BIG DATA AND PREDICTIVE ANALYTICS IN RENEWABLE ENERGY INVESTMENT

Executive summary

- Predictive Analytics and Big Data are currently two significant disruptive forces across the global economy, a trend which looks set to continue for the foreseeable future. Organisations that are sophisticated enough to leverage these tools are developing acute competitive advantages based on their nuanced insights and accurate forecasts.
- The implementation of a strategy incorporating Big Data or Predictive Analytics poses several technical challenges that need to be reviewed in detail to build out a suitable infrastructure to reliably grow and utilise the vast reserves of data required for actionable business insights.
- Within Renewable Energy Investment, the use of Predictive Analytics has not been as widespread as in other industries. One reason for this could be due to the lack of large datasets available to analyse. Despite this, there are currently applications of predictive maintenance, asset performance management and automated deal sourcing all leveraging different analytical tools and sources of data. Embracing these trends will allow investment funds to refine their processes and boost investor returns, providing a key differentiating factor between competitors.
- As data availability is expected to improve in the future it is advisable for organisations to develop their competencies in this area early so they can begin answering their most crucial business questions and locating their main performance drivers to maintain their competitive positions.

Author:

Saul Butt
Alternative Investment Associate
saul.butt@aquila-capital.com

Lars Meisinger
Head | Sales Management & Business Development
lars.meisinger@aquila-capital.com

The increasing democratisation of data and technology is making Predictive Analytics and Big Data pertinent themes for future business growth. They are becoming significant disruptive forces that are changing the fortunes and strategies of organisations throughout the economy. Fundamentally, they allow organisations to incorporate an adaptive quantitative basis to their decision-making, which is crucial in helping firms maintain strong competitive positions in an increasingly complex and dynamic world. Conversely, failing to develop competencies in this area could result in an organisation operating at an informational disadvantage to competitors. In order to appropriately leverage these tools, it is imperative that business leaders have a thorough understanding of how Big Data and Predictive Analytics will impact their respective industries and prepare accordingly.

The use of Big Data and Predictive Analytics is highly flexible, with applicability to almost any business context. However, the question of Big Data and Predictive Analytics' applicability in Renewable Energy Investment is a novel and challenging one. This brief paper will look to introduce the two concepts of Big Data and Predictive Analytics independently before attempting to describe their combined utility within the renewable energy space. The sections below will initially offer a conceptual overview of Big Data and Predictive Analytics interspersed with examples of business applications before progressing to more technical information concerning the challenges posed by their implementation.

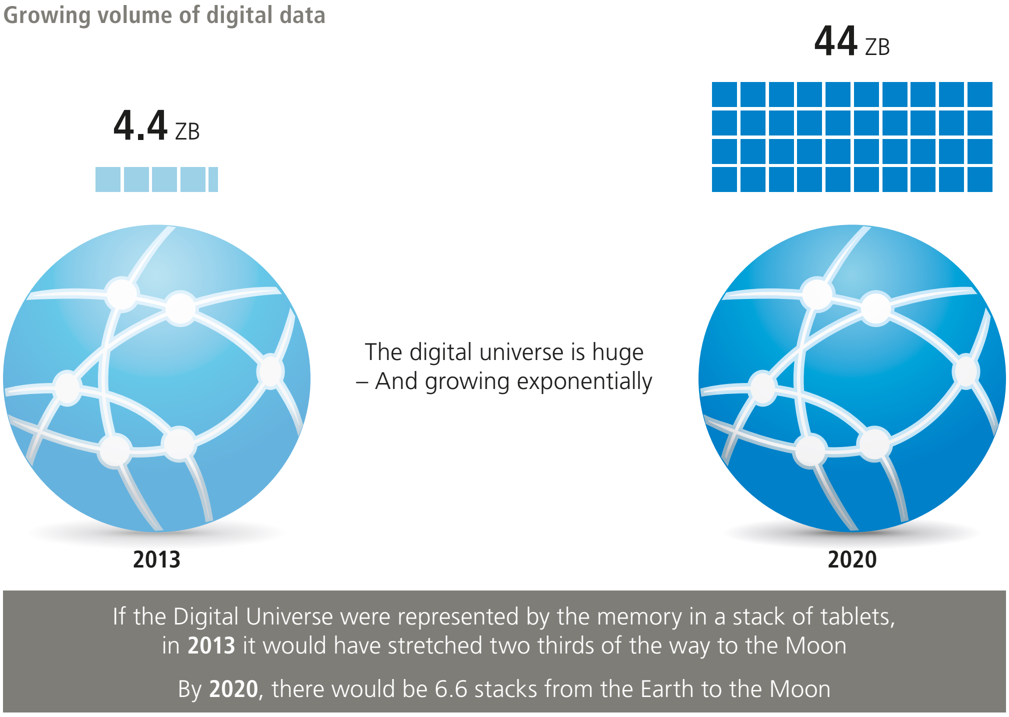
1. Big Data – Overview

Big Data is a term used to describe the vast scale of new datasets being assembled by organisations. Formally, there remains no consensus regarding what exactly is classified as big data, as the notion of scale in relation to data is continuously changing. However, a useful rule of thumb for Big Data would be data that is too large to store or practically manage within an Excel spreadsheet.

Data's exponential growth results from increasing levels of digitalisation and internet connectivity. IBM estimates that 90% of data in existence was created in the past two years[1]. This has coincided with ever more powerful computer hardware and distributed storage technology (i.e. The Cloud), which has followed Moore's Law of doubling in power every two years. This has provided organisations with the foundation to begin measuring and storing vast arrays of data points on a scale unseen in the past.

The intricacy and impact of insight derived from algorithmic, statistical and visualisation tools have grown significantly with the arrival of Big Data. Due to the value of these insights many organisations are attempting to adopt a more data-centric strategy. However, the process of developing a Big Data initiative poses several challenges, which require careful technical planning to overcome.

Growing volume of digital data



Source: https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

[1] *https://www.ibm.com/developerworks/community/files/form/anonymous/api/library/054c2ab9-ea33-4c70-b0c6-b5bb2482a098/document/7de665ff-2327-41a8-b7b0-5f0b-ba97356f/media/BIG%20DATA%20%2B%20MAINFRAME.pdf*

Some of the characteristics of Big Data are summarised in the four 'Vs':

Volume	The scale or quantity of data. Increasing the amount of data allows for more accurate modelling and insight into business questions. However, more data also raises significant obstacles to ordering, processing and storing information efficiently.
Velocity	The speed or frequency at which data is created. Greater internet speeds and digital connectivity have allowed for real-time data streaming.
Variety	The different types of data being reviewed. New techniques have allowed information to be obtained from text, speech and satellite images.
Veracity	The accuracy or quality of the data. The value of large datasets will be severely undermined if the quality of the data is poor.

Source: http://www.ibmbigdatahub.com/infographic/four-vs-big-data

A further extension to this framework similarly considers the value associated with data. This is a function of both its underlying quality and what insights the data can bring.

How has Big Data been used to create value?

Examples of Big Data acting as a source of value creation can be seen with online platforms such as Facebook, which have the ability to consolidate large datasets and develop acute insights into their user-base to leverage and support sales propositions. More generally, having large datasets has enabled analysts to better record the past and delve into complex interactions and systems, such as those between companies, humans and markets.

Case study – Applications of Big Data

Investment Management: It was recently reported that Quantitative Hedge Funds will surpass USD 1tn AUM[2]. Automated data-driven solutions are more scalable and accurate than traditional discretionary fund strategies. The attainment of these achievements is based on the ability to operate flexibly within a complex environment and transform vast amounts of data into accurate investment signals.

Credit Scoring: A range of data has allowed for increasingly reliable credit scoring. As opposed to assembling a credit score based solely on individuals' past financial behaviour, many organisations are tapping into alternative Big Data sources such as social media accounts. The facets of one's social media account are often indicative of personal behaviour traits, which can be highly informative of future actions and credit-worthiness. This, among other applications, has made the provision of social media data a highly profitable business.

Healthcare: The advances in healthcare analytics over the coming decades will be phenomenal. The quantity of data being created through healthcare apps and by consolidating healthcare records are providing unparalleled insight into individuals' lifestyles and wellbeing. This is being investigated in regards to individuals' health and has created the opportunity to forecast the probability of contracting specific diseases and to personalise preventative healthcare treatments to best mitigate them.

[2] *https://www.ft.com/content/ff7528bc-ec16-11e7-8713-513b1d7ca85a*

2. Big Data – Technical considerations

What are the challenges with managing Big Data?
Managing large databases efficiently is a complex task and often requires specialist knowledge. Some of the common issues associated with managing database structures are:

Data Formatting	The correct formatting of data is required to avoid data duplication, which can cause information to grow unnecessarily large. Further, the speed at which data can be accessed and processed will be depends on its format.
Storage space	Organisations need to decide how to store data, whether through acquiring servers in-house or leasing server space from external organisations. Both of these options have different associated cost structures and benefits.
Processing power	To access or process Big Data, a substantial amount of computational power is required. This can either come from more powerful computers, which will suffice up to a point. However, distributed systems are far more scalable and increasingly affordable.
Data quality	The volume of Big Data is presently a significant point of discussion in strategy development. However, if the quality of the assembled data is poor, then its value will be substantially diminished.
Data security	There is a trade-off in providing employees with significant accessibility of corporate data in a centralised format and the potential severity of a data breach. It is imperative that data is stored using secure servers, strong passwords and advanced level encryption to avoid the loss of customer and business data.

Source: http://pwc.blogs.com/analytics_means_business/2017/09/big-data-big-deal.html

'Digital exhaust' is a term often used to describe data created by firms as a by-product of their operations. The challenge of data storage is amplified for organisations that produce significant amounts of digital exhaust and wish to store it for future analysis. In this case, it is required to have a detailed and forward-looking data architecture such that new data sets can be seamlessly integrated into an existing system. Without an appropriate architecture it is likely that the ability to efficiently store and access data will be constrained in the future.

What are the common approaches to data storage?
In most circumstances, relational databases are generally suitable for data storage requirements. This structure allows information to be maintained in an efficient and easy-to-recover format. The foundation of this format is based on data inputs, tables and relationships between the tables. The core technology behind this is typically a free and well-established SQL structure that can be easily replicated and stored in the cloud to improve accessibility. In addition to this there are corporate services that provide customer support and guarantee efficient design[3,4,5].

In the case of Big Data, more novel storage methods are often based on NoSQL, such as MongoDB. This does not rely on the traditional relational structures between data tables and allows for a far quicker and more dynamic way to store and retrieve data. This approach is generally built on flexible JSON inputs; however, this data solution is likely to be excessively complex for most organisations without vast data reserves.

Early stage data strategies
When developing a strategy transitioning towards more empirical decision-making, organisations often already possess large quantities of highly valuable information internally. Usually, this information has not been appropriately stored or formatted in a useable form. In order to make this information accessible for analysis, collating the data within a centralised data warehouse is a key step in beginning an initiative of data-driven decisions and automation.

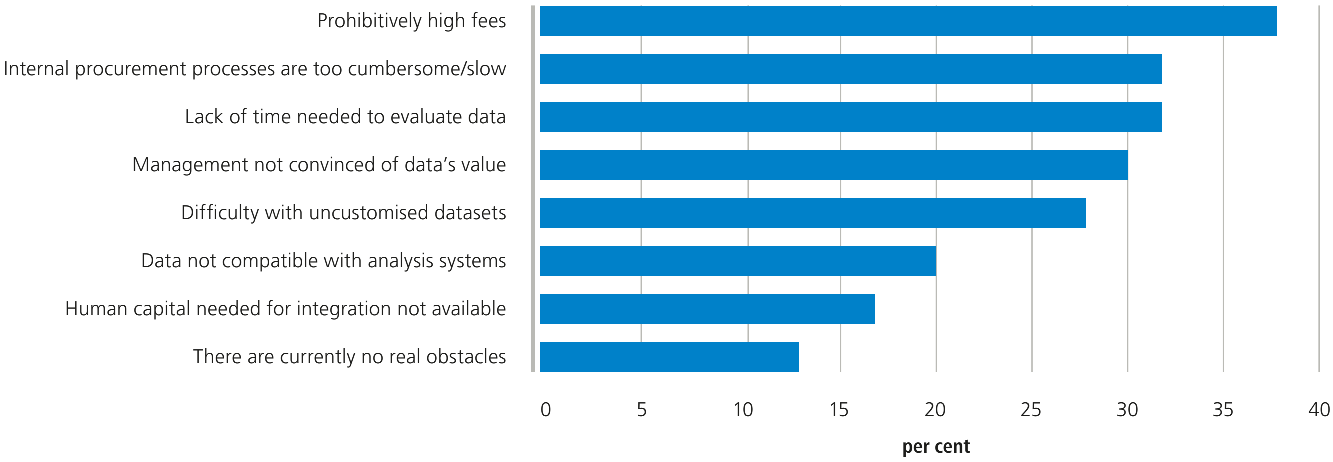
The standardisation of information is fundamental. It is estimated that data analysts spend around 80% of their time cleaning, structuring and reviewing data, with the time demands of applying value-adding analytics being relatively trivial[6]. Having data stored in a standardised format would greatly improve the ability to provide-broad-based and detailed analytics. Furthermore, with data easy to access, the automation of tasks such as reporting and performance indexing becomes a far more feasible proposition.

In addition to this, information concerning an organisation's past performance will often prove to be highly valuable in developing an understanding of its core performance drivers. Therefore, prior to delving into acquiring exotic data sets, it is cost-effective to collect the low-hanging fruit provided by collating internally-stored data.

Advanced stage data strategies
Many organisations do not possess 'Big Data', which can put them at a competitive disadvantage. If an organisation already owns a developed data architecture, one option to overcome this would be to acquire or lease external datasets to derive further insights. A complication with this approach is that the valuation of datasets is a novel and developing field[7,8]. It is often difficult to estimate where and how much value will be created by a specific dataset prior to its acquisition, which makes the market pricing of data highly subjective.[9]

[3] http://www.oracle.com/technetwork/topics/entarch/oracle-wp-big-data-refarch-2019930.pdf
[4] https://tech.winton.com/2017/03/data-technologies-at-winton/
[5] https://tech.winton.com/2017/09/creating-a-scalable-data-ingestion-process/
[6] www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says
[7] https://hbr.org/2016/09/do-you-know-what-your-companys-data-is-worth
[8] https://svds.com/valuing-data-is-hard/
[9] https://sloanreview.mit.edu/article/whats-your-data-worth/

Obstacles to use of alternative data by asset managers and hedge funds



Source: https://www.ft.com/content/d86ad460-8802-11e7-bf50-e1c239b45787

Opportunities to develop datasets

It is possible to grow an organisation's data resources without acquiring datasets from third-party providers by scraping the internet for information. This approach allows firms to grow their datasets by focusing on the key data they require for their decision-making. Some tools to simplify the web-scraping process are import.io and mozenda.com.

Combining datasets and looking externally

Novel insights have often been found by combining multiple datasets and uncovering new correlations. Considering Renewable Energy investments, there are several useful external datasets that provide market data that could be utilised. Some examples include:

Renewables. ninja (weather data)	Provides hourly forecasts of weather conditions around the globe.
OPEN.ei (Energy Production)	A range of free datasets concerning energy and renewables. The information is peer-verified to establish the quality.
Energy Demo (Solar Energy Production)	Open source data providing the location of numerous solar producing assets.
Quandl (Alternative Data)	A combination of free and paid datasets, this is perhaps the most established and comprehensive source of alternative investment information.

Take away

Big Data is changing organisations and is a key source of competitive advantage. It is providing the foundations for more empirically-based decision-making.

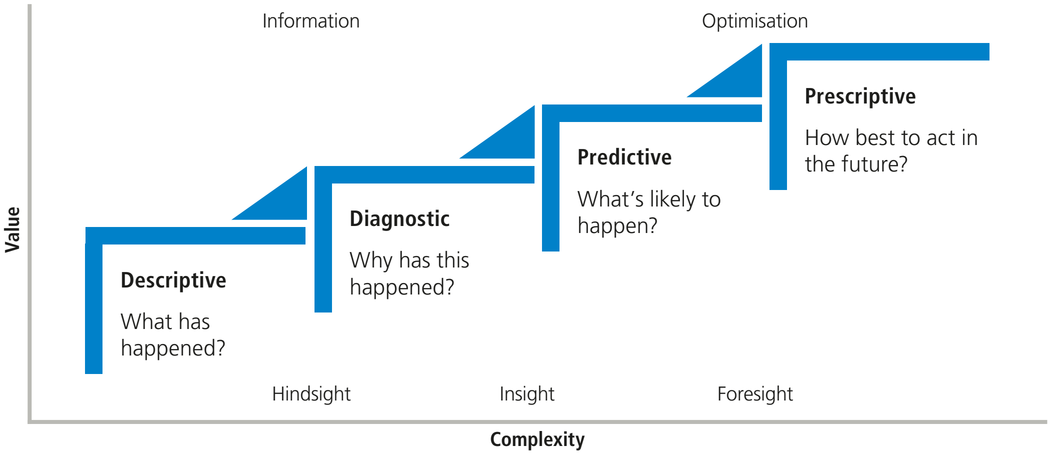
The storage and management of growing amounts of data will often require significant planning and investment to develop and maintain an effective data infrastructure.

For organisations beginning to transition towards empirically-assisted decision-making, a robust centralised database and processing historic operational data will often provide significant value.

3. Predictive Analytics – Overview

Analytics is broadly defined as the systematic processing of data to generate insight. The forms of analytics can be categorised into the following branches:

Stages of analytical analysis



Forms of analytics

Descriptive analytics: the most commonly applied form of analytics. It is used to provide insight into what has happened in the past.

Diagnostic analytics: seeks to understand why particular events occurred. This is often based on testing correlations between various relevant variables.

Predictive analytics: attempts to develop forecasts. Probability and statistics are employed to understand the likelihood of an outcome occurring based on past information. This sees the complexity of predictive analytics rising significantly relative to earlier stages of analytics.

Prescriptive analytics: attempts to estimate the best responses to forecasted events. When operating within a stochastic context this requires complex computational capabilities and will be based on expected value calculations. The benefit of accurately modelling best response decisions can be significant. This form of analytics is primarily dominated by Machine Learning and Artificial Intelligence.

Predictive Analytics are changing the way that organisations approach decision-making. They allow firms to not just look at the past and evaluate their performances but to forecast what will likely happen. The tools available to apply predictive analytics are extremely broad and flexible, which makes defining a clear business problem crucial to identifying the appropriate solution. When defining a business problem, it is often useful to go through each stage in the above graph. Additionally, incorporating visualisation tools into the analytics process can help to quickly communicate quickly the narrative behind a problem and develop actionable insights.

Common techniques applied in Predictive Analytics

Regressions	Regressions provide insights into how highly correlated variables are to a particular outcome or event. Once established, the model inputs can be adjusted to understand the expected outcome in different scenarios.
Time Series Models	This is an extension to a simple regression model, with the variable of time considered in order to understand how trends will develop and progress. This is key to creating insightful forecasts.
Geospatial Models	This is an extension to a simple regression model, with the variable of space considered to show how trends develop in areas. This is a particularly powerful tool when the element of time is additionally considered.
Machine Learning	Utilises statistical methods along with computer science principles to allow programs to learn independently. The algorithms developed for machine learning are useful in uncovering patterns within data. In the machine learning space there is a plethora of tools available of differing complexity. The most basic models tend to be linear, while more advanced methods can model non-linear relationships.
Artificial Intelligence	These are programs that are often more advanced than machine learning and look to learn independently and respond to environmental stimuli. These programs can be extraordinarily useful for prescriptive analytics when run in simulations. Common tools are reinforcement learning and artificial neural networks.

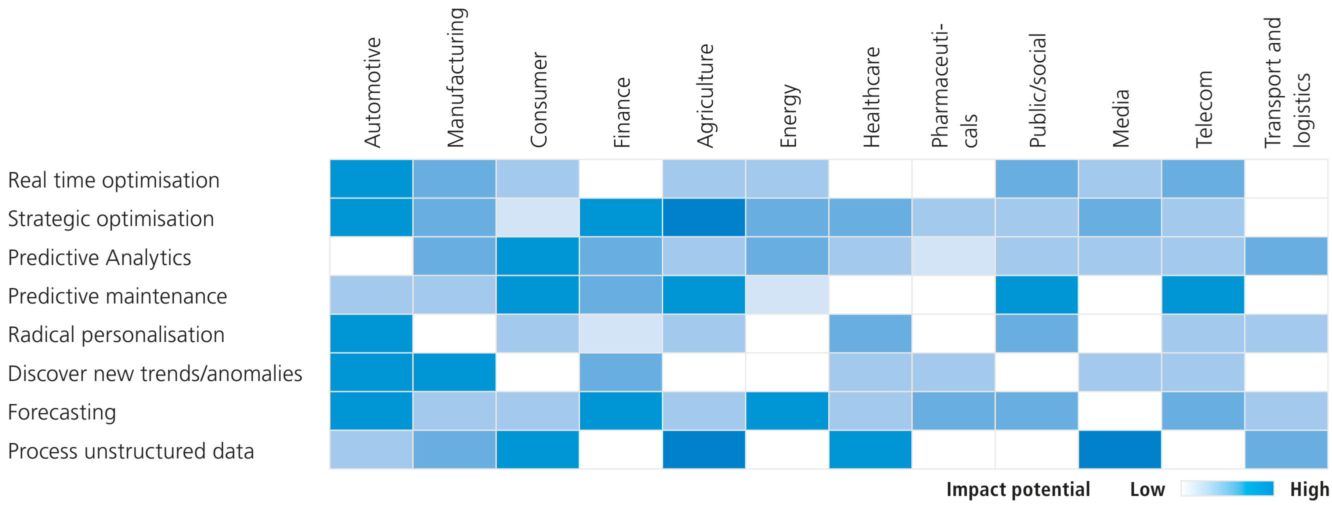
What research questions can Predictive Analytics be applied to?
1) Categorisation – Programs look to decipher whether an observation belongs in one category or another. An example of this would be whether to categorise a stock as a buy or a sell, based on its fundamentals.

2) Numerical estimation – This type of estimation is generally slightly more complex, where a program will be used to quantify an outcome. An example of this would be: based on a stock's changing value, what is the probability that it will continue to rise or fall?

3) Correlations – Models looking to uncover correlations are often very useful in understanding the interrelatedness of variables within a complex environment.

What is the interaction between Big Data and Predictive Analytics?
Big Data has been integral in developing the accuracy and efficacy of predictive analytics. This has been achieved through the breadth and depth of data now available. The increasing digitalisation of society has led to a plethora of new data points by which to refine the modelling of complex relationships. This much broader array of information that can be incorporated into models has shed new light on desired research topics. Further, it is clear that having more complete population data will allow a model to more accurately replicate reality. Therefore, as the depth and coverage of Big Data increases, the reliability of models will likely improve.

Machine learning has the potential to be applied across many industries



Source: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/The%20age%20of%20analytics%20Competing%20in%20a%20data%20driven%20world/MGI-The-Age-of-Analytics-Full-report.ashx

Case Study – Innovative Uses of Predictive Analytics

Asset maintenance: Rolls Royce operates within the aviation maintenance business, through leasing jet engines. As opposed to simply leasing the assets, they have augmented their service by providing advanced predictive analytics based on data collected across all their jet engines. This allows Rolls Royce to forecast when their customers' engines require maintenance, which lowers repair costs and downtime. This new business model structure would not be possible without the developments in storing and forecasting using large datasets.

Pricing Forecasts: Kayak, an online travel agent, provides forecasts for customers to identify the ideal time to purchase flight tickets at the lowest prices. This has been achieved through collating a vast database from airlines' daily ticket pricing and uncovering pricing patterns.

Risks associated with Predictive Analytics

There are a number of risks associated with implementing predictive models, with a lack of training data and transparency in black box models being two key issues. Poor or inadequate data being used to fit a model will often result in forecasts that misrepresent reality. A means to overcome this is to include more predictive variables. However, this can lead to overfitting and increasing model complexity, making it extremely difficult to interpret the fundamental reasoning behind a model's output.

Additionally, analytical models are particularly accurate in stable environments. However, when there are fundamental changes in what a model is measuring, there is no way to accurately account for this. This makes it important to maintain models with current data and understand when to disregard their output to avoid fallacious assumptions.

4. Predictive Analytics – Technical considerations

Challenges of establishing Predictive Analytics
For analytics to be valuable there must be a suitable quantity and variety of data to process. Without such data it is impossible to generate statistical analyses that have the required accuracy or reliability upon which to base business decisions.

It is also necessary to invest in hiring appropriate talent: an analyst must have the right technical capabilities, along with a strong mathematical and computer science background. Sector knowledge is crucial to identify the key business issues and to structure research accordingly.

What are the Common Tools Utilised in Predictive Analytics?

Python	Python is perhaps the programming language that is most synonymous with Data Science. It includes the most advanced packages available for Machine Learning and has strong data management tools. The flexibility of the language allows for efficient algorithms to be constructed.
R	Similarly to Python, R is a high-level programming language. It has extremely advanced statistics and visualisation libraries.
Tableau	Allows users to quickly create elegant visualisations and conduct data analysis to understand the stories within large datasets.
SQL	As previously discussed, SQL is a database software. The language is declarative, meaning that it will select the optimal way to execute a task, making it an efficient data-querying language.
Jupyter Notebooks	A reporting technology that allows analysts to incorporate code, text and graphs together. This allows for transparency in methodology and encourages collaboration.
GitHub	An online code repository. This allows individuals to share code and approaches to solving problems. This is the core platform to collaborate on software development and can host a great deal of open-source software.
Azure	The distributed computational platform from Microsoft. It is a user-friendly means by which to parallelise big computational tasks over large datasets.
LaTeX	A report processing language that allows the user to design reports with significant additional flexibility in relation to traditional text applications such as Word.

Large technology companies such as Google and Facebook are continuously publishing open-source software libraries that enable analysts to conduct ever more advanced computations. Additionally, the existing software is being constantly developed to make them more user-friendly. These two trends mean it is easier for organisations to undertake valuable analytics.

Take away
Predictive Analytics can provide potential solutions to a range of business problems.

Simple diagnostic analytics using visualisation tools can Often create substantial value in understanding business problems.

The accuracy of a model's output is highly correlated to the quantity of data available to train it.

5. Big Data and Predictive Analytics – current applicability within renewable energy

The use of Big Data in the management of financial asset classes has produced excellent results for some organisations. These successes has been based on leveraging the three trends of: increasing computer power, data proliferation and improving analytical methodologies[10]. Within Renewable Energy investment, the use of analytics has been slightly more subdued due to the low availability of data. This is largely due to the nature of the asset class being privately held with a highly-fragmented ownership structure, meaning the opportunities to assemble consolidated market data is extremely limited. Despite the sparse data environment within the Renewable Energy investment industry, there exists opportunities to undertake valuable analytics.

Asset Performance Management - APM entails collating past performance information from a portfolio of assets and uncovering correlations between certain variables and a target metric, such as maintenance spending relative to profit. Based on this, a forecast will be made to calculate how performance can be adjusted to optimise on the target metric. This is often a balancing act in looking to manage the asset conservatively, while also attempting to maximise energy output.

This is often an ideal research project for organisations with a large backlog of historic data. After establishing past relationships, predictive analytics methods such as regression analysis can be applied to find the optimal conditions to run a project.

Data related to Asset Performance is growing significantly. One of the main drivers of this has been the incorporation of sensors used to measure the operating efficiency of an asset. Once data has been recorded for a great number of assets it can be generalised into a model and used for forecasts.

Case Study – General Electric

GE has become a market leader in asset performance management for renewable energy production. Through using the principles of the Internet of Things and installing sensors with internet connection to small SQL databases, GE can assist in assembling personalised information for energy producers.

On top of GE's ability to assist asset owners in collecting performance data, they have similarly used their scale of insight in renewable assets to provide performance and optimisation recommendations to increase profitability and reduce downtime.

APM becomes a more compelling proposition when combining past operational data with alternative datasets to hold external factors constant and understand how much of a project's success is dependent on the asset management's decision-making. One growing form of large datasets, which could be significant in this case, is weather data. This is being assembled using asset sensors, satellite imaging and independent weather stations. It provides time series data on progressing weather patterns, which can then be used to measure the impact of weather on asset performance.

Case study – IBM

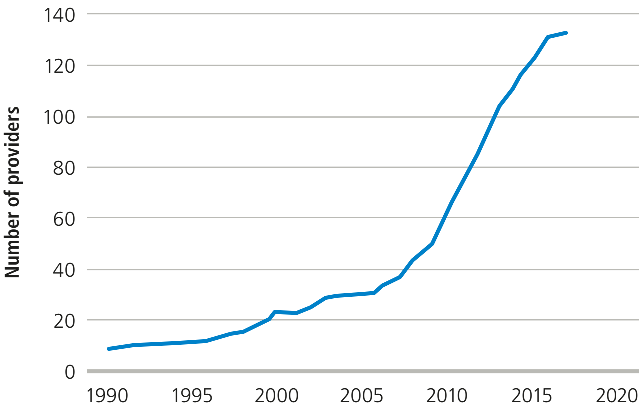
Weather data and modelling has advanced significantly since IBM's recent acquisition of The Weather Company, which provides information to Watsons Artificial Intelligence platform. The system is currently collecting "four gigabytes of data per second from 40 million mobile phones, 147,000 sensors in weather observing stations, as well as 50,000-plus daily flights that collect data on temperature, turbulence and barometric pressure." The insights this will generate in collaboration with IBM's Artificial Intelligence capabilities will allow renewable energy producers to better forecast energy production and plan maintenance works[11].

Case Study – Energy storage and market pricing

Using predictive analytics to model the conditions for an asset's optimal operating efficiency and energy output is a valuable tool in improving investor returns. However, the requirement of this service will become far greater with the implementation of smart grids and storage facilities. In order to avoid oversupply to the grid because renewable energy sources produce simultaneously[12], it is likely that batteries will be used increasingly to regulate the supply of electricity. In order to achieve this, optimisation models forecasting the appropriate times to release energy into the grid will become a significant differentiator in the profitability of renewable energy assets.

[10] *https://www.man.com/the-rise-of-machine-learning*
[11] *https://www.cio.com/article/3006300/big-data/weather-company-forecasts-more-big-data-for-ibm-watson-analytics.html*
[12] *https://www.renewableenergyworld.com/articles/2014/02/the-interconnection-nightmare-in-hawaii-and-why-it-matters-to-the-u-s-residential-pv-industry.html*

Alternative data provision is a fast-growing niche



Source: https://www.bcg.com/publications/2017/principal-investors-private-equity-fund-strategy-operations-digital-deal-sourcing-private-equity.aspx

Reliability-centred management & predictive maintenance – The aviation industry initiated the preliminary study of Reliability Centred Management (RCM) in the 1960s to combat their assets' escalating O&M costs. This resulted in a broad and collaborative research initiative concerning the maintenance and durability of mechanical components to reduce costs associated with component defects, replacement and asset downtime. The resultant adaptation to maintenance spending was to allocate expenditure based on a component's relative importance to the asset and replacement cost, which was developed into a more effective weighted maintenance spending model.

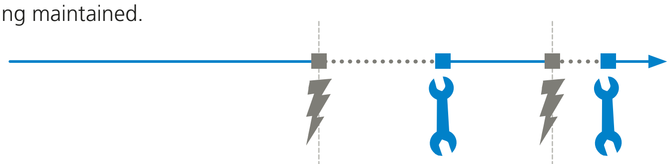
Traditionally, manufacturers would assist asset operator RCM by providing a recommended preventative maintenance program for its components. Specifically, they would provide suggestions for when to undertake maintenance that would typically be based on time, events or meter readings. An example of this would be having a wind turbine gearbox serviced every 18 months. These measures have proven highly effective in reducing maintenance costs. However, these preventative measures are generally not specific to the context in which an asset is used. For example, a wind farm in Portugal exposed to extremely hot and dry weather will likely face different maintenance issues from an identical asset based in Sweden with freezing temperatures and high precipitation.

A recent development in RCM that can account for the variance in assets' maintenance requirements is predictive maintenance. This approach utilises various data sources to predict when maintenance expenditure is required, indicate what actions to undertake and forecast the benefit of the expenditure. When completing predictive maintenance, traditional factors such as the age of components can be used. However, novel factors such as thermal imaging and vibration monitors along with external factors such as weather are becoming increasingly popular to develop a clear picture of what issues an asset is likely to incur. This approach, utilising regression analysis and optimisation models, allows for specific maintenance programs to be created for each component and will flag issues prior to a major defect occurring, resulting in better allocation of O&M spending. This novel method can combine both predictive and prescriptive analytics with the benefits of this approach estimated to be over 30% cheaper than traditional preventative measures[13].

[13] *https://www.coresystems.net/blog/the-difference-between-predictive-maintenance-and-preventive-maintenance*

Maintenance strategies in an industrial context

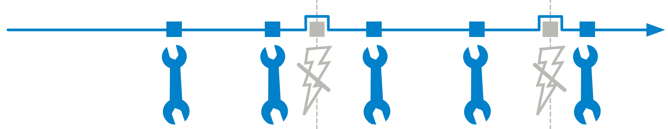
Reactive maintenance
Assets fail before being maintained.



+
- no analytics or sensors needed
- spare parts are fully used up

-
- no analytics or sensors needed
- spare parts are fully used up

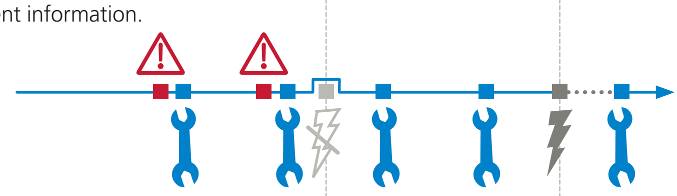
Preventive maintenance
Systems are maintained at fixed intervals to ensure continuous availability.



+
- no in-process sensors needed
- reduces unplanned downtimes

-
- waste in life cycles of spare parts
- excessive maintenance and downtime

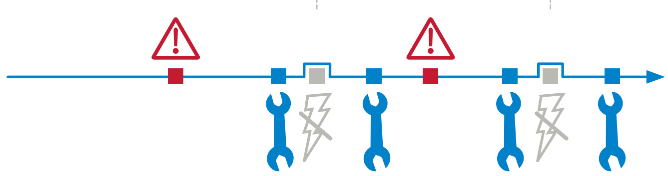
Condition-based maintenance
Systems are maintained based on simple rules using equipment information.



+
- good usage of equipment life cycle
- limits unplanned downtimes

-
- limited failure forecasting accuracy
- medium risk of failures remains

Predictive maintenance
Systems are maintained before failure, but run as long as possible without interruption.



+
- optimal usage of equipment lifetime
- minimises unplanned downtimes

-
- upfront investment necessary
- expert knowledge required

Source: https://www2.deloitte.com/content/dam/Deloitte/de/Documents/deloitte-analytics/Deloitte_Predictive-Maintenance_PositionPaper.pdf

In order to improve the validity of predictive maintenance, a range of historic data must be available to refine forecasts of component failure. One means to achieve this has been to encourage industries to centralise their data such that each market participant will have a wider and potentially longer backlog of historic data. T-Systems recently indicated that even small degrees of corporate collaboration through sharing data on a single component can yield significant results[14]. A good example of data sharing can be found in Norway, whereby due to the extensive cross-ownership of hydroelectric producing assets, Energy Norway has been able to centralise a database of between 90 – 140 TWh of Hydropower production data a year. With this vast reserve of data, a research project called *MonitorX* has begun, which looks to use machine learning and artificial intelligence in collaboration with Big Data to model component lifetime optimisation based on a variety of surrounding conditions and uncertainties[15]. This collaboration includes 11 Norwegian and Swedish power producers and two suppliers that are working together in testing new analytics methodologies to better evaluate their maintenance expenditure and processes[16].

Deal sourcing & due diligence – Combining web-scraping with quantitative analysis methods such as natural language processing has made it possible for managers to translate substantial reserves of information online into insights on potential acquisition targets. With web-scraping programs it is possible to trawl the internet to find specific assets that have been mentioned in the news or on social media. Once an asset has been located, relevant information can be found and compared against a set of criteria specified by the investment manager. If the asset meets a specific set of requirements it can then be flagged as worthy of greater research. In a recent BCG report, the importance of leveraging proprietary information from web-scraping and data analytics to guide both deal sourcing and due diligence were highlighted as being vital to avoid being at a permanent information disadvantage to competing funds[17].

Case study – Private equity due diligence developments
The application of automated due diligence through assembling data with web-scraping is providing a significant advantage to tech-savvy private equity funds. According to a recent article[18], the information found through trawling the internet is providing new insights that support private equity funds in their increasingly competitive bidding processes. Additionally, the automation of the research is reducing the time and costs associated with typical due diligence processes, creating a leaner and more detailed analysis.

Further applicable datasets
Macroeconomic data - Gaining insight into detailed macroeconomic data can be a key driver in strategic and corporate level decision-making. Understanding the exact point at which one stands in the business-cycle can be highly informative as to the impending economic forces to which a business and its clients will be exposed. In relation to this, the investment manager can begin preparing the appropriate investment products that clients will require over subsequent months and years. This allows the fund managers to make forward-looking decisions as opposed to reacting to current market trends. Recently there have been a number of innovations in macro-economic forecasting with inflation forecasts being run in real time using online data such as the Billion Price Project or satellite imaging to infer economic activity and GDP growth. Leveraging these tools could allow managers to develop more nuanced and timely economic indicators to support their decision-making.

[14] *https://www.t-systems.com/de/en/industries/automotive/sales-aftersales-solutions/remote-support/predictive-maintenance-379704*
[15] *https://www.energinorge.no/energiforskning/fornybar-energiproduksjon/vannkraft/maskin-elektro/pagaende-maskin-elektro/monitorx/*
[16] *https://www.energinorge.no/energiforskning/nyheter/2016/verktoy-for-a-unnga-havari/*
[17] *https://www.bcg.com/publications/2017/principal-investors-private-equity-fund-strategy-operations-digital-deal-sourcing-private-equity.aspx*
[18] *https://www.forbes.com/sites/baininsights/2017/04/07/data-mining-your-way-to-better-due-diligence-in-private-equity/#6cda573c243e*

6. Summary

Big Data and Predictive Analytics in combination are providing a competitive edge in many industries. The cost of accessing data is falling, technology is ever more powerful and analytical methods are increasingly sophisticated. However, the scarcity of relevant data in the renewable energy investment industry remains the limiting factor in the application of further predictive analytics.

At present, the renewable energy industry can develop detailed descriptive analytics of its past performance by combining assets' operating information with external datasets such as weather conditions. However, with the persistent proliferation of available information through data collection technologies and large corporate data providers it is expected that data will become sufficiently abundant in the near future to support more elaborate predictive analytics methods. This seemingly secular trend in data growth will support the development of intriguing new insights in Renewable Energy Investments along with accurate forecasts of asset returns and performance. It is therefore crucial to begin developing the requisite capabilities to manage and process data, such that Renewable Energy Investment funds can gain early insight into their performance drivers and business problems. Once these key variables and issues have been isolated, organisations can undertake deep analyses into these topics when suitable datasets become available. This will provide a platform to uncover best practices and maintain superior decision-making ahead of competitors.

For more information, please contact:

Aquila Group
Valentinskamp 70
20355 Hamburg
Germany
Tel.: +49 (0)40 87 50 50-199
E-mail: info@aquila-capital.com
Web: www.aquila-capital.com

Follow us on

Hamburg · Frankfurt · London · Luxembourg · Madrid · Oslo · Zurich · Invercargill · Singapore · Tokyo

A publication of Aquila Capital Investmentgesellschaft mbH. As of July 2018. Author: Saul Butt, Lars Meisinger

