

The Value of Information in Managing the Electricity System

JEREMY BUDD¹, CHRIS DENT², DANIEL MAXWELL³, ROB TOVEY⁴,
FRANCIS G. WOODHOUSE,⁵ AMY WILSON² and STAN ZACHARY^{2†}

¹ *Delft University of Technology, Netherlands*

² *University of Edinburgh, UK*

³ *University of Exeter, UK*

⁴ *University of Cambridge, UK*

⁵ *University of Oxford, UK*

(Communicated to MIIR on 31 October 2021)

Study Group: ESGI145, 8–12 April, 2019, Cambridge, UK

Communicated by: Stephen Cowley

Industrial Partner: National Grid Electricity System Operator

Presenter: Andrew Richards

Team Members: Jeremy Budd, Delft University of Technology, Netherlands; Chris Dent, University of Edinburgh, UK; Daniel Maxwell, University of Exeter, UK; Rob Tovey, University of Cambridge, UK; Francis G. Woodhouse, University of Oxford, UK; Amy Wilson, University of Edinburgh, UK; Stan Zachary, University of Edinburgh, UK;

Industrial Sector: Energy/Utilities

Key Words: Electricity transmission networks, Renewable energy, Data Uncertainty

MSC2020 Codes: 62, 68

Summary

National Grid (NG) Electricity System Operator is responsible for balancing supply and demand in the GB electricity transmission system. To do so it requires accurate real-time information on both generation and demand. However, much smaller-scale and highly variable renewable generation is embedded directly in the distribution networks. NG does not have direct metering of how much power this generation is providing at any given time. This lack of knowledge increases the constraint and other costs to keep the system balanced. The problem posed is that of estimating the potential savings of having accurate real-time knowledge of such embedded renewable generation. This report considers two approaches to making such estimates and briefly explores a statistical approach to the modelling of non-constraint costs. Suggestions are made as to what further information might be collected by NG to make improved estimates of the benefits in system operation of direct metering of embedded resources.

† Corresponding author: s.zachary@gmail.com

Report authors

Chris Dent, Amy Wilson, Stan Zachary (University of Edinburgh)
Francis G. Woodhouse (University of Oxford)
Jeremy Budd (Delft University of Technology)
Rob Tovey (University of Cambridge)
Daniel Maxwell (University of Exeter)

Executive Summary

National Grid (NG) Electricity System Operator is responsible for balancing supply and demand in the GB electricity transmission system. To do so it requires accurate real-time information on both generation and demand. However, much smaller-scale and highly variable renewable (wind and solar) generation, together with some non-weather generation, is “embedded” directly in the lower-level electricity distribution networks. NG does not have direct metering of how much power this generation is providing at any given time, and for renewable generation must rely on forecasts which are typically several hours old. This lack of knowledge increases the constraint and other costs which NG incurs in order keep the system balanced. The problem posed is that of estimating the savings to be made by having accurate real-time knowledge of such embedded renewable generation.

This report considers two approaches to making such estimates. The first uses data made available by NG and relates embedded wind generation in Scotland at times of peak demand to the corresponding constraint costs on the critical boundary between Scotland and the rest of GB. Given a lack of detailed knowledge of the NG Control Room reserve setting procedures, this enables a necessarily very rough estimate to be made of the relationship between the above two variables and finds modest savings to be made by a better real-time knowledge of the embedded wind generation. The second approach considers more generally the value of improved information in decision making, and describes a general probabilistic framework within which, given sufficient data, it is possible to evaluate more precisely the economic benefits – in terms either of economic cost or of control of risk – of having improved information.

The report also briefly explores a statistical approach to the modelling of non-constraint costs, but finds no evidence in such extremely limited data as are available that better real-time knowledge of embedded generation would reduce these. However, much more data is needed here.

Finally, suggestions are made as to what further information might be collected by NG to make improved estimates of the benefits in system operation of direct metering of embedded resources.

Contents

1	Introduction	4
2	Problem Description	6
2.1	The question	6
2.2	The data	8
3	Modelling approaches	9
3.1	A preliminary estimate of the potential reduction in constraint costs	9
3.2	The value of improved information in decision making	12
3.3	Preliminary exploration of non-constraint costs	15
4	Conclusions	18

1 Introduction

National Grid (NG) Electricity System Operator operates the national electricity transmission system in the whole of Great Britain¹. A diagram of the transmission system is shown in Fig. 1. The transmission network consists of lines at 275 kV and 400 kV in England and Wales, and additionally at 132 kV in Scotland. This carries bulk power around the country, and large generators and the very largest demands will be directly connected to it. The local distribution networks provide connections to smaller generators and most end users. Generators are separately owned from the networks.

In GB, the system operator directly manages the market from 1 hour ahead of each 30 minute ‘settlement period’ (the 48 windows per day into which time is divided for energy trading purposes). Ahead of this, generators and demands trade energy bilaterally, or nearer real time through independent power exchanges. For the last hour and during the settlement period itself, generators and demand submit to NG bids (to increase demand or reduce generation) and offers (to do the opposite) which NG uses to balance the system in real time (Fig. 2), termed the Balancing Mechanism (BM). In addition, because forward trading is carried out as if there were no limits on the size of network flows, in the BM NG also have to ensure that flows on all lines are below their allowed limits to prevent, for example, lines sagging below safe heights above ground.

Accurate knowledge of supply is critical to optimising the BM. NG has direct metering of all generation connected directly to the transmission network, and also of all distribution-connected generation of capacity above 100 MW in England and Wales, and over lower limits in Scotland. Historically, the majority of the generation had been large transmission-connected, but in recent years there has been a rapid growth of smaller distribution-connected generation, termed ‘distributed’ or ‘embedded’ resources (Fig. 3). All solar generation is distributed: the largest solar farm in England and Wales is about 70-80 MW, and for obvious reasons it is uneconomic to build large solar farms in Scotland. Distributed wind has grown relatively slowly, as much recent wind development has been large scale (including offshore), and 43% of total wind capacity is in Scotland, and 24% of the Scottish wind generation is not metered by NG. Non-weather distributed resources are particularly tricky, as even if there is knowledge of location and capacity there may be no knowledge of technology, and some generation might be ‘behind the meter’ co-located with demand in which case it may be entirely invisible to the network (e.g. backup generation for industrial sites which also runs to sell energy when prices are high).

The consequences for the effective demand which NG must meet are illustrated in Fig. 4. ‘National demand’ is the effective demand which National Grid must meet, defined as the

¹See <https://www.nationalgrideso.com/>. The transmission network in England and Wales is owned by National Grid Electricity Transmission, <https://www.nationalgridet.com/>, which is a separate company within the National Grid Group. The transmission network in Scotland is owned by Scottish Power and Scottish and Southern, but operated by National Grid Electricity System Operator.

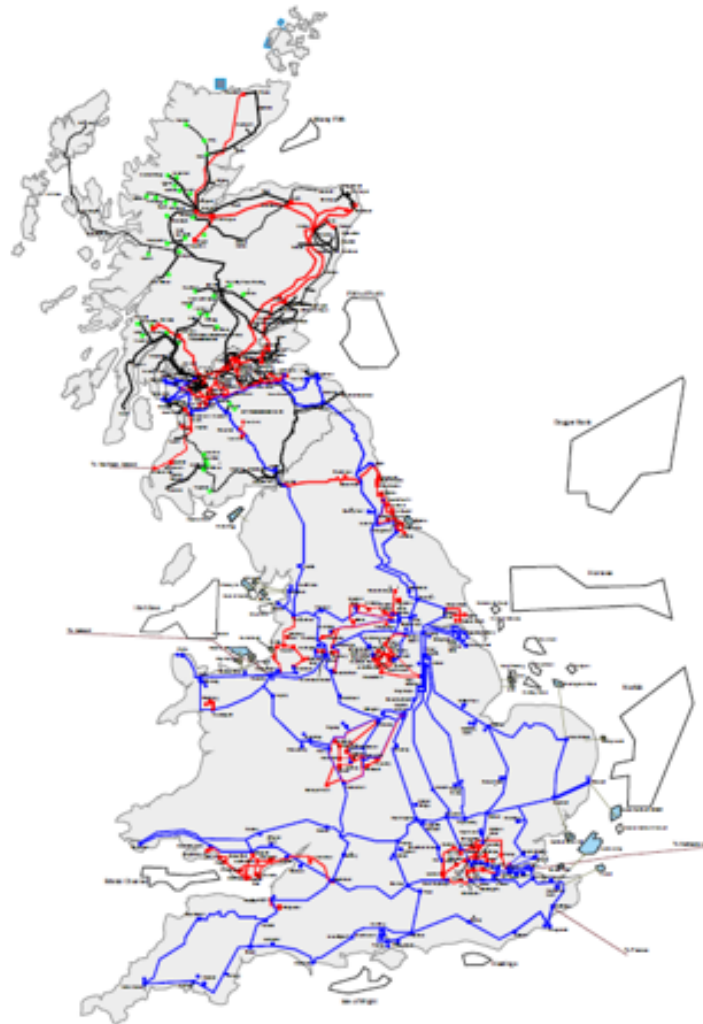


Figure 1: Diagram of the GB electricity transmission system.



Figure 2: Timeline of markets in the GB energy system.

[sum of the outputs of generators²] [nominal load for generator operation (500MW in Summer, 600MW in Winter)] [pumped storage demand] [export on interconnectors to other systems], and is the black line in the figure. The figure also illustrates the ‘raw’ demand when estimates of embedded wind and solar are added back on, though this raw demand is still net of embedded non-weather generation, on which information is insufficient to make outputs specific to any one historic point in time. This illustrates the extent to which embedded generation, which NG does not meter, can change both the level and daily pattern of the demand which NG must manage.

The next section will describe NG’s challenge of quantifying the value to consumers of NG gaining access to metered data from embedded generation.

2 Problem Description

2.1 The question

The challenge from National Grid Electricity System Operator (ESO) is stated as follows, the text here being a slightly expanded paraphrase of the slides presented at the start of the Study Group:

²Depending on the purpose, the set of generators considered may be slightly different. For system operation it would be those metered by NG, for settlement of ex post payments it would be all BM units. Details of this distinction are a second order issue for the analysis in this report.

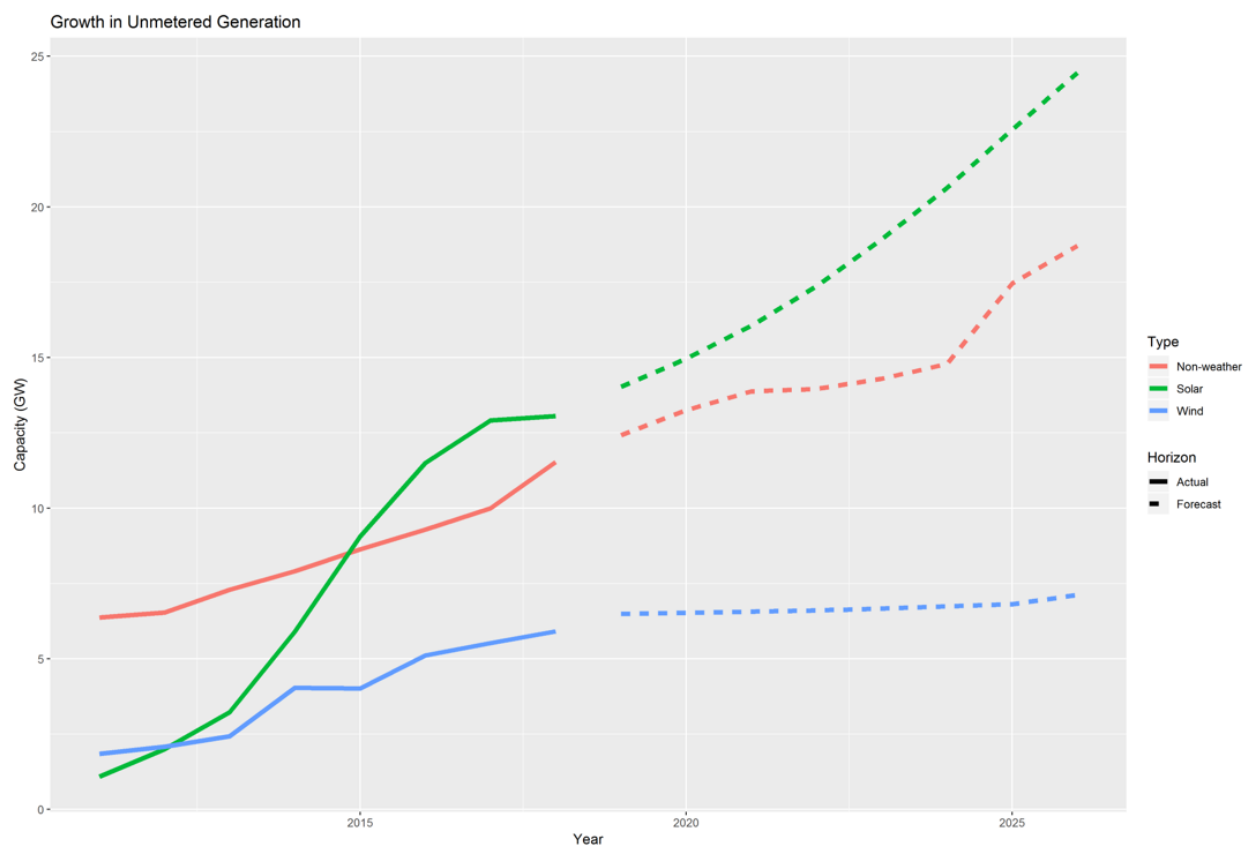


Figure 3: Growth of unmetered distributed generation in the GB system.

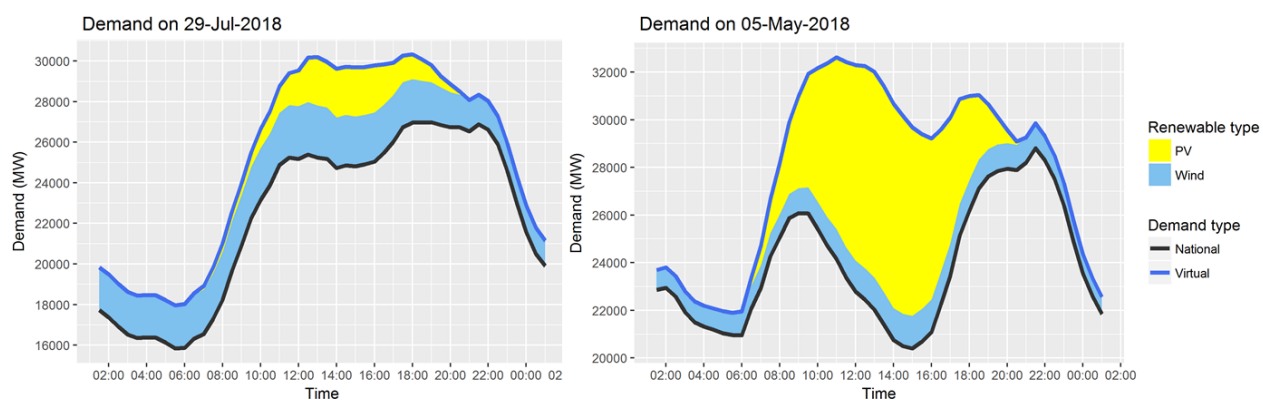


Figure 4: 'National' demand and estimate of raw demand on two sample days.

- Increasing the amount of distributed generation (DG), particularly renewables, has a clear benefit to society.
- Because of the lack of information flowing to the ESO about the DG, this increase in the amount of DG increases the uncertainty which the ESO must manage.
- To cover the risks of increased uncertainty, the ESO undertakes more balancing actions, which in turn increases the cost to the consumer.
- The key question is thus: ‘Can we build a model to quantify these costs, in order to predict future increases in these costs, and then use the model to inform government and the regulator about how to optimise the regulatory framework around distributed generation?’

The costs which NG faces in balancing the system are divided for recording purposes into two categories: costs incurred through redispatching generation due to network constraints; and other costs. This division is not precise, as it is common for a single BM action to serve more than one purpose. Non-constraint costs include

- Fast acting response, which covers against sudden losses of infeed from generators;
- Slower acting operating reserve, which is used to manage uncertainty in short term forecasting (of wind, solar and demand), and in mechanical availability of generators.

No simple formulae exist for the volume of reserve or response procured in given conditions, as there is a large element of professional judgement of control room engineers in these decisions. Due to the complexity of the system and the sensitive dependence of prices on generation and demand background, a bottom-up system modelling approach was deemed impractical and so a top-down data-driven approach was chosen. This aims to derive a statistical link between cost and the differences between forecasts and reality, and then consider how more information improves forecast accuracy and reduces costs.

The primary direct benefit of having better observations of distributed generation is better understanding of forecast uncertainty. For instance, a better understanding of forecasts might allow less reserve to be held, or if it is known why a particular network flow is lower than forecast (e.g. if wind output in Scotland is lower than expected) then NG can take advantage of this spare capacity with greater confidence.

2.2 The data

NG provided four years of historic data, split into 30-minute settlement periods yielding 48 settlement periods per day (except on the two clock-change days). For each settlement period the following data were available:

- The costs incurred by NG, split into network constraint costs and non-constraint costs, both as described above,
- The forecast and actual metered wind generation. We use metered generation to refer to generation that is either transmission-connected generation or distribution-connected generation above 100MW. The forecasts are two-day ahead. The actual metered wind generation is that which occurred in the settlement period in question.
- The forecast and estimated actual non-metered wind and solar generation. As described above, the non-metered, or embedded generation is distribution-connected generation of less than 100MW. As for metered generation, one-day ahead forecasts were available. The estimated actual non-metered generation is an estimate of that which occurred in the settlement period.
- Various other variables, e.g. energy prices, forecast and actual national energy demand etc.

The data were aggregated by taking the mean of the variables over the four settlement periods between 1700 and 1900 each day. This aggregation was done to smooth the measurements and hence improve the fit of any statistical models. The period between 1700 and 1900 was chosen as this is the period of peak demand each day in the UK and so we expect constraint costs to be higher during this period. One consequence of choosing this time period is that there is little solar generation at this time. For this report we have therefore focused on wind generation.

3 Modelling approaches

This section contains two approaches for estimating the reduction that could be achieved in constraint costs if the actual embedded wind generation was known. The first approach is a preliminary approach that uses only the data available during the data study group. The second approach explicitly considers how the flows down lines impact the constraint costs. This second approach requires additional data. Note that we focus on constraint costs in this section.

3.1 A preliminary estimate of the potential reduction in constraint costs

Based on discussions with NG it was clear that the largest constraint costs are associated with high wind in Scotland. Much of the wind generation in GB is in Scotland but most of the demand is in South-East England. When there is high wind generation this can put pressure on the transmission lines between Scotland and England and result in the need

to constrain Scottish wind generation to avoid overloading these transmission lines. Lack of knowledge of embedded generation can mean that costly constraint actions are taken unnecessarily because control room engineers anticipate that transmission line limits might be breached by an increase in embedded wind generation. If the actual embedded wind generation was known in each 30 minute period then forecast error in future periods would be reduced and so some of these costly constraint actions could be avoided.

In this section we attempt to estimate the effect of having real-time updates of non-metered wind generation. Our approach is to estimate the relationship between the constraint costs and the actual wind generation by estimating the function

$$\text{constraint costs} = f(\text{actual wind}). \quad (1)$$

In practice, the actual wind on the system is unknown as the amount of embedded wind generation is unknown. Control engineers will attempt to estimate the actual wind using a combination of the actual metered wind and the forecast of the non-metered wind (or probably some quantile of this forecast). Thus, to estimate the reduction in constraint costs as a result of better information we need to estimate the improvement to the argument of the function as a result of better information.

The function (1) can be estimated using metered wind data and the measured constraint costs. In an ideal world this function would be estimated using the estimate of the actual wind used by the control engineers. As all control engineers will determine this differently this was not practical with the available data.

We expect the function (1) to be piece-wise linear because costs are only incurred when the net generation that needs to pass down the transmission line exceeds the maximum line rating. As a result, the function will have approximate form:

$$f(w; l) = \max(w - l, 0), \quad (2)$$

where l is the threshold for the transmission lines between Scotland and England and w is the actual wind. The threshold l is uncertain and varies between settlement periods. It depends on the line rating, the demand in Scotland and (if using metered wind to estimate the function) the non-metered wind generation. Since l is unknown and variable, the function $f(w)$ (as a function of w alone) is convex and increasing. After excluding a small number of data points where constraint costs were zero over the entire day's 2-hour window, examination of the data suggested that in the region of interest, $\log f(w)$ was approximately linear. We therefore fit the following model:

$$\log f(w) = \alpha + \beta w + \epsilon,$$

with α and β constants and ϵ the unexplained variation. Least-squares regression on the log-costs yields the coefficients α and β , giving the form

$$\text{constraint costs} = \gamma \exp(\beta w + \epsilon) \quad (3)$$

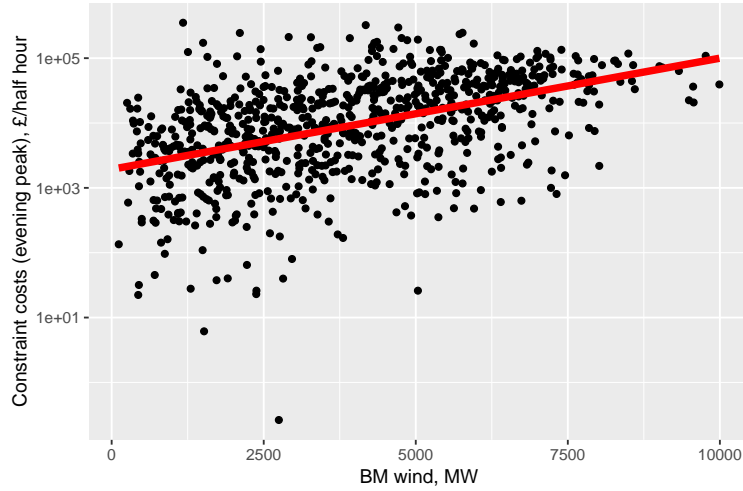


Figure 5: A plot of the constraint costs (on the log scale) against the actual wind measurements with the fitted line as given in Eq. (3).

where $\gamma = e^\alpha$. This approximates the form in Eq. (2), but is smooth to account for the variation in l . A plot of the constraint costs against the actual wind with the fitted function above overlaid is given in Fig. 5. For more details on the rationale of Eq. (2) see the next section.

The next step was to estimate the improvement in the argument of Eq. (1) if the real-time embedded generation could be measured. The opportunity for saving costs is when the forecast is larger than the actual generation as this is when unnecessary actions are taken. If we assume that control engineers currently use the forecast of embedded generation as the basis on which to take control actions, then the improvement in the argument of Eq. (1) is given by

$$\max(\text{forecast}, \text{actual}) - \text{actual}.$$

That is, if the forecast of the embedded wind is larger than the actual embedded wind the reduction in the argument is given by the difference between the forecast and the actual. If the forecast of the embedded wind is smaller than the actual embedded wind then there is no cost saving. Thus, using Taylor's theorem we can estimate the approximate reduction in cost by

$$\frac{df}{dw}(\text{forecast}) \times [\max(\text{forecast}, \text{actual}) - \text{actual}] \quad (4)$$

It seems unlikely that control engineers will use the forecast of the embedded generation to take decisions; in practice they may be more conservative to avoid damaging the transmission lines. Thus, a more realistic estimate of the reduction in costs might be achieved by replacing the forecast in Eq. (4) by the 90% quantile of the forecast.

To estimate the savings that might be possible we implement this approach using the data provided by National Grid. The derivative in (4) is

$$\frac{df}{dw}(w) = \beta\gamma \exp(\beta w). \quad (5)$$

The parameter estimates are $\beta = 0.0004$ and $\gamma = 1932$ (based on the actual metered wind generation measurements in GB). Equation (4) is then used to estimate the savings that would have occurred in the last year of the dataset (from 28 January 2018 to 27 January 2019 inclusive). For each day in this year the estimated reduction in cost is calculated by substituting the forecast of the embedded wind generation and the estimate of the actual embedded wind generation into (4). We use the estimate of the actual embedded wind as we do not know the true embedded wind.

The result of the calculation above is a saving of approximately £70,000. This is the total over a full year for an average half hour period between 1700 and 1900. Scaled up, this would be approximately £280,000 for a full year over the period from 1700 to 1900 (or 0.3% of the incurred constraint costs for those settlement periods). This saving is essentially negligible given other fluctuations. Operational expertise from NG ESO suggests that they expect to be able to make far greater savings than this given better system knowledge. Some possible reasons for this discrepancy include:

- The model may be missing an underlying more complex relationship.
- Forecasts of embedded wind generation are not the driving factor behind increasing costs to NG ESO.
- The averaging over the period from 1700 to 1900 might be smoothing out larger differences between the forecast of embedded wind and the estimated actual embedded wind.
- Control engineers might take conservative decisions that are based on summary statistics other than the mean forecast (e.g. including some estimate of the error in the forecast).

3.2 The value of improved information in decision making

We give a general probabilistic framework within which, given sufficient data, it is possible to evaluate more precisely the economic benefits of having improved information. In particular this is applicable to the reduction of constraint costs, where the improved information would consist of improved knowledge of line flows, e.g. the replacement of estimates by metered values of wind or solar generation. The reason why we have not been able to implement this more formal approach within the current project is mainly that there is not sufficient data

collected with which to estimate the probability distributions involved—in particular the distribution of loads on lines, and the conditional distributions of these loads given improved information. The collection of such data would seem to be necessary in order to make better progress. However, the framework below does indicate the necessary way of proceeding.

We consider a decision variable x —e.g. the level of power constraint to be obtained in the market—and we suppose that the cost of implementing the decision x is $c(x)$ for some function c . The decision variable x is to be optimally chosen. This may be done *either*:

1. **unconditionally**, i.e. with respect to minimal knowledge, represented by some (unconditional) probability model, with a probability measure \mathbf{P} , *or*,
2. **conditionally**, i.e. *with respect to further knowledge* (perhaps obtained by waiting for, or collecting, more information, such as better estimates of renewable generation); we assume this further knowledge to be represented by a knowledge of the value of some parameter μ , which within the earlier probability model would be a random quantity; we denote the corresponding conditional probability measure, given μ , by \mathbf{P}_μ .

In either case, we assume that there is some random variable Z_x (e.g. the power flowing across the constraint boundary), parametrised by the decision variable x (e.g. the level of constraint applied to that power flow), which is as far as possible to be kept out of some “risk” region A (e.g. the region of power flow which exceeds the safe limit).

We consider two possible approaches—which mirror National Grid’s possible approaches to other decision-making problems.

Minimisation of an overall economic cost. We assume that there is incurred some given, presumably *large*, cost K if the random variable Z_x falls within the risk region A (e.g. power exceeds the safe limit).

1. **Unconditional model.** Under this model x is to be chosen (unconditionally) so as to minimise the overall economic cost

$$g(x) := c(x) + K\mathbf{P}(Z_x \in A). \quad (6)$$

We denote the optimal solution (that minimising (6)) by \bar{x} . The corresponding minimised cost is then $g(\bar{x})$.

2. **Conditional model.** Here x is to be chosen as a function of the revealed additional information μ so as to minimise the (conditional) overall economic cost

$$g_\mu(x) := c(x) + K\mathbf{P}_\mu(Z_x \in A). \quad (7)$$

We denote the optimal solution by \bar{x}_μ . The corresponding conditional minimised cost is then $g_\mu(\bar{x}_\mu)$ and the unconditional minimised cost is $\mathbf{E}g_\mu(\bar{x}_\mu)$, where the expectation is taken over the distribution of the random quantity μ .

We now have

$$\begin{aligned}\mathbf{E}g_\mu(\bar{x}_\mu) &= \mathbf{E} \min_x g_\mu(x) \\ &\leq \min_x \mathbf{E}g_\mu(x) \\ &= \min_x g(x) \\ &= g(\bar{x}),\end{aligned}$$

so that the expected cost under the **conditional** approach (i.e. when further knowledge is obtained) is always less than the expected cost under the **unconditional** approach, and may be readily calculated.

Control of risk. Here the objective is to choose the decision variable x to control the *risk* of the event $\{Z_x \in A\}$ (e.g. line overload) so that it occurs with at most some very small, but usually nonzero, probability p . For simplicity of exposition, we assume that x is chosen so as to control the probability of the event $\{Z_x \in A\}$ to be exactly p .

1. **Unconditional model.** Here x is chosen so that

$$\mathbf{P}(Z_x \in A) = p. \quad (8)$$

We denote the solution to (8) (usually unique, but otherwise that minimising $c(x)$) by \hat{x} . The corresponding cost is then

$$c(\hat{x}) + Kp. \quad (9)$$

2. **Conditional model.** Here x is chosen as a function of the revealed additional information μ so that

$$\mathbf{P}_\mu(Z_x \in A) = p. \quad (10)$$

We denote the (unique or most economical) solution to (10) by \hat{x}_μ . The corresponding conditional cost is then

$$c(\hat{x}_\mu) + Kp \quad (11)$$

and the unconditional minimised cost is then $\mathbf{E}c(\hat{x}_\mu) + Kp$, where again the expectation is taken over the distribution of the random quantity μ .

The saving obtained by using the conditional model (i.e. making use of the available additional information) is therefore

$$c(\hat{x}) - \mathbf{E}c(\hat{x}_\mu). \quad (12)$$

The determination of the above quantity is now a matter of calculation, and it is not guaranteed that this quantity is always positive. However, this will generally be the case, as we

now show. Suppose that—as is the case for line overload and is likely to be the case more generally—the event $\{Z_x \in A\}$ is a tail event of the random variable Z_x which, by suitable definition of the latter, we may without loss of generality take to be the event $\{Z_x > 0\}$. Suppose further that the decision variable x corresponds to a locational shift for the distribution of Z_x —as will typically be the case in practice, e.g. where x is the level of power constraint on a line—and that μ is similarly a *location parameter* for this distribution—as may well be the case at least to a first approximation. Then

$$\mathbf{P}_\mu(Z_x > 0) = \bar{F}(x - \mu), \quad (13)$$

for some *tail distribution function* \bar{F} (where $\bar{F}(z) = 1 - F(z)$ for some distribution function F).

It now follows, using (8) and (13), that, within the **unconditional** model, the optimal decision \hat{x} is the solution of

$$p = \mathbf{P}(Z_{\hat{x}} > 0) = \mathbf{E}\mathbf{P}_\mu(Z_{\hat{x}} > 0) = \mathbf{E}\bar{F}(\hat{x} - \mu). \quad (14)$$

Within the **conditional** model, the optimal decision \hat{x}_μ , given μ , is the solution of $p = \mathbf{P}_\mu(Z_{\hat{x}_\mu} > 0) = \bar{F}(\hat{x}_\mu - \mu)$, and so also

$$p = \bar{F}(\mathbf{E}\hat{x}_\mu - \mathbf{E}\mu). \quad (15)$$

The tail function \bar{F} is almost certainly convex in the region of interest (it is typically close to exponential and is always at least eventually convex), so that, from Jensen's inequality, $\bar{F}(\hat{x} - \mathbf{E}\mu) \leq \mathbf{E}\bar{F}(\hat{x} - \mu)$, and hence, from (13) and (14) and using the fact that \bar{F} is decreasing, $\hat{x} \geq \mathbf{E}\hat{x}_\mu$. Finally, in the case where costs are linear in the decision variable x , so that $c(x) = cx$ for some $c > 0$ (since decisions which further reduce risk cost more), it follows that $c(\hat{x}) \geq \mathbf{E}c(\hat{x}_\mu)$. Hence, as might be expected, the use of the additional information provided by the conditional model reduces overall costs.

3.3 Preliminary exploration of non-constraint costs

In this section some brief exploratory ideas on the modelling of non-constraint costs are discussed. As described above, these non-constraint costs relate mainly to the costs of keeping reserve on the system to deal with unexpected events (e.g. a sudden drop in wind). The volume of reserve is set based on the judgement of control room engineers, underlying control room policies are reviewed periodically (at least once a year).

Figure 6 shows the variation of non-constraint costs through time in the available dataset. It is clear that there are several very large outliers. When non-constraint costs were compared with electricity prices there was a strong correlation between the two variables. This suggests that these outliers are a result of unusual events on the system and will not be reduced by improving knowledge of the embedded generation on the system. It further suggests benefit

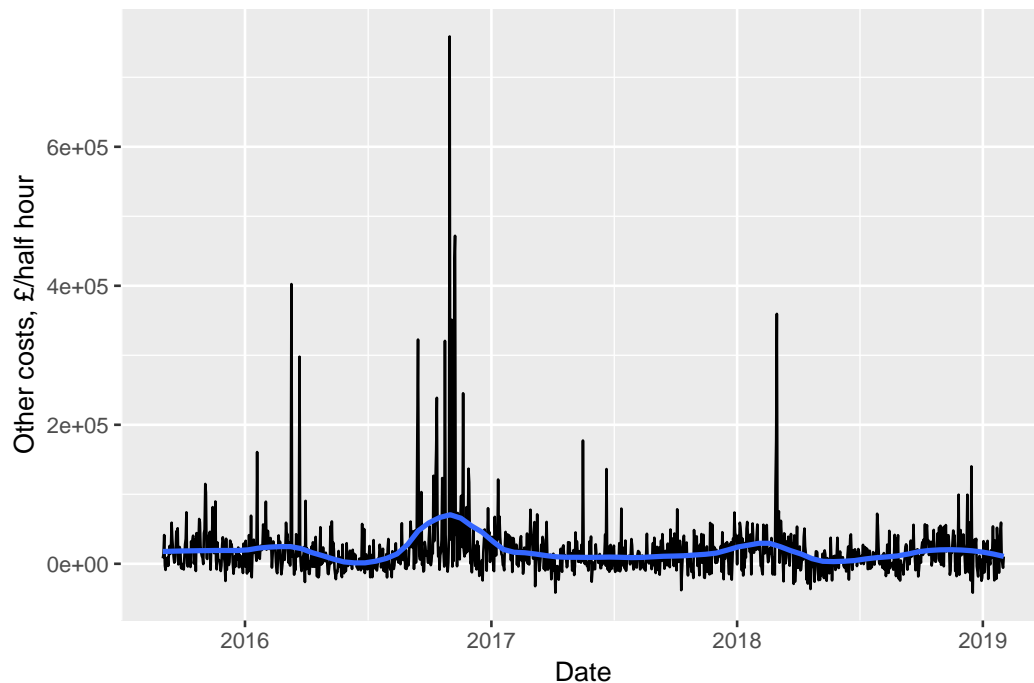


Figure 6: The variation in non-constraint costs through time, with LOESS smoothing line in blue.

from excluding these outliers from any analysis, particularly if they can be associated with known severe events. There is no evidence in the plot of any increase in non-constraint costs over time as the capacity of embedded generation increases.

Figure 7 presents a scatter plot of the non-constraint costs against the metered wind, with no substantial trend visible as the metered wind output increases. If there were any substantial dependence of non-constraint costs on embedded generation forecast errors, one would expect also some statistical dependence on outcomes (as errors are likely to be larger when the volume of wind generation is higher).

There is thus no evidence found *in the data* that better forecasting or knowledge of embedded generation can reduce non-constraint costs. However, issues with the information available mean that there could still be some benefit of better forecasting. Most importantly, better knowledge of how the control room decides on the volume of reserve to procure – this could provide very direct evidence of where improved information on embedded information could reduce costs under current practices, and could reveal opportunities to take additional factors associated with forecast error into account beyond current practices. On one specific point, the group understands that the reserve setting practices are updated approximately quarterly, and knowledge of what is updated and how might help with analysis of the historic data.

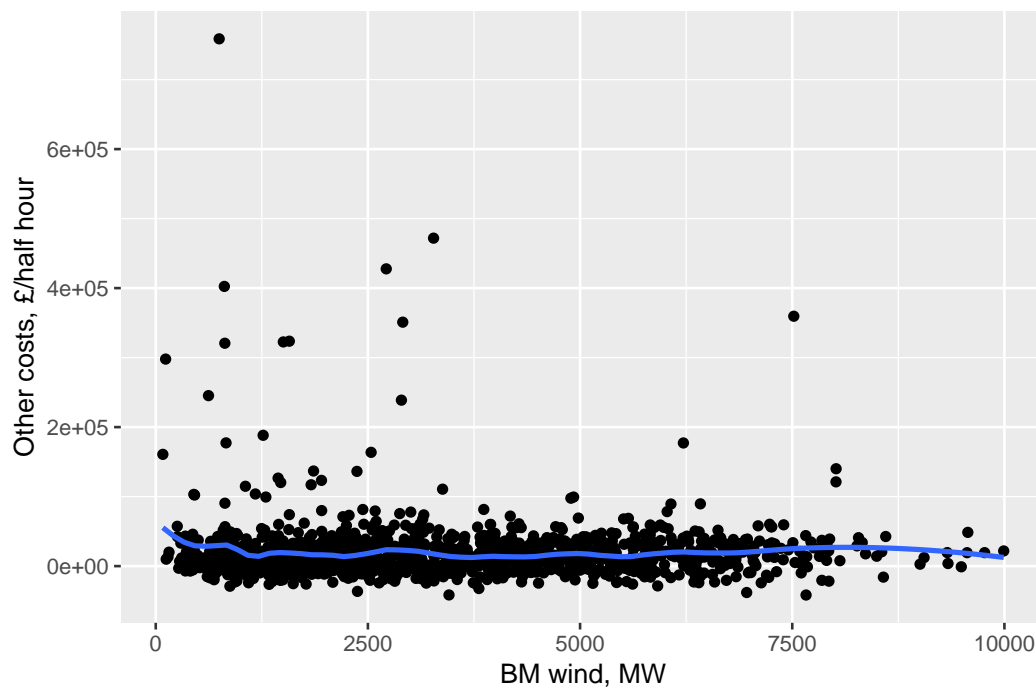


Figure 7: A scatter plot of the non-constraint costs against the metered wind generation, with LOESS smoothing line in blue.

4 Conclusions

The group has performed statistical analysis of the relationship between constraint costs and actual outcome wind generation, which suggests a positive but weak relationship between improved knowledge of embedded generation and reduction in costs. There has also been preliminary investigation of the relationship between non-constraint costs and improved information on embedded generation, which revealed no evidence of benefit of improved information.

It is possible, however, that with different data and better knowledge of control room practices, better assessments of the value of improved information could be made. The group also proposes the following as the basis of any further study of value of better forecasts and knowledge of distributed resources:

- Currently NG does not archive forecasts of boundary flows. If it were possible, at reasonable expense and for a limited period, to archive forecasts and outcomes of boundary flows on key boundaries (e.g. SCOTEX and SEIMP) this would provide a more direct means of assessing the potential value of improved forecast information.
- Access to demand data and forecasts by region (e.g. Scotland) would improve detail of analysis of constraint costs, as would additional information on when some specific underlying effect (e.g. a major maintenance or fault outage of a transmission line) pushes constraint costs into a different regime.
- Information regarding how the distribution network is managed, in particular regarding any Active Network Management (ANM) schemes that may be in place. ANM schemes control how much local generation goes into the distribution networks, in order to prevent those networks from failing. An understanding of how these schemes operate would reduce:
 - Forecasting error, as NG could predict the states of the distribution networks better.
 - Real-time error, as the actions of these schemes affects the output of the non-metered generation; therefore knowing about ANM actions when they occur could lead to a better estimate as to how much supply is actually being delivered to that network—e.g. if NG knew when an ANM had switched off most non-BM generation, it could take actions knowing that the supply to that network was just the BM generation being delivered to it.
- The choice of instructions in the balancing mechanism is (formally) a matter of judgement by control room engineers. Better knowledge of the logic behind these actions, and how this might change given additional information on distributed resources, would clearly improve analysis of the value of those resources. Most specifically, knowledge

of how operating reserve levels depend on the wind and solar forecasts would provide a direct means of assessing the reduction in reserve volumes due to better information.