transition
Moving to a smart future

Real-time Forecast Optimisation Report

Authors: Sia Partners in partnership with Electralink and NERDA

Scottish & Southern Electricity Networks

Scottish & Southern Electricity Networks SIAPARTNERS NERDA Near Real-time Data Access ElectraLink

Contents

Table of Figures

Executive summary

The Ofgem NIC-funded TRANSITION[1] project aims at further developing DNO experience and capability with deploying local flexibility markets as part of the DSO function, enabling non-DSO services such as peer-peer capacity trading as well as enlightening broader system coordination efforts. The TRANSITION project trials are being undertaken in the Oxfordshire region of SSEN SEPD DNO license area, as part of a collaboration with the Local Energy Oxfordshire (LEO) industrial project[2].

During the first phase of the project, Sia Partners implemented a Load Forecasting solution for the purpose of predicting the expected load at each substation, feeder and generator included in the TRANSITION Project. The models developed in this initial phase relied on historical weather and network load data and therefore represented a behaviour of the asset in this particular point in time. With COVID-19 lockdowns and network topology evolutions, the way electricity is consumed, where and when, can change regularly. The expectation that the period on which the models were calibrated is a representative period for future forecast is no longer a valid assumption.

Therefore, Phase 2 of this project has focused on the use of real-time data to further refine the load forecasts made at each asset within the TRANSITION scope. This report, produced in collaboration between SSEN and Sia Partners presents how Sia Partners' Operational Load Forecasting solution has been further enhanced for the TRANSITION project, and summarises the main outputs and key lessons learnt from this second phase of the innovation project.

NERDA[3] is an SSEN innovation pilot project focusing on open data and creating system/network data access portals for wider energy system stakeholders. It has been identified as the key source for network load data measurements for transformers and feeders. Electralink[4] acts as the responsible party for the settlement of all generators contributing to the BMU and has been used to capture real outputs from identified generators. Both solution were connected to the Load Forecasting Solution via API to retrieve most recent real-time DSO system data on an ongoing basis, what can further improve the forecasting process.

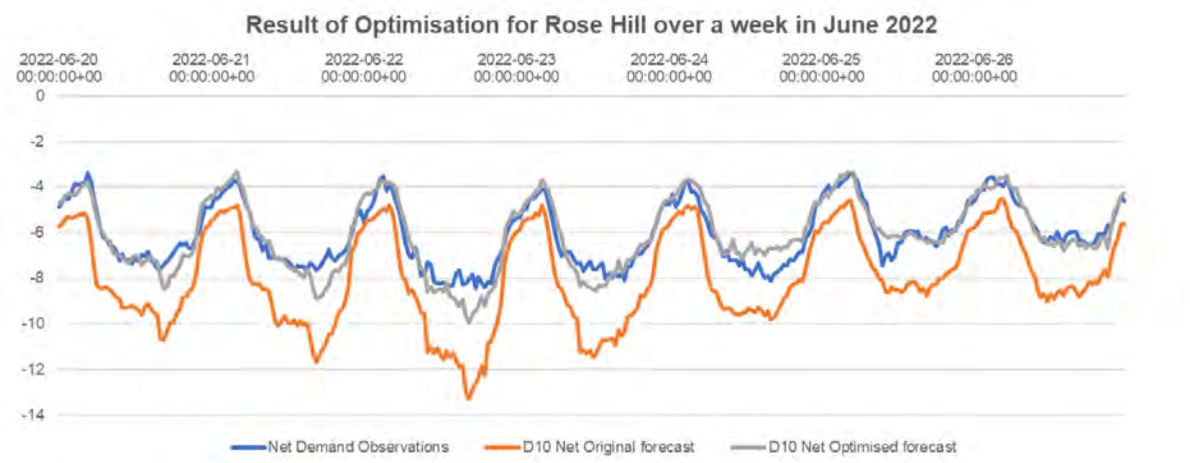
The real observations of net demand and generation output were used to optimise the original load forecast provided by the calibrated models. It focused on reducing the error between the initial forecast and the observations over the recent days and weeks to reallocate this error on future forecast horizon. The methodology applied both volume and shape correction sequentially to provide an 'optimised forecast' at each point of the network in scope.

The results saw an impressive 32% reduction of the error (MAPE*) on average across all EHV primary substations and 43% improvement at HV feeder level with specific models.

Below is the table summarising the performance of the forecasts before and after optimisation over the period of August 2021 to November 2022. Furthermore, Rose Hill Primary has been used throughout the report as the Primary substation example to demonstrate the results captured across all substations and feeders. Below is the view of both original and optimised forecast over a week in June 2022.

[1] SSEN Transition (ssen-transition.com)
[2] Home - Project LEO (project-leo.co.uk)
[3] NIA SSEN 0050: Near Real-time Data Access (NeRDA) | SSEN Innovation (ssen-innovation.co.uk)
[4] Home - ElectraLink

	Original Forecast			Optimised Forecast		
HV Group (Forecast D10)	MAPE	MAPE*	RMSE	MAPE	MAPE*	RMSE
Arncott	31%	31%	1.133	21.4%	21.4%	0.628
Berinsfield All Feeders individually	21%	21%	3.398	13.1%	13.1%	1.656
Bicester	14%	14%	4.452	7.5%	7.5%	0.953
Bicester North Primary	9%	9%	1.238	7.5%	7.5%	0.998
Deddington All Feeders individually	10%	10%	0.064	9.7%	9.7%	0.06
Eynsham	10%	10%	0.691	7.1%	7.1%	0.317
Kennington	19%	19%	0.207	12.0%	12.0%	0.095
Milton	32%	32%	32.311	21.3%	21.3%	19.731
Oxford Primary	15%	15%	3.324	6.3%	6.3%	0.693
Rose Hill	20%	20%	2.857	7.7%	7.7%	0.579
University Parks	6%	6%	0.92	4.7%	4.7%	0.511
Yarnton Primary	9%	9%	1.335	7.9%	7.9%	1.233



The introduction of real-time data allows to understand the true accuracy of the forecasts, as opposed to the quality of the models initially calibrated. It demonstrated that the introduction of real-time data was essential to capture short-term and longer term variations of consumptions, both in magnitude and in shape. These new forecasts have been made available using the existing automatic API solution to the Power System Analysis tools in use for the ongoing Technical Trials of the TRANSITION project, with very limited adaptation needed. They will serve as the new basis for capturing flexibility requirements during the final phases of the trials. The final outputs of the TRANSITION project, including the impacts of these new forecasts, will be shared with the ENA and more broadly with the Industry as the project comes to a close.

Access to real-time data also allowed to capture the relatively low impact of the horizon on the forecast. It demonstrated that the week ahead demand forecast was as accurate as the day ahead forecast for the primary substations considered. The error on the horizon was solely borne by the error of generation forecasts. This could be demonstrated at the scope of the TRANSITION project and would require further investigation when scaling the solution to a network-wide solution.

Further learnings were captured on the ability to scale this solution to an entire network. The solution design, data models and end-to-end data processes supporting the production of optimised load forecasts would need to be adapted to replicate the capabilities at scale.

Finally, the forecasting solution front-end/visual User Interface has also been reviewed in this Phase 2 work to display real time forecast data, to facilitate learnings from the project and will support future discussions for transition to BAU.

1 Introduction

1.1 Project background

TRANSITION is an Ofgem Electricity Network Innovation Competition (NIC) funded project, led by SSEN in conjunction with project partners ENWL, CGI, Origami and Atkins. The TRANSITION project works in conjunction with Project LEO within SSEN, and also collaborates very closely with the other two Ofgem NIC funded projects Electricity Flexibility and Forecasting System (EFFS) from WPD, and FUSION from SPEN. Together, these three projects form the T.E.F. collaborative forum, with the overall aim of coordinating innovation requirements, sharing key learnings in the DSO space, and broadening the application of this knowledge to trials and testbeds in a variety of UK DNO regional settings and better inform the wider ENA Open Networks Project activity.

The TRANSITION project is now coming to the end of its Trial Period phase where the market dynamics associated with contracted local flexibility, as well as enabling non-DSO services such as peer-peer capacity trading and coordinating the whole system view with the ESO services are being tested to provide evidence-based learnings to the industry.

Currently the flexibility requirements are based on the Load Forecasting solution developed in the first phase of the project. The Load Forecasting Solution, developed by Sia Partners working with the TRANSITION team, has been detailed in a previous report available on the TRANSITION website[5]. The solution aimed to develop Demand and Generation models for all asset across the Trial geographical area. These models were calibrated offline based on **historical** load and weather data over 4 years, between 2017 and 2020. They represent the relationship between historical measurements (load or generation) and the weather at that point in time. On a day-by-day basis, these models are then applied and evaluated with the present real time weather forecast to determine expected demand and generation across the area over the coming hours and days ahead of real-time.

The phase 1 work of this project identified that one of the key learnings was the potential drawbacks to relying on models that are based on historical data alone. For example, since 2020 the COVID-19 pandemic may have changed significantly how electricity is consumed at some substations and feeders on the Transition scope. The lockdown due to COVID-19 pandemic saw a drop of consumption of around 30% across the country with different levels of drops locally. The lifting of lockdown restrictions brought new ways of working and living. Therefore, relying on 4-5 years old consumption patterns to determine the next few days' flexibility requirements might not be optimal. Also, the DNO network topology is always evolving based on reconfigurations and new customer additions, hence, demand patterns through substations may occasionally abruptly change.

The new phase of work for the TRANSITION project's forecasting solution described in this report looks to understand how accurate the original historical data-based forecasts are and furthermore, how they can be refined using connectivity to real time data sources to ensure that DSO flexibility requirements are accurate.

[5] https://ssen-transition.com/wp-content/uploads/2021/11/TRANSITION-Load-Forecasting-Dissemination-Report-Final-V3.pdf

1.2 Purpose of the document

The purpose of this document is to provide a comprehensive and detailed summary of Sia Partners' phase 2 evolution of the original Load Forecasting solution implemented in the TRANSITION Programme.

It aims to:

- provide a greater understanding of the setup of real-time and automated data pipelines from other sources and of the optimisation forecast methodology used to adapt to them,
- provide clarity around the improvements the forecast optimisation has brought to the flexibility requirements and how they can be further applied ,
- and shed light on the various types of implementation challenges encountered during this new development phase.

It will also provide focus on a number of studies aiming to inform on the broader contextual value of forecasts, a future upscaling strategy and potential application of TRANSITION's results (which naturally focused on a small locational footprint), to a much wider regional basis potentially including entire DNO license area in future.

2 Sia Partners updated forecasting process

2.1 Geographical scope

The geographical scope of the optimisation forecast is the same part of the SSEN Oxfordshire network as where Load forecasting capabilities were developed in the first Trial Phases. Future load on the network is expected at a set of Bulk Supply Points (BSP), Primary substations, their associated feeders and all Generation assets contributing to the reduction of Demand drawn from the Transmission Network.

It is in line with the geographical scope of the TRANSITION Programme. The BSPs and associated Primaries in scope are listed below:

GSP Code	GSP	BSP Code	BSP	Prim Code	Primary
COWL	Cowley	OXFO	Oxford	OXFO	Oxford Primary
COWL	Cowley	OXFO	Oxford	UNIP	University Parks
COWL	Cowley	DRAY	Drayton	MILT	Milton
COWL	Cowley	HEAD	Headington	ARNC	Arncott
COWL	Cowley	COLO	Cowley Local	BERI	Berinsfield
COWL	Cowley	COLO	Cowley Local	ROSH	Rose Hill
COWL	Cowley	COLO	Cowley Local	KENN	Kennington
COWL	Cowley	COLO	Cowley Local	WALL	Wallingford
COWL	Cowley	YARN	Yarnton	YARN	Yarton Primary
COWL	Cowley	YARN	Yarnton	DEDD	Deddington
COWL	Cowley	YARN	Yarnton	EYNS	Eynsham
COWL	Cowley		Witney	NA	NA
ECLA	East Claydon	BICN	Bicester North	BICN	Bicester North Primary
ECLA	East Claydon	BICN	Bicester North	BICE	Bicester

Figure 1 - List of Primaries and BSPs in scope of TRANSITION

NB: The 13 Primary substations are part of the trial area in Oxfordshire, but do not cover the entire scope of the BSPs.

2.2 Overview of the updated forecasting process

The original forecasting process has been developed in the first trials phases. The models developed relies on static historical data and were evaluated with weather forecast data. The diagram below summarises the various forecasting stages

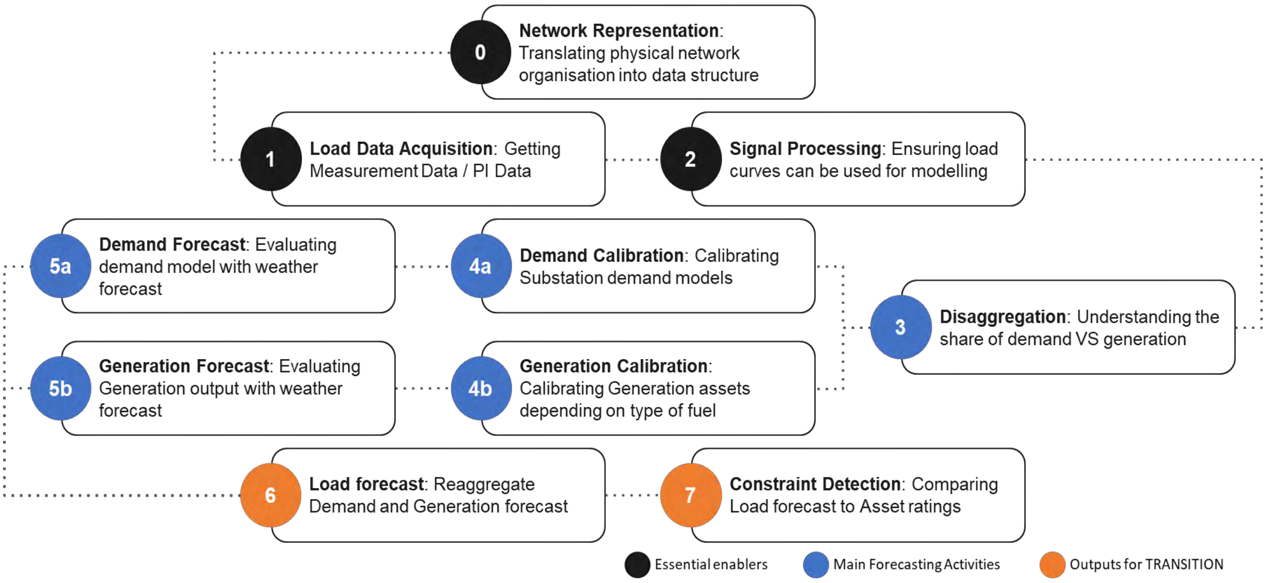


Figure 2 - High level Original Forecasting Process

The revised forecasting process includes the integration of 2 new real-time data sources (see dedicated sections for more details):

Nerda: SSEN's Real-time system based on PI data

Electralink: Independent body responsible for the management of settlement data

The addition of the real-time realised data will provide an understanding of the accuracy of the models on an ongoing basis. Furthermore, instead of being solely based on historical load behaviour, it also allows the opportunity to identify and capture more recent trends and rapid changes of consumption.

For instance, demand models were calibrated over a period that span between 2017 to end of February 2020 to avoid the disruption brought by the lockdown over the COVID-19 pandemic. Since this event, the ways of living and consuming electricity might have changed drastically at local level. The DNO network in the Oxfordshire area has also been reconfigured with feeder transfers etc over the years too. Therefore, the understanding of the most recent measured load at each asset on the network can be used to optimise future forecasts on this asset.

The general profile of weather-dependent renewable generators is not impacted by such sudden changes, and it is assumed that the models calibrated on historical data are still valid. The objective is therefore to capture changes in demand patterns, taking into account the real outputs of the generation and net flows measured at substations and feeders.

The diagram below explains how these 2 new streams of information will be included in the original forecasting process

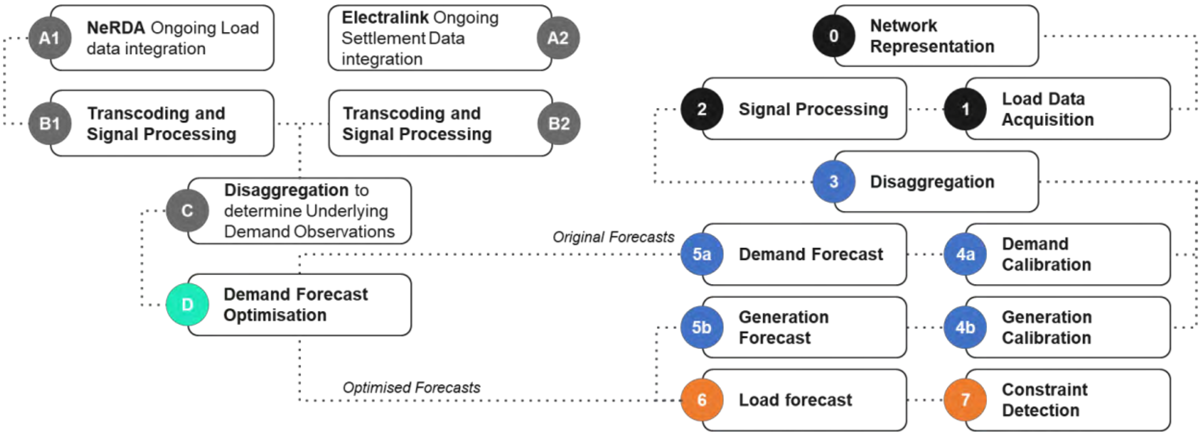


Figure 3 - High level Forecasting Process with Demand Optimisation

Nerda Pipeline:

A1: NeRDA Ongoing Load data integration: Establish a connection between the Load Forecasting Solution and SSEN's load platform NeRDA. For each SSEN asset on the network (transformer and feeders), retrieve the latest data points latest data points and upload the raw values in the Load Forecasting Solution

B1:NeRDA Transcoding and Signal Processing: Map the NeRDA data model to the Load Forecasting Data model, in order to populate the existing tables. This includes mapping the representation of measurements in NeRDA to the assets in the Load forecasting solution, as well as transforming the signal into half-hourly values. Finally for each raw load curve a set of filters is applied. The objective is to clean the various signals and receive a representative ongoing net load curve for the signal measured.

Electralink Pipeline:

A2: Electralink Ongoing settlement data integration: Establish a connection between the Load Forecasting Solution and Electralink settlement solution. For each generator in scope for which settlement is provided, retrieve the latest data points latest data points and upload the raw values in the Load Forecasting Solution.

B2: Electralink Transcoding and Signal Processing: Map the Electralink data model to the Load Forecasting Data model, in order to populate the existing tables. This includes mapping the representation of settlements in Electralink to the assets in the Load forecasting solution. Finally for each raw load curve a set of filters is applied. The objective is to clean the various signals and receive a representative ongoing net load curve for the signal measured.

C: Disaggregation: Similarly to the historical branch to fit the models, this crucial activity allows to identify, within a representative net load curve, the share of demand and the share of generation. The 'pure generation' signal is composed of the sum of outputs of each generator, coming for the

Electralink branch or estimated by generic modelling. The output is the underlying demand signal representing the total demand connected to the asset

D: Demand Forecast Optimisation: By comparing the forecasted demand and the 'measured' underlying demand, we can determine the error on the past recent days and understand how future forecasts need to be corrected. The optimisation of the demand forecast takes specific parameters which are defined individually for each asset of the network and used when a new forecast is produced.

2.3 Reminder of the weather forecast data sources used in TRANSITION for model evaluation

For each asset on the network, pure demand and pure generation models have been calibrated based on the historical underlying demand and generation load curves, obtained at the disaggregation stage. These individual models are used to determine the expected underlying demand and generation outputs in the future.

These models are function of explanatory variables, namely calendar and weather variables. The models are evaluated with the forecast of these explanatory variables to determine the expected demand and generation forecast.

The Load Forecasting solution relies on 2 sources of weather data:

- **ICON-EU-EPS (called D4 in this report):** through the collection of 40 individual members, it provides the probabilistic view of the weather forecast over the next 120 hours. This brings 40 individual weather forecasts.
- **MOSMIX (called D10 in this report)** which provides a single deterministic view of the weather forecast over the next 240 hours. This forecast is seen as scenario #41.

These weather forecasts are made available by DWD and captured, processed, and cleaned by Sia Partners' proprietary solution, before being made available to the TRANSITION Load Forecasting solution.

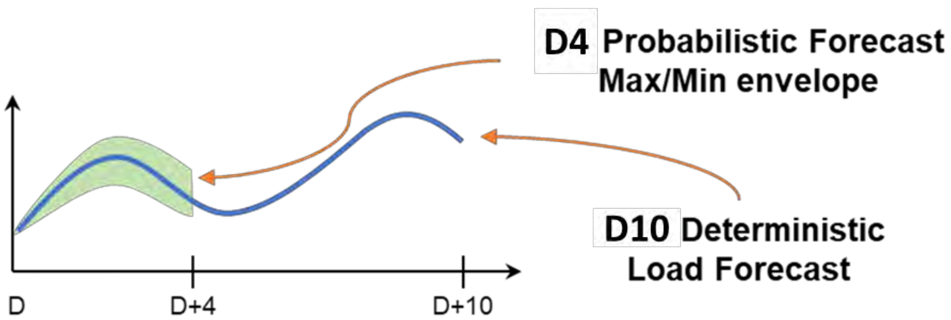


Figure 4 - Illustration of the weather data sources used for load forecasting

3 Nerda interface & pipeline

3.1 Description and objectives

The SSEN Near Real-time Data Access (NeRDA) [6]Projects aims to examine how near real-time data can be best used by stakeholders.

With an increasing number of consumers forecast to shift towards low carbon technologies, such as electric vehicles and heat pumps, the energy system is becoming smarter to accommodate this uptake. The near real-time data identifies areas where flexibility could be used or where network reinforcement will be needed.

Flexibility enables the movement of timing and location of both the consumption and generation of energy; with more technologies participating in flexibility, having an up-to-date view of the energy system will be key in ensuring an efficient and cost. SSEN have been working with Open Grid Solutions to deliver the NeRDA project over the last two years and it is anticipated that it will be transferred to business as usual.

It includes a visual user interface portal, but also an API solution to allow bulk download and consumption of real time system data.

Each analogue on the network is identified by a measurement_id in Nerda. A data point is created every 3 minutes and based on the values in PI Historian. The Load Forecasting solution aims to use the measurements at transformers and feeders to develop the ongoing net demand at those assets on the network.

For the purpose of the TRANSITION project, a number of dedicated Nerda APIs have been developed

transition_static: the endpoint returns all measurement_ids in TRANSITION project

nerda_from and *nerda_between*: For a given measurement id the endpoints return the values of the measurement_id over the period selected, or from the date specified.

3.2 Connexion to Nerda

Members of the TRANSITION Project have been granted access to the Nerda User Interface as well as the API Catalogue on a restricted scope. Access to the interface allows to run a number of consistency tests between automatic querying from the Solution and visual results from the User Interface.

In order to query on a regular basis the data from Nerda, authorisation levels have been provided with dedicated secure API Key. Moreover, the Load Forecasting Solution IP address has been fixed in order to monitor connexion

3.3 Data mapping

With different purposes, Nerda and the Load Forecasting Solution have been designed with 2 distinct Data models. In order to collect the data from Nerda on an ongoing basis, the first objective was to ensure we could align the data models and reconcile the ongoing sources of data.

[6] NIA SSEN 0050: Near Real-time Data Access (NeRDA) | SSEN Innovation (ssen-innovation.co.uk)

In Nerda, each analogue on the physical network is represented by an individual measurement_id. In the Load Forecasting solution, each asset on the physical network is represented individually, with a set of load curves.

Each measurement_id is linked to a Tx (Transformer) or a line (Feeder) which represents where the analogue is connected on the network. The following information are available to locate the measurement_id on the network. The tables below describe the fields available and an example for Rose Hill T2 and Rose Hill E5L5 respectively.

Field in nerda_static	Description	Example for Rose Hill T2
tx_name	Name of the Transformer	2
nerda_tx_uuid	Unique Nerda transformer identifier	_f0a65ba8-8da3-4b8d-80e2-c7f164e99e15
nerda_measurement_id	measurement_id	_226ff537-b1c6-4c29-b04d-f534ae1e904b
name	Name of the measurement ID	ROSE HILL E2T0 AMPS
pi_tag	PI Historian Identifier	ROSH~E2T0~AMPS~AI

Field in nerda_static	Description	Example for Rose Hill E5L5
line_name	Name of the Feeder	E5L5
nerda_line_uuid	Unique Nerda feeder identifier	_5213dd1c-ccf9-498e-82d0-02cfcacc4864
nerda_measurement_id	measurement_id	_0ebc7cbf-ca29-4aab-8a30-08eb6feab1d0
name_vals	Name of the measurement ID	ROSE HILL E5L5 Amps
pi_tag	PI Historian Identifier	ROSH~E5L5~AMPS~AI

Also, each measurement_id has some attribute which provides some context to the value recorded. The tables below describe the fields available and an example for Rose Hill T2 and Rose Hill E5L5 respectively

Field in nerda_static	Description	Example for Rose Hill T2
measurementType	Type of analogue recorded	LineCurrent
unitSymbol	Unit of the value	A
unitMultiplier	Multiplier	none

Field in nerda_static	Description	Example for Rose Hill E5L5
measurementType	Type of analogue recorded	LineCurrent
unitSymbol	Unit of the value	A
unitMultiplier	Multiplier	none

The identification of the lines and transformers can be matched to the transformers and feeders in the Load Forecasting solution. The identification of the unit of the measurement_id allowed to map with the appropriate load curve (MW, MVAr, Amps, kV)

This results in the following mapping table between Nerda and the Load Forecasting Solution. The example below is for Bicester Primary substation transformers

object_id [PK] integer	object_type text	measurement_id [PK] text	multiplier double precision	measurement_unit text	polarity integer
2101	transformer	_e1b7d238-6390-49dd-a2c5-71232ac5fbe0	1	amps	1
2101	transformer	_ae005978-974d-46c9-9899-ed407c24c88f	1000	kv	1
2101	transformer	_4e33b867-3a94-4323-9187-8c410da608c3	1000000	mvar	1
2101	transformer	_d4c42232-60cd-4fc6-a10e-df33f339e761	1000000	mw	1
2102	transformer	_94a9cc7e-f482-4526-b01a-51d31665ab96	1	amps	1
2102	transformer	_4564facc-3961-4870-ad9f-55e65fa73b22	1000	kv	1
2102	transformer	_8a630f02-5665-48ff-ac70-5286ff6c439d	1000000	mvar	1
2102	transformer	_b7c16e7e-57c8-430f-b2d2-8eb218fc9200	1000000	mw	1
2103	transformer	_42b23d7f-884c-43a5-bce8-30ecfd5043cf	1	amps	1
2103	transformer	_6d101c7c-cc51-484c-8e22-d80fa2eb4308	1000	kv	1
2103	transformer	_7c071ea0-0816-4600-b216-07c912c7b64a	1000000	mvar	1
2103	transformer	_0df95719-9eee-4efe-8735-651861994791	1000000	mw	1

Figure 5 - Nerda Mapping Table for Bicester Primary substation Transformers

3.4 Data pipeline

The diagram below provides an overview of the ongoing NeRDA pipeline that has been implemented to capture transformers and feeders net flows.

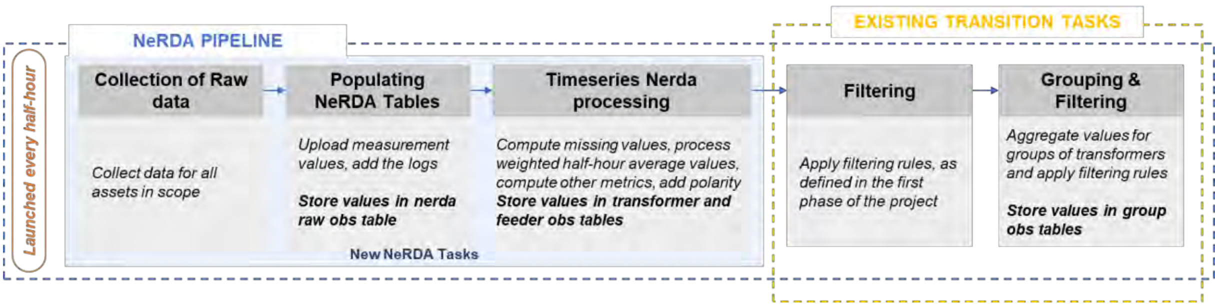


Figure 6 - NeRDA Pipeline

3.4.1 Collection of raw data

Two endpoints have been made available to the TRANSITION project.

Nerda_after: For a given measurement_id, all values captured in Nerda after the timestamp specified will be returned

Nerda_between: For a given measurement_id, all values captured in Nerda between the 2 timestamps specified will be returned

On an ongoing basis, Nerda is scheduled to be queried every half-hour to capture the most recent measured load data on the network. For each measurement_id recorded and mapped to the Load Forecasting Solution data model, the solution queries the nerda_after endpoint, using the latest value recorded and uploaded in the Load forecasting solution as a starting point. All values returned by Nerda are uploaded.

The original development of the Load Forecasting solution relied on PI network data until 31/12/2020. In order to have a continuous set of data points, the nerda_between endpoint has been queried for all available measurement_ids between 01/01/2021 and mid 2022 when the pipeline was operational.

For each query and each measurement_id the response from the API is also captured in a logs table. This allows to understand quickly the reasons for potential errors or missing data. All data points returned by Nerda are captured in a dedicated table, acting as a raw data table, before being processed to fit the Load Forecasting data model.

For any given measurement_id, each timeseries returned by Nerda is processed to ensure it is fit for the Load Forecasting solution. It consists of 2 key stages

Creating a half-hourly timeseries

Affecting the timeseries to the correct analogue (asset x unit measured)

3.4.2 Timeseries processing of raw data

Nerda is an application still under development and the first version went live in February 2022. The timeseries held were therefore of different granularity.

For the historical data, prior to June 2021, information stored in PI Historian was available. To limit the amount of data being recorded, SSEN PI Historian is set up to record data by exception, according to some pre-determined *jitter factor*. In other words, PI works based on dead-bands of how big/small a change in the data needs to be for it to get recorded in the system. Hence, there is no minimum time resolution for the data. In order to transform the signal to a half-hourly signal, a **'forward-fill' methodology** was applied. A value is valid until a new one replaces it.

For ongoing data, since February 2022, Nerda queries PI Historian on a 3 minute basis and retrieve the instantaneous value for the specific measurement_id. This means that approximately 10 values are recorded for a single half-hour timestamp. Depending on communication reliability, some calls to PI Historian could be missed. Therefore, a **'time-weighted average methodology'** was implemented, coupled to the forward fill.

The diagram below details how the methodology has been implemented

Data could go back at least 5-10 years for those substations with old enough communications systems. Even though this would have to be supplemented by accurate knowledge of any network reconfiguration that may have impacted the measurements

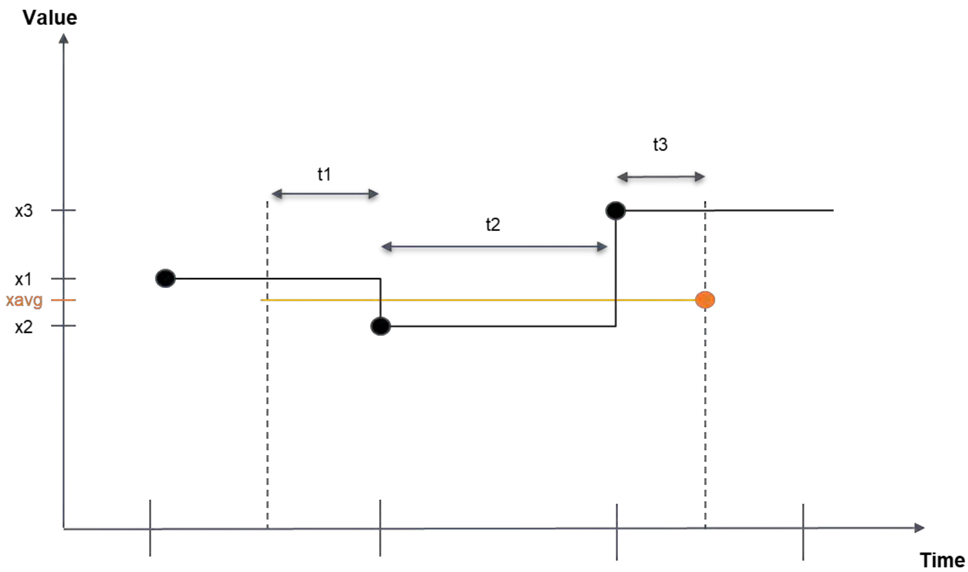


Figure 7 - Forward-fill and time-weighted methodologies for half-hourly signal

3.4.3 Integration in LFS data model

For each measurement_id, we use the mapping table developed to determine whether a specific measurement is expected for a given asset in the Load Forecasting table. Not all transformers and feeders have a full suite of measurements (Amps, MW, MVAr, kV).

For each asset in scope of the Load forecasting solution, data points from the expected measurement_ids are collected and uploaded in their corresponding unit to create a single timestamp based record.

These records are uploaded in the respective feeder_obs and transformer_obs tables

3.4.4 Load forecasting data processing – use of existing TRANSITION tasks

The Load Forecasting solution already holds a number of processing rules which have been put in place in the first developments. Those rules have been developed over historical datasets and aimed to

1. Filter signals at transformer or feeder level to ensure that the resulting signal observed at the point of forecast is of good quality. Transform signals to align units for MW and MVA signals
2. Aggregate load flows at group level to have a representative point of forecast. Filter the resulting signal with dedicated rules
3. Disaggregate any generation impacting the net demand at group or feeder level
4. Based on filter results from net demand and generation, determine the group / feeder filtered underlying demand.

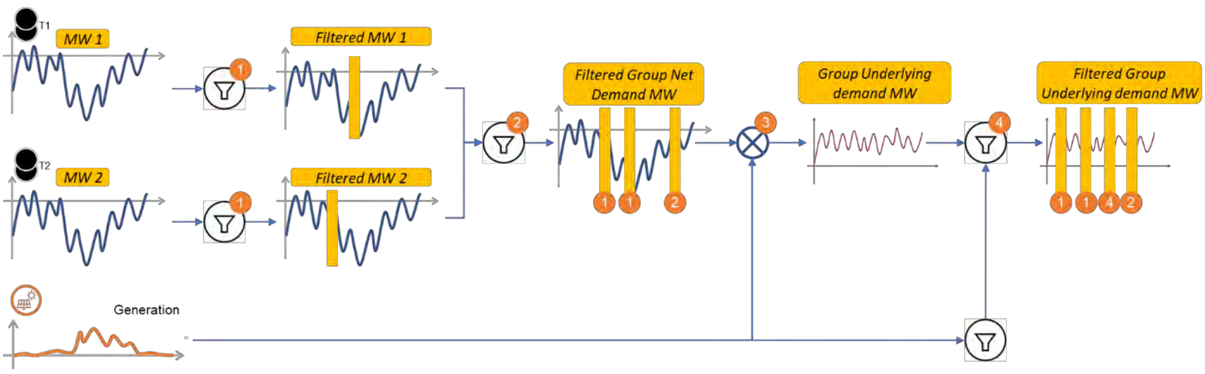


Figure 8 - Load Forecasting Solution Signal Processing

The generation part of the data pipeline is provided by Electralink (see dedicated section). The records created from the Nerda pipeline therefore stop at the net demand signal for all feeders, transformers and groups in scope.

In order to keep the solution consistent, the exact same rules are being applied to newly collected data points. The existing processes had to be readjusted to account for the origin, frequency and period of the data collected.

3.5 Challenges and solutions

3.5.1 Missing measurement_id

The NeRDA solution was not specifically designed for the TRANSITION project and several measurement_ids were missing from the original scope. Several alignments were required to ensure at least 1 Amp or MW measurement_id was available for each transformer and feeder on the network.

In some instances, newly created PI Tags had not been created on NeRDA. This meant the connection between PI Historian and NeRDA was not operational and the Load Forecasting solution could not query the dedicated measurement_ids.

3.5.2 Duplicated measurement_id

In a number of instances, multiple measurement_ids were available for the same asset on the same measurement_unit. This corresponded to multiple objects in PI Historian or NeRDA such as the analogue and the switch. The introduction of the PI Tag allows to clearly identify which tag was recording the values.

3.5.3 Data gap September June 2021 to February 2022

The NeRDA solution became operational in February 2022. Prior to this date the solution was working off a static historical data set which finishes in June 2021. This means that there was data missing between June 2021 and February 2022.

This data gap did not impact specifically the ongoing management of the pipeline. However, it was impact the research of optimisation parameters (see dedicated section). Moreover, for the sake of consistency and continuity it was essential to fill in the data gaps and provide users a full historical dataset.

3.5.4 Change of analogues

For a specific substation, the analogues on one of the transformers have been replaced. Originally, only Amps measurements were available at T1. However, the readings were constant for more than a year, and thus not useable. The faulty analogue has been replaced by a MW analogue, which only started in October 2022.

This results in a lack of depth of history for the substation as a whole and consequences for the selection of best optimisation parameters (see dedicated section)

4 Electralink interface & pipeline

4.1 Description and objectives

ElectraLink[7] was created in 1998 by the UK's electricity Distribution Network Operators (DNOs) to provide an independent, secure and low-cost service to transfer data between the participants in the deregulated UK electricity market. The company continues to operate the regulated Data Transfer Service (DTS) that underpins core processes that are critical to a competitive energy market.

Specifically of interest to the TRANSITION project, Electralink captures and make available "close to real time" settlement data from generators participating in the Energy market, notably small-medium scale generators embedded deep within the HV network not monitored by DNOs. These settlements data will be used to understand the real observations from all generators in scope of TRANSITION.

The settlements provided by Electralink represent the energy that the generator has injected on the network. It is provided as a timeseries of individual periods. On daylight saving days 46 or 50 periods are populated, whereas all other days have 48 periods populated.

4.2 Data security requirements

Electralink holds the settlement data of a huge number and wide range of electricity consumption customers. Care had to be taken in this Phase 2 project to ensure that access to this data for SIA, a 3rd party in the DNO -> Electralink relationship, was (i) limited (ii) targeted (iii) controlled and (iv) used in a manner solely for the specific intended purpose.

No customer sites of a personal data / GDPR sensitivity were relevant for the TRANSIITON trials scope of interest which reduced any risks accordingly, however, strong controls were nevertheless put in place and correct procedure implemented to ensure that the SIA access process adhered to these basic principles.

SIA Partners only needed access to the Electralink settlement data from a small number of generation sites in Oxfordshire area of trials. Furthermore the access to this data was controlled by a "whitelisting" process whereby the DNO was a 'keyholder' intermediary that controlled the list of sites open to API access. SIA were not able to see the MPAN numbers at any point, only the pseudo-identifiers that were provided by SSEN and the API query settlement data returned. Furthermore, only a defined list of IP address machines form the SIA side were approved to connect to the API.

This process, along with proper scrutiny of the solution by IT experts on all 3 sides, ensure the solution was fit for purpose in this respect.

Network flows from Nerda could have been used to capture those observations. However, in order to prevent the direct identification of generators to third parties, the network data has not been made available to the Load Forecasting solution. Instead, Electralink was able to provide the settlements associated with each generator.

[7] Home - ElectraLink

4.3 Connection requirements

4.3.1 Pentest requirements

In order to guarantee the safety of the infrastructures, a Cyber security assessment of the Load Forecasting solution has been performed, by an independent and recognised third-party. While 3 minor potential vulnerabilities have been identified, the overall security of the solution was evaluated as Good and on par with industry standards.

The audit has highlighted the existence of good security practices on the application. The application authorization scheme correctly restrict access to protected data and could not be bypassed. In addition, the application uses a strong layer of cryptography to protect the confidentiality and the integrity of the communication

Once the requirements have been fulfilled and demonstrated to SSEN's internal Cyber team and Electralink, the connection to the APIs could be tested.

4.3.2 IP whitelisting & static IP

Electralink provided access to both their Production and Development environments. For each of those environments a list of accepted IP had to be shared and whitelisted by SSEN and Electralink to allow connection.

The Load Forecasting Solution initially used a Dynamic IP. For the purposes of the project, a mechanism has been put in place to fix the IP seen by Electralink. Each of the Load Forecasting solution environments (PROD and DEV) had a dedicated static IP, which could then be whitelisted to gain access to Electralink.

For each IP an APIKey was generated and communicated securely to authenticate on Electralink's service. These credentials have been fixed and communication successfully and securely established.

4.3.3 Endpoints available

For the purpose of the project, Electralink customised existing endpoints to retrieve the settlements of the generators in scope. The endpoints were made available to SSEN, and the Load Forecasting solution. Access to Electralink swagger allows to tests those endpoints individually.

The Load Forecasting solution used a dedicated endpoint which allows to retrieve the settlement data for a given pseudonymised generator between 2 given dates. The result is based on the existing service provided by Electralink in its Data Transfer Catalogue (DTC) D0036: *Validated Half Hourly Advances for Inclusion in Aggregated Supplier Matrix*. It provides the Half Hourly consumption values for use in Supplier and Distributor billing and for submission to HHDA.

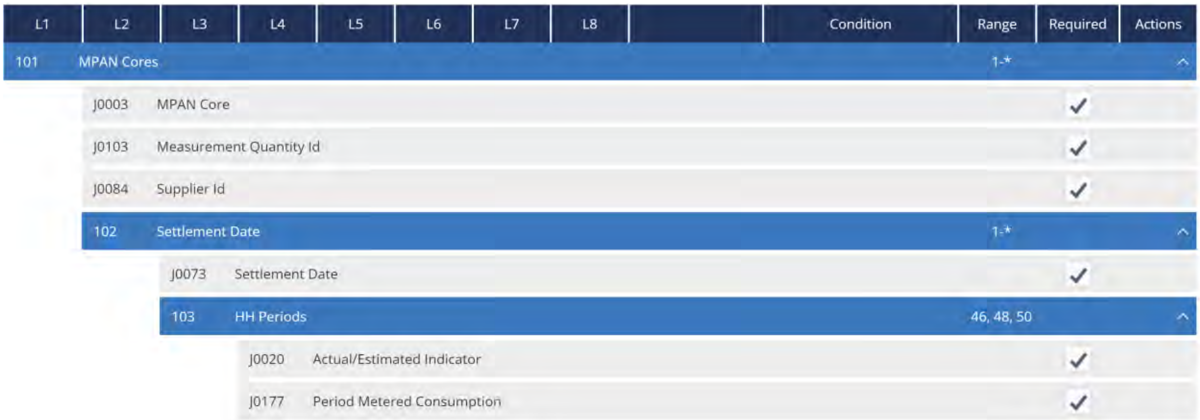


Figure 9 - Data structure of Electralink endpoint payload

4.4 Data mapping

Electralink and the SIA Load Forecasting Solution have been designed with 2 distinct Data models. In order to collect the data from Electralink on an ongoing basis, the first objective was to ensure we could align the data models and reconcile the ongoing sources of data.

In Electralink, each generator is represented by a set of 2 MPAN, capturing respectively the Export and Imports measured at the site. For data privacy considerations, the Load Forecasting solution cannot have access to the MPANs. Therefore a mapping table has been developed between Electralink and the Load Forecasting Solution where SSEN played the partial role of segregating the information on each side. It included the *MPAN* (for Electralink) and a *site* and *description* for the Load Forecasting solution.

The table was uploaded in Electralink and constitutes the list of available MPANs the Load Forecasting solution was able to query. Several updates of the table were required to align on the entire scope

The *site* field was populated with the CIM Name of the Asset, and the *description* determine if it was exporting or importing MPANs. The couple (site, description) was then matched to their respective generation asset in the Load Forecasting solution. The example below shows the mapping for a subset of generation asset.

site character varying	description character varying	electralink_asset_id [PK] integer	asset_id integer
BAT[redacted]	Import MPAN	310[redacted]	310[redacted]
BAT[redacted]	Export MPAN	310[redacted]	310[redacted]
BOO[redacted]	Import MPAN	310[redacted]	310[redacted]
BOO[redacted]	Export MPAN	310[redacted]	310[redacted]
CUL[redacted]	Import MPAN	310[redacted]	310[redacted]
CUL[redacted]	Export MPAN	310[redacted]	310[redacted]
FFC[redacted]	Import MPAN	310[redacted]	310[redacted]
FFC[redacted]	Export MPAN	310[redacted]	310[redacted]
HOV[redacted]	Import MPAN	310[redacted]	310[redacted]
HOV[redacted]	Export MPAN	310[redacted]	310[redacted]
LAN[redacted]	Import MPAN	310[redacted]	310[redacted]
LAN[redacted]	Export MPAN	310[redacted]	310[redacted]

Figure 10 - Electralink Mapping table example

4.5 Data pipeline

The diagram below provides an overview of the ongoing Electralink pipeline that has been implemented to capture transformers and feeders net flows.

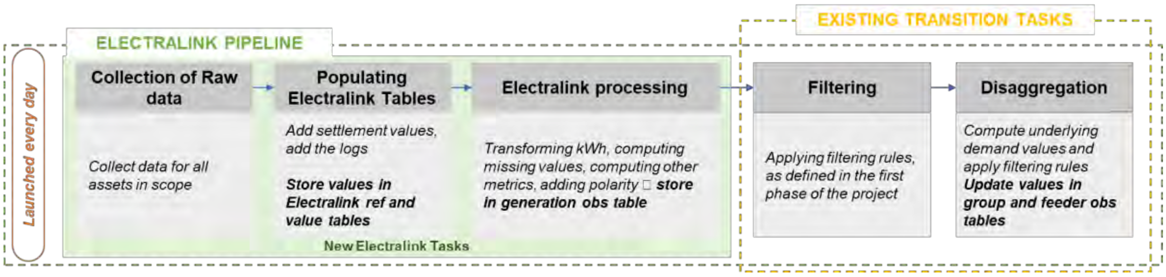


Figure 11 - Electralink Pipeline

4.5.1 Collection of raw data

Settlements from generators are collected and managed by Electralink. The half-hourly measurement follow an internal process unknown to the Load Forecasting solution. However, they impact the date the Load Solution will receive, and therefore how Electralink should be queried. 2 key features have been identified:

1. <u>Delay in availability of settlements:</u>

Electralink settlements are made available on a daily basis early in the morning for 3 days before. The diagram below show when data is available when Electralink is queried in the morning of the day D.

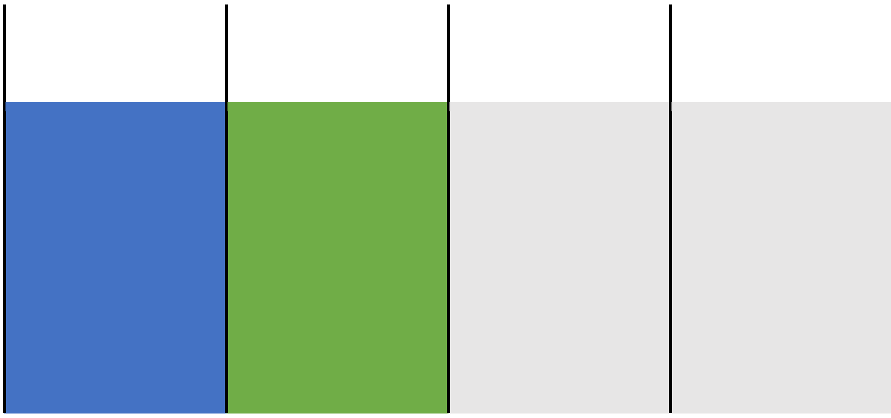


Figure 12 - Availability of Electralink Settlements

2. <u>Actual VS Estimated</u>

Every half-hourly settlement , from the most recent day available or in the past, can either be a true measure or an estimation of the output of the generator. This is captured by Electralink processes and provided as an information. For an estimated half-hourly settlement, when the *Actual* value is integrated by Electralink, it will automatically replace the *Estimated* value.

These 2 features had an impact on the design and schedule of the solution. The endpoint available to the Load Forecasting solution allowed to query for a specific generation asset the available settlements between 2 dates or from a specific date. On an ongoing basis, Electralink was therefore scheduled to be queried every day to capture the most settlements for the generation assets. For each asset recorded and mapped to the Load Forecasting Solution data model, the solution queries the Electralink endpoint, using the latest day recorded and uploaded as a starting point. All values returned by Nerda are uploaded.

It was understood that Estimated half-hourly settlement were not common and usually due to communication issues. So when estimated, an half-hourly settlement would normally receive the actual measurement fairly quickly afterwards. It was therefore decided that for each generator, the last 7 days would be queried on a daily basis. If a measurement was *Actual*, then the new value would be the same and if *Estimated* it would be replaced by a more accurate value.

Finally, Electralink endpoint (as presently implemented) allows a depth of 14 moths for history for the settlements. Since the original development of the Load Forecasting solution relied on PI network data until 31/12/2020, the last 14 months of generation settlements were queried for all generation assets in scope. This left a gap generation data of ~ 8 months between 01/01/2021 and 01/08/2021.

4.5.2 Processing of raw data

For each generation asset, for each day a full week of half-hourly settlements are integrated for both Export and Import MPANs from Electralink. The values of those settlements are provided in kWh. In order to align with the Load Forecasting solution data model which uses MW, every half-hour settlement is transformed in an average Active Power over the half-hour, assuming constant instantaneous power. No Reactive Power has been considered throughout.

Both Import and Export MPANs are available. After a number of test to compare historical generation data and ongoing settlement data, it was decided to only use Export data to represent the output of the generator on the network.

4.5.3 Integration in LFS data model

For each generation asset, we use the mapping table developed to determine whether settlements are expected from Electralink. Not every generator will have settlements, especially LV generation asset.

For each asset where settlements from Electralink are expected, all settlements from all Export MPAN relating to this asset are captured, aggregated and uploaded in the generation_obs table already developed in the first phase of the project. Moreover, to ensure the best representation of the settlements, the entire previous week is updated based on the most recent settlements.

4.5.4 Load forecasting data processing – use of existing TRANSITION tasks

The Load Forecasting phase 1 solution already holds a number of processing rules which have been put in place in the first developments. Those rules have been developed over historical datasets and aimed to

1. Filter outputs signals at generation asset level to ensure that the resulting signal observed at the point of forecast is of good quality.
2. Compute the underlying demand at Group and feeder level to determine the real level of consumption across the network
3. Filter the underlying demand signal based on generation and net demand filters

In order to keep the solution consistent, the exact same rules are being applied to newly collected data points. The existing processes had to be readjusted to account for the origin, frequency and period of the data collected.

Since the Electralink pipeline has already created the records, the Disaggregation task will update the underlying demand and its filtering in the group_obs and feeder_obs tables.

4.6 Challenges and solutions

4.6.1 Assets with multiple MPANs

The Load Forecasting solution relied in the first instance on PI data provided for each generator asset. In some instances, a single load timeseries was provided as a representation of the output of a generator. However, the situation on-site showed that multiple MPANs were installed to record the output of multiple installations. For instance a PV farm with multiple stages of development or several Wind farms.

The mapping between the 2 solution therefore needed to account for the multitude of MPANs for a single asset

4.6.2 Actuals vs estimated

Since the data provided by Electralink is always the best view of the settlements, running a study on the occurrence of estimated settlements is difficult. Indeed, once a settlement has received actuals we lose the knowledge of previous estimated measures.

To ensure we always bring back the best view, the daily process could look to query the entire history of available settlements. However, from a processing point of view this takes much longer and brings additional delay in the availability of the settlements and the rest of the pipeline.

Defining the right period to capture was therefore based on the data available at present and expert view and information from Electralink. A period of 7 days was therefore selected, assuming that all estimated settlements would be resolved by then

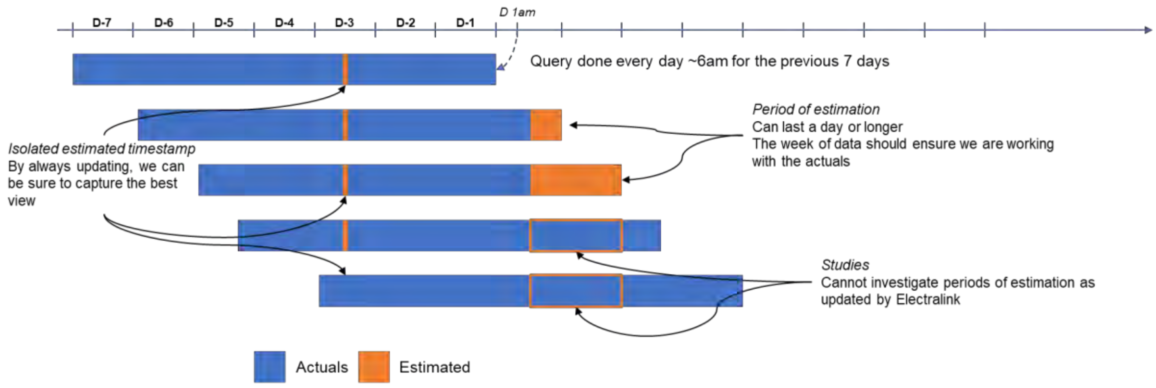


Figure 13 - Impact of Actuals VS Estimated in the data collection process

4.6.3 Delay in data availability

It was always acknowledge that the data received by Electralink was a daily dump for the previous 48 half-hourly settlements. Therefore some delay was anticipated to a maximum of 1 day.

However, after connection to Electralink and investigation of the data received, there is a minimum of 2 days delay as mentioned at the beginning of the section. This impacted the design of the data pipeline (see next section) and the design of the Forecast Optimisation module (see dedicated section)

5 Real-time connexion pipeline

5.1 Overview of the new pipeline

The 2 individual pipelines developed have been brought together to form a new Load Forecasting Pipeline for Observations. This does not impact the original pipelines developed to evaluate demand and generation models with weather forecast and provide a Load Forecast.

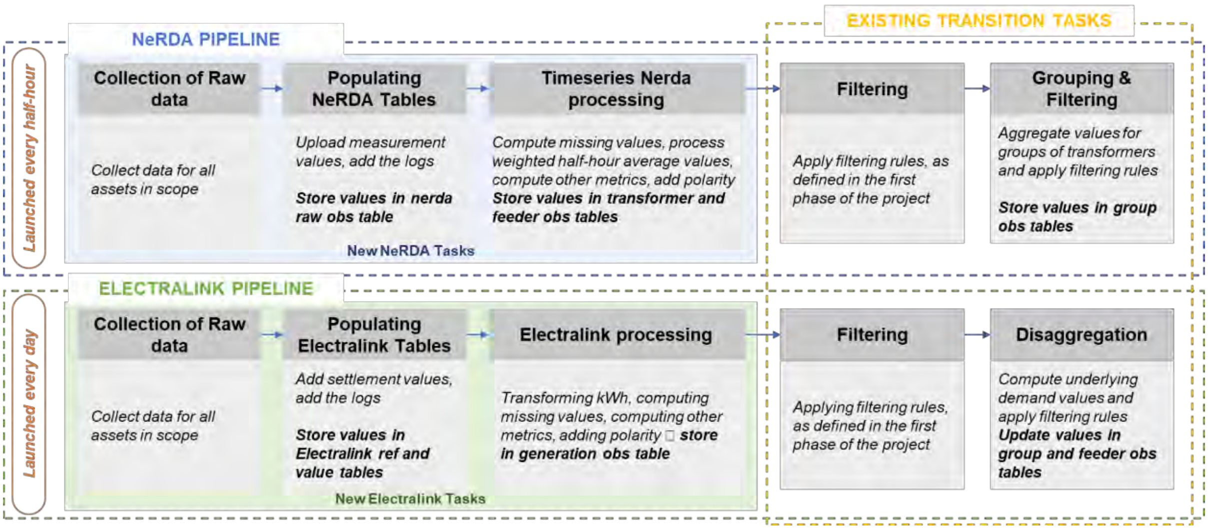


Figure 14 - Real-time Connection Pipeline

5.1.1 Pipeline scheduling

Due to the availability of the data from each source, the alignment of the 2 data pipeline needs to be carefully managed. The frequency of Electralink data provision being the limiting factor, several options were considered for the Nerda pipeline

1. Option 1: Collect all data (NeRDA & Electralink) once day, in line with Electralink
2. Option 2: Collect NeRDA data at a higher frequency than daily and Electralink once a day

Each option has advantages and disadvantages

	Option 1: All together	Option 2: Separate schedules
Pros	**Simplicity** and accuracy: 'accurate' information is displayed consistently across the scope of assets Net demand, underlying demand and generation outputs are made available for all assets at the same time	**Speed**: available data are displayed more rapidly Net demand for all groups and feeders can be displayed throughout the day, but not underlying demand and generation assets
Cons	**Delay**: all data are displayed with delay **Processing:** larger datasets are processed simultaneously	**Lack of clarity**: not all assets (generation, feeders and groups) will have the same data available throughout the day

Following the realisation that a minimum of 2 days will be observed in Generation data, it was decided to implement Option 2 and provide users with the most recent data, despite having a complete scope.

5.1.2 Impact on disaggregation

Because the 2 pipeline are not synchronous, there is an impact on the availability of the data, both to be displayed and to be used in the Forecast Optimisation modules.

The underlying demand is computed by removing the generation from the net demand measured at transformer and feeder levels. However, part of the generation is only available much later due to Electralink availability. The additional complexity is that not all feeders and groups rely on Electralink data for the purpose of disaggregation. Some LV generators are do not hold settlements and therefore waiting for Electralink would not be necessary.

The following table shows that most feeders do not have Generators connected for which we expect Electralink settlement data. However, almost all groups are impacted by the delay.

	Electralink generation	11kV	33kV	Total	%
Feeder	*Yes*	14	10	24	12%
	No	110	65*	175	88%
Group	*Yes*	9	6	15	79%
	No	4	0	4	21%

** Impact of HV generator with settlement cannot be measured as the Load Forecasting solution does not hold topology*

Following investigation, it was decided have a consistent approach for all assets and wait for Electralink to trigger the Disaggregation for all assets.

5.1.3 Task orchestration

Based on the 2 previous observations, the pipeline are therefore running independently from each other. Each step is triggering the next stage of processing. In the Nerda pipeline the final stage is the creation of new records in group_obs and feeder_obs tables populated and filtered on the net demand.

For the Electralink pipeline, the final stage is the update of the group_obs and feeder_obs tables by adding the underlying demand computation (based on settlement and modelled generation) and filtering. The generation_obs table has been populated a step before.

5.2 Alignment of the sources of data

5.2.1 Difference PI average / PI max & alignment with Electralink

With the addition of ongoing settlement data, it was possible to form of view of the accuracy of the generation models. The initial results were surprising and concerning. Multiple investigation took place to identify the poor performance of the models, despite good calibration.

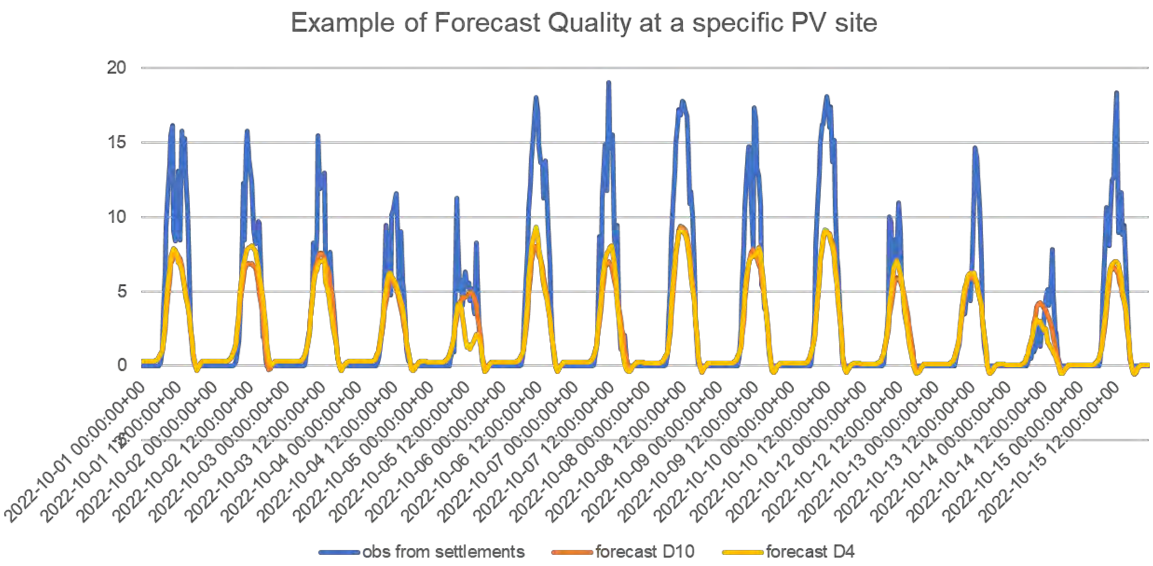


Figure 15 - Example of Poor Forecast quality for a PV generator

1. **Are the weather sources ERA5, Mosmix and Icon materially different?**

Conclusion: No evident bias to support a discrepancy in the forecast

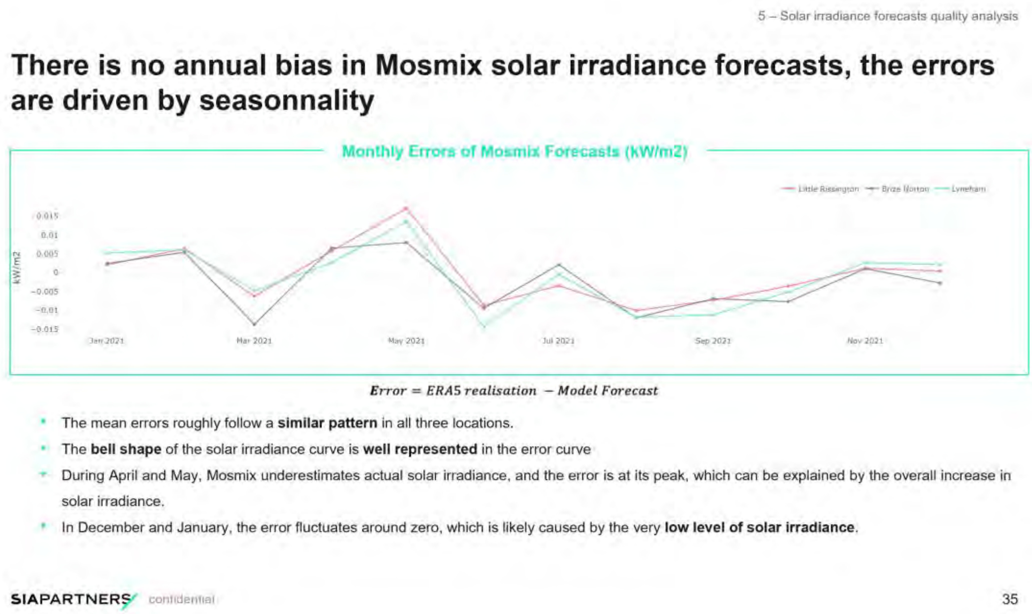


Figure 16 - Summary conclusion of discrepancy between weather sources for solarirradiance

2. **Are there errors in the calibration and evaluation scripts?**

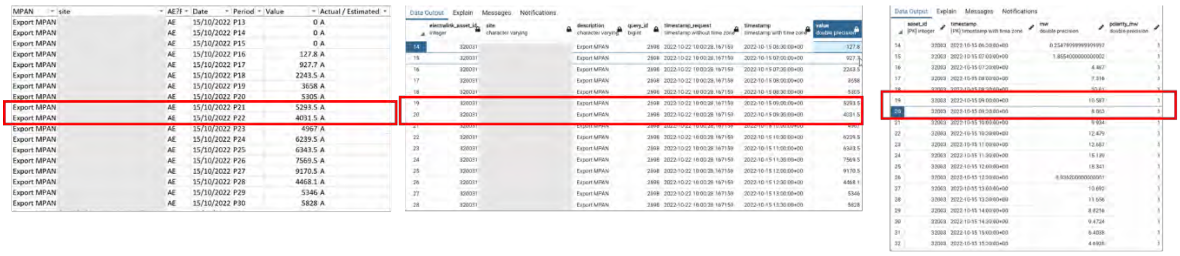
The calibration script has not been impacted. The source of data is correct and we could demonstrate that the datasets were appropriately merged.

On the forecast evaluation the models were replayed on the calibration dataset. The objective was to ensure we were have the same results through the calibration script and the model evaluation script.

Conclusion: Both scripts are sane and behave appropriately

3. Is the Electralink Pipeline working correctly?

The Electralink pipeline has been retraced manually from start to finish and could not be faulted



Raw extract from Electralink (all in kWh and actuals) *Integration in Load Forecasting Database in kWh* *Processing and upload in Generation DB in MW*

Figure 17 - Evidence of the functioning of the Electralink Pipeline

4. Is there a way to improve the models?

Various approaches have been retested to improve the quality of the models themselves. Additional variables, review of the splines in GAM and new methodologies were implemented. This marginally improved the quality of the models but material enough to conclude on the large differences in forecast.

5. Is there a difference between historical and ongoing data sets

The assumption throughout the project was that the PI network flows provided to calibrate the model was of the same nature as the Electralink settlements from the ongoing process. As mentioned in a previous chapter, there is a gap between the Electralink data and the historical PI data. Therefore, the 2 sources could not immediately compared. Additional extractions of PI were required to match the historical Electralink Data set for a specific Generation asset.

The results showed a great difference between the 2 sources, but not consistently throughout the year.

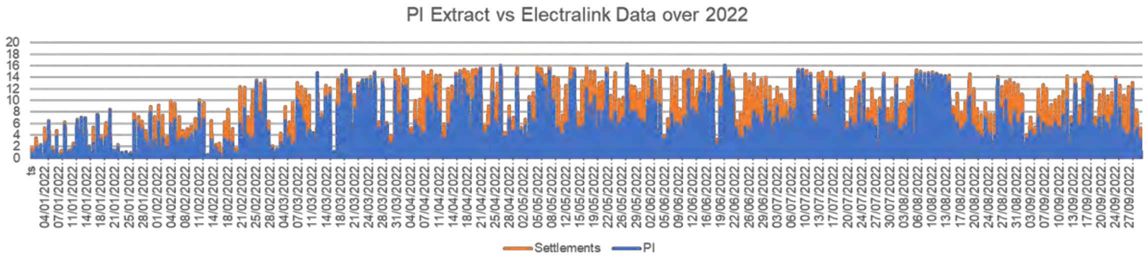


Figure 18 - PI VS Electralink over 2022

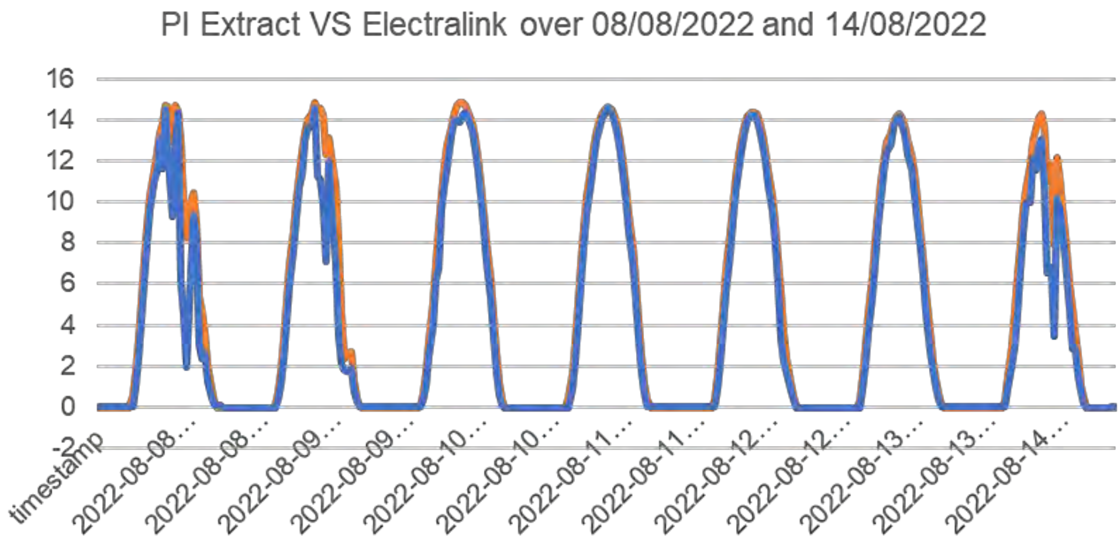


Figure 19 - Figure 17 - PI VS Electralink over a week in August 2022

PI allows the possibility to operate transformations in the data before extraction. The assumption that Average would be the correct method of extraction was questioned and both 'PI Average' and 'PI Max' were extracted and compared to Electralink settlements.

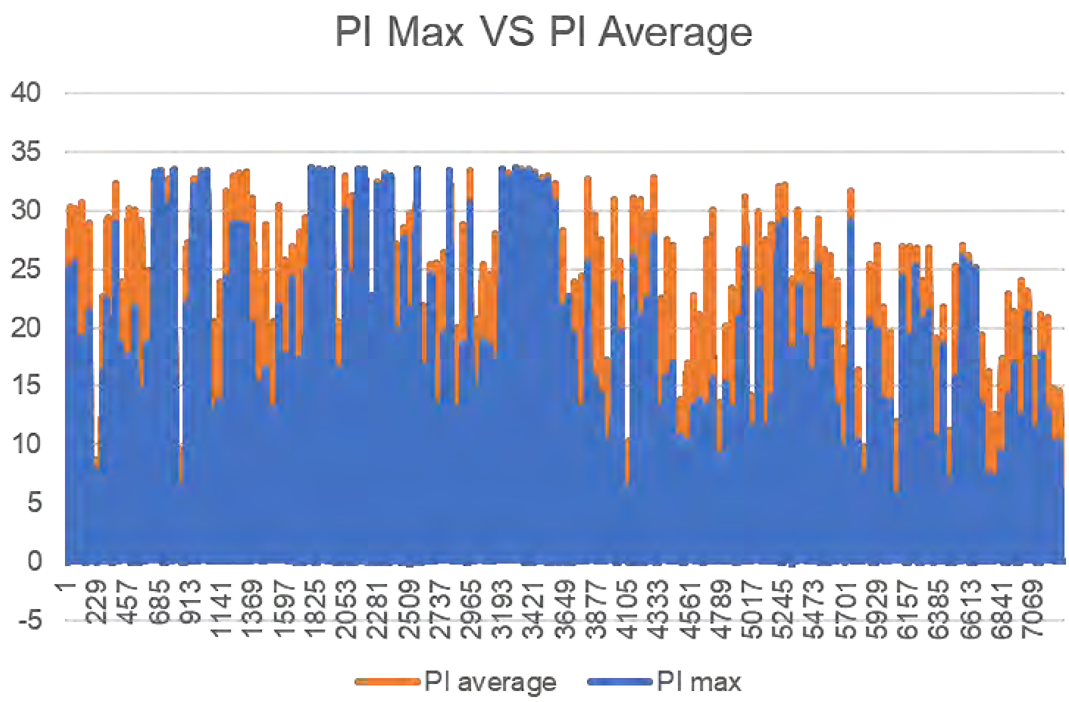


Figure 20 - Comparison PI Max VS PI Average

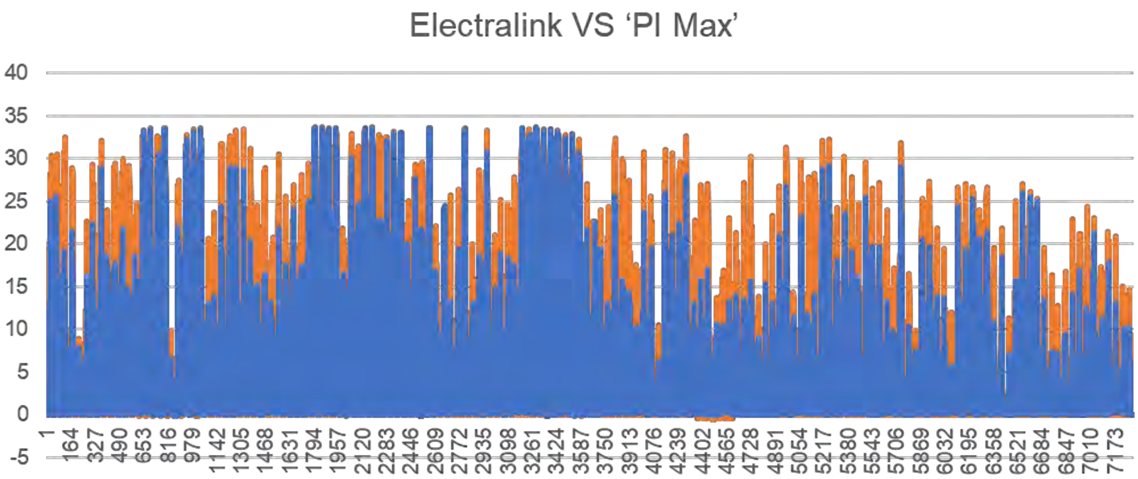


Figure 21 - Comparison Electralink VS PI Max

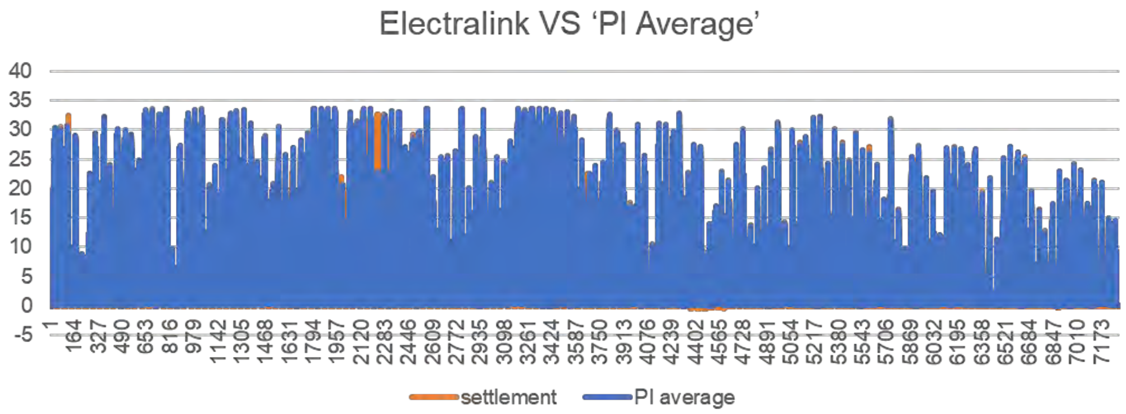


Figure 22 - Comparison Electralink VS PI Average

Conclusion

There is a clear difference between PI Max and PI Average. PI Average is aligned with Electralink and therefore should be considered as the most realistic signal for network load data measurements. From the quality of the generator models, it is assumed that the network load data provided in the first stage of the project was PI Max and not PI Average.

All generation assets have been recalibrated with newer Electralink settlement data to ensure the generation forecasts are representative of the actual outputs of the generators.

On the demand side, we can assume that the models have been computed using PI Max as well. This will naturally introduce further error in the quality of the demand forecast. The 14 months of history provided by Electralink does not provide a large enough dataset to calibrate new demand models. Indeed, since demand is highly correlated to time of year, in order to accurately capture the behaviour, models need to be trained on several years. Otherwise, a single year would be considered representative, which is not acceptable.

However, with the introduction of the Forecast Optimisation, realignment to both changes of demand and PI Max to PI Average can be expected. While it is anticipated that the forecast will yield better accuracy, the origin of the error of the original forecast cannot be determined.

6 Demand forecast optimisation methodology

6.1 Objectives

The models calibrated can only reflect the behaviour of the historical period they have been calibrated on. They are evaluated with short-term weather forecasts used to model the weather sensitivity of the demand and generation. However, if a change of behaviour in demand happens after the calibration, the model cannot take this into account.

Recalibrating the model frequently would only partially mitigate this problem as more recent history would capture the change of behaviour. However, considering the depth of history needed in the training dataset (3 years), a change of behaviour happening in the last few months of the training dataset would be diluted in the remaining years of data with the previous behaviour. This change of behaviour would therefore be captured gradually after several years. Moreover, running a full calibration of all the models is a long and computation-intensive process that needs to be run carefully.

The Forecast Optimisation implemented aims at using the most recent measured network load data to inform future forecasts. It relies on the comparison of the most recent network load data with the original forecast, determine the error and defines the way to correct future forecast, knowing this error. <u>It does not aim at recalibrating the historical data-based models, but instead adapting/offsetting their outputs.</u>

The solution retained is based on 2 steps

Firstly a **Volume Correction** to capture the change in energy delivered

Secondly an **Instant Correction** to capture the changes in the shape of the demand

For each step the depth of history used is independent for each asset and defined ahead of the evaluation of the optimised forecast. The models and Optimisation process can be summarised as below:

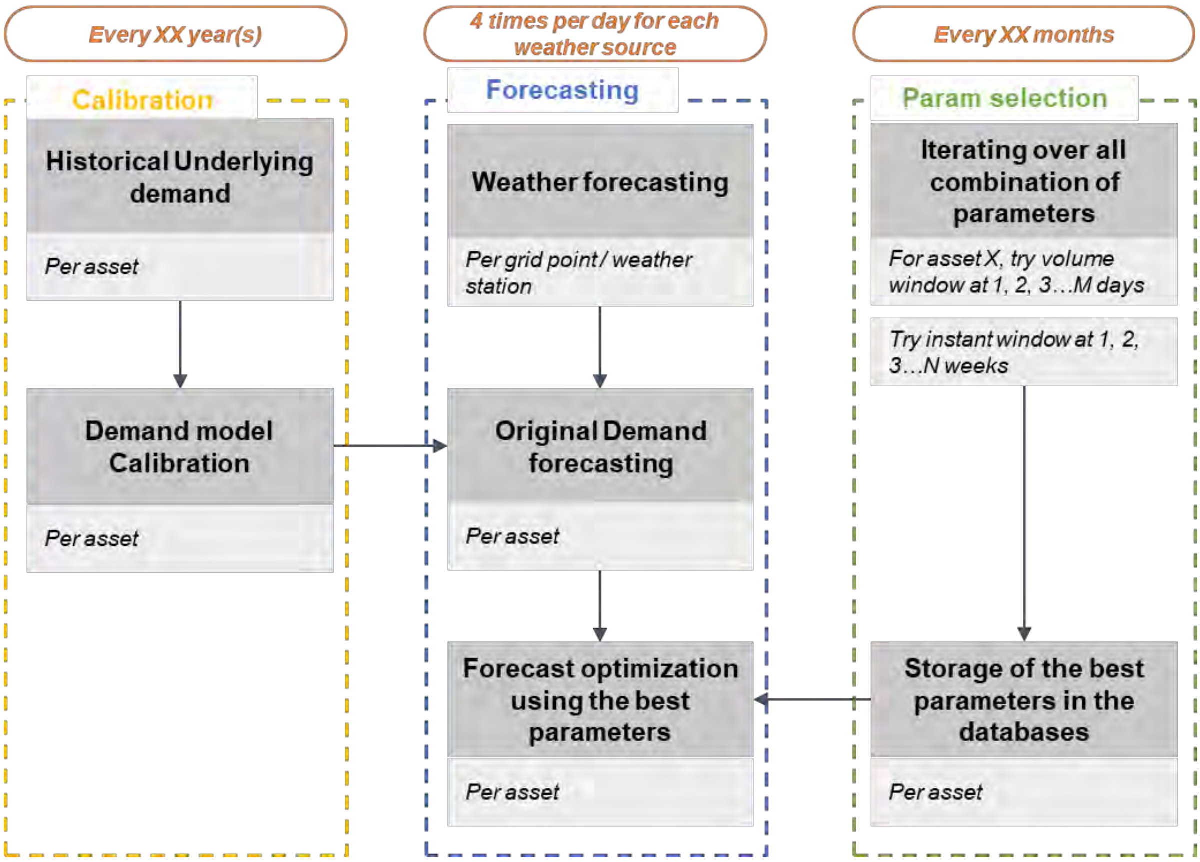


Figure 23 - Model and Optimisation Parameters processes

6.2 Description of the proposed solution

6.2.1 First order – volume correction

The **Volume correction** approach focuses on the difference between the total energy forecasted and the total energy consumed (observed underlying demand) over the previous days. The correction is based on the difference of the two volumes.

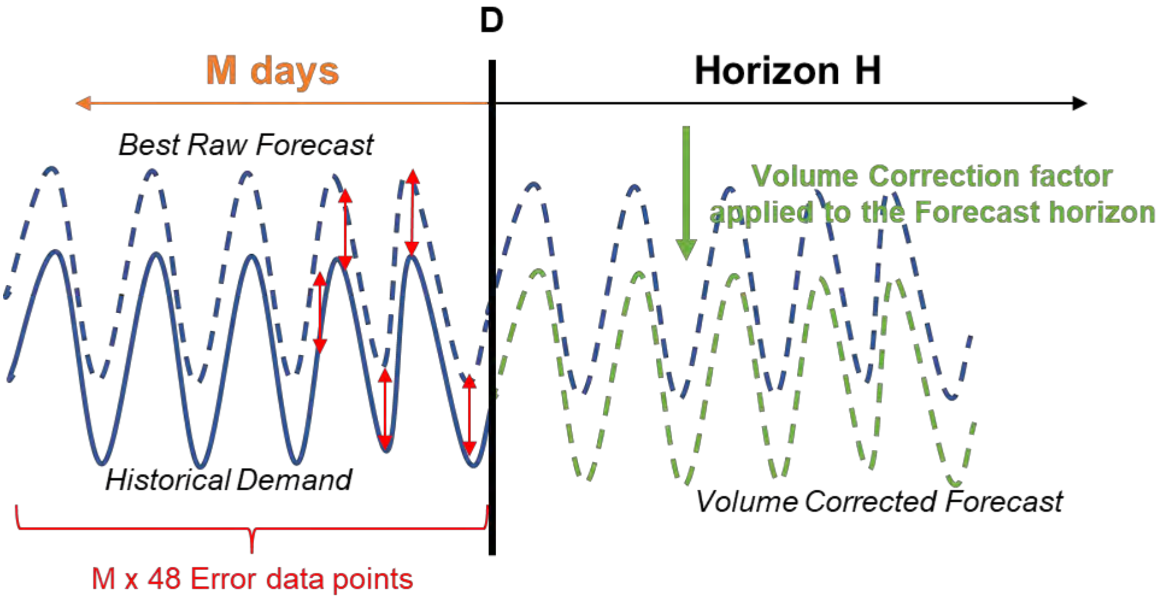


Figure 24 - Volume Correction Methodology

Over the previous **M days**, the error between the Best Forecast and the observed underlying demand is computed. Historical underlying demand flows are computed using Nerda and Electralink data. Each **asset K** (group, feeder) has a dedicated demand model, calibrated over historical data. Each **forecast run D** determines the expected future demand over the **Horizon H**.

The error computed over the previous M days becomes a *Volume Correction Factor* (scaling factor), valid for each asset and recomputed at each Forecast run:

$$Volume\_correction\_factor_{K,M,D} = \frac{\sum_{i=1}^{48*M} Hist\_Underlying\_Dmd_K(D-i)}{\sum_{i=1}^{48*M} Best\_Original\_Forecast\_Dmd_{K,D}(D-i)}$$

After Volume Correction, for a given asset K, the *Volume Corrected Forecast* D run then becomes:

$$Volume\_Corrected\_Forecast_{K,D} = \begin{bmatrix} Volume\_Corrected\_Forecast_{K,D}(1) \\ \dots \\ Volume\_Corrected\_Forecast_{K,D}(H) \end{bmatrix}$$
$$= \begin{bmatrix} Best\_Original\_Forecast\_Dmd_{K,D}(1) \\ \dots \\ Best\_Original\_Forecast\_Dmd_{K,D}(H) \end{bmatrix} * Volume\_correction\_factor_{K,M,D}$$

6.2.2 Second order – instant correction

Definition: An instant is defined by a specific half-hour timestamp of the week. There are therefore 48 * 7 = 336 instants per week.

The Instant correction will look to capture the longer-term changes at specific time of the week. As each day of the week has its own specific pattern of demand, it is important to compare it to the same day from the recent historical data as the day that the predictions are being made for. For example, it investigates the error made on Monday 08:00am of the previous weeks to correct the forecast of Monday 08:00am in the coming forecast.

The correction therefore aims at correcting the shape of the forecast without changing the volume delivered (which has been corrected in the previous step).

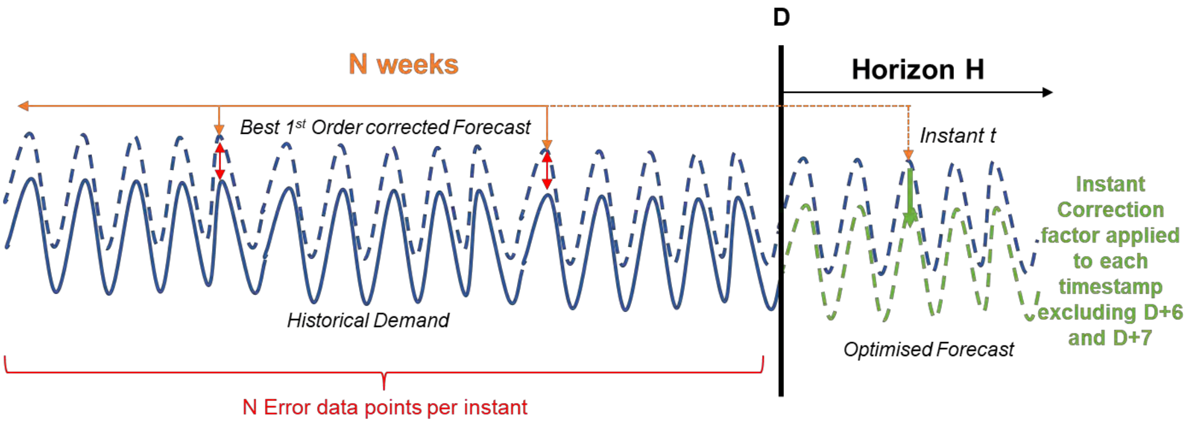


Figure 25 - Instant Correction Methodology

Over the previous **N weeks**, the error between the *Volume Corrected Forecast* (from previous step) and the observed underlying demand is computed for each instant. Historical underlying demand flows are computed using Nerda and Electralink data. Each **asset K** (group, feeder) has a dedicated demand model, calibrated over historical data. Each **Volume Corrected Forecast run D** determines the expected future demand over the **Horizon H**.

For each instant t of each Volume Corrected Forecast run D of each asset, the *Instant Correction Factor* (scaling factor) is computed, by taking the average of the errors over the previous N weeks:

$$\forall t \in [1, H], Instant_{correction_{Factor}\,K,M,N,D}(t) = \frac{\sum_{i=1}^{N} \frac{Hist_{Underlying_{Dmd}\,K}(t - i * 48 * 7)}{Volume_{Corrected_{Forecast}\,K,M,D}(t - i * 48 * 7)}}{N}$$

After Volume Correction and Instant Correction, for a given asset K, the *Optimised Forecast* D run then becomes:

$$Optimised\_Forecast_{K,M,N,D} = \begin{bmatrix} Optimised\_Forecast_{K,M,N,D}(1) \\ \cdots \\ Optimised\_Forecast_{K,M,N,D}(H) \end{bmatrix}$$
$$= \begin{bmatrix} Volume_{Corrected_{Forecast}\,K,D}(1) * Instant_{correction_{Factor}\,K,M,N,D}(1) \\ \cdots \\ Volume_{Corrected_{Forecast}\,K,D}(H) * Instant_{correction_{Factor}\,K,M,N,D}(H) \end{bmatrix}$$

6.2.3 Overall methodology – sequencing

The two corrections are applied sequentially to ensure the instant correction benefits from the realignment of the volume correction. This allows to limit distortion instant by instant.

6.3 Research of best parameters for each asset

Each asset on the network (group, feeder) has different elements connected to it. Therefore, the profile of the demand connected will change and the impact of more recent data points will be independent and quite locationally specific, network asset by network asset.

In order to best capture the behaviour at each asset, the depth of history considered to correct future forecasts (M days and N weeks) is selected to create the optimal optimised forecast. An exhaustive gridsearch is applied to each asset to determine what is the best combination/couple of parameters for this specific asset.

For each asset, the solution will look to apply the above-defined methodology with variable M and N parameters and evaluate the quality of the resulting Optimised forecast over a given period, compared to the measurements. The couple of parameters which provides the most accurate optimised forecast is selected to be the new set of default parameters for the future evaluation of the Optimised Forecast.

The period on which the parameters have been looked for in August 2021 to November 2022. This is to ensure measurements on both Generation and Net demand were available.

6.4 Challenges and solutions

6.4.1 Availability of Electralink data

On an ongoing basis, Electralink data is only available after 2 days (see dedicated section). Therefore, when applying both Volume and Instant correction methodologies, the observed underlying demand of the previous 2 days, composed on observed net demand from Nerda and settlements from Electralink, is not available.

When looking for the best parameters for each asset, the processes applied for determining the best couple need to reflect the situation that they will be used in. Each of the 2 methodologies have been tweaked to reflect this impact.

<u>Volume Correction</u>

The 2-day delay means that the ongoing volume correction will look to correct with past data not available. For the research of the best parameter M, the previous 2 days will therefore be omitted in the volume correction part of the optimisation.

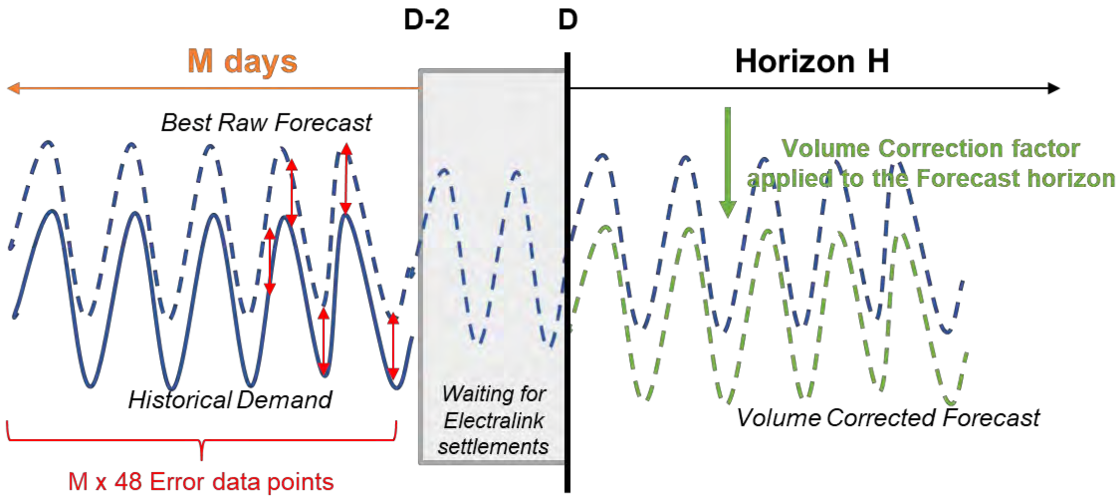


Figure 26 - Revised Volume Correction Methodology with Electralink Impact

<u>Instant by instant Correction (focus on the shape):</u>

The 2-day delay means that the ongoing instant correction will look to correct with past data not available for any instant in D+6 and D+6. However, for D to D+5 there is no impact. Therefore, instead of shifting all comparisons by a full week, it is assumed that the process for finding the research of the best instant correction parameter N, the methodology has not been changed to limit the impact on other days.

6.4.2 Filtered data and impact on each step

In order to ensure the correction is applied to a representative data set, the definition of the best parameters will be taken into account only with a enough valid data points. The historical measured net flows from Nerda and the settlements from Electralink and filtered through the defined processes. This defines which underlying demand data points are valid.

Moreover, since the instant correction method is based on the outcomes of the volume correction, the instant correction will not be carried out if the volume correction has not been successful.

6.4.3 Definition of the quality metrics to select best parameters

In order to select the best parameter for a specific asset, some quality metrics have been implemented to rank them. 3 Metrics have been proposed for evaluation and computed:

Mean Absolute Percentage Error (MAPE): Measures the average of relative error between optimised forecast and observed underlying demand over the period

$$\mathrm{MAPE} = 100\% * \frac{1}{N}\sum_{t=1}^{N}\left|\frac{Error(t)}{Realised(t)}\right|$$

Mean Absolute Percentage Error * (MAPE*): Similar to MAPE but compares the error to Maximum realised over the period analysed. This metric is particularly more representative for dataset with values close to 0

$$\text{MAPE}^* = 100\% * \frac{1}{N} * \frac{\sum_{t=1}^{N} |Error(t)|}{|Max(Realised)|}$$

Root Mean Square Error (RMSE): Measures the error between forecast and realised expressed in MW. This provides a sense of absolute magnitude of the error:

$$\text{RMSE} = \sqrt[2]{\frac{1}{N}\sum_{t=1}^{N}(Error(t))^2}$$

For the selection of the best parameters, it has been decided to use the MAPE* as the key indicator of quality. Therefore the couple (M, N) of parameters with the lowest MAPE* will be selected to be applied on the ongoing evaluation of the Optimised forecast.

However, all 3 metrics are important to look in order to get a good understanding of the accuracy of the forecast. Indeed, a very high MAPE or MAPE* does not necessarily mean very poor results if the asset has very low load.

6.5 Results

6.5.1 Primary substations

The research of best parameters has been run for all HV groups. For each 'couple' of parameters the MAPE, MAPE* and RMSE metrics have been computed over the period August 2021 to December 2022. Below is the view of the accuracy metrics for each group without any real time data flow correction. Thus it compares Best forecast and real observations signals.

	Forecast D10			Forecast D4		
HV Group	**MAPE**	**MAPE***	**RRMSE**	**MAPE**	**MAPE***	**RMSE**
Arncott	31%	31%	1.133	33%	33%	1.232
Berinsfield All Feeders individually	21%	21%	3.398	22%	22%	3.643
Bicester	14%	14%	4.452	13%	13%	3.604
Bicester North Primary	9%	9%	1.238	9%	9%	1.173
Deddington All Feeders individually	10%	10%	0.064	10%	10%	0.08
Eynsham	10%	10%	0.691	10%	10%	0.636
Kennington	19%	19%	0.207	19%	19%	0.217
Milton	32%	32%	32.311	33%	33%	33.513
Oxford Primary	15%	15%	3.324	17%	17%	3.819
Rose Hill	20%	20%	2.857	21%	21%	3.179
University Parks	6%	6%	0.92	6%	6%	1.059
Yarnton Primary	9%	9%	1.335	9%	9%	1.397

Figure 27 - Forecast Accuracy of historical data-based models without correction i.e. without any real time data flow

We can note that D4 and D10 have similar performance which demonstrates alignment between the 2 weather models.

In order to appreciate the improvement of the forecast accuracy, below is an example of the distribution function of the error. This represents the difference between the original D10 forecast and the observed underlying demand on one side, and between the optimised D10 forecast and the observed underlying demand on the other. The study has been produced between 01/11/2022 and 28/02/2023. We can see that the original forecast was clearly overestimating the demand. On the other hand the optimised forecast error is both centred and more dense around zero.

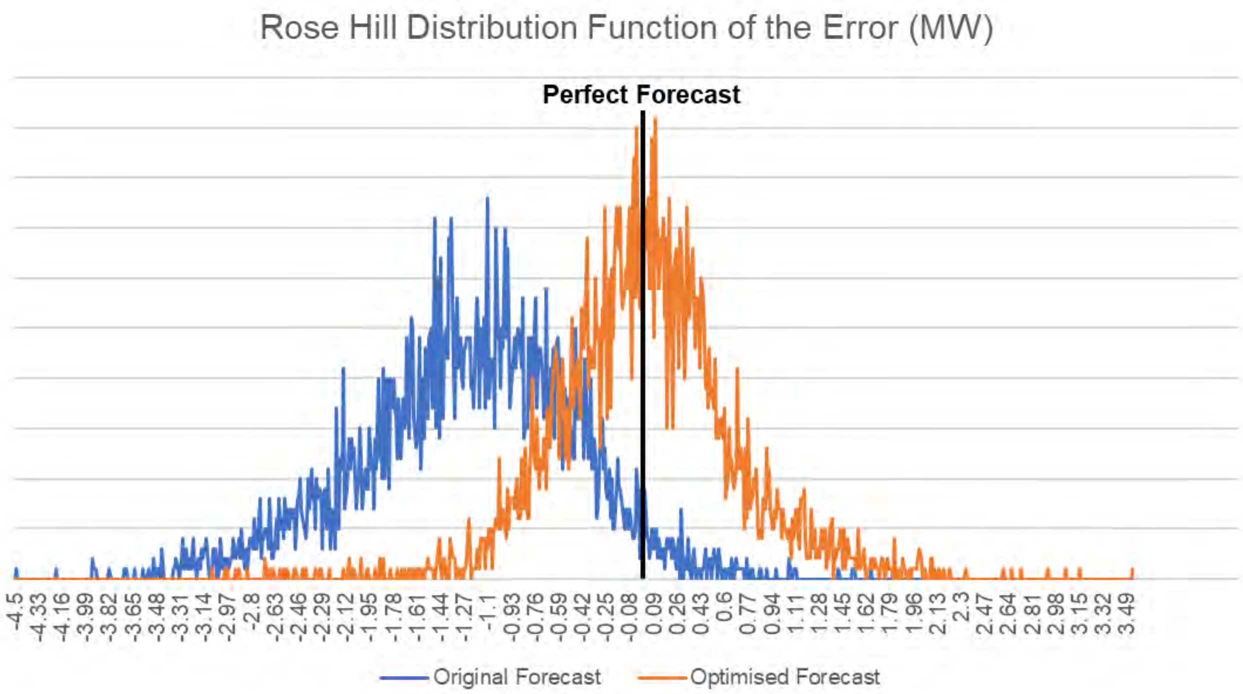


Figure 28 - Rose Hill Error distribution (MW)

The research of best parameters for each HV group gives the following results:

	Forecast D10		Forecast D4	
HV Group (Forecast D10)	**Volume Parameter**	**Instant Parameter**	**Volume Parameter**	**Instant Parameter**
Arncott	7	0	7	0
Berinsfield All Feeders individually	7	28	7	28
Bicester	7	28	7	28
Bicester North Primary	7	28	7	28
Deddington All Feeders individually	7	28	7	28
Eynsham	2	28	2	28
Kennington	7	28	7	21
Milton	2	21	2	21
Oxford Primary	7	28	7	21
Rose Hill	7	28	7	28
University Parks	7	28	7	28
Yarnton Primary	7	21	7	21

Figure 29 - Best Forecast Correction parameters for HV groups

After applying the correction of the original forecasts based on the real time data flow, the new metrics show great improvement on all groups:

	Forecast D10			Forecast D4		
HV Group (Forecast D10)	**MAPE**	**MAPE***	**RMSE**	**MAPE**	**MAPE***	**RMSE**
Arncott	21.4%	21.4%	0.628	21.6%	21.6%	0.642
Berinsfield All Feeders individually	13.1%	13.1%	1.656	13.4%	13.4%	1.693
Bicester	7.5%	7.5%	0.953	7.4%	7.4%	0.948
Bicester North Primary	7.5%	7.5%	0.998	7.2%	7.2%	0.901
Deddington All Feeders individually	9.7%	9.7%	0.06	9.7%	9.7%	0.076
Eynsham	7.1%	7.1%	0.317	7.2%	7.2%	0.323
Kennington	12.0%	12.0%	0.095	11.8%	11.8%	0.096
Milton	21.3%	21.3%	19.731	22.0%	22.0%	20.981
Oxford Primary	6.3%	6.3%	0.693	6.3%	6.3%	0.713
Rose Hill	7.7%	7.7%	0.579	7.7%	7.7%	0.601
University Parks	4.7%	4.7%	0.511	4.8%	4.8%	0.565
Yarnton Primary	7.9%	7.9%	1.233	8.0%	8.0%	1.561

Results for all set of parameters can be found in appendix.

6.5.2 HV Feeders

The research of the best short term correction parameters has run for all HV feeders individually. For each couple of parameter the MAPE, MAPE* and RMSE have been computed over the period August 2021 to December 2022.

The table below presents the results of the best parameters for all HV feeders. Out of the 124 HV feeders currently in scope of Transition:

For 88% (109/124), the research for optimisation parameters provides a result (see below)

11% (14/124) of feeders are normally open points. Therefore, no correction will be applied, as previous days and weeks would not give more insights on the expected load over the next few days

Only 1 feeder did return any result as no observation data was available.

For the feeders with best parameters found, a volume parameter of 7 days provides the best results and an instant by instant correction over 4 weeks (28 days) provides best results as well. This is consistent with the results found on HV groups.

N.B. : A bespoke set of parameters is computed for each feeder.

		Number of HV Feeders (D10 Forecast)				
		Instant Correction Parameter				
		0	**7**	**21**	**28**	**Total**
Volume Correction	**0**	3	1	0	0	**4**
	1	6	0	6	6	**18**
	2	7	1	3	11	**22**
	7	19	0	10	36	**65**
	Total	**35**	**2**	**19**	**53**	**109**

Figure 30 – Distribution of HV feeders by set of parameters (D10 Forecast)

		Number HV Feeders (D4 Forecast)				
		Instant Correction Parameter				
		0	**7**	**21**	**28**	**Total**
Volume Correction	**0**	3	1			**4**
	1	7		7	5	**19**
	2	7	1	2	11	**21**
	7	18	1	19	27	**65**
	Total	**35**	**3**	**28**	**43**	**109**

Figure 31 - Distribution of HV feeders by set of parameters (D4 Forecast)

The table below show overall results before and after optimisation on HV Feeders for which a specific model was developed in the first phase of the project. The search for best parameters shows that the optimisation yield significant improvement in the accuracy of the forecast. D10 and D4 have similar performances.

	D10 Forecast	
	No Optimisation – i.e. historical data-based model only	Best Parameters - with real time correction applied
Min	10.4%	5.6%
Average	33.3%	18.9%
Median	21.7%	11.6%
Max	275.6%	208.1%
Feeders with MAPE* <20%	45	87
Feeders with MAPE* <25%	64	90
Feeders with MAPE* <30%	74	92
Total Feeders	104	104

	D4 Forecast	
	No Optimisation – i.e. historical data-based model only	Best Parameters - with real time correction applied
Min	10.2%	5.7%
Average	33.9%	18.7%
Median	22.2%	11.6%
Max	277.2%	216.6%
Feeders with MAPE* <20%	42	88
Feeders with MAPE* <25%	61	90
Feeders with MAPE* <30%	73	92
Total Feeders	104	104

The charts below show the improvement of the MAPE* metric over the period with or without optimisation for each Forecast

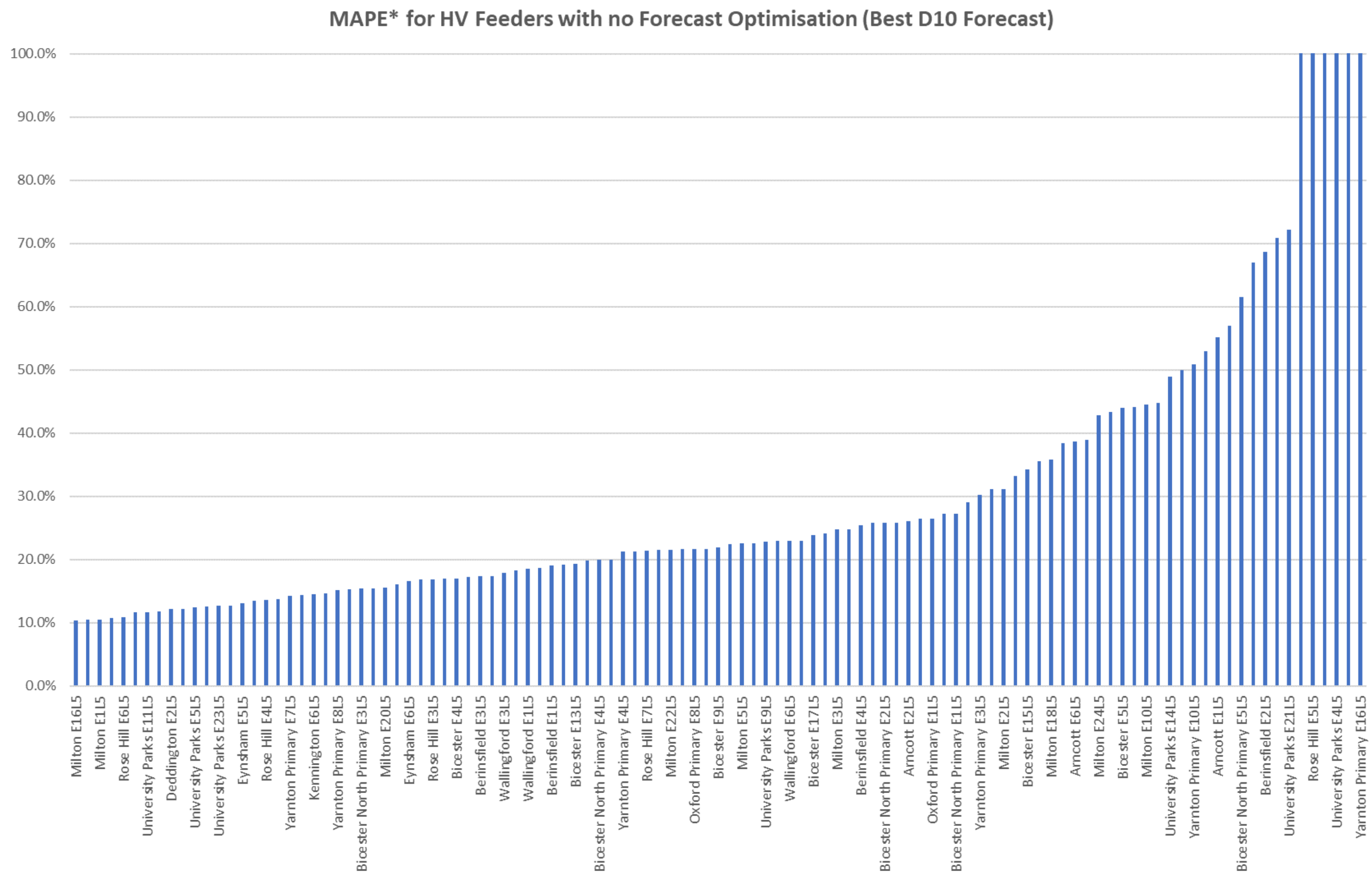


Figure 32 – MAPE for HV Feeders with no Forecast Optimisation (Best D10 Forecast)*

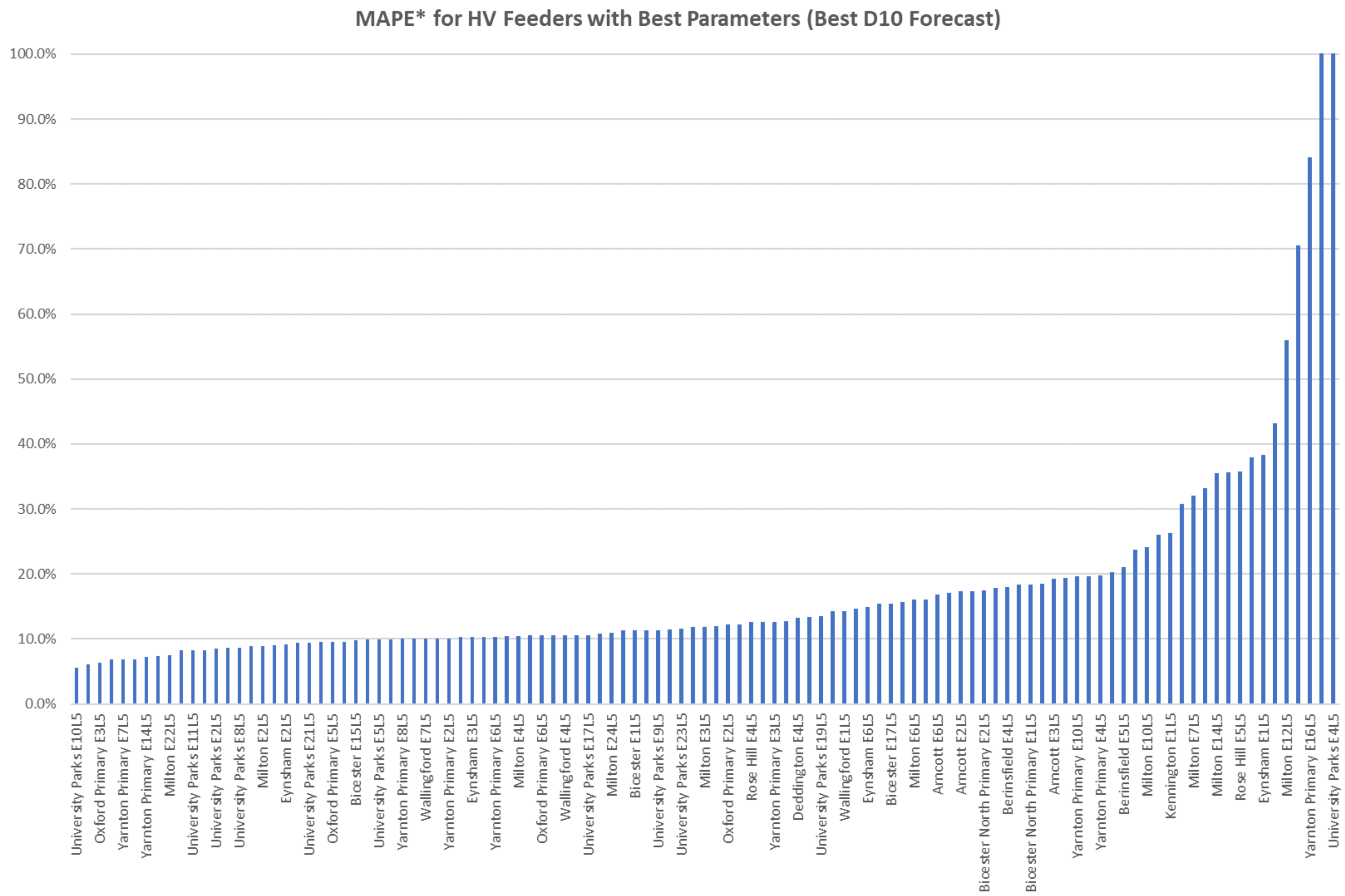


Figure 33 - MAPE for HV Feeders with Best Parameters (Best D10 Forecast)*

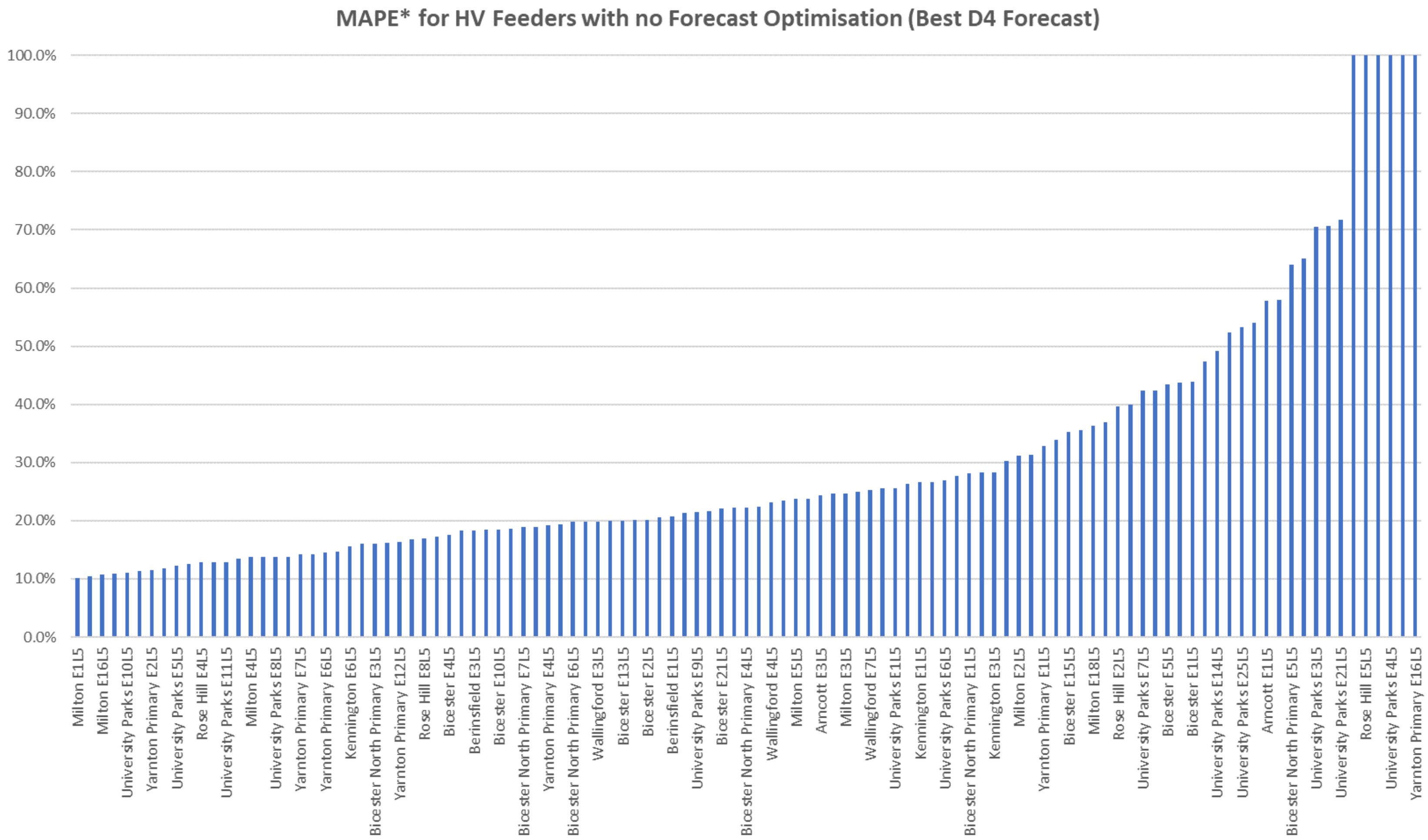


Figure 34 - MAPE for HV Feeders with no Forecast Optimisation (Best D4 Forecast)*

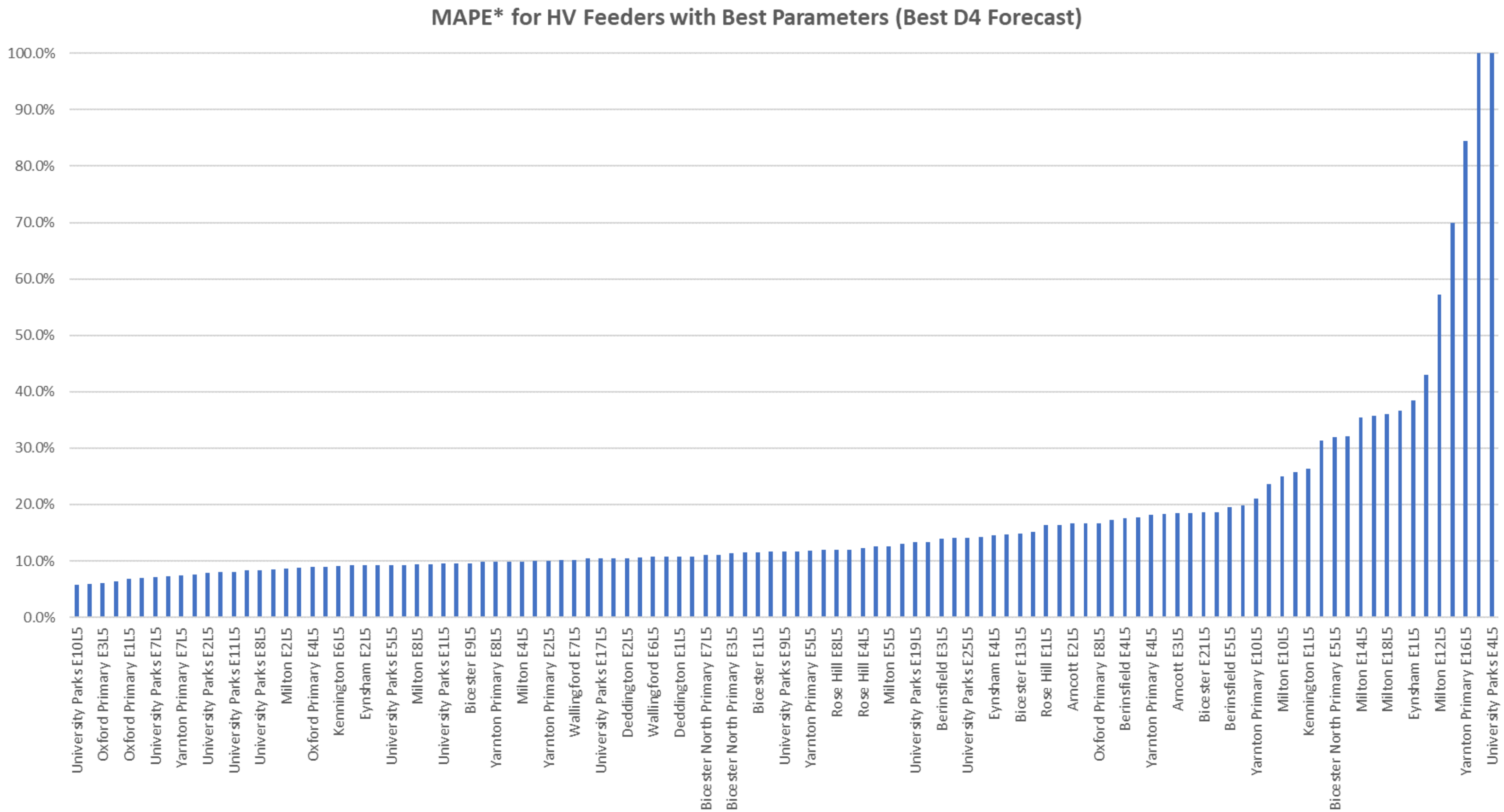


Figure 35 - MAPE for HV Feeders with Best Parameters (Best D4 Forecast)*

7 Evaluation of the optimised forecast

7.1 Description and objectives

The existing processes provide load forecasts at group, feeder and generators 4 times a day for the deterministic D10 and probabilistic D4. These 'original forecasts' are not been changed.

Following each Forecast run, the optimisation forecast task is triggered. It uses the parameters defined in the previous section, and applies the 2-step methodology based on the most recent forecast run and the observations from Nerda and Electralink.

7.2 Challenges and solutions

7.2.1 Impact of ongoing filtered data

In order to ensure the correction is applied to a representative data set with enough data points, a number of controls are being put in place.

On an ongoing basis, the data received from Nerda and Electralink is filtered through automatic processes which capture both data quality and Business rules. These processes ensure that the signal observed is of good quality.

For each stage of the forecast optimisation, there will need to be a minimum of 75% of non-filtered data points. Moreover, since the instant correction method is based on the outcomes of the volume correction, the instant correction will not be carried out if the volume correction has not been successful.

7.2.2 Impact of 2-day delay on instant-by-instant correction

On an ongoing basis, Electralink data is only available after 2 days (see dedicated section). Therefore, when applying both Volume and Instant correction methodologies, the observed underlying demand of the previous 2 days, composed on observed net demand from Nerda and settlements from Electralink, is not available.

The research of best parameters has been developed to take this effect into account. Therefore when evaluated the Optimised forecast, specific rules are applied:

Volume Correction: The 2-day delay means that the ongoing volume correction will look to correct with past data not available. For the optimised forecast, the previous 2 days are omitted in the volume correction part of the optimisation. Only the previous M days before D-2 are observed.

Instant Correction: The 2-day delay means that the ongoing instant correction will look to correct with past data not available for any instant in D+6 and D+6. However, for D to D+5 there is no impact. While the research of best parameters has not been changed to limit the impact on other days, the evaluation has to take into account the missing 2 days.

Different processes have been implemented for the evaluation, based on availability of data

For the ongoing evaluation of D to D+5, the optimization has not been changed

For the ongoing evaluation of D+6 and D+7, the optimization with look at N+1 weeks instead of N

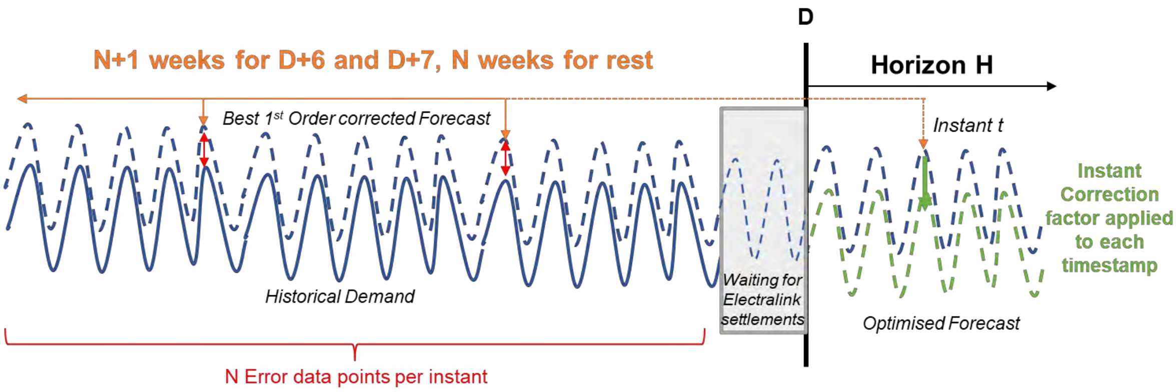


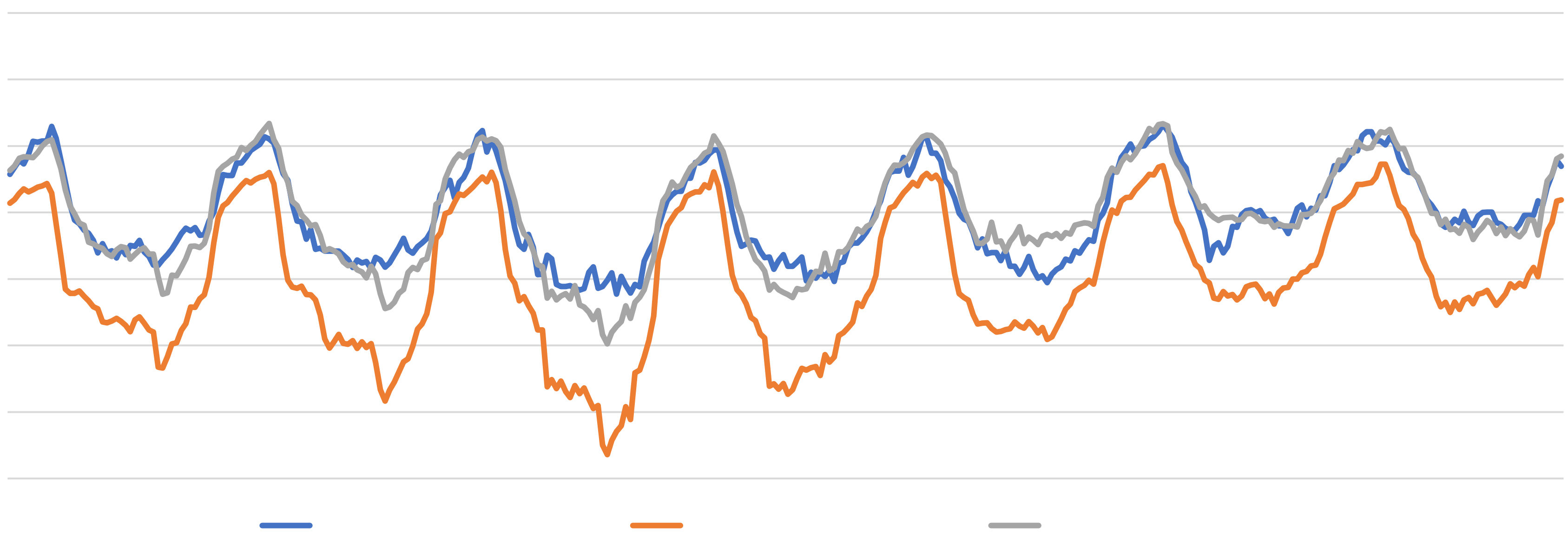
Figure 36 - Revised Instant Correction Methodology with Electralink Impact

7.3 Results of forecast optimisation (example on Rose Hill)

The chart below provides a view of the Original and Optimised forecast replayed in operational conditions against the measurements over a week in June 2022. It takes into account both the underlying demand forecast at group level and generation forecasts from all generators connected to Rose Hill to compute the net demand forecast.

Over this particular week, **MAPE** drops **from 31.5% for the Original Forecast** to **6.5% for the Optimised forecast**. This represents a drop of RMSE from 2.1MW to 0.5 MW (a **74% improvement**).

Result of Optimisation for Rose Hill over a week in June 2022



8 Results dissemination

8.1 User stories and identification of requirements

Apart from the clear need for Operational Load Forecasting capability within the TRANSITION project, such forecasting activities are an essential part of a broader suite of DSO competencies expected from DNO's in the near future. Such operational forecasting tools will enhance system monitoring and analysis capabilities in the DNO control room operational timeframes.

The current forecasting solution is used by TRANSITION stakeholders:

- The **SSEN Innovation team**
- **Administrators and Developers** of the tool
- The **whole system coordination (WSC) and PSA/PowerFactory tools used for automated network analysis workflows**

In the future though, **Control Room engineers** - dispatchers and back-office operators - are likely to become users too.

The list of user stories below has been identified through workshops with the identified groups of users for the solution and validated across the TRANSITION project team.

Discussions have then been conducted across the project team and internally at Sia Partners, to identify responses to the use cases, assess the corresponding workload of cost of development of the responses and prioritise use cases to be addressed within this phase of the project.

User stories that align with the objectives of the current innovation project to demonstrate feasibility, explore scalability and provide learnings for future projects have been prioritised while user stories related to the monitoring of processes and user accesses, and to the modification or addition of parameters and data, have been judged more relevant for operational solutions utilised in production environment on a full network scope. Similarly, user stories for Control Room engineers can be answered in a later stage of the project, when they become actual users of the tool.

Users	#	I want to ...	So I can...	proposed response
SSEN Innovation Team	a	Understand the accuracy of the forecast	Define the optimal window for flexibility requirements (study on past forecasts	study
	b	Understand the accuracy of the forecast	Have confidence in future forecasts	platform screen(s)
	c	Understand the impact of real-time data on optimised forecas	Understand the reduction of forecasting errors	study
	d	Understand the impact of real-time data on optimised forecas	Monitor the reduction of forecasting errors	platform screen(s)
	e	Understand the impact of real-time data on optimised forecas	Ensure the forecast adapt to network reconfiguration	study
	f	Understand the benefits of probabilistic weather forecasting	Design future forecasting requirements	study
	g	Investigate scalability of the solution	Plan for the deployment of the forecasting tool to entire SSEN network	study
SSEN tool administrator	h	Have access to user management functionalities	Manage authorisations and access logs	not prioritised
	i	Monitor the processes of the solution	Ensure everything is working fine	not prioritised
	j	Add new generators	Ensure generation forecast is accurate	not prioritised
	k	Monitor the integration with NeRDA	Investigate when there is an error	not prioritised
	l	Monitor the integration with Electralink	Investigate when there is an error	not prioritised
	m	Monitor the integration with WSC	Investigate when there is an error	not prioritised
	n	Modify specific parameters (ex PF)	Assess impacts on forecasts	not prioritised
Sia Partners developer	o	Have access to user management functionalities	Manage authorisations	existing feature
	p	Monitor the processes of the solution	Ensure everything is working fine	not prioritised
	q	Monitor the integration with NeRDA	Investigate when there is an error	existing feature
	r	Monitor the integration with Electralink	Investigate when there is an error	existing feature
	s	Monitor the integration with WSC	Investigate when there is an error	existing feature
WSC	t	Retrieve optimised forecasts for all assets	Use the most accurate forecasts in the PSA	updated endpoint
Control room engineer	u	Visualise the forecasted load on the network	Manage constraints and fault	not prioritised
	v	Visualise the forecasted load on the network	Dispatch flexibility in the future	not prioritised
operator	w	Visualise the forecasted load on the network	Dispatch flexibility in the future	not prioritised
Other	x	Forecast accuracy study for Ofgem		study

Figure 37 - User stories and proposed responses (dissemination options)

8.2 Dissemination options developed

8.2.1 Studies

To answer the user stories listed above, Sia Partners has developed five studies, included in this dissemination report. The scope of the studies is:

Study	Scope	User stories answered	Report section
1. Study on the impact of real-time data integration and forecast optimisation	Introduction to accuracy calculation Investigation of initial forecasts accuracy and optimised forecasts accuracy Comparison of both in various use cases	**c**	
2. Study on the deterioration of forecast accuracy across time and the impacts on flexibility dispatching	Investigation of forecast accuracy for different horizons of time (forecast at D+1, 2…10) in various use cases Conclusion on optimal horizons of forecasts depending on use cases Application of findings to flexibility dispatching	**a, x**	
3. Study on forecast accuracy in case of network reconfiguration	Identification of network changes use cases within the scope of the innovation project Investigation of the evolution of forecast quality for initial and optimised forecasts for these use cases	**e**	
4. Weather study	Investigate quality of the different weather forecasts used in the innovation project	**f**	
5. Learnings for scaling the forecasting solution to the whole SSEN network	Explore all other learnings of the pilot project (e.g., regarding interfaces, screens, infrastructure etc.) Draw conclusions and recommendations for scaling the solution to the whole SSEN network Include considerations about smart meter and low voltage data	**g**	

Table 1 - Studies developed for dissemination purposes

8.2.2 Platform screens

To answer user stories b and d listed above, it was proposed to update one existing User Interface/graphical screen of the forecasting platform front end solution, and to create one new screen.

Wireframes of these two screens have been designed by one of Sia Partners' UX expert (see below), refined through discussions with the whole project team, and finally developed by Sia Partners' wed developers.

Updated 'Load Forecast' Tab, in the 'Substation View' module

For Groups of transformers and feeders, the corresponding load charts have been updated to allow for the display of additional load curves:

- ✓ **Optimised forecast:**
  - Underlying demand (deterministic forecast)
  - Underlying demand min and max of probabilistic forecast
  - Net demand (deterministic forecast)
  - Net demand min and max of probabilistic forecast
  - Generation (deterministic forecast)
  - Generation min and max of probabilistic forecast
- ✓ **Real time:**
  - Underlying demand
  - Net demand
  - Generation

For Generation assets, the same has been done for the real time generation curve.

Taking into account the high number of load curves available for display on each chart (17 for groups of transformers and feeders), the screen has also been modified with the addition of side-panel allowing users to display/hide curves as they want, by simply selecting/unselecting them. This enables users to easily compare 'by eye' realised data with forecasts, and optimised forecasts with initial forecasts, thus answering use cases b and d.

New 'Forecast quality' Tab, in the 'Substation View' module

To provide a more detailed and quantified understanding of the quality of forecasts, and of the improvement brought by the introduction of real time data and optimised forecasts, a 'Forecast quality' tab has been added to the 'Substation view' module. This tab provides four quality indicators (RMSE, RMSE rolling, MAPE* and MAPE* rolling, see explanations below), for both deterministic and probabilistic forecasts, initial and optimised, at Group, feeder and generation asset levels, as described in the figure below.

These indicators are displayed in table with drop down lists enabling the users to see/hide the ones of interest for him/her (each block of four indicators – shown in grey in the figure below – can be shown or hidden), e.g., enabling the comparison of Group optimised deterministic forecast indicators with Group initial deterministic forecasts indicators.

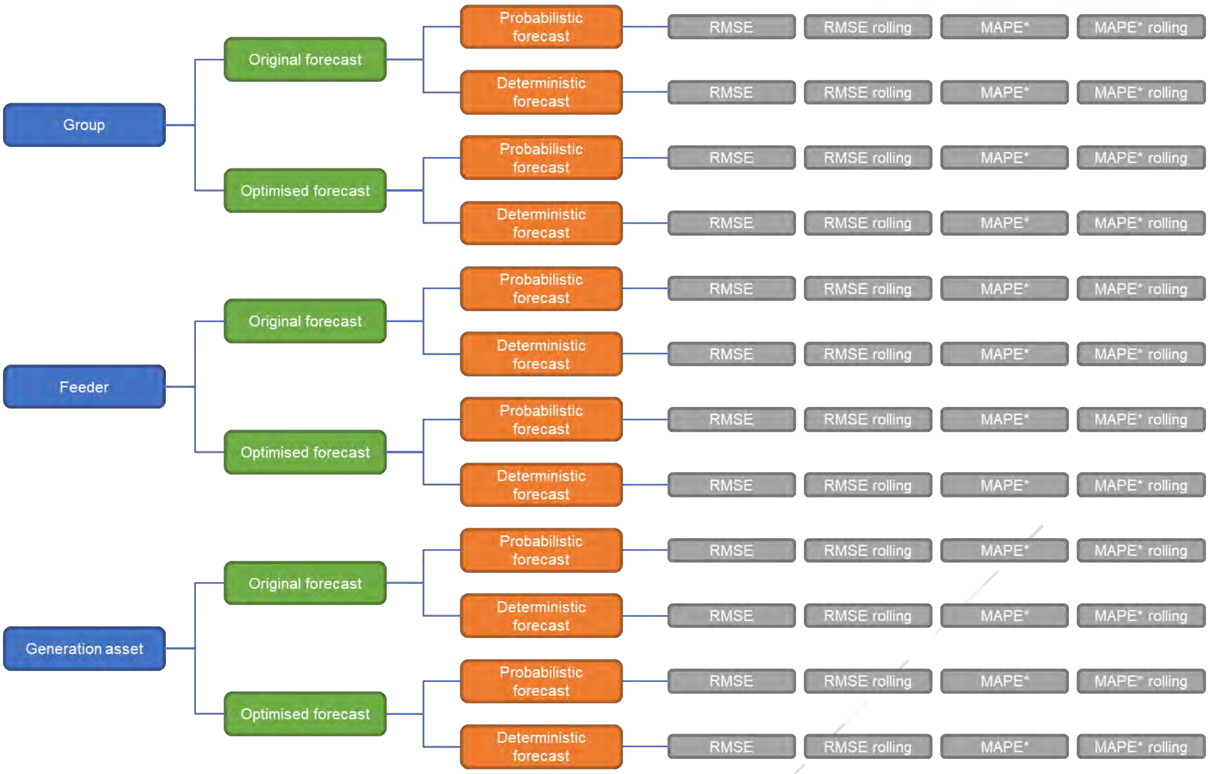


Figure 38 - Schematic view of the quality indicators produced

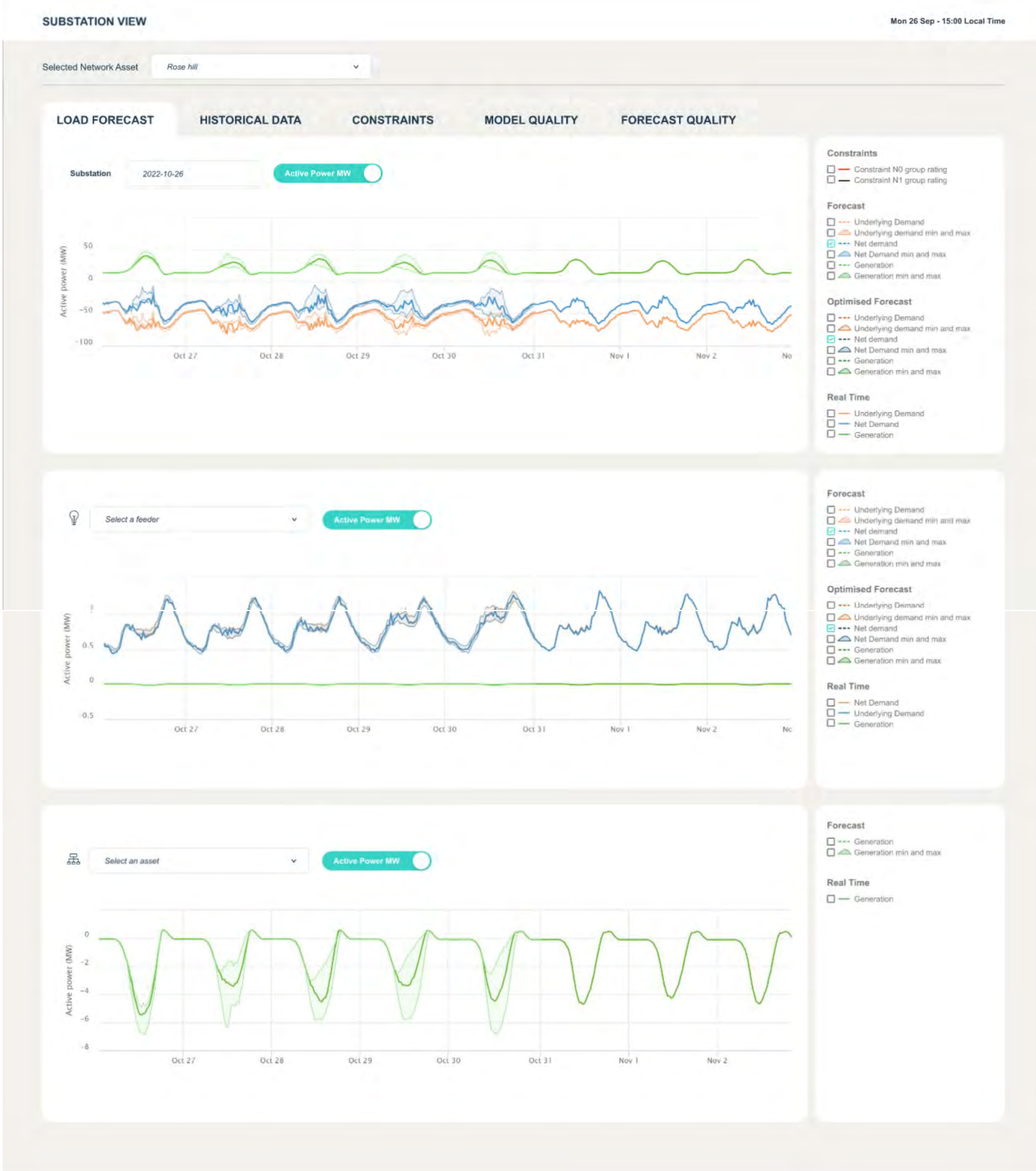


Figure 39 - Wireframe for the update of the 'Substation View' screen

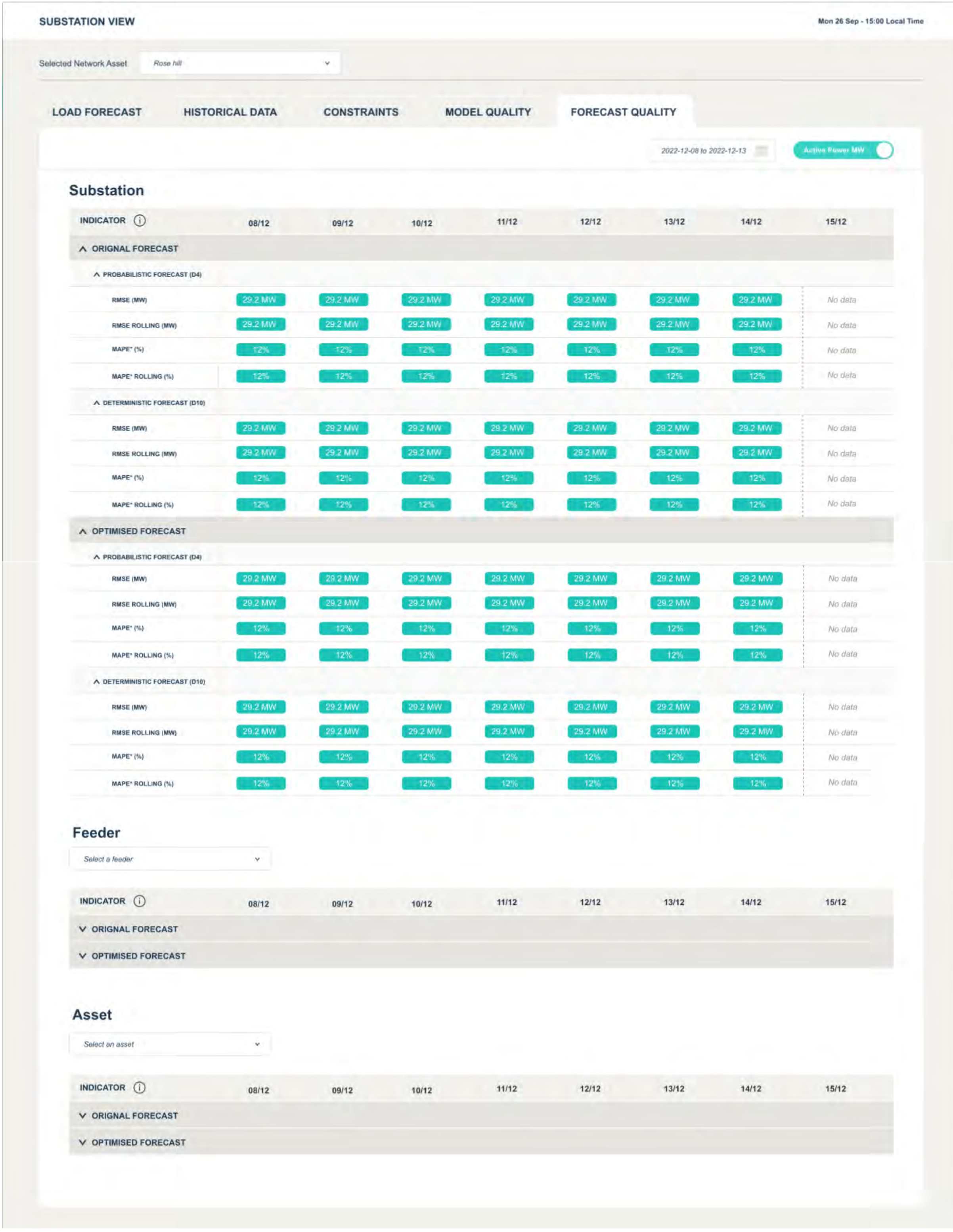


Figure 40 - Wireframe for the new 'Forecast quality' screen

The development of the user interface resulted in new screens for users, capturing the requirements developed above:

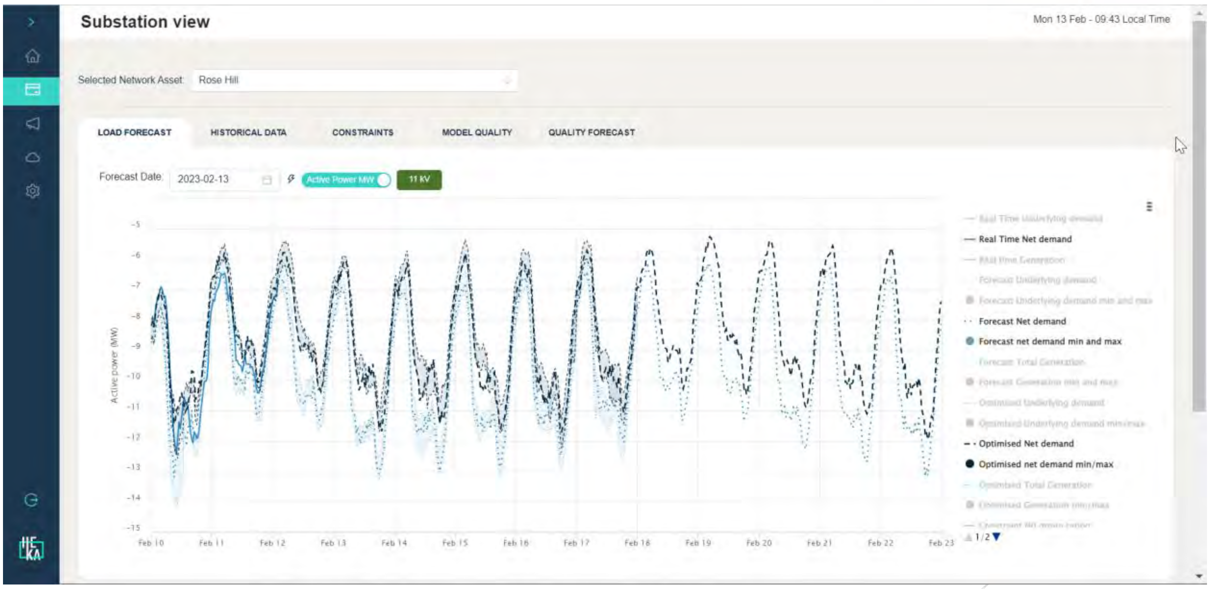


Figure 41 - User Interface update - Load Forecast Tab

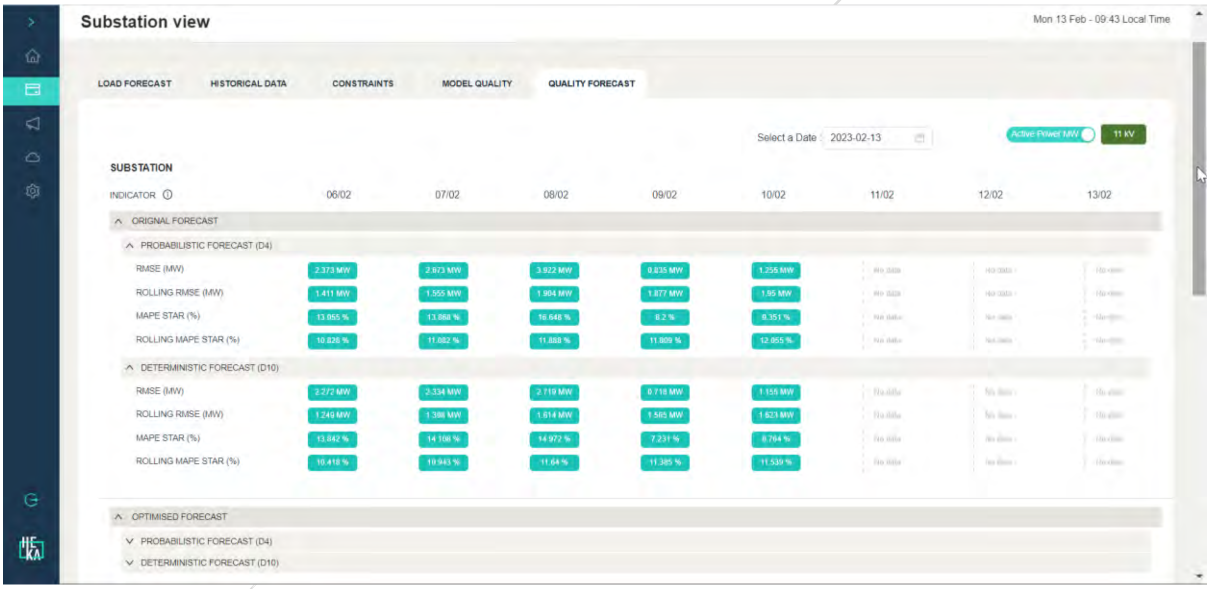


Figure 42 - User Interface update - Model Quality tab

9 Results on net demand forecast accuracy

The integration of observed data, both for demand and generation, allows to assess forecast quality. Four quality indicators are calculated for each asset (group, feeder, generator), and for all forecasts (D10 and D4, original and optimised).

The research of best parameters already computes those indicators to determine the best combination of parameters to use to optimise the **underlying demand forecast**. The forecast accuracy study focuses on the accuracy of the **net demand forecast** at group and feeder level, and **generation** for all generator assets. The study focuses on the month of June 2022.

These four indicators are:

For demand:
- MAPE (see section 7.4.3 for explanations), for values above 0.5MW
- RMSE (see section 7.4.3 for explanations)
- Rolling MAPE (average MAPE calculated over a 7-day period)
- Rolling RMSE (average RMSE calculated over a 7-day period)

For generation:
- MAPE* (see section 7.4.3 for explanations), using the installed capacity of the generator as reference
- RMSE (see section 7.4.3 for explanations)
- Rolling MAPE* (average MAPE calculated over a 7-day period)
- Rolling RMSE (average RMSE calculated over a 7-day period)

Results are presented in the sub-sections below.

9.1 Group net demand forecast accuracy

9.1.1 Summary view

Average values for the four accuracy indicators over the month of June 2022 are provided in the tables below for each group of primary substations in scope, for D4 and D10, for original and optimised forecasts, in order to allow for comparison.

	Original D10				Original D4			
	MAPE*	RMSE	Rolling MAPE*	Rolling RMSE	MAPE*	RMSE	Rolling MAPE*	Rolling RMSE
Arncott	108%	7.68	98%	6.19	109%	7.77	99%	6.24
Berinsfield	33%	6.33	27%	4.94	27%	5.01	27%	5.01
Bicester	16%	2.33	14%	1.95	16%	2.18	14%	1.81
Bicester North Primary	13%	1.58	12%	1.45	12%	1.55	12%	1.44
Deddington *	10%	0.03	10%	0.03	9%	0.03	9%	0.03
Eynsham	8%	0.45	8%	0.44	8%	0.43	8%	0.41
Kennington	18%	0.15	18%	0.15	17%	0.13	18%	0.14
Milton	10%	4.06	10%	3.70	9%	3.49	9%	3.18
Oxford Primary*	22%	4.78	22%	4.78	22%	4.83	22%	4.83
Rose Hill	26%	3.47	26%	3.42	27%	3.79	26%	3.69
University Parks	9%	1.79	9%	1.80	9%	1.77	9%	1.82
Yarnton Primary	15%	2.18	14%	2.02	15%	2.05	14%	1.87

Table 2 – Original forecast accuracy indicators for group demand in June 2022

	Optimised D10				Optimised D4			
	MAPE*	RMSE	Rolling MAPE*	Rolling RMSE	MAPE*	RMSE	Rolling MAPE*	Rolling RMSE
Arncott	85%	5.52	78%	4.52	85%	5.52	78%	4.51
Berinsfield All Feeders individually	25%	4.76	21%	3.88	27%	5.01	27%	5.01
Bicester	9%	1.26	9%	1.20	9%	1.21	9%	1.15
Bicester North Primary	9%	0.93	9%	1.04	8%	0.90	9%	1.00
Deddington All Feeders individually	10%	0.03	10%	0.03	9%	0.03	9%	0.03
Eynsham	6%	0.25	6%	0.24	6%	0.25	6%	0.24
Kennington	12%	0.07	14%	0.10	12%	0.07	14%	0.10
Milton	33%	24.80	31%	23.08	33%	24.45	31%	22.70
Oxford Primary	25%	5.85	25%	5.85	25%	5.96	25%	5.96
Rose Hill	8%	0.45	9%	0.62	8%	0.46	9%	0.62
University Parks	3%	0.25	4%	0.41	3%	0.23	4%	0.42
Yarnton Primary	10%	1.37	10%	1.31	10%	1.37	10%	1.31

Table 3 - Optimised forecast accuracy indicators for group demand

These high-level results show good forecast quality at 11kV group level, with only three groups with MAPE* above 25% for the original forecasts.
The optimisation process also brings significant value, significantly improving quality levels, especially for groups with high initial levels of error.

Detailed charts showing the evolution of the daily four metrics across the month of June 2022 for D10 original and optimised forecasts are provided for each group in Appendix 3 – Forecast Accuracy Primary substation charts over June 2022.
* Charts have not been produced for two groups (Deddington and Oxford Primary), as too much data was below 0.5MW and therefore filtered.

Below is the example for Rose Hill over June 2022. For each day of the month, the forecast quality indicators have been computed:

MAPE*: The MAPE* of the day D
RMSE: The MAPE* of the day D
Rolling MAPE*: The average MAPE calculated over the period between D and D-6 included
Rolling RMSE (average RMSE calculated over the period between D and D-6 included

9.1.2 Detailed view – Example on Rose Hill

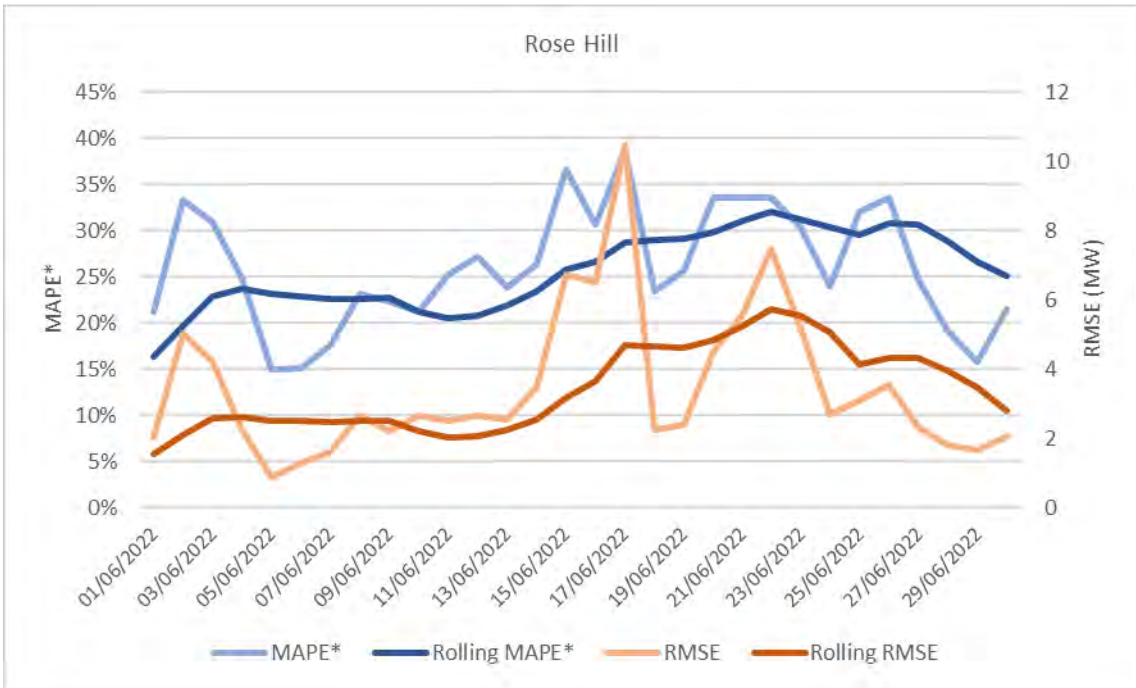


Figure 43 – Original forecast accuracy indicators for group demand in June 2022 – Rose Hill

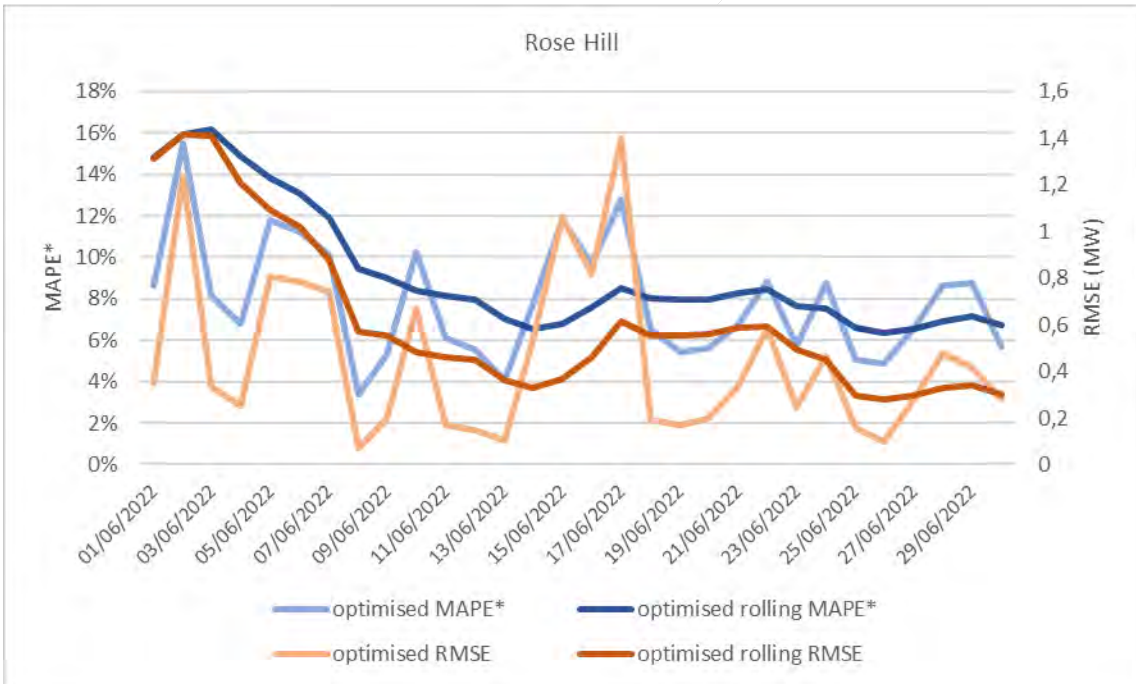


Figure 44 – Optimised forecast accuracy indicators for group demand in June 2022 – Rose Hill

9.2 Feeder net demand forecast accuracy

Average values for MAPE* and rolling MAPE* for the month of June 2022 are calculated for each feeder in scope.

An overview of the results is provided in the table below, for D4 and D10, and for original and optimised forecasts, in order to allow for comparison.

	Original				Optimised			
	D10		D4		D10		D4	
	MAPE*	Rolling MAPE*	MAPE*	Rolling MAPE*	MAPE*	Rolling MAPE*	MAPE*	Rolling MAPE*
Min	6.3%	6.6%	6.7%	7.1%	3.9%	5.2%	5.7%	7.1%
Average	31%	31%	32%	31%	24%	24%	30%	30%
Median	22%	22%	22%	23%	12%	14%	20%	20%
Max	149%	142%	222%	222%	533%	468%	222%	222%
Feeders with MAPE* <20%	60	58	61	59	100	97	68	66
Feeders with MAPE* <25%	80	77	79	77	112	109	84	82
Feeders with MAPE* <30%	94	93	92	94	117	116	96	98
Total number of feeders	145	145	145	145	145	145	145	145

Table 4 - Original and optimised forecast accuracy indicators for feeder demand in June 2022

The chart below provides the view of the average rolling MAPE* per feeder for the month of June 2022, without any optimisation

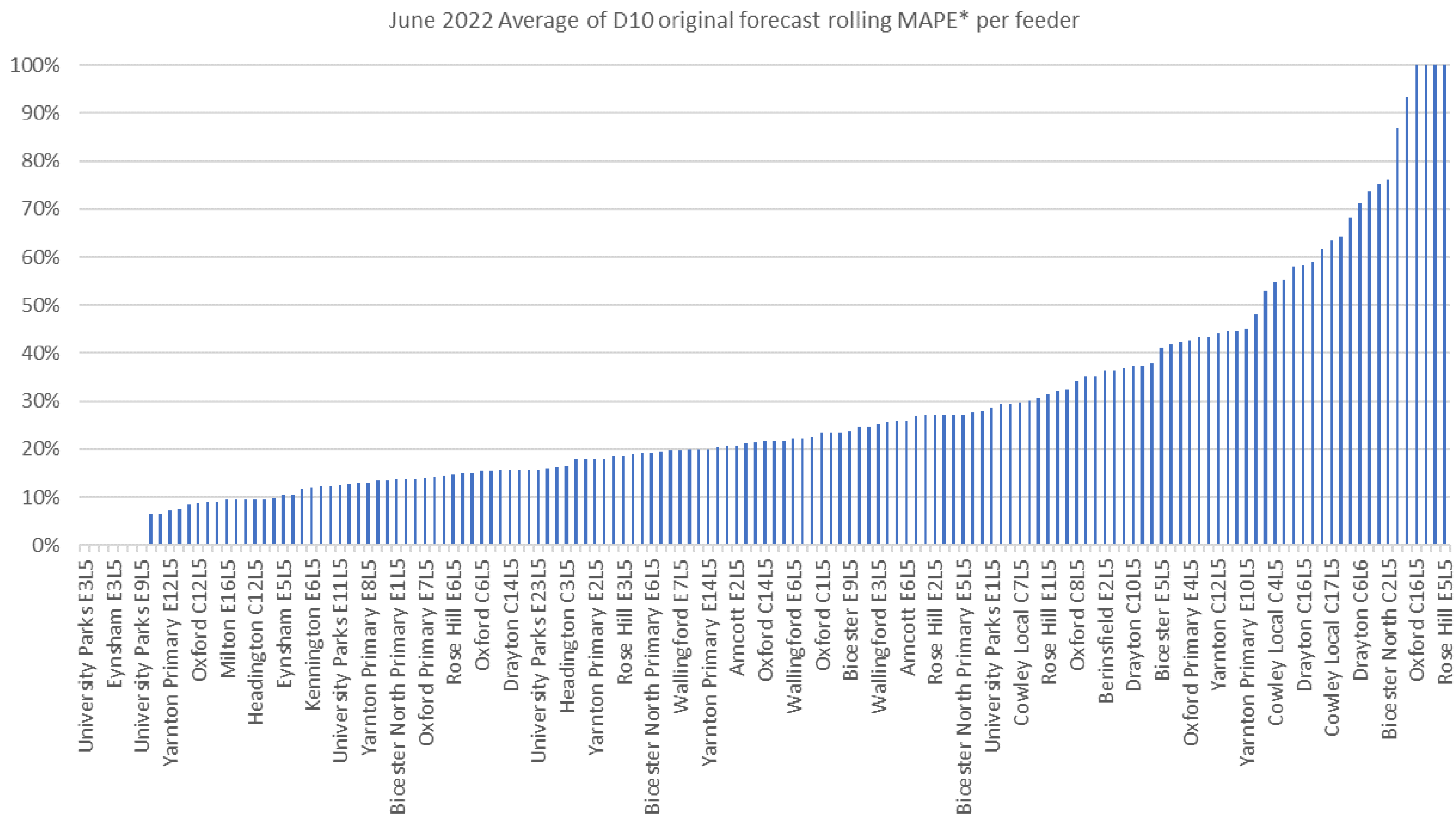


Figure 45 - June 2022 average of D10 original forecast rolling MAPE per feeder*

The chart below provides the view of the average rolling MAPE* per feeder for the month of June 2022, after optimisation with the best parameters

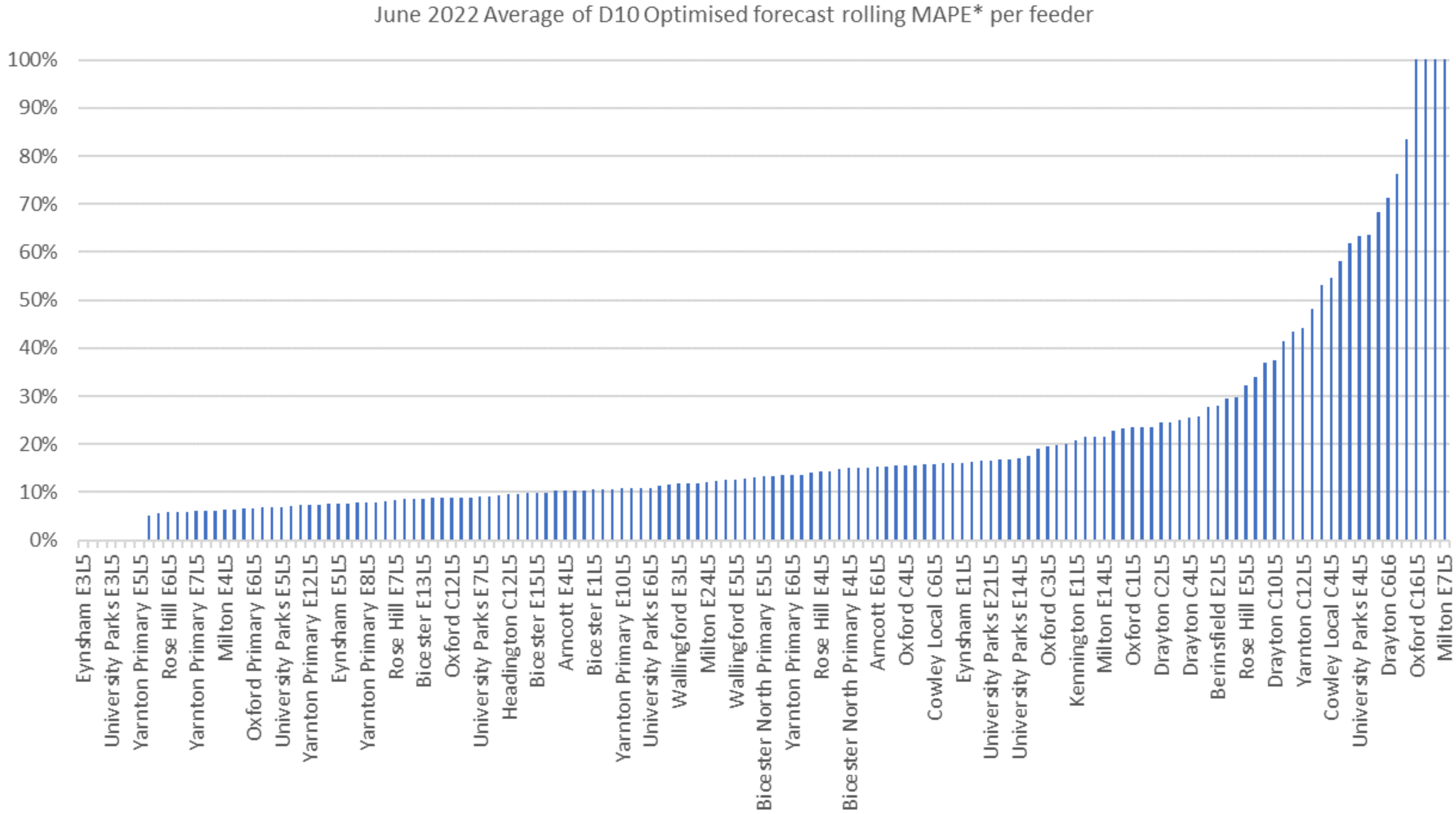


Figure 46 - June 2022 average of D10 optimised forecast rolling MAPE per feeder*

Forecast quality at feeder level is satisfying, with two thirds of feeder with MAPE* below 30%, considering the quality of input data.

Forecast optimisation also brings significant improvement at feeder level, with around 20 more feeders below the 30% error threshold (for D10).

9.3 Generation forecast accuracy

Average values for the four accuracy indicators for the month of June 2022 are provided in the tables below for each generation asset in scope, for D4 and D10 in order to allow for comparison.

N. B.: Generation forecasts are not optimised, so results are only provided for original forecasts.

	Original D10				Original D4			
	MAPE*	RMSE	Rolling MAPE*	Rolling RMSE	MAPE*	RMSE	Rolling MAPE*	Rolling RMSE
XXX EFW	29%	15,05	29%	15,15	29%	15,05	29%	15,15
XXX AD	3%	0,10	3%	0,13	3%	0,10	3%	0,13
XXX CHP	5%	0,32	5%	0,33	5%	0,32	5%	0,33
XXX Diesel	0%	0,01	0%	0,01	0%	0,01	0%	0,01
XXX PV	8%	0,87	8%	0,90	8%	0,85	8%	0,87
XXX Hydro	14%	0,06	14%	0,06	14%	0,06	14%	0,06
XXX PV	6%	2,23	7%	2,48	7%	2,42	7%	2,57
XXX PV	6%	1,23	6%	1,33	6%	1,29	6%	1,31
XXX PV	6%	0,46	7%	0,49	7%	0,49	7%	0,50
XXX PV	5%	0,90	5%	0,99	6%	0,95	6%	1,01
XXX PV	6%	0,47	6%	0,54	6%	0,50	6%	0,53
XXX PV	6%	1,26	6%	1,40	7%	1,37	7%	1,44
XXX PV B	5%	1,48	5%	1,62	6%	1,60	6%	1,73
XXX PV	4%	0,05	4%	0,06	4%	0,05	4%	0,06
XXX PV	6%	1,68	6%	1,85	6%	1,69	6%	1,79
XXX PV	6%	1,83	7%	2,09	7%	1,94	7%	2,14
XXX PV	6%	3,73	7%	4,13	7%	3,95	7%	4,21
XXX PV	6%	0,19	6%	0,21	6%	0,21	7%	0,21
XXX PV	6%	1,36	6%	1,46	6%	1,44	6%	1,50
XXX PV	5%	1,07	5%	1,15	6%	1,11	6%	1,19
XXX PV	6%	1,08	7%	1,19	7%	1,19	7%	1,24
XXX PV	5%	0,00	5%	0,00	6%	0,00	6%	0,00
XXX Hydro	20%	0,06	15%	0,05	257%	0,01	257%	0,01
XXX Synchronous	0%	0,00	0%	0,00	0%	0,00	0%	0,00
XXX PV	5%	0,88	6%	1,03	6%	0,95	6%	1,06
XXX Bio Gas	1055%	4,77	1051%	4,75	1055%	4,77	1051%	4,75
XXX Landfill	55%	7,14	55%	7,12	55%	7,14	55%	7,12
XXX Diesel	0%	0,00	0%	0,00	0%	0,00	0%	0,00
XXX CHP	27%	0,22	27%	0,23	27%	0,22	27%	0,23
XXX PV	4%	0,10	5%	0,11	5%	0,10	5%	0,11
XXX PV	5%	1,10	5%	1,22	6%	1,17	6%	1,24
XXX Windfarm	11%	0,84	11%	1,11	23%	1,72	21%	1,92

Table 5 – Generation forecast accuracy indicators

The average values of rolling MAPE*for D4 and D10 forecasts across June 2022 are also displayed on the chart below for each generation asset.

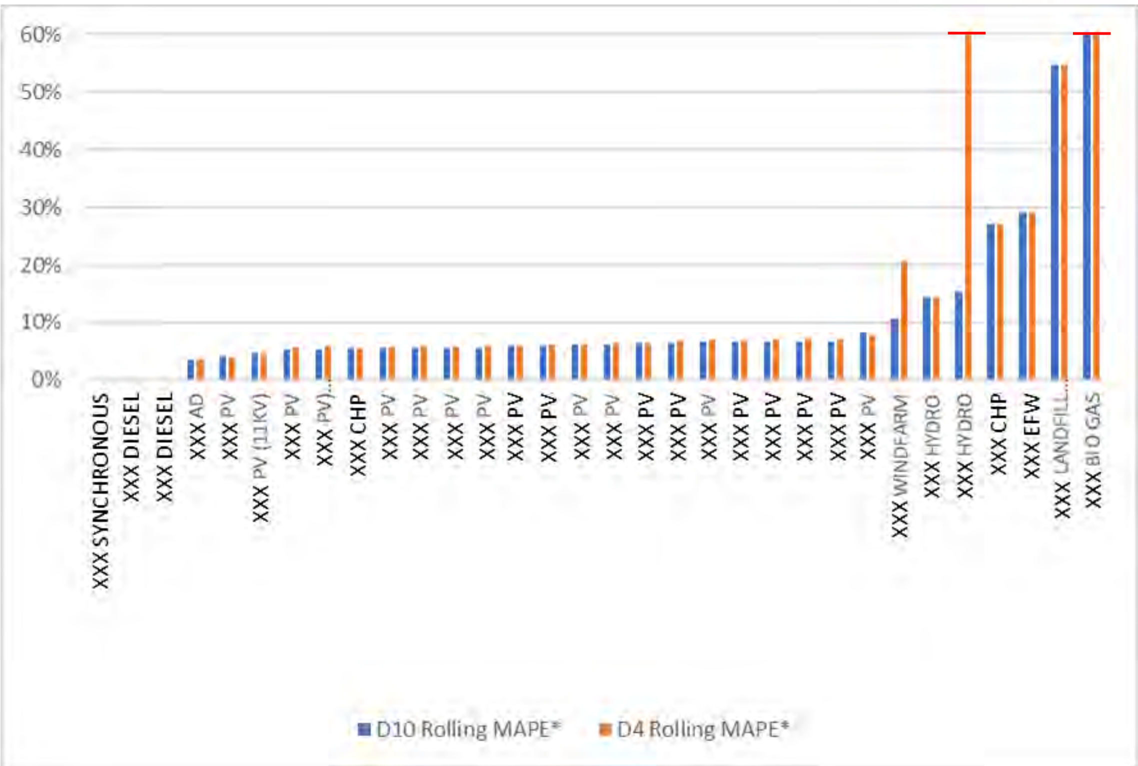


Figure 47 - Generation forecasts average rolling MAPE for June 2022*

Forecasting quality for generators is very good, especially for PV. Only seven assets have MAPE* higher than 10%, which are mainly market-driven assets, and only 2 are above 30%.

9.4 Impact of the forecast optimisation – other special cases

9.4.1 Very low values

The initial project had identified a number of feeders for which the load observations were very small. Due to Jitter factor and modelling requirements, it was not possible to determine a specific profile for those feeders. Instead, considering the limited variability around the Load data, it was decided to apply a constant model which correspond to the yearly average over the history of the feeder.

Below is the table summarising the feeders and the constant model applied:

Feeder Name	Constant Model applied
Milton E3L5	-0.075 MW
Milton E10L5	-0.070 MW
Milton E14L5	-0.055 MW
Yarnton Primary E5L5	-0.065 MW
Yarnton Primary E16L5	-0.045 MW

Following the research for best parameters, a set of parameters could be found for some of those feeders in order to capture the most recent behaviours from the Load observations. Below are the resulting tables of best parameters for each feeder.

	Forecast D10		Forecast D4	
HV Group (Forecast D10)	Volume Parameter	Instant Parameter	Volume Parameter	Instant Parameter
Milton E3L5	2	0	7	0
Milton E10L5	7	0	7	0
Milton E14L5	No Parameters found		No Parameters found	
Yarnton Primary E5L5	0	0	0	0
Yarnton Primary E16L5	7	0	7	0

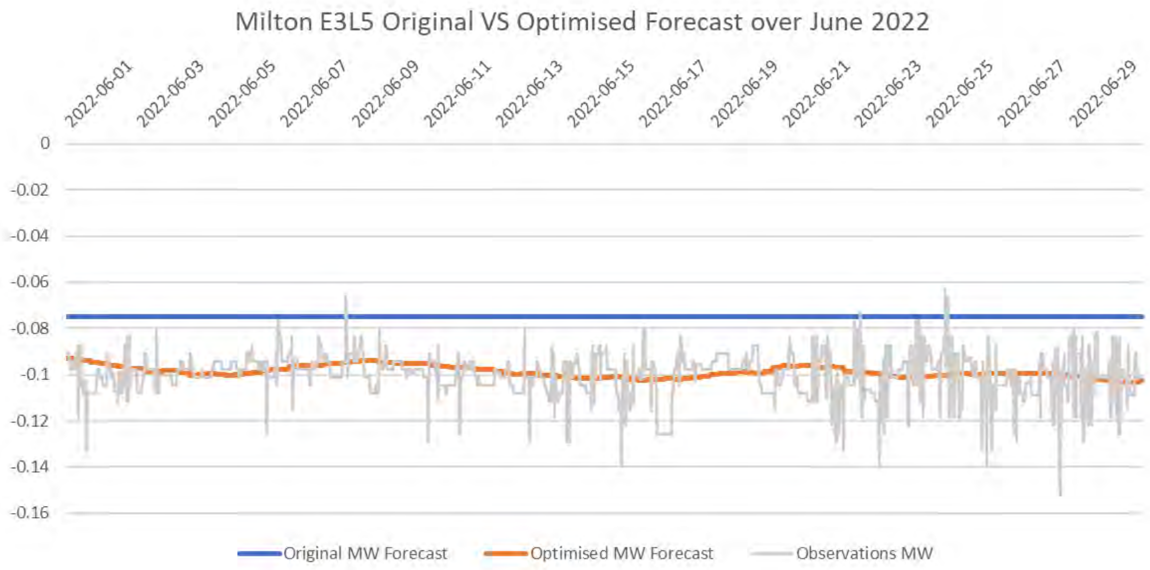
For 3 feeders, applying a new set of parameters would improve the demand load forecast of those feeders. For Milton E14L5, No parameters could be found due to a very limited amount of available data. Finally for Yarnton Primary E5L5, the best forecasts was found to be the original forecast.

Below are the Metrics tables without optimisation and with best parameters applied. Note that MAPE is not representative for those feeders

	Original Forecast D10			Original Forecast D4		
HV Group (Forecast D10)	MAPE	MAPE*	RMSE	MAPE	MAPE*	RMSE
Milton E3L5	-	24.8%	0.001	-	24.7%	0.001
Milton E10L5	-	44.5%	0.007	-	47.3%	0.012
Milton E14L5	-	35.5%	0.004	-	35.5%	0.004
Yarnton Primary E5L5	-	11.8%	0.004	-	11.8%	0.004
Yarnton Primary E16L5	-	1388.4%	0.002	-	1382.5%	0.002

	Optimised Forecast D10			Optimised Forecast D4		
HV Group (Forecast D10)	**MAPE**	**MAPE***	**RMSE**	**MAPE**	**MAPE***	**RMSE**
Milton E3L5	-	11.8%	<1kW	-	11.5%	<1kW
Milton E10L5	-	24.1%	0.006	-	25.0%	0.007
Milton E14L5	-	35.5%	0.004	-	35.5%	0.004
Yarnton Primary E5L5	-	11.8%	0.004	-	11.8%	0.004
Yarnton Primary E16L5	-	84.1%	<1kW	-	84.5%	<1kW

For example on Milton E3L5, over the month of June 2022, we can clearly see the positive impact of the Forecast optimisation



9.4.2 Continuous changes of demand behaviour

The initial project had identified a number of feeders for which the load observations were showing a gradual change of demand behaviour. This lead to poor model quality as the evaluation of the model would on a on a different demand behaviour than what the model was calibrated on.

Following the research for best parameters, a set of parameters could be found for all those feeders in order to capture the most recent behaviours from the Load observations. Below are the resulting tables of best parameters for each feeder.

	Forecast D10		Forecast D4	
HV Group (Forecast D10)	**Volume Parameter**	**Instant Parameter**	**Volume Parameter**	**Instant Parameter**
Bicester E5L5	1	21	1	21
Berinsfield E2L5	1	0	1	28
Milton E24L5	2	28	1	28
Rose Hill E5L5	7	0	7	0
University Parks E14L5	2	0	2	0
University Parks E21L5	1	28	2	21

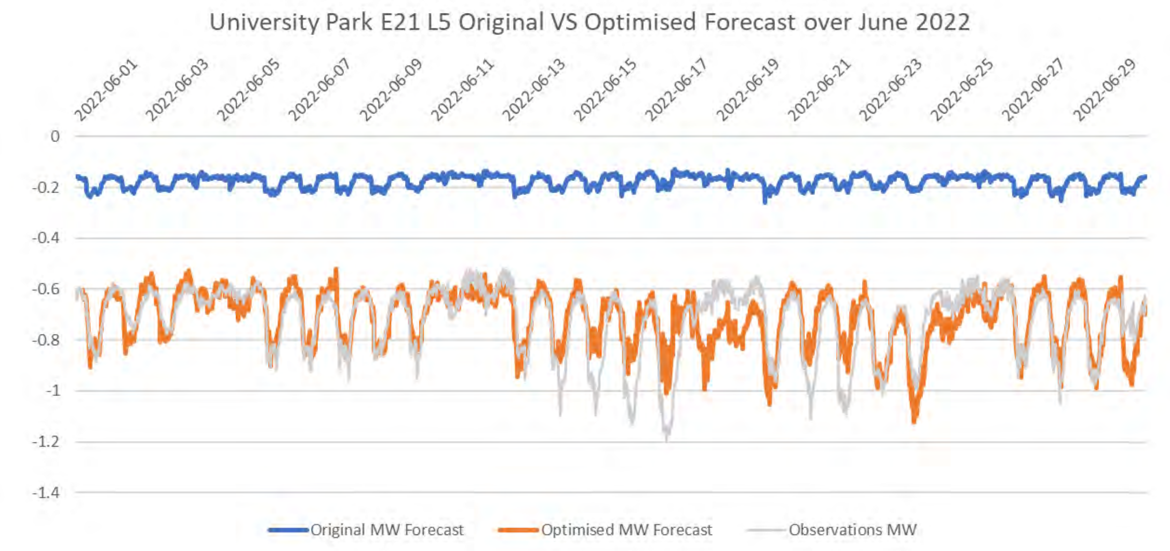
For 3 feeders, applying a new set of parameters would improve the demand load forecast of those feeders. For Milton E14L5, No parameters could be found due to a very limited amount of available data. Finally for Yarnton Primary E5L5, the best forecasts was found to be the original forecast.

Below are the Metrics tables without optimisation and with best parameters applied. Note that MAPE is not representative for those feeders

	Original Forecast D10			Original Forecast D4		
HV Group (Forecast D10)	MAPE	MAPE*	RMSE	MAPE	MAPE*	RMSE
Bicester E5L5	44.0%	44.0%	0.202	43.4%	43.4%	0.198
Berinsfield E2L5	-	68.7%	0.227	-	70.7%	0.242
Milton E24L5	42.8%	42.8%	0.186	42.4%	42.4%	0.184
Rose Hill E5L5	127.5%	127.5%	1.265	128.3%	128.3%	1.271
University Parks E14L5	48.9%	48.9%	0.035	49.2%	49.2%	0.036
University Parks E21L5	72.1%	72.1%	0.266	71.8%	71.8%	0.263

	Optimised Forecast D10			Optimised Forecast D4		
HV Group (Forecast D10)	MAPE	MAPE*	RMSE	MAPE	MAPE*	RMSE
Bicester E5L5	10.5%	10.5%	0.017	9.5%	9.5%	0.019
Berinsfield E2L5	-	43.2%	0.185	-	43.0%	0.19
Milton E24L5	11.0%	11.0%	0.018	10.8%	10.8%	0.016
Rose Hill E5L5	35.8%	35.8%	0.207	35.7%	35.7%	0.206
University Parks E14L5	30.8%	30.8%	0.021	32.1%	32.1%	0.022
University Parks E21L5	9.4%	9.4%	0.011	9.0%	9.0%	0.01

For example on University Park E21L5, over the month of June 2022, we can clearly see the positive impact of the Forecast optimisation



10 Forecast horizon accuracy

10.1 Understanding the degradation of the forecast with time horizon

For each half-hour of any given day D, the solution produces 11 forecasts. The first one is produced 10 days ahead of the given day (called D+10), the second 9 days ahead (D+9) … until the final forecast (called D) produced just a few hours before the given half-hour.

It is well known that the accuracy of weather forecasts decreases as the forecast horizon is further away in time. The objective of the present study is to determine whether this is equally true for the load forecasts produced by the solution.

The study has been conducted for the month of June 2022. For this month and for each group of primary substations, the eleven forecasts produced by the solution have been compared to the observed load by calculating and plotting the daily average relative error (MAPE).

The same study has been conducted for one PV generation asset, for the month of July 2022 (for data availability reasons). The eleven forecasts produced by the solution for the asset have been compared to the observed generation load recorded, by calculating and plotting the daily MAPE*.

The detailed results are provided for each substation over the month of June 2022 in Appendix 4 – Forecast Horizon Accuracy study. Below are the detailed results for Rose Hill. We can see from the optimised forecast charts that the optimisation realigns the forecast on a daily basis. There are clear benefits from the optimisation in general. We also realise that both in the original and optimised forecast, the error of D+10 compared to D is relatively similar.

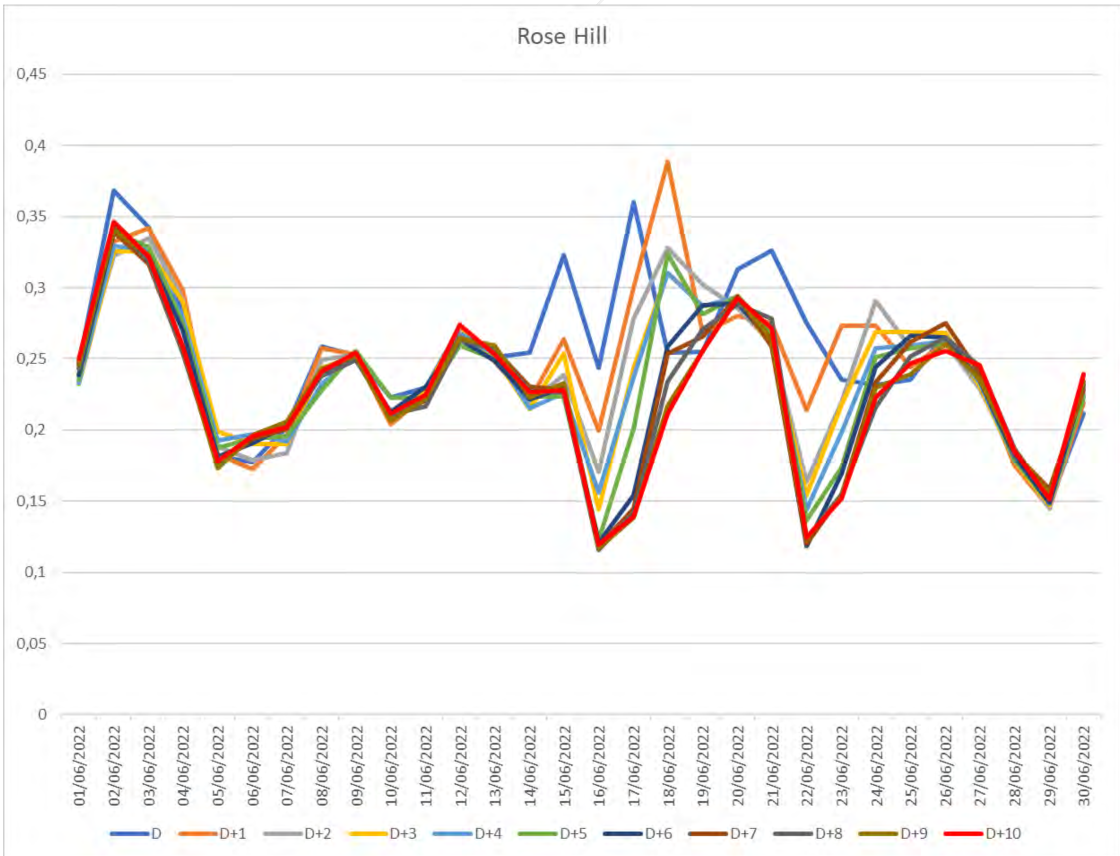


Figure 48 - Daily average relative error for all original forecasts in June 22 – Rose Hill

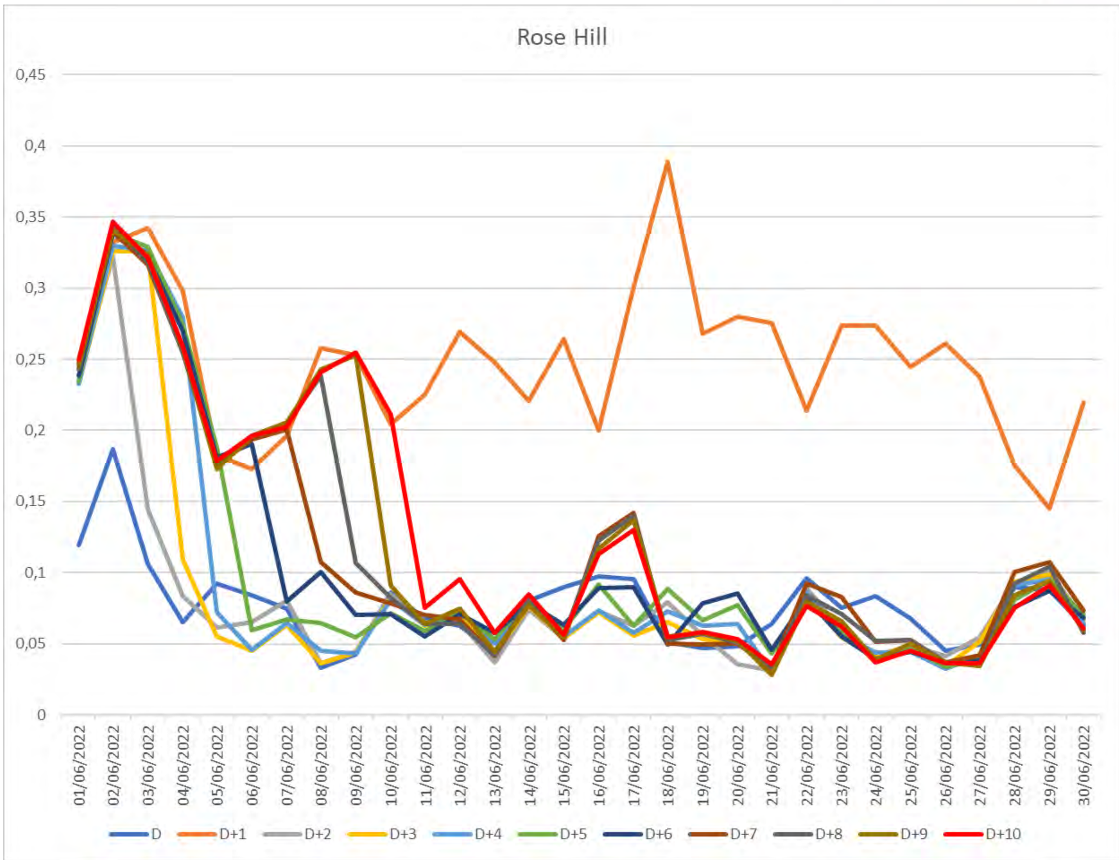


Figure 49 - Daily average relative error for all optimised forecasts in June 22 – Rose Hill

In the table below the results for each substation are summarised. It demonstrates that in general there is no obvious distortion of the forecast with horizon, i.e., D+10 forecasts are (except for Arncott whose D+10 forecast seems to be an outlier) not substantially of poorer quality than D+2, D+1 or D forecasts. This means that the 10 day forecast can be trusted and is of good enough quality to make business decisions.

It is also important to note that this study has been carried out over the single month of June 2022. Further studies can be carried out at various times of year to understand the dependencies to seasonality or special days such as Christmas etc.

The study also focuses independently on underlying demand forecast and generation forecast. For each different type of generation, a separate study could be undertaken retrospectively to understand the impact of horizon on the accuracy of the forecast. However, for non-weather dependent generators, the impacts of the markets could play a key role and would introduce further uncertainty.

	D	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+8	D+9	D+10
Arncott	30%	30%	30%	30%	29%	29%	29%	29%	29%	29%	69%
Berinsfield	31%	29%	28%	28%	28%	28%	28%	27%	27%	27%	27%
Bicester Primary	16%	14%	13%	12%	12%	12%	11%	11%	11%	11%	11%
Bicester North Primary	13%	13%	13%	13%	13%	14%	14%	14%	14%	14%	14%
Deddington	20%	19%	19%	19%	19%	18%	18%	18%	19%	19%	19%
Eynsham	8%	8%	8%	8%	8%	8%	7%	7%	7%	7%	7%
Kennington	19%	19%	18%	19%	19%	19%	19%	19%	19%	19%	19%
Milton	39%	38%	38%	37%	37%	37%	37%	37%	37%	36%	36%
Oxford Primary	27%	27%	26%	26%	26%	26%	25%	25%	25%	25%	25%
Rose Hill	25%	25%	24%	24%	24%	23%	23%	23%	23%	22%	23%
University Parks	8%	8%	8%	7%	7%	7%	7%	7%	7%	7%	7%
Yarnton	14%	13%	11%	11%	10%	10%	10%	10%	10%	9%	9%

Table 6 - Daily average error (MAPE) for all primary substations original forecasts in June 22

	D	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+8	D+9	D+10
Arncott	18%	19%	19%	20%	20%	20%	20%	20%	20%	21%	21%
Berinsfield	22%	20%	20%	21%	22%	22%	22%	23%	25%	26%	26%
Bicester Primary	9%	9%	9%	9%	9%	9%	8%	8%	8%	8%	8%
Bicester North Primary	8%	8%	7%	7%	8%	8%	8%	8%	9%	9%	9%
Deddington	19%	19%	19%	19%	19%	18%	18%	18%	19%	19%	19%
Eynsham	6%	7%	8%	9%	9%	10%	10%	10%	10%	10%	11%
Kennington	10%	11%	11%	12%	13%	13%	13%	13%	13%	13%	13%
Milton	6%	8%	10%	11%	13%	13%	14%	14%	15%	15%	16%
Oxford Primary	25%	25%	25%	25%	24%	24%	24%	24%	24%	24%	24%
Rose Hill	8%	25%	8%	8%	9%	10%	11%	11%	12%	12%	13%
University Parks	3%	4%	4%	5%	5%	5%	6%	5%	6%	6%	6%
Yarnton	10%	9%	9%	10%	10%	10%	10%	9%	8%	8%	8%

Table 7 - Daily average error for all primary substations optimised forecasts in June 22

	D	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+8	D+9	D+10
PV Asset	13%	14%	14%	15%	15%	15%	15%	14%	14%	14%	14%

Table 8 - Daily average error (MAPE) for one PV generator original forecasts in July 22*

Figure 50 - Daily average relative error for all original generation forecasts in July 22 – Elms PV

10.2 Impact on flexibility process

The conclusion of this study limited to a month of evaluation is encouraging and promising for the flexibility procurement process. For demand-driven flexibility needs, it shows that flexibility requirements can be assessed as soon as a first forecast is issued with a similar level of accuracy as what will be ultimately achieved with final forecasts.

11 Consideration for scalability and rollout to entire SSEN network

This section aims to provide some areas of considerations when deploying the Load Forecasting solution to a wider scale, ie BAU on a full-scale DNO license area Network. It is based on the observations made while running the existing version of the Load Forecasting solution.

The key difference is the number of assets, and associated volume of data to be managed at network level, compared to the Innovation project. Since the Load Forecasting solution has been developed for the purposes of the TRANSITION project, there are a number of decisions that have been made because of the specific size of the network.

11.1 Impact on the data model

Due to the limited number of substations, generators and feeders on the network, the data model is very simple. All assets of a same category are together and relationships between different voltage levels are included directly in the same table.

With a much larger network, it would be recommended to separate the assets in different voltage levels. This would allow data tables to be further split and limit the volume of each table.

11.2 Impact on data pipeline and performance of the processes

A study of the performance of the processes with the current scope has been performed to understand the availability of new forecasts.

Below is a view of the sequencing of a typical weather run and associated performances

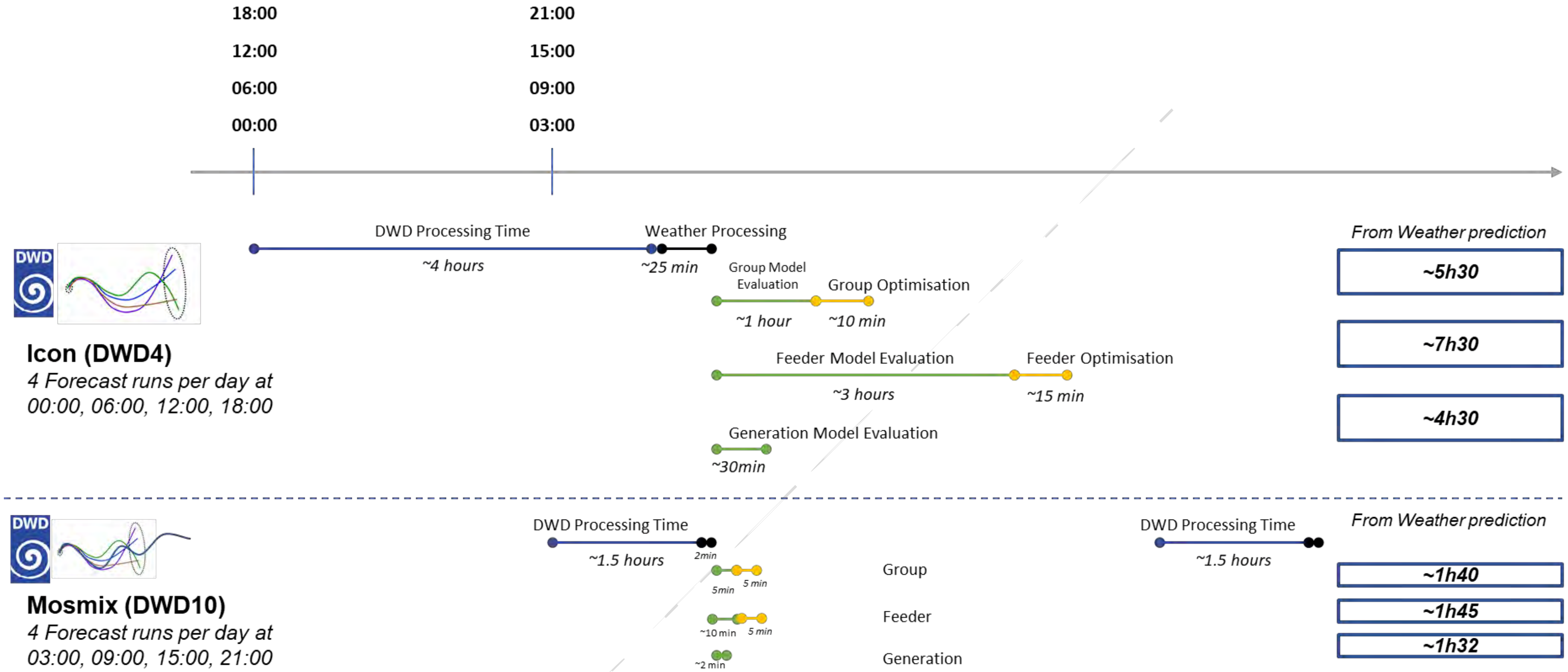


Figure 51 - Processes performance for a weather run

For each weather run (4 per day each weather source), the weather agency run the weather model to produce the new forecast and makes this new forecast available to its users. Depending on the weather this can take ~4h for Icon or ~1.5h for Mosmix. The weather data is then processed and integrated in the Load Forecasting solution.

Once the new weather run is uploaded it triggered the computation of new Load forecasts for groups, for feeders and for generators in parallel. The evaluation of the model against a new set of weather data is independent from the weather source. However, since Icon provide 40 different weather runs, the models are evaluated 40 times more than for Mosmix, which explains the large difference in computing time. Each demand forecast (group and feeder) is then optimised based on the methodology explained previously.

Over an entire day, this could lead to making new Load forecasts available fairly late if done for a full-scale network. Below is a daily view of the Load forecasts processes for TRANSITION

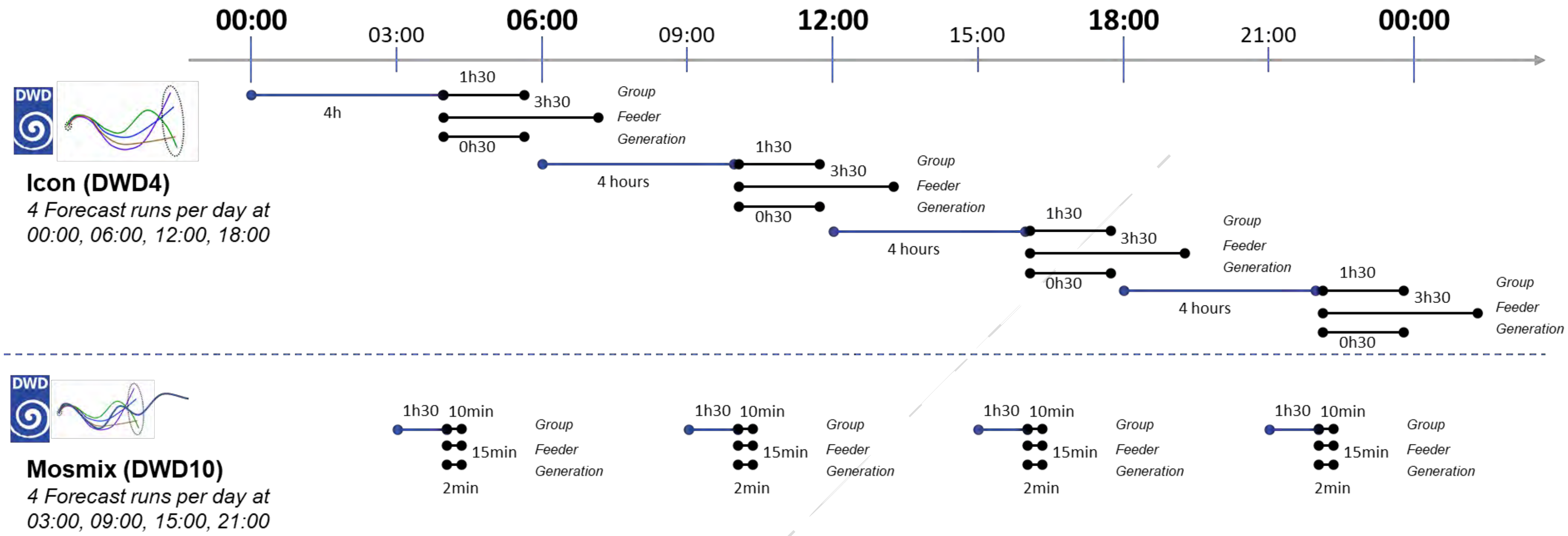


Figure 52 - Processes performance over a typical day

Two considerations are therefore important when scaling the Load Forecasting solution:

Is Probabilistic Forecast essential? If so, a strategy to reduce the overall time to produce load forecasts is necessary. Technical considerations around parallelisation can be implemented to minimise the total time, but this would not change the unit time to evaluate the models.

The individual optimised forecast activities could run after each original forecast. Currently this is set up as a different process, optimising the forecast for all assets at once – i.e. in parallel on a suitable scaled computational setup. It was design in this way to prevent impacting original forecasts. It would be more relevant and likely quicker to operate the optimisation straight after the computation of the original forecast.

The data pipeline for Real-time measurements (Nerda and Electralink) is very efficient as it stands. The 2-day delay imposed by Electralink makes the criticality of the process time less of a material issue. However, if Nerda were to be used as the single source of information for the measurements (PI Measurements of the generators for example), the half-hourly process would need to be carried out under 30minutes to ensure representativity.

11.3 Impact on performance of the interface

With a much larger network comes much larger tables in the database. The interfaces is connected to the database and interrogates the various tables to display the required information. Querying those much larger tables will undoubtedly result in longer waiting times for the information to be displayed. This was not the case with a very limited scope.

A strategy to clearly defined the most use cases will have to be undertaken to ensure the User interface provides an appropriate user experience. Otherwise, the use of the Load Forecasting solution would be minimised.

For instance, if 90% of the queries relate to the latest forecast available, it might be worth considering building specific views refreshed regularly, database maintenance processes, or dedicated smaller tables to ensure the information is provided quickly.

11.4 Impact on the weather sources

The study of the differences between weather sources (see Appendix 2 – Study on the importance of locational weather sources for Load Forecasting -Focus on weather forecast quality in Oxfordshire) has highlighted the importance of the weather source selection. As recommended in the study, it is advised to undertake a study of the most appropriate weather source based on the needs of the network. In the specific case of SSEN, the 2 license areas operated are widely different in terms of weather exposure, so the study will have to cover both geographies.

11.5 LV monitoring impact

As network operators transition to the DSO world, Load forecasting capabilities will become standard BAU activities which need to be rolled out for license area wide networks. It could also expand to include provision of forecasts for LV network asset levels. This would mean adding the secondary substation asset level, below the primary substation level and HV feeder head, currently the lowest voltage covered by the solution in TRANSITION. It would also require a suitable LV data source to train such models upon.

11.5.1 Data availability and volume of data

Data in kW, AMPS, kV and kVAr for both transformers and feeders of the secondary substations, at 30-minute time step would be required (same as the data used for primary substations).

Based on sample data provided by SSEN for three secondary substations (within which there has been some LV monitoring installed as part of TRANSITION work) it appears that data is likely to be available:

- ✓ At 10-minute time stamp
- ✓ For feeders (for each phase) and transformers
- ✓ In active, reactive and apparent power

Sufficient data therefore seems to be available to include LV assets in the solution, for a limited scope of the network, however.

This would mean an increase in volumes of data stored and processed (depending on the actual scope of data availability), which may require some work to make tasks and processes more efficient, but the solution has been designed to be scalable, so we do not foresee any blocking point in this regard.

11.5.2 Impacts on data model, forecasting models and data pipelines

The solution's data model has been designed to be fully scalable. Therefore, adding secondary substations in the solution would be very easily addressed by:

- Creating new transformers in the **Transformer_static** table (see below), with corresponding new groups (children of existing 11kV Groups).
- Creating new feeders in the **Feeder_static** table (see below), attached to the newly created groups of transformers.
- Creating new generation assets in the **Generation_static** table (see below), attached to the newly created groups of transformers and feeders.

Transformer_static
Transformer_id
Transformer_name
Group_id
Group_name
Parent_id
Parent_group_name
NRN
Lat
Lon
Voltage
Calib_var_mw
Calib_var_mva
Model_type_mw
Model_type_mva
Flip_threshold
Rating
UUID
Platform Display

Feeder_static
Feeder_id
Feeder_nrn
Feeder_name
Group_id
Voltage
Calib_var_mw
Calib_var_mva
Model_type_mw
Model_type_mva
Flip_threshold
Rating
UUID
Platform Display

Generation_static
Asset_id
Asset_name
Group_id
Feeder_id
Dedicated_cb_id
Date_commissioned
Voltage
Fuel
Capacity
Calib_var_mw
Calib_var_mva
Model_type_mw
Model_type_mva
CIM Name
UUID
Platform Display

Figure 53 - Illustration of static tables in the existing solution data model

Existing demand models would be able to use new LV demand data. Similarly, most of the new LV generation data should be covered by existing generation models (wind, solar, etc.). However, it is possible that some new types of fuels exist at LV level, which would require the introduction of new models.

LV feeder data comes from the SSEN's ENEIDA portal, which would need to be interfaced with the forecasting solution, in a similar manner to NeRDA and Electralink.

11.5.3 Quality

It is common to notice a loss of quality in load forecasts for lower levels of voltage, due to increased volatility (as multiplicity decreases on smaller lines) and sometimes lower data quality linked to low sensitivity sensors.

This would need to be confirmed with SSEN LV data but should be expected and accounted for when developing use cases.

11.5.4 Use Cases

LV load forecasting can answer a variety of use cases including:

- ✓ Detection of local consumption anomalies (e.g., fraud)
- ✓ Modelling consumption and/or generation at large consumers' sites (retailer, hospital, industry etc.)
- ✓ Modelling new household power usage (heat pumps, EV charging, PV)

12 Conclusion

The connection of real-time data has proven to be extremely valuable in refining the production of the Load forecast at each substation and feeder level. Using internal network load data from the network allows to compare the outputs of original forecast with actual measurements on the ground. On the other side, the connection to Electralink is deemed accurate but the 2 day delay introduce complexity and relative latency in the optimisation. This causes the optimisation to constantly be behind its optimal position.

Yet the project has demonstrated great results in the improvement of the accuracy of the net demand forecast at HV group and feeder levels. It further confirms the importance of real-time data but recognises that a model calibrated on a historical data set is still required to provide optimal outputs.

The demand forecast capabilities implemented are strong and reliable. The week ahead forecast accuracy is of good quality and opens a lot of opportunities for securing flexibility. At a larger scale the solution will need to be adapted thought to ensure speed of provision of the load forecast at each point of the network in a robust manner.

13 Appendixes

In this section are a number of tables and charts resulting from the studies undertaken as part of the project.

13.1 A1 – Tables of optimisation parameters

		Arncott – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 31.2% MAPE* = 31.2% RMSE = 1.133MW	MAPE = 34.3% MAPE* = 34.3% RMSE = 1.671MW	MAPE = 32.4% MAPE* = 32.4% RMSE = 1.292MW	MAPE = 32% MAPE* = 32% RMSE = 1.238MW
	1	MAPE = 21.9% MAPE* = 21.9% RMSE = 0.676MW	MAPE = 24.6% MAPE* = 24.6% RMSE = 1.042MW	MAPE = 23.1% MAPE* = 23.1% RMSE = 0.786MW	MAPE = 22.8% MAPE* = 22.8% RMSE = 0.75MW
	2	MAPE = 21.8% MAPE* = 21.8% RMSE = 0.665MW	MAPE = 24.5% MAPE* = 24.5% RMSE = 1.032MW	MAPE = 23% MAPE* = 23% RMSE = 0.775MW	MAPE = 22.7% MAPE* = 22.7% RMSE = 0.739MW
	7	MAPE = 21.4% MAPE* = 21.4% RMSE = 0.628MW	MAPE = 24.1% MAPE* = 24.1% RMSE = 0.989MW	MAPE = 22.6% MAPE* = 22.6% RMSE = 0.738MW	MAPE = 22.3% MAPE* = 22.3% RMSE = 0.702MW

		Arncott – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 32.9% MAPE* = 32.9% RMSE = 1.232MW	MAPE = 35.8% MAPE* = 35.8% RMSE = 1.763MW	MAPE = 34% MAPE* = 34% RMSE = 1.382MW	MAPE = 33.6% MAPE* = 33.6% RMSE = 1.327MW
	1	MAPE = 22.3% MAPE* = 22.3% RMSE = 0.7MW	MAPE = 24.4% MAPE* = 24.4% RMSE = 1.054MW	MAPE = 23% MAPE* = 23% RMSE = 0.799MW	MAPE = 22.7% MAPE* = 22.7% RMSE = 0.763MW
	2	MAPE = 22.1% MAPE* = 22.1% RMSE = 0.683MW	MAPE = 24.2% MAPE* = 24.2% RMSE = 1.036MW	MAPE = 22.8% MAPE* = 22.8% RMSE = 0.782MW	MAPE = 22.5% MAPE* = 22.5% RMSE = 0.745MW
	7	MAPE = 21.6% MAPE* = 21.6% RMSE = 0.642MW	MAPE = 23.6% MAPE* = 23.6% RMSE = 0.988MW	MAPE = 22.3% MAPE* = 22.3% RMSE = 0.741MW	MAPE = 21.9% MAPE* = 21.9% RMSE = 0.706MW

Berinsfield All Feeders individually – D10 Forecast					
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 21.3% MAPE* = 21.3% RMSE = 3.398MW	MAPE = 21.8% MAPE* = 21.8% RMSE = 3.763MW	MAPE = 21.5% MAPE* = 21.5% RMSE = 3.424MW	MAPE = 21.4% MAPE* = 21.4% RMSE = 3.392MW
	1	MAPE = 14.3% MAPE* = 14.3% RMSE = 1.886MW	MAPE = 14.6% MAPE* = 14.6% RMSE = 2.188MW	MAPE = 13.9% MAPE* = 13.9% RMSE = 1.941MW	MAPE = 13.9% MAPE* = 13.9% RMSE = 1.912MW
	2	MAPE = 14.4% MAPE* = 14.4% RMSE = 1.964MW	MAPE = 14.9% MAPE* = 14.9% RMSE = 2.238MW	MAPE = 14.2% MAPE* = 14.2% RMSE = 1.984MW	MAPE = 14.2% MAPE* = 14.2% RMSE = 1.956MW
	7	MAPE = 13.6% MAPE* = 13.6% RMSE = 1.617MW	MAPE = 13.8% MAPE* = 13.8% RMSE = 1.932MW	MAPE = 13.1% MAPE* = 13.1% RMSE = 1.688MW	MAPE = 13.1% MAPE* = 13.1% RMSE = 1.656MW

Berinsfield All Feeders individually – D4 Forecast					
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 22.3% MAPE* = 22.3% RMSE = 3.643MW	MAPE = 22.9% MAPE* = 22.9% RMSE = 3.967MW	MAPE = 22.5% MAPE* = 22.5% RMSE = 3.664MW	MAPE = 22.5% MAPE* = 22.5% RMSE = 3.633MW
	1	MAPE = 14.4% MAPE* = 14.4% RMSE = 1.902MW	MAPE = 14.8% MAPE* = 14.8% RMSE = 2.192MW	MAPE = 14.3% MAPE* = 14.3% RMSE = 1.963MW	MAPE = 14.2% MAPE* = 14.2% RMSE = 1.933MW
	2	MAPE = 14.4% MAPE* = 14.4% RMSE = 1.966MW	MAPE = 15% MAPE* = 15% RMSE = 2.233MW	MAPE = 14.4% MAPE* = 14.4% RMSE = 2.003MW	MAPE = 14.3% MAPE* = 14.3% RMSE = 1.971MW
	7	MAPE = 13.7% MAPE* = 13.7% RMSE = 1.658MW	MAPE = 14.1% MAPE* = 14.1% RMSE = 1.952MW	MAPE = 13.5% MAPE* = 13.5% RMSE = 1.726MW	MAPE = 13.4% MAPE* = 13.4% RMSE = 1.693MW

		Bicester Primary – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 14.2% MAPE* = 14.2% RMSE = 4.452MW	MAPE = 14.1% MAPE* = 14.1% RMSE = 4.605MW	MAPE = 13.9% MAPE* = 13.9% RMSE = 4.523MW	MAPE = 13.9% MAPE* = 13.9% RMSE = 4.504MW
	1	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.054MW	MAPE = 8% MAPE* = 8% RMSE = 1.104MW	MAPE = 7.8% MAPE* = 7.8% RMSE = 1.031MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.022MW
	2	MAPE = 8% MAPE* = 8% RMSE = 1.034MW	MAPE = 7.9% MAPE* = 7.9% RMSE = 1.085MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.011MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.002MW
	7	MAPE = 7.8% MAPE* = 7.8% RMSE = 0.993MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.038MW	MAPE = 7.5% MAPE* = 7.5% RMSE = 0.963MW	MAPE = 7.5% MAPE* = 7.5% RMSE = 0.953MW

		Bicester Primary – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 12.6% MAPE* = 12.6% RMSE = 3.604MW	MAPE = 12.6% MAPE* = 12.6% RMSE = 3.905MW	MAPE = 12.4% MAPE* = 12.4% RMSE = 3.661MW	MAPE = 12.3% MAPE* = 12.3% RMSE = 3.632MW
	1	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.084MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.317MW	MAPE = 7.8% MAPE* = 7.8% RMSE = 1.084MW	MAPE = 7.8% MAPE* = 7.8% RMSE = 1.057MW
	2	MAPE = 8% MAPE* = 8% RMSE = 1.051MW	MAPE = 8% MAPE* = 8% RMSE = 1.301MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.052MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.025MW
	7	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.975MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.221MW	MAPE = 7.4% MAPE* = 7.4% RMSE = 0.975MW	MAPE = 7.4% MAPE* = 7.4% RMSE = 0.948MW

		Bicester North Primary – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 9.1% MAPE* = 9.1% RMSE = 1.238MW	MAPE = 9.4% MAPE* = 9.4% RMSE = 1.33MW	MAPE = 9% MAPE* = 9% RMSE = 1.241MW	MAPE = 9% MAPE* = 9% RMSE = 1.229MW
	1	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.162MW	MAPE = 8.6% MAPE* = 8.6% RMSE = 1.279MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.177MW	MAPE = 8% MAPE* = 8% RMSE = 1.162MW
	2	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.146MW	MAPE = 8.5% MAPE* = 8.5% RMSE = 1.265MW	MAPE = 8% MAPE* = 8% RMSE = 1.162MW	MAPE = 8% MAPE* = 8% RMSE = 1.147MW
	7	MAPE = 7.6% MAPE* = 7.6% RMSE = 1.001MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.111MW	MAPE = 7.6% MAPE* = 7.6% RMSE = 1.012MW	MAPE = 7.5% MAPE* = 7.5% RMSE = 0.998MW

		Bicester North Primary – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 8.9% MAPE* = 8.9% RMSE = 1.173MW	MAPE = 9% MAPE* = 9% RMSE = 1.236MW	MAPE = 8.7% MAPE* = 8.7% RMSE = 1.148MW	MAPE = 8.7% MAPE* = 8.7% RMSE = 1.138MW
	1	MAPE = 8% MAPE* = 8% RMSE = 1.102MW	MAPE = 8.2% MAPE* = 8.2% RMSE = 1.18MW	MAPE = 7.8% MAPE* = 7.8% RMSE = 1.079MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.069MW
	2	MAPE = 7.9% MAPE* = 7.9% RMSE = 1.06MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.139MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.038MW	MAPE = 7.6% MAPE* = 7.6% RMSE = 1.027MW
	7	MAPE = 7.5% MAPE* = 7.5% RMSE = 0.935MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 1.01MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.912MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.901MW

		Deddington All Feeders individually – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.064MW	MAPE = 10.6% MAPE* = 10.6% RMSE = 0.07MW	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.064MW	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.063MW
	1	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.065MW	MAPE = 10.7% MAPE* = 10.7% RMSE = 0.071MW	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.066MW	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.065MW
	2	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.063MW	MAPE = 10.6% MAPE* = 10.6% RMSE = 0.069MW	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.064MW	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.063MW
	7	MAPE = 9.8% MAPE* = 9.8% RMSE = 0.06MW	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.066MW	MAPE = 9.8% MAPE* = 9.8% RMSE = 0.061MW	MAPE = 9.7% MAPE* = 9.7% RMSE = 0.06MW

		Deddington All Feeders individually – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.08MW	MAPE = 10.6% MAPE* = 10.6% RMSE = 0.086MW	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.08MW	MAPE = 10.2% MAPE* = 10.2% RMSE = 0.08MW
	1	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.082MW	MAPE = 10.7% MAPE* = 10.7% RMSE = 0.088MW	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.081MW	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.081MW
	2	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.08MW	MAPE = 10.5% MAPE* = 10.5% RMSE = 0.086MW	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.079MW	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.079MW
	7	MAPE = 9.8% MAPE* = 9.8% RMSE = 0.076MW	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.083MW	MAPE = 9.7% MAPE* = 9.7% RMSE = 0.076MW	MAPE = 9.7% MAPE* = 9.7% RMSE = 0.076MW

		Eynsham – D10 Forecast			
		Instant Correction Parameter			
		0	**7**	**21**	**28**
Volume Correction Parameter	**0**	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.691MW	MAPE = 10.6% MAPE* = 10.6% RMSE = 0.76MW	MAPE = 10.4% MAPE* = 10.4% RMSE = 0.699MW	MAPE = 10.3% MAPE* = 10.3% RMSE = 0.694MW
	1	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.321MW	MAPE = 7.6% MAPE* = 7.6% RMSE = 0.396MW	MAPE = 7.3% MAPE* = 7.3% RMSE = 0.329MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.323MW
	2	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.316MW	MAPE = 7.6% MAPE* = 7.6% RMSE = 0.393MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.323MW	MAPE = 7.1% MAPE* = 7.1% RMSE = 0.317MW
	7	MAPE = 7.3% MAPE* = 7.3% RMSE = 0.332MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.409MW	MAPE = 7.4% MAPE* = 7.4% RMSE = 0.34MW	MAPE = 7.3% MAPE* = 7.3% RMSE = 0.334MW

		Eynsham – D4 Forecast			
		Instant Correction Parameter			
		0	**7**	**21**	**28**
Volume Correction Parameter	**0**	MAPE = 9.9% MAPE* = 9.9% RMSE = 0.636MW	MAPE = 10.1% MAPE* = 10.1% RMSE = 0.721MW	MAPE = 9.9% MAPE* = 9.9% RMSE = 0.638MW	MAPE = 9.8% MAPE* = 9.8% RMSE = 0.634MW
	1	MAPE = 7.3% MAPE* = 7.3% RMSE = 0.333MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.422MW	MAPE = 7.3% MAPE* = 7.3% RMSE = 0.333MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.328MW
	2	MAPE = 7.3% MAPE* = 7.3% RMSE = 0.328MW	MAPE = 7.6% MAPE* = 7.6% RMSE = 0.422MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.329MW	MAPE = 7.2% MAPE* = 7.2% RMSE = 0.323MW
	7	MAPE = 7.5% MAPE* = 7.5% RMSE = 0.345MW	MAPE = 7.8% MAPE* = 7.8% RMSE = 0.437MW	MAPE = 7.4% MAPE* = 7.4% RMSE = 0.346MW	MAPE = 7.4% MAPE* = 7.4% RMSE = 0.341MW

Volume Correction Parameter		Kennington – D10 Forecast			
		Instant Correction Parameter			
		0	**7**	**21**	**28**
	0	MAPE = 18.8% MAPE* = 18.8% RMSE = 0.207MW	MAPE = 19% MAPE* = 19% RMSE = 0.217MW	MAPE = 18.9% MAPE* = 18.9% RMSE = 0.211MW	MAPE = 18.8% MAPE* = 18.8% RMSE = 0.21MW
	1	MAPE = 13.7% MAPE* = 13.7% RMSE = 0.119MW	MAPE = 13.9% MAPE* = 13.9% RMSE = 0.126MW	MAPE = 13.7% MAPE* = 13.7% RMSE = 0.121MW	MAPE = 13.6% MAPE* = 13.6% RMSE = 0.12MW
	2	MAPE = 12.6% MAPE* = 12.6% RMSE = 0.101MW	MAPE = 12.8% MAPE* = 12.8% RMSE = 0.108MW	MAPE = 12.6% MAPE* = 12.6% RMSE = 0.103MW	MAPE = 12.5% MAPE* = 12.5% RMSE = 0.102MW
	7	MAPE = 12.1% MAPE* = 12.1% RMSE = 0.094MW	MAPE = 12.4% MAPE* = 12.4% RMSE = 0.1MW	MAPE = 12.1% MAPE* = 12.1% RMSE = 0.096MW	MAPE = 12% MAPE* = 12% RMSE = 0.095MW

Volume Correction Parameter		Kennington – D4 Forecast			
		Instant Correction Parameter			
		0	**7**	**21**	**28**
	0	MAPE = 19% MAPE* = 19% RMSE = 0.217MW	MAPE = 18.9% MAPE* = 18.9% RMSE = 0.226MW	MAPE = 18.7% MAPE* = 18.7% RMSE = 0.217MW	MAPE = 18.7% MAPE* = 18.7% RMSE = 0.217MW
	1	MAPE = 14.3% MAPE* = 14.3% RMSE = 0.133MW	MAPE = 14.2% MAPE* = 14.2% RMSE = 0.139MW	MAPE = 14% MAPE* = 14% RMSE = 0.133MW	MAPE = 14.1% MAPE* = 14.1% RMSE = 0.132MW
	2	MAPE = 12.9% MAPE* = 12.9% RMSE = 0.107MW	MAPE = 12.8% MAPE* = 12.8% RMSE = 0.114MW	MAPE = 12.6% MAPE* = 12.6% RMSE = 0.108MW	MAPE = 12.6% MAPE* = 12.6% RMSE = 0.107MW
	7	MAPE = 12.1% MAPE* = 12.1% RMSE = 0.096MW	MAPE = 12% MAPE* = 12% RMSE = 0.102MW	MAPE = 11.8% MAPE* = 11.8% RMSE = 0.096MW	MAPE = 11.8% MAPE* = 11.8% RMSE = 0.095MW

		Milton – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 32.1% MAPE* = 32.1% RMSE = 32.311MW	MAPE = 32% MAPE* = 32% RMSE = 32.578MW	MAPE = 32% MAPE* = 32% RMSE = 32.497MW	MAPE = 32% MAPE* = 32% RMSE = 32.438MW
	1	MAPE = 22.6% MAPE* = 22.6% RMSE = 20.988MW	MAPE = 22.4% MAPE* = 22.4% RMSE = 21.116MW	MAPE = 22.4% MAPE* = 22.4% RMSE = 21.045MW	MAPE = 22.4% MAPE* = 22.4% RMSE = 21.008MW
	2	MAPE = 21.4% MAPE* = 21.4% RMSE = 19.674MW	MAPE = 21.3% MAPE* = 21.3% RMSE = 19.795MW	MAPE = 21.3% MAPE* = 21.3% RMSE = 19.731MW	MAPE = 21.3% MAPE* = 21.3% RMSE = 19.689MW
	7	MAPE = 22.1% MAPE* = 22.1% RMSE = 20.578MW	MAPE = 22% MAPE* = 22% RMSE = 20.706MW	MAPE = 21.9% MAPE* = 21.9% RMSE = 20.644MW	MAPE = 22% MAPE* = 22% RMSE = 20.599MW

		Milton – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 32.6% MAPE* = 32.6% RMSE = 33.513MW	MAPE = 32.5% MAPE* = 32.5% RMSE = 33.73MW	MAPE = 32.5% MAPE* = 32.5% RMSE = 33.611MW	MAPE = 32.5% MAPE* = 32.5% RMSE = 33.575MW
	1	MAPE = 23.2% MAPE* = 23.2% RMSE = 22.291MW	MAPE = 23.2% MAPE* = 23.2% RMSE = 22.41MW	MAPE = 23.1% MAPE* = 23.1% RMSE = 22.323MW	MAPE = 23.1% MAPE* = 23.1% RMSE = 22.298MW
	2	MAPE = 22.1% MAPE* = 22.1% RMSE = 20.953MW	MAPE = 22.1% MAPE* = 22.1% RMSE = 21.069MW	MAPE = 22% MAPE* = 22% RMSE = 20.981MW	MAPE = 22% MAPE* = 22% RMSE = 20.953MW
	7	MAPE = 22.5% MAPE* = 22.5% RMSE = 21.353MW	MAPE = 22.4% MAPE* = 22.4% RMSE = 21.467MW	MAPE = 22.4% MAPE* = 22.4% RMSE = 21.38MW	MAPE = 22.4% MAPE* = 22.4% RMSE = 21.352MW

Volume Correction Parameter		Oxford Primary – D10 Forecast Instant Correction Parameter: 0	7	21	28
	0	MAPE = 15.3% MAPE* = 15.3% RMSE = 3.324MW	MAPE = 15.4% MAPE* = 15.4% RMSE = 3.319MW	MAPE = 15.4% MAPE* = 15.4% RMSE = 3.268MW	MAPE = 15.4% MAPE* = 15.4% RMSE = 3.262MW
	1	MAPE = 6.6% MAPE* = 6.6% RMSE = 0.749MW	MAPE = 6.6% MAPE* = 6.6% RMSE = 0.755MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.718MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.713MW
	2	MAPE = 6.5% MAPE* = 6.5% RMSE = 0.735MW	MAPE = 6.6% MAPE* = 6.6% RMSE = 0.742MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.705MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.7MW
	7	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.726MW	MAPE = 6.5% MAPE* = 6.5% RMSE = 0.733MW	MAPE = 6.3% MAPE* = 6.3% RMSE = 0.697MW	MAPE = 6.3% MAPE* = 6.3% RMSE = 0.693MW

Volume Correction Parameter		Oxford Primary – D4 Forecast Instant Correction Parameter: 0	7	21	28
	0	MAPE = 16.5% MAPE* = 16.5% RMSE = 3.819MW	MAPE = 16.5% MAPE* = 16.5% RMSE = 3.803MW	MAPE = 16.5% MAPE* = 16.5% RMSE = 3.754MW	MAPE = 16.5% MAPE* = 16.5% RMSE = 3.753MW
	1	MAPE = 6.7% MAPE* = 6.7% RMSE = 0.789MW	MAPE = 6.6% MAPE* = 6.6% RMSE = 0.774MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.738MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.737MW
	2	MAPE = 6.7% MAPE* = 6.7% RMSE = 0.779MW	MAPE = 6.6% MAPE* = 6.6% RMSE = 0.763MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.727MW	MAPE = 6.4% MAPE* = 6.4% RMSE = 0.727MW
	7	MAPE = 6.6% MAPE* = 6.6% RMSE = 0.764MW	MAPE = 6.5% MAPE* = 6.5% RMSE = 0.748MW	MAPE = 6.3% MAPE* = 6.3% RMSE = 0.713MW	MAPE = 6.3% MAPE* = 6.3% RMSE = 0.712MW

		Rose Hill – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 19.6% MAPE* = 19.6% RMSE = 2.857MW	MAPE = 19.7% MAPE* = 19.7% RMSE = 2.968MW	MAPE = 19.6% MAPE* = 19.6% RMSE = 2.869MW	MAPE = 19.6% MAPE* = 19.6% RMSE = 2.856MW
	1	MAPE = 8.1% MAPE* = 8.1% RMSE = 0.639MW	MAPE = 8.6% MAPE* = 8.6% RMSE = 0.708MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 0.638MW	MAPE = 8% MAPE* = 8% RMSE = 0.63MW
	2	MAPE = 8.1% MAPE* = 8.1% RMSE = 0.629MW	MAPE = 8.5% MAPE* = 8.5% RMSE = 0.698MW	MAPE = 8% MAPE* = 8% RMSE = 0.628MW	MAPE = 8% MAPE* = 8% RMSE = 0.62MW
	7	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.588MW	MAPE = 8.2% MAPE* = 8.2% RMSE = 0.657MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.587MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.579MW

		Rose Hill – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 20.6% MAPE* = 20.6% RMSE = 3.179MW	MAPE = 20.7% MAPE* = 20.7% RMSE = 3.256MW	MAPE = 20.6% MAPE* = 20.6% RMSE = 3.161MW	MAPE = 20.6% MAPE* = 20.6% RMSE = 3.147MW
	1	MAPE = 8.3% MAPE* = 8.3% RMSE = 0.686MW	MAPE = 8.6% MAPE* = 8.6% RMSE = 0.736MW	MAPE = 8.2% MAPE* = 8.2% RMSE = 0.67MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 0.661MW
	2	MAPE = 8.2% MAPE* = 8.2% RMSE = 0.671MW	MAPE = 8.5% MAPE* = 8.5% RMSE = 0.721MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 0.655MW	MAPE = 8% MAPE* = 8% RMSE = 0.645MW
	7	MAPE = 7.9% MAPE* = 7.9% RMSE = 0.627MW	MAPE = 8.2% MAPE* = 8.2% RMSE = 0.677MW	MAPE = 7.8% MAPE* = 7.8% RMSE = 0.611MW	MAPE = 7.7% MAPE* = 7.7% RMSE = 0.601MW

		University Parks – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 5.8% MAPE* = 5.8% RMSE = 0.92MW	MAPE = 5.8% MAPE* = 5.8% RMSE = 0.905MW	MAPE = 5.6% MAPE* = 5.6% RMSE = 0.838MW	MAPE = 5.6% MAPE* = 5.6% RMSE = 0.834MW
	1	MAPE = 5.4% MAPE* = 5.4% RMSE = 0.675MW	MAPE = 5.2% MAPE* = 5.2% RMSE = 0.676MW	MAPE = 5.1% MAPE* = 5.1% RMSE = 0.611MW	MAPE = 5.1% MAPE* = 5.1% RMSE = 0.607MW
	2	MAPE = 5.3% MAPE* = 5.3% RMSE = 0.651MW	MAPE = 5.1% MAPE* = 5.1% RMSE = 0.652MW	MAPE = 5% MAPE* = 5% RMSE = 0.587MW	MAPE = 5% MAPE* = 5% RMSE = 0.583MW
	7	MAPE = 5% MAPE* = 5% RMSE = 0.577MW	MAPE = 4.9% MAPE* = 4.9% RMSE = 0.58MW	MAPE = 4.7% MAPE* = 4.7% RMSE = 0.515MW	MAPE = 4.7% MAPE* = 4.7% RMSE = 0.511MW

		University Parks – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 6.1% MAPE* = 6.1% RMSE = 1.059MW	MAPE = 6.1% MAPE* = 6.1% RMSE = 1.043MW	MAPE = 6% MAPE* = 6% RMSE = 0.994MW	MAPE = 6% MAPE* = 6% RMSE = 0.992MW
	1	MAPE = 5.4% MAPE* = 5.4% RMSE = 0.707MW	MAPE = 5.2% MAPE* = 5.2% RMSE = 0.705MW	MAPE = 5.1% MAPE* = 5.1% RMSE = 0.657MW	MAPE = 5.1% MAPE* = 5.1% RMSE = 0.655MW
	2	MAPE = 5.3% MAPE* = 5.3% RMSE = 0.684MW	MAPE = 5.2% MAPE* = 5.2% RMSE = 0.682MW	MAPE = 5% MAPE* = 5% RMSE = 0.633MW	MAPE = 5% MAPE* = 5% RMSE = 0.631MW
	7	MAPE = 5.1% MAPE* = 5.1% RMSE = 0.617MW	MAPE = 4.9% MAPE* = 4.9% RMSE = 0.615MW	MAPE = 4.8% MAPE* = 4.8% RMSE = 0.567MW	MAPE = 4.8% MAPE* = 4.8% RMSE = 0.565MW

		Yarnton Primary – D10 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 8.7% MAPE* = 8.7% RMSE = 1.335MW	MAPE = 8.6% MAPE* = 8.6% RMSE = 1.365MW	MAPE = 8.5% MAPE* = 8.5% RMSE = 1.306MW	MAPE = 8.5% MAPE* = 8.5% RMSE = 1.304MW
	1	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.331MW	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.363MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.305MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.305MW
	2	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.309MW	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.338MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.28MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.279MW
	7	MAPE = 8.2% MAPE* = 8.2% RMSE = 1.265MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.29MW	MAPE = 7.9% MAPE* = 7.9% RMSE = 1.233MW	MAPE = 7.9% MAPE* = 7.9% RMSE = 1.231MW

		Yarnton Primary – D4 Forecast			
		Instant Correction Parameter			
		0	7	21	28
Volume Correction Parameter	0	MAPE = 9% MAPE* = 9% RMSE = 1.397MW	MAPE = 9.1% MAPE* = 9.1% RMSE = 1.759MW	MAPE = 8.9% MAPE* = 8.9% RMSE = 1.686MW	MAPE = 8.9% MAPE* = 8.9% RMSE = 1.63MW
	1	MAPE = 8.4% MAPE* = 8.4% RMSE = 1.347MW	MAPE = 8.4% MAPE* = 8.4% RMSE = 1.701MW	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.644MW	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.586MW
	2	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.31MW	MAPE = 8.3% MAPE* = 8.3% RMSE = 1.682MW	MAPE = 8.2% MAPE* = 8.2% RMSE = 1.604MW	MAPE = 8.2% MAPE* = 8.2% RMSE = 1.546MW
	7	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.268MW	MAPE = 8.1% MAPE* = 8.1% RMSE = 1.656MW	MAPE = 8% MAPE* = 8% RMSE = 1.561MW	MAPE = 8% MAPE* = 8% RMSE = 1.502MW

13.2 Appendix 2 – Study on the importance of locational weather sources for Load Forecasting -Focus on weather forecast quality in Oxfordshire

13.2.1 Context of the study

Using accurate weather data is essential to drive accurate load forecast and determine the flexibility requirements. During the project a specific phenomenon was observed at a specific Wind Farm generator. While the generator is not part of the Transition project, it was included as a representative of Wind power modelling across the network. However, it appeared that the Mosmix deterministic forecasts provided at the windfarm regularly fell outside of the probabilistic envelope of ICON forecasts for the same generation asset.

A study has therefore been carried out to understand why the forecast were so different. The study has been limited to the geographical scope of the Transition project which is limited to 3 weather stations: Little Rissington, Brize Norton and Lyneham. It investigates the quality of Mosmix' and ICON's forecasts specifically for the wind variable, and concludes on the solar irradiance variable as well.

The full study has been made available to the Transition Project

13.2.2 Key takeaways

1. The distance between windfarms and weather stations do not explain the differences between the weather models:

   There is only ten kms between the exact location of the windfarm and its closest weather station, Brize Norton, where Mosmix forecasts are calculated

   There is only two kms between Brize Norton weather station and its closest ICON grid point.

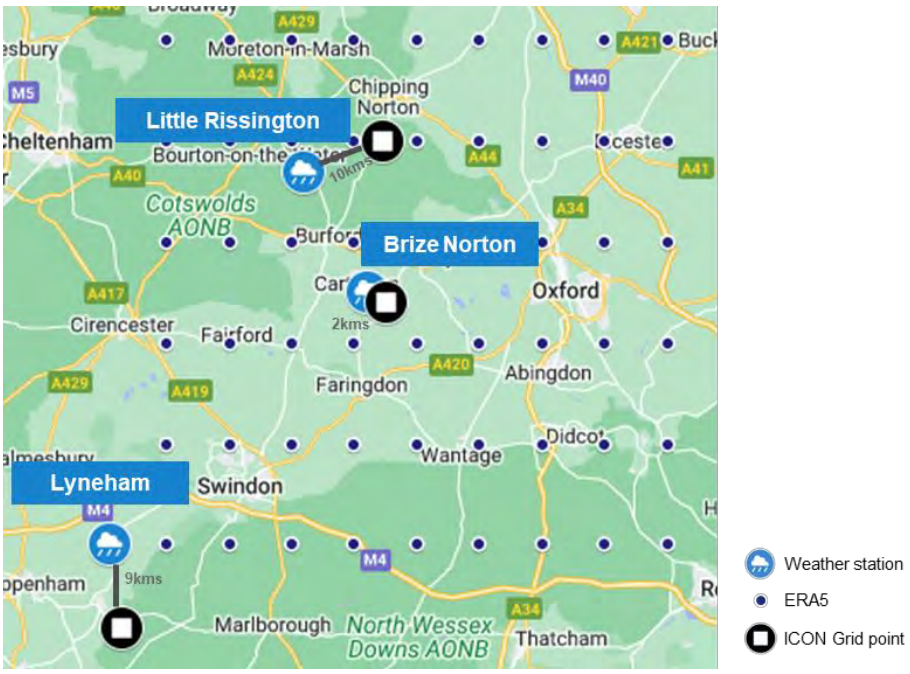


Figure 54 - Geographical scope of the weather study

Conclusion: The difference observed between the two model forecasts cannot be inputted to the distance between their prediction location

2. ICON is the better model for wind speed prediction

    Taking the mean of ICON forecasts is a robust method. Compared to Mosmix, error metrics are much lower and within a close interval

    The mean of ICON presents a slight overestimation bias, which can be recalibrated.

    Its performance did not change during Summer 2021, when winds were under normal levels

	Annual mean Error (m/s)		Annual WAPE (%)	
Weather Station	**Icon (mean)**	**Mosmix**	**Icon (mean)**	**Mosmix**
Little Rissington	**- 0.53**	- 1.25	**13%**	33%
Brize Norton	**- 0.39**	0.56	**10%**	16%
Lyneham	- 0.58	**- 0.23**	14%	**8%**

Figure 55 - Annual Errors of ICON and Mosmix for Wind speed

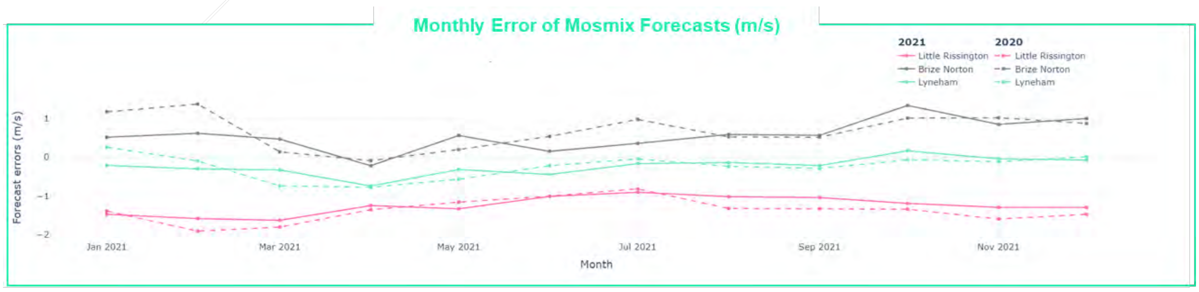
Conclusion: Icon is more consistent and resilient than Mosmix

3. Mosmix quality variability is an inherent weakness

    The quality of Mosmix forecasts is variable: both the direction of errors and their magnitude significantly differ depending on the location.

    Mosmix errors can reach 40% of actual wind speed, an accuracy red flag.

    Same levels of performance were found for 2020



$$Error = ERA5\ realisation - Model\ Forecast$$

Figure 56 - Monthly Errors (m/s) for Mosmix

Conclusion: Mosmix local statistical training is not relevant for all stations. The potential overfitting of Mosmix is a black box.

4. Mosmix' poorer prediction performance is specific to wind speed

    Its performance is significantly less volatile for solar irradiance than for wind speed.

    Error metrics are both smaller and within a close interval from one location to the other

    Thus, it is still a viable model for the global irradiance

	Solar Irradiance		Wind Speed (reminder)
Weather Station	**Error (kW/sqm)**	**WAPE (%)**	**WAPE (%)**
Little Rissington	- 0.0005	5%	33%
Brize Norton	- 0.0022	7%	16%
Lyneham	- 0.0013	7%	8%

Conclusion: While solar irradiance is forecasted thanks to sound and widely tested theoretical formulas, wind speed is intrinsically more uncertain

13.2.3 Conclusion and next steps

The study concludes that on this scope, Mosmix wind forecasts are inaccurate and not reliable. The poor quality of forecast is not seen on global irradiance which is reassuring. Considering no windfarm currently participates in the Transition project, the impact on the quality of the demand and generation forecasts is limited. Mosmix can still be used to determine flexibility requirements on the network. On the short-term, ICON seems to be overall more robust and consistent

For the future development of short-term forecasting capabilities, it is recommended to undertake a wider study across the license areas. The objective would be to determine the appropriate weather source SSEN should use, independently from the calibration of the models.

13.3 Appendix 3 – Forecast Accuracy Primary substation charts over June 2022

13.3.1 Arncott

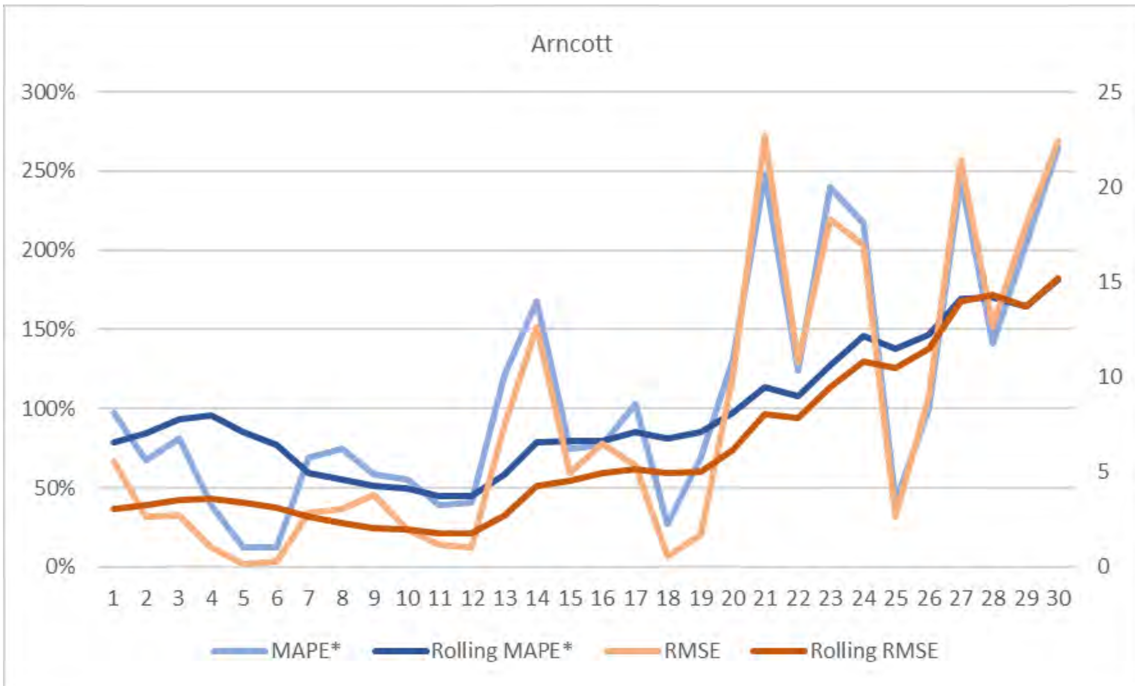


Figure 57 – Original forecast accuracy indicators for group demand in June 2022 – Arncott

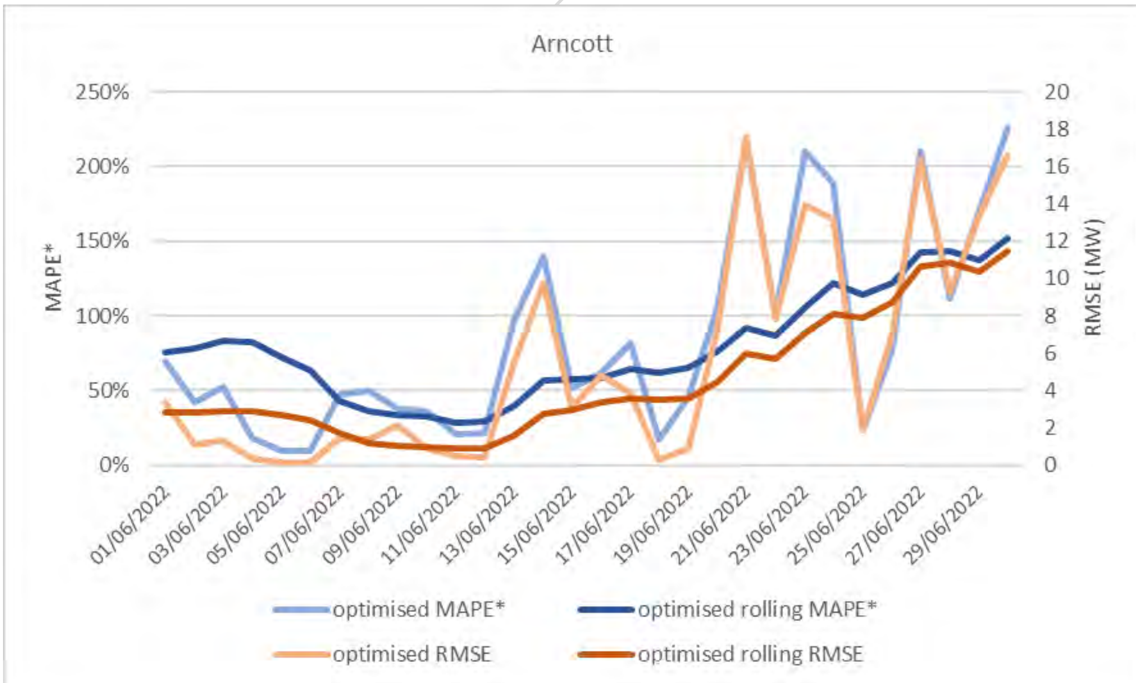


Figure 58 – Optimised forecast accuracy indicators for group demand in June 2022 – Arncott

13.3.2 Berinsfield

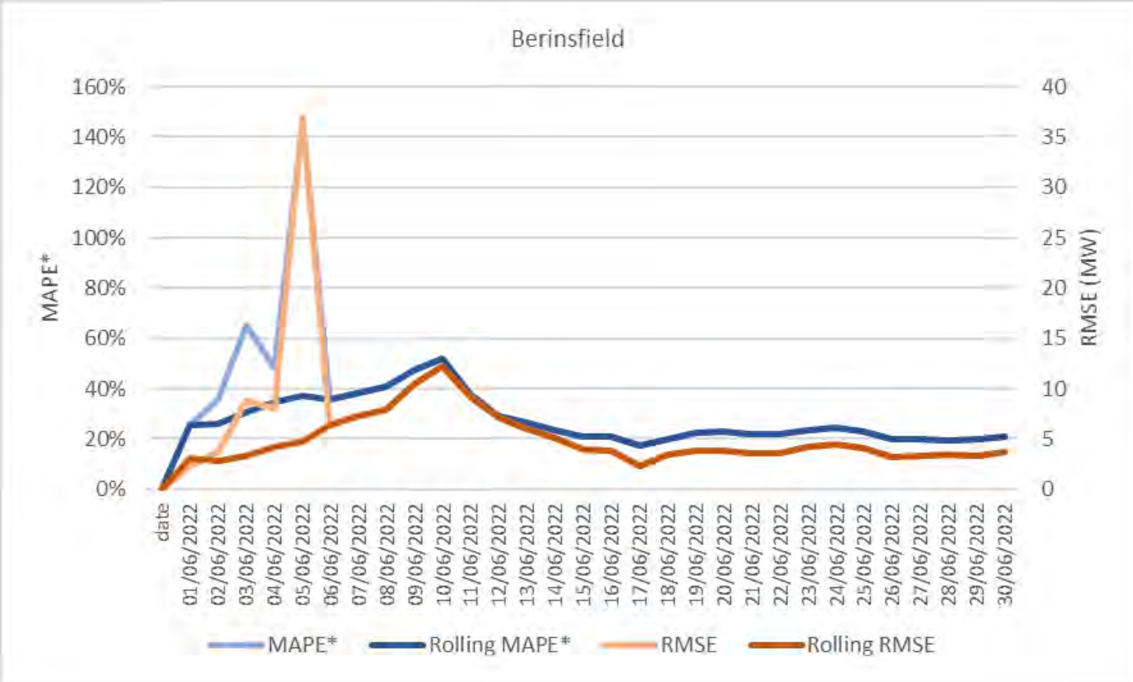


Figure 59 – Original forecast accuracy indicators for group demand in June 2022 – Berinsfield

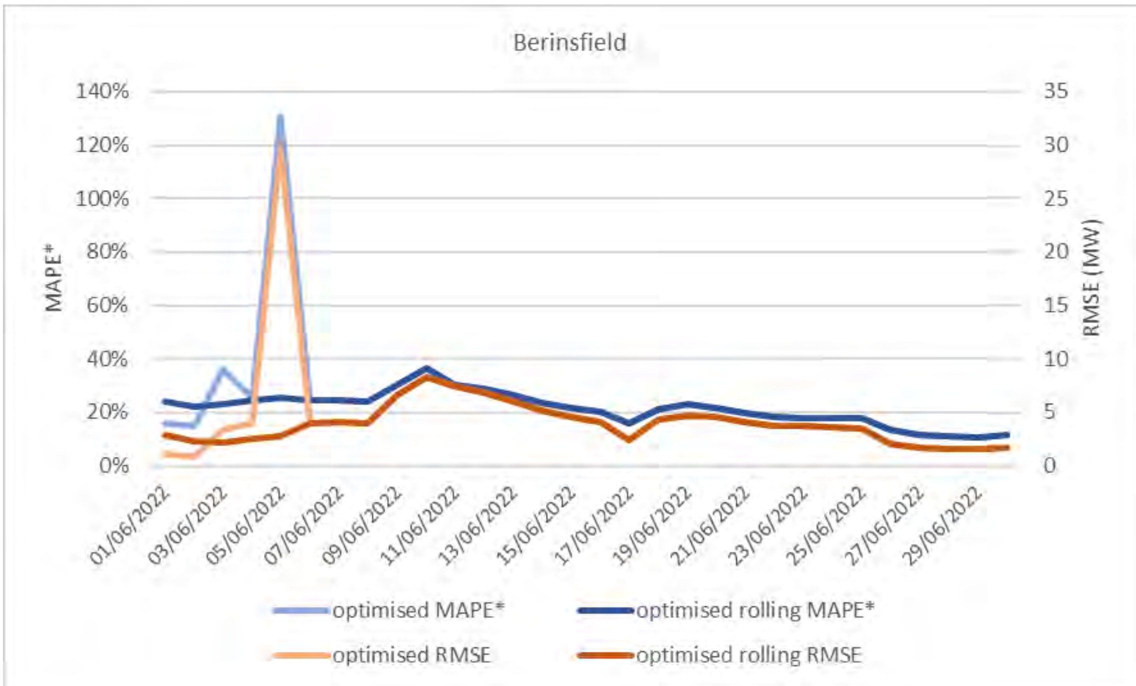


Figure 60 – Optimised forecast accuracy indicators for group demand in June 2022 – Berinsfield

13.3.3 Bicester

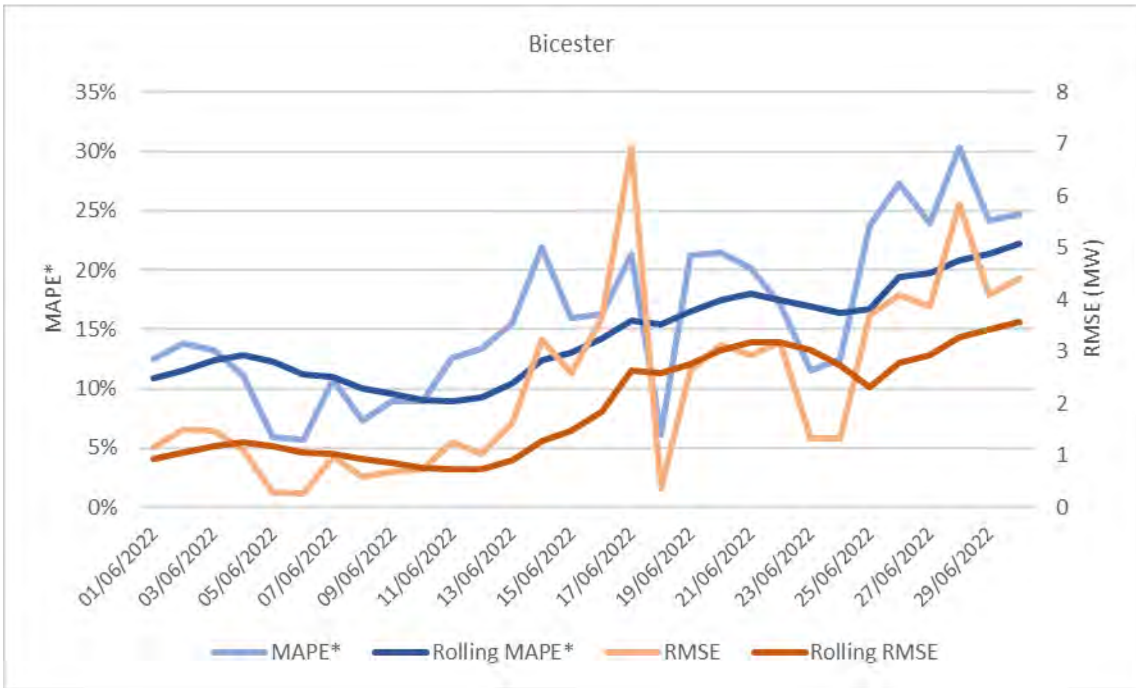


Figure 61 – Original forecast accuracy indicators for group demand in June 2022 – Bicester

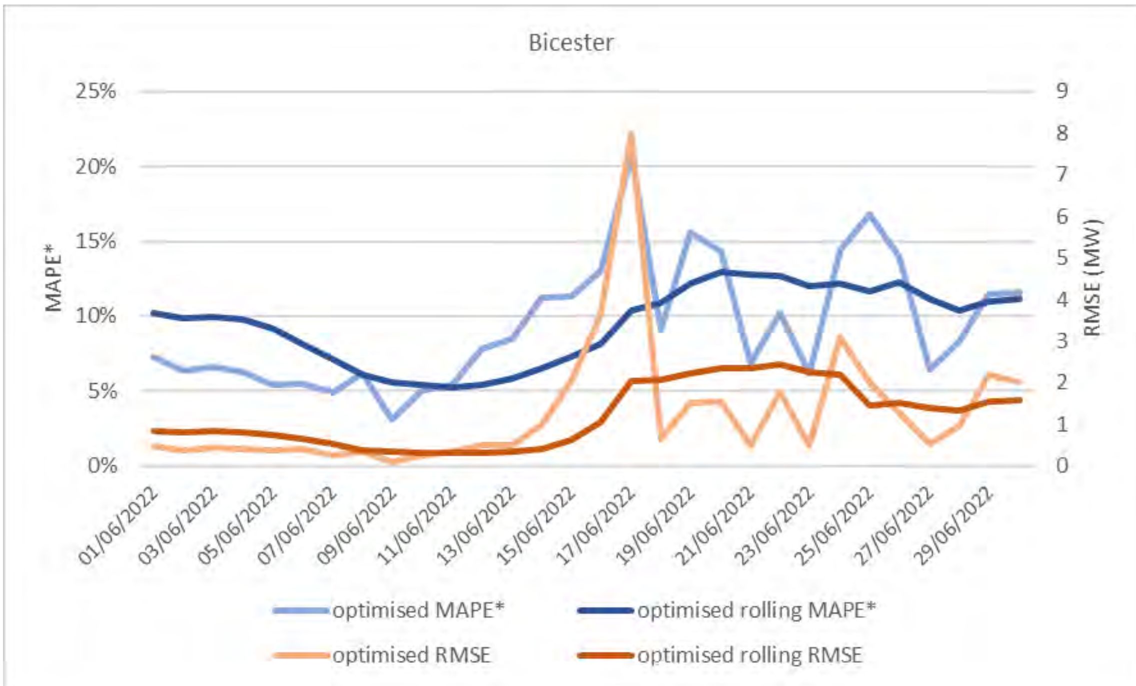


Figure 62 – Optimised forecast accuracy indicators for group demand in June 2022 – Bicester

13.3.4 Bicester North Primary

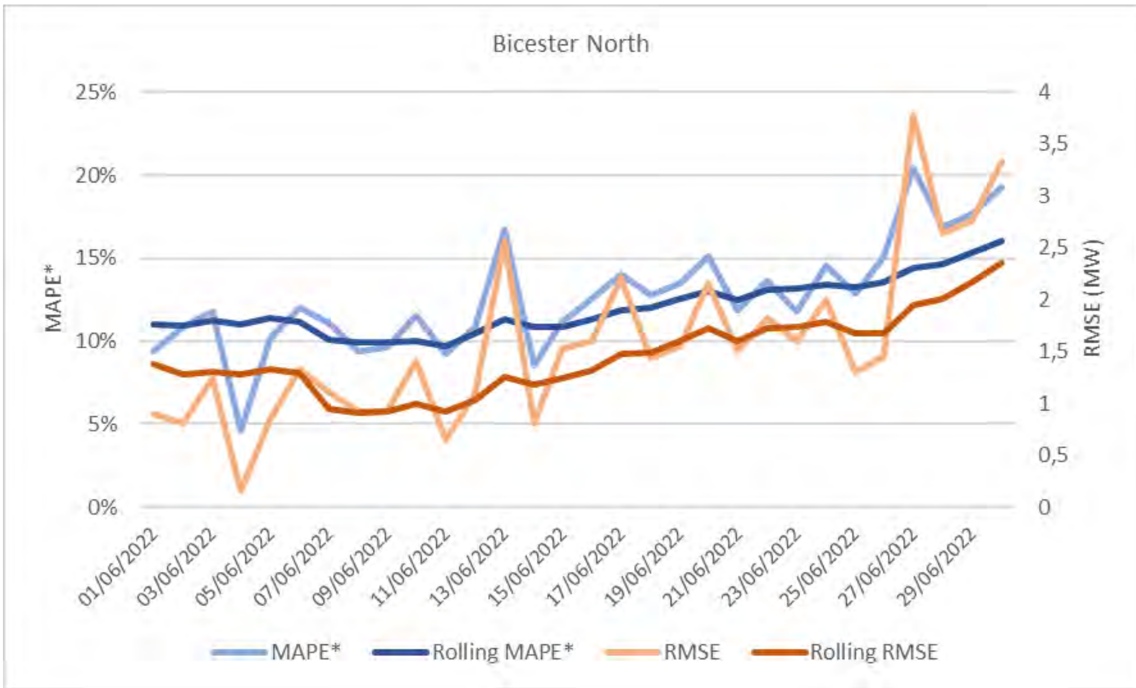


Figure 63 -Original forecast accuracy indicators for group demand in June 2022 – Bicester North Primary

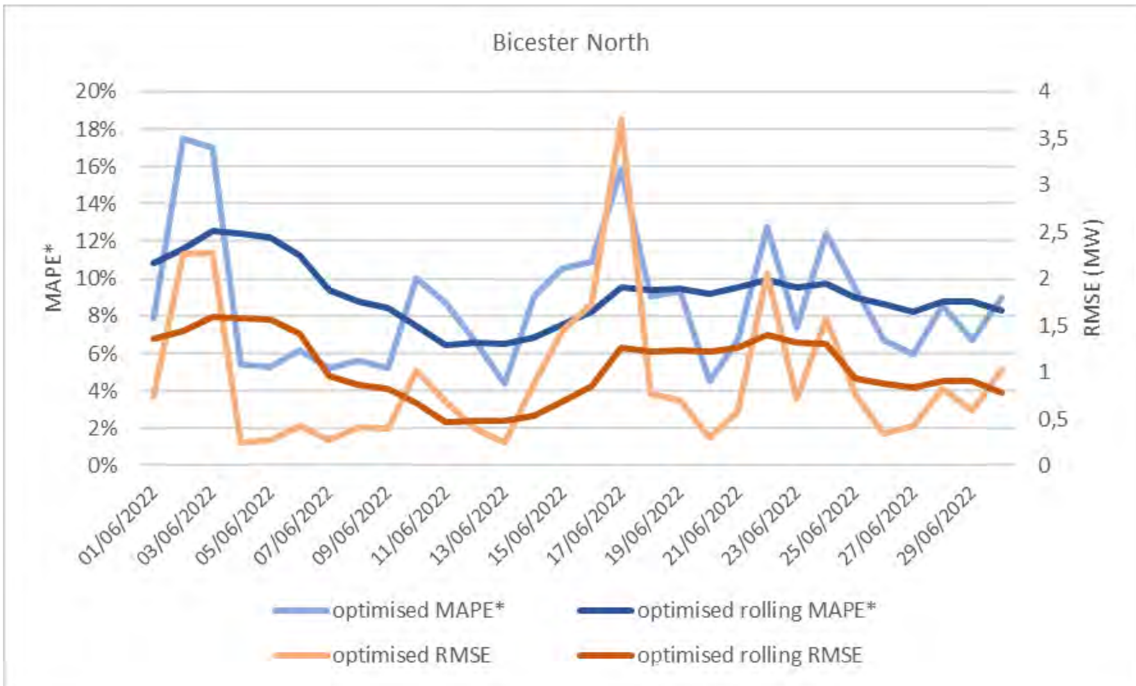


Figure 64 – Optimised forecast accuracy indicators for group demand in June 2022 – Bicester North Primary

13.3.5 Eynsham

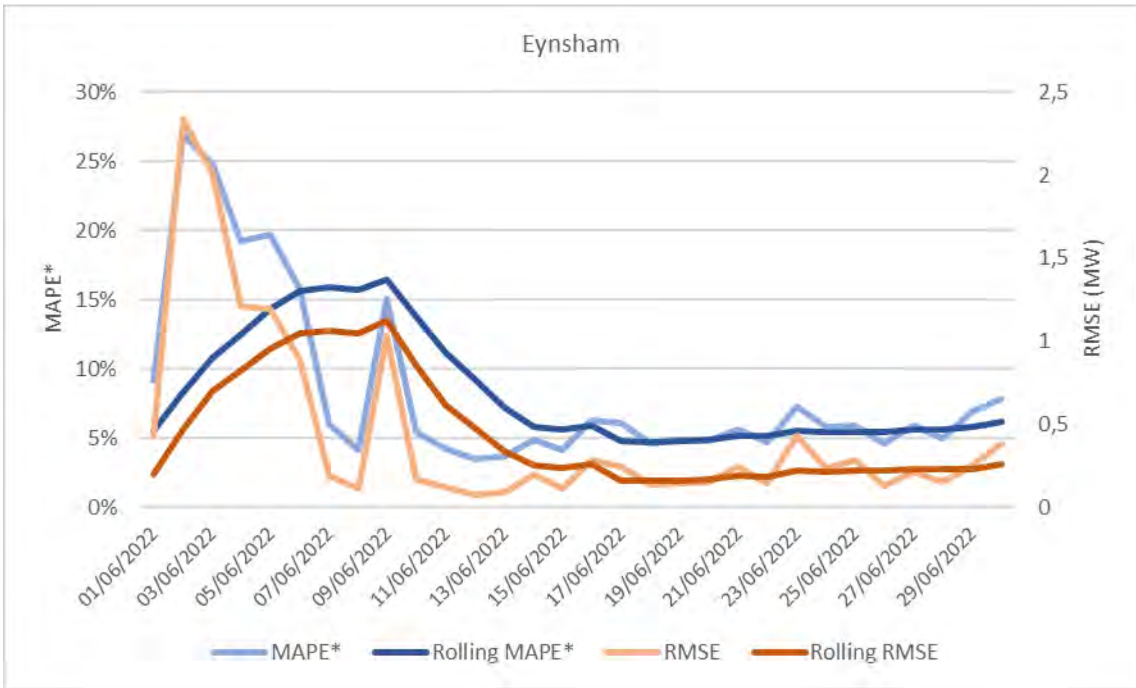


Figure 65 – Original forecast accuracy indicators for group demand in June 2022 – Eynsham

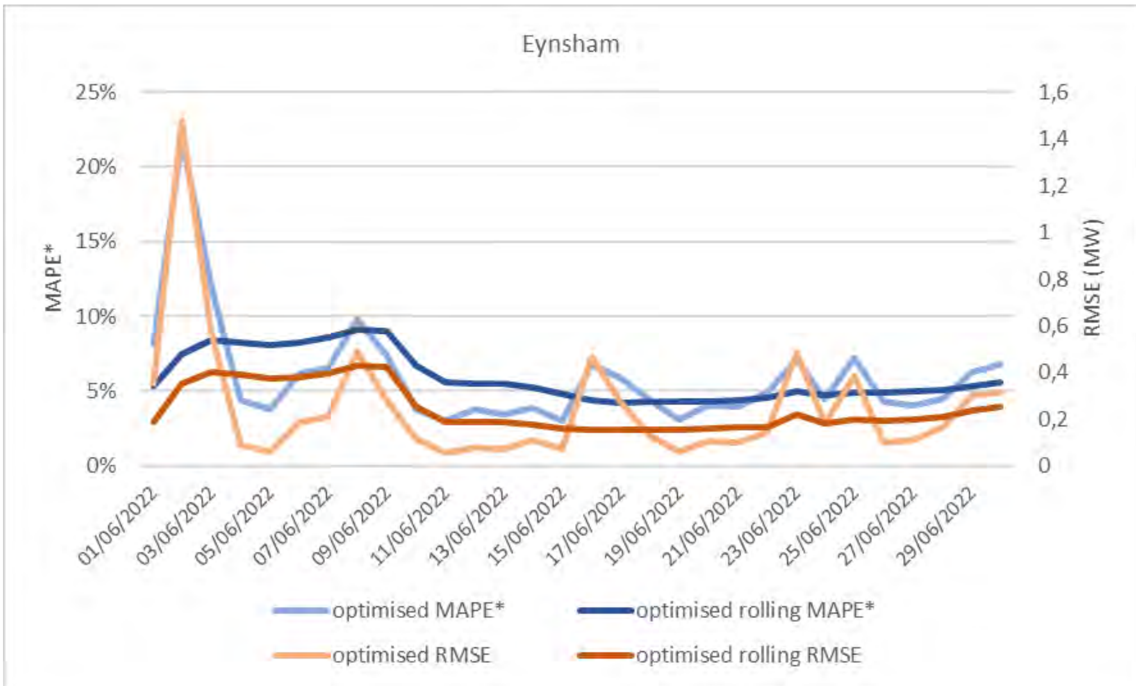


Figure 66 – Optimised forecast accuracy indicators for group demand in June 2022 – Eynsham

13.3.6 Kennington

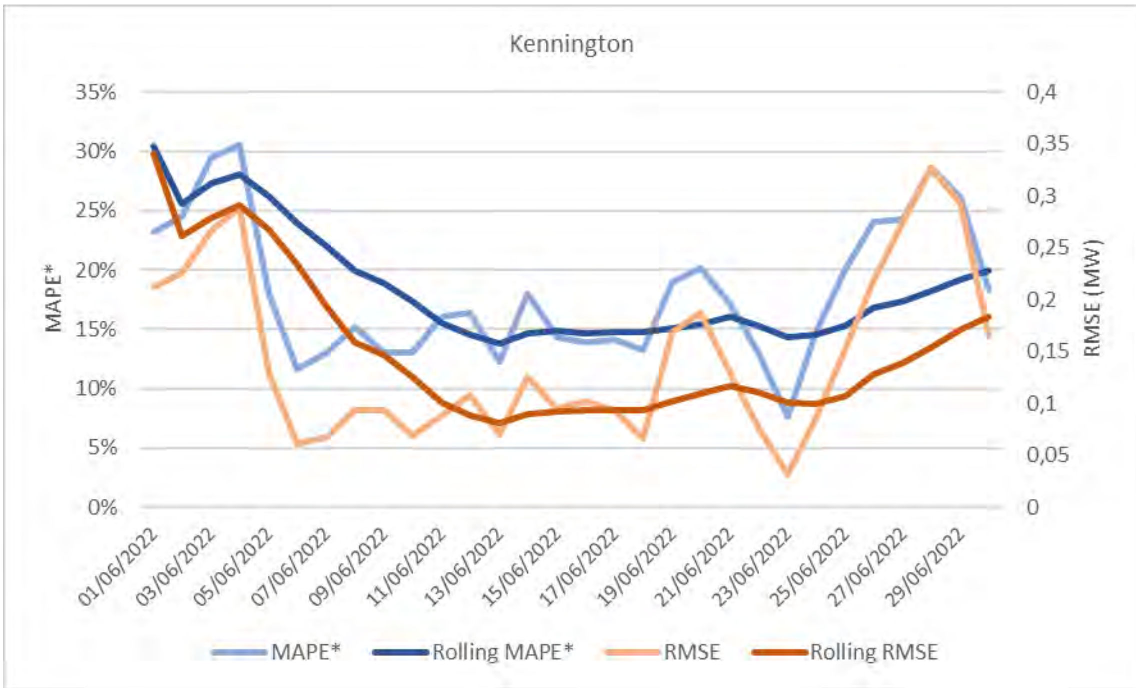


Figure 67 – Original forecast accuracy indicators for group demand in June 2022 – Kennington

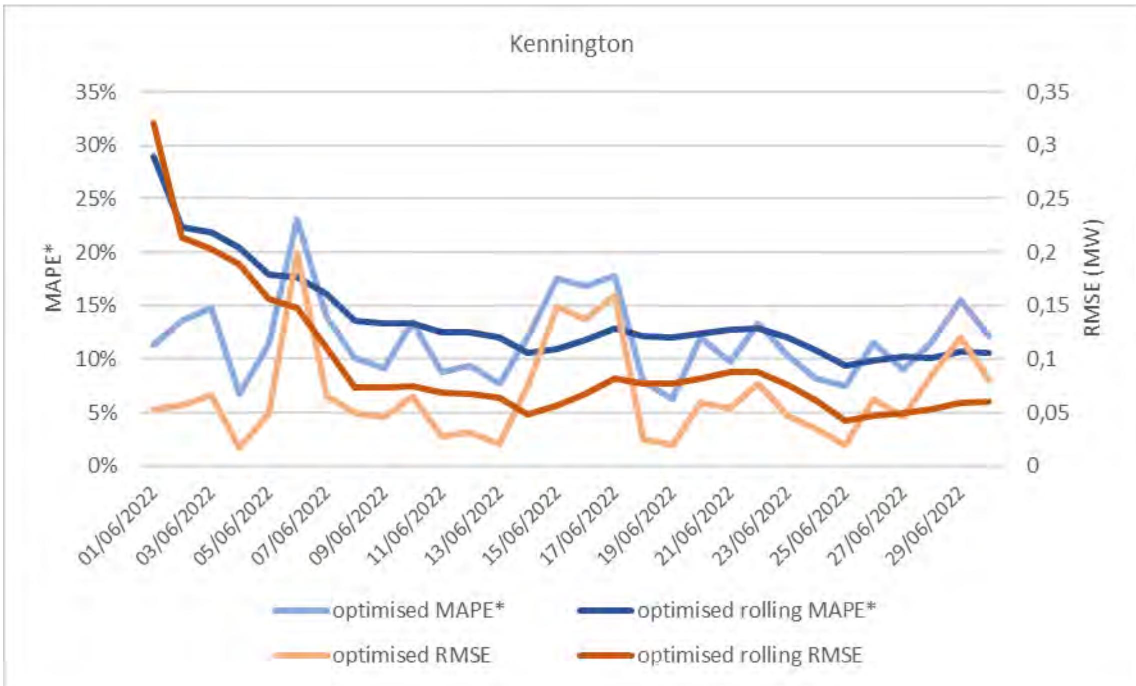


Figure 68 – Optimised forecast accuracy indicators for group demand in June 2022 – Kennington

13.3.7 Milton

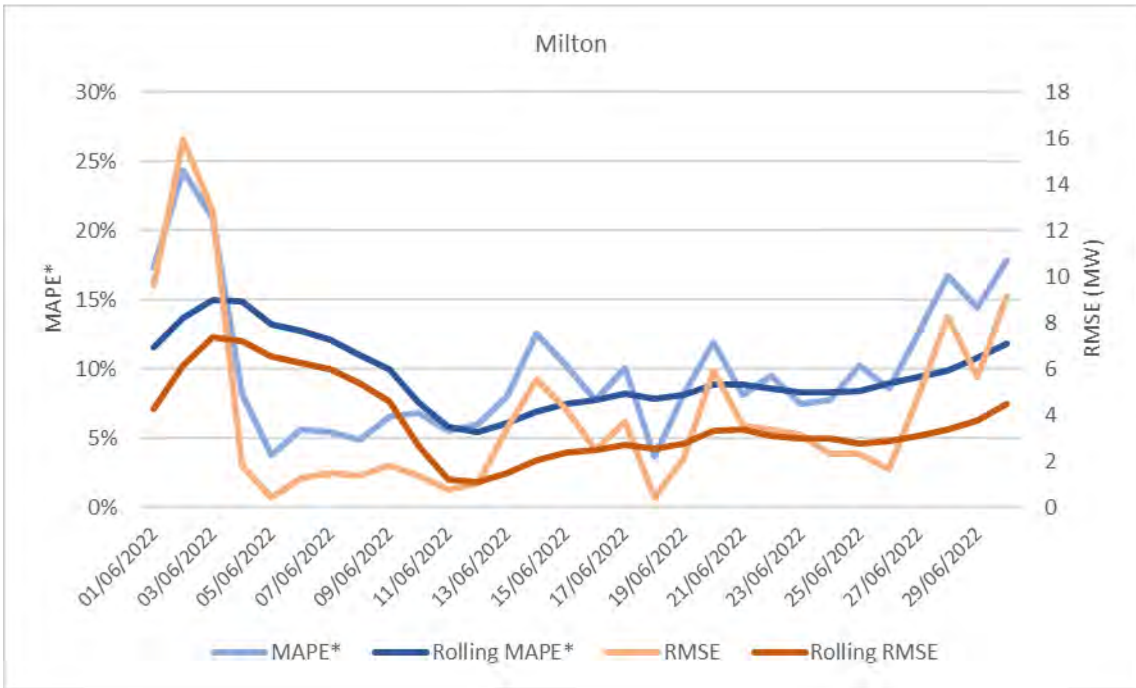


Figure 69 – Original forecast accuracy indicators for group demand in June 2022 – Milton

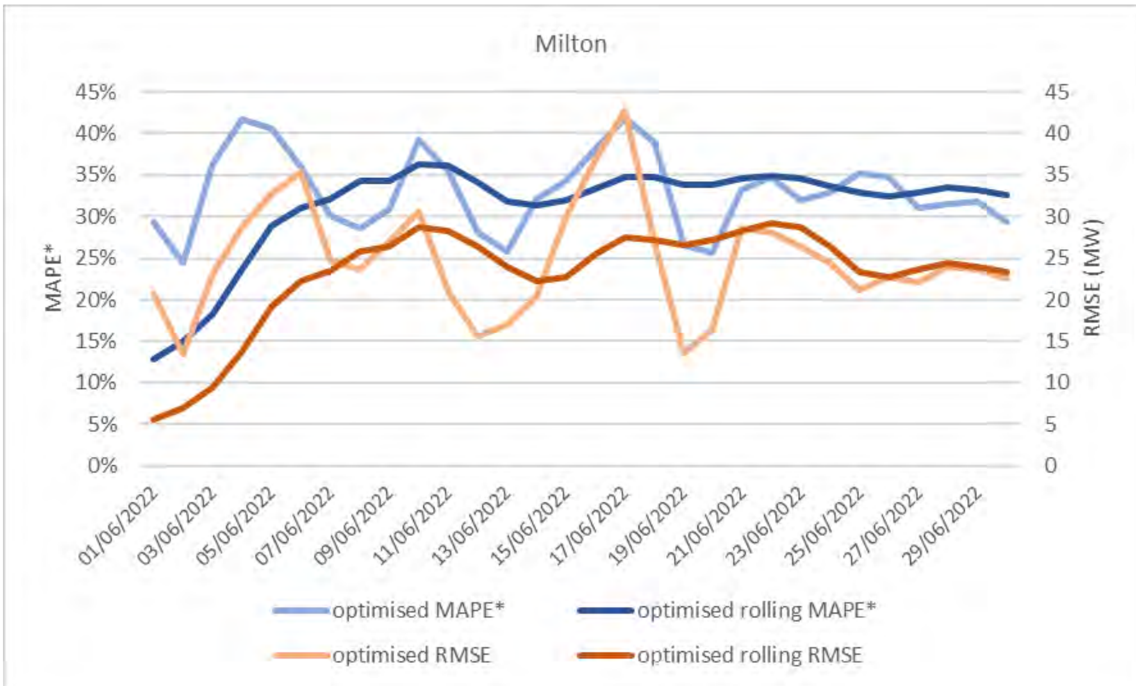


Figure 70 – Optimised forecast accuracy indicators for group demand in June 2022 – Milton

13.3.8 Rose Hill

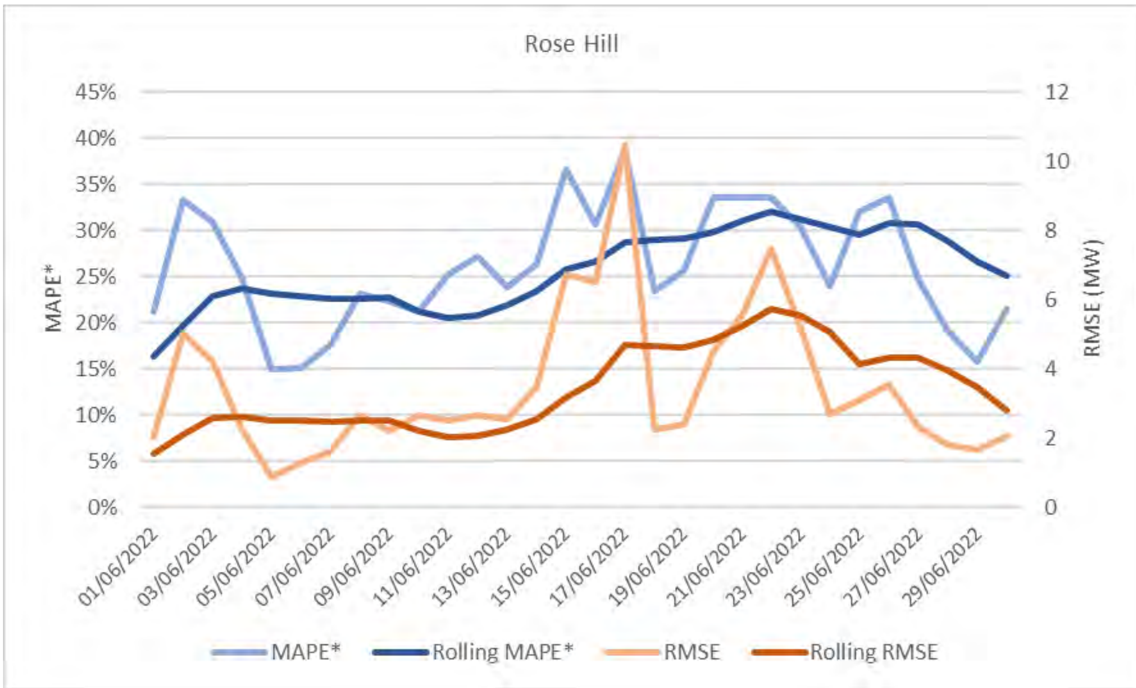


Figure 71 – Original forecast accuracy indicators for group demand in June 2022 – Rose Hill

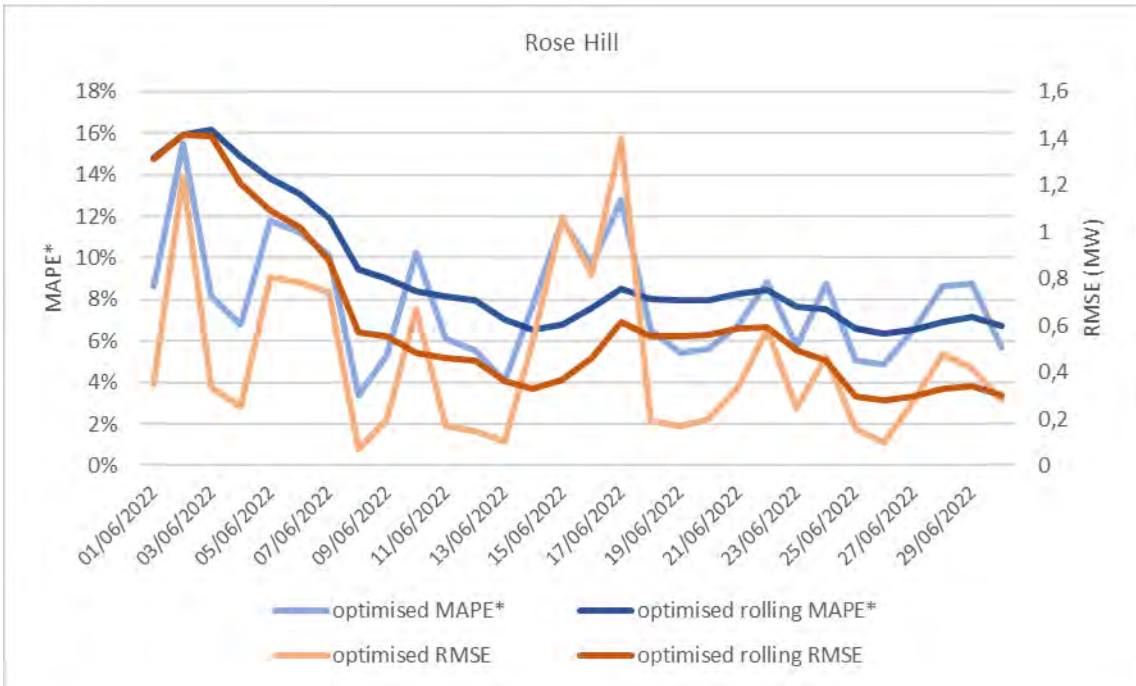


Figure 72 – Optimised forecast accuracy indicators for group demand in June 2022 – Rose Hill

13.3.9 University Parks

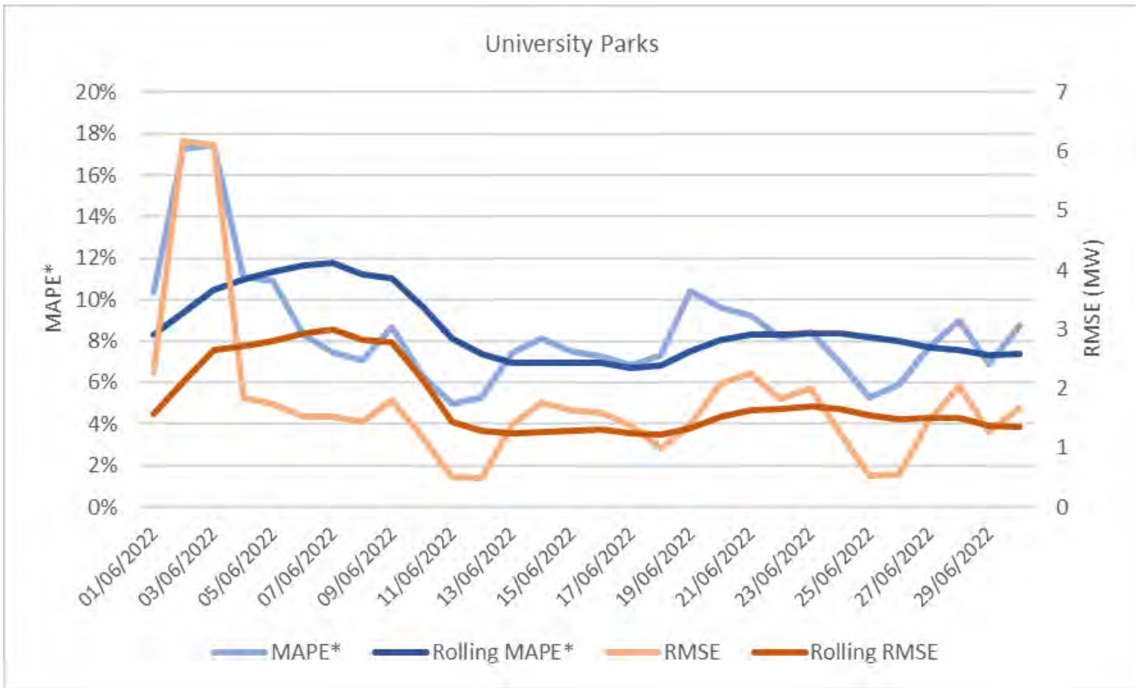


Figure 73 – Original forecast accuracy indicators for group demand in June 2022 – University Parks

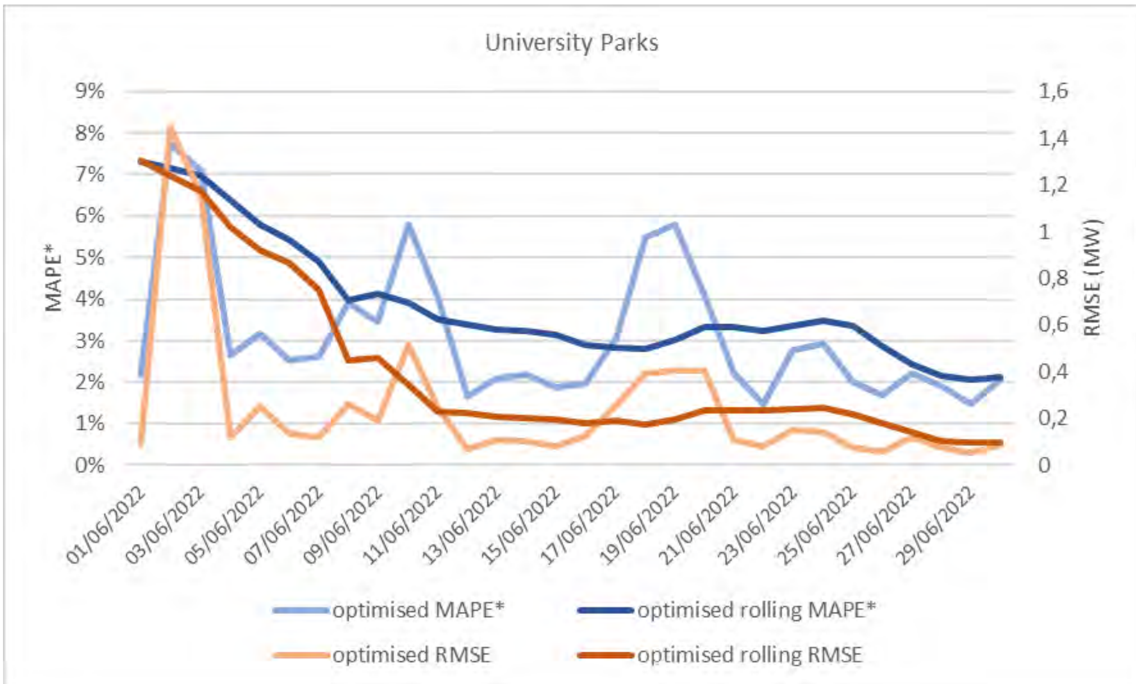


Figure 74 – Optimised forecast accuracy indicators for group demand in June 2022 – University Parks

13.3.10 Yarnton Primary

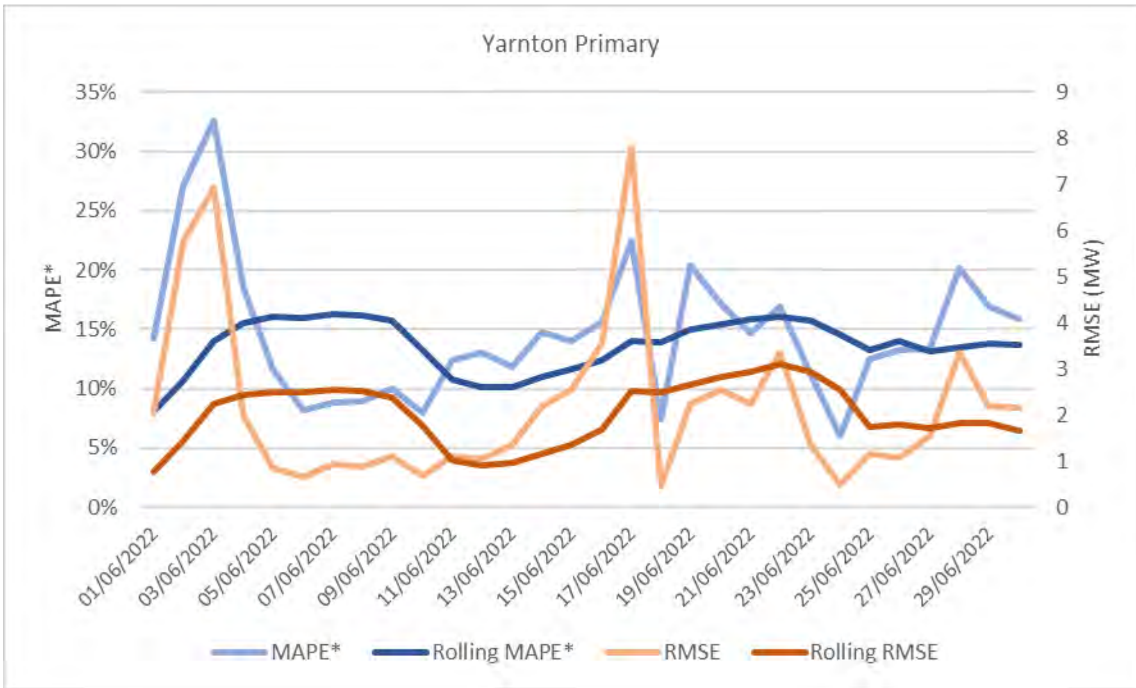


Figure 75 – Original forecast accuracy indicators for group demand in June 2022 – University Parks

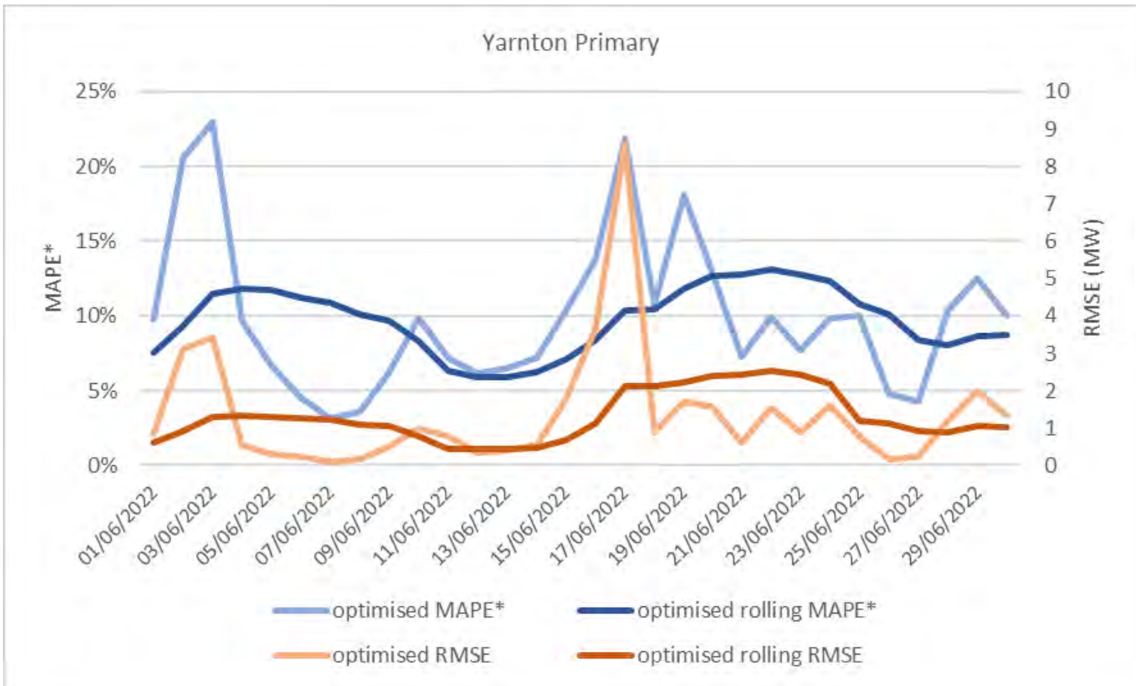


Figure 76 – Optimised forecast accuracy indicators for group demand in June 2022 – Yarnton Primary

13.4 Appendix 4 – Forecast Horizon Accuracy study

13.4.1 Original Demand Forecast - Results at primary substation

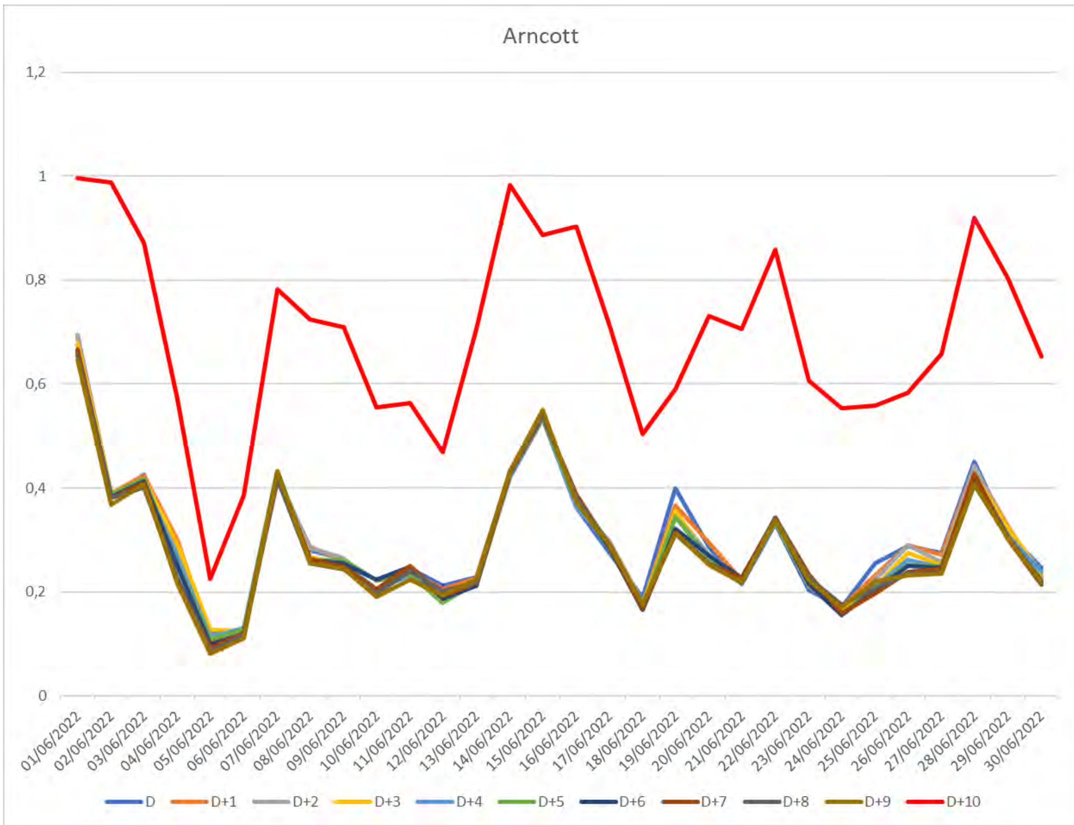


Figure 77 - Daily average relative error for all original forecasts in June 22 – Arncott

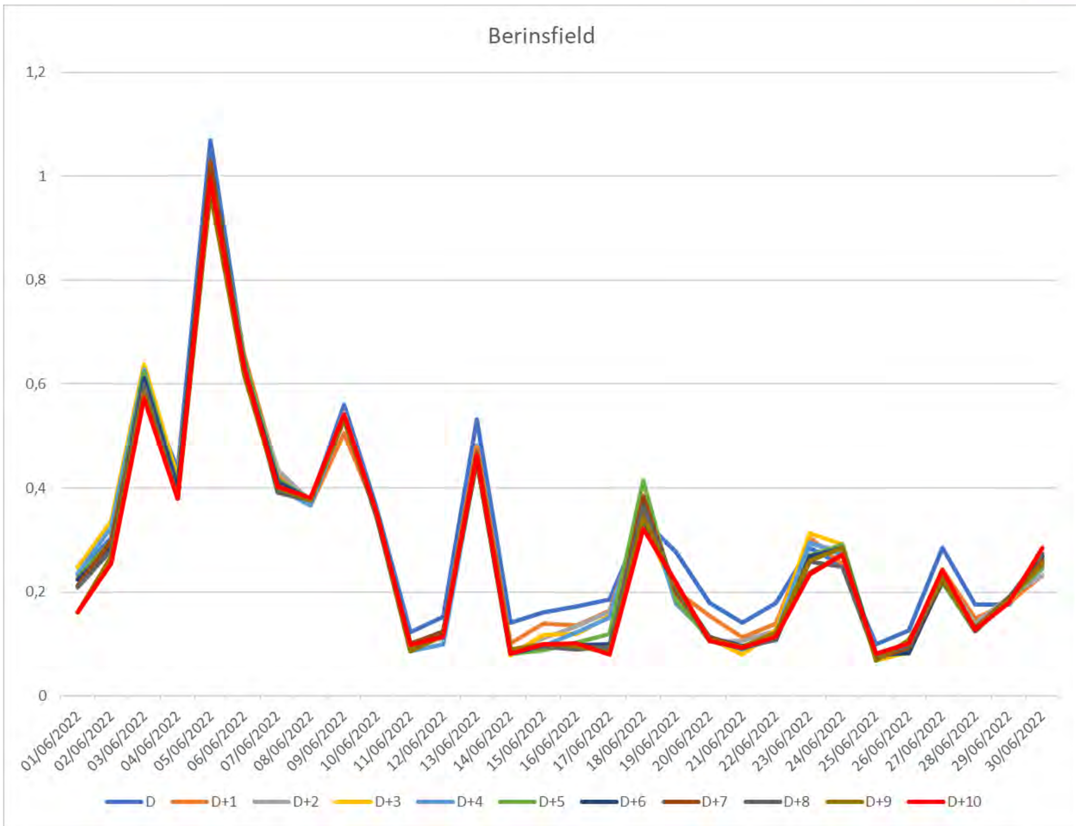


Figure 78 - Daily average relative error for all original forecasts in June 22 – Berinsfield

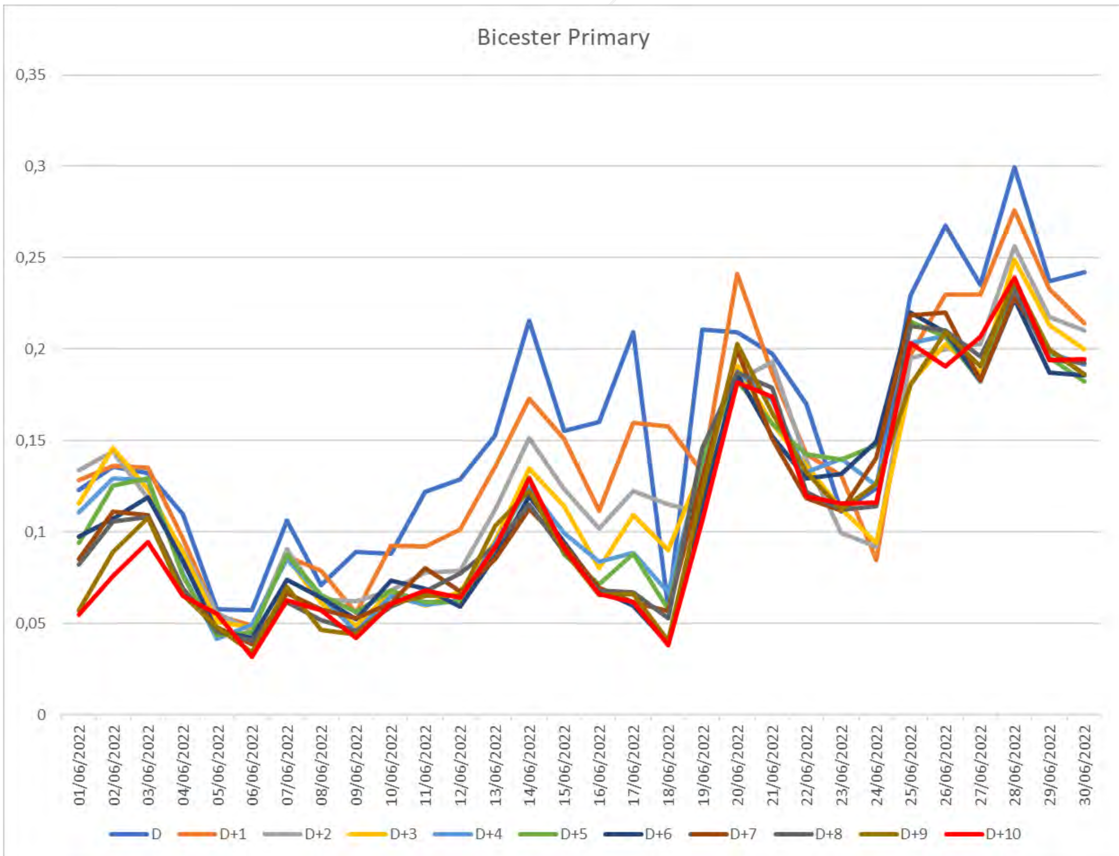


Figure 79 - Daily average relative error for all original forecasts in June 22 – Bicester Primary

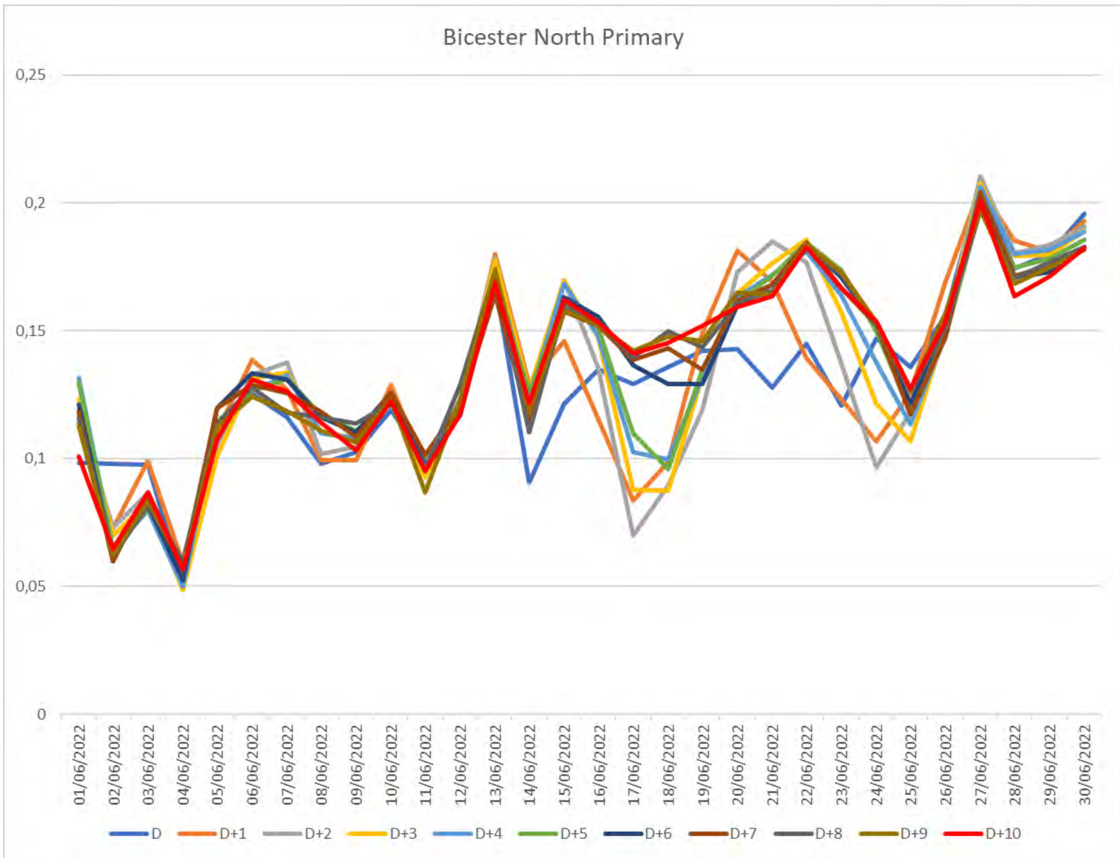


Figure 80 - Daily average relative error for all original forecasts in June 22 – Bicester North Primary

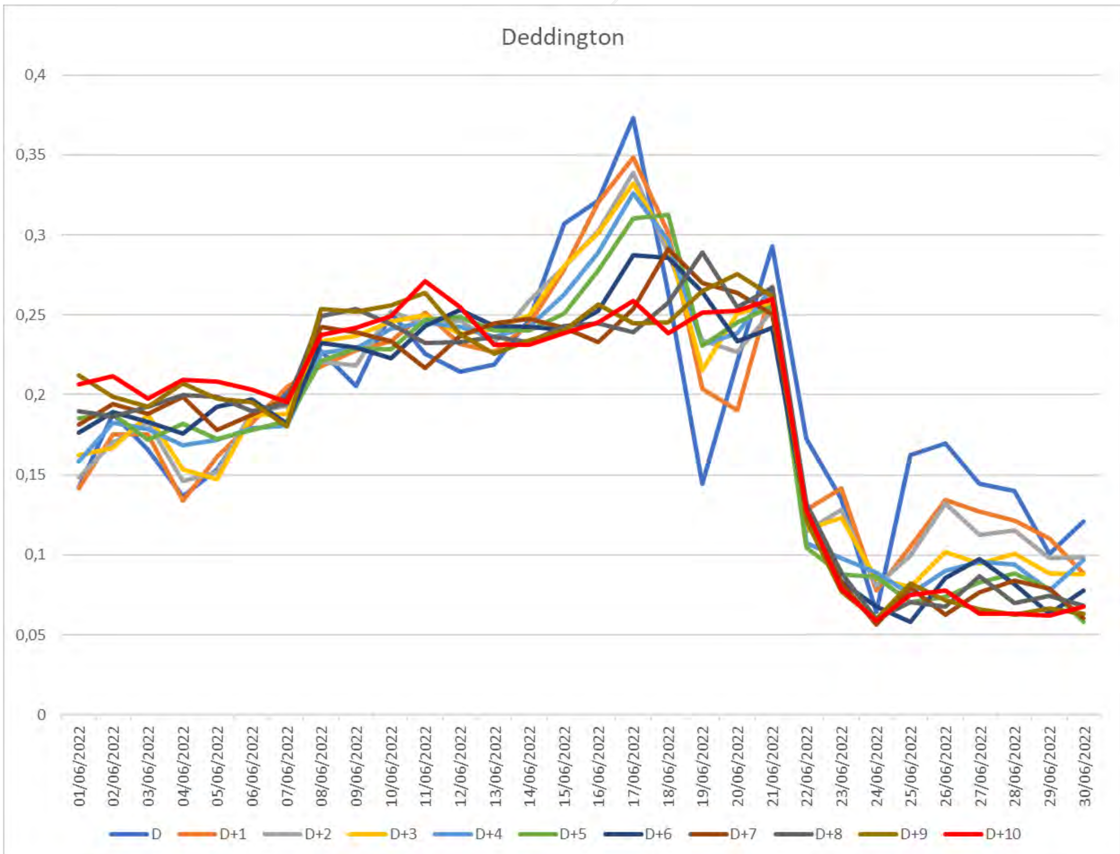


Figure 81 - Daily average relative error for all original forecasts in June 22 – Deddington

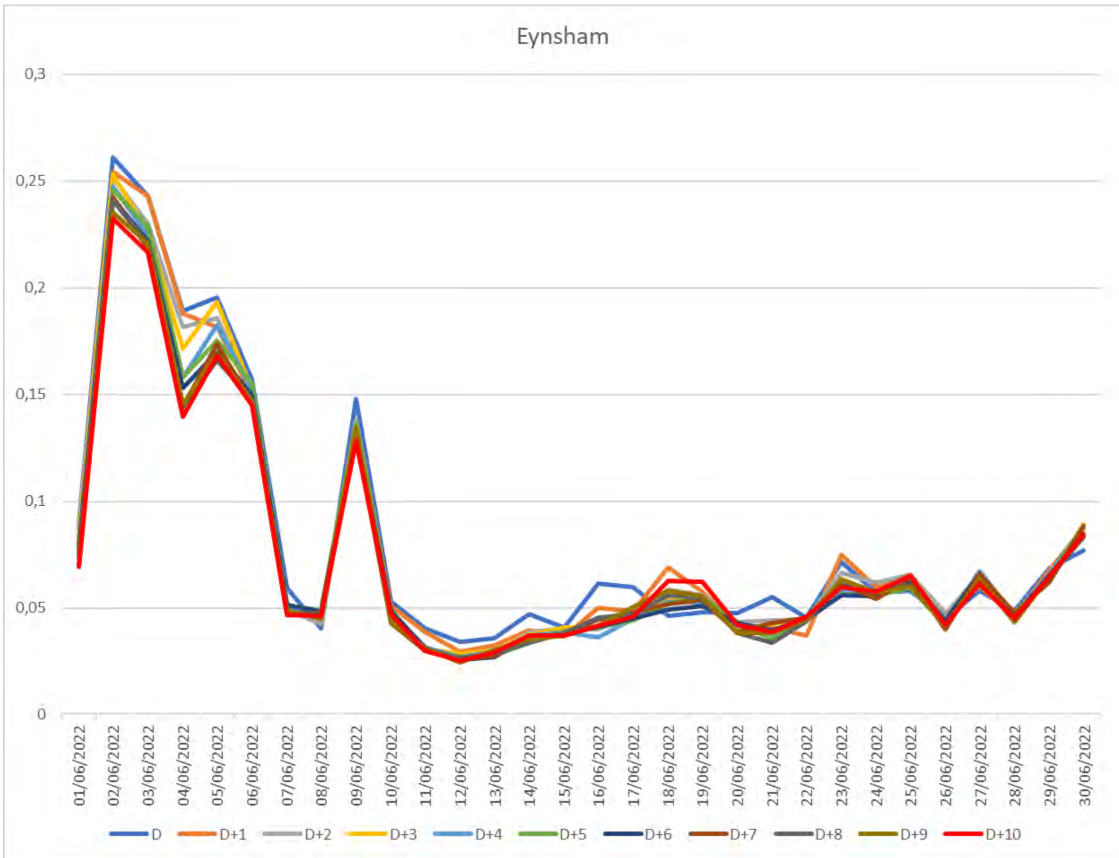


Figure 82 - Daily average relative error for all oiginal forecasts in June 22 – Eynsham

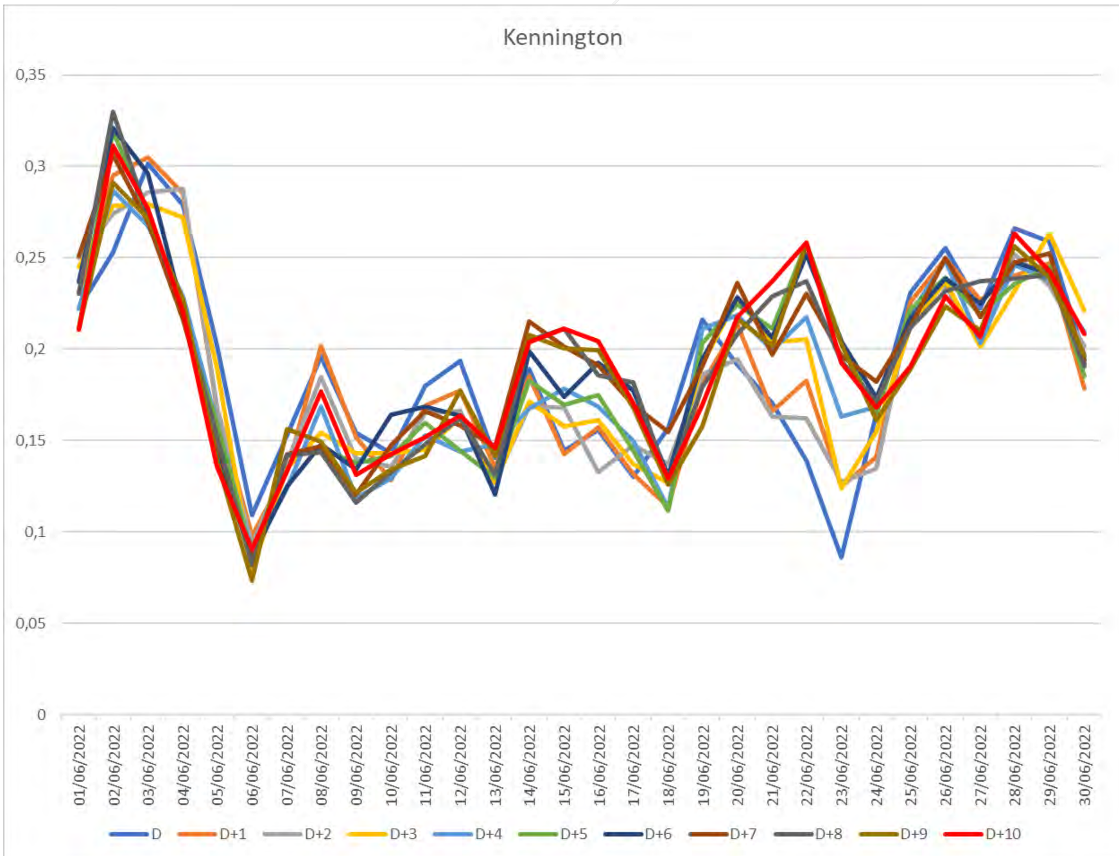


Figure 83 - Daily average relative error for all original forecasts in June 22 – Kennington

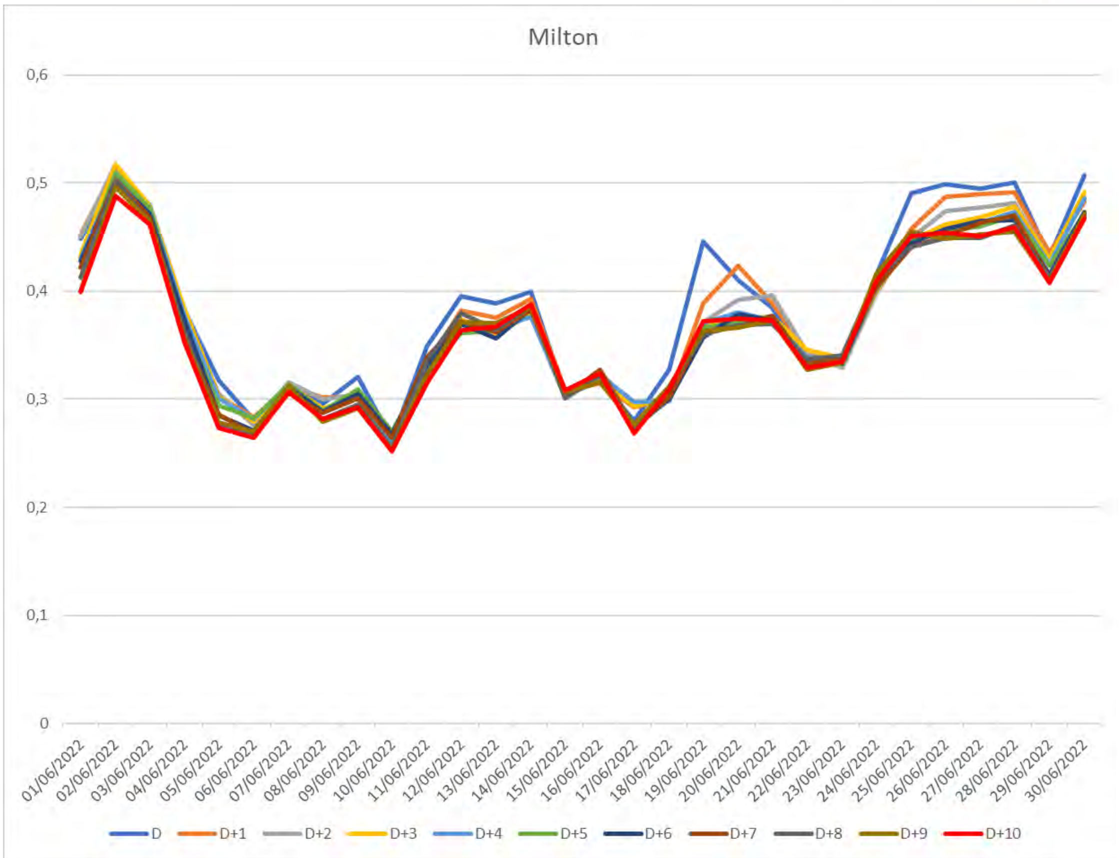


Figure 84 - Daily average relative error for all original forecasts in June 22 – Milton

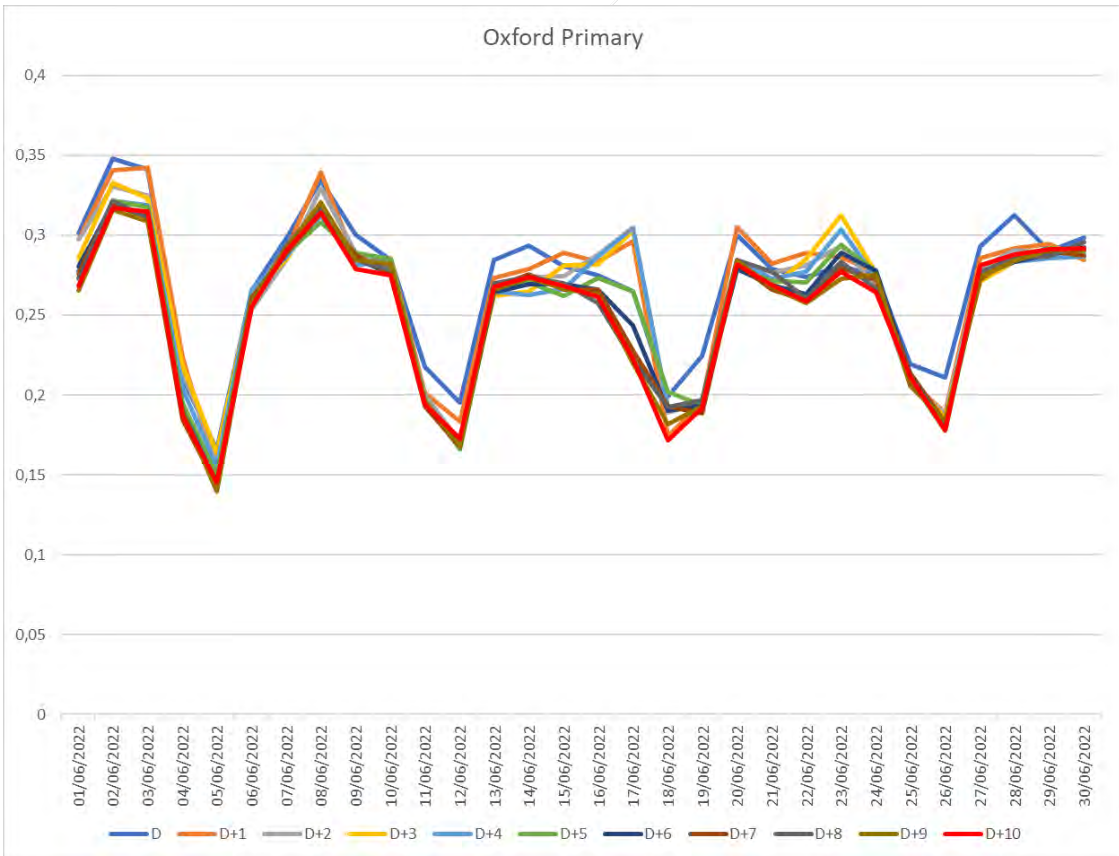


Figure 85 - Daily average relative error for all original forecasts in June 22 – Oxford Primary

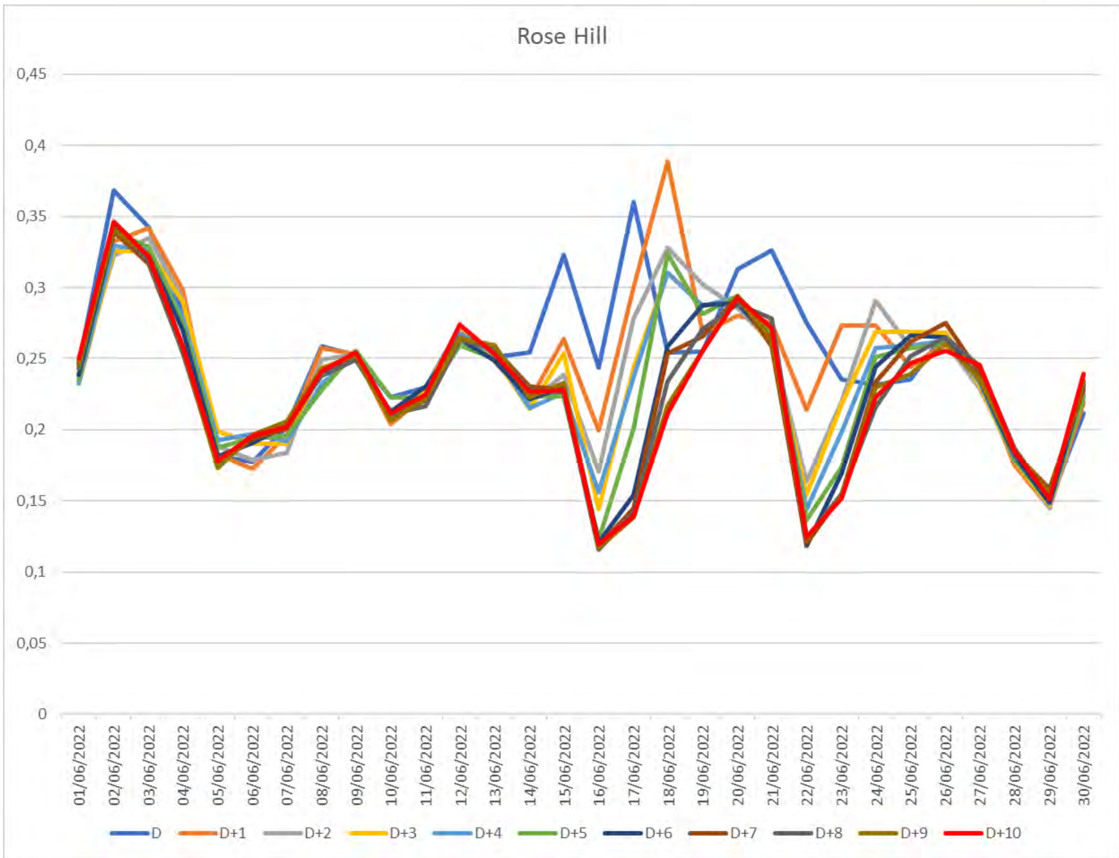


Figure 86 - Daily average relative error for all original forecasts in June 22 – Rose Hill

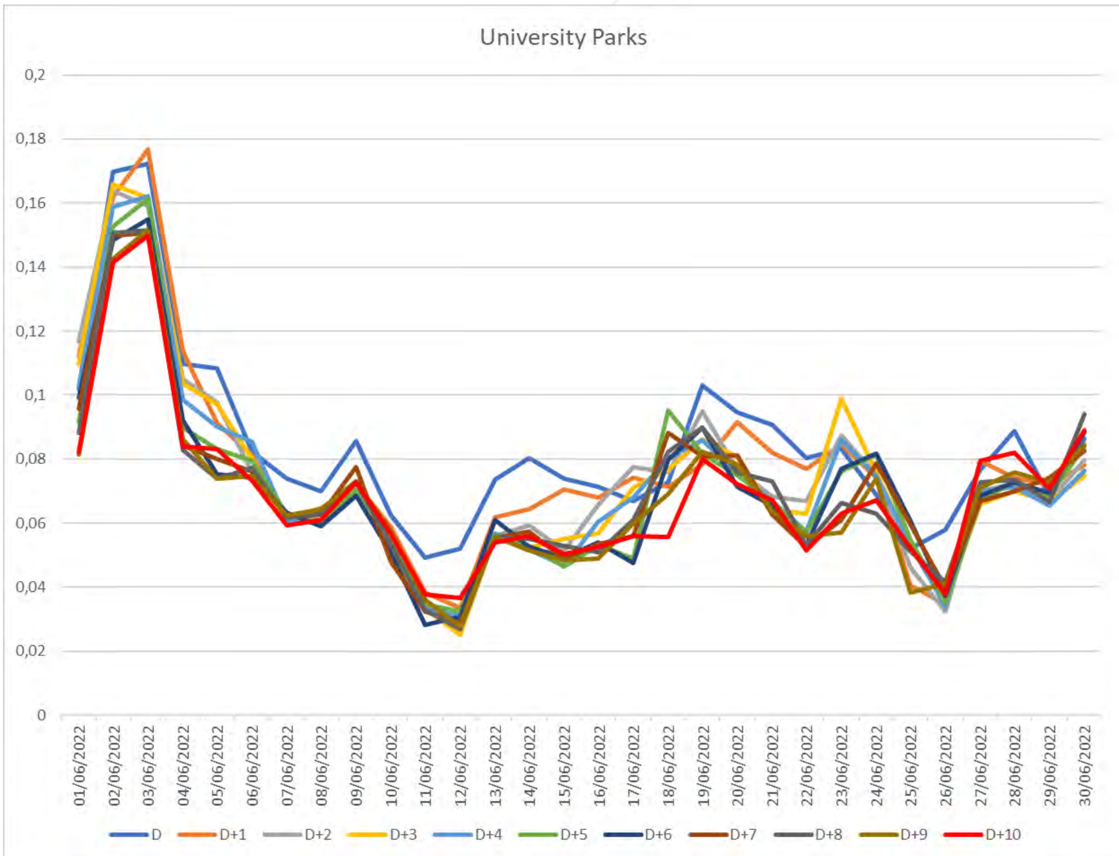


Figure 87 - Daily average relative error for all original forecasts in June 22 – University Parks

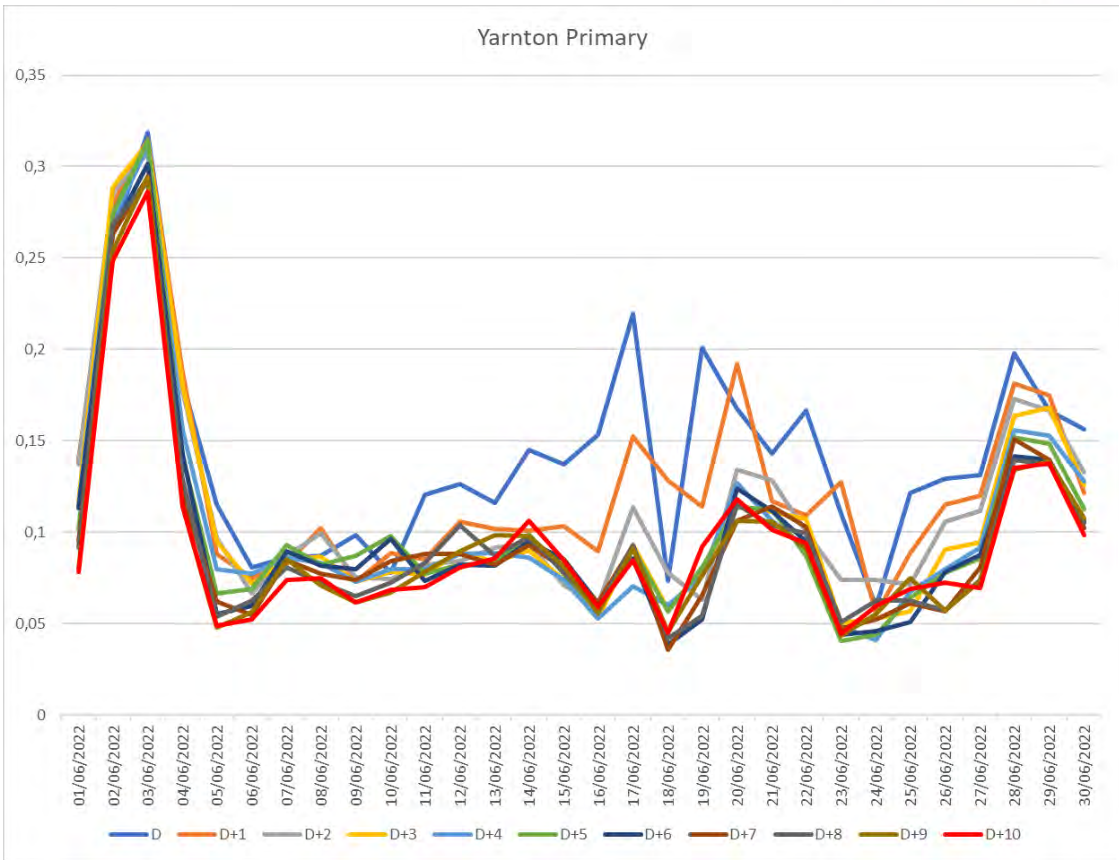


Figure 88 - Daily average relative error for all original forecasts in June 22 – Yarnton Primary

13.4.2 Optimised Demand Forecast - Results at primary substation

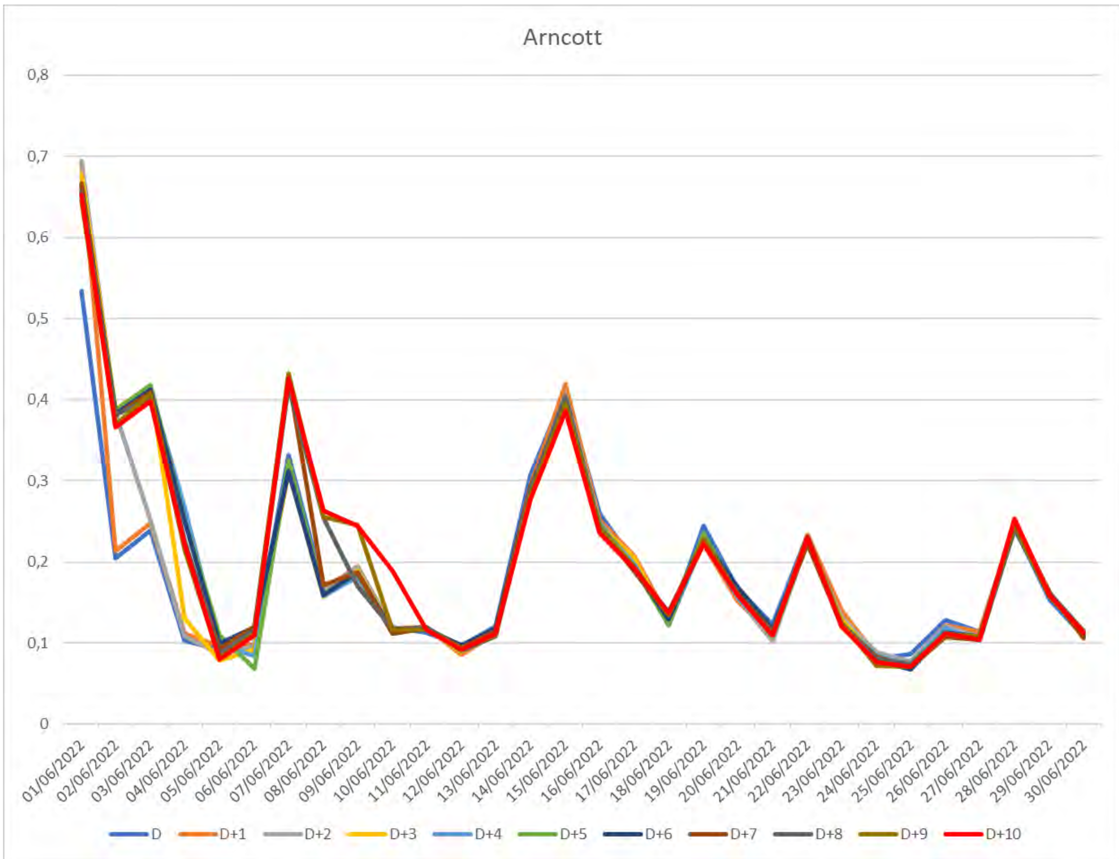


Figure 89 - Daily average relative error for all optimised forecasts in June 22 – Arncott

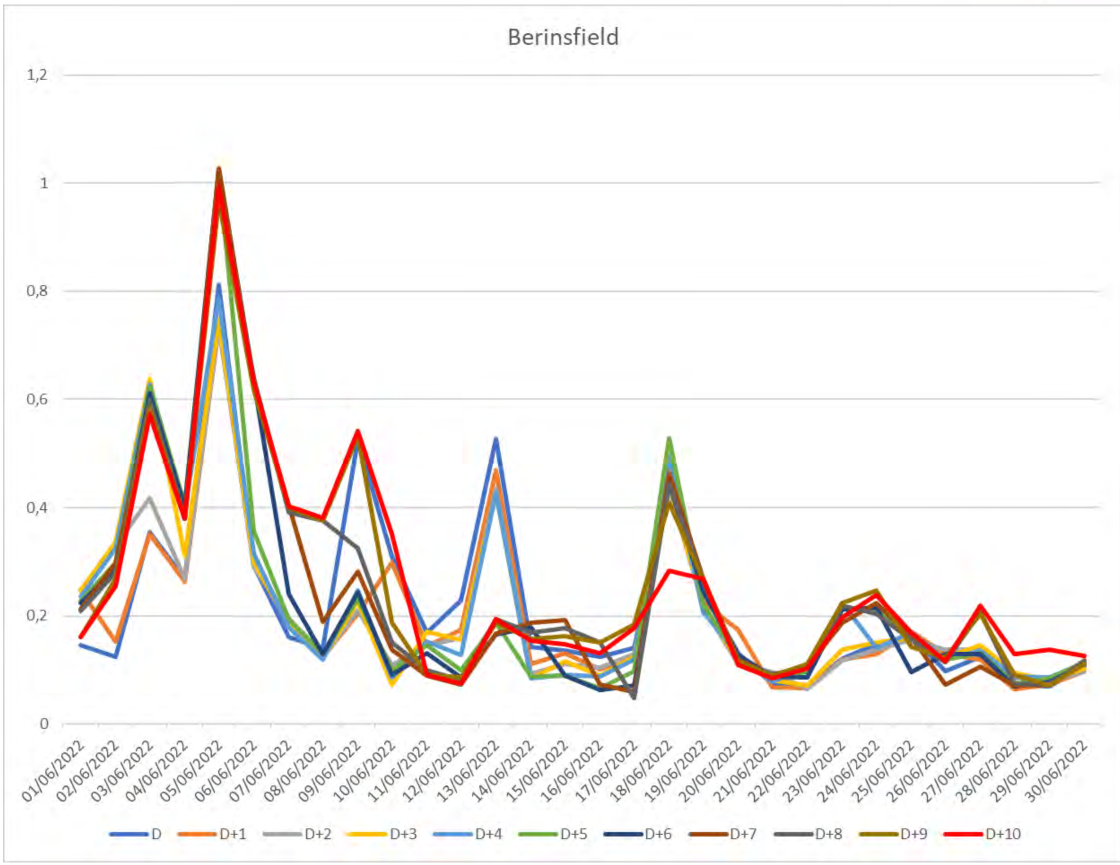


Figure 90 - Daily average relative error for all optimised forecasts in June 22 - Berinsfield

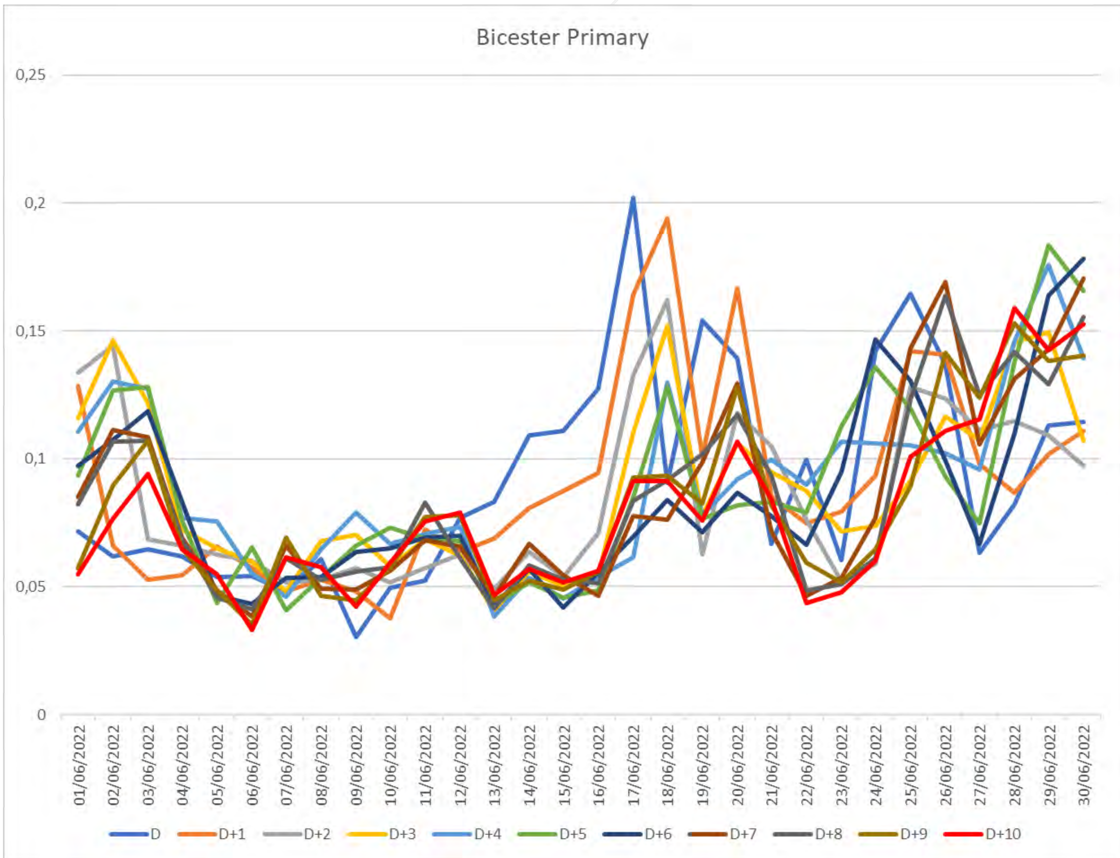


Figure 91 - Daily average relative error for all optimised forecasts in June 22 – Bicester Primary

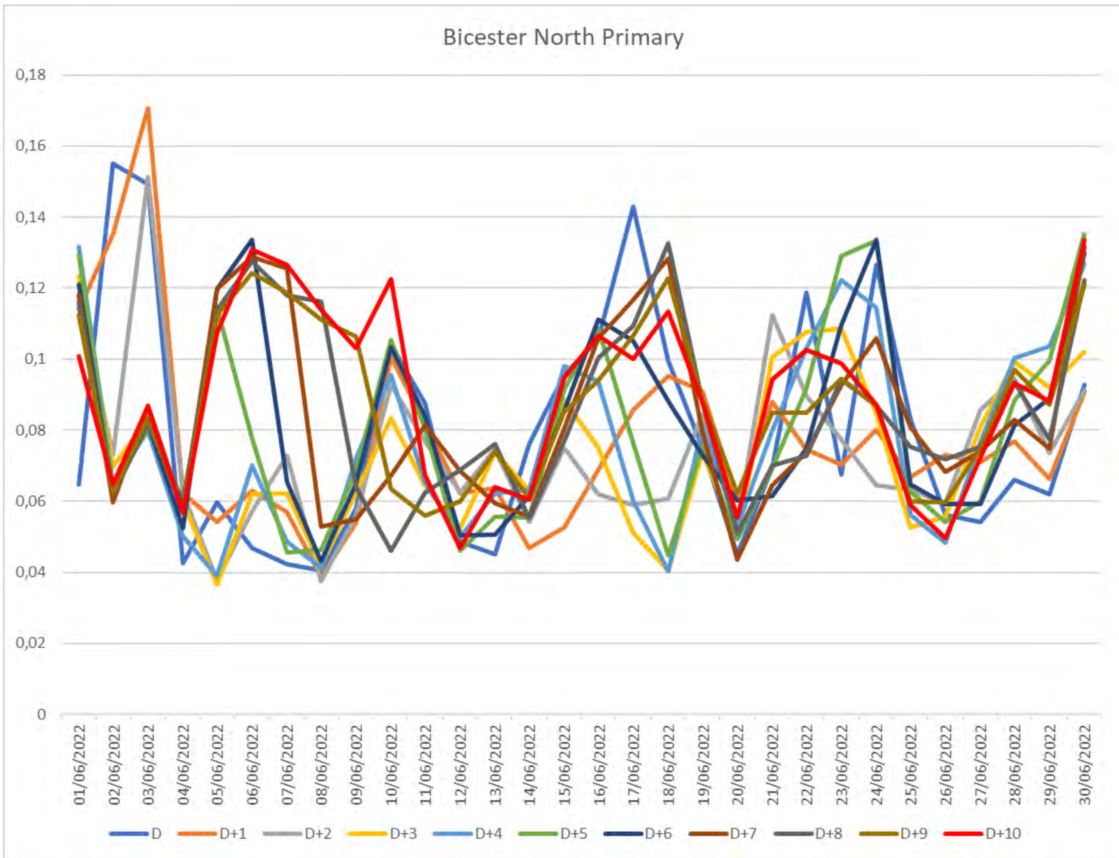


Figure 92 - Daily average relative error for all optimised forecasts in June 22 – Bicester North Primary

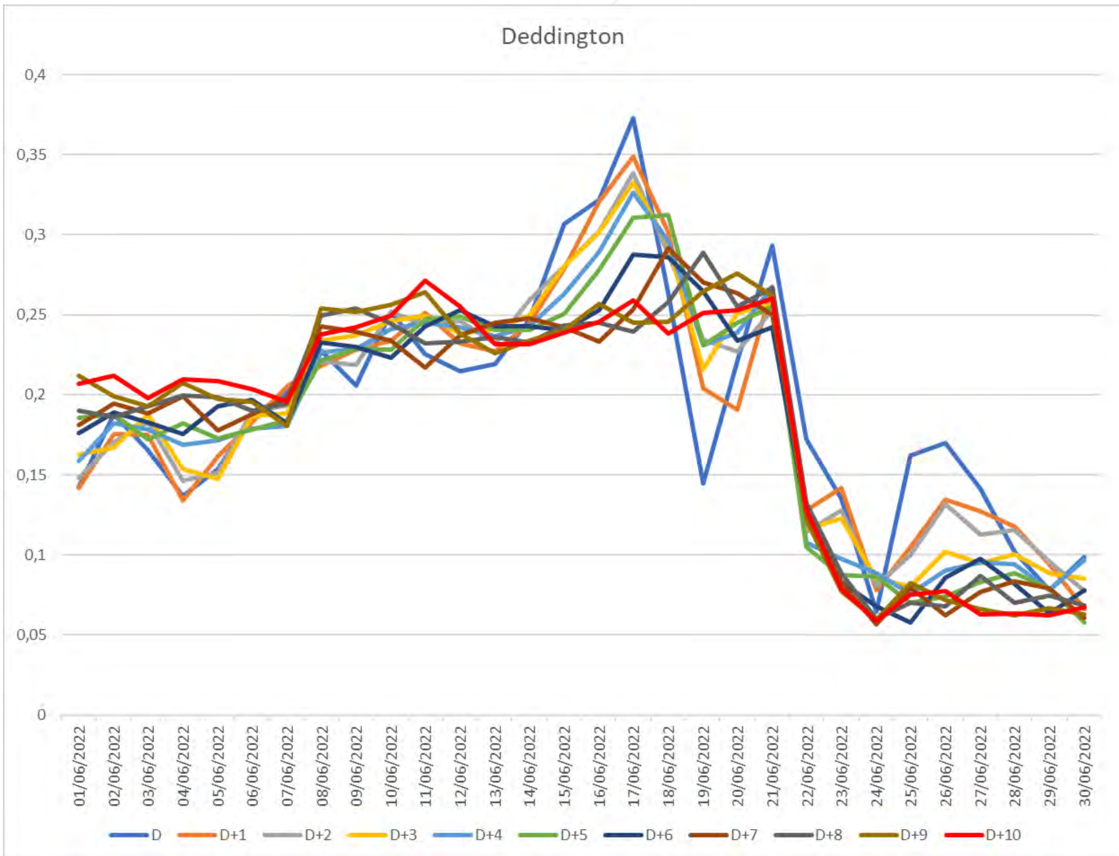


Figure 93 - Daily average relative error for all optimised forecasts in June 22 – Deddington

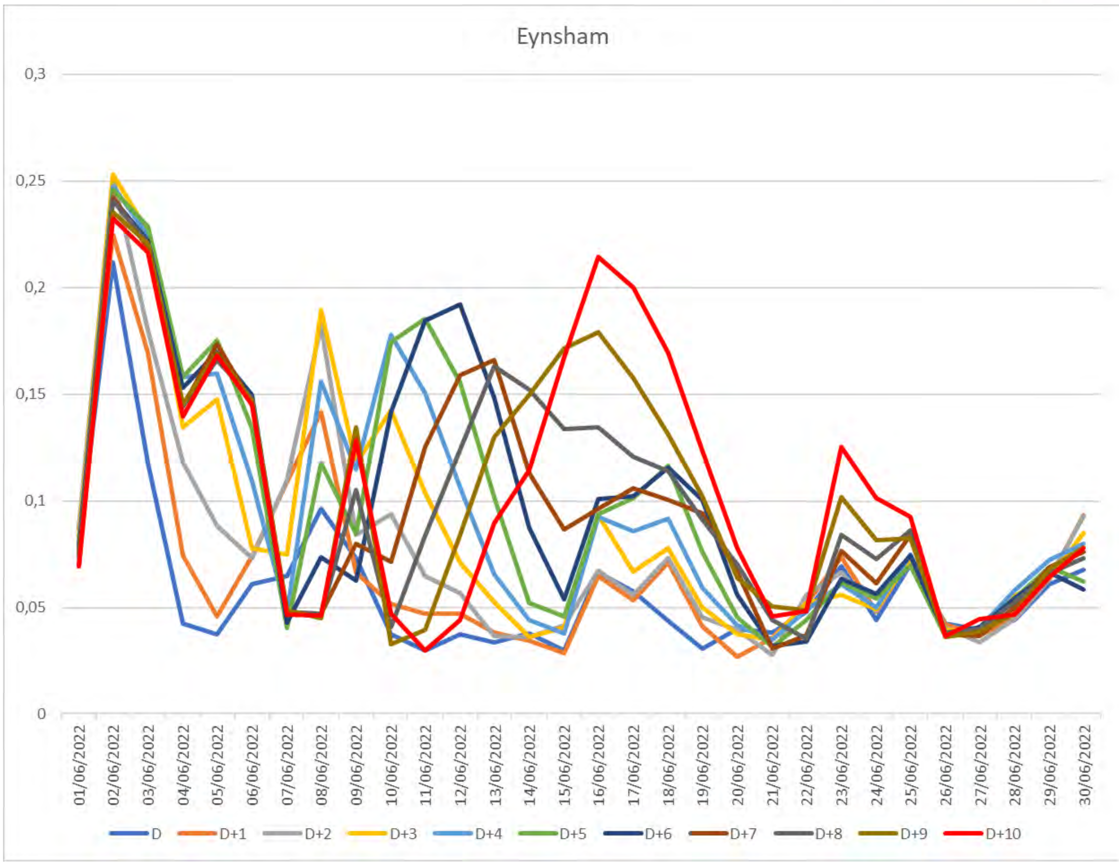


Figure 94 - Daily average relative error for all optimised forecasts in June 22 – Eynsham

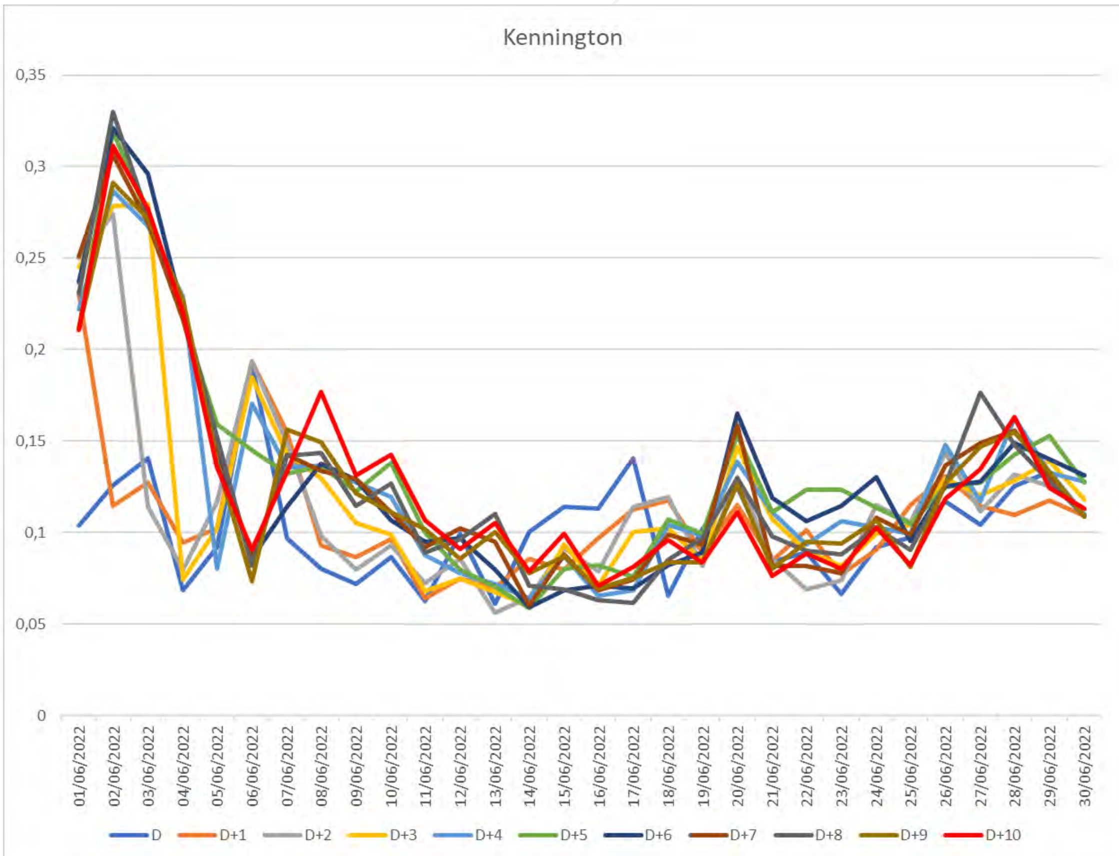


Figure 95 - Daily average relative error for all optimised forecasts in June 22 – Kennington

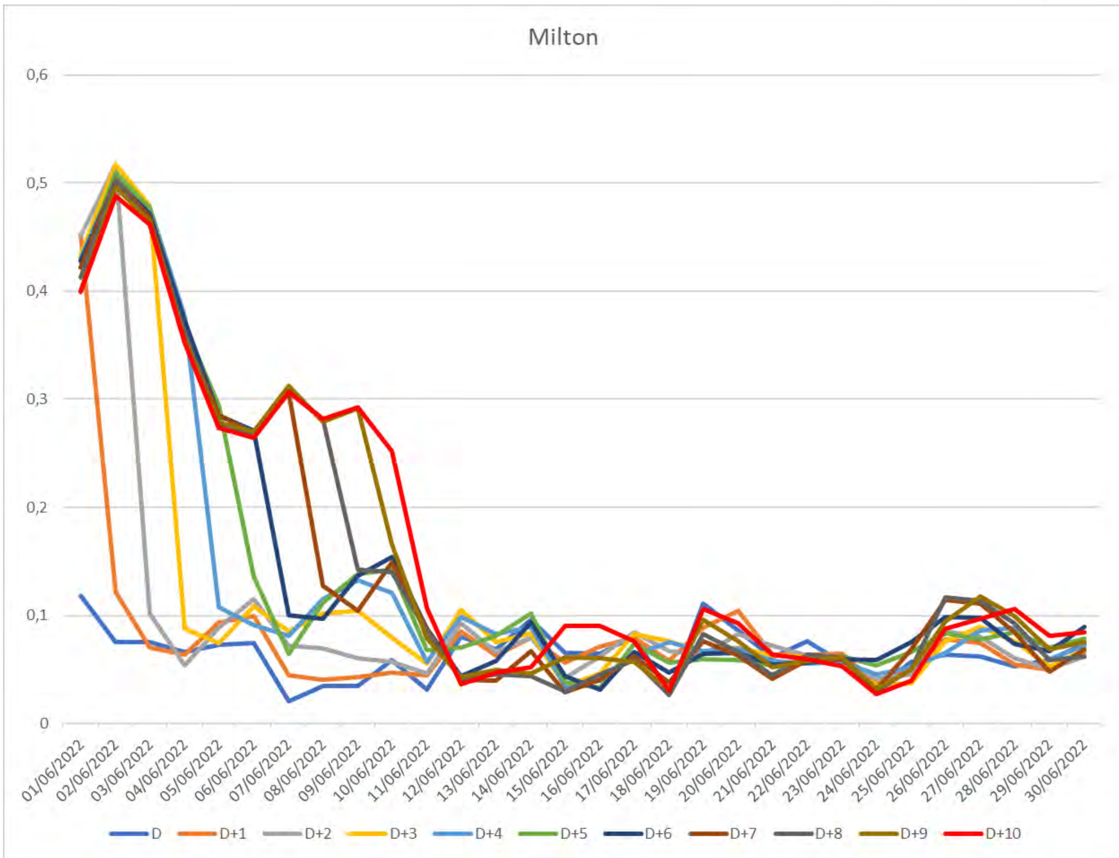


Figure 96 - Daily average relative error for all optimised forecasts in June 22 – Milton

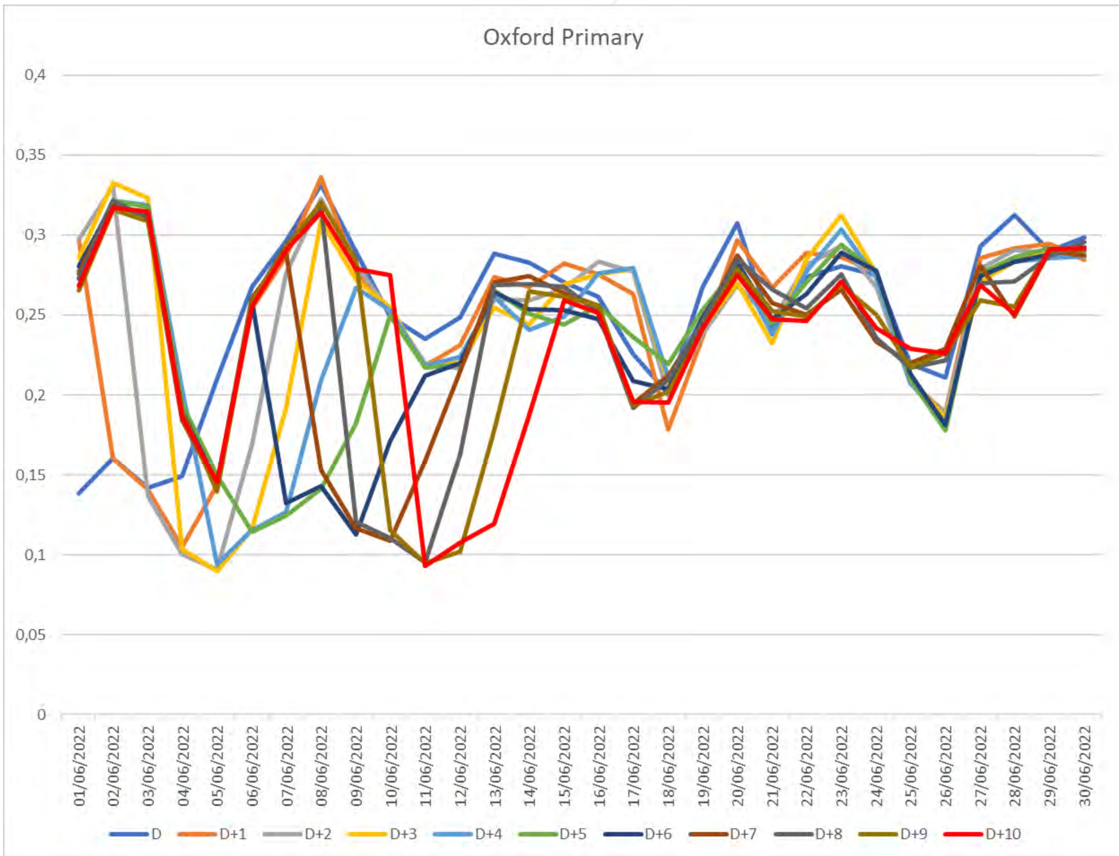


Figure 97 - Daily average relative error for all optimised forecasts in June 22 – Oxford Primary

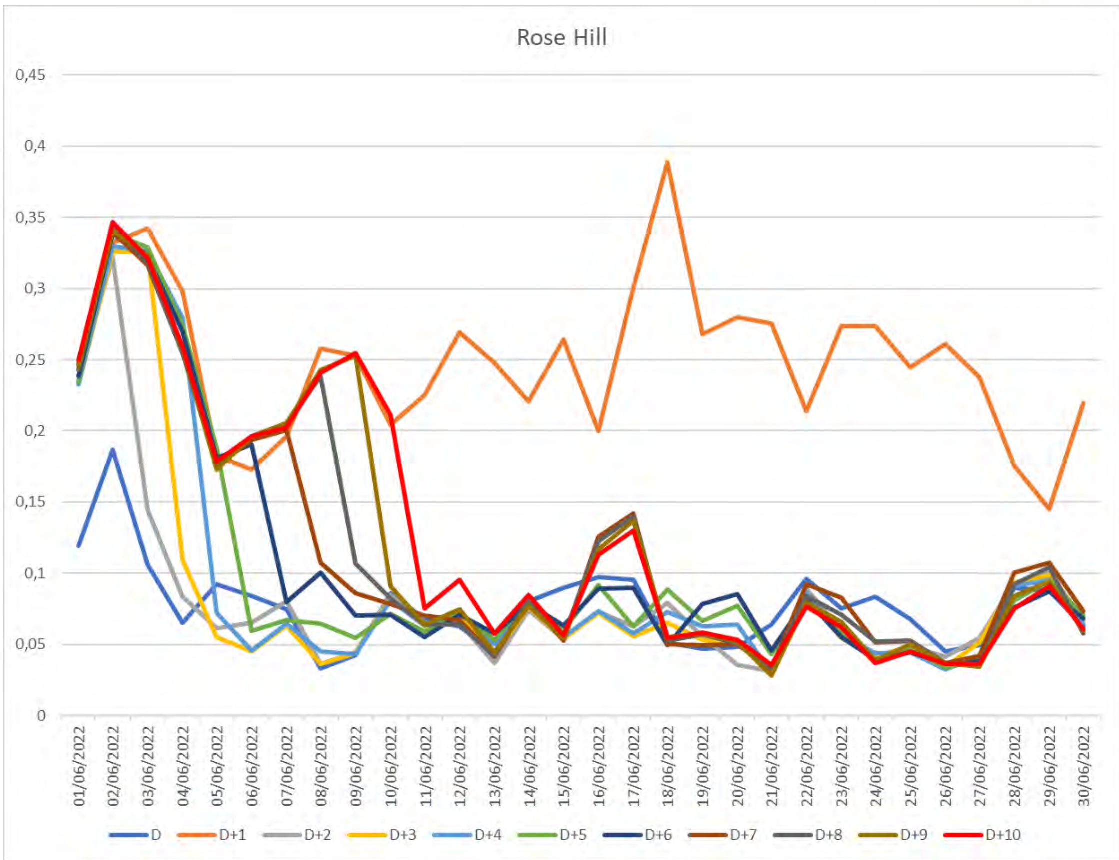


Figure 98 - Daily average relative error for all optimised forecasts in June 22 – Rose Hill

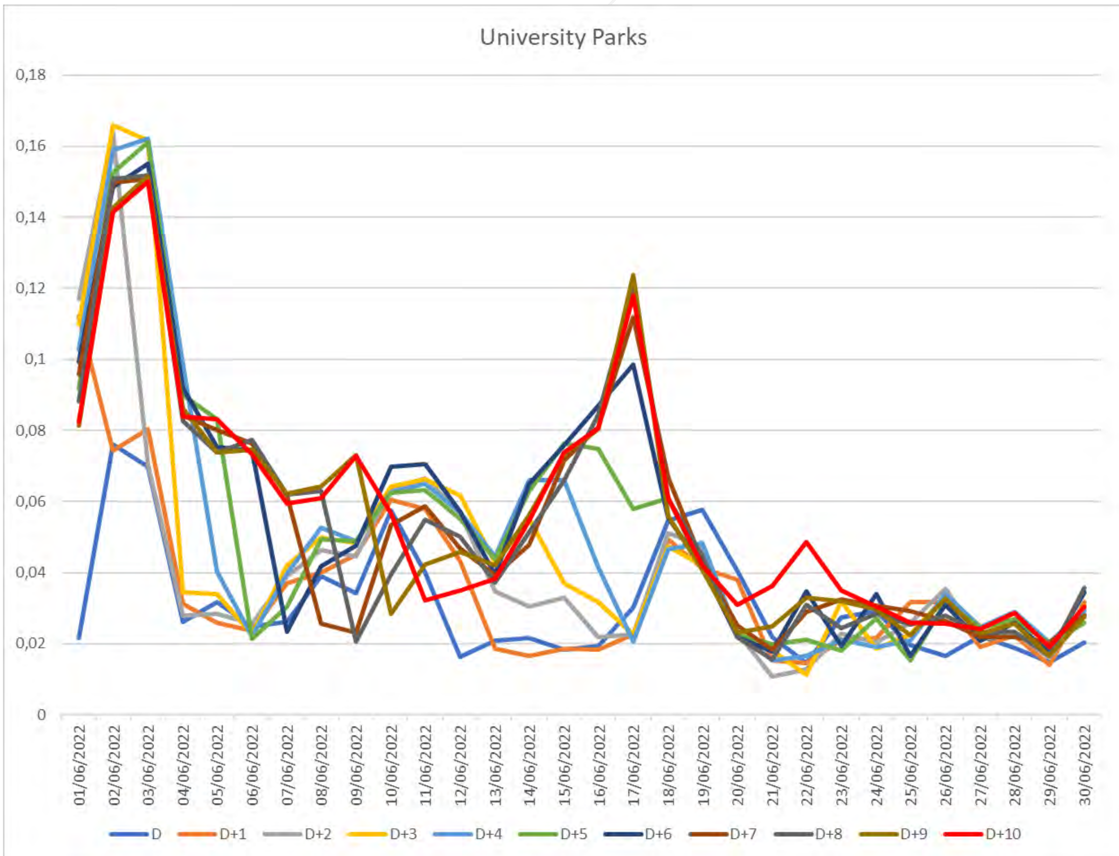


Figure 99 - Daily average relative error for all optimised forecasts in June 22 – University Parks

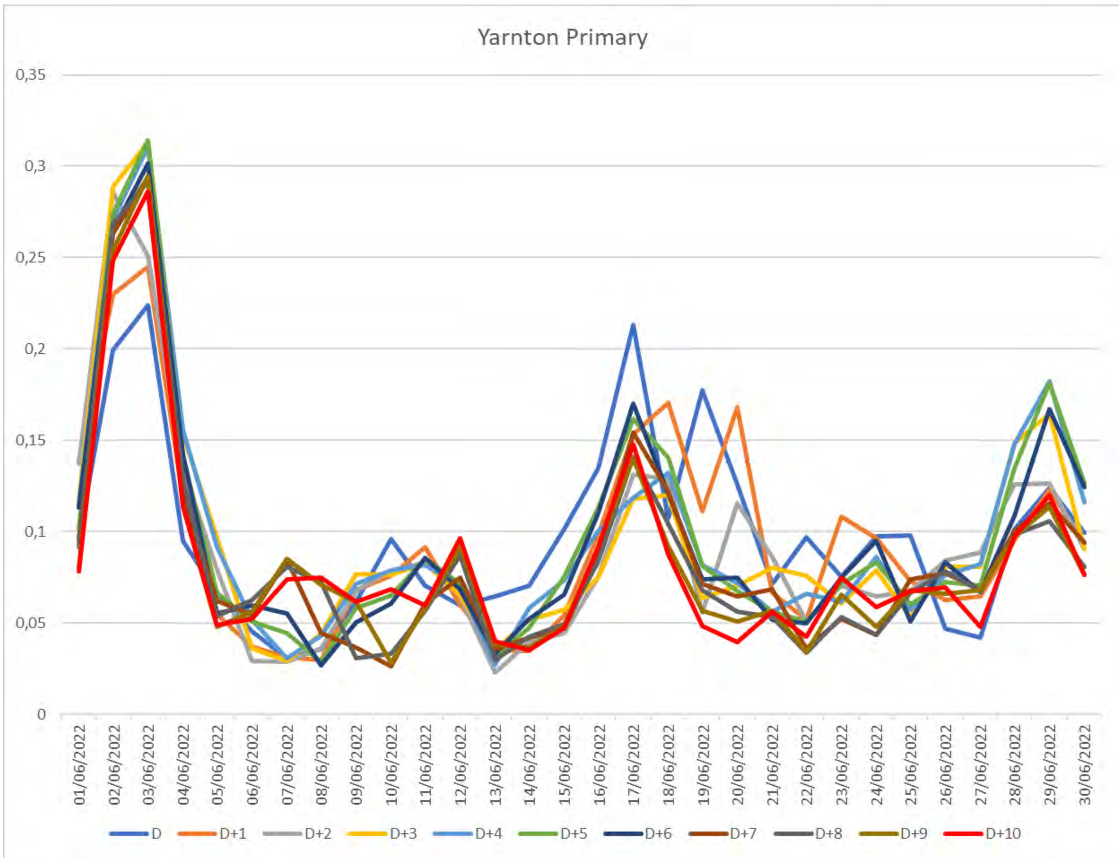
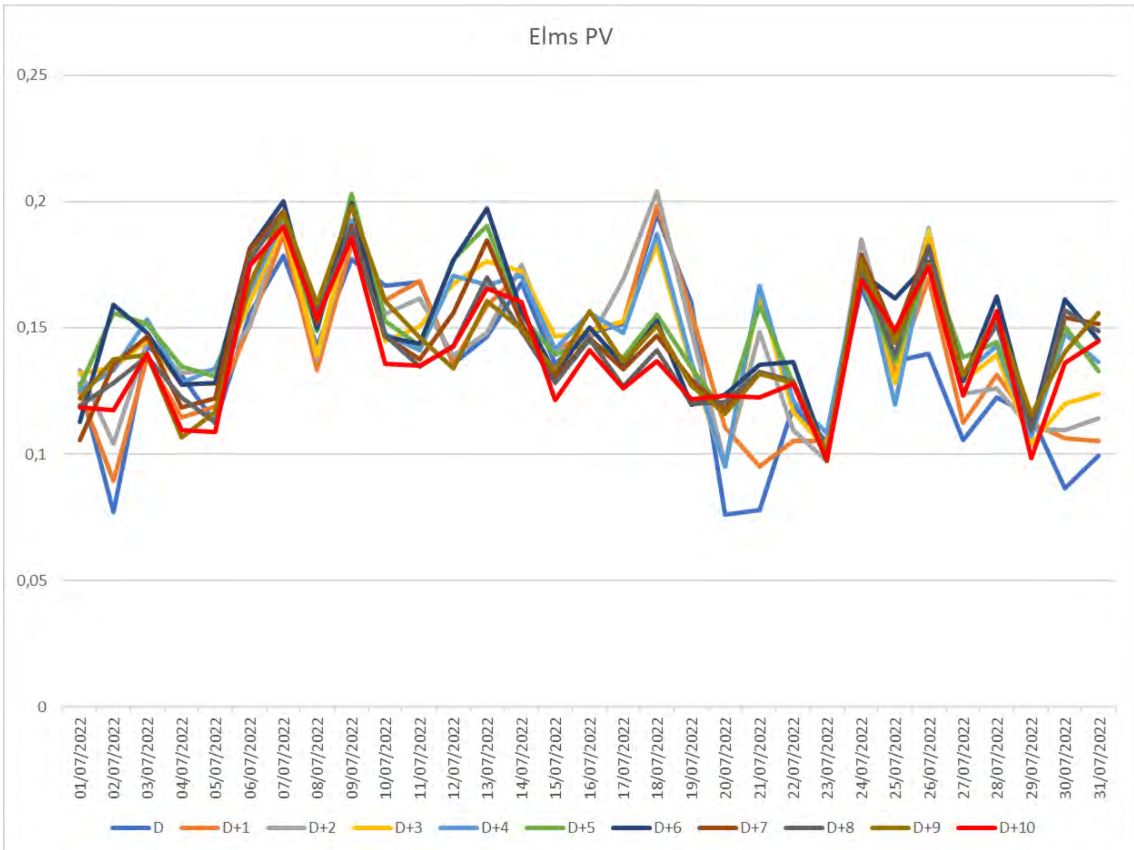


Figure 100 - Daily average relative error for all optimised forecasts in June 22 – Yarnton Primary

13.4.3 Results for generation original forecasts



- Abu Dhabi
- Belgium
- Canada
- France
- Germany
- Hong Kong
- Ireland
- Italy
- Japan
- Luxembourg
- Morocco
- Netherlands
- Panama*
- Qatar
- Saudi Arabia
- Singapore
- United Kingdom
- United Arab Emirates
- United States

* Sia Partners Panama, a Sia Partners member firm

SIAPARTNERS

About Sia Partners.

Sia Partners is a next-generation management consulting firm and pioneer of Consulting 4.0. We offer a unique blend of AI and design capabilities, augmenting traditional consulting to deliver superior value to our clients. With expertise in more than 30 sectors and services, we optimize client projects worldwide. Through our Consulting for Good approach, we strive for next-level impact by developing innovative CSR solutions for our clients, making sustainability a lever for profitable transformation.

www.sia-partners.com

transition

Moving to a smart future

Designing, developing, demonstrating and trialling new tools, data and system architecture required to enable a flexible energy network that works for everyone

https://ssen-transition.com

SSENTransition@sse.com