweden, 2009.
24. Zhai C, Lafferty J. A study of smoothing methods for language models applied to information retrieval. *ACM Transactions on Information Systems* 2004; **22**: 179–214.