SANDIA REPORT
SAND2010-8800
Unlimited Release
Printed January 2011

Using Wind Plant Data to Increase Reliability

Bridget L. McKenney, Alistair B. Ogilvie, Valerie A. Peters

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online order:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2010-8800
Unlimited Release
Printed January 2011

Using Wind Plant Data to Increase Reliability

Bridget L. McKenney, Alistair B. Ogilvie, Valerie A. Peters

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87123

Abstract

Operators interested in improving reliability should begin with a focus on the performance of the wind plant as a whole. To then understand the factors which drive individual turbine performance, which together comprise the plant performance, it is necessary to track a number of key indicators. Analysis of these key indicators can reveal the type, frequency, and cause of failures and will also identify their contributions to overall plant performance. The ideal approach to using data to drive good decisions includes first determining which critical decisions can be based on data. When those required decisions are understood, then the analysis required to inform those decisions can be identified, and finally the data to be collected in support of those analyses can be determined. Once equipped with high-quality data and analysis capabilities, the key steps to data-based decision making for reliability improvements are to isolate possible improvements, select the improvements with largest return on investment (ROI), implement the selected improvements, and finally to track their impact.

Contents

Contents 5
Figures..... 6
1. Introduction..... 7
1.1. Importance of Understanding Reliability Performance 9
1.1.1. Financial Performance 9
1.1.2. Public Sector Support and Policy 10
1.2. Current Industry Status 10
2. Data-Based Decision-Making..... 13
3. Data Collection 15
3.1. Design, Install, and Upgrade Data Collection Systems 15
3.2. Collect and Store Data 16
3.2.1. SCADA – Time Series Data 16
3.2.2. SCADA – Alarms 18
3.2.3. Computerized Maintenance Management Systems – Work Orders 18
3.2.4. Other Systems 20
3.2.5. Data Processes 21
3.3. Integrate Data..... 22
4. Data Analysis 23
4.1. Establish Baseline Performance..... 23
4.1.1. Tools 23
4.1.2. Analysis..... 24
4.2. Identify Performance Drivers 26
4.3. Determine Root Cause 27

Figures

Figure 1: Wind Industry Maturity 8
Figure 2: Wind Industry Partner Maturity Model.. 11
Figure 3: Data-Based Decision Making Loop 13
Figure 4: Data-Based Decision Making Example 14
Figure 5: Notional Data Analysis – Time Allocation 25
Figure 6: Notional Data Analysis – Power Curve 25
Figure 7: Notional Data Analysis – Unavailability Pareto 26

1. Introduction

The Department of Energy (DOE) published a 2008 report describing a scenario in which wind energy provides 20% of the United States electricity production by the year 2030. In order for wind-generated energy to reach high electrical market penetration, customers must have confidence in fleet operating performance levels. These are judged both by energy delivery and operating costs.

Operating performance is characterized by a series of key metrics which are calculated from actual operating data. Due to the number of wind turbines in a typical wind plant and the number of data points which are collected by a turbine's Supervisory Control and Data Acquisition (SCADA) system, the sheer size of available data represents a daunting task for collection, storage, validation, and analysis. In addition, turbine and balance of plant (BOP) operational data must be correlated to maintenance and repair data in order to understand and correct performance issues. Although wind industry standards for performance measurement are gradually evolving through the efforts of a number of national and international energy organizations, there is not yet a de facto standard for definition and calculation of key wind metrics. In addition, a wind turbine is a complex machine with hundreds of parts and a detailed taxonomy. There is little current guidance on the depth and frequency of needed data points required to enable the level of analysis which can accurately guide complex operations decisions.

The current maturity level of the industry focuses on adherence to preventive maintenance schedules and reaction to failures and faults. The logistics of wind generation reinforce a reactive maintenance paradigm. Wind plants are often in remote locations due to the amount of land required, wind patterns, and negative perception of wind plant physical characteristics. This requires a small maintenance force to support repairs that are spread over large distances. Difficulty of access to the turbine itself, availability of replacement parts, the requirement to work with large-scale parts, and technician availability drive lengthy repair times. Figure 1 illustrates the current and anticipated maturity of the wind industry.

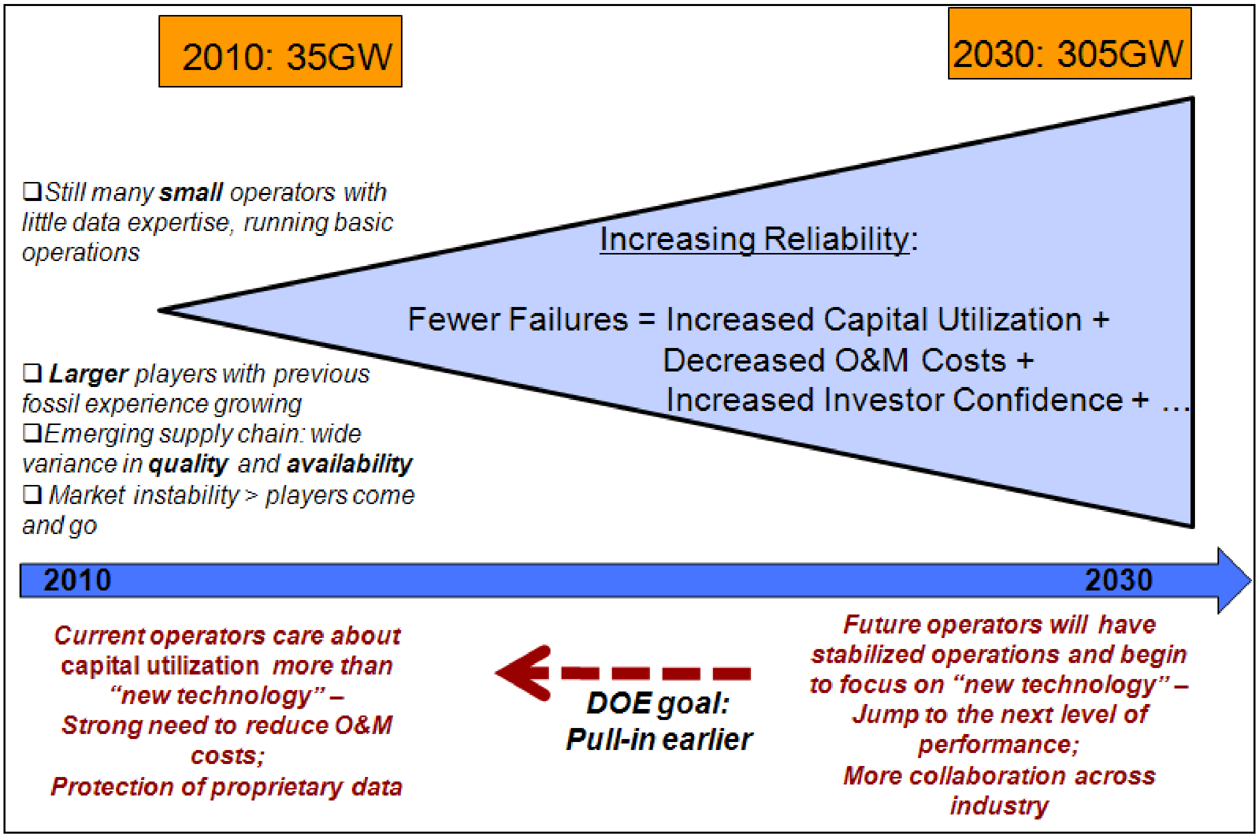

Figure 1: Wind Industry Maturity

If operational and maintenance data were readily available for analysis by support personnel equipped with robust analytical tools and reliability models, it would be possible to improve operational efficiency, with resulting increases in energy delivery and decreases in costs. Creation of failure Paretos and identification of root causes of those failures would result in a number of operational benefits, including definition of more effective preventive maintenance schedules and identification of the more frequent and/or expensive failures. This understanding would enable focus on those improvement efforts with the largest payback potential. As failure modes are better understood, operations focus could move from reactive to proactive behavior, by identifying trends and early signals which indicate likelihood of failure in advance. The highest O&M (operations and maintenance) costs are incurred post-failure, due to lost generation opportunity and repair costs. These same costs are minimized if developing problems can be addressed prior to actual failure.

The costs associated with data collection and analysis are not trivial, but are far outweighed by the benefits gained through an ability to predict future failures and minimize their impact. This paper presents recommended practices in wind plant data collection and analysis necessary to understand and utilize the performance information which is available to any owner/operator. These learnings come from Sandia National Laboratories' reliability collaboration and system analysis activity in support of the Department of Energy (DOE) mission to enable continuous reliability improvement of wind turbines in the United States. Sandia is establishing a national reliability database, to contain a sufficiently large sample of wind plants to benchmark the operation and maintenance experience of the US wind turbine fleet.

Sandia has a long and successful history supporting the wind industry, with DOE investments in wind energy research at Sandia dating back to the early 1970s. Sandia's Wind Power Technologies department has a variety of core competencies, with programs ranging from basic DOE-supported research in materials and blade concepts, to industry partnerships in reliability and manufacturing, to university systems integration studies with other laboratories and universities. Sandia's reliability program is tasked with improving the reliability of the existing wind technologies and also guiding future designs and investments.

1.1. Importance of Understanding Reliability Performance

Reliability is defined as "the ability of an item to perform a required function, under given environmental and operational conditions and for a stated period of time" (ISO 8402). Wind turbines are rated by their manufacturers for specific performance, including power output for a given wind speed (power curve). Individual turbines will exhibit statistical variance in their performance across a range of indicators. Operators should begin with focus on the performance of the wind plant as a whole, which is normally measured by power output during periods of wind availability. To understand the factors which drive individual turbine and Balance of Plant (BOP) performance, which together comprise the plant performance, it is necessary to track a number of key indicators. Analysis on these key indicators will reveal the type and frequency of high-level failures and their contributions to overall plant performance. To understand the root cause of a given failure, it is necessary to collect data at a sub-component level. In order to anticipate future failures, and replace parts prior to actual failure, it is also necessary to track and examine the conditions leading up to a failure of interest.

1.1.1. Financial Performance

One of the key reasons to track and improve reliability is to improve the plant's financial performance. Reliability can improve financial performance through cost avoidance, including operations costs and overhead costs, reducing the cost of energy. Reductions in operations costs can come from a variety of activities, such as decreasing unscheduled maintenance and translating costly parts replacements into less costly scheduled maintenance. Overhead decreases can result from benefits such as reduced insurance costs based on decreased performance uncertainty and reduced investment in spare parts inventories. Improvements in reliability can also benefit the bottom line by increasing revenue, which maximizes the plant's capital investment (e.g., by increasing turbine availability during useful wind).

Secondary financial impacts can include such varied benefits as freeing up capital to allow investments in other programs, providing accurate input to inform long-term investment decisions, and improving performance in other areas of the company by establishing a culture of using data to drive improvements. For example, a good understanding of the impact of curtailment may impact the location choice for a company's next wind plant or understanding the causes of plant-to-plant variability could inform the next turbine selection.

1.1.2. Public Sector Support and Policy

Capturing high-quality reliability data has benefits beyond the plant. It enables automated reporting to various regulatory entities, including the North American Electric Reliability Council (NERC) and Independent System Operators (ISO). It also creates a foundation for reporting to partners, such as parties in a Power Purchase Agreement (PPA) or joint-venture.

In addition to the benefits for an individual plant or operator, collective reliability data for the industry can bolster support for wind-friendly policy from federal and states government, public utility commissions, tax and rate payers, and the insurance and investment community. Also, an understanding of national wind industry reliability performance can guide standards, Original Equipment Manufacturers' (OEM) research and development (R&D) investments, and investments from government organizations such as DOE.

1.2. Current Industry Status

Although the wind industry in the US has existed for almost forty years, it only recently began to experience rapid gains in a maturing approach to the collection, management, and analysis of data. As a collective industry, there has been a marked change in the last several years in progression from manual methods of data collection and minimal attention to maintenance data, to automated data collection and linkage of SCADA systems and Computerized Maintenance Management System (CMMS) within a data historian. A growing number of owner/operators are realizing that a deep understanding of operational performance data supports improved power output and reduced O&M costs. These data "pioneers" in the wind industry are investing in data management processes, automation, and analytic personnel.

There are several major elements in the use of data to better understand the reliability performance of a wind plant. Ensuring the validity of data is still a challenge in the industry, with issues ranging from basic recording of accurate SCADA data to more complicated challenges in data transfer and storage. Once data is collected, it must be understood and analyzed before it can be used for decision making. The industry must support the emergence of tools to facilitate improved data gathering, aggregation, and analysis. Once the analysis is performed, actions can be driven by data-based decisions which can facilitate improvement efforts in those areas with the highest potential return on investment (ROI).

There has been a growth in maturity of the wind industry with regards to data collection and analysis, but there is still significant variability in the data infrastructure and analysis abilities across the industry. Paper work orders, other manual practices, and a lack of automation (such as not storing SCADA data) are still seen, though maturing companies have shown progress in implementing SCADA historians and CMMS. The most mature companies are correlating failures and faults seen in the SCADA data with work orders and checking SCADA data for accuracy (and cleaning it where possible). As the industry further matures, the goals should be to have work orders automatically generated by SCADA systems; enter searchable electronic work orders with symptom, corrective action, and root cause; and compare turbine and plant performance to an industry benchmark. The evolving data maturity of wind plants is illustrated in Figure 2.

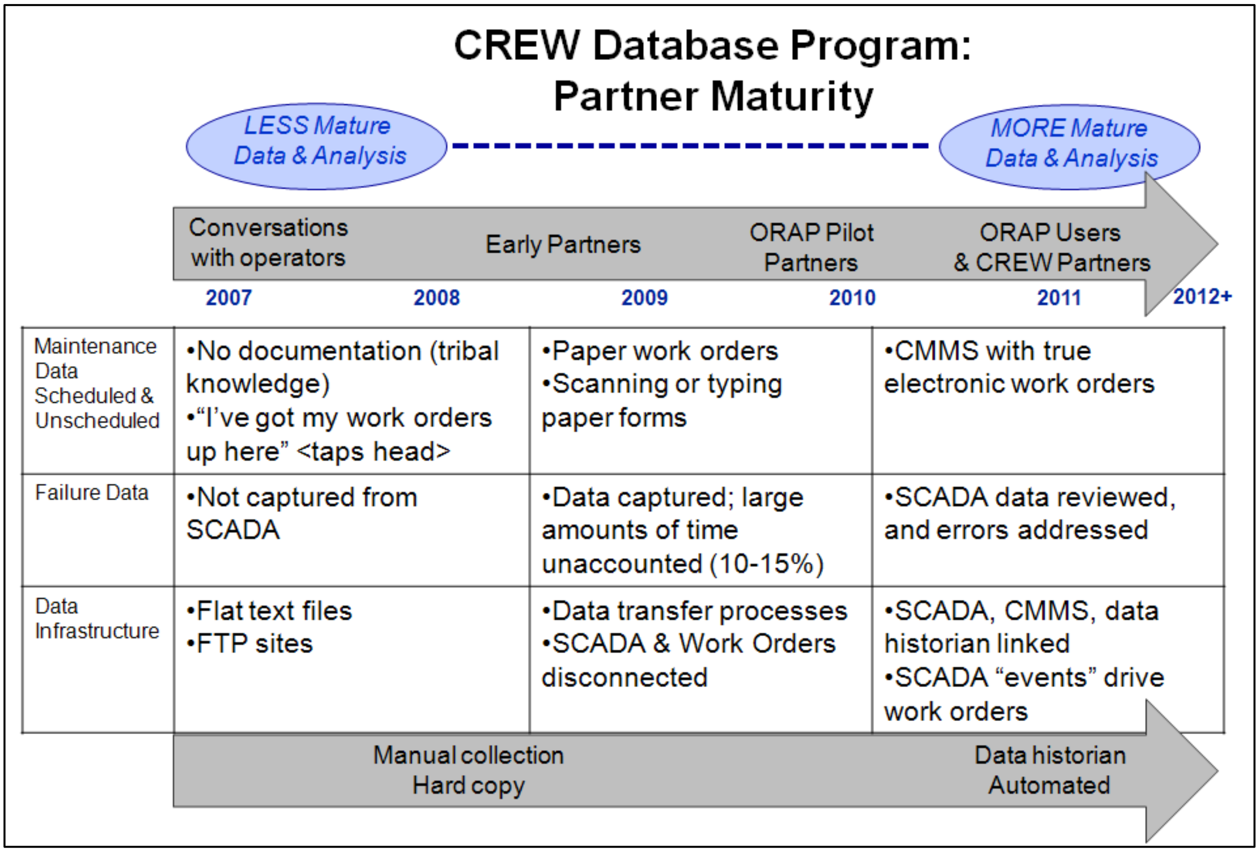

Figure 2: Wind Industry Partner Maturity Model

The infrastructure available to the operator will directly influence the way their data collection systems will be implemented, how the systems are linked, and how analysis is performed. Sandia's experience with wind plants over the past years have shown that many information technology (IT) systems and tools are available to the wind industry, but as with many other industries, the use of the available systems and tools varies greatly. Companies struggling with these tools have practices such as not storing SCADA feeds, not recording work orders even in a basic electronic logbook, and abandoning implementations of SCADA historians and CMMS before completion. These practices eliminate a company's ability to use their SCADA and/or CMMS data for information mining and analysis. On the other hand, companies successfully storing and using their reliability data are supporting their business through activities such as analysis of data that find trends and causes for downtime, and failure forecasting that drives preventive maintenance and repair schedules.

The basis for information-rich decision making is the selection of useful data points from the numerous data collection systems, at the correct level of granularity and frequency, stored centrally for effective and efficient retrieval, by knowledgeable staff, equipped with the right analysis and reporting tools. As a whole, the wind industry has beneficial, robust practices in collecting, analyzing, and reporting high-level safety and financial information for wind plants. In addition, plant-level reliability metrics are generally captured and reported at regular intervals. On the other hand, the industry has significant opportunities for improvement. Most plants do not adequately collect the information needed for root cause analysis, making identification and correction of recurring failure modes much more difficult. Monitoring the impacts of key reliability issues is sometimes lacking, and forecasting the occurrence of such problems is rare. Many plants do not monitor and learn from turbine-to-turbine variability, though most plants do at least have the ability to track performance at a turbine level.

2. Data-Based Decision Making

Collecting good data can allow a company to answer numerous questions, including ones that may not yet be articulated. The ideal way to approach the challenge of data collection is to first determine which critical decisions can be based on data in order to identify those improvements with the highest return. Once those required decisions are understood, then one can identify what analysis is required to inform those decisions, and finally identify the data to be collected to support those analyses. If the process is done the other way around (first collect data, then figure out what to do with it), the result can be a lot of non-value added work and an inability to collect the information critical to informing key O&M decisions. The data-based decision making cycle is illustrated below in Figure 3, and **Error! Reference source not found.**Figure 4 conceptually illustrates how this process would work for identifying the ideal timing of a parts replacement.

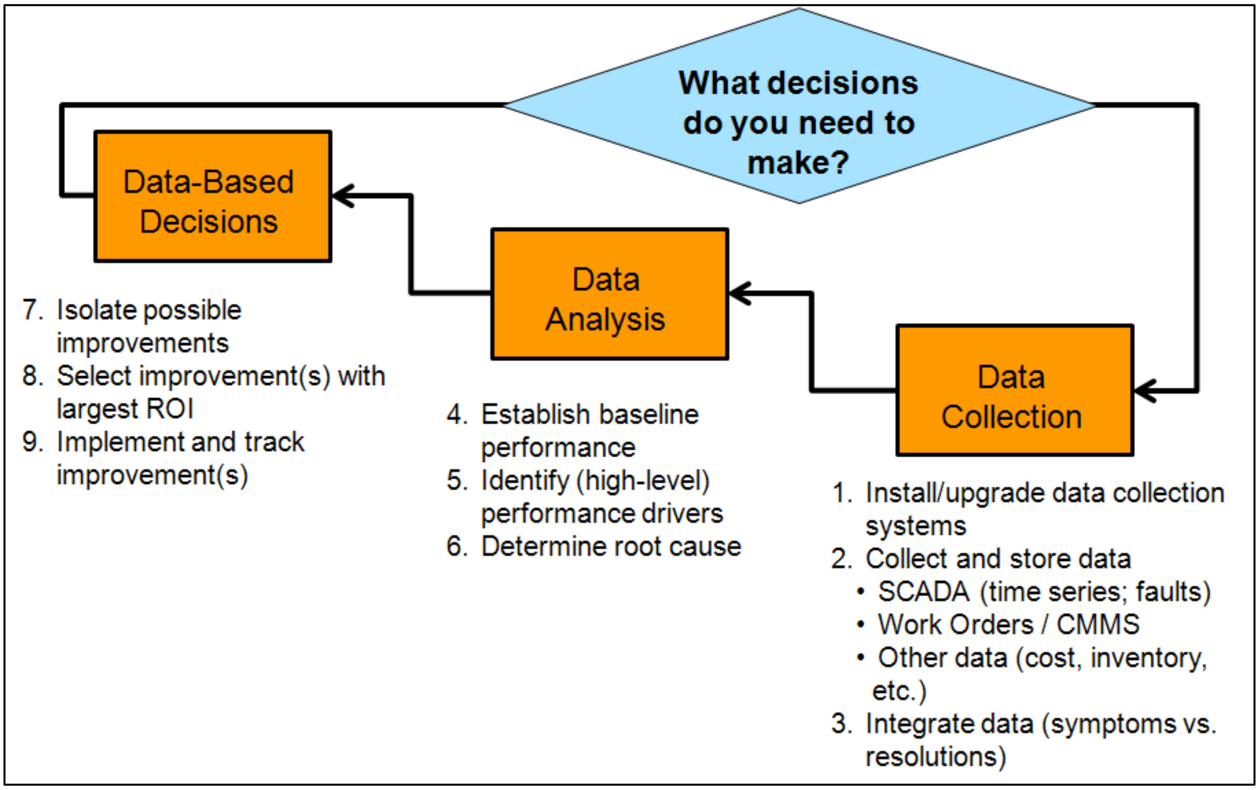

Figure 3: Data-Based Decision Making Loop

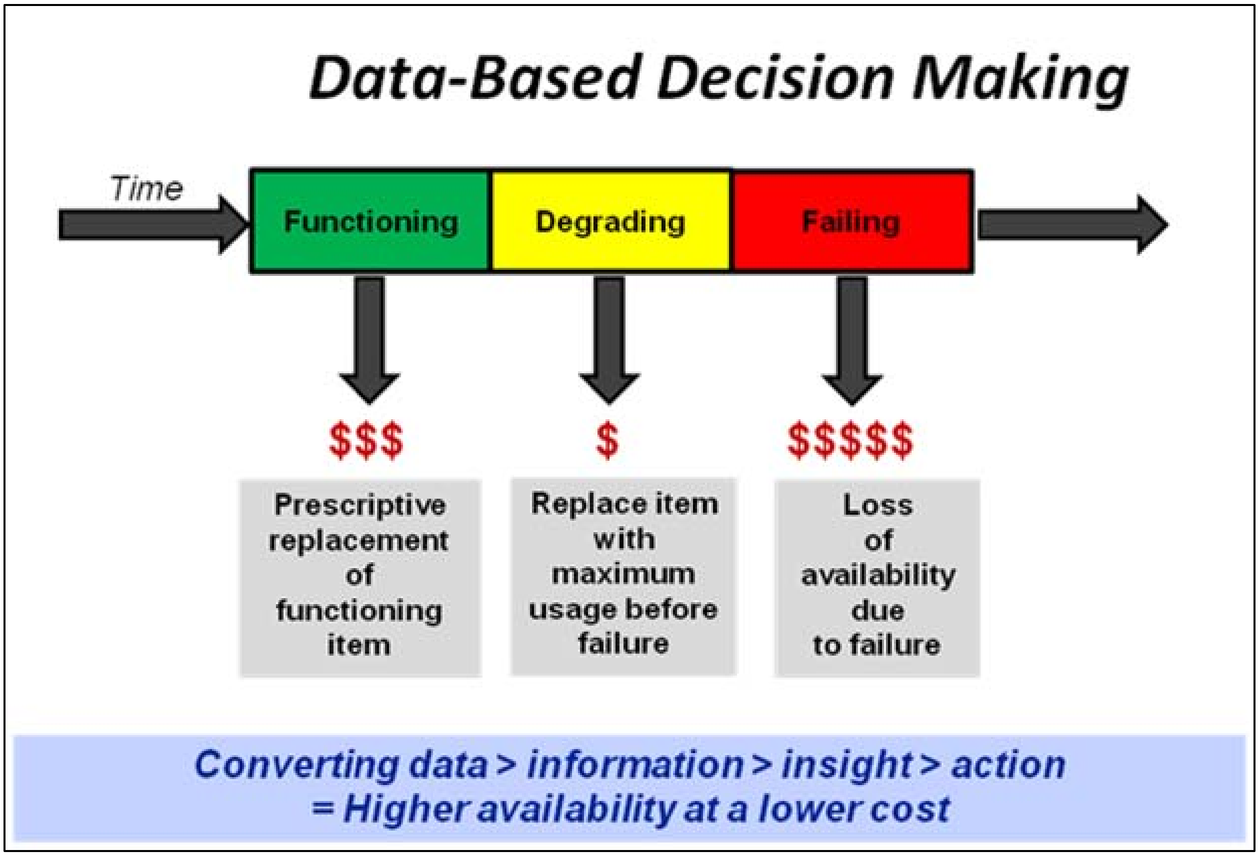

Figure 4: Data-Based Decision Making Example

Business decisions are made in a company every day, whether or not they are (or even could be) based on data. The key to successful data-based decision making is to both improve practices which have been identified as lacking and discover new areas that may not have been considered before for assessment. The decisions that need to be made that should be based on data will vary from company to company and moment to moment. Examples of such decisions include: identifying ideal preventive maintenance schedules to minimize the costs of spare parts and consumables, deciding whether or not to invest in a capital improvement project, determining which parts need to be purchased from the OEM and which can be acquired as a commodity, or changing vendors for a component that has a large contribution toward poor availability. In terms of decisions that center on reliability, the key steps are to first isolate possible improvements, select the improvements with largest ROI, implement the selected improvements, and finally to track the impact of the improvements.

3. Data Collection

The first step to information-rich decision making is storage of the vast amounts of data generated by each turbine, BOP equipment, and plant. As with any storage solution, the initial step in the design process is the most important: deciding the type and uses of information that needs to be "extracted." Designing a data collection infrastructure with the end use in mind allows for the eventual retrieval, analysis, reporting, and displaying of information in a far more efficient and effective manner.

3.1. Design, Install, and Upgrade Data Collection Systems

Wind plant operators want tools to drive their decisions with solid data. After determining what systems are most appropriate and the general type of data that should be captured, implementation is the next step. Yet implementation is not a trivial task, and requires committed and knowledgeable staff and executives. Incomplete and ineffectual implementations result in high-cost systems with few benefits, often requiring replacement or upgrade before any return on investment can be achieved. Additionally, one of the most important aspects of a data architecture is that is needs to evolve. Many good systems are eventually tossed aside or misused because they do not evolve and grow as the business changes. For companies that own or operate multiple turbine technologies, the initial design stage also needs to include an assessment of the data available from each technology and a way to map these data points so that equal comparisons can be made.

The design of the data collection and storage systems can be approached from two directions: internal resources and external consulting services. Often the use of internal skills and systems facilitates a custom approach that best suits the company and its existing infrastructure. However, realistic assessment of the skills available versus the skills required is an important first step. A hybrid approach of using external consultants to fill any gaps in in-house knowledge is also an effective solution. Once it is established that the skills needed are available, a detailed cost assessment should then be done to give a realistic cost assessment for the project's implementation.

Collecting data and collecting useful data are not the same, and this distinction is often the defining characteristic of a successful implementation versus an unsuccessful one. The vast amount of the data available can paradoxically make collecting useful data more challenging. Wind plants produce staggering amounts of data – estimated annual storage of essential SCADA data from a plant with 100 modern wind turbines exceeds 138 Gigabytes of data annually and this figure increases dramatically if every SCADA tag is stored at the highest possible frequency. Collecting too many or too few pieces of data can both result in inefficient systems that do not produce the analysis results that are expected. Decisions need to be made in advance of collection to establish the types of analysis needed, thereby insuring the collection of the data needed to complete the analysis.

While data collection for wind turbines is important, to fully understand plant reliability, data must be collected for turbines and other equipment in the Balance of Plant (BOP). Data from meteorological towers, the substation, and the electrical collection system are absolutely necessary to understand the reliability of the turbines and the whole plant. For example, turbine availability can be 100%, but if the substation is down the plant is not producing. Failing to capture such a situation will lead to large blind spots in any reliability analysis.

3.2. Collect and Store Data

In any data storage scheme, the structure of the whole is as important as the structure of the individual parts. Each piece of data needs a single place to be stored and should be stored as the kind of data that it represents (e.g. a whole number should be stored as an integer, dates and times should be stored as DateTime fields). Proper storage will make calculations and retrieval on large data sets more efficient. Each field is then stored in a table which logically groups like data together (e.g. plant name and its address can be stored together). These tables are then linked to each other by their relationships (e.g. plant “XYZ” is linked to company “ABC”) to create the database as a whole. The structure, or schema, of the database is especially important as the data set grows and fast retrieval of information is needed. Note that a schema will be needed within each database, and also across systems.

While the data architecture is being designed, built, tested, and made ready to receive and store the data, the different data collection systems need to be assessed. Each of the systems needs to be reviewed for content and usefulness. Experience has shown two common approaches to wind plant data storage; relational databases and data historians. Relational database products, like Microsoft SQL Server and Oracle Database, offer storage of large data sets using relationships between the data pieces to solve storage and access management. This type of database is widely used in many implementations throughout many industries. Data historian products, like OSIsoft PI System and Invensys Wonderware, are information technology systems that store time-series data, allowing the storage of large data streams at high speed while using compression to manage the hard drive storage space needed for these millions of pieces of information. Data historians are commonly found in manufacturing, pharmaceutical, and utility industries, including wind.

3.2.1. SCADA – Time Series Data

One of two main types of information captured by SCADA is time series data on turbine, BOP, and environmental conditions. For turbines, this time series data creates the “heartbeat” of the machine. It is collected almost continuously (typically once per second or more often) and is stored in regular intervals – typically once per second, minute, ten minutes, or hour. The various data streams that are captured are sometimes referred to as “tags.” (For those familiar with traditional databases, a tag is like a database field.) These data points record the operating conditions of the turbine and its parts, as well as the environmental conditions in which the turbine is operating. Many plants choose to archive their SCADA data in a data historian.

A multitude of time series data is available from a wind turbine SCADA system, enabling a great variety of analysis. As an example, the set of tags necessary for basic reliability analysis for a turbine is:

- Turbine Status or Operating State
    - Terminology can vary widely from operator-to-operator or OEM-to-OEM, but some basic examples include: up and running, available but idle, down for repair, curtailed, manually stopped at turbine.
    - This value can be stored as a text field (usually with abbreviated versions of the state descriptions), or as an integer (with a given number mapping to a specific description). Care should be taken that if this value is stored as an integer, that it is not “translated” to a real-valued number in a historian or other database. If 1 means “Up and generating” and 2 means “Down for maintenance,” a value of 1.62 is not very useful.
    - For turbine status, a high resolution of data (data captured very frequently versus less often) is necessary to determine turbine status over the full course of a day, week, month, or year. When turbines are coming online and offline frequently, data summarized at too high a level does not provide enough visibility into the turbine’s true state. This limitation can be overcome either by collecting this data at a higher frequency, or by only capturing this data point upon change.
- Power Generated
    - Typically stored in kW, the power generated by the turbine is very useful for reporting on turbine production. It can also be a valuable “sanity check” when various data sources are conflict regarding the turbine’s actual status.
    - Most SCADA systems offer more than one Power metric (gross, net, etc.). It is important to be clear which is being reported.
- Wind Speed
    - All analyses, reliability or otherwise, of a wind plant’s operations needs to consider wind speed. At a minimum, knowledge of whether the wind is between Cut In and Cut Out wind speed or not is required for reliability analysis. Ideally, the actual wind speed (usually measured in meters per second) should be captured.
    - Typically, there will be at least two sources of wind speed data – the turbine’s anemometry and the meteorological tower. Both sources can be useful to understand what is really happening at a turbine.

Beyond those listed above, many of the other turbine, BOP, and environmental tags in the SCADA time series data will be useful at some point for root cause reliability analysis. In particular, tags that are generally useful for root cause analysis include measures of temperature (including ambient air temperature and the temperature of components) and measures of other air conditions (including wind speed and direction, air pressure or density, and turbulence).

3.2.2. SCADA – Alarms

The second type of SCADA data relates to alarms at the turbine. Events, alarms, and faults are collected when they occur (not continuously, as with the time series data). With this kind of data, information is only stored when something interesting happens – namely, events are recorded when the operating or environmental conditions of the turbine and its parts fall outside of specific boundaries. Combined with Work Orders, alarm information can help provide a complete set of downtime events for each turbine, BOP equipment, and/or plant. Ideally, any alarm that requires human intervention will also have a Work Order associated with it. As an example, turbine alarms should contain the following information as a minimum:

- Turbine Identifier
  - Recording a turbine ID links each alarm with a specific turbine.
- Event Identifier
  - Most SCADA systems have a list of a few hundred alarm types. Capturing an identifier for each alarm, then cross-referencing the meaning from a complete list of alarms and their attributes, provides much information about what was going wrong. Attributes can include useful information, such as whether an alarm can be automatically or remotely reset, and whether the alarm was triggered automatically or by human intervention.
- Alarm Start Date & Time
  - Date and time when the alarm begins.
- Alarm End Date & Time
  - Date and time when the alarm ends.

3.2.3. Computerized Maintenance Management Systems – Work Orders

Beyond SCADA storage, more advanced companies have also implemented Computerized Maintenance Management Systems (CMMS) for their Work Orders. A CMMS enables access to work order data for trend analysis, detailed parts tracking, and root cause analysis. A CMMS is a crucial, but frequently overlooked, aspect of a data collection architecture; paper work orders and technician “tribal knowledge” are ineffective sources of information about turbine, BOP, and plant performance. One of the largest analysis challenges facing the wind industry is the current dependence on manual maintenance and repair documentation processes. These are not scalable and deprive owners and operators of the crucial corrective action information that is necessary for root cause analysis. Well-written work orders can provide a gold-mine of information for a company; poorly-written work orders can be a waste of valuable technician time.

One of the cardinal rules of a CMMS (or any other data entry system requiring human input) is that it needs to be as painless as possible to do data entry. The people involved in work order data entry can vary widely, but often include technicians, administrative staff at the plant, and employees in an Operations Command Center (OCC). Other important aspects to keep in mind in designing a data entry systems is that optional fields tend to remain blank and

"Miscellaneous" is a popular choice. As an example, high-quality work orders for a turbine will contain:

- Turbine Identifier
  - Recording a turbine ID links each maintenance event with a specific turbine. Events that do not tie to a specific turbine can still be captured, but this should be clearly specified. Ideally, there will be options to choose specific BOP equipment, in addition to specific turbines.
- Event Type
  - Event type captures, at a high level, what kind of work is being performed (e.g., component failure, preventative maintenance, inspection).
  - All types of downtime and maintenance events should be recorded, including inspections and other scheduled maintenance events. Even inspections and scheduled maintenance that are relatively short in duration, relatively infrequent, and/or can occur while the system is running are crucial to understanding the availability, reliability, and financial performance of a system.
- Affected Component
  - Ideally, the affected component would be chosen from a standard breakdown of the turbine (e.g., taxonomy or equipment breakdown structure). This value may not be initially known with certainty, so a good CMMS needs to allow for updates, editing, and refinement as more knowledge is gained.
  - In order to conduct real root-cause analysis, it is also useful to capture a brief description of the failure mechanism and/or the external event that caused the downtime or maintenance (e.g., curtailment, chipped gear tooth, dirty oil).
  - For relevant event types, the source of parts is also a useful piece of information. This includes parts acquired through non-standard methods (e.g., swapped from another turbine, purchased outside supply system, machined on site). Identifying the source of parts allows for more accurate cost calculations and will allow more advanced CMMS uses such as parts and inventory tracking.
- Equipment Status
  - Not all maintenance events will stop a turbine from generating, for example, some inspections are allowed when the turbine is running. Suggested choices for equipment status include Offline, Degraded, and Online, though other options may prove useful.
- Event Start Date & Time
  - Date and time when the status of the turbine changes (or, if the turbine status does not change due to the start of the event, the date and time the maintenance event begins).
- Event End Date & Time

    - Date and time when the status of the turbine changes (or, if the turbine status does not change due to the end of the event, the date and time the maintenance event ends).
- Downtime
    - There are many ways to measure downtime. From the event start and end times, the Total Duration of the downtime or event can be captured. Other useful measures include:
        - Active Maintenance Time: The total amount of time maintenance was being actively performed on the turbine.
        - Person-Hours: The total number of person-hours required to complete the maintenance action. Note that this may be very different (greater than or less than) total downtime, and may be greater than the Active Maintenance Time if more than one technician was needed.
        - Waiting Time: Ideally, this can be broken into time spent waiting for a technician to become available, waiting for a part from supply, waiting for a piece of support equipment to become available, or waiting on other administrative delays.
- Description/comments
    - Though free-text comments can be difficult to use in an automated way, allowing technicians to capture anything unexpected or unusual about a maintenance event can be quite useful when delving deeply into specific events or types of events. In addition, this field can be helpful to support the collection of additional data while the CMMS is being upgraded to capture it in a more appropriate field.

3.2.4. Other Systems

In addition to the data that is captured from SCADA and Work Orders, supplemental turbine and plant information is also needed. Enterprise Resource Planning (ERP) systems with cost data and other information such as taxonomies provide the supplemental information needed to support data-based decision making.

Ideally, cost information would include component-level repair costs, component-level replacement costs, consumables costs (for example, the price of a liter of gearbox oil), costs associated with technician time, and costs associated with overhead (such as administrative time) if such overhead is linked to maintenance or downtime. Additionally, some plants also look at lost revenue from generation or penalties assessed for not generating.

Information on turbine and BOP configurations is another essential aspect of comparing “apples to apples” when performing analysis at system, sub-system, component-group and component levels, especially across multiple plants or turbine technologies. A hierarchical equipment breakdown structure (EBS) or taxonomy divides the turbine and BOP into their generalized parts in a parent-child relationship that allows sub-parts to be rolled up into parts, sub-systems, and systems. Once a general taxonomy is developed, then each of the turbine technologies or plants can be mapped to it, creating a standard that allows comparison. Also, a detailed description of the equipment (make, model, manufacturer of major components, presence/absence of optional systems such as de-icing equipment or condition monitoring, etc.) is important for comparisons. Lastly, documented system knowledge, such as turbine specifications or substation fault trees, can provide the basis for more advanced reliability analysis.

3.2.5. Data Processes

In the wind industry, many larger companies have implemented operations command centers (OCCs) with real-time operating data flowing from plants to a centralized monitoring and control center. The real-time data is then stored in a single large database covering multiple plants. This approach requires a robust and reliable connection from each turbine and plant to the OCC. An alternative, seen at smaller operators and plants, is where a storage system is implemented on-site and stores a finite time-period of SCADA data. Some of these implementations allow for sub-sets of the data to be sent to a central office for storage and analysis.

Those companies that do store their SCADA data consider it highly proprietary and treat it as intellectual property (IP). This adds a requirement for encryption and security during the transfer of data from the plant, and access levels and controls that restrict who can view the data once in storage. If there is transfer of the data from a plant to OCC, a T1 line seems to be the most common approach used by larger companies. This creates a "dedicated" connection between the wind plant and the OCC, making it a good choice for carrying large amounts of data, as it is both reliable and secure. Once data is stored at the OCC, the use of Microsoft Windows Authentication can fulfill all the needs for controlling access to the data.

Once the data is stored and accessible to those with the rights to see and use it, data protection becomes a primary task of the data administration staff. Design, implementation, and maintenance of a backup and recovery plan are essential to preventing the loss of data through accident, data corruption, hardware failure, or natural disaster. The plan should include levels of importance for the data, projected recovery timeframes, scheduling and monitoring of backups, on-going validation testing of the backups, and the media choices on which the backups will be stored.

In addition to backups, with the amount of data being stored for each turbine and plant, an archiving strategy is necessary to manage the size of the database and maintain a high-functioning retrieval system. One approach is to archive the raw data but to retain calculated values. Another approach is to store the data using special compression techniques that reduce the number of data pieces stored without losing the meaning of the data. Data historians are especially designed with this type of compression in mind. With the cost of storage coming down over the last decade, archiving strategies should be periodically re-evaluated to ensure that the correct levels of data are available for analysis.

When setting up transfer and storage protocols, the data to be stored must be determined. Because some data will invariably NOT be stored, there is the potential to reduce the value of the data set. This concern is especially relevant when looking at the need for summarizing the voluminous SCADA time series data. For some analysis, such as monthly or yearly performance metrics, summarized data is sufficient. However, more detail may be needed for root cause analysis, making compression highly beneficial when compared to summarizing. Compression means reducing the number of electronic bits that represent a piece of data, thus reducing the number of bits that need to be transferred from the plant and stored. For example, only storing values when they change can save a great deal of space if that data does not change often.

One of the other aspects of data integrity is addressing missing or illogical data, with data validation serving as an essential aspect of any data collection system. When there is only a single piece missing, it will likely have little to no impact on analysis. But when larger amounts of data are missing, perhaps covering hours, the loss may be important. The practice of data editing, or filling in the data with realistic values, can assist in creating a complete data set. For illogical data, values for a piece of data can be compared to previous values or sets/ranges of acceptable values, allowing an unrealistic value to be identified. Care must be taken with data editing, however as it can reduce confidence in the data as a whole. Among other challenges, important signals can be missed if unexpected, but accurate, values are overwritten; also filling in unknown values can mask a data communications problem.

For all of these data integrity concerns, their impact can be reduced by implementing good business processes and procedures, where all employees follow the same process when dealing with the data. Whether the employee is at the plant, or in the corporate IT department, or in the analysis group, business processes allow for the same methodology to be implemented, and for necessary improvements to be implemented systematically. Part of these remedies should be a standard approach to the use and interpretation of data. This creates an environment where comparisons between turbines and plants can easily be made, because the analysis is based on the same assumptions about the data.

3.3. Integrate Data

One of the greatest challenges in using CMMS and SCADA data to perform reliability analysis is in matching work orders, SCADA time series, and SCADA alarm data. This linking of symptom (for example, high SCADA temperature recordings followed by a gearbox over-temperature alarm) to corrective action (a work order to replace a lubrication oil pump) allows for the beginning stages of root cause analysis, parts tracking, and trending. An automated (or semi-automated) method for performing this linking will greatly improve the detail and accuracy of reliability analysis, but it is not an easy process. Challenges in linking data can include conflicts between CMMS and SCADA regarding turbine status, incomplete work orders, and missing SCADA data; additionally, real situations that are difficult to interpret will appear, such as curtailment, overlapping work orders, and back-to-back alarms.

4. Data Analysis

Just as any good data collection system will evolve, the overall process of using data for analysis and decision making will also change and grow. New questions will arise that require new data or analysis methods, new data will become available that would be ideal for answering existing questions, or better business and analytical processes will be developed.

With reliability analysis, or any analysis that attempts to improve the understanding of the current situation and make predictions, it is tempting to jump right into problem-solving. But, a better method is to "peel the onion" – first establish baseline performance to understand what the current situation is, then identify performance drivers, and finally determine root causes.

4.1. Establish Baseline Performance

4.1.1. Tools

Though analysis tools are not frequently discussed as a key aspect of a reliability program, they are an integral part of the decision making process. Their accessibility, ease of use, inherent assumptions, and flexibility will all affect the outcome of any analysis performed. In general, there are three types of analysis tools used for any kind of analysis: 1) basic tools, 2) custom analysis systems, and 3) commercial-off-the-shelf (COTS) systems.

Basic, flexible analysis tools include general programs such as Microsoft Excel, Wolfram Mathematica, and desktop databases. While these tools are very flexible, they are not pre-configured to assist with specialty analyses such as reliability analysis. These kinds of tools are best for developing new ways of exploring and reporting on data.

Custom analysis systems that are built and maintained for a specific company allow for analysis and reporting to be conducted in exactly the way that the company wants. Their major downside is that they require a significant investment of time, money, and mental capital – a company must have access to knowledgeable staff and contractors that have both the IT and analysis knowledge to build a useful, accurate, robust system.

COTS systems (such as PTC's Relex, ARINC's Raptor, and Reliasoft's suite of tools) can be a great way to get started with more advanced analysis once a company has exceeded the capabilities of the basic tools, but is not ready or is not able to invest in a custom analysis system. The benefit of COTS systems is that another organization takes on the overhead of building and maintaining the systems. The downsides are that users will have less knowledge about what is happening "under the hood," will not have large influence on improvements to the system, and will have to work harder to maintain or grow in-house analysis knowledge about when and how the tools should be applied. It has been Sandia's experience that care must be taken in using COTS systems for analysis of wind plants. The uniqueness of power generation, especially variable generation such as wind energy, can result in illogical output from COTS systems if they do not have appropriate assumptions and data interpretation.

4.1.2. Analysis

Successfully answering questions such as "What is the current performance?" and "How good is it?" is the first step to making improvements. This will point toward problematic areas on which to focus. Skipping this step may lead to solving "flavor-of-the-month" problems, versus fundamental problems with the turbines or other systems. Examples of questions and analysis related to baseline performance point include:

- What is the baseline performance?
  - Calculate basic operations and reliability metrics, such as Availability, MTBE (Mean Time Between Events), and Mean Downtime.
- How does the plant performance compare to the OEM or financial expectations?
  - Create a time allocation graph, to identify how a typical turbine spends its time (what percent of the time is it running, idle and available, down for scheduled maintenance, curtailed, etc.); be sure to identify when the turbine state cannot be determined (such as when SCADA communication is lost or the historian briefly stops recording). (See Figure 5)
  - Draw a power curve, and use both the "pretty data" and the outliers. (See Figure 6)
  - Calculate turbine capacity factors.
- Are the data aspects of the operations and maintenance business processes well understood?
  - Make and document assumptions about the data being gathered, and how it is gathered, stored, and used for analysis. (It can be surprising when one department has totally different assumptions about the data. For example, if 5% of SCADA records are not usable, there are multiple assumptions an analyst could make including assuming that the turbine was up, assuming it was down, or excluding that time all together.)

As an example of Basic Analysis, the "Time Allocation" graph in Figure 5 shows how a typical turbine spends its time. This particular graph is notional (i.e., it does NOT contain data from any real plants), but serves to illustrate the kind of insights that can be found through establishing baseline performance. Some pieces of the pie (green, yellow, and orange) show generation that is generally well-understood (though some of the slices may raise questions, such as Low Generation for Rated Wind). The blue and red pie pieces represent non-generating time; these pieces break down the status of the turbine and the wind speed during times of non-generation. The three black/gray pieces represent time that cannot be fully accounted for, including when the turbine is transitioning from one state to another with infrequent SCADA data collection, when the turbine has an illogical state or environmental condition, and time that is simply missing from the data. Another example of Basic Analysis is a Power Curve. As the notional data in Figure 6 shows, both areas of underperformance and overperformance can be important to understand.

Usually a good power curve points to areas of inconsistency (and therefore potential improvement). Occasionally, a power curve reveals data collection inaccuracies.

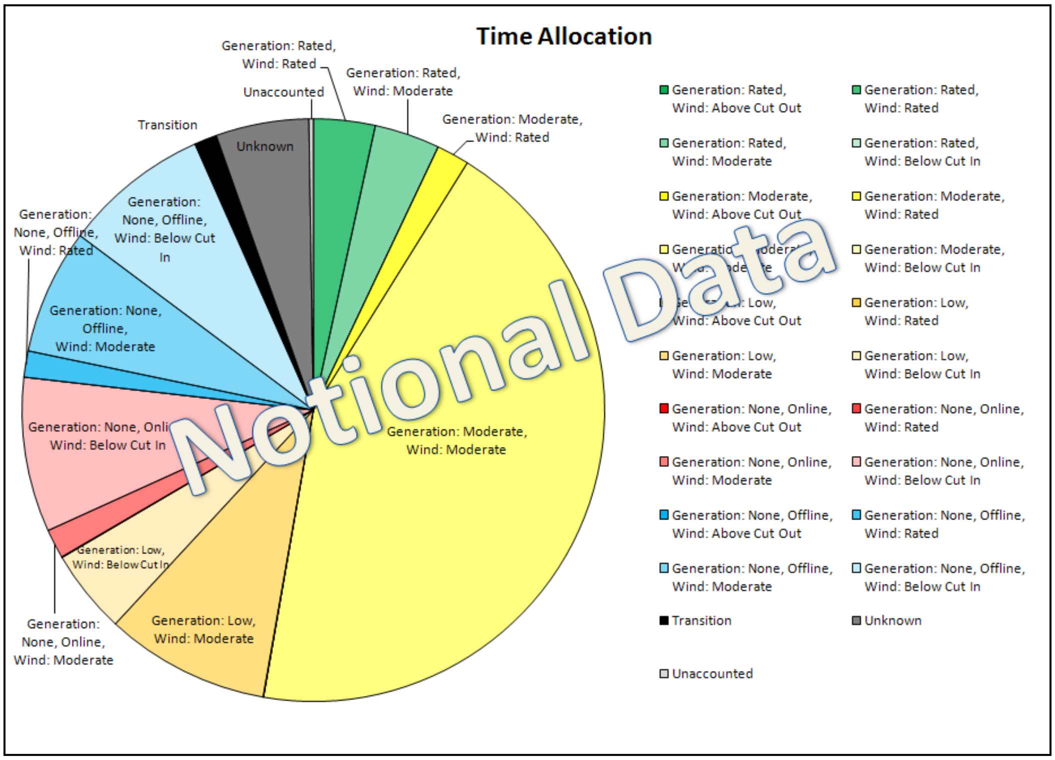

Figure 5: Notional Data Analysis – Time Allocation

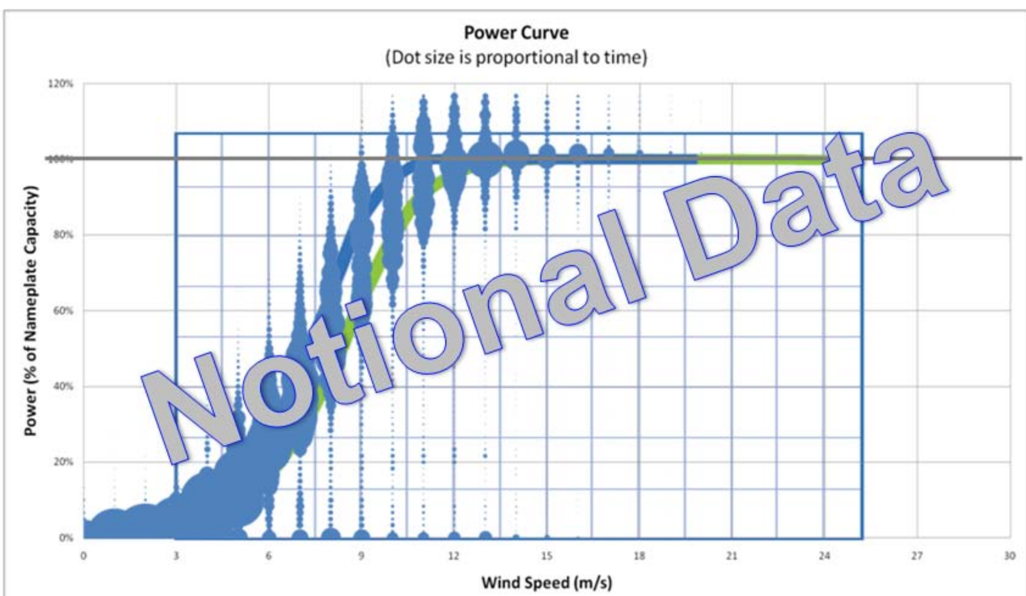

Figure 6: Notional Data Analysis – Power Curve

4.2. Identify Performance Drivers

Once baseline performance is understood, then performance drivers can be found. Methods for identifying these drivers can include exploring trends, outliers, good performance, and surprising results. Examples of questions that can be answered at this point and their related analyses include:

- Which are the good turbines? Which are problematic?
  - Identify underperforming and overperforming turbines, looking at all the various metrics, graphs, and analysis developed as part of the baseline.
  - Identify turbines that cause problems with data or interpretation of data.
- Which types of downtime events are driving poor performance?
  - Identify key contributors to low generation, unavailability (downtime), etc. Pareto analysis is a simple, but very powerful method for identifying large contributors. (See Figure 7)
  - Compare multiple metrics, such as a graph of event frequency versus event duration by component or generation versus turbine wind speed. The outliers are especially interesting in these types of graphs.
- Where is performance roughly the same? Where is there great variability?
  - Explore turbine-to-turbine performance and variability in all the various aspects covered in "The Basics."
  - Explore trends (daily, weekly, monthly, and seasonal). Plot graphs of the metrics over time. Look at the whole plant, and also look at individual turbines or individual event types.
- Where are the business and data processes different?
  - Addressing any inconsistencies in data processes and assumptions, including determining if there a valid reason for doing things differently.
  - Understand limitations in the data systems, analysis/modeling, and reporting.

As an example of Identifying Performance Drivers, Figure 7 illustrates a Pareto chart of components contributing to unavailability. While this type of analysis may seem simple, it helps ensure that resources are being spent on improving the most significant contributors.

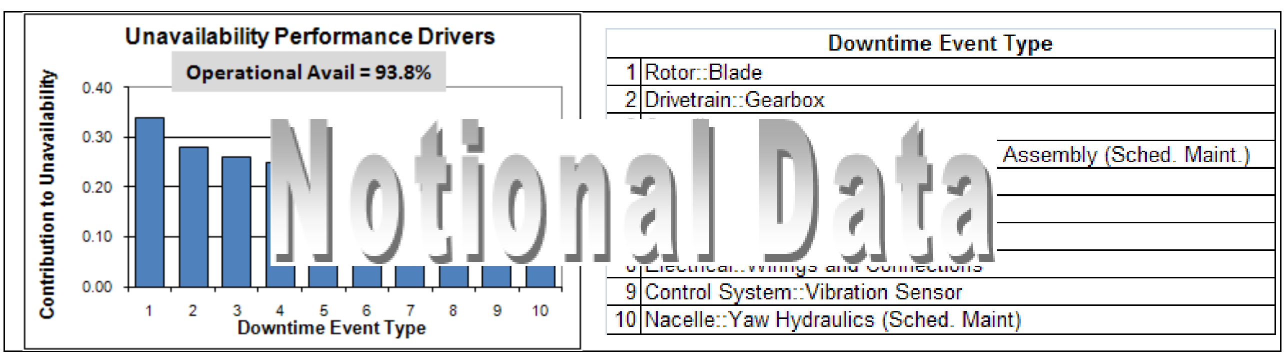

Figure 7: Notional Data Analysis – Unavailability Pareto

4.3. Determine Root Cause

After understanding baseline performance and identifying some of the key drivers, then root causes can be identified to solve problems. Examples of questions and analysis that can be addressed at this point include:

- Why are the good turbines "good"? Why are the bad turbines "bad"?
    - Explore unexpected results, both good and bad. Look for commonalities among the turbines that perform well.
- Why are certain aspects of operations (e.g., turbines or groups of turbines, months or days of the week, types of scheduled maintenance) having such a negative impact?
    - Investigate de-rates and periods of unexplained performance.
    - Interpret unexpected patterns.
- What are the root causes of the top problems?
    - After identifying the downtime events that are most impact, select one and conduct a root cause investigation. It is especially important for leadership (management, engineers, analysts, etc.) to follow through on this activity. Simply identifying potential root causes is not enough. Real fixes have to be developed, tested, implemented, and *assessed*. It will take time and other resources to determine if changes are having the positive impact that is needed.

Acronyms

- BOP: Balance of Plant
- CMMS: Computerized Maintenance Management System
- COTS: Commercial Off-The-Shelf
- DOE: Department of Energy
- EBS: Equipment Breakdown Structure
- ERP: Enterprise Resource Planning
- IP: Intellectual Property
- ISO: Independent System Operator
- IT: Information Technology
- NERC: North American Electric Reliability Council
- O&M: Operations and Maintenance
- OCC: Operations Command Center
- OEM: Original Equipment Manufacturers
- PPA: Power Purchase Agreement
- R&D: Research and Development
- ROI: Return On Investment
- Sandia: Sandia National Laboratories
- SCADA: Supervisory Control and Data Acquisition
- US: United States

DISTRIBUTION

Kurt Anderson
Upwind Solutions
3555 Lear Way
Medford, OR 97504

Larry Barr
enXco
17298 Commerce Way
Tracy, CA 95377

Dr. Phillip Barry
NA WRTC
Mesalands Community College
911 South Tenth Street
Tucumcari, NM 88401

Rashi Bates
Shell Wind Energy
150 N. Dairy Ashford
Bldg C, Room 334C
Houston, TX 77079

Fred "Doc" Beasom
NextEra Energy
7021 Oak Creek Road
Mojave, CA 93501

William Beecher
AES Wind Generation
P.O. Box 581010
North Palm Springs, CA 92258-1010

Benjamin Bell
Garrad Hassan America, Inc.
45 Main Street, Suite 302
Peterborough, NH 03458

Glen Benson
AWS Truewind LLC
463 New Karner Road
Albany, NY 12205

Daniel W. Bernadett
AWS Truewind LLC
463 New Karner Road
Albany, NY 12205

Roy Blackshear
AEP
Desert Sky Wind Farm Lp
PO Box 518
Iraan, TX 79744

Eckart Bodenbach
Winergy Drive Systems Corporation
950 Tollgate Road
Elgin, IL 60123

Ken Bolin
XCEL Energy
1225 17th Street, #483
Denver, CO 80202-5534

Dan Brake
NextEra Energy
700 Universe Blvd
Juno Beach, FL 33408

Dreama Brower
Duke Energy
71 Allen Street, Suite 101
Rutland, VT 05701

Sandy Butterfield
Boulder WindPower
209 Boulder View Lane
Boulder, CO 80304

Luis Cerezo
EPRI
1300 West WT Harris Blvd
Charlotte, NC 28262

Craig Christensen
Clipper Windpower Technology, Inc.
6305 Carpinteria Ave., Suite 300
Carpinteria, CA 93013

Michael Courville
Shell Wind Energy
150 N. Dairy Ashford
Bldg C, Room 334C
Houston, TX 77079

Martin C. Crotty
Upwind Solutions
3555 Lear Way
Medford, OR 97504

Mike Curley
NERC
116-390 Village Blvd.
Princeton, NJ 08540

Ted De Rocher
Caithness Operating Company, LLC
9790 Gateway Dr., Suite 220
Reno, NV 89521

Tracy Deadman
AES Wind Generation
4300 Wilson Boulevard
Arlington, VA 22203

Salvatore A. DellaVilla
Strategic Power Systems, Inc.
11016 Rushmore Drive
Frenette Building Suite 275
Charlotte, NC 28277

Edgar DeMeo
Renewable Energy Consulting Services
2791 Emerson St.
Palo Alto, CA 94306

John R. Dunlop
American Wind Energy Association
448 Morgan Avenue South, Suite 300
Minneapolis, MN 55405

Scott Eatherton
AES Wind Generation
14740 Altamont Pass Road
Tracy, CA 95391

Neal Emmerton
Bluarc Management Group LLC
777 Tahquitz Canyon Way, Suite 200-13
Palm Springs, CA 92262

Jeff Gibbons
AES Wind Generation, Inc.
P.O. Box 581010
North Palm Springs, CA 92258-1010

Ben M. Givens
American Electric Power
1423 CR 131, Trent Wind Farm, L.P.
Trent, TX 79561-3029

Miguel Angel Gonzalez-Posada
Gamesa
2050 Cabot Blvd. West
Langhorne, PA 19047

Randall Grayson
enXco
215 111th Street
Lake Wilson, MN 56151

Dave Harkness
Xcel Energy
414 Nicollet Mall (MP9)
Minneaplois, MN 55401

Dave Hightower
Xcel Energy
1800 Larimer Street, 8th Floor
Denver, CO 80202

Bill Holley
GE Wind
GTTC, 200-D
300 Garlington Road
Greenville, SC 29615

Thomas Jonsson
iQwind LLC
8208 NW 30th Terrace
Miami, FL 33122

Ninochska Maldonado
enXco
3896 Begonia Street
Palm Beach Gardens, FL 33410-5608

Mark McDade
NREL/NWTC
1617 Cole Bldv. MS 3811
Golden, CO 80401

Amir Mikhail
Clipper Windpower Technology, Inc.
6305 Carpinteria Ave., Suite 300
Carpinteria, CA 93013

Laura Miner
Invenergy LLC
One South Wacker Drive
Suite 1900
Chicago, IL 60606

Jim Morgan
Mesalands Community College
911 South Tenth Street
Tucumcari, NM 88401

Ahmed Morshed
Chevron
1500 Louisiana, 36062B
Houston, TX 77002

Juan Nasiff
NextEra Energy
700 Universe Blvd
Juno Beach, FL 33408

Ross Newlin
enXco
3818 Aspenwood Drive
Richmond, TX 77469

NWTC Library (5)
NREL/NWTC
1617 Cole Boulevard
Golden, CO 80401

T.S. Obdam
PO Box 1
Petten, 1755
The Netherlands

Terry Oswald
Horizon Wind Energy
808 Travis St., Suite 700
Houston, TX 77002

Tom Paff
Duke Energy
400 S Tryon St
Charlotte, NC 28285

Robert Poore
DNV-Global Energy Concepts
1809 7th Ave., Suite 900
Seattle, WA 98101

Pep Prats
Ecotecnia
Roc Boronat 78
8005 Barcelona
Spain

Mirlo Previsic
Re vision Consulting, LLC,
6630 14th Street
Sacramento, CA 95831

Kenda Ransom
Strategic Power Systems, Inc.
11016 Rushmore Drive
Frenette Building Suite 275
Charlotte, NC 28277

Dave Roberts
OSI Soft
777 Davis Street, Suite 300
San Leandro, CA 94577

Paul Rowan
MLS Electrosystem
9333 Timber Trail
Pittsburgh, PA 15237

Manny Sanchez
NextEra Energy
700 Universe Blvd
Juno Beach, FL 33408

Shawn Sheng
National Renewable Energy Laboratory
1617 Cole Bldv. MS 3811
Golden, CO 80401

Robert Steele Jr.
Strategic Power Systems, Inc.
11016 Rushmore Drive
Frenette Building Suite 275
Charlotte, NC 28277

Cece Sterling (10)
Wind & Hydropower Technologies
EE-2B Forrestal Building
U.S. Department of Energy
1000 Independence Ave. SW
Washington, DC 20585

Bahvahnie Suppiah
OGE Energy Corp.
P.O. Box 321, M/C 504
Oklahoma City, OK 73101-0321

Kedian Taborn
Strategic Power Systems, Inc.
11016 Rushmore Drive
Frenette Building Suite 275
Charlotte, NC 28277

Paul Veers
NREL/NWTC
1617 Cole Bldv. MS 3811
Golden, CO 80401

Chris Walford
Puget Sound Energy
PO Box 97034, PSE-09S
Bellevue, WA 98009-9734

Eric White
AWS Truewind LLC
463 New Karner Road
Albany, NY 12205

Simon Wright
Industrial Evolution
Deer Valley Financial Center, Suite 100
22601 North 19th Ave
Phoenix, AZ 85027

INTERNAL DISTRIBUTION:

MS	0734	E. Stechel, 6383	
MS	1033	A. Ellis, 6112	
MS	1033	J. Granata, 6112	
MS	1033	C. Hanley, 6112	
MS	1124	T. Ashwill, 6121	
MS	1124	D. Berg, 6121	
MS	1124	S. Gershin, 6121	
MS	1124	R. Hill, 6121	
MS	1124	W. Johnson, 6121	
MS	1124	B. Karlson, 6121	
MS	1124	D. Laird, 6121	
MS	1124	B. McKenney, 6121	(15)
MS	1124	A. Ogilvie, 6121	(5)
MS	1124	J. Paquette, 6121	
MS	1124	M. Rumsey, 6121	
MS	1124	J. Zayas, 6121	
MS	1124	Wind Library, 6121	(5)
MS	1188	B. Thompson, 6133	
MS	1188	T. Brown, 6133	
MS	9406	D. Manley, 8114	
MS	9406	V. Peters, 8114	(5)
MS	0899	Technical Library, 9536	(Electronic)