Quantifying the potential for improved management of weather risk using subseasonal forecasting: the case of UK telecommunications infrastructure

Article

Published Version

Open Access

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH ARTICLE

Quantifying the potential for improved management of weather risk using sub-seasonal forecasting: The case of UK telecommunications infrastructure

David J. Brayshaw[1,2] (iD) | Alan Halford[1,3] | Stefan Smith[3,4] | Kjeld Jensen[5]

[1]Department of Meteorology, University of Reading, Reading, UK

[2]National Centre for Atmospheric Science, University of Reading, Reading, UK

[3]Technologies for Sustainable Built Environment Centre for Doctoral Training, University of Reading, Reading, UK

[4]School of the Built Environment, University of Reading, Reading, UK

[5]Applied Research, British Telecommunications plc, London, UK

Correspondence
David J. Brayshaw, Department of Meteorology, University of Reading, Reading, RG6 6BB, UK.
Email: d.j.brayshaw@reading.ac.uk

Funding information
BT (British Telecommunications plc); Engineering and Physical Sciences Research Council (EPSRC), Grant/Award Number: EP/G037787/1

Abstract

Reliable and affordable telecommunications are an integral part of service-based economies, but the nature of the associated physical infrastructure leads to considerable exposure to weather. With unique access to observational records of the UK fixed-line telecommunications infrastructure, an end-to-end demonstration of how extended range forecasts can be used to improve the management of weather risk is presented, assessing forecast value on both short-term "operational" (weeks) and long-term "planning" timeframes (months/years). A robust long-term weather-related fault-rate climatology is first constructed at weekly resolution, based on the ERA-Interim reanalysis. A clear dependence of winter fault rates on large-scale atmospheric circulation indices is demonstrated. The European Centre for Medium-Range Weather Forecasts (ECMWF) sub-seasonal forecast system is subsequently shown to produce skilful forecast of winter weekly fault rates at lead times of three to four weeks ahead (i.e. days 14–20 and 21–28). Forecast skill at a given lead time is, however, a necessary rather than a sufficient condition for improved risk management. It is shown that practical decision-making leads to dependencies across multiple forecasts times that cannot be modelled using traditional "cost-loss matrix" methods as errors in previous forecasts influence the value of subsequent forecasts. A parsimonious model representing operational decision-making for fault repair scheduling is therefore constructed to show that fault-rate forecast skill does improve both short- and long-term management outcomes (in this case meeting performance targets more often in the short term, or reducing the resources required to achieve these targets in the long term). Consequently, it is argued that new methods are needed for forecast skill assessment in complex decision environments.

KEYWORDS
extended range, forecast value, skill, sub-seasonal, telecommunications

1 | INTRODUCTION

Telecommunication networks are an integral part of secure and competitive societies where commerce and services depend on low-cost and reliable communications. In the United Kingdom, an estimated net economic contribution of £33 billion/year (or 1.5% of gross domestic product (GDP)) is attributable to telecommunications infrastructure (Kelly,

2015). However, as with other aspects of infrastructure, the exposed nature of the fixed-line telecommunications network leads to weather risk (e.g. transport and electricity; e.g. McColl *et al.*, 2012; Palin *et al.*, 2013). Indeed, Openreach—a division of BT plc responsible for almost 90% of the UK's fixed line telecommunications infrastructure—highlights weather as a contributor to increased line fault rates associated with service delays, disruptions and challenging conditions in each of its annual reports from 2013 to 2018 (BT, 2013, 2014, 2015, 2016, 2017, 2018). The quantification, prediction and management of weather-related line fault rates on the UK fixed-line network (hereafter, the "network") is therefore an important problem, with each aspect presenting distinct challenges.

First, the quantification of weather impacts on line fault rates is difficult due to the rapidly evolving nature of the infrastructure, with different line types having different exposures to weather. While the overall number of fixed lines for which Openreach is responsible has only increased slightly from 2012 to 2017, the mixture of line types (copper *versus* fibre, voice only *versus* voice-and-broadband) has changed dramatically. A large-scale network weather-hardening programme intended to improve its weather resilience also took place in the period 2009–2011. As a consequence, relevant line fault rate data are available only over a short period, presenting challenges for identifying weather-drivers of fault rates. This problem, which is also faced by other types of critical infrastructure, has recently been tackled through the creation of "synthetic" historical data sets derived from meteorological reanalyses (e.g. electricity; Ely *et al.*, 2013; Kubik *et al.*, 2013; Cannon *et al.*, 2015; Sharp *et al.*, 2015; Pfenninger and Staffell, 2016; Santos-Alamillos *et al.*, 2017; Troccoli *et al.*, 2018).

Second, the prediction of line faults days to months in advance is useful for day-to-day decisions required to ameliorate the impact of disruptive weather on network performance (e.g. overtime, transferring engineering resources between regions, temporarily contracting additional engineers). Surveys and semi-structured interviews conducted within BT reveal that these actions typically need to be taken several days or weeks in advance (Halford, 2018). The development of climate services, using skilful meteorological predictions weeks to months in advance (e.g. Lynch *et al.*, 2014; Scaife *et al.*, 2014; Clark *et al.*, 2017; Beerli *et al.*, 2017; Buontempo *et al.*, 2018; Troccoli *et al.*, 2018), may therefore enable improvements in the performance of a given system (i.e. the network and its associated set of maintenance resources) on a day-to-day "operational" basis. Moreover, these operational improvements may translate into better long-term "planning" decisions by changing the dimensions of the system itself (e.g. improvements in

operational decision-making could enable the same network to be managed with fewer resources without performance loss).

It is beyond the scope of the present paper to discuss sub-seasonal-to-seasonal (s2s) forecasting in detail. It is, however, noted that s2s systems are typically probabilistic, consisting of an ensemble of multiple realizations of possible future weather, and skill is derived from this ensemble rather than a single deterministic forecast (Richardson, 2000). Moreover, s2s systems are typically better at predicting the evolution of large-scale atmospheric patterns (about 100–1,000 s of km) rather than localized surface properties. For simplicity, the North Atlantic Oscillation (NAO) is used here to indicate the dominant pattern of large-scale winter atmospheric circulation over Western Europe (positive NAO states are associated with warm, wet, windy winters in the UK and *vice versa* for negative; Hurrell *et al.*, 2003).

Finally, a forecast is useful insofar as it provides value for a particular end-user. For a forecast to provide value, it must relate to a quantity of concern to a particular user and enable better decision-making compared with alternative strategies (Murphy, 1985; Richardson, 2000). In the present context, line fault rate forecasts are not the primary concern; instead, it is ability to meet performance targets for timely fault *repair* that matter most directly to the end user.

With unique access to observational records of the UK telecommunications infrastructure, the present paper addresses all three aspects described above, providing an end-to-end demonstration of how meteorological information can be used in an important practical setting, addressing forecast value over both day-to-day "operational" and longer term "planning" timeframes. It does not seek to produce the best possible forecast, instead using simple techniques to highlight the processes involved and demonstrate the potential value. The end-user application, which shares many similarities with applications in energy-systems operations and planning, represents a shift in the way weather and climate forecasts and simulations are used and evaluated: from a situation where forecast skill is assessed primarily in terms of meteorology (upon which user decisions are subsequently taken) towards one where decisions are explicitly included in the assessment process.

The paper begins with a description of the data and methodologies, and the creation of a long-term synthetic reconstruction of line fault rates (Sections 2 and 3). The resulting data set is used to quantify the extent to which large-scale winter atmospheric circulation patterns influence line fault rates and demonstrates that predictive skill is achievable from the current generation of numerical weather prediction systems up to four weeks ahead (Section 4). Finally, a

parsimonious model of the decision-making process is used to identify the conditions under which the fault rate forecast can provide user value (Section 5). Section 6 provides concluding discussion.

2 | ESTABLISHING A LONG "FAULT RATE CLIMATOLOGY"

2.1 | Observed faults

Observed fault data from the telecommunications network is provided by Openreach from April 2011 to December 2017. A fault is defined as an unintended interruption in service on Openreach's network that Openreach is required to repair. Faults are reported to Openreach from their customers *via* their communications service provider and thus there is often a delay of hours, days or even weeks between the fault occurring and it being recorded by Openreach. In recognition of this unknown and variable delay, weekly rather than daily fault rates are used.

Faults arise from a range of causes. Where possible, non-weather-related faults associated with "early life" issues (within the first 28 days from installation) are removed from the data set, along with faults within exchange buildings and customer premises (which are not part of Openreach's responsibility).

Observed fault rate data are available at several spatial scales and across four different line types (Halford, 2018). These line types are known as VOICE, VOICE_BB, VOICE_NGA and MPF, employing different network technologies and different combinations of communications products. Here, fault rates are aggregated nationally, with three distinct line types considered (VOICE, VOICE_BB and MPF), broadly corresponding to a set of predominantly copper-based lines capable of carrying voice-only (VOICE) and voice-plus-broadband (VOICE_BB and MPF) services. Collectively the three types correspond to over 80% of the installed lines in 2016, and large numbers of each line type are present across the whole record. The remaining line type (VOICE_NGA corresponding to newer "Next Generation Access" lines) is excluded because only a small number of lines of this type were present before about 2014/15. Each line type is modelled individually before aggregation across the line types because, although the three line types have similar average weekly fault rates *per* installed line (Halford, 2018), each has a distinct set of technological characteristics and quantitatively different weather responses.

Neither actual fault rates nor numbers of lines installed are presented (for commercial sensitivity); instead, fault rates are normalized with respect to a reference period (discussed below).

2.2 | Meteorological data

Raw 6 hr surface meteorological data from ERA-Interim (Dee, 2011; about 80 km resolution, 1979–2017) are extracted for land-points only in the domain 12° W–4° E, 48° N–61° N. This is aggregated over area and time to weekly resolution for use in the fault rate regression model (Section 3).

An NAO index is defined as the first empirical orthogonal function (EOF) of weekly mean sea level pressure (MSLP) over the North Atlantic domain (20–80° N, 80° W–40° E). The resulting spatial pattern resembles other similar NAO calculations, typically performed on monthly mean MSLP (e.g. Hurrell *et al.*, 2003; Zubiate *et al.*, 2017), and explains a similar amount of the variance (36%).

Archived meteorological hindcasts are taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) extended range forecast system (Vitart *et al.*, 2008), *via* the S2S database (Vitart and Robertson, 2018), corresponding to twice-weekly (Monday and Thursday) forecast launch dates over December 2016 to February 2017. These are chosen as they correspond to a single, recent, version of the operational ECMWF forecast model (Cy43r1). Each launch date produces an 11-member hindcast ensemble corresponding to the same calendar date occurring in each of the previous 20 years (i.e. the December 1, 2016, forecast produces hindcasts for December 1, 1996, 1997, …, 2015).

For each hindcast ensemble member, a weekly NAO index is calculated by projecting the NAO spatial pattern derived from reanalysis onto the weekly mean MSLP pattern derived from the forecast model. To account for model bias and drift, a lead time-dependent bias correction is applied to the NAO values calculated from the hindcast using a "leave-one-year-out" method. In effect, the bias of each individual ensemble member's NAO hindcast for a particular year [$x \in X$] is estimated by comparing the ensemble-mean time-mean NAO averaged over all the remaining years (i.e. $X \setminus \{x\}$) and comparing this with the observed time-mean NAO from ERA-Interim (averaged over the same years, $X \setminus \{x\}$). This correction is applied to each ensemble member for the year x. This correction is applied separately for each lead time, for each launch date and for each year (following Lynch *et al.*, 2014). As this correction is applied for each launch date separately, it removes the impact of any sub-seasonal drift in the climatological-mean NAO (i.e. the forecast cannot produce skill by predicting the climatological evolution of the NAO across the winter season; e.g. Keeley *et al.*, 2009).

Forecast "week 1" is defined as days 0–6, "week 2" as days 7–13, and so forth. Owing to the fixed launch dates and the focus on forecasting the same period (December–February) there are fewer hindcast launch dates available for

longer lead times (e.g. for the first week of December, the first hindcast launch date December 1 provides a week 1 prediction but the corresponding week 2, 3, …, 6 hindcasts are not available as earlier launch dates are excluded). Reduced sample sizes at longer lead times lead to wider confidence ranges in the forecast skill assessment but do not affect the conclusions of the present paper.

Several standard forecast skill metrics are applied to both NAO index forecasts and fault-rate forecasts: anomaly correlation co-efficient (ACC), rank-probability skill score (RPS), continuous rank probability skill score (CRPS); root mean square error (RMSE) and mean absolute error (MAE). Definitions can be found in standard statistical textbooks (e.g. Wilks, 2011). In most cases, these metrics are presented as skill scores (i.e. normalized with respect to a reference forecast such as climatological expectation) such that the numerical values are dimensionless.

3 | FAULT-RATE MODEL

Following similar approaches in weather-related electricity demand modelling (Taylor and Buizza, 2003; Bloomfield et al., 2016), a multiple linear regression model is constructed for each line type to characterize the link between the observed fault rates and selected trial weather variables during the observational period (listed in Table 1). The trial weather variables are based on BT's experience of fault repair and prevention, recorded through interviews and expert elicitation (Halford, 2018). The model fitting procedure is described, using the VOICE line type as an example.

For commercial sensitivity reasons, the observed fault rates for VOICE lines are first divided by their average fault
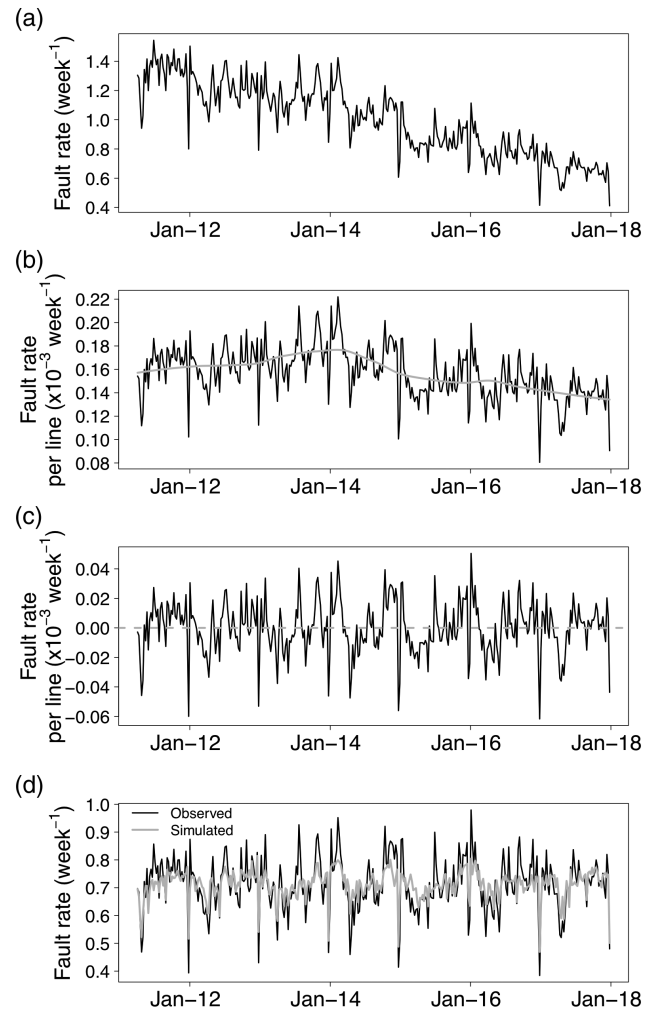
TABLE 1   Fault-rate model variables

Variable	Description
PS	Weekly total precipitation (three week running mean)
PT	Weekly total precipitation over threshold (binary value; threshold 105 mm·week^{-1})
T	Weekly mean temperature
W	Square of weekly mean 10 m wind speed
WT	Weekly mean wind speed over threshold (binary value; threshold 15 m·s^{-1})
RHT	Flag to indicate if three or more consecutive days occur with relative humidity over a threshold (binary value; threshold 85%)
HOL	Public holidays in a week
$\varepsilon(0, \sigma)$	Residual/"noise" term; normal distribution

Note: For each weather variable where a smoothing window or threshold is used, a range of windows/thresholds was tested before the final selection presented here.

FIGURE 1   Fault-rate model for VOICE lines: (a) observed fault rate (divided by the average fault rate over the entire period); (b) observed fault rate *per* installed line (black line), with "background trend" (smooth grey line)—the 2017 reference point is identified; (c) observed fault-rate anomalies (with the "background trend" removed); and (d) comparison between observed and simulated expected fault rates over the observational period

rate over the observational period (hereafter referred to as $\overline{FR}_{\mathrm{obs}}^{\mathrm{VOICE}}$). The resulting data are presented in Figure 1a, and are subsequently normalized by the number of VOICE lines installed ($NLINES^{\mathrm{VOICE}}$) to account for the installation and removal of lines (Figure 1b).

The resulting fault rate record, however, still contains year-to-year trends associated with non-weather-related effects, such as network degradations, preventative maintenance, changes in processes and working practices, and changes in customer expectations. A smooth locally estimated scatterplot smoothing (LOESS) curve is fitted to the fault data (window of 156 weeks; Cleveland, 1979), to identify these inter-annual trends in fault numbers. This trend line is referred to as the "background fault rate" (BFR^{VOICE}).

Fault-rate anomalies with respect to the trend line are then calculated (Figure 1c). The value of the long-term trend at the last week in 2017, $BFR^{\text{VOICE}}|_{2017}$, is also noted for later use as a reference point in the network's physical configuration (i.e. the y-axis value of the right-hand most point of the smooth grey line in Figure 1b).

A stepwise regression (Burnham, 2004) is performed on the fault-rate anomalies, linking them to selected weather parameters. This tests all possible combinations of the trial weather parameters, seeking the minimal combination with the best model fit (measured by minimizing the Akaike's information criterion score; Akaike, 1974). The resulting model takes the form (see Table 1 for definitions; the unit is normalized fault rate *per* line *per* week, i.e. line^{-1} week^{-1}):

$$EFRA^{\text{VOICE}} = \alpha_0 + \alpha_1 PS + \alpha_2 PT + \alpha_3 T + \alpha_4 W + \alpha_5 WT + \alpha_6 RHT + \alpha_7 HOL, \tag{1}$$

where $EFRA^{\text{VOICE}}$ is the expected fault-rate anomaly on VOICE lines; and $\alpha_{1,\dots,7}$ are the fitted regression co-efficients (Table 2). The residuals (i.e. model minus observed fault-rate anomalies) are near normal with weak autocorrelation and therefore a normally distributed random number, $\varepsilon_i(0, \sigma)$, is optionally added to produce an individual realization, i, of the fault-rate anomaly FRA_i:

$$FRA_i^{\text{VOICE}} = \alpha_0 + \alpha_1 PS + \alpha_2 PT + \alpha_3 T + \alpha_4 W + \alpha_5 WT + \alpha_6 RHT + \alpha_7 HOL + \varepsilon_i(0, \sigma) \tag{2}$$

such that in the limit of a very large set of individual realizations:

TABLE 2 Parameters in the fault-rate models (see the text for a discussion)

Co-efficient (×10⁶)	VOICE	VOICE_BB	MPF
α_0 (line^{-1} week^{-1})	−36.30	−37.30	−20.60
α_1 (line^{-1} week^{-1}/ mm·week^{-1})	1.37	1.29	0.75
α_2 (line^{-1} week^{-1}/ mm·day^{-1})	n.s.	n.s.	n.s.
α_3 (line^{-1} week^{-1}/°C)	0.90	0.92	0.34
α_4 (line^{-1} week^{-1}/m^2·s^{-2})	0.08	0.10	n.s.
α_5 (line^{-1} week^{-1}/m·s^{-1})	n.s.	n.s.	1.07
α_6 (line^{-1} week^{-1}/%)	0.93	1.06	1.17
α_7 (line^{-1} week^{-1}·days^{-1})	−17.50	−11.60	−11.90
σ (line^{-1} week^{-1})	10.60	8.56	8.03

Note: Entries marked "n.s." indicate the corresponding variable was not selected by the stepwise regression. Units of each co-efficient are marked in parentheses for convenience.

$$EFRA^{\text{VOICE}} = \frac{1}{N} \sum_{i=1}^{N \to \infty} FRA_i^{\text{VOICE}}. \tag{3}$$

Once the regression co-efficients have been calculated, the fault-rate anomalies (FRA_i^{VOICE} and $EFRA^{VOICE}$) are converted back to normalized fault rates for the line type assuming a steady-state network equivalent to late 2017, that is, using the number of installed lines and background fault rate from the end of 2017. The resulting model for VOICE lines performs well, with an $R^2 = 0.64$ and RMSE $= 0.079$ week^{-1}, with a clear correspondence between the simulated expected fault-rate and the observed fault rate (Figure 1d; note that the observed fault rate has been similarly adjusted to match a network corresponding to late 2017, but both the simulated and observed fault rates remain normalized by the long-term average fault rate, $\bar{FR}_{obs}^{VOICE}$, for commercial sensitivity reasons). A similar process follows for the other two line types (MPF and VOICE_BB), with $R^2 = 0.59$ and 0.64 and RMSE $= 0.11$ and 0.069 week^{-1} respectively (for model co-efficients; Table 2).

To evaluate the fault rates on the network as a whole, that is, the total faults across the three line types, the normalized fault-rate anomalies for each line type are converted back to actual fault rates for the line type before summing over the line type. For example, a single realization of the simulated total fault rate across the three line types is TFR_i:

$$TFR_i = \sum_{linetypes} FRA_i^{linetype} \cdot BFR^{linetype}\big|_{2017} \cdot NLINES^{linetype}\big|_{2017} \cdot \bar{FR}_{obs}^{linetype} \tag{4}$$

with an equivalent expression for the simulated expected total fault rate $TEFR$:

$$TEFR = \sum_{linetypes} EFRA^{linetype} \cdot BFR^{linetype}\big|_{2017} \cdot NLINES^{linetype}\big|_{2017} \cdot \bar{FR}_{obs}^{linetype}. \tag{5}$$

Finally, the values of TFR_i and $TEFR$ are normalized by averaging over a baseline period in the simulated expected fault rates (for commercial sensitivity reasons). The reference chosen is arbitrary but corresponds to the long-term average over the entire *observational* record (April 2011–December 2017) once calibrated to a network state of 2017. In other words, as in the fault-rate models, the observations for each line type is divided by $NLINES^{linetype}$ and the long-term trend $BFR^{linetype}$ subtracted, then the 2017 "network state" is applied (add $BFR^{linetype}|_{2017}$ and multiply by $NLINES^{linetype}|_{2017}$), before summing over the three line types.

In summary, *TEFR* is considered to represent the weekly fault rate that would have been expected to occur due to the weather conditions in some historic period *if the network at the end of 2017 had existed at that point in time*. Similarly, by using multiple realisations of TFR_i, a probability distribution of fault rates can be constructed under the same network assumptions. Here, 1,000 realizations of TFR_i, are used to estimate each probability distribution.

Figure 2 demonstrates the resulting models. Figure 2a indicates a good match between the modelled and observed fault rates, though the model has a slight tendency to overestimate low fault numbers and underestimate high fault numbers. Overall, the national model aggregated over the three line types has an $R^2 = 0.67$ and RMSE = 0.074 week^{-1}. The spread around the 1:1 line in Figure 2a is well captured by the "residual" $\varepsilon_i(0, \sigma)$ term: Figure 2b shows that the observed fault rate values typically lie within the 5–95% confidence interval of the stochastic model.

The resulting model has many applications in terms of establishing a "climatological" perspective (i.e. 1979–2017) on weather-related faults. Figure 3a provides a simple indication of this: the expected range of fault rates across the annual cycle. The distribution of fault rates encompasses not only the effects of the residual in the fault-rate model but also an estimate of the weather uncertainty (i.e. the fault rates that could be experienced given a different weather year). A clear annual cycle is visible, with fault rates peaking in late autumn into winter (about 0.9–1.2 week^{-1}) compared with lower values in late spring and early summer (about 0.7–1.1 week^{-1}), though the latter is somewhat influenced by public holidays.

Having demonstrated the model's overall performance, it is now possible to focus on the impact of weather on fault rates. Some terms—in particular public holidays, $\alpha_7 HOL$, and the residual, $\varepsilon_i(O, \sigma)$—have no meteorological significance such that a simplified version of the model can be written (with these terms neglected):

$$EFRA^{linetype} = \alpha_0 + \alpha_1 PS + \alpha_2 PT + \alpha_3 T + \alpha_4 W + \alpha_5 WT + \alpha_6 RHT.$$

(6)

Figure 3b shows that the annual cycle in this simplified model still has a clear fault rate peak in winter. In the remainder of the present paper, the resulting simplified modelled total expected fault rate (*TEFR*), after normalization by the long-term mean, is referred to simply as the "fault rate" and symbolically represented as *FR*.

4 | THE NAO AND WINTER FAULT RATES

The strong seasonal cycle of fault rates means that winter is an important period for network and resource management. A key driver of UK and European winter weather is the NAO, with recent studies suggesting potential predictability weeks to months ahead.

Two components are required for the construction of a weather forecast-based fault-rate prediction system. First, there must be a strong relationship between a weather predictand and the fault rate; and second, the chosen weather
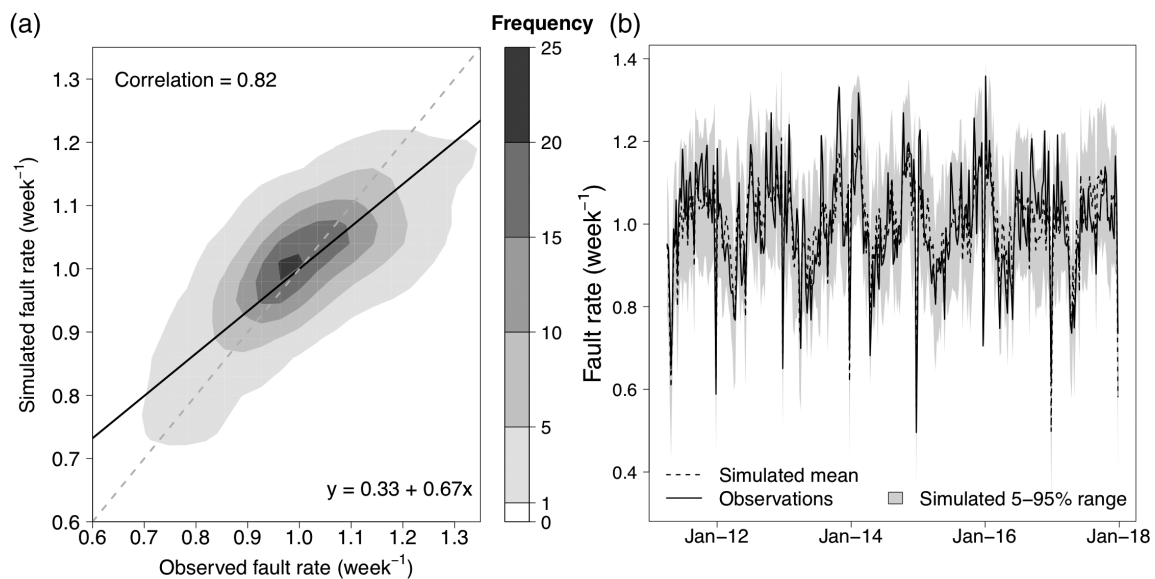

FIGURE 2 Comparison between the fault-rate simulation and observed fault data illustrated as (a) frequency density and (b) time series. In each case, the observed fault data are adjusted to correspond to a constant 2017 network state (see the text for a description). In (a), the best linear fit is shown by the black curve, and a 1:1 line is provided (grey dashed)
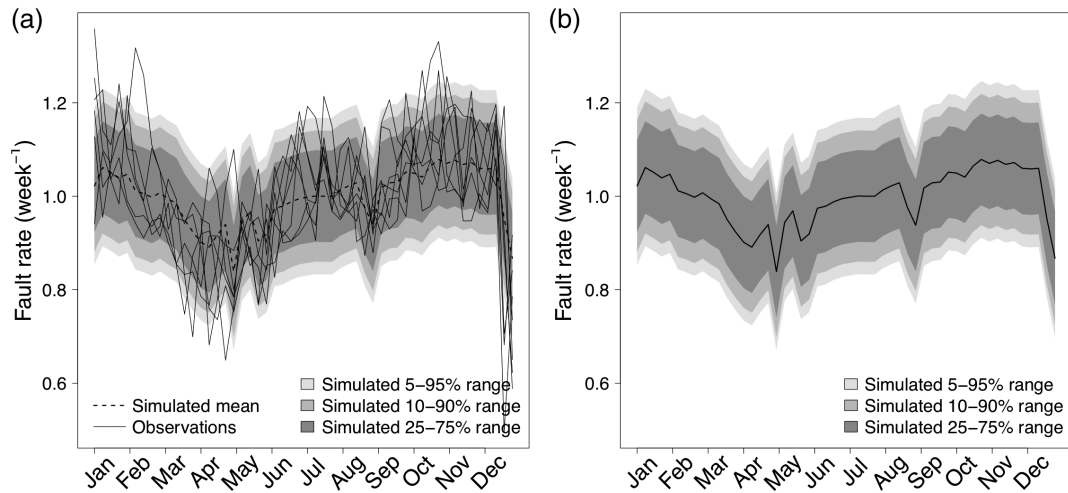
FIGURE 3 Annual cycle of fault rates: (a) the "full model" (Equation (4); includes the bank holiday term); and (b) the "no bank holiday model" (with α_7 in Equation (4) set to zero). Each panel shows the simulated mean (central line) and expected range of fault rates (shading) that could be expected under different weather years. The available whole years of observed fault data (2012–2017) are shown in (a) as thin solid lines

predictand must be skilfully predicted by the weather forecast system. These two components are discussed in the following sections separately, before the performance of a complete "fault rate-prediction" system is presented. For simplicity, the only weather predictand considered is the NAO, but other weather predictands such as direct measures of pressure gradients may lead to higher overall levels of fault rate forecast skill (Zubiate *et al.*, 2017; Thornton *et al.*, 2019).

4.1 | Historic relationships between the NAO and winter faults

Figure 4 shows the correlation of the NAO with simulated historic fault rates during winter. Positive NAO values generally lead to higher than normal fault rates, though there is also considerable scatter (Figure 4a). This is consistent with the well-known impact of the NAO on Northern European

winter climate, whereby positive NAO states lead to warm, wet and windy conditions.

The continuous NAO index is divided into three roughly equal bins (in terms of frequency of occurrence) in Figure 4b. A positive NAO state is associated with higher than normal fault rates, and *vice versa* for NAO negative, though there is some overlap between the distributions. Therefore, there are many different possible UK/European weather states (and therefore fault rates) associated with the same value of the NAO state, highlighting the importance of viewing the relationship between the NAO, European weather and fault rates as probabilistic rather than deterministic. Nevertheless, the strong relationship between NAO and fault rates suggests that a skilful meteorological forecast of the NAO could potentially provide valuable information.
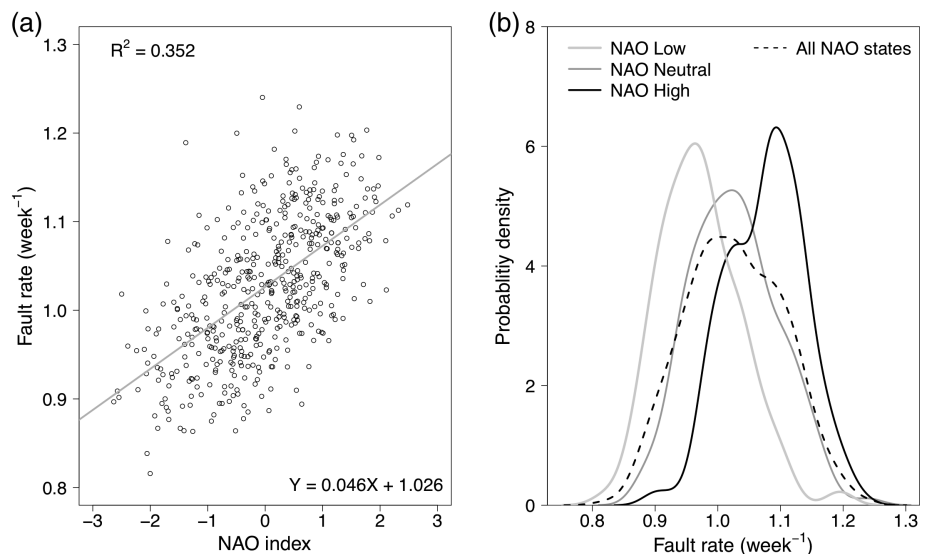

FIGURE 4 Impact of the North Atlantic Oscillation (NAO) state on winter fault rates: (a) NAO index *versus* fault rate; and (b) fault-rate distributions contingent on NAO "states", defined as: "high" (NAO > 0.45), "low" (NAO < −0.45) and otherwise NAO "neutral". These thresholds correspond to three divisions with each containing approximately one-third of the time series

4.2 | NAO forecast skill

Figure 5 shows three metrics of NAO-forecast skill, expressed as a score relative to a climatological forecast (i.e. the mean NAO value observed for that week, averaged over the whole ERA-Interim record 1979–2017). A score of unity represents a perfect forecast, positive values indicate the model is outperforming climatology, and negative values worse than climatology. As expected, the skill reduces with increasing lead time and the following:

- ACC: a measure of whether the ensemble-mean NAO can forecast the *sign* of the NAO suggests good skill in weeks 1–3 (about 1.0 dropping to about 0.5), with a small measure of positive skill across even out to week 6 (about 0.3).
- RPS: a measure of the ability to forecast ranked categorical "states" of the NAO (i.e. positive/neutral/negative) that decreases rapidly from about 0.7 in the first week to about 0.1 in week 3. In week 4 and beyond, skill cannot be statistically detected.
- Similar to the RPS, the CRPS: a measure of the skill using the full probability distribution for the NAO that decreases to modest values in week 3 (about 0.15), and is not statistically significant in week 4 onwards.

In summary, the ECMWF extended range forecast system contains demonstrable forecast skill. The weekly mean
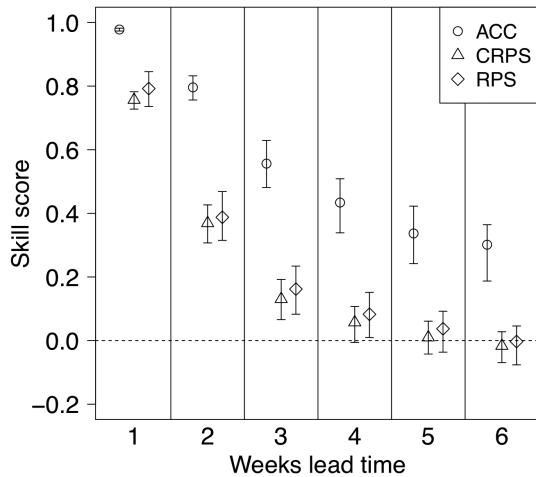
NAO probability distribution produced by its ensemble members has skill in week 3 (days 14–20) while the ensemble mean NAO forecast has skill even in week 6 (days 35–41).

This is not an exhaustive analysis of forecast skill: these results only illuminate the performance of a single forecast model over a relatively short hindcast period. It is, however, expected that the NAO skill results presented are a lower bound on the achievable forecast skill in the operational forecast model for three reasons: first, the operational forecast contains more ensemble members than the hindcast (51 compared with 11); second, it is known that the RPS is low-biased for small ensembles (Weigel *et al.*, 2007); and third, newer versions of the forecast now in operational service may offer improvements over the version analysed here.

4.3 | Fault-rate forecasts

Given the meteorological forecast skill for the NAO, two strategies are tested to convert forecasts of NAO state information into fault rate estimates (deterministic and probabilistic).

4.3.1 | Deterministic fault-rate forecasts

A deterministic forecast is a single estimate of fault rate. In this case, a fault rate forecast for calendar week w at a forecast lead time l is given by:

$$FR_{det}^{w,l} = FR_{clim}^w + \frac{1}{M}\sum_{j=1}^{M} F\bar{R}A\left(NAO_j^{w,l}\right), \qquad (7)$$

where FR_{clim}^w is the climatological-mean fault rate; $NAO_j^{w,l}$ is the NAO state forecast (positive/neutral/negative) from ensemble member j; $F\bar{R}A(\cdot)$ is the mean fault-rate anomaly associated with a given NAO state; and M is the ensemble size. The fault rate predicted is the climatological value for the relevant week of the year, plus the mean fault-rate anomaly associated with the forecast NAO state, weighted by the number of ensemble members predicting the occurrence of each NAO state. This is hereafter referred to as the deterministic "operational" fault rate forecast.

The performance of this forecast method is compared against two benchmarks:

- Deterministic "climatological" fault rate forecast where Equation (7) reduces to $FR_{det,clim}^{w,l} = FR_{clim}^w$ for all lead times, l.
- Deterministic "perfect NAO" fault rate forecast where Equation (7) reduces to $FR_{det,perf}^{w,l} = FR_{clim}^w + F\bar{R}A\left(NAO_{observed}^{w,l}\right)$ for all lead times, l.

FIGURE 5 Skill scores for North Atlantic Oscillation (NAO) anomaly forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) sub-seasonal forecast: (a) anomaly correlation co-efficient (ACC); (b) CRPS; and (c) RPS. Skill scores are referenced with respect to a time-evolving climatological NAO-anomaly forecast (i.e. the climatological NAO anomaly observed in each week of winter). A skill of unity represents a perfect forecast; zero/negative values represents no additional skill or a degradation in performance compared with the reference climatology forecast. Error bars indicate the 90% confidence band (i.e. 5–95% range)

The ECMWF forecasts are launched on Mondays and Thursdays, but are assessed on their performance to simulate weekly mean fault rates. To ensure that the two sets of forecasts are compared at consistent lead times, the fault-rate climatology (described in Section 3) is recalculated for each forecast set using week start dates consistent with the launch days (i.e. the Monday-launched forecasts are compared against a weekly fault rate climatology with weeks starting on Mondays, and similarly for Thursdays).

4.3.2 | Probabilistic fault-rate forecasts

A probabilistic fault rate forecast is constructed in a similar manner to the deterministic forecast described above, but provides a probability distribution of fault rates rather than a single value. Instead of using the *mean* fault-rate anomaly associated with each NAO state, i.e. $F\bar{R}A(\cdot)$, the corresponding empirical probability distribution corresponding to the NAO state is used, here denoted $F\tilde{R}A(\cdot)$. Thus a probabilistic fault rate forecast for calendar week w at a forecast lead time l is given by:

$$FR_{prob}^{w,l} = FR_{clim}^w + \frac{1}{M}\sum_{j=1}^{M} F\tilde{R}A\left(NAO_j^{w,l}\right). \quad (8)$$

This is hereafter referred to as the probabilistic "operational" fault rate forecast.

The performance can then be evaluated against two benchmarks:

- Probabilistic "climatological" fault rate forecast. Equation (8) reduces to $FR_{prob,clim}^{w,l} = FR_{clim}^w + F\tilde{R}A_{allNAO}$ for all lead times, l, where $F\tilde{R}A_{allNAO}$ is the fault-rate anomaly distribution across all NAO states.
- Probabilistic "perfect NAO" fault rate forecast where Equation (8) reduces to $FR_{prob,perf}^{w,l} = FR_{clim}^w + F\tilde{R}A\left(NAO_{observed}^{w,l}\right)$ for all lead times, l.

4.3.3 | Fault-rate skill

Figure 6 shows the relative skill of each method. In each case, the MAE (deterministic) or CRPS (probabilistic) is expressed as a dimensionless skill score with respect to a deterministic climatological forecast (see Section 2.2). Positive scores therefore imply an improvement on deterministic climatological information.

The perfect NAO forecasts provide an upper bound of the skill that can be derived from NAO information alone. The skill scores of about 0.15 (deterministic) and about 0.4 (probabilistic) represent statistically significant improvements on a purely climatological fault rate forecast, thus it is clear that a
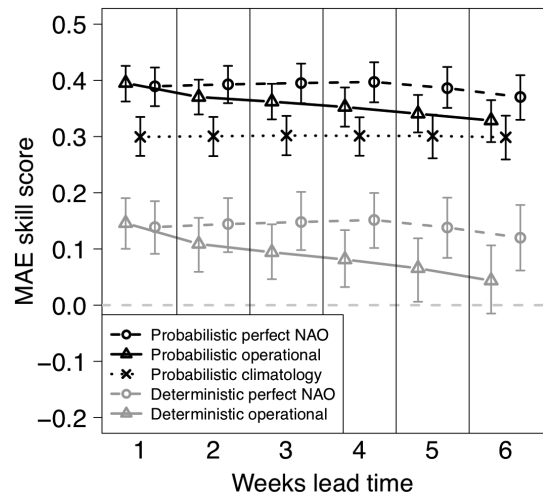

FIGURE 6 Skill scores for fault rate forecasts (see the text for a description). Skill scores are referenced with respect to a time-evolving deterministic climatological fault rate forecast (i.e. the climatological expected fault rate observed in each week of winter). Probabilistic skill scores are measured using CRPS, whereas deterministic skill scores use mean absolute error (MAE). A skill of unity represents a perfect forecast; zero/negative values represents no additional skill or a degradation in performance compared with the reference climatology forecast. Error bars indicate the 90% confidence band (i.e. 5–95% range)

skilful forecast of the NAO provides skill in terms of fault rate. The operational forecasts (i.e. using ECMWF forecasts of the NAO) confirm this, with comparable skill to the perfect NAO fault rate forecast in week 1, with skill dropping at longer lead times (at week 5 skill scores of around about 0.07 (deterministic) and about 0.35 (probabilistic) represent very marginal improvements over their respective climatological equivalents; beyond this any skill improvement over climatology is not detectable).

While this overall pattern of skill decay with lead time is consistent with the skill in forecasting the NAO-forecast itself (Figure 5, Section 4.2), the "operational" fault rate forecasts' skill decays with lead time more slowly than the NAO RPS/CRPS skill scores might suggest (there is still some skill over climatology in weeks 4–6 whereas the NAO RPS score suggests no skill beyond week 3). This is initially surprising, given that the fault rate forecast is based on predicting the NAO states and therefore might be expected to resemble the NAO's RPS skill. However, the fault rate forecast methods (Sections 4.3.1 and 4.3.2) are more closely related to the forecast skill of the ensemble mean NAO (and hence ACC) rather than its distribution (and hence RPS). This can be seen by considering the deterministic fault rate method, where the predicted fault-rate anomaly is the sum of the individual NAO state anomalies predicted by each ensemble member (i.e. effectively a weighted ensemble mean with the weights corresponding to the strength of the fault anomalies in each NAO state). The skill of the fault-

rate forecast therefore relies on the skill in the ensemble mean prediction of the NAO, not in the NAO probability distribution.

Though it is difficult to compare quantitatively the performance of probabilistic and deterministic forecasts the skill measures used (MAE for deterministic and CRPS for probabilistic) are comparable insofar as CRPS reduces to MAE in the deterministic limit. On this basis, the probabilistic fault rate forecast outperforms the deterministic fault rate forecast at all lead times.

5 | THE VALUE OF FAULT-RATE FORECASTING

A major internal performance indicator for Openreach is timely fault repair. Depending on the service level agreement between Openreach and its customers (i.e. communication service providers), reported faults typically must be fixed within one to three working days. The fraction of successful fault clears within these deadlines is a key operational measure, which we will refer to here as RD3, following the naming used internally within Openreach. Openreach has RD3 targets imposed by the industry regulator, known as OfCom, and penalties are incurred if these targets are not met (OfCom, 2014).

Although in practice there are many subtleties associated with, for example, timing, geography, technology and the skillsets of particular engineers, repair can be viewed as drawing upon a pool of engineering resource (a set of trained engineers) which can be allocated to repairs. The fault rate is therefore an ingredient in the work stack that must be completed by the engineering resource in each period, but improvements in fault-rate prediction do not guarantee improved performance against an RD3 target. Moreover, the performance against the RD3 target depends upon operational decisions (themselves based on work inflows and forecasts), and thus it is not possible to write a function linking weather inputs to an estimate of RD3 (in particular, unfixed faults are carried over as a workstack thus RD3 at any instant in time, t_0, is potentially influenced by not only the instantaneous fault rate but also the fault rate at all previous timesteps $\{t_{-N},...,t_{-2}, t_{-1}\}$). To understand whether fault rate forecasts can help to meet RD3 targets, it is necessary to simulate the decision-making process.

This decision can be approximated as follows. Under normal conditions, a minimum level of engineering resources (engineers) are retained: corresponding to a set of appointment slots that can be allocated to repairing lines. In situations where the normal level of engineering resource is unable to meet the RD3 target, a series of actions can be performed (at a cost) to temporarily increase resource levels by a few

tens of a per cent (e.g. overtime, delaying non-essential actions, and issuing short-term contracts for additional external resource). These actions, however, take time to implement and so decisions on resource levels are typically taken based on *forecast* fault rates and locked in days to weeks ahead.

In the following section, a parsimonious model is presented to mimic this decision process, enabling the time evolution of the engineering resource (and subsequent failures to meet a defined RD3-like target) to be simulated for a given fault rate forecast.

5.1 | Decision model

The real scheduling problem involved is complex with actions taken across multiple lead times and balancing resource between multiple objectives (Halford, 2018). For simplicity, scheduling decisions are reduced to a two-step framework corresponding to forecast weeks 1 and 2:

- In analogy to the RD3 target, a fraction of each week's incoming faults is specified as a target repair threshold. These must be repaired within the week, otherwise a "target failure", α, is recorded for each excess unrepaired fault. A threshold of 70% is used, broadly consistent with recent OfCom RD3 targets (sensitivity tests confirm qualitatively insensitivity to moderate variations, not shown).
- Unrepaired faults carry over to the next week. This work stack must be repaired before any new faults are repaired.
- Engineering resources, r, are scheduled for week 2 based on the work stack and incoming fault rate, and may take any value within the range $[r^{\min}, r^{\max}]$. Additional engineering resource incurs a finite but small cost.
- The number of foreseeable target failures is always minimized.
- Consistent with current practice, only deterministic fault rate forecasts are considered.

The problem is implemented as an iterative programme on a weekly time step. The target failure rate, α, and stack, s, at the end of week k are:

$$
\begin{aligned}
\alpha_k &= s_{k-1} + FR_k^{actual} \cdot (1-\lambda) \\
&\quad - r_k s_k = s_{k-1} + FR_k^{actual} - r_k,
\end{aligned}
\tag{9}
$$

where FR_k^{actual} is the observed fault rate in week k, and $(1-\lambda)$ is the target threshold parameter (set to 0.7, or 70%); and r_k corresponds to the engineering resource allocation, which must be calculated 1 week in advance (i.e. the value is calculated for week $k+1$, based only on information available at the start of week k). Mathematically, calculation of r for each week is a linear programme:

$$\min\left(\alpha_{k+1}^{forecast} \cdot c_t + r_{k+1} \cdot c_r\right)$$

where:

$$\alpha_{k+1}^{forecast} = \max\left(s_k + FR_{k+1}^{forecast} \cdot (1-\lambda) - r_{k+1}, 0\right)$$

subject to:

$$r^{\min} \leq r_{k+1} \leq r^{\max}, \tag{10}$$

where c_t and c_r are the unit costs of failing to meet a repair target and increasing the engineering resource respectively such that $c_t \gg c_r$ (provided $c_t > c_r$ is satisfied the solution to each individual optimization problem is insensitive to the values of c_t and c_r). $FR_{k+1}^{forecast}$ and r_{k+1} are the forecast fault rate and the resulting engineering resource allocated for week $k+1$, respectively. For simplicity, it is assumed that very high-quality short-range forecast of the fault rates in the first week (i.e. week k) is available such that $FR_k^{forecast} = FR_k$, and hence s_k and α_k are perfectly known (see Equation 9). However, in the second week (i.e. week $k+1$), the fault rate forecast is assumed to be imperfect, such that $\alpha_{k+1}^{forecast}$ corresponds to a *forecast* failure rate which may be different to the *actual* failure rate, α_{k+1}. This assumption therefore seeks to highlight the role of forecast information in the second week and is consistent with the observation that short-term forecasts (less than 1 week ahead) tend to be much more skilful than longer range forecasts (greater than 1 week ahead).

The fault rate is presented in normalized dimensionless units (i.e. a fraction of the long-term fault rate). Consequently, r and s are in similar normalized non-dimensional units: a value of $r = 1$ week^{-1}, for example, indicates engineering resources sufficient to meet the long-term average fault rate.

Only the period for which week 2 sub-seasonal hindcasts are available are considered (the second week of December 1996 through to the last week of February 2016). The model is run continuously over all winter weeks across all years (e.g. the last week in February 2010 is followed immediately by the second week in December 2010). The model is initialized with $s_1 = 0$ week^{-1} and $r_1 = r^{\min}$. The first and last years are removed as spin up and incomplete data respectively, leaving 18 years of 14-week winters data for each simulation.

A sensitivity test of a similar two-timestep decision model based on the week 4 fault forecast was performed (e.g. for the week starting January 22, the week 2 forecast for $FR_{k+1}^{forecast}$ which was launched on January 15 is replaced with a week 4 forecast launched on January 1). The results

were consistent with the week 2 discussion presented here, though with weaker skill levels (not shown).

5.2 | Failure rates

The decision model includes two parameters, r^{min} and r^{max}, corresponding to the lower and upper bounds on the engineering resource, r, with the difference between them ($r^{max} - r^{min}$) representing the schedulable contingency. These two parameters can therefore be viewed as controllable through appropriate long-term business planning, whereas the carryover from week to week (λ) is imposed externally by the regulator, OfCom.

The concern is to identify the extent to which fault-rate forecasts translate to:

- Reductions in expected target failure rates given different prevailing business conditions (i.e. different r^{min} and r^{max}).
- Potential for reducing resource levels (i.e. r^{min}) without increasing the long-term expected target failure rate.

Figure 7 addresses the first question, showing the impact of fault-forecasts on the target failure rate under differing r^{max}. The value of $r_{ref}^{min} \approx 1.04$ week^{-1} in each simulation is
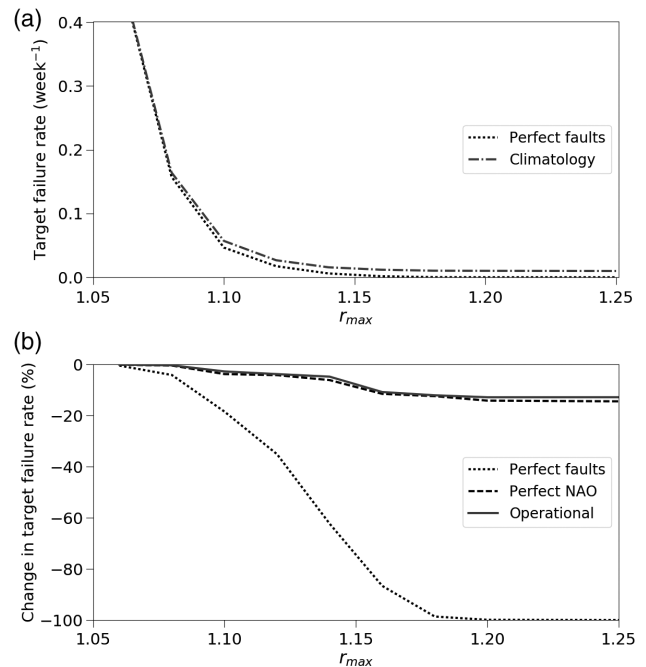

FIGURE 7 Impact of r_{max} (the upper bound on "engineering resources") on fault repair target failure rates using different fault-rate prediction schemes, for a given r_{min} (about 1.04 week^{-1}), as simulated by the simplified decision-model outlined in the text: (a) absolute failure rates; (b) change (reduction) in the target failure rate for each scheme with respect to the climatological forecast

set to the average weekly fault rate in the winter period, broadly consistent with current practice (Halford, 2018). Three fault rate forecasts (climatology, perfect NAO and operational week 2) are shown along with a true perfect forecast (where the actual fault rate is used in place of a forecast fault rate). For all forecast methods, small r^{max} values lead to high target failure rates (Figure 7a, about 0.3 week^{-1}). This drops to low values for $r^{max} \gtrsim 1.1 - 1.2$ week^{-1} indicating that most target failures can be avoided with a modest flexibility in the overall resource (about 15% points, corresponding to roughly 2 standard deviations (SD) in the wintertime weekly fault rate). This suggests that overall character of target failure rates is dominated by the flexibility to add engineers when required to deal with fault spikes. As flexibility reduces, target failures necessarily increase.

Figure 7b shows the value of different forecast schemes. For $r^{max} \lesssim 1.1$ week^{-1}, the system has insufficient flexibility in the upper bound of engineering resource to respond to improvements in fault rate forecasts: despite knowing a fault rate spike will occur, little additional engineering resource can be obtained. For $1.1 \lesssim r^{max} \lesssim 1.2$ week^{-1}, fault rate forecasts begin to demonstrate advantages over climatology and for $r^{max} \gtrsim 1.2$ week^{-1}, the benefits of fault rate forecasts saturate: sufficient contingency engineering resource can be obtained such that any backlog of unrepaired faults can be rapidly cleared. In the large r^{max} limit, the perfect forecast leads to an almost 100% target failure rate reduction, whereas the "operational" week 2 forecast leads to an about 12% reduction (similar to the perfect NAO forecast; an operational week 4 forecast offers a about 5% reduction, not shown). Operational forecast methods offer a relatively consistent fraction of the total improvement offered by the perfect fault rate forecast across the whole range of r^{max} (i.e. about 10% for week 2 and about 5% week 4, respectively, week 4 not shown).

Figure 8 addresses the second question: the extent to which a fault rate forecast can reduce r^{min} while maintaining the same expected target failure rate. Based on the analysis above and stakeholder interviews (Halford, 2018), a contingency level of 15% is assumed ($r^{max} - r^{min} = 0.15$ week^{-1}) and the target failure rates evaluated for each forecast system under a range of r^{min}.

Figure 8a shows that, as r^{min} decreases, the target failure rate increases, as expected (there is less engineering capacity to fix faults so more target failures occur). Figure 8b also shows that the benefit of all forecasts (including the perfect forecast) decreases for low r^{min}, reinforcing the observation that forecast value is limited by the decision maker's ability to respond. As before, however, the perfect NAO and operational week 2 forecasts offer some reduction in target failure rates compared with a climatological forecast (typically 5–10% reduction).
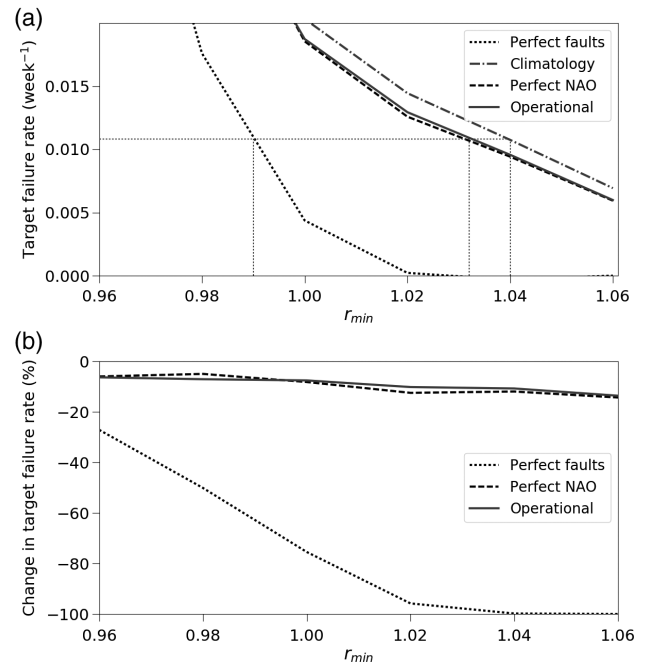

FIGURE 8 Impact of r_{min} (the lower bound on "engineering resources") on fault repair target failure rates using different fault-rate prediction schemes, for a given contingency ($r_{max} - r_{min} = 0.15$ week^{-1}), as simulated by the simplified decision-model outlined in the text: (a) absolute target failure rates; and (b) change (reduction) in target failure rate for each scheme with respect to the climatological forecast. The dotted lines in (a) are discussed in the text

Figure 8a further shows that forecasts enable a reduction in r^{min} while maintaining the same level of target failure rate risk, and the same level of access to contingency resources. For example, a system with r^{min} about 1.04 week^{-1} using a climatological forecast has the same expected target failure rate as a system with r^{min} about 1.03 week^{-1} using an operational week 2 forecast: using the operational week 2 forecast therefore enables a about 1% reduction in permanently held engineering resources. For comparison, a perfect forecast could achieve a 5% reduction in permanently held engineering resources for the same expected target failure rate.

In summary, the usefulness of fault rate forecasts depends upon the ability of decision-makers to respond. Nevertheless, in this simple illustrative decision model, fault rate forecast skill can either reduce the frequency with which repair target failures occur (by about 5–10% using the present week 2 forecast system) or reduce the amount of permanently held engineering resources (by about 1%).

6 | DISCUSSION AND CONCLUSIONS

A fault rate forecast system for the Openreach telecommunications network in the UK is demonstrated and a

simple decision model used to explore the consequent operational and planning benefits. Four aspects are highlighted:

- *Quantifying baseline weather and climate risk*: a 38 year historic reconstruction of fault rates is created using meteorological reanalysis. This enables more robust characterization of the "climatology" of risk compared with observational records.
- *Identification of climate drivers of fault rates*: winter is a particularly challenging period for faults, with higher fault rates than summer. The winter North Atlantic Oscillation (NAO) is shown to affect faults, with positive NAO typically leading to higher rates compared with NAO negative conditions.
- *Skill is found in ECMWF sub-seasonal NAO forecasts and can translate into skilful fault rate forecasts weeks in advance*: statistically significant skill is found for the NAO at lead times of several weeks (6 weeks for ACC; about 3–4 weeks for RPS/CRPS) and consequently in NAO-tercile based fault rate forecasts.
- *Skilful fault rate forecasts could enable better fault management in Openreach's UK telecommunications network*: improved fault rates forecasts could improve the rate at which performance targets are met (i.e. reduce the number of faults that cannot be repaired within a given time window) or reduce the cost of maintaining the system at a given level of risk.

From a business perspective, the simplified decision model suggests that sub-seasonal numerical weather predictions offer considerable potential advantages. While the reductions in failure rates and engineering resources appear modest (about 5% and 1%, respectively, for a week 2 forecast), they correspond to significant potential savings. For context, the maximum cost of repair target failures is about £1 million·day^{-1}, while an engineering workforce with on the order of 10,000 staff implies annual staffing costs of around £500 million: a 1% saving in engineering resource could therefore correspond to a potential saving of about £1 million. These estimates are, of course, upper bounds insofar as there are competing demands for engineering resources and limitations on the decision-maker's ability to respond. Nevertheless, with more advanced methods, particularly making more use of probabilistic forecasts and decision-making or inclusion of other meteorological ingredients (e.g. the East Atlantic or Scandinavian patterns; Zubiate *et al.*, 2017; Halford, 2018), it is likely possible to achieve reductions towards about 5% achievable with a perfect forecast system.

Consistent with the experiences of other recent climate service activities (e.g. Buontempo *et al.*, 2018; Troccoli

et al., 2018), this research highlights the importance of understanding the decision-making context in evaluating weather forecast performance. In this example, it is possible to recognize two timeframes, operations and planning, both of which can be informed by the forecast system.

On the short operational timeframe, the user objective can be viewed as making the best use of predetermined set of resources. In the present context, the resources correspond to the permanently held engineering resources (r^{min}) and the contingency available ($r^{max} - r^{min}$). Here, one seeks to minimize the amount of contingency resource used ($r_t - r^{min}$) and the failure rate for each time step t such that $r^{min} < r_t < r^{max}$. In this manner, the forecast's value should be expressed in a reduction in both failure rates and operating costs ($\sum_t (r_t - r^{min})$).

On the long planning timeframe, the user objective is different: it can be viewed as determining the optimal set of resources to achieve a desired balance between cost and risk (here, the expected target failure rate). That is, a forecast is skilful if it enables reductions in permanently held engineering resources (r^{min}) or the maximum contingency available ($r^{max} - r^{min}$) while maintaining the same expected performance against given targets.

Neither of these forecast value assessments can, however, be made without explicitly including a representation of the decision process. Though there are examples of decision models in weather/climate forecasting studies (e.g. Sonka *et al.*, 1987; Kim and Palmer, 1997), and though climate service projects have engaged strongly with users in a wide range of sectors and applications (e.g. Buontempo *et al.*, 2018), most assessment of s2s forecasts remains focused in the meteorological domain (e.g. evaluating forecast skill for meteorological properties or simple transformations of them, rather than the value added given a particular decision-making process). The use of an explicit decision model in the present paper contrasts with standard meteorological impact assessments where impact I is viewed as a transfer function from meteorological conditions $f(M)$ and "forecast value" is determined from an $N \times N$ "cost/loss" decision matrix model (Murphy, 1985; Richardson, 2000). That approach is insufficient for situations where no transfer function f exists ($\nexists f : I = f(M)$; Brayshaw, 2018) and decision outcomes and forecast errors compound over time. The inclusion of decision-making represents a different paradigm in forecast assessment. The impact model used to convert weather variables to decision outcomes becomes integral to the forecast assessment (rather than being added afterwards), and may potentially use forecasts across many different lead times. This suggests the need not only for continuing in-depth sectoral engagement in climate service development (as in, for example, EUPORIAS and its many successors;

Buontempo *et al*., 2018) but also for new generalizable tools and methods to represent better complex decision-making in weather and climate forecast assessment.

ORCID

David J. Brayshaw https://orcid.org/0000-0002-3927-4362

REFERENCES

Akaike, H. (1974) A new look at statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Beerli, R., Wernli, H. and Grams, C.M. (2017) Does the lower stratosphere provide predictability for month-ahead wind electricity generation in Europe? *Quarterly Journal of the Royal Meteorological Society*, 143, 3025–3036.

Bloomfield, H.C., Brayshaw, D.J., Shaffrey, L.C., Coker, P.J. and Thornton, H.E. (2016) Quantifying the increasing sensitivity of power systems to climate variability. *Environmental Research Letters*, 11, 124025.

Brayshaw, D.J. (2018) The nature of weather and climate impacts in the energy sector. In: Troccoli, A. (Ed.) *Weather & Climate Services for the Energy Industry*. Basingstoke: Palgrave Pivot, pp. 151–160.

BT. (2013). *BT Group plc Annual Report and Form 20-F*. London: BT Group plc. Available at: https://www.btplc.com/Sharesandperformance/Financialreportingandnews/Annualreportandreview/pdf/2013_BT_Annual_Report_smart.pdf [Accessed 31st May 2019].

BT. (2014). *BT Group plc Annual Report and Form 20-F*. London: BT Group plc. Available at: https://www.btplc.com/Sharesandperformance/Financialreportingandnews/Annualreportandreview/pdf/2014_BT_Annual_Report_smart.pdf [Accessed 31st May 2019].

BT. (2015). *BT Group plc Annual Report and Form 20-F*. London: BT Group plc. Available at: https://www.btplc.com/Sharesandperformance/Financialreportingandnews/Annualreportandreview/pdf/2015_BT_Annual_Report.pdf [Accessed 31st May 2019].

BT. (2016). *BT Group plc Annual Report and Form 20-F*. London: BT Group plc. Available at: https://www.btplc.com/Sharesandperformance/Financialreportingandnews/Annualreportandreview/pdf/2016-Annual-Report.pdf [Accessed 31st May 2019].

BT. (2017). *BT Group plc Annual Report and Form 20-F*. London: BT Group plc. Available at: https://www.btplc.com/Sharesandperformance/Financialreportingandnews/Annualreportandreview/pdf/2017_BT_Annual_Report.pdf [Accessed 31st May 2019].

BT. (2018). *BT Group plc Annual Report and Form 20-F*. London: BT Group plc. Available at: https://www.btplc.com/Sharesandperformance/Financialreportingandnews/Annualreportandreview/pdf/2018_BT_Annual_Report.pdf [cited 31st May 2019].

Buontempo, C., Halon, H.M., Soares, M.B., Christel, I., Soubeyroux, J.-M., Viel, C., Calmanti, S., Bosi, L., Falloon, P., Palin, E.J., Vanvyve, E., Torralba, V., Gonzalez-Reviriego, N., Doblas-Reyes, F., Pope, E.C.D., Newton, P. and Liggins, F. (2018) What have we learnt from EUPORIAS climate service prototypes? *Climate Services*, 9, 21–32.

Burnham, K.P. (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33, 261–304.

Cannon, D., Brayshaw, D.J., Methven, J., Coker, P.J. and Lenaghan, D. (2015) Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. *Renewable Energy*, 75, 767–778.

Clark, R.T., Bett, P.E., Thornton, H.E. and Scaife, A.A. (2017) Skilful seasonal predictions for the European energy industry. *Environmental Research Letters*, 12, 024002.

Cleveland, W.S. (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74, 829–836.

Dee, D.P. Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bor mann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Ely, C.R., Brayshaw, D.J., Methven, J., Cox, J. and Pearce, O. (2013) Implications of the North Atlantic oscillation for a UK–Norway renewable power system. *Energy Policy*, 6, 1420–1427.

Halford, A. (2018). *Forecasting weather impacts on the United Kingdom telecommunication network*. PhD Thesis, University of Reading, p. 233.

Hurrell, J.W., Kushnir, Y., Ottersen, G. and Visbeck, M. (Eds.) (2003) An overview of the North Atlantic oscillation. In: *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*. *Geophysical Monograph*, Vol. 134. Washington, DC: American Geophysical Union, pp. 1–35.

Keeley, S.P.E., Sutton, R.T. and Shaffrey, L.C. (2009) Does the North Atlantic oscillation show unusual persistence on intraseasonal timescales? *Geophysical Research Letters*, 36, 1–5.

Kelly, S. (2015) Estimating economic loss from cascading infrastructure failure: a perspective on modelling interdependency. *Infrastructure Complexity*, 2, 7.

Kim, Y.-O. and Palmer, R.N. (1997) Value of seasonal flow forecasts in Bayesian stochastic programming. *Journal of Water Resources Planning and Management*, 123, 327–335.

Kubik, M.L., Brayshaw, D.J., Coker, P.J. and Barlow, J.F. (2013) Exploring the role of reanalysis data in simulating regional wind generation variability over Northern Ireland. *Renewable Energy*, 57, 558–561.

Lynch, K.J., Brayshaw, D.J. and Charlton-Perez, A. (2014) Verification of European subseasonal wind speed forecasts. *Monthly Weather Review*, 142(8), 2978–2990.

McColl, L., Palin, E.J., Thornton, H.E., Sexton, D.M.H., Betts, R. and Mylne, K. (2012) Assessing the potential impact of climate change on the UK's electricity network. *Climatic Change*, 115, 821–835.

Murphy, A. (1985) Decision making and the value of forecasts in a generalized model of the cost-loss ratio situation. *Monthly Weather Review*, 113, 362–369.

OfCom. (2014) *Fixed Access Market Reviews: Wholesale Local Access, Wholesale Fixed Analogue Exchange Lines, ISDN2 and ISDN30—Consultation on the Proposed Markets, Market Power Determinations and Remedies*, p. 409.

Palin, E.J., Thornton, H.E., Mathison, C.T., McCarthy, R.E., Clark, R.T. and Dora, J. (2013) Future projections of temperature-related climate change impacts on the railway network of Great Britain. *Climatic Change*, 120, 71–93.

Pfenninger, S. and Staffell, I. (2016) Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*, 114(1), 1251–1265.

Richardson, D.S. (2000) Skill and relative economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 126, 649–667.

Santos-Alamillos, F.J., Brayshaw, D., Methven, J., Thomaidis, N.S., Ruiz-Arias, J.A. and Pozo-Vazquez, D. (2017) Exploring the meteorological potential for planning a high performance European electricity super-grid: optimal power capacity distribution among countries. *Environmental Research Letters*, 12(11), 114030.

Scaife, A.A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R.T., Dunestone, N., Eade, R., Fereday, D., Folland, C.K., Gordon, M., Hermanson, L., Knight, J., Lea, D.J., Maclachlan, C., Maidens, A., Martin, M., Peterson, A.K., Smith, D., Vellinga, M., Wallace, E., Waters, J. and Williams, A. (2014) Skillful long range prediction of European and North American winters. *Geophysical Research Letters*, 41, 2514–2519.

Sharp, E., Dodds, P., Barrett, M. and Spataru, C. (2015) Evaluating the accuracy of CFSR reanalysis hourly wind speed forecasts for the UK, using in situ measurements and geographical information. *Renewable Energy*, 77, 527–538.

Sonka, S.T., Mjelde, J.W., Lamb, P.J., Hollinger, S.E. and Dixon, B.L. (1987) Valuing climate forecast information. *Journal of Climate and Applied Meteorology*, 26, 1080–1091.

Taylor, J.W. and Buizza, R. (2003) Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, 19, 57–70.

Thornton, H.E., Scaife, A.A., Hoskins, B.J., Brayshaw, D.J., Smith, D., Dunstone, N., Stringer, N. and Bett, P.E. (2019) Skillful seasonal prediction of winter gas demand. *Environmental Research Letters*, 14, 024009.

Troccoli, A., Goodess, C., Jones, P., Penny, L., Dorling, S., Harpham, C., Dubus, L., Parey, S., Claudel, S., Khong, D.-H., Bett, P.E., Thornton, H., Ranchin, T., Wald, L., Saint-Drenan, Y.-M., De Felice, M., Brayshaw, D., Suckling, E., Percy, B. and Blower, J. (2018) Creating a proof-of-concept climate service to assess future renewable energy mixes in Europe: an overview of the C3S ECEM project. *Advances in Science and Research*, 15, 191–205.

Vitart, F., Buizza, R., Balmaseda, M.A., Balsamo, G., Bidlot, J.-R., Bonet, A., Fuentes, M., Hofstadler, A., Molteni, F. and Palmer, T. (2008) The new VarEPS-monthly forecasting system: a first step towards seamless prediction. *Quarterly Journal of the Royal Meteorological Society*, 134, 1789–1799.

Vitart, F. and Robertson, A.W. (2018) The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate and Atmospheric Science*, 1, 3.

Weigel, A.P., Liniger, M.A. and Appenzeller, C. (2007) Generalization of the discrete brier and ranked probability skill scores for weighted multimodel ensemble forecasts. *Monthly Weather Review*, 135, 2778–2785.

Wilks, D.S. (2011) *Statistical Methods in the Atmospheric Sciences*, 3rd edition. Oxford: Elsevier, p. 676.

Zubiate, L., McDermott, F., Sweeney, C. and O'Malley, M. (2017) Spatial variability in winter NAO - wind speed relationships in western Europe linked to concomitant states of the East Atlantic and Scandinavian patterns. *Quarterly Journal of the Royal Meteorological Society*, 143, 552–562.