Forecasting 101 for the Electrical Networks Sector

A paper by TNEI

Released 24.02.2021

Forecasts have always been essential to the operation of electrical power networks and have been used for many decades for purposes such as deciding the optimal sizing of cables, the required rating for transformers, and the amount of generation that should be dispatched at any given time.

However, their role is rapidly becoming ever more important and the way they must be used is becoming ever more complex. Quantities being forecast range from the output of a wind farm 10 minutes ahead to the maximum active power demand on a transformer during the next winter, to the way most people will heat their homes in 30 years. Some of the reasons behind this rising importance include:

- The shift towards more decentralised generation;
- The shift of many customers at all voltage levels from passive consumers to active prosumers;
- The shift from dispatchable generation to variable and only partially predictable renewable generators;
- The current expectation of a much leaner system than several decades ago, in terms of spare capacity; and
- The explosion in data availability and computing power to allow much more wide-spread, granular and frequently updated forecasts.

The general notion of a forecast is obviously familiar to everyone – particularly in a country famously obsessed with the weather.

"The new prominence of forecasting, particularly as part of a probabilistic perspective, is increasingly demanding a degree of expertise from many power system engineers"

However, the new prominence of forecasting, particularly as part of a probabilistic perspective, is increasingly demanding a degree of expertise from many power system engineers. As a result, it may be the case that some industry professionals might not feel entirely clear or confident about some of the more specialised concepts and terminology associated with the production and evaluation of forecasts. To address this, we have produced this 'Forecasting 101' article that we hope de-mystifies the topic, explaining the concepts in clear and simple terms. We will also cover some aspects that are particularly relevant to our sector but, perhaps most importantly, will not bog you down in any mathematical detail. It will hopefully make you think about forecasting in a new light, and we hope you find it helpful and interesting.

What is a forecast?

We all have some understanding of what forecasts are – they're part of the fabric of life. We might describe them in everyday terms as something like '*a prediction by experts of something that's **likely to happen***', or we might even say '*a prediction by experts of something that **they believe** is likely to happen*'. Even though we know that forecasts are never perfect, they still represent our best guess of what will happen i.e. our belief about the outcome in question, and we may take actions or decisions on the basis that our belief is true, or at least reasonably close to being true.

Of course, it's not only experts that can make forecasts – if you knew there was going to be a football match between FC Barcelona and your local pub team, you wouldn't need to study the players' recent match statistics to forecast that Barcelona would deliver a thrashing, and to have a very strong belief in that outcome. To give an example relevant to our sector, you wouldn't need a sophisticated mathematical model to help you forecast that the daily national demand peak on a sunny Sunday in early August is going to be among the lowest of the year. However, you are likely to need such a model to predict the exact number of Gigawatts (GW) four days ahead.

Although functional, the definitions of forecasts given above don't quite get to the bottom of the matter. A much more formal but insightful definition could be: '***an expression of belief as to the outcome of an uncertain event, or the value of an uncertain quantity***'.

That definition definitely needs some unpacking!

Let's break it down by returning to our example of that rather hypothetical football match. Upon hearing that this match will take place, it would be reasonable for you to issue the forecast: "FC Barcelona are very likely to win!" That would be an expression of your belief about an uncertain outcome – the winner of the match – and life experience certainly justifies that belief. You may also feel inclined to offer a forecast for the final score, and in that case you would be justified in predicting that the goal difference would be very large, but would also realise that there are a rather large number of plausible outcomes, and the probability of getting it exactly right is not very high. For example, you may forecast a result of 50-0, but acknowledge that the probabilities of it being 51-0 or 49-0 are also pretty similar, while the probabilities of the results 0-1 or 10,000-2 are both much lower. By setting out several such justified beliefs about the result, you could formulate a fairly detailed expression of beliefs about the uncertain outcome.

Returning to the example of daily peak demand, you could be asked to guess which day in 2020 had the lowest peak. A notable difference here is that all daily peak demand values already exist, since all days have already occurred. However, it is highly unlikely for anyone to know the answer, even if you work for NGESO, so in terms of expression of belief it is an uncertain outcome until you either look it up in a dataset or someone tells you the correct day. It therefore makes sense to think of (i) forecasts for events that haven't happened yet and (ii) any outcome or value that you do not yet happen to know, as being exactly the same. Like the previous example, you might forecast that it occurred on the 1st Sunday in August, while believing that the probability of it being the 2nd Sunday of that month is only slightly lower, but that the probability of it being the 1st Monday in February is much lower.

How are forecasts generated?

Some people might generate a forecast by reading tea leaves – this is certainly consistent with the definition of a forecast, even if the wisdom of trusting such forecasts is questionable.

In our two examples above we considered forecasts that definitely had stronger justifications, but were also rather limited by being based purely on 'common sense' and general life experience. However, when operating a transmission or distribution network, or trading in an energy futures market or any of the many roles connected to the energy system that rely on forecasts, you clearly want a much more rigorous and precise basis for generating them. Indeed, it is much better to use forecasts generated by a physical or statistical model, or ideally a combination of both.

"Statistical models are in increasingly common use for electricity network/ energy system applications"

With a physical model, the main application is forecasting the weather from several hours to several days ahead. This is based on obtaining the best possible representation of the current physical state of the atmosphere and calculating what the laws of physics tell us about the way that physical state will change. Weather forecasts derived from such Numerical Weather Prediction (NWP) like these are very commonly used by network operators.

Statistical models are in increasingly common use for electricity network/ energy system applications. These are models that have been fitted to appropriate data using mathematically justified and tested methods. Such models are implemented by forecasting algorithms that collect relevant input data, such as the outcomes of recent events, and uses them to produce a prediction in accordance with the statistical model. This may be an ongoing process whereby the algorithm constantly collects fresh data and outputs new forecasts.

Types of forecast

For many forecasts, the outcome or quantity we are making predictions about can only take one of a limited number of values – e.g. 'heads' or 'tails' for a coin toss, or the numbers 1, 2, ..., 6 for dice. Considering a hypothetical situation where we are interested in an uncertain outcome that may be A, B or C, the simplest type of forecast simply chooses which outcome is most likely. The forecast may state that 'the outcome will be B' – and there is a tacit understanding that this means that B is believed to be most likely. Alternatively, if the options are ordered (e.g. small, medium, large) it could be that the forecaster believes that the likelihoods of each outcome are similar and it makes sense to choose the central option. If the outcome indeed turns out to be B, then we say the forecast was correct, while if the outcome turns out to be A, we say the forecast was wrong. We might be able to say that a statistical model generating such forecasts is correct e.g. 90% of the time, in the long run.

> *"A more sophisticated forecast will be explicit in its expression of beliefs about the outcome. These are known as probabilistic forecasts"*

A more sophisticated forecast will be explicit in its expression of beliefs about the outcome. These are known as ***probabilistic*** forecasts. For example, it might state: 'there is a 10% chance the outcome will be A, an 85% chance it will be B and a 5% chance it will be C'. This forecast clearly presents the outcome in the form of a ***probability distribution*** – i.e. a probability associated with each possible outcome, where they must sum to 100%. In this context, regardless of the outcome, we cannot ever say that any individual forecast was 'correct' or 'wrong'. If the outcome of the event for our example probabilistic forecast turns out to be C, we can say that this was quite unlikely given our forecast, but not that the forecast was wrong.

Although we can never say that individual forecasts were right or wrong, we can assess the quality of the (usually statistical) model that produces them. This involves mathematically analysing many forecasts generated by the model – ideally 100s or 1000s of them – which allows us to come to firm conclusions about the ***skilfulness*** of the forecasting model, and of course to compare the skill of different models when using the same data set.

Other quantities such as peak MW demand or thermal utilisation of a cable or height have continuous values, which means there is an infinite number of possibilities for each outcome, even if there is a limited range in which it must lie. That is because a demand value could turn out to be e.g. 7.813286 MW or 7.813287 MW – the point being that there's no limit as to how close those distinct values can be, so there are an infinite number of possible values. In this case, the simplest type of forecast will say 'the outcome will be X'. Such a forecast is, necessarily, ***always slightly wrong***, but the skill of the forecasting model can be measured by how far apart the outcomes and predictions are on average. This distance is measured using metrics such as the root mean squared error (RMSE.) or mean absolute error (MAE).

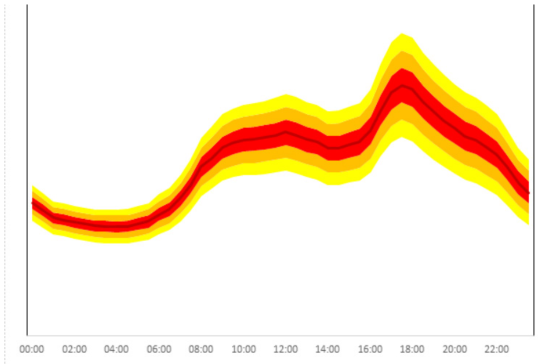
A more sophisticated type of forecast for continuous valued quantities (i.e. a probabilistic forecast) take the form: '***there is a 95% chance*** that the outcome will ***lie between*** A and B'. This is clearly another type of forecast that constructs a probability distribution as an expression of the full set of beliefs about an uncertain outcome or value. In the continuous case, there are many alternative ways in which the probability distribution may be presented. For example, the forecast could simply state: 'the expected value for the outcome is X', or 'the median value for the outcome is X' – meaning that there is a 50% chance that the outcome will be smaller than X. Another option is to present plausible worst cases, e.g. 'there is only a 5% chance that the maximum current experienced this winter is bigger than X Amps'.

The value and suitability of different types of forecast

The existence of two very different types of forecasts described in the previous section, i.e. those described as the simple types and the more sophisticated alternatives, can be attributed to a fundamental difference in the assumed purpose of forecasting. The first assumed purpose is that forecasts should simply present the forecasting model's (or human's) ***'best guess' of what will happen.*** This is typically paired with the assumption that any decision-making that relies on the forecasts should be taken on the basis that the outcome will certainty match the forecasts perfectly. This is despite knowing that the model's forecasts are, in reality, not completely reliable (for discrete variables), or always slightly wrong (for continuous variables). There are some notable exceptions, such as NGESO reserve setting, but the control room does embrace probabilistic thinking (if not in a fully mathematically robust way).

The more sophisticated versions described in the previous section are used when it is assumed that the purpose of a forecast is, rather, to provide ***a full summary of what is known about, or perhaps informed opinions about, uncertain future events.*** Such full summaries are represented mathematically by probability distributions, as in the examples provided. It has been established by mathematicians, particularly those working in the fields of decision theory or operational research, that such full summaries are necessary to make truly optimal decisions, when those decisions must be made under uncertainty and where the costs of underestimating are different to the costs of overestimating. By uncertainty here, we mean that a sufficient degree of uncertainty exists whereby ignoring it can lead to significantly sub-optimal decisions, with material consequences. It is not always obvious, of course, when this is or isn't true – but some clear examples are e.g. the annual peak demand that a primary transformer may experience in 10 years' time, or the power output of a wind farm in two days' time.

Probability distributions provide decision makers with much more valuable information than



Example of a `fan chart': showing the central forecast for a number of look-ahead times, surrounded by shaded areas representing uncertainty, and where darker/ stronger colours represent higher probability. For example, there might be a 50% probability of the outcome being within the red band, and a 90% probability of being within the orange (and red) band.

simply the 'best guess'. They can be a simple set of numbers or a mathematical function characterised by one or several a number of parameters, but either way they are simply a formal expression of everything the person who creates or uses the forecast believes to be true about the outcome or quantity in question. It therefore makes sense that mathematicians have established that, if a forecast is being used to identify or calculate optimal decisions, it is essential that the forecaster passes on all their beliefs to the decision maker, by constructing a full probability distribution.

However, it is important to note that providing a full probability distribution may only be beneficial if a decision maker using the forecast has a mathematically derived process for making decisions that require it. Providing a single value, such as the median or a specific percentile may be more useful instead. Providing a particular percentile may be particularly appropriate in situations where the decision maker must be very risk averse. Although the forecast being provided is simply a number, producing a probabilistic forecast first is a prerequisite for ensuring that it is the most helpful number for a specific decision-making process, rather than a central estimate or best guess.

It is also worth noting that probabilistic forecasts allow the development of useful visualisations to support decision-making, even if the forecasts themselves are single numbers or categories. In fact, it may be possible to produce several complementary visualisations from a single probabilistic forecast e.g. separately visualising the uncertainty in the magnitude and timing of a peak demand level.

Essential characteristics of forecast models

So far, we have looked in depth at one of the key distinctions between forecasts: whether they chose a single value or option to be taken as 'what will happen', or whether they associate probabilities with every possible outcome. However, there are many other attributes and parameters that distinguish one forecast from another. These key characteristics are listed below:

- **The look-ahead time** for a forecast is the length of time between the period (time step) for which the forecast is being made, and when it is issued – e.g. day-ahead or week-ahead.
- Whether the forecast model only issues forecasts with ***a single, fixed look-ahead time, or instead issues forecasts for a number of consecutive time steps*** – e.g. each of the 48 time-steps between 1 half hour ahead and 24 hours ahead. In that case, the largest look-ahead time (24 hours) is called the **forecast horizon.**
- **The temporal (and spatial) resolution** of a forecast is the length of each time step, e.g. 30 minutes. Forecasts might be concerned with both time and space, in which case it might make separate predictions for e.g. a set of 1km-by-1km squares. In that case the spatial resolution would be 1km.

"The most sophisticated type of forecast provides full summaries of what is known about, or informed opinions about, uncertain future events"

- **The rate of updating**, or **refreshing**, refers to how often forecasts are issued, e.g. every hour, twice a day. If the time between updates is less than the forecast horizon, it is possible for the forecast horizon to either be fixed or variable. For example, a forecast might be issued at midnight, and make predictions for the next day, i.e. the next set of 48 half-hourly periods. An updated forecast might be issued at mid-day that either includes predictions for the next set of 48 time-steps, or the remaining 24 time-steps until the next midnight forecast. It is worth noting that it is possible to regularly update statistical forecasts e.g. every half hour, even if a large contributor to that forecast is issued less frequently e.g. weather forecasts released once or twice a day combined with data about very recent conditions updated every hour, or potentially every minute.

Forecasts versus scenarios

One rather subtle and sometimes complicated topic associated with forecasting in the electricity networks sector relates to long-term planning and the production of high-level scenarios to represent possible pathways for the evolution of the system. In Great Britain this essentially means the Future Energy Scenarios (FES) produced annually by National Grid ESO, and the more granular versions developed by the DNOs.

The interesting thing about these, in the context of this article, is that they are ***not forecasts.*** Rather, they represent several plausible pathways that could turn out to be approximately true, designed to represent the breadth of possibilities without being too extreme. If a planner or policymaker were to decide that one of the scenarios should be interpreted as the most likely out-turn, with the others being alternatives that 'could still happen', then they would be turning the scenarios into a forecast – because they now represent a belief as to what will happen.

> *"Energy system scenarios are not forecasts, rather they represent several plausible pathways that could turn out to be approximately true"*

The same also holds true if probabilities are associated with each scenario – either by those issuing the scenario or added as an expression of expert opinion by a user. It must be noted, however, that allocating probabilities to the scenarios is not necessarily a bad move, and it is interesting to note that applying a least-worst regrets to scenarios without probabilities has been shown to be equivalent to a certain set of probabilities in terms of the timing and scale of investments chosen.

Contact

I hope that you found this article interesting, and that it has cleared up a few of the finer points of what forecasting is really about, along with some of the terminology commonly used. If you would like to ask any questions related to forecasting, or to discuss any of the content, we'd love to hear from you! If you're interested in developing any ideas for a forecasting-based project, even better.

Gruffudd Edwards
Senior Data Scientist

Get in touch

Gordon McFadzean

Gordon has a broad background covering energy, engineering and economics, with extensive experience of techno-economic modelling and cost benefit analysis of power systems and innovative technologies. He has worked with network operators, regulators and governments to consider issues related to system planning, price controls, and network charges. Recently, he has focused on topics like low voltage networks, demand side response, electrification of heat and transport and smart meters.

Email: gordon.mcfadzean@tneigroup.com

Gruffudd Edwards

Gruffudd is a senior data scientist with considerable experience in mining, analysing, interpreting, predicting and visualising data. His academic background spans statistics, electrical engineering and physics, and his research has often bridged these disciplines, covering topics like generation capacity adequacy, the operation of energy storage, and the role of smart meters in low voltage planner. He has strong expertise in stochastic power system modelling and the application of operational research methods to power system problems.

Email: gruffudd.edwards@tneigroup.com

TNEI is an independent specialist energy consultancy providing technical, strategic, environmental and consenting advice to organisations operating within the conventional and renewable energy sectors. Our consultants have industry leading expertise in grid code compliance studies, noise assessment and analysis of innovative, smart grid technologies.

Our clients range from large utilities and network operators to regulators and community groups We are a specialist, independent company, and that's why we can offer a flexible, personal service and help our clients quickly and efficiently. And most importantly of all, we love to solve problems.

To find out more about how we can assist in this sector or to share feedback on this article, please contact us:

info@tneigroup.com
+44 (0)141 428 3180

7th Floor
80 St. Vincent Place
Glasgow
G1 2DT

tnei

www.tneigroup.com
info@tneigroup.com