Short-term solar forecasting for utility solar sites: core business or distraction?

Forecasting | Solar farm operators are coming under growing pressure to help maintain network stability by providing accurate generation forecasts. As Nick Engerer writes, innovations in short-term forecasting are providing operators with new options for making this task less of a distraction from their core business

Credit: ARENA

Short-term forecasting is becoming an increasing burden for operators of utility solar farms, such as Australia's Broken Hill

As the rate of growth in installed solar capacity accelerates globally, along with it has come a new suite of challenges to integrate the incredible volumes of new capacity into our energy markets and networks. With just shy of 100GW of new solar added in 2017, following on from 75GW in 2016, many energy markets around the world have experienced a breakdown in the long-established processes and methods for forecasting supply and demand, and planning and operating electrical networks. Load forecasting models have begun to stop working. Increasing volatility is being seen in ancillary services markets. Energy market trading strategies are shifting. The long-term business models dependent on wholesale electricity prices are being challenged in new ways. Such are the growing pains of our global transition to depending on the weather conditions to supply our energy needs. However, there is one segment of the solar energy industry that is beginning to bear the responsibility of these consequences in an unexpected way, as often, the response of the local energy market operator or system operator is to pass the responsibility for delivering accurate forecasts to the market or network onto the owners/operators of the solar farm. This is a trend that is just getting started, and likely to expand significantly over the coming decade.

For a solar farm operator, this can be a significant distraction from core business. I've rarely met a solar farm owner, operator, asset manager or O&M provider who wasn't actively managing several projects, in various stages of their life cycle. Prospecting, planning and modelling, completing connection agreements, reaching financial close, commissioning and then the day-to-day operations are not an easy set of tasks to manage. Plus these tasks all tend to arrive at different times across a growing pipeline of projects, then interspersed by unexpected problems. When these teams are required by their local market or system operator to start delivering a forecast for their solar facility, yet another responsibility is added to an already full agenda.

What's more, inaccurate forecasting can often be tied to financial penalties, imposed by the market or system operator, as well as lost revenue and adverse impacts upon the bottom line of the project as a whole. For example, in Australia, solar farm operators are required to pay fees for ancillary services to offset forecasting errors. In India, state distribution companies are now due to enforce penalties for inaccurate forecasts, while in Europe and the United States, forecasting services have been required for a number of years already, and grid operators there are increasingly using constraint or curtailment in lieu of direct financial penalty. The exact forecasting requirements and consequences for solar farm operators vary significantly by region, but one aspect of this transition is clear: as penetrations increase, accurate forecasting becomes more important, particularly at shorter timescales (minutes to hours ahead). This is apparent in the increasing importance of the 15-minute-ahead forecast in India, the 5-minute-ahead forecast in Australia and 5-minute market operating in California, each of which has seen very fast growth in utility scale solar farm sites.

Short-term solar forecasting is fundamentally different from hours-ahead or day-ahead forecasting

At these forecasting horizons (minutes to hours ahead), another challenge appears, this time for the solar forecasting technologies themselves. Over the past decade, day-ahead forecasting for unit commitment and hourly forecasting for load balancing across a given day have been the status quo. And for these purposes, modelling approaches based on numerical weather models and machine learning techniques have been suitable. Weather models, with rather coarse resolutions of tens of kilometers, produce forecast outputs in three-hourly increments (with some models now producing hourly outputs), and generate estimates of moisture content by layer of the atmosphere. These can be tweaked to represent cloud forecasts, and make general estimates of the future solar radiation resource. Machine learning techniques are often applied against actual measured generation from solar facilities or solar radiation measurement sites, to generate regressions to 'fit' them to the weather model data.

However, these approaches are far from suitable for making a short-term prediction of solar farm power output on the timescale of minutes ahead, for one fundamental reason: none of them know where the clouds actually are at any given time. Contrary to popular belief, weather models do not actually model cloud conditions directly, they instead rely on parameterisations and other tricks to handle them at a broad scale. Their purpose is to capture rainfall development and to calculate overall incoming and outgoing radiation balance so as to make a prediction of temperature and relative humidity. They aim to make predictions of the weather that are suitable for a wide variety of purposes, and are not made to track or forecast actual cloud cover features. Furthermore, they update only every 3-6 hours, and rely on outdated cloud data sources or proxies to do so. Even with the best machine learning methods available, if the forecasting approach doesn't actually use knowledge of local cloud cover conditions, it will fail to make accurate short-term forecasts.

Satellite based solar forecasts: better, but not always

This is where satellite based imagery and forecasting can offer significant improvements. Weather satellites are placed in geostationary orbits where they record images of the Earth, which includes the surface and cloud cover. These images update every 10-15 minutes, and are available resolutions as fine as 1km^2, which is exactly the kind of data input short-term solar forecasting methods should be utilising.

Yet, not all satellite-based forecasting services are the same. Just because one can pull in this imagery data does not actually mean they are able to identify clouds or track their motion with precision, let alone model the amount of solar radiation arriving at the Earth's surface or the eventual power output from a solar facility.

Take, for example, making an estimate of the energy generation at a solar farm facility for the current time (no forecasting involved). To do this, there are several important modelling steps in between. Satellite images must be decomposed into cloud versus no-cloud regions. Cloud opacity (its opaqueness to light transmission) must be estimated. The location of the cloud shadow needs to be projected spatially. Appropriate application of solar radiation modelling tools is then required, as well as PV power plant modelling to create the energy generation estimate. This is not an easy set of tasks, which each rely on domain expert knowledge to implement. This type of expertise is hard to come by.

Assuming these steps can be competently managed, using satellite imagery to produce a forecast is an entirely different application. Clouds are always changing; whether moving their position, forming, decaying or changing shape, cloud cover is a rapidly evolving phenomenon, whose underlying microphysics are not yet completely understood. They are affected by terrain and the types of land surface beneath them, and move through the atmosphere according to the wind velocities at their given level of the atmosphere. This last point is particularly important, as wind shear, the change of wind direction and intensity with height, means that clouds at a lower level of the atmosphere are often moving in a different direction to clouds at higher levels. Unfortunately, nearly every satellite-based forecast methodology in practice oversimplifies these problems, extrapolating a motion vector from previous cloud movement and projecting the cloud positions forward all as one mass, in the same direction, to make a forecast.

Incorporating solar forecasts into utility-scale operations

So let's loop back to our overall storyline here. We have incredible growth in solar installations globally, primarily driven by our hard-working solar farm owners and operators. In response to the large volumes of solar being added to energy markets and networks, they are increasingly being tasked with delivering short-term predictions of the energy generation from their facility to their local market or network. This is outside core business, and is arguably a distraction from their mission. The short-term solar forecasting they are tasked with acquiring depends on expertise and techniques that are difficult to acquire, with many forecasting providers using outdated approaches, and possibly altogether lacking the expertise required to engage with a complex meteorological problem. Where does this leave asset managers, O&M providers, solar farm owners and operators, but in a difficult position, with exposure to financial penalties and a responsibility to choose a forecast provider in a time-poor environment? Is that a recipe for building the solar powered future, at an ever increasing pace? Not in my opinion.

Making solar forecasting easy

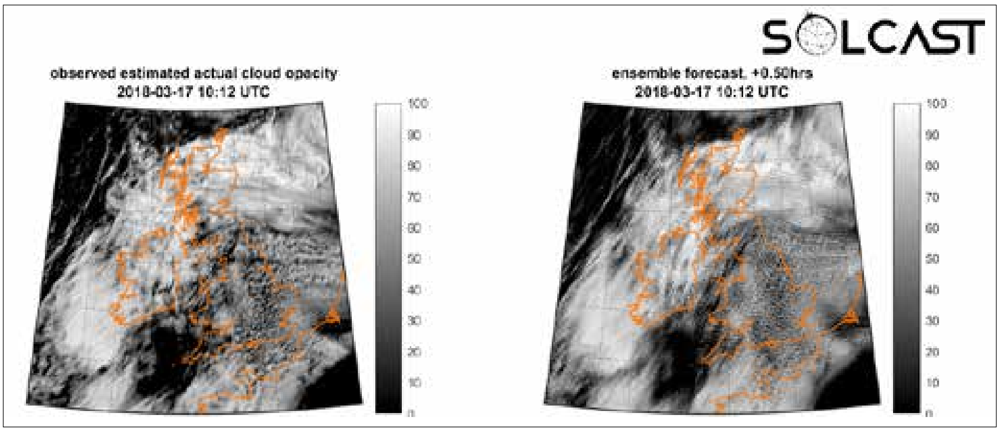
This is challenge is precisely why I co-founded Solcast along with our CEO James Luffman. As lifelong meteorologists, we have both been long fascinated by cloud cover and the processes that drive it. With all that we have learned about meteorology over the years of studying, practising it, teaching it and managing companies that retail products based on forecasting services, we have come to see cloud cover forecasting as a place where we can dedicate our talents, skills and interest to further the accomplishments of the solar energy industry. As a result it has become our team's goal to enable solar farm owners and operators to accomplish their core mission: planning, building and operating their facilities, by making solar forecasting services easy to access and assess. By applying our domain knowledge to these goals, we have taken on the tasks of managing the incredible volumes of data produced by weather satellites and sorting through the complexity of deploying high quality solar radiation and PV plant modelling, so that our solar farm owners and operators can get back to hard work of building the solar powered future.

Reflecting on that journey to date, our team has found many of the barriers to generating satellite based solar forecasts and making them easy to access quite remarkable. Along the way, these have forced us to be quite creative, and to turn to new solutions such as cloud-based computing infrastructure, APIs and machine learning.

Challenge: Generating satellite-based solar forecasts globally

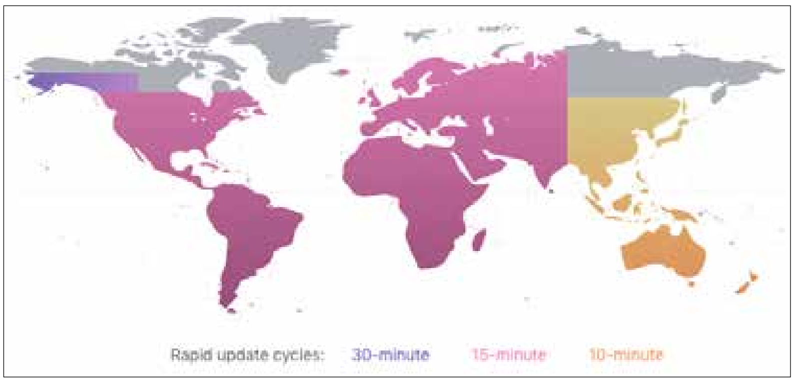
There are several geostationary weather satellites in orbit around the planet, and altogether Solcast currently pulls in data from five of these. Several of these are next-generation weather satellites that provide new imagery of the Earth's surface and its cloud cover every 10-15 minutes, at 1-2km^2 spatial resolution – including the GOES-17 satellite, which is likely to have begun operations over western North and South America and the Pacific by the time of the publication of this article. Altogether, this stacks up to 30-40GB of satellite imagery data each day that need to be downloaded, processed, quality controlled and then delivered to the forecasting algorithms. Some quick mathematics will tell you that adds up to more than 12TB of storage just for the satellite images, each year! After this data pipeline is managed, algorithms must then be applied to track the actual locations, determine the characteristics of cloud cover in real-time, and estimate how thick they are with respect to sunlight. Each of these processing steps are dealing with geospatial data in the order of hundreds of megabytes from the satellites alone, with several additional gigabytes of data added once information from numerical weather models is also incorporated.

All of that, and we still haven't even crunched a single forecast! In fact, it is the forecast computation and delivery that really pushes the computational limits, as each satellite is sending Solcast new data every 10 to 15 minutes. At each scan interval, we have to re-forecast for every location on the planet, our prediction of cloud cover and solar radiation over the next four hours. At each forecast step, that involves applying machine learning, to review the previous round of forecasts, in order to determine how to adjust the next lot of forecasts according to the local weather conditions (are clouds fast changing? Slow moving? etc.). Looking at these numbers, a total of 600 million forecasts are generated every hour, adding up to more than 9 billion forecasts each day. When our team was first planning on deploying forecasting services to a global audience, these numbers shocked us. In fact, they were about double that estimate when we first worked it out on paper. But along the way, we found several ways to be smarter about managing our total volume of data, but more importantly, we also discovered the answer to the problems of deploying a global cloud forecasting service – by, ironically, moving them to the cloud!



Rapid update forecasts over the UK

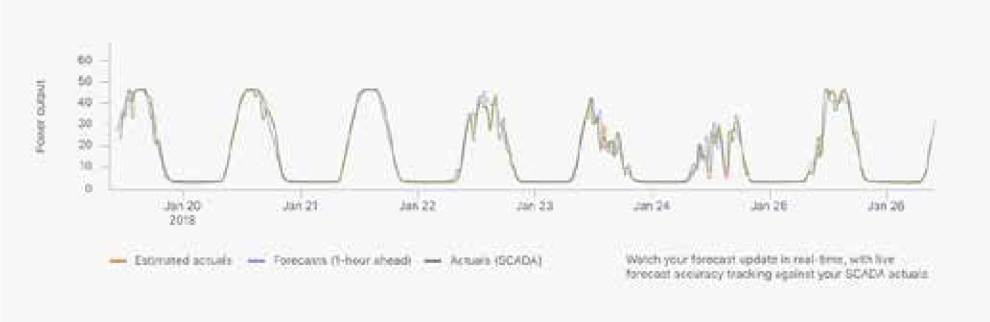
In response to the fantastic volumes of forecasting data required to deploy a global solar forecasting service, Solcast made an early stage decision to operate from the Amazon Web Services (AWS) platform. There, we were able to take advantage of new technologies like Redis databases and the large storage volumes of S3 and long-term storage solutions like Glacier to reach the levels of performance required for operational services, whilst not totally upending our operational budget. The flexibility of spinning up new instances for compute and memory power according to our current needs was also a powerful option. What's more, this decision then allowed us to pivot into another AWS-based solution, this time confronting another challenge: making those solar forecasts easy to access.



Global coverage map

Challenge: delivering solar forecasting data

By switching over to AWS, were able to take advantage of a whole host of solutions that



make deploying an API framework much easier. This includes tools like load-balancers which can accommodate a sudden influx of new API requests, and Elastic Compute Cloud (EC2) instances to power the API itself. We elected to move ahead with a REST API running on Microsoft's .NET Core technology, which was made very easy by using the ServiceStack framework. Servicestack made it straightforward for Solcast to deliver solar forecasts via HTML, JSON, XML, csv and jsv formats. We highly recommend this approach for other software engineers in the solar energy industry looking to deliver data via API.

One of the best things about an API-based service is that, unlike FTP or other similar approaches, it allows a two-way flow of data. And that two-way flow of data, via API, allowed us to tackle yet another challenge in delivering utility scale solar farm forecasts….

Challenge: every solar farm is different

It's one thing to deliver a forecast of solar radiation; it is a completely different task to deliver a forecast of the power output from a utility-scale solar farm. Yet this isn't traditionally how this problem is approached. Standard practice in the industry is to use a forecast of Global Horizontal Irradiance (GHI) and other related solar radiation or weather parameters, as inputs to a PV power plant model. And while we're happy to provide solar radiation forecasts, Solcast has hardly been satisfied with this solution. It's just not easy enough.

And that is because utility-scale solar installations are often quite complicated! Even with perfect execution against a construction plan, once commissioned, solar facilities encounter a host of real-world challenges that can impact performance in many ways. Modules and other hardware are split amongst many different arrays, each potentially impacted by topography, and having widely varying orientations. They are often differentially shaded, are always accumulating soiling and dust, and also experience different rates of degradation. Where does this leave solar farm owners and operators, but with another round of headaches? However, as alluded to earlier, APIs can provide us a clever solution to this problem.

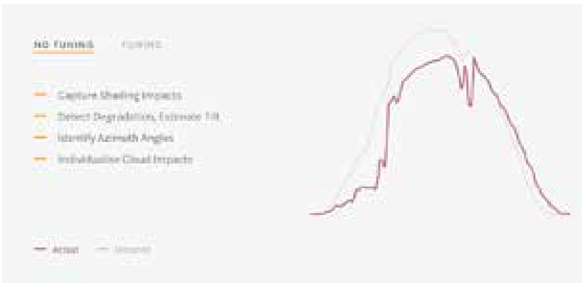
Utility-scale solar farm forecasts can be tracked against SCADA data

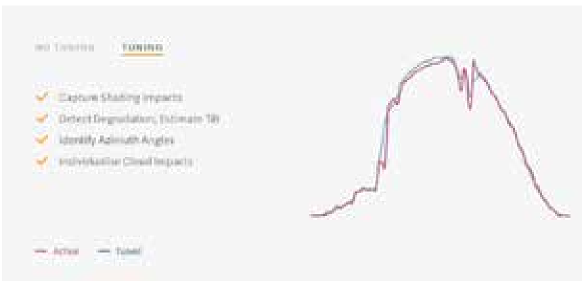
APIs are a two-way street. Users can use a GET request to retrieve their forecasting data, which is where the traditional solar forecast customer relationship usually ends (e.g. grab solar radiation data from an FTP), but with an API, users can also POST request, meaning they send data to the API. We've used this opportunity to allow our customers to send recent SCADA data measurements back to Solcast, which then allows us to confront the issues of uniqueness in each solar plant, with a little machine learning magic.

To take advantage of the ability users have to POST data to the API, we've deployed PV Tuning technologies which apply machine learning to measured power output data in order to learn how the utility-scale solar farm responds to a given set of radiation conditions. By matching up the performance of the solar farm site with solar irradiance and weather data, a 'tuned' forecast that represents that specific facility is produced. Altogether it allows the forecast to:

Tuning technologies learn how utility solar farms respond under certain radiation conditions

- Capture the impacts of shading on your system, including vegetation, topography and surrounding infrastructure
- Detect overall impact of the varying orientations of the arrays in a PV site (azimuth and tilt angles)
- Sense the degradation of your utility scale site, assigning it a loss factor
- Individualise the way your solar farm respond to a given level of cloud cover
- Use these learned parameters to provide you with an improved solar forecast specific to that asset
- Accommodate both fixed and single-axis tracking sites

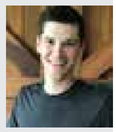




Closing thoughts

Upon review, solar forecasting at scale isn't easy. Simply put, it is a "big data" problem. But new solutions such a cloud computing resources and APIs are enabling a new suite of solutions for solar forecasting, which is good news for our hard-working solar farm owners and operators. These folks are deploying impressive amounts of solar all around the world, building the solar future on our behalf. But as time goes on, that task becomes increasingly challenging, as they are tasked with delivering short-term power forecasts to their local market operator. Since this is decidedly outside of core business, as it is not aligned with their mission, they deserve solutions which remove the pain of having to source and implement them. But with many forecasting providers using outdated approaches to generate, deliver and assess their forecasts, we are confident there's room for improvement. Hopefully we've done our part to make your life a little easier. ■

Author



Dr. Nick Engerer is the CTO of Solcast. He is an expert in the field of solar radiation modelling and forecasting, and has co-founded Solcast out of a desire to enable others to build the solar-powered future. Nick displays an unusual level of passion for developing applied research projects and collaborating with industry to make BIG ideas a reality. Follow him on Twitter @nickengerer or connect on LinkedIn. For further information on Solcast's technologies and details on a free trial of its project forecasting service, visit http://solcast.com.au