Open Networks Project

Operational Forecasting

20 November 2020 | Version 2.2

WS1B P3

DOCUMENT CONTROL

Authorities

Version	Issue Date	Authorisation	Comments
1	23/09/2020	John West	Initial template / skeleton report for product team review.
2	02/11/2020	ENA P3 Stakeholders	DNOs and ESO added description of their Operational Forecasting methodologies.
2.1	20/11/2020	Ali Ahmadi & John West	Further updates to provide report for WS1B review.
2.2	28/11/2020	Courtney Madden	Formatting

ACKNOWLEDGEMENTS

We would like to thank P3 stakeholders for their contributions to this report and wider WS1B P3 deliverables including Ali R. Ahmadi, Matthew White, John West, Kanan Ganakesavan, Charlotte Grant, Steve Davenport, Harshil Sumaria, Steve Gough, Grant Mcbeath, Malcolm Grisdale and Daniel Burke.

TABLE OF CONTENTS

1 Introduction

1.1 About ENA

Energy Networks Association (ENA) represents the owners and operators of licenses for the transmission and/or distribution of energy in the UK and Ireland. Our members control and maintain the critical national infrastructure that delivers these vital services into customers' homes and businesses.

ENA's overriding goals are to promote UK and Ireland energy networks ensuring our networks are the safest, most reliable, most efficient and sustainable in the world. We influence decision-makers on issues that are important to our members. These include:

- Regulation and the wider representation in UK, Ireland and the rest of Europe
- Cost-efficient engineering services and related businesses for the benefit of members
- Safety, health and environment across the gas and electricity industries
- The development and deployment of smart technology
- Innovation strategy, reporting and collaboration in GB

As the voice of the energy networks sector, ENA acts as a strategic focus and channel of communication for the industry. We promote interests and good standing of the industry and provide a forum of discussion among company members.

1.2 Our members and associates

Membership of Energy Networks Association is open to all owners and operators of energy networks in the UK.

- Companies which operate smaller networks or are licence holders in the islands around the UK and Ireland can be associates of ENA too. This gives them access to the expertise and knowledge available through ENA.
- Companies and organisations with an interest in the UK transmission and distribution market are now able to directly benefit from the work of ENA through associate status.

1.2.1 ENA members

1.2.2 ENA associates

- Chubu
- EEA
- Guernsey Electricity Ltd
- Heathrow Airport
- Jersey Electricity
- Manx Electricity Authority
- Network Rail
- TEPCO

1.3 Open Networks Project

The Open Networks (ON) Project is a major industry initiative that will transform the way our energy networks operate, underpinning the delivery of the smart grid. The project looks to change how the networks operate to facilitate the transition to a smart, flexible energy system. A key objective is to bring consistency in approaches across networks through existing and new processes to support the transition to Distribution System Operations, interactions with each other and interactions with customers. Open Networks is being delivered through a number of Workstreams and Products.

Workstream 1B of the ON project focuses on Planning and Forecasting activities. These include Product 3 on Real Time Data Exchange which is intended to cover the data exchange requirements to manage whole system activities including Operational Tripping, Connect and Manage arrangements and Service Conflict Management. Workstream 1B Product 3 is also considering how improved operational forecasting could support these and related whole system network operation activities.

1.4 Scope of this report

This report covers Workstream 1B Product 3's 2020 work on Operational Forecasting including the assessment of further data exchange between network companies to improve the delivery of whole system operation activities including Operational Tripping, Connect and Manage arrangements, Service Conflict Management, and service procurement.

The 2020 Project Initiation Document for the ON project separated the work on Operational Forecasting into 3 activities and recognised that the work would be to some extent contingent on progressing the wider development and implementation work on Operational Tripping, Connect and Manage arrangements and Service Conflict Management. As a minimum, the 2020 report is intended to address activity 1 in the Project Initiation Document and consider the forecasting methodologies being used by the NG ESO and DNOs and the data transfers for demand and generation forecasting.

As work on Operational Tripping, Connect and Manage arrangements and Service Conflict Management has not progressed to implementation during 2020, this report largely focusses on activity 1 by considering the existing operational forecasting methodologies used by network companies and identifying opportunities for improvement through further data exchange. A number of operational forecasting Use Cases are also developed to understand where increased data exchange could improve whole network operational activities.

Figure 1 – Extract from 2020 Project Initiation Document

Ref	Product Element	Activities	Deliverables
3	Advise on the ESO/DSO data exchange and the methodology for operational forecasting.	1. Describe ESO forecasting methodology, confirm the information and data required to forecast demand and generation. Ensure consistency between ESO and DSO forecasting methodologies. 2. Following real-time delivery of N-3 OTS, C&M and Service Conflicts empowered by data exchange and ESO/DSO system coordination, assess: a. what data exchange is required to achieve desired forecasting results. b. how operational forecasting can better inform ESO/DSO whole system operation. c. inputs/outputs of information/data required for forecasting algorithms. (Sub-deliverable 2 is dependent on sub-deliverables 1B and 1C.) 3. Assess how the availability of network forecasts empowered by data exchange could change the ESO & DSO procurement of services from DERs.	Report by Dec 2020 to achieve part 1 as a minimum.

1.5 Structure of this report

As well as this introductory section, the report is structured into 6 further sections. Section 2 provides a high level description of Operational Forecasting and recaps related work carried out in Workstream 1B Product 3 during 2019.

Section 3 describes the how operational forecasting is carried out in NG ESO and also the developing capabilities in DNOs.

Section 4 considers a number of Use Cases where improved operational forecasting and data exchange could improve particular operational activities.

Section 5 further considers some of the tools and algorithms being used for Operational Forecasting.

Section 6 considers how Operational Forecasting might be improved through further data exchange between network companies. It also considers how the data exchanges identified for the Use Cases in Section 4 might be taken forward.

In line with network company undertakings to consider where data can be made more transparent, Section 7 considers whether the data being handled for Operational Forecasting would have wider value to industry stakeholders and the extent to which the data can be shared more widely.

2 Operational Forecasting - Introduction

2.1 Introduction to Operational Forecasting

Operational Forecasting is already carried out by network companies for example to enable energy balancing, to help plan network activities (e.g. outages) and to identify the requirements for services.

In the ESO, operational forecasting is relatively mature as it is a key activity for energy balancing. Operational forecasting is increasingly being considered DNOs as well, as they look to manage increasing levels of connected generation with constrained networks.

The timeframes for operational forecasting could be several months ahead (e.g. forecast requirements for the coming winter) up to day-ahead or on-the-day forecasting of requirements.

2.2 Recap of 2019 Work

The work carried out in by Workstream 1B Product 3 during 2019 focussed on the data exchanges to support short term (48 hours ahead) operational forecasting and considered 3 worlds being considered as part of Workstream's 3 DSO development work. Areas of data exchange and preferred methods of communicating data were identified. Figure 2 is a summary of the potential data exchanges identified in the 2019 Workstream 1B Product 3 final report.

Figure 2 – Forecasting data exchange signals and methods of communications. (From 2019 report)

	World A – DSO		World B – Joint Procurement		World D – ESO	
	DSO to ESO	**ESO to DSO**	**DSO to ESO**	**ESO to DSO**	**DSO to ESO**	**ESO to DSO**
What Data to Exchange	•DSO every half hour Forecast (Demand and Generation at 11kV and 33kV feeder level). Presented to ESO at GSP level.	•ESO National Forecast (Demand and Generation). •Interconnector Forecast Flows.	•DSO every half hour Forecast (Demand and Generation at 11kV and 33kV feeder level). Presented to ESO at GSP level. •DSO) Outage. •E&W forecasted 132kV active flows.	•ESO National Forecast (Demand and Generation). •Interconnector Forecast Flows.	•N/A	•N/A
	Data Types	**Method of Communications**	**Data Types**	**Method of Communications**	**Data Types**	**Method of Communications**
How to Exchange the Data	•Web with CIM data exchange format	•Web with CIM data exchange format	•Web with CIM data exchange format	•Web with CIM data exchange format	•Web with CIM data exchange format	•Web with CIM data exchange format

3 Description of Ongoing Operational Forecasting Capability

3.1 Operational Forecasting in NG ESO

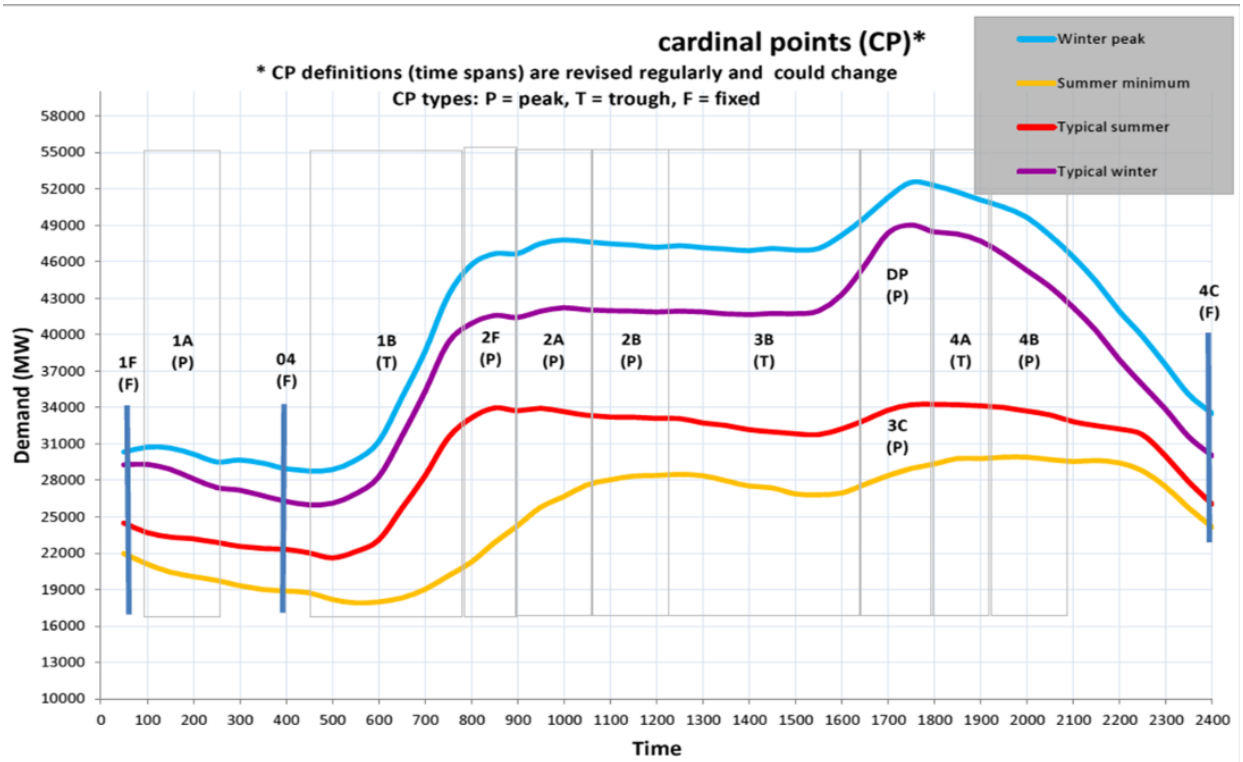
NG ESO forecasts transmission system demand on a national basis. These forecasts cover several time periods during the day and are published to the Market to assist balancing the electricity transmission system as safely and economically as possible. Usually, local Grid Supply Point (GSP) demand forecasts are produced depending on specific network management needs, though this is changing as part of a strategic project to replace the existing forecasting system (this is described later).

Transmission system demand is lower than total UK consumption due to embedded generation connected on the distribution network, and NG ESO uses various modelling techniques and external data regarding embedded generation capacities, to derive demand forecasts.

3.1.1 Demand Forecasting Process

The typical shape of the daily demand profile does not change significantly from day to day, with the main changes to the profile being the timing and intensity of key critical points – e.g. peak demand, daytime minimum etc, which are referred to as 'Cardinal Points'. Therefore, the forecasting methodology focusses on a chosen set of 'Cardinal Points' and a profile is interpolated between these points based on historic demand.

Below is a diagram showing typical seasonal demand profiles, pre-COVID 19.



The demand at each Cardinal Point (CP) is broken down into separate forecast components according to the following formula:-

$$\begin{aligned} &\mathit{Forecast\ Demand} \\ &\quad = \mathit{Basic\ Demand} + \mathit{Weather\ Correction} - \mathit{Time\ and\ day\ of\ Week\ Correction} \\ &\quad - \mathit{Embedded\ Wind\ and\ Solar} + \mathit{Special\ Events}\ (e.g.\ 6\ \mathit{Nations}, \mathit{Wimbledon\ etc}) \end{aligned}$$

where

- Basic Demand is the underlying normalised total UK consumption which is influenced by long term changes e.g. economy, efficiencies, human behaviour.
- Weather, time of day and day of the week correction variables are derived from mathematical models that consider temperature, illumination, windspeed, sunset time, adjustments for trends on earlier demands etc. Separate models have been developed for each CP.
- For embedded wind and solar, generation capacities and locations are found from BEIS/public databases. Wind turbine models, basic and enhance statistical models for solar, along with machine learning for national solar power are used.

3.1.2 Met Office Data

Data from the Met Office is factored into the mathematical models; each forecast received is for the next 14 days ahead and at hourly resolution. The weather variables that come through are temperature, solar radiation, wind speed and wind direction.

To derive the renewable generation output, data from the nearest weather station to the generator is used by the models. There are approximately 98 weather stations to represent the GB wind fleet with around 52 locations for radiation forecasts for solar PV.

3.1.3 Research and development

Over several years, NGESO has developed projects with universities and mathematical modelling institutes to improve forecasting techniques, though a major change to the forecasting system is currently underway. This will provide more accurate forecasting on a half -hourly, per GSP basis, allowing easier integration with other systems and programs that are utilised across the business. This change to our forecasting system is described in the following section.

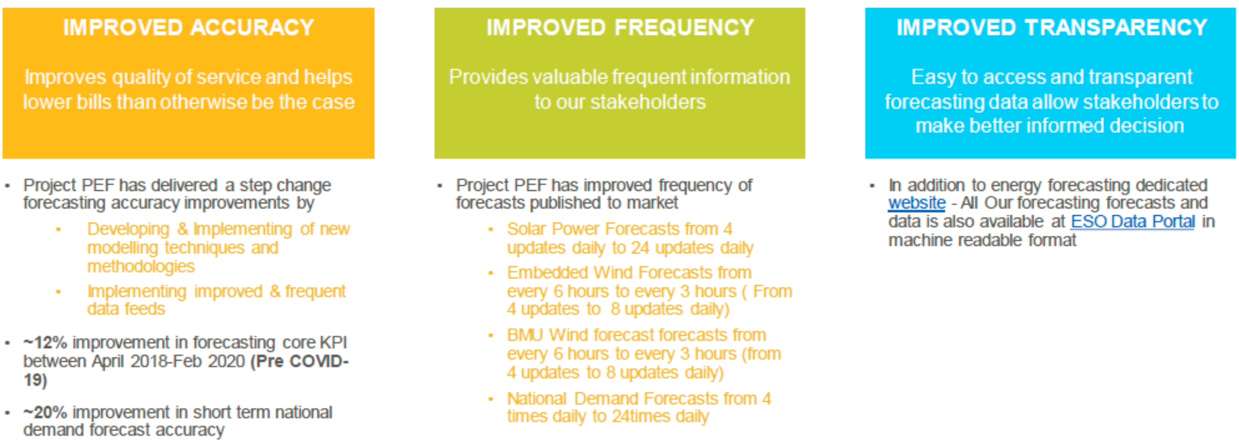
3.1.4 Platform for Energy Forecasting Strategic project

Two years ago, NGESO developed a strategic forecasting project with the aim to replace the existing energy forecasting system with an advanced cloud-based platform for energy forecasting (PEF) that would focus on

- designing and improving forecasting models, methodologies,
- the application of advanced statistical learning, machine learning modelling techniques and automation.

There are several phases to the project that also includes delivering more frequent and accurate forecasts so that market participants can balance their position ahead of time, reducing the requirement for balancing services and residual balancing actions.

Aspects that have been developed so far are shown in the figure below.



The platform for energy forecasting will significantly enhance the forecasting capability of NG ESO as it can run tens of thousands of concurrent forecasting models using some of the latest machine learning, integrated with statistical approaches. It is also self-adapting to the changes in the energy system and will continue to morph to a zero-carbon model.

3.1.5 National Demand Forecast

The mathematical model for national demand forecasting for each cardinal point from now to 365 days ahead has been re-designed. This new modelling approach can capture behavioural patterns to generation more efficiently and has resulted in approximately 20% improvement across the accuracy of all cardinal points. The same cardinal point methodology is used and interpolated to derive half-hourly demands, then forecasts for unmetered embedded wind and solar power generation are subsequently included to arrive at a National Demand forecast.

3.1.6 Grid Supply Point Forecast

A completely new machine learning Grid Supply Point (GSP) forecast modelling approach has been developed. This methodology encompasses net demand, unmetered embedded solar and wind generation forecasts for every GSP in GB.

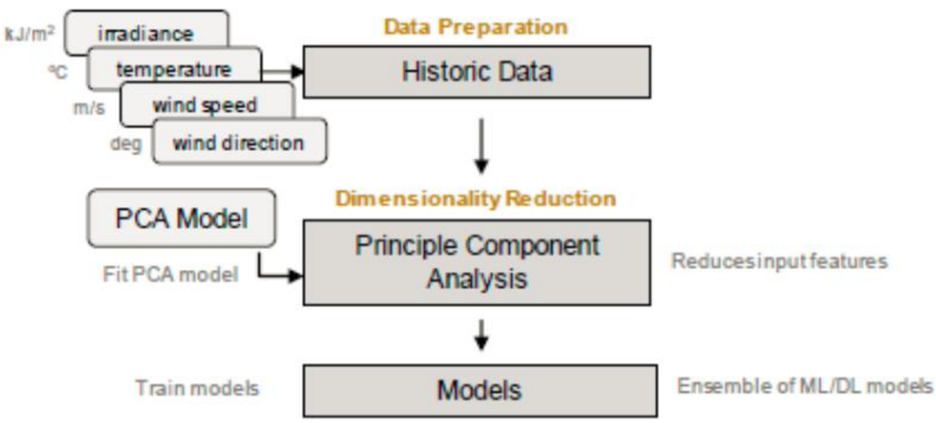
There are several elements that are incorporated to produce a GSP forecast and how machine learning is applied to train the model that are outlined below.

GSP Model Training:

Training the GSP model involves

- Approximately 10 deep learning and machine learning models per GSP
- 7000+ models re-trained regularly with latest data

The figure below gives an overview of the principles of the machine learning model training for a single GSP.



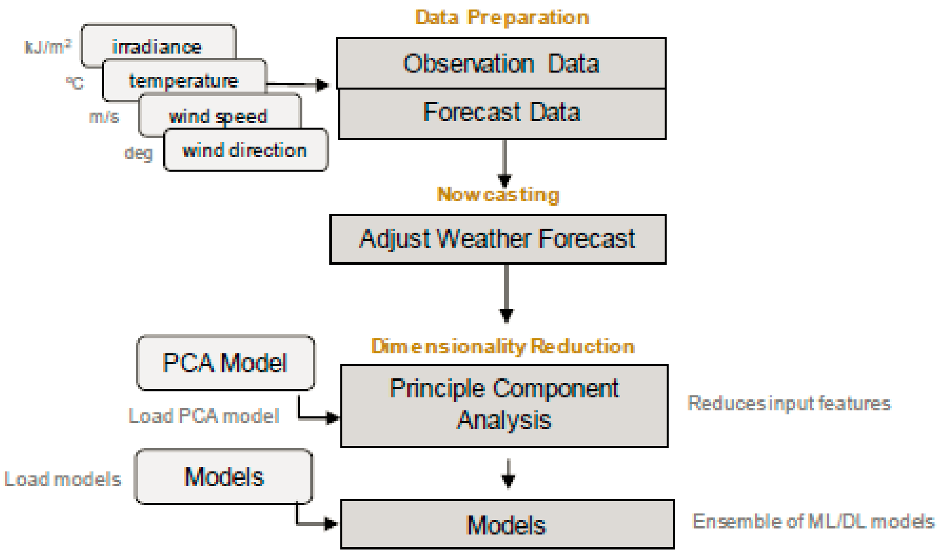
GSP Model Prediction:

The prediction part of the model encompasses a high-level approach that

- Updates every hour with new data that's available
- Makes predictions for each settlement period produced from now to 14 days ahead
- Uses weather data to correct for errors in forecasts throughout the day
- Has the potential to increase the forecasting regularity and granularity dependent on data frequency and performance of the model

In addition to weather data, historic metering is used, along with Electralink data for over 1900 embedded wind sites and GSP level solar outturn data from Sheffield Solar.

The principles of the machine learning model prediction for a single GSP is shown below.



Validation of these GSP forecasts has indicated a step change in improvement in forecasting accuracy compared to previous techniques.

3.1.7 Timelines for the PEF project

From now until the end of this financial year, the focus of the project is to develop the national demand machine learning model and methodology, improving the medium (2-14 days ahead) and long term (2-52 weeks ahead) and wind forecasts.

During summer 2021, subsequent phases will look at the potential to incorporate additional forecasts for other generation fuel types and improve forecasting efficiency through innovation. This could involve bespoke forecasts for individual units where frequency, resolution and forecast horizon can be varied depending on stakeholder needs that can benefit consumers.

Further information about PEF and the strategic roadmap is available on the NG ESO website

3.2 Operational Forecasting in UKPN

This section is based on our provisional research, analysis and results using a generic ANN algorithm. Operational Forecasting tools will allow engineers to transition away from using worst-case operational scenarios for modelling purposes to using forecast load and generation data. Worst-case scenarios in many areas across EPN and SPN consider maximum generation in coincidence with minimum demand. With an increasing amount of distributed generation connected, the network is sometimes considered to be constrained in these worst-case scenarios. In certain planned outage situations, generators need to be curtailed to mitigate against potential network violations. Using these scenarios will not always be fully representative of actual running conditions and thus forecasting tools will potentially allow more generators to operate unconstrained during planned outages.

The two main input data to the Operational Forecasting Engine which were used for KASM, Power Potential and UKPN’s in-house forecasting tools are presented below:

Metered data: The metered values for the system load points, wind and solar generators are retrieved from the UKPN OSIsoft PI Historian, as a one-off to train the forecasters, and are periodically updated. The time granularity of the metered data will be half-hourly (i.e. at 1hour intervals and 30-minute intervals) and the values will be in real power or current measurement terms.

Met Office weather data: The Met Office weather data includes weather forecasts and observations. The Forecaster weather data is delivered to the Forecaster platform to a designated folder via UKPN service bus architecture. The data is fetched periodically from a designated folder which can be accessed by the Forecaster. The Met Office weather data is updated every six hours (at 00:00, 06:00, 12:00 and 18:00 every day). Each update provides five-day-ahead (120-hour) weather forecasts and previous six-hour observations from a list of predefined weather stations. The data includes values of following weather measurements:

- Temperature in Celsius degree (C)
- Humidity (%)
- Radiation in W/m2
- Wind direction in degree (0-359 degrees)
- Wind speed in m/s
- High cloud cover
- Medium cloud cover
- Low cloud cover
- Total cloud cover

The Forecasting Engine modules take historical system measurements and historical and real-time weather forecasts as inputs to train the forecaster algorithms, and utilises advanced adaptive analytics to produce accurate forecasts. The forecasting engine produces forecasts of:

1. Metered system load points
2. Outputs of wind generators
3. Outputs of solar generators

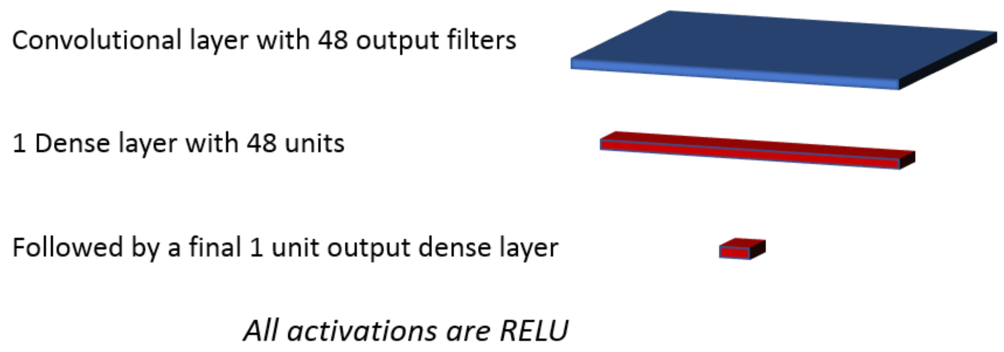
The Forecaster Model Engine is the module which produces the desired forecasts. The input data is retrieved from the database. The outputs (the responses of the Forecaster Model Engine to the inputs) are used to produce the final forecasts for load, wind and solar generation in time. The Forecasting Engine implements a portfolio of predictive analytic models, including artificial neural networks (ANNs) and relevant optimisation techniques. These analytic models and computing techniques are integrated to produce the predictive analytics.

Initial results show that the forecasts provide similar results when benchmarking against wider industry forecasts and actual recorded data. It is noticeable that a few load data point forecasts can have high errors, which could be due to a number of factors including: masked embedded generation, limited metering data or poor quality metering data. The overall results show that the majority of forecasts lie within the acceptable range.

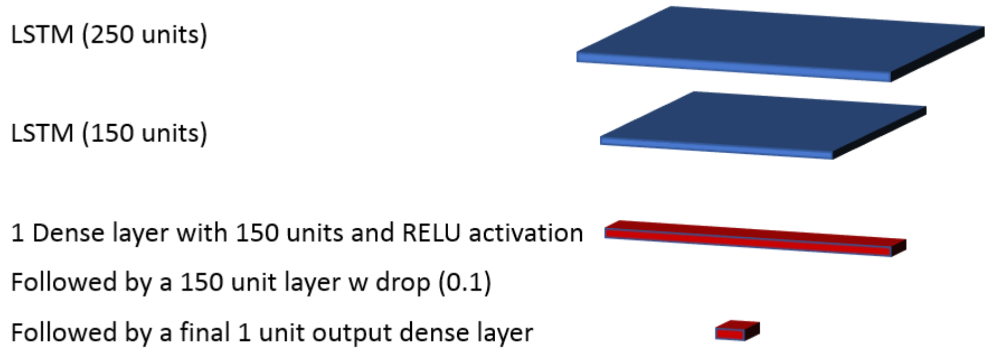
There is currently work ongoing to develop operational forecasting further using the latest advances in machine learning and deep learning. Using the previously described datasets, the Metered data and Met Office Weather data, development is underway to determine the best architecture to use to develop load and generation forecasts. The aim is to develop a proof of concept model by comparing Convolutional Neural Networks (CNNs), Long Short Term Memory Recurrent Neural Networks (LSTM RNNs) and Extreme Gradient Boosting (XGBoost) for substation load and embedded generation forecasts. The benefit of using these is that the data does not need to be stationary, which is a requirement for methods such as Seasonal Autoregressive Integrated Moving Average (SARIMA), whilst the models are most robust than traditional methods. The CNN and the LSTM RNN are both neural network methods, utilising deep learning techniques which are widely used across many industries. CNNs utilise convolutions to extract key features which the algorithm learns from. The LSTM is a type of RNN, which feeds data from previous stages as well as the current stages into each training step, thereby creating memory, which is essential for time series forecasting. Each LSTM cell has a series of gates, and the function within each gate defines whether the data is kept in memory or discarded for the next step. XGBoost algorithm is a tree based algorithm, and uses an ensemble of trees to generate forecasts. It uses boosting, which is an ensemble method which reduces bias and variance in training. It creates a series of weak trees which focus on the weakness of each previous tree, and then readjusts the weights, which overall leads to a strong learner.

The architectures used for the neural networks are shown below.

CNN:



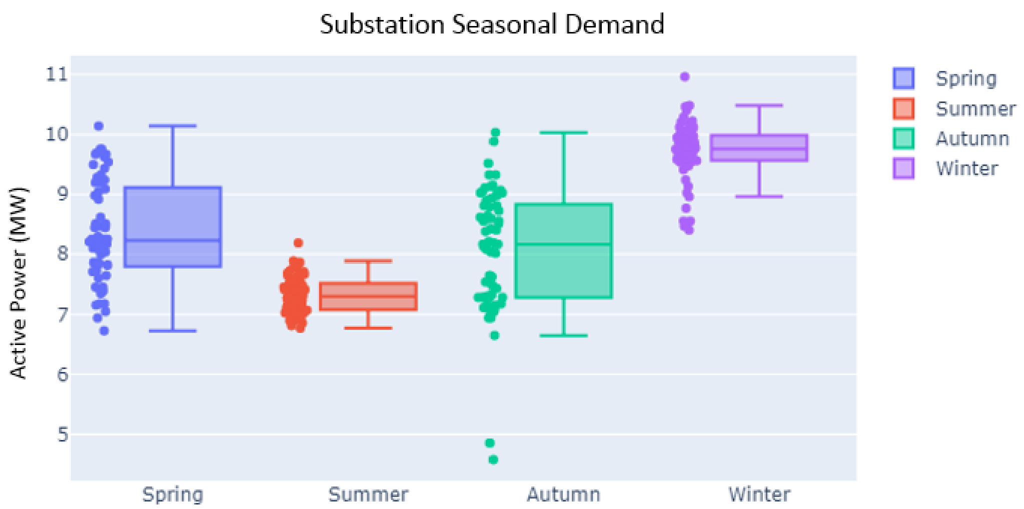
LSTM:



All activations are RELU

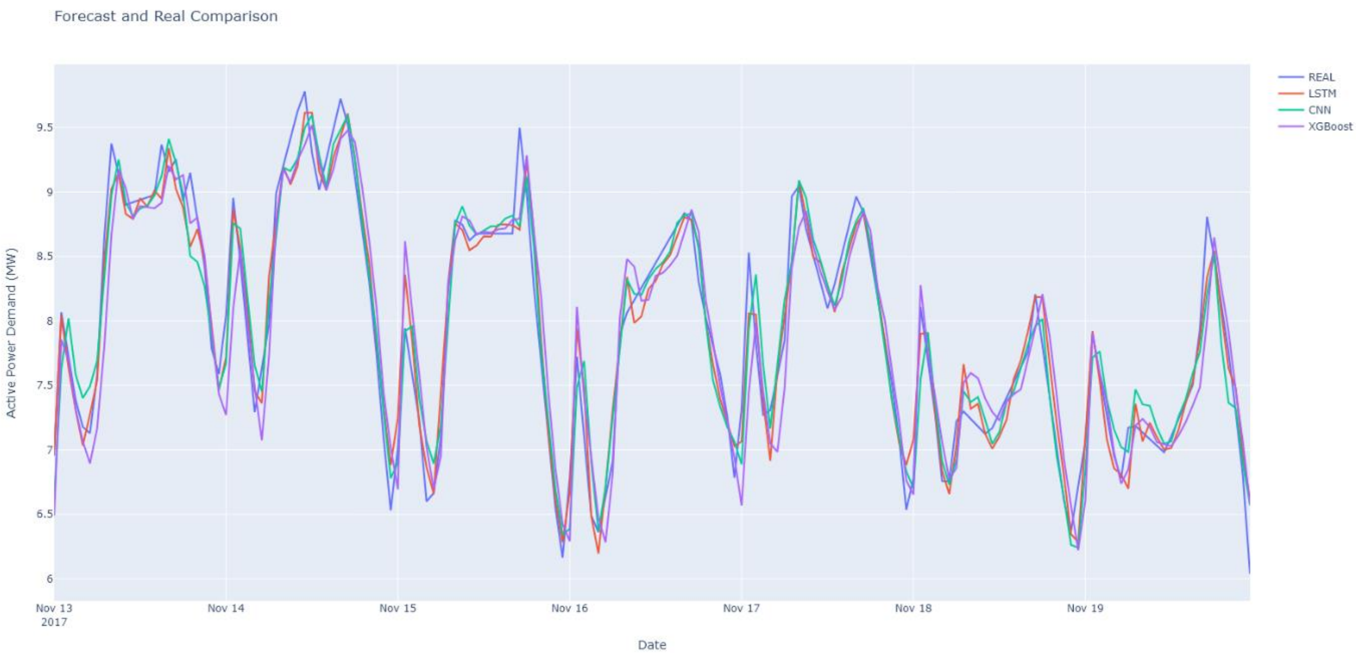
An example substation and a windfarm have been selected based on their proximity to weather stations. Four years of metered data for each site was extracted. For the substation this consisted of both Active and Reactive Power, and for the windfarm only Active Power was considered.

When performing an exploratory data analysis of the substation demand data some load transfers were found. To avoid over-complication of the initial model, any periods with load transfers were not considered when training the model. A simplistic approach was taken to deal with missing measurements by setting them to zero. This could be improved in the future by using other measurements to interpolate and fill in the gaps. There was a clear seasonal trend in the data which was explored using the below box plot. A monthly trend was also explored, and due to the clear monthly trend, both seasonal and monthly features were added to the dataset. There was also found to be a relation between the type of day (i.e. weekday, weekend and holiday) as well as the time of day. Therefore these were also added as features to the dataset. The historic weather data had an hourly granularity and as such the substation demand data was down sampled to match this. Any missing values were again replaced with zeroes, and the two datasets were then joined on the date time axis.



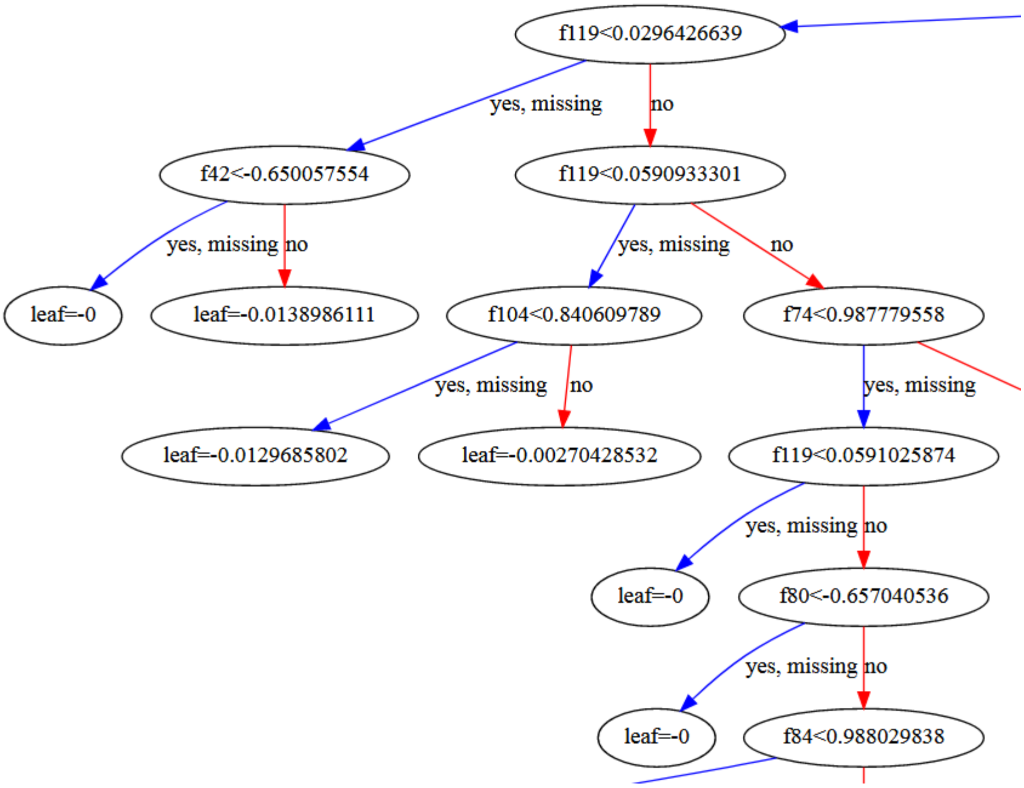
The data was then fed into each of the models described previously. The data was split into a training set, a validation set and a test set. The algorithm trained on the training set, and after every epoch of training it determined how well it had learnt by testing on the validation set. Once there were no improvements on the validation set the training stopped. The epoch of training which returned the lowest validation error was reloaded and used to forecast a year of data using the historic weather as input.

This has shown to be extremely successful for Active Power, producing errors of less than 3%, and 12% for Reactive Power. The below graphs show each forecast in comparison to the real measured data for a randomly chosen week of the year.

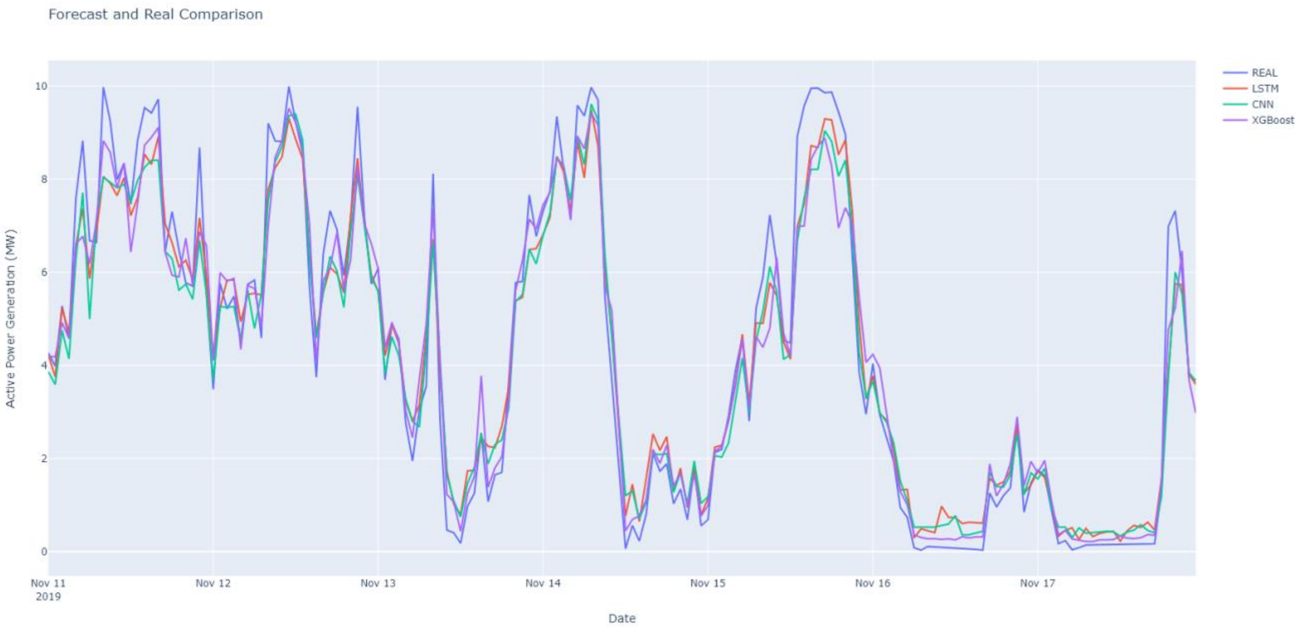


Further development is underway to extend this to generation. A windfarm in close proximity to a weather station was chosen, and the active power generation data was extracted. A similar method was followed to prepare the data, and a training, validation and test set were created. The CNN, LSTM and XGboost models were trained and then used to predict a years' worth of generation output, and the initial results were promising. The root mean square error (RMSE) was found to be 12.81% for the LSTM, 13.42% for the CNN and 13.89% for the XGBoost model. This was using a basic architecture for each algorithm, and can be improved with hyper-parameter optimisation in each case, as well as exploring more complicated models.

A sample of the model created by training the XGBoost algorithm can be seen below. The overall model had many large trees, so a snapshot of a single tree is shown below.



The results for a particular week are shown for each algorithm below:



In all cases, the active power demand, the reactive power demand and the active power generation, the LSTM RNN performed best, although it was closely matched by the CNN and the XGBoost algorithms.

It is important to note that these methods are limited by the quality of the data. In the case of the weather data, methods must be investigated in order to obtain more geographically granular data. This is required due to the hyper-localisation of weather, and to truly create a relationship between demand and generation and the weather, the weather data used must originate from a location sufficiently close to the site in question.

The next steps for the project are to further verify the results using other demand and generation sites. Once these are verified, a methodology will be developed to extend this to non-weather dependent generation, substations with high penetrations of generation and substations with load transfers. Each of these pose a unique challenge and may require different algorithms. It will be unlikely that there will be a single algorithm which can tackle all of these situations, the overall solution will likely be one composed of multiple algorithms.

To enable whole system development, the DNOs would greatly benefit from the sharing of data and forecasts from NGESO including the daily GSP forecasts currently being developed. This would allow for DNOs to better manage the distribution transmission boundary, and therefore allow the ESO to run a more efficient network operation.

Additionally, there is development underway on a long-term forecasting tool. The Strategic Forecasting System (SFS) is a project being carried out in partnership with Element Energy (EE), Imperial College London (ICL) and CGI. The aim is to create a load, generation, network constraints and LRE forecasting system which is embedded into UKPN’s BAU planning process, to facilitate efficient network investment and provide robust evidence for UKPN’s ED2 LRE submission. The initial stage of the project is well underway, aiming to finish at the end of 2020. The project will bring together a lot of existing tools and outputs and improve them. It will do so by adopting a stringent data cleaning methodology, feature the updating and maintenance of models and validate the results extensively. An early output has been the Distribution Future Energy Scenarios (DFES) which have been well received by stakeholders and will be used in future development to inform the growth scenarios used in the project.

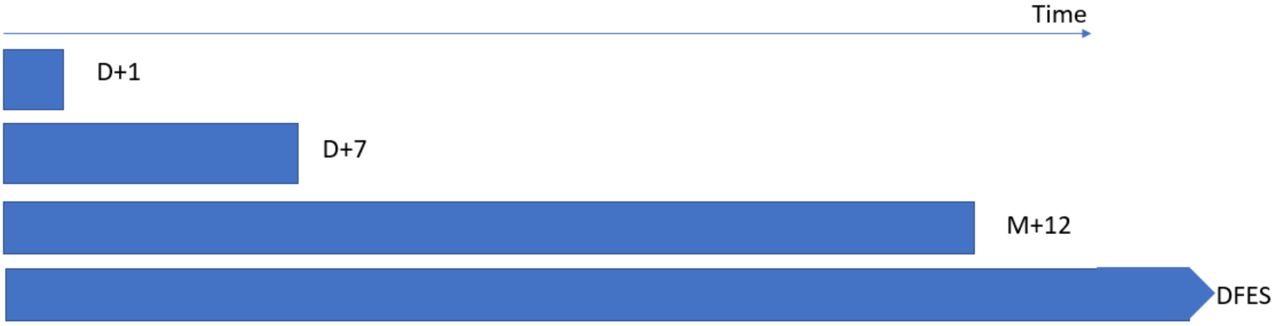
The SFS is bringing together multiple datasets such as MPAN Demand and Generation, Pi metered data, weather data, network topology amongst others to produce a robust long-term forecast, enabling UKPN to efficiently plan and run the network. The approach is DSO ready as it will include the hours over firm, the enable flexibility assessments as well as the traditional cost benefit analysis.

3.3 Operational Forecasting in SSEN

There has been limited deployment of operational forecasting in SSEN however there has been some foundational work for the business process and building on planning timescale forecasting which has informed the broader forecasting framework SSEN aims to deploy. Much for the practical learning for operational forecasting will be developed through the Transition Project.

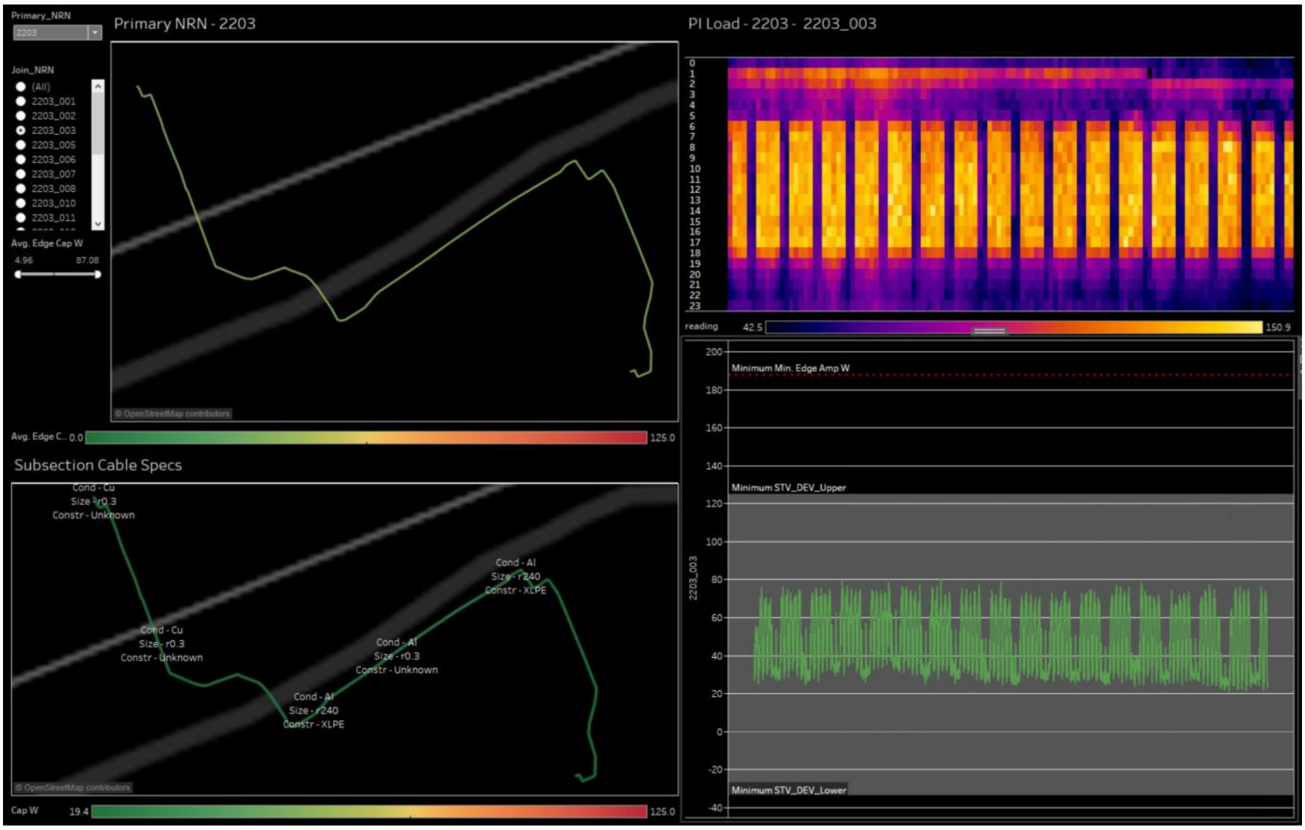
3.3.1 Forecasting Framework

Four key forecasting time horizons have been identified, focusing on slightly different business requirements and looking at differing granularity of forecasts.



- Operational, looking to inform and validate control room activity. It will also be used to call on flexibility assets. D+1 / rolling hourly, 48hrs 10min granularity
- Short term commercial, looking to validate planned running arrangements and prime flexibility assets and procure any predicted shortfall of flexibility. D+7 / rolling daily 7 days, 30min granularity.
- Long term commercial, looking to set up mitigating actions through striking flexibility contracts and operational techniques plus identifying and instigating any additional reinforcement requirements. M+12 / rolling monthly, 12 months daily peak and trough granularity or 365 by 48HH for peer to peer facilitation view
- Planning, identifying longer term constraints and used for signposting. Driven by the DFES and incorporated into Network Development Plans. A + 30 / rolling annual, 30 years with indicative seasonal 48HH profiles.

In addition to the timescales and business processes considered, the is the level of disaggregation of the forecasts are important to consider. There are a few key underpinning inputs that have been recognised. Firstly, the locations for all customers and DERs needs to be correctly identified so a full and accurate network connectivity model is critical. Secondly, the disaggregated baseline demand needs to be understood. Through the planning forecasting work a Disaggregated Load Model is being developed with plans to apply these methodologies to nearer real time. Fundamentally this approach uses demographic and local DER data to build up customer profiles per LV feeder then using a combination Smart Metering data and LV monitoring to train a Machine Learning algorithm to predict the HH load profile. This is scaled and validated with monitored HV feeder data. This work is being underpinned with plans to incorporate aggregated Smart Meter data, LV monitoring and the outputs for the Disaggregated Load Model into the connectivity model to support business wide forecasting requirements.



3.3.2 Transition Project Forecasting Requirement

Project TRANSITION is an Ofgem NIC funded innovation project which is designed to inform the wider evolution of the DNOs to DSO. It is led by SSEN but also has ENWL as a collaborative partner. It is also closely linked to the other NIC funded projects EFFS at WPD as well as FUSION at SPEN via the TEF forum.

One of the objectives of TRANSITION will be to implement a market for flexibility in order to resolve network congestion issues with flexibility contracts from DNO connected assets. This will involve real-world physical trials within the Oxfordshire area.

Therefore, a significant task as part of the TRANSITION process chain will be to calculate the MW/MWh requirement for flexibility at different points in the distribution network. To achieve this, a forecasting tool will need to be put in place that, when linked to a power systems analysis tool (via a Whole System Coordinator and a Neutral Market Facilitator platform), will analyse the congestion levels in the network at a number of different timeframes corresponding to the commercial and operational aspects of the market implementation.

Regarding forecasting horizon timeframe, it is still to be confirmed but there will likely be a system assessment required ~ 1 year ahead and a few days ahead for flexibility commercial contractual tasks, as well as down to operational real-time for flexibility asset arming and physical dispatch.

An important point to consider in the TRANSITION forecasting objective is that flexibility contract considerations are mainly motivated by a requirement to protect the network from potential overload conditions, and therefore the forecasting solution will need to be cognisant of extreme peak conditions rather than typical or “expected” peak flow conditions. Therefore, the forecasting tool objective needs to be aimed at assessing the range of values of what possibly could happen that the network will have to accommodate as opposed to providing a single best view of what will happen at each time stage of interest. Forecast uncertainty management will therefore be key.

The main project activity in this forecasting area will be ongoing in late 2020 and throughout 2021. At present, the TRANSITION team has been developing a scope of works for the forecasting tool, as well as performing market research to see what providers and consultants may be able to offer solutions. The learnings from the EFFS project have been especially useful in this respect to guide the direction required for TRANSITION.

Inputs to the forecasting tool are likely to be weather (wind speed, solar irradiance, hydro inflows for renewable generation profiling, and temperature for any demand dependency relationship issues) as well as calendar issues related to season, weekday, time of day, special days (e.g. Christmas, Bank Holidays, etc).

Forecasting of dispatchable renewable generation, non-renewable small scale conventional generation and storage is perceived at this time as a significant challenge. These generation sources are driven by non-weather trends and a complex mixture of commercial signals across a number of markets, and their likely value of generation within the interval of max and min generation may be hard to quantify accurately at any given forecast time horizon, and thus their profile may have a substantial element of uncertainty related.

Other perceived challenges that will need to be addressed are potential data gaps, visibility of the demand and generation sources on the network and disaggregation of some generation from demand that may not be fully clear to the DNO, and requiring a suitably long a history profile of data for statistical robustness of any forecasting model tuning.

More information may be seen on the transition project website at https://ssen-transition.com/

3.4 Operational Forecasting in SPEN

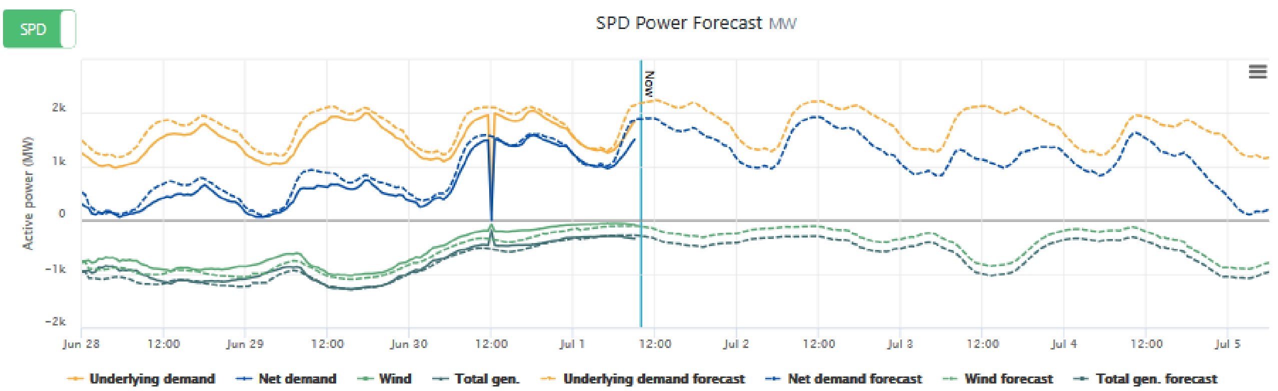
SP Energy Networks forecast the Distribution electricity demand and embedded generation in the SPD & SPM areas to better understand where congestion may arise. Understanding where congestion may occur will allow the DSO to take mitigating actions to ensure assets are not overloaded and customer security of supply is maintained. The forecasting platform PRAE delivers a 4 day-ahead underlying demand, embedded generation and net flow forecast at license area, District, GSP. Primary and 11kV feeder levels.

In a future DSO environment flexibility in new connections will need to be considered as an alternative to traditional forms of network investment. Flexible connections potentially will result in distribution networks being operated near or above full load capability during periods of maximum demand or generation output. During these periods, where the network potentially could be operated out-with its network capability, constraints will need to be issued to ensure network congestion is avoided. It is essential the DSO understands where congestion will occur in advance of real-time otherwise overloaded networks could lead to unplanned interruptions to customers. The DSO will need to understand where, when, how much and for how long will congestion occur to ensure flexibility can be despatched to operate the network within operational limits. PRAE forecasting platform has the capability of alerting operators of future congestion and how much flexibility and when would need despatched to mitigate the issue.

The PRAE forecasting model utilized 3 years-worth of ½h hourly average analogue data at GSP & Primary S/S levels to train the model to forecast demand and generation in the future at a granular level. SP Energy Networks collaborated with Sia Partners to deliver this initiative and demonstrated the value of Utilities existing data sets and the benefits this data can deliver by exploiting data science. SIA Partners use a number of different modelling techniques and approaches to forecast demand and generation. Modelling techniques referred to as GAM, RandomForest, XGboost and Hybrid are all undertaken to establish the most accurately forecasting approach. MAPE (mean average percentage error) is a prediction of forecasting error against each of the modelling approaches. It is expected the MAPE will further reduce with increased data sets and further model tuning over time. The statistical approach and data anomaly detection will ignore poor or missing data if this could detrimentally skew the forecast prediction.

The functionality of the operational forecasting tool is explained below:

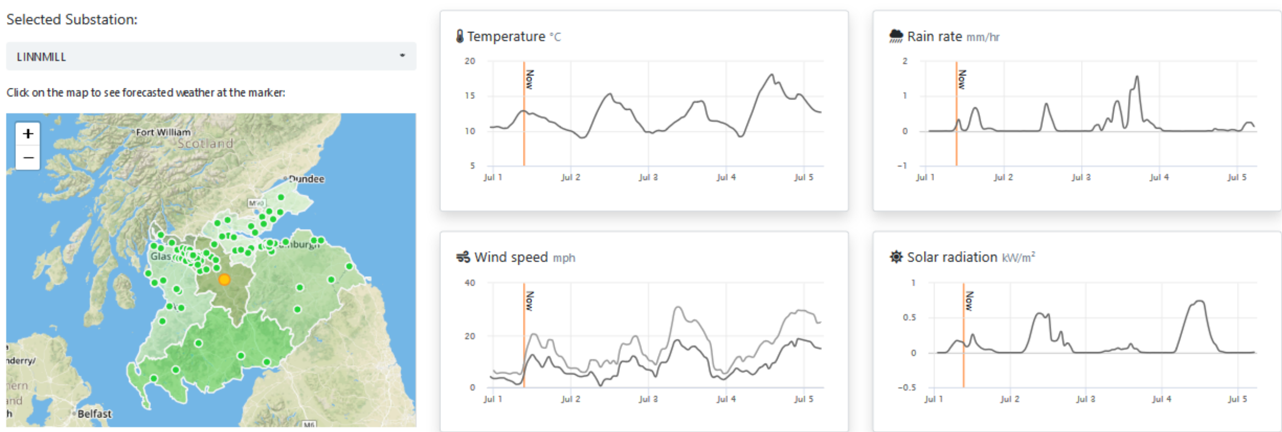
- Forecasting platform for both SPD / SPM licenses.
- 4 day ahead weather forecasts used to:
    - Forecast EHV / HV generation ½hrly output
    - Forecast ½hrly EHV / HV underlying demand
    - Forecast data reconstructed to Net flow data
- Constraint Analysis
    - Informs where, when & how much flexibility req'd
- Model Accuracy
    - Reports accuracy of forecast against actual
    - Reports data quality issues

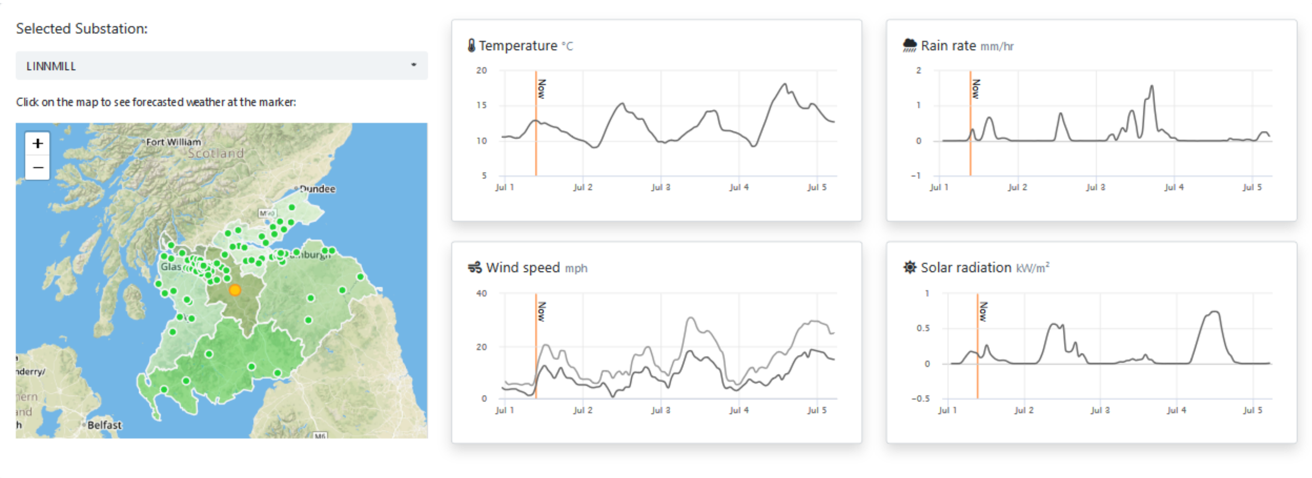


Forecast	Now	01/07 12h	01/07 18h	02/07 0h	02/07 6h	02/07 12h	02/07 18h	03/07	04/07
Underlying demand	2177	2244	2182	1556	1378	2229	2118	1273 | 2131	1223 | 1958
Net demand	1895	1899	1707	1138	1065	1928	1616	835 | 1372	496 | 1633
Wind generation	106	167	294	245	138	117	316	201 | 858	126 | 886
Total generation	282	345	475	418	312	301	502	374 | 1036	310 | 1060
Constraints 132kV	●	●	●	●	●	●	●	●	●
Constraints 33kV	●	●	●	●	●	●	●	2 1	3
Constraints HV	1	1	1	1	1	1	1	1	1

Benefits include:

- Increased awareness of COVID impact on demand
- Validated SPEN LFDD & demand disconnection plans
- Assisted SPEN / NGESO ODFM assessment
- Assisted SPEN develop GEN disconnection schedules
- SPEN requested to demo forecasting platform at NGESO Summer Outlook session June 2020

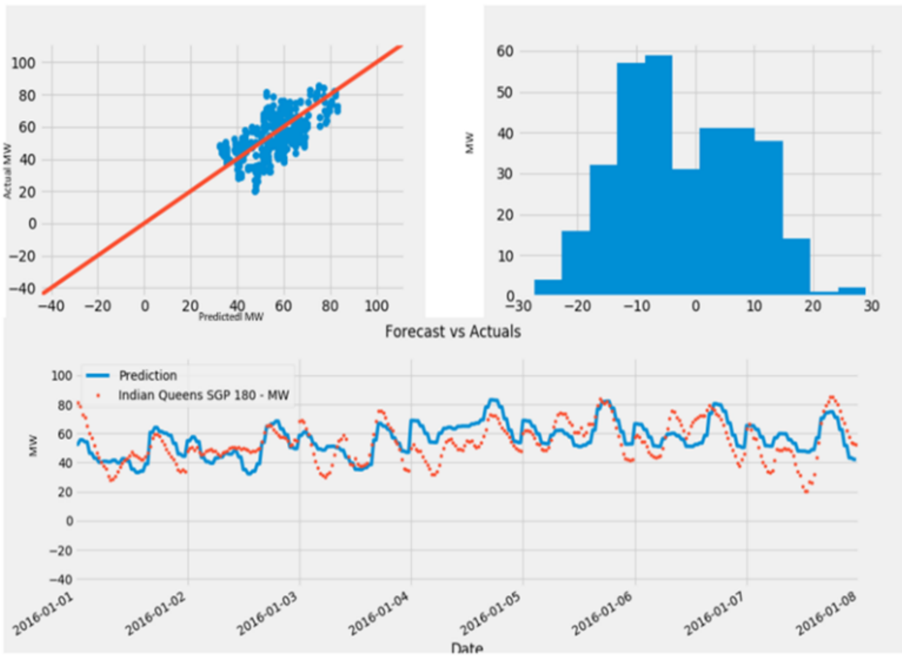




3.5 Operational Forecasting in WPD

Forecasting is a module that sits within the EFFS Networkflow solution. There have been two pieces of work commissioned to explore and develop the scripts within this module, both completed by SGS:

- CR0 was first completed to explore and test the accuracy of various modelling methods.
- CR1 was a follow-on piece of work to explore how weather forecasts can be utilised as well as engineering models.



Example Analysis: India Queens T1 Week Ahead results for real power, 1-7 January 2016.

CR0 explores the forecasting of real and reactive power flows at Primary substations, Bulk Supply Points (BSP), and Grid Supply Points (GSP) over different time horizons: six months ahead; one month ahead; one week ahead; one day ahead; and one hour ahead.

The following methods investigated:

- **Auto-Regressive Integrated Moving Average (ARIMA).** A classic statistical modelling approach for building time-series forecasting models.
- **Long Short Term Memory (LSTM) Artificial Neural Networks.** A specific type of deep-learning neural network for learning patterns in time-series data.
- **Extreme Gradient Boosting (XGBoost).** A machine-learning technique based on decision trees that has performed well in recent machine learning and forecasting competitions.

Key conclusions include:

1. For the majority of test cases, Extreme Gradient Boosting outperformed the other methods tested.
2. Techniques based on historical data work best on short time horizons (hour ahead and day ahead).
3. For the Primary and BSP cases with low penetration of wind and solar, relative to yearly demand, a feature set containing only temporal trends will provide predictions with acceptable levels of accuracy; for higher penetrations of renewables, predictions benefit from the addition of weather features to meet accuracy requirements.
4. For the GSP power flows, the stochastic nature of the renewable generation made it more challenging to identify trends/patterns from historical data. By forecasting on an individual transformer basis and then aggregating the forecasts yielded better results.
5. The results for the performance of the techniques themselves are difficult to decouple from the skill of the data scientist building the model. In the EFFS analysis, metric such as training time, tuning time and forecasting time were considered to give an indication of what would be involved to use these techniques at scale.

As per recommendations in the CR0 work, CR1 explored:

- The performance of the XGBoost model with new feature data, i.e. weather forecast data. The XGBoost models, previously constructed in the earlier phases of EFFS, which included historical weather data only, were updated be able to use historical weather data and short term weather forecast data – the short-term forecast weather data to be used for forecasts up to and including a week in advance, where weather forecasts are expected to be available and more useful than seasonal average values.
- The performance of engineering models of WPD generator sites are driven by weather forecast data. New engineering models were constructed that use mathematical expressions of how a generator behaves, expressing how external weather stimuli interact with the physical construction and limitations of PV and Wind sites to produce MW and MVar export.

The utilisation of weather forecasts in addition to historic weather data does increase the accuracy of machine learning models:

- Machine Learning (ML) performs better in general because site operational coefficients are embedded in ML decision trees, whereas EM model a baseline coefficient applied across board. ML tunes itself from the data.
- Engineering Models (EM) Peak prediction better than ML. ML is trying to minimise error across the whole time series and not just the peaks and appears to smooth out some peaks.
- The comparing of both has provided insight that a possible ensemble approach (using ML and EM models together) to forecast could best reduce error when procuring services.

Implementing Forecasting Methods: an Open and Reusable Approach:

- To allow others to use, reproduce or even improve on the results of the UK-customer funded work in this project, the underlying forecasting tool-chain used by the project's forecasting methods partner has been detailed in available reports. This has been done at a suitable level to allow the TRANSITION and FUSION projects to implement specific forecasting models based on the same techniques for their licence areas.

4 Operational Forecasting Use Cases

A number of Use Cases have been considered to determine where improved operational forecasting might benefit ESO or DNO system operation and where data exchange between network companies might support this. These are described below.

Each Use Case includes a description, considers what information is needed by the network company (this could be a DNO or the ESO), and considers whether there are items of data that might be exchanged between network companies to improve the accuracy of forecasts.

4.1 Use Case 1 – Identifying Flexible Service Requirements at Distribution

Increasingly, DNOs are using flexible services to manage network constraints. In forecasting what types and volumes of services are likely to be required, it may be that the ESO can provide data that will improve forecasts.

Use Case	Identifying Flexible Service Requirements at Distribution
Description & potential benefits.	DNOs are seeking to accurately identify the requirements for flexible services to manage distribution network capacity limitations. With good information, DNOs may also be able to identify more cost effective options for flexibility services. If flexible services are used effectively, networks can be operated more securely and DER curtailment can be reduced.
Timeframe	2 weeks ahead up to day ahead.
DNO Forecasting – Data requirements.	• Local demand & local generation availability. • Weather conditions to assess likely wind and solar generation levels. Also to understand potential networks risks through high winds. • Distribution network conditions including planned outages and distribution network limits. • Potential for local flexibility services. What service providers are likely to be available?
Potential data exchange from ESO to DNO	Are wider transmission limitations likely to be active that might impact the local area? What are the predicted GSP interface powerflows? Is there potential for network throughflows? Are there DER assets locally that are likely to be used for balancing or ancillary services (e.g. STOR, Fast Reserve services)? Does the ESO have access to better information on likely weather conditions?

For this Use Case, a better understanding of ongoing transmission network conditions and potential impacts on the distribution network, and any DER assets that may be providing ESO services could enable DNOs to more accurately forecast flexible service requirements.

4.2 Use Case 2 - Increasing DG Output Capacity under Outage Conditions

Increasingly, Distributed Generation (DG) is being connected to distribution networks where network capacity limitations may constrain DG output at times. Often, new DG connections are dependent on Active Network Management (ANM) arrangements.

More accurate forecasts of local network conditions would allow network limits and limits to be better understand so that DG output can be maximised when constraints are active.

In addition, when there are communication losses to DG installations that are part of ANM arrangements, DG output will be reduced to zero, or limited, to ensure that the local network limits are not exceeded and the network remains secure. If accurate forecasts for local networks are available under these circumstances, DG output could be increased under these circumstances.

In forecasting local network conditions, it may be that the ESO can provide data that will improve the accuracy of DNO forecasts.

Use Case	**Increasing DG Output Capacity under Outage Conditions (and where there are Communication Losses)**
Description & potential benefits:	To better understand local network conditions so that distribution network outages and temporary capacity limitations can be better managed. This could enable increased outputs for local generation. DER curtailment might be minimised by taking preventive actions such as coordination with energy storage.
Timeframe	A few days ahead up to day ahead.
DNO Forecasting requirements	• Local demand. Local generation availability. • Weather conditions to assess wind and solar generation. • Local outages and distribution network limits including equipment ratings. • ANM failsafe actions could use forecast flows during communication loss to not instantly disconnect DERs. This will lead to advanced failsafe. • Identifying potential interactions between ESO services, DNO flexibility requirements and ANM.

Potential data exchange from ESO to DNO	Are wider transmission limitations likely to be active that might impact the local area? What are the predicted GSP interface powerflows? Are there DER assets locally that are likely to be used for balancing or ancillary services (e.g. STOR, Fast Reserve)? Does the ESO have access to better information on likely weather conditions?

Again, for this Use Case, a better understanding of transmission network conditions, and any DER assets that may be providing ESO services could enable DNOs to reduce DG output limitations.

4.3 Use Case 3 - Identifying and Managing Increased Network Risks

With increasing levels of distributed generation, whole system risks affecting the wider GB system are more difficult to identify and more complex to manage. It may be that the provision of data from DNOs to the ESO can improve visibility of distribution network conditions and resources and could enable the ESO to improve whole network security.

Use Case	Identifying and Managing Increased Network Risks
Description & potential benefits:	Wider network risks being managed by the ESO might be exacerbated by unusual local network conditions and/or unusual demand or generation conditions. By taking account of distribution network conditions, increased risks can be identified and addressed. Wider transmission system actions (e.g. operational tripping) can be better focussed.
Timeframe	Several months ahead to real-time.
ESO Forecasting requirements	• Demand and generation nationally. • Demand forecast by GSP. Local generation availability. • Weather conditions to assess wind and solar generation. • Use of OTS including generation available to tripped.
Potential data exchange from DNO to ESO	What generation is available locally, and what are potential distribution network restrictions? What DER assets are being utilised for distribution network services? What DER are not available for flexibility services?

4.4 Use Case 4 - Improved GB Energy Balancing

With more complex local networks, improved information from DNOs to the ESO on DNO networks and assets could enable better GB energy balancing.

Use Case	Identifying and Managing Increased Network Risks
Description & potential benefits:	To improve national energy balancing through more accurate data on local network conditions. By taking account of distribution network forecasts and conditions, a more accurate national position can be developed.
Timeframe	A few days ahead up to day ahead.
ESO Forecasting requirements	• Demand and generation nationally. • Local generation availability. • Weather conditions to assess wind and solar generation.
Potential data exchange from DNO to ESO	Local generation availability and potential restrictions due to network outages. Possible use of local DER assets for distribution network services.

5 Forecasting Methodologies, Tools and Techniques

This section explains the algorithms used for operational forecasting. The following basic questions are answer to capture key characteristics of the forecasting engines:

1. What are the forecasting methodologies/algorithms being used or developed by network companies?
2. Can we draw any initial conclusions on preferred methodologies?
3. How far into the future?
4. Is it useful to look back?
5. How far back?
6. Is 5-minutely overkill?
7. Do you want a probabilistic forecast?
8. If so, how do you want probabilities expressed?
9. How would you convince yourself that a new forecast system is useful?

5.1 UKPN response

Operational Forecasting; from 48hours ahead to 4 weeks ahead. Medium/Long Term Forecasting; 6month/1year ahead. Looking back helps us to determine trends and identify event occurrence/magnitude.

We regularly look back at historical data patterns in electricity systems. We can look ≥10 years back.

This is not an overkill. Most of the power system studies we perform are at steady state which requires the pattern of load and generation to be modelled at 1/2hourly resolution. However, having higher resolutions e.g 5-minutely will increase the accuracy of the simulations and results. Power systems are becoming more complex and smart, having higher resolution data will help better analyse the behaviour of networks and optimise/effective system operations.

Probabilistic forecasts can outperform more classic forecasting algorithms because they exploit the precise structure of the probabilities estimated for future demand/generation. Probabilistic forecasts could provide additional quantitative information on the uncertainty associated with demand/generation. Modern computing systems can handle complex calculations for probabilistic forecasts. In addition probabilistic forecasts can be combined with machine learning algorithms to help train their algorithms in line with changes in weather/patterns of demand/generation usage. As the pattern of demand/generation becomes more intermittent, the use of probabilistic forecasting becomes more applicable and accurate. Probabilistic forecasting can determine those unexpected patterns/events which can impact investment decisions.

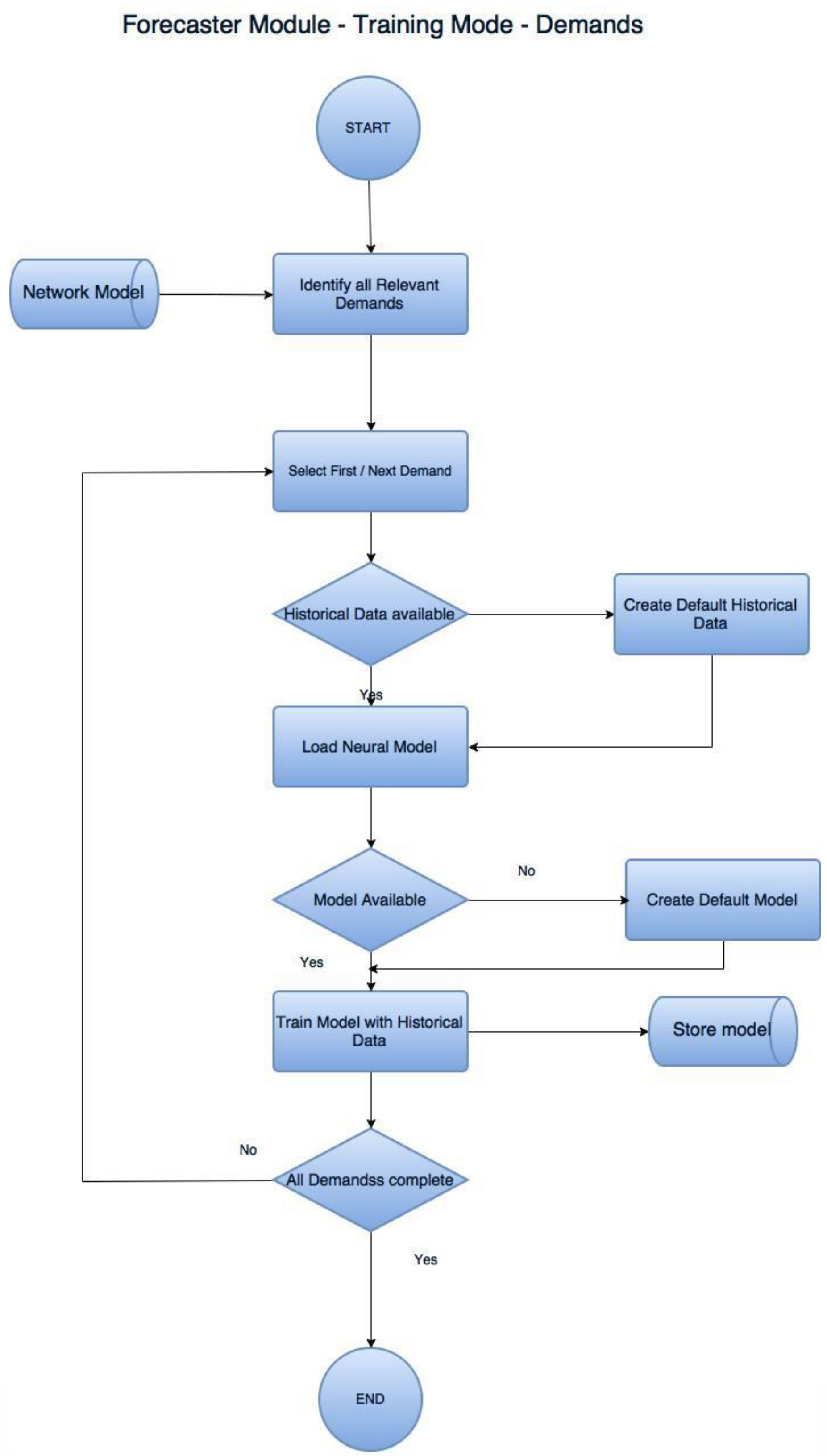
The most likely value of demand and generation every 5-minutely. As well as the 2nd-5th most likely values can also be presented. The approach for algorithm implementation can be based on best practice – further investigation is needed. What is the best probabilistic algorithm approach for distribution networks.

Having an accurate forecasting system will help our deterministic network management tools operate automatically and efficiently.

5.1.1 UKPN's Artificial Neural Network – Machine Learning

The DERMS forecaster module in Power Potential project design for operational time frames (48 hours ahead), relies on a machine-learning algorithm using variables such as historical power measurements, numerical weather data, SCADA data, calendric information etc, to produce a fully automated predictive model.

The automatic training of the forecast model extrapolates the correlation of the DER production with the relevant variables and will then use this correlation in the future to predict the renewable production. The Demands forecasting process is illustrated below.


Forecaster Module - Training Mode - Demands
START
Network Model
Identify all Relevant Demands
Select First / Next Demand
Historical Data available
Create Default Historical Data
Yes
Load Neural Model
No
Model Available
Create Default Model
Yes
Train Model with Historical Data
Store model
No
All Demandss complete
Yes
END

A forecast model is assigned to each specific Demand according to type. During the training process, historical data is loaded from database for a given Demand and the forecast model for that specific Demand is trained with that data.

The Demand forecast presents seasonality patterns on three different levels: yearly, weekly and daily that are weighted with calendric numerical information and are calculated during the training process. In addition, there are weather related influences that may show up as relevant during the training process.

The most appropriate variables for this load forecast are reported below:

Calendric variables:

- Month
- Date (Day of Month)
- Daytype (Monday, Tuesday,…, Sunday)
- Time of Day
- Holiday calendar: public holiday (with bridge day, after/before holiday) and school holiday

Weather numerical variables:

- Temperature: impact on power heating and conditioning
- Humidity: impact on perceived temperature

Forecasts are carried out on a 30mins interval. The process is the same for forecasting both DER Production and Demands. The forecaster module relies on a machine-learning algorithm using variables such as historical power measurements, numerical weather data, SCADA data, calendric information and any other numerical variable, to produce a fully automated prediction of the forecast value based on the inputs provided.

5.2 ENW response

Looking back should take into account the evolution of domestic and non-domestic loads through time. Therefore, it’s risky to consider horizons over 4-5 years.

Temporal resolution / is 5-min an overkill: At high level half-hourly resolution is useful for network planning, whereas going down to 5-min or even 1-min and lower makes sense for operational planning to fine tune the control logic of any controller (eg, passive control for OLTCs traditionally and more advanced in ANM era). Higher resolution is expected to bring higher errors and given it’s focused on operations rather than procurement of flexible services it makes more sense if it focuses on up to 1 week horizon (fair to say that even 48 hrs is too much).

Probabilistic modelling: it can provide more information around uncertainties, but especially when it comes to using machine learning techniques we need to highlight the accuracy aspects in accordance with the time horizon. For example, an ARMA technique can bring significant accuracy in short term (day ahead/week ahead forecasts), but it’s computational heavy with significant errors if it goes for several months horizons.

How we'd like probabilities to be expressed: apart from the most likely values, the most extreme values with associated probabilities are useful. For example if the most likely value is 10MW with an associated probability of 50% and a 15MW value has a 10% probability, then the latter value could not be negligible in some planning/operational processes.

Use of normal distribution/gaussians etc: these are mathematical/statistical tools that mean nothing outside an engineering context. So, a forecasting methodology should clearly show how these statistical tools are used to provide a meaningful forecast with a clear engineering use. For example, if the analysis of historical demand has shown that there is a normal distribution in MW values at an X-date/time then this can be used in the model to assign higher/lower probabilities around MW values closer/further away from the peak of the curve, respectively.

Final question: "how would you convince yourself that a new forecast system is useful?" – I agree that accuracy is a top priority. But for operational purposes, ie the short term forecast of 24 hrs to 1 week ahead, accuracy should apply not only on the most likely values but on the extremes. For example, if it's critical to know the max/min values that have at least 5 or 10% probability to occur, then the forecast needs to show that in a large number of forecasts the % of values that are away from the most likely prediction are showing a similar % with the probability in the model.

6 Conclusions and Proposals for Data Exchange/Next Steps

This report provided a detailed view of key operational forecasting methodologies implemented by DNOs and NGESO. There are similarities in the algorithms adopted for operational forecasting between network/system operators with Artificial Neural Networks (ANNs) classified as the most accurate, reliable and efficient solution for providing demand and generation forecasts across transmission and distribution levels. The two methods of Extreme Gradient Boosting (XGBoost) and Long Short Term Memory (LSTM) with latter performing better with lower errors in forecasted output results.

This section would also recommend the following data transfers between DNOs and ESO:

- GSP forecast data to be forecasted by NGESO and shared with DNOs. These also include:
    - National Demand and Gen forecast.
    - Interconnector Flows.
- Aggregated forecasted demand and generation at DNO GSP level to be shared with NGESO per technology types; DER= Wind, PV, Storage.
    - Storage Forecast driven by commercial markets to be further investigated.
- DNO Grid Outage Plans to be exchanged with NGESO, these are those outages which result in demand transfer on the DNO. It is important to note that outage plans are already being exchanged between DNOs and the ESO. However, to facilitate the benefit of the proposed use cases, greater emphasis is given to highlighting the impact to DG availability and quantifying the likely volume of DG due to network outages.

- 132kV interconnector flows to be shared with NGESO from DNO. This depends on having a functional loadflow engine and network model to intake the forecasted demand/generations to represent flows on 132kV branches using the numerical model's network characteristics data.
- Data on ESO services indicating which assets are contracted with ESO to provide services. Data from the ESO (across all relevant time scales) from the outcomes of their Ancillary Services/Flexibility markets/auctions would be very useful, where those services are procured from resources that sit at DNO level. This would allow DNO's Operational Forecasting tools to assess the likely network flow impacts of these resources either (i) being held in reserve for, or (ii) acting in response to, an ESO trigger for the relevant service. The kind of data required might be e.g. network location on the DNO system, amount in MW/MWh/MVAr of the resource, and the technical details of the service that is procured.

Further work is required to develop the use cases outlined in section 4 to understand how the proposed data transfers can be achieved (e.g. in what format) across planning and operational timescales. The impact on business processes, resources and IT systems needs to be assessed before implementation.

7 Data Transparency Aspects

This report has reviewed the existing and developing operational forecasting capability within network companies and has identified potential opportunities for data exchange between network companies to improve operational forecasting capability and accuracy.

With respect to data transparency, some of the forecasts produced by the ESO are already made available to stakeholders. Whilst WS1B P3 focused on defining the signals real-time and operational forecasting data exchange which would benefit whole system coordination and control between ESO and DNO/DSO, DNOs have been considering and progressing the feasibility of sharing information with other industry stakeholders such as Embedded Capacity Register and other operational data [1-3]. As DNO's further develop their operational forecasting capability, making these forecasts more widely available to stakeholders will be considered.

8 References

[1] https://www.westernpower.co.uk/live-data-feed-application-map

[2] https://www.current-news.co.uk/news/wpd-releases-real-time-power-flow-data-to-aid-operational-decisions-as-it-increases-digitalisation

[3] https://www.ukpowernetworks.co.uk/electricity/distribution-energy-resources/the-embedded-capacity-register

Energy Networks Association
4 More London Riverside
London SE1 2AU
t. +44 (0)20 7706 5100

w. energynetworks.org
@EnergyNetworks

Energy Networks Association Limited is a company registered in England & Wales No. 04832301
Registered office: 4 More London Riverside, London, SE1 2AU