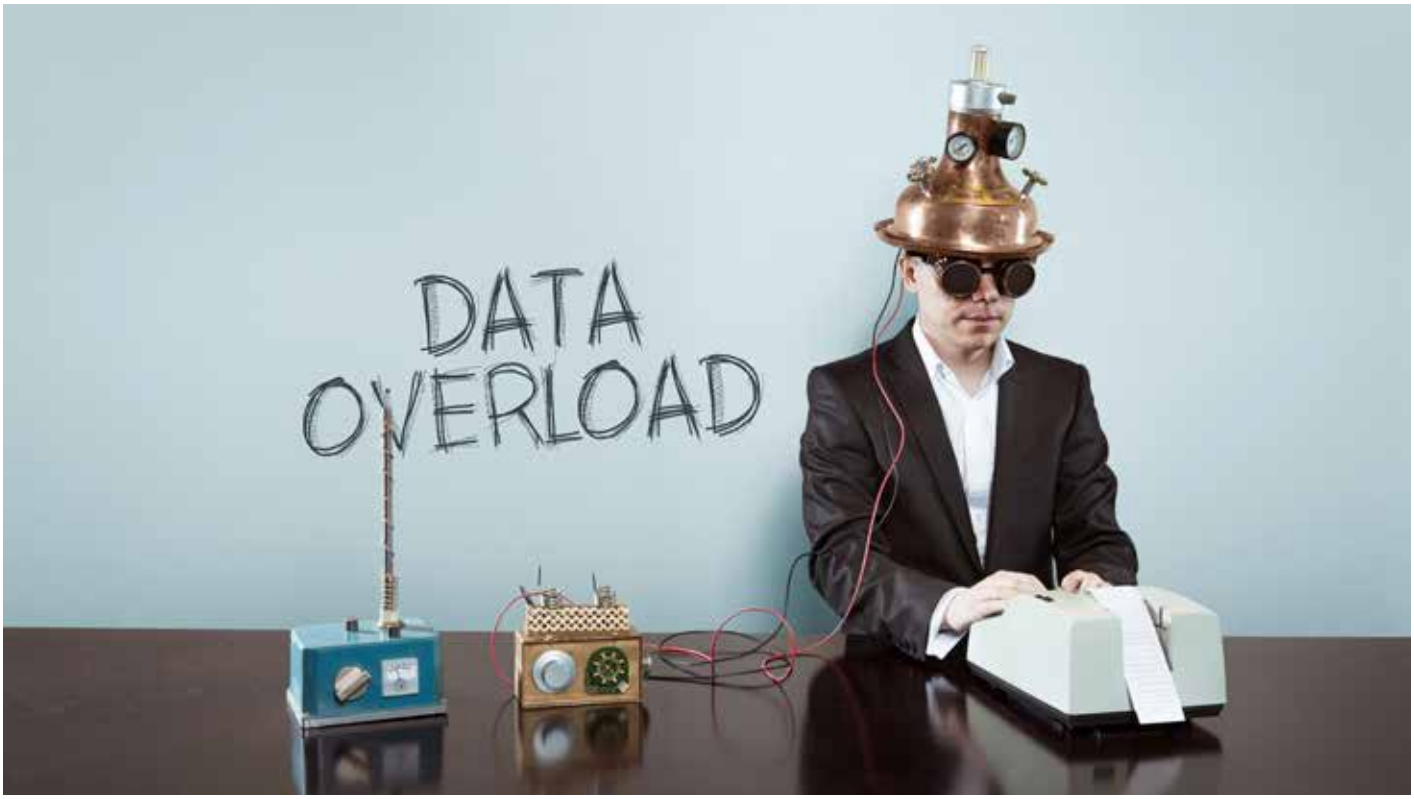
Big data under the microscope

Monitoring | The term "big data" has recently become something of a buzzword in PV circles. But how are companies using this data and is it generating useful intelligence or merely information overload? Catherine Early reports



In a world where government financial support for PV technology is becoming increasingly rare, the need to optimise plants and make them more efficient has become critical. There is no longer any margin for underperformance - losing even just a small amount of energy production due to faulty components could destroy a plant's business case.

One solution being actively pursued by many in the sector is the collection and analysis of "big data" – large and diverse data sets that, when combined, could hold the key to innovation that could reduce costs, such as predicting when equipment needs to be maintained rather than reacting after a fault has become apparent.

One company that is already using big data both internally and with clients is Greek solar asset management and O&M firm Alectris. Its Enterprise Resource Planning (ERP) software, known as Actis, incorporates all elements involved in operating assets, such as maintenance services, power purchase agreement billing, project management and financial and technical reporting. Developed as an internal tool seven years ago, the firm uses the software to help its decision making, and manage its costs.

The tool is now gaining traction externally with clients, especially in the US, with some in Europe and Asia also showing interest in the past few months, according to Vassilis Papaeconomou, the firm's managing director. It means that the firm can share all the information it has about an asset directly with its owners, enabling far more detailed reporting.

"If we have to charge our customer for some maintenance, they can see exactly what activity has taken place to justify that – which component, what type of equipment failure and how this has impacted the overall performance of a plant," Papaeconomou explains. "It gives the ability to look holistically at the data and not see a static report, which would just say 'US$800 – fixing an inverter'."

Big data is becoming increasingly prevalent in PV but question marks remain over how best to use it

The firm can also use this data to identify the most common failure relating to the impact of the plant, and from that, to make decisions about whether certain elements need replacing. "You may have 100 failures of a certain type, but they have no impact on the plant. Or you may have another 10 events, but the impact could be substantial – the software enables you to focus on the right elements, and make informed decisions," he says.

What is "big data"?

A solar PV plant is rich in data, from the components themselves, to financial information and weather statistics. But big data analytics is relatively new in the PV sector, and definitions are somewhat fluid. Vassilis Papaeconomou, managing director at O&M firm Alectris, says that care is needed in using the term "big data" in reference to the PV industry. "In solar, we're not talking so much about the volume of data, that's nothing compared with Google and other data companies," he says.

Lars Landberg, director of renewables, strategic research and innovation at certification and consulting firm DNV GL, says that big data is about far more than just the volume of information. There are three other criteria, which he explains in terms of the PV sector.

These are velocity, where a lot of data is coming at you all the time such as high resolution data from SCADA monitoring systems; variability, which refers to different sources of data, such as SCADA software, the stock exchange, and financial transactions; and veracity, which refers to how data from some sources, such as a tweet from US president Donald Trump, could be misleading.

Clients can look at their assets and identify the top five causes for loss of output, and how much more energy could be produced if these problems were fixed, and therefore calculate the business case for doing so, he adds.

Project advisory consultancy Greensolver meanwhile, is using big data in its Greensolver Index tool. The software has been available for the wind industry for a few years, but has recently been launched for the PV sector. The tool can measure performance, irradiation, capacity factor, plant availability and energy generation.

Alexander Harssema, partner at the French firm, explains that the software enables analysis of the average time between failure of an asset, the performance of a contractor in terms of responding to faults, the failure rate of a particular brand of technology.

"We can help clients negotiate better maintenance contracts; for example, we could say that a particular inverter has a failure rate that could be improved if the maintenance provider had a local service station, or more staff on site. That could be the commercial result of this data," he says.

Many firms are carrying out research and development on other applications of big data. One major area under investigation is predictive maintenance, where data could be analysed to pick up unusual activity that could indicate an underperforming asset. Time and money could then be saved by sending out maintenance teams to fix a fault ahead of it affecting productivity of a plant.

Machine learning, where a device would spot patterns in behaviour, could be applied to this use of big data, says Lars Landberg, director of renewables, strategic research and innovation at certification and consulting firm DNV GL. The case for such use of data is rising along with the installed capacity of solar technology. "If you have one power plant, it's easy to detect when something is wrong. But with thousands of PV panels, it would be a huge task," he says.

Dr Günter Maier, managing director and chief operating officer of Alternative Energy Solutions, which specialises in PV monitoring data analysis, says that this could be taken a step further, with asset owners using predictive maintenance to improve the efficiency of spare part management, and hold parts in stock centrally rather than having masses of parts at every plant. "In other industries this is common sense, but the PV sector still has some way to go before everything is optimised," he says.

Optimising use

But Edmee Kelsey, chief executive of asset management company 3megawatt, is

How will big data, machine learning and AI impact jobs in the PV sector?

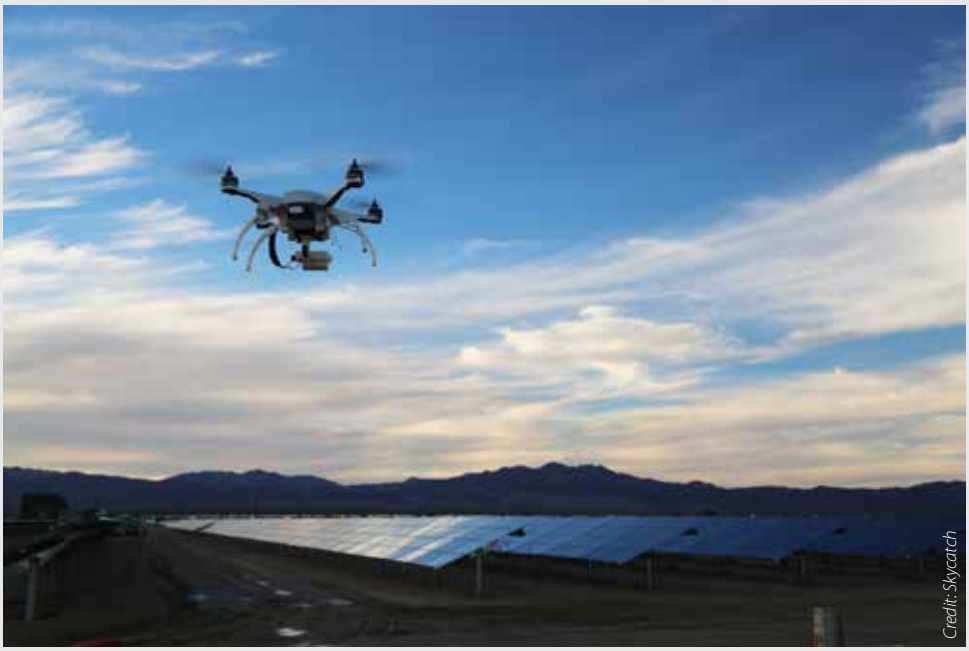
A common-held fear is that as more tasks are automated or carried out by drones or robots, the need for human workers will fall. But commentators believe that in the PV sector, the impact is more likely to be that jobs become more productive and rewarding, as workers' time is freed up from laborious tasks.

"There are so many things that can be fixed, professionalised and industrialised using AI. It will help us with tasks that are labour intensive, such as using drones to check solar panels. Think of a worker checking each panel on a large plant, it would take days. If drones and infrared cameras are used instead, you still need someone to pilot the drones," says Cazacu.

Kelsey notes that many jobs in the PV sector are very highly skilled, such as electrical engineers and lawyers. Rather than replacing these types of jobs, digitalisation will free employees up to take a step back and use their experience to analyse problems identified by machines. "In the beginning, systems will still be learning, so you have to have someone who looks at the results and asks if they make sense. Jobs will change and will be more fun and won't have to do repetitive tasks," she says.

Papaeconomou agrees that jobs will become more fulfilling rather than be replaced. He adds: "Workers will become more efficient and can do more complex tasks. If you have a team of five, the number of things they can accomplish a day increases from five to 15."

Khorana notes that there has typically been a shortage of people who want to maintain PV plants in remote areas, despite wages for these roles increasing. "People would rather sit in an office and optimise plants remotely than be away from home for days at a time in the field," he says.



Drones and the data they collect represent a way of easing labour-intensive tasks

sceptical about the practicalities of predictive maintenance. "I can see the value that could come with artificial intelligence (AI) to make sense of patterns of behaviour. But it has to be actionable to make it work. For example, if you have a big power plant and something goes wrong, you still have to send someone out there, so it has to be a big enough problem to justify this."

Calculating the financial case for sending a technician out is increasingly complicated in the post-subsidy era, where power plants are selling energy on the open market, and might also include other elements such as storage, she says.

"You have to calculate the full picture in order to get permission to send someone out to fix a fault. I've no doubt we'll get there with predictive maintenance, but it's very early days," she says.

German plant management firm BayWa r.e is using big data to improve asset management. In July, it signed a partnership with Canadian company PowerHub to develop a cloud-based platform that will provide clients with a near real-time view of operating and financial information, and performance data of renewable energy assets. It also automates workflows, reporting, document and task management and ticketing.

However, despite all this research and innovation, the PV sector is still learning how best to make the use of big data, and where the pitfalls lie. Commentators agree that there is a risk that asset owners and maintenance companies end up with data overload, and that collection of data must be focussed on practical application to be of use.

"One of the big challenges for the industry is that no-one wants to throw any data away in case its useful one day, but no-one knows what it will be useful for," says Kelsey.

"There are lots of potential applications from big data, but if you don't know what you're going to get out of it, how can you say that you need it? All these companies are doing fancy things like sending out drones to take thermographic scans, collecting information every minute, or every 15 minutes – it's good to do that, but you have to know what you're going to use it for because you need to justify the cost," she says.

Data can also be unreliable, and contain gaps. "It's not the case that operators haven't got the teams to analyse all this data. It's more that a significant amount of data is unreliable or wrong. Big data, machine learning and artificial intelligence are all based on data, so the first thing you need to do is make sure the data is correct," says Maier.

Kelsey agrees. "You can spend a lot of time cleaning up data so you have a good dataset that you can go and analyse," she says.

Kelsey also raises a question mark over the usefulness of data given the relative youth of the industry. "Everyone tells me that it gets interesting when plants are 10 years old, because that's when components start to fail. But we don't have experience of that yet, because plants are not old enough. It's a challenge that comes with a young industry, but eventually we'll get there.

Harssema agrees that the data must be made meaningful. "We say that we take data and turn it into information – it's the information that comes from the data that is useful."

However, according to Papaeconomou, there is no such thing as too much information. As long as PV companies standardise the data, and have the appropriate infrastructure to deal with it, it can be made useful, he says. "Once you have these two things, the possibilities are endless," he says.

"Regardless of how much data you have, you design a report to take an abstract of that – then you don't care if there are terrabytes of data," he says.

Data protection

Use of big data brings up questions of data ownership and privacy. Virgil Cazacu, expert for digital transformation at BayWa r.e, says: "It's important to know what data you can touch, and what you can expose to your customers and partner networks, especially since the EU's General Data

Protection Regulation [which came into effect in May this year]. You also need to make the system secure to prevent hacking," he says.

Prashant Khorana, a power and renewables consultant at analysts Wood Mackenzie, believes that data protection issues could be one of the reasons why use of big data is still in the early stages in the PV industry. Asset owners may not want to give a third party access to data, he says.

In any case, it is in the asset owners' interest to keep track of their own data, rather than a third party, says Harssema. "Every solar PV plant will need refinancing, and having operational data is crucial in determining the value of a plant," he says.

While use of big data in the PV sector remains at the early stages of research and experimentation, experts foresee that much greater use is not far in the future. Papaeconomou believes it will not be too long until systems cover the whole lifetime of an asset, from design, through construction, commissioning and operation, rather than multiple systems, as is the case currently.

"Imagine a platform that will cover all your needs for the whole lifetime of a renewable energy asset. It's going to take time to get there, but that's the way we see things developing," he says.

> "All these companies are doing fancy things like sending out drones to take thermographic scans, collecting information every minute, or every 15 minutes – it's good to do that, but you have to know what you're going to use it for because you need to justify the cost"

Harssema sees more advanced uses developing for big data in predicting the failure rate of technical components, in particular, degradation rates in different climates around the world. "We're benchmarking different PV plants to identify performance in different geographical regions. We know that PV components deteriorate over time, and temperature is an important part of it, so we want to know what could be the geographical effect of a low temperature area to a high temperature area in the failure rate of a component."

Big data will also find more uses in combination with AI, for example, to balance the grid using storage, or to operate drones to scan PV equipment for faults, commentators predict.

Landberg sees using big data and machine learning together to improve the accuracy of grid forecasting, and also the levelised cost of electricity. Combining information from power plants with weather data and tweets from relevant organisations or politicians could lead to more accurate forecasts of energy prices, he believes.

"We are in the hype phase now, everyone wants to use big data and machine learning on everything, which is good because then we can explore what it really can be used for," he says. "Things will calm down and by having gained experience, people will know when to use it and when not to use it." ■

Catherine Early is a freelance journalist.