Smart Network Design Methodologies

Closedown Report

March 2020

Contents

0 Document control

0.1 Document history

Version	Status	Issue Date	Authors
1	First issue	11/02/20	Francis Shillitoe

0.2 Document review

Name	Responsibility	Date
Alan Creighton	Technical Lead (NPg)	18/02/20

0.3 Document sign-off

Name	**Responsibility**	**Date**
Mark Nicholson	Project Sponsor	12/03/20

1 Executive summary

Smart Network Design Methodologies (‘SNDM’) is a Network Innovation Allowance project which ran from February 2018 until January 2020. The project’s aims were to develop improved methodologies for LV network design, primarily by using smart meter data, and to develop methodologies for multi-voltage network models (EHV/HV/LV), as well as test networks, such that voltage can be studied holistically across multiple voltage levels. The project work was primarily carried out by two contractors: TNEI and Capgemini. TNEI led the novel analysis, and multi-voltage level modelling work. Capgemini led the smart meter data analytics work. The project budget was £399,510 which rose to £411,910 by the end of the project.

Smart meter data analytics

Due to the unavailability of smart meter data, arising from delays in the smart meter programme, the data analytics was carried out using the smart meter data set collected during the Northern Powergrid (NPg) CLNR project comprising approximately 8,000 half-hourly consumption time series collected between January 2011 and December 2013. The data analytics focused on three areas:

- Consumption data aggregation – Distribution Standard Licence Condition 10A (SLC 10A) requires that DNOs cannot access electricity consumption data from a single domestic customer relating to a period less than one month. We found that when consumption data from two customers is aggregated together, this sufficiently anonymises the data on days of both high and low consumption; any higher aggregation level would introduce far higher levels of error with no practical benefit. The smart meter data privacy plan submitted by NPg to Ofgem in 2019 was based on an aggregation level of two and the findings of this project were used in that submission.
- Single phase LV customer phase identification – We investigated the effectiveness of voltage correlation techniques to identify customer phase connectivity. The literature review indicated promising methods using K-means clustering, the hypothesis being that customers connected on the same phase will have similar voltage characteristics. We tested these methods using various statistical markers or characteristics (such as mean, median, standard deviation etc), but could not produce satisfactory results. One of the reasons for this could be that the input consumption data from CLNR, which was used to generate voltage data via modelling, was half-hourly. Previous research used much smaller time steps. The default time step for recording average voltage measurements on GB SMETS2 meters is 30 minutes, although this can be set to a lower value. It may be worth investigating this analysis in the future but only when: there is a large enough deployment of SMETS2 smart meters on a range of LV networks to be able to gain sufficient data, the systems are in place for the DNO to reduce the recording time interval for recording the voltage, voltage measurements across different smart meters are clock synchronised and the voltage measurement error is known.
- Generation of probability distributions for customer load – We produced probability density functions and cumulative density functions that allow the estimation of the most likely values of the shape parameters used to define the parametric distributions that model the consumption characteristics of groups of customers for a given time of day and season. These were subsequently used within the novel analysis techniques at LV workstream.

Novel analysis techniques at LV

Existing techniques to model LV networks, such as the ACE49 and the After Diversity Maximum Demand (ADMD) approaches, have a number of weaknesses. They assume that all customers of a specific type exhibit the “average” behaviour for that type of customer, whereas trial data shows this not to be the case. Another weakness is that power flows on the network are often not studied, or studied in a simplified way using a statistical approach such as ADMD.

We have developed innovative techniques which use smart meter data and are rooted in the use of a statistical model to reflect both the variability and uncertainty in demand on LV networks. More specifically, we have explored the use of a Bayesian statistic model for representing customer demand. Our method proposes that initial, or prior, probability distributions should be formed based on existing data sets from projects such as CLNR. Bayesian statistics allows for the initial prior beliefs to reflect the uncertainty which exists when trying to understand customer demand at the LV level, without specific local customer demand data. When smart meter data becomes available it can be used to update these prior probability distributions according to a technique known as Bayesian updating, to form a posterior probability distribution. It is expected that this will reduce the uncertainty in the demand estimate. This procedure can be repeated indefinitely every time new data becomes available. Eventually, the initial prior belief will have very little influence on the modelled demand. Our method also captures the impacts which these customer demands will have on the network, in terms of thermal utilisation and voltage excursion, to inform network planning and new connection designs, based on detailed AC power flow modelling. However, rather than attempting to run thousands of AC power flow simulations using Monte-Carlo sampling every time a study is run, we have proposed a method which decouples the AC power flow modelling from the modelling of demand. In addition to modelling the thermal utilisation and network voltages, the approach also estimates the probability that user defined utilisation and voltage thresholds will be exceeded.

The level of statistical knowledge required to thoroughly understand the techniques proposed is probably too advanced for the majority of engineering staff within a DNO. However, our functional specification for a minimum viable product for a new LV design tool hides this complexity from end users who are only interested in whether a thermal or voltage excursion is beyond a pre-defined risk level. We have written our LV novel analysis techniques report in a way that should be understandable by degree-educated electrical engineers, with the mathematical theory, aimed at statisticians, described in the appendices.

Multi-voltage level modelling

A methodology was developed to build Multi-Voltage Level (MVL) network models from existing data sources such as IPSA models, DINIS models, GIS data and SCADA data. Two test network models were built:

- A Northeast rural network: Norton (GSP) - Leeming Bar (132/33kV) – Thirsk (33/11kV) – Sinderby (11kV/LV) – Sinderby (LV)
- A Yorkshire urban network: Creyke Beck (GSP) – Beverly (132/33kV) – First Avenue (33/11kV) – Cranwood (11kV/LV) – Cranwood (LV)

A number of voltage management solutions were identified, assessed and then, for a subset, analysed using the test network models to resolve voltage issues arising from a set of defined scenarios including increased levels of embedded generation and demand. Solutions focussed on resolving specific design issues whilst also informing strategic voltage policy including plant specifications and target operating voltages. Whilst it is recognised that only two NPg networks were modelled and hence there are some limitations in terms of wider rollout, the results provide some key observations, learning and recommendations on how this approach could provide significant value if applied as a business as usual approach. The main findings are set out below:

- The MVL modelling approach provides most value for HV and LV models; the addition of EHV networks to the models does not provide significant additional value apart from enabling a more detailed review of the compliance of a network with the requirements set out in Engineering Recommendation P10. For example, a HV and LV MVL model allows the cross-voltage level impact of connecting generation at HV combined with high uptake of PV connected at LV to be more accurately modelled as well as testing the effectiveness of potential voltage management solutions.

- The present NPg voltage policy is suitable for existing and future loading conditions under credible operational regimes, although it should continue to be reviewed as network loading changes. The allocation of voltage drop across the HV and LV voltage levels is broadly appropriate for both urban and rural networks. However, with heat and transport electrification and increased penetration of embedded PV, rural networks are likely to be more affected by voltage issues and may require a more detailed assessment and/or a review of the design principles. There is potentially a case for different voltage design policies for rural and urban networks, however, this can lead to lack of clarity for networks that are difficult to classify.
- The tap range of primary transformers can generally accommodate a wide range of loading conditions although, for the representative networks analysed, the primary transformer tap changers are typically operating towards the higher end of the tapping range, potentially reducing the amount of generation headroom. It is recognised that this can be modified by reducing EHV operating voltages at the next of transformation.
- The assumptions related to the modelling of tapchanger deadband can be material in those situations where network voltages are approaching their upper or lower limits. The risk associated with tapchanger relays spending a significant amount of time at the extremes of the deadband range should be considered in more detail.
- The specification of and the loading on supply point transformers is such that the requirements of Engineering Recommendation P10 are satisfied.
- The modelling of NPg networks could be improved to better capture voltage behaviour e.g. by including susceptance values in the network models.

There are several voltage management solutions that can be implemented readily using equipment that is already installed or currently being installed as part of NPg’s ED1 smart grid enablers programme. These include:

- Load drop compensation could be implemented to cover various future demand and generation conditions at HV and LV although the application would be network specific and would need to consider the loading patterns of all feeders connected to those networks – LDC is more effective where loading patterns are fairly similar under normal and under contingency conditions.
- Target voltage settings on new AVC relays fitted with remote communications could be changed on a seasonal basis. The MVL modelling approach can be deployed to develop an improved understanding of when to deploy target voltage changes and/or LDC, how application of these solutions might interact including during contingency conditions, and how application of these solutions could best be reflected in both the planning and design processes.

2 Project background

The present design and modelling tools for LV systems are more simplistic than those used for HV & EHV planning. This was acceptable when the LV network was load centric, passive in nature and accurately monitored (SCADA or Half-Hour metering) datasets were not available. Modelling tools are either spreadsheet based solutions that consider a typical end user demand (e.g. After Diversity Maximum Demand (ADMD)) or software based like the LV DEBUT tool that uses annual consumption figures. These tools are based on a single "worst case" maximum demand scenario rather than considering how demand varies for customers in different geographic locations, across different seasons, across socio-economic groups and whether low carbon technologies are present. All these factors combined can lead to inaccurate and over-engineered solutions as they tend to be based on pessimistic assumptions of the LV system utilisation and operating conditions.

LV network modelling has typically been undertaken on individual LV feeders using these simplistic spreadsheet models or simplified probabilistic modelling approaches. Whilst probabilistic modelling of customer demand can provide a good understanding of the impact of varying demand profiles and embedded generation rather than behaviour at peak loading, it doesn't use actual power flow modelling to assess the thermal or voltage impact of the customer demand on networks. It also does not enable analysis of the wider LV network and the capture of any interdependencies across voltage levels from 132kV down to LV.

With the advent of LV monitoring and smart meter data at LV and growth of different low-carbon technologies, the methodologies/assumptions employed by the current analysis tools are rapidly becoming out-of-date. For instance, when carrying out voltage studies downstream of the final voltage controlled busbar (i.e. downstream of the primary substation transformer) several uncertainties are present such as the representation of busbar voltage variation with load within the local and wider network. These uncertainties have led designers to use rudimentary and deterministic assumptions when assessing voltage regulation. There is anecdotal evidence, from on-site measurements, that suggest these assumptions may lead to over-reinforcement in urban areas and under-reinforcement in some rural networks. Also, when assessing the holistic effects of advanced voltage control techniques such as Load Drop Compensation (LDC), the way that existing design tools are used is not fit for purpose as each voltage level is typically modelled independently.

3 Scope and objectives

The original scope split the project into the following five individual workstreams:

- Workstream 1 - Horizon scanning: Will conduct the required literature review to gather learnings from previous and present innovation projects in regards to this project. Selected use cases will be mapped against the requirements of this project and test networks will be identified.
- Workstream 2 - LV Network Model methodology: Will undertake the model build for the test networks identified previously, addressing the challenges of customer load definition and phase connectivity based on the various data inputs, especially smart metering data.
- Workstream 3 - Multi-Voltage Level methodology: Will enable a more holistic assessment of the impact of a wide range of network loads/states on power flow and voltages, leading to improved recommendations on voltage control and management.

- Workstream 4 - Smart Meter Data Analytics: Will define and articulate how data and analytics can assist in dealing with the challenges of utilising smart metering data in network design and planning.
- Workstream 5 - Novel Analysis technique: Will explore and compare different novel network modelling analysis techniques that could be applied and from this, develop a set of user requirements to inform a future functional specification for new power system software.

The original objectives of the project were to:

- Deliver recommendations on improved network EHV/HV/LV network build and holistic network analysis under a range of conditions.
- Provide recommendations to improve network planning and implementation of design solutions.
- Provide recommendations on how to deal with challenges for smart meter data utilisation.
- Validate of equipment specifications.
- Create a set of requirements for future functional specifications of new power system software.

4 Success criteria

At the outset of the project the following success criteria were defined:

- The network methodologies actually propose potentially acceptable solutions for smart metering challenges which will significantly improve the design and planning assumptions especially at LV.
- The network methodologies allow the modelling of more innovative solutions due to the improved knowledge and visibility of the holistic operation of the combined networks.
- The network methodologies assist in achieving the potential £5m of network reinforcement benefit due to use of smart metering data.

5 Details of the work

The project was led by Northern Powergrid's Smart Grid Implementation Unit and comprised a project manager, technical lead and smart grid engineers as required. The bulk of the work was carried out by two sub-contractors: TNEI (workstreams 1, 2, 3 and 5) and Capgemini (workstreams 1 and 4). All the publicly available reports have been reviewed by the project manager and technical lead.

5.1 Workstream 1 – Horizon Scanning

Start: March 2018, Finish: June 2018.

5.1.1 Literature review

The Literature Review is available at:

https://www.northernpowergrid.com/asset/0/document/4772.pdf

The literature review identified more than twenty relevant academic and industry papers including Low Carbon Networks Fund project reports that provided valuable insights and learning that were used in the development of the project.

5.1.2 Use cases

The project use cases are available at:

https://www.northernpowergrid.com/asset/0/document/4773.pdf

We shortlisted and completed a set of eleven use cases. These uses cases were considered as a 'high priority' given their high potential to improve the planning and design of the distribution network by leveraging advanced network modelling using smart meter data. These Use Cases have helped to identify where the novel analysis techniques developed in later phases have focused, and to inform the functional specification of the LV modelling tool.

ID	Use cases	Benefits summary
1.1	Identify and mitigate thermal violations - LV	- Improved customer satisfaction enabled by a better network performance - Improved selection of appropriate reinforcement solutions - e.g. smart solutions - Facilitates a more co-ordinated and economical approach to managing thermal violations
1.2	Identify and mitigate thermal violations - HV/EHV	
2.1	Identify and mitigate voltage violations - LV	- Reduced customers complaints - Improved customer satisfaction enabled by a better network performance - Facilitates a more co-ordinated and economical approach to managing voltage violations - Improved selection of appropriate reinforcement solutions - e.g. smart solutions
2.2	Identify and mitigate voltage violations - HV/EHV	
3.1.1	Model new connections - Generation - LV	- Improved customer services through more efficient connection process and the identification the most economical intervention where required - Improved identification of capacity for the connection of Low Carbon Technology - Improved workflow for connection application
3.1.2	Model new connections - Generation - HV/EHV	
3.2.1	Model new connections - Demand Load - LV	
3.2.2	Model new connections - Demand Load - HV/EHV	
4	Maintain the network model - HV/LV	- Enabling the delivery of benefits deriving from the other use cases
5	Monitor and manage alternative supply arrangements - LV	- Improved customer satisfaction enabled by a better application of alternative supply arrangements to minimise outages / potential outages - Operational planning activities made more efficient
6	Perform Strategic Network Modelling Analysis	- More informed decision making on long-term network investment strategy - Facilitates the development of an efficient, economical and co-ordinated system - Development of investment scenarios enabling to better deal with increased uncertainty - including the evolution toward a more devolved system operation role

Prior to completing the use cases with 'high priority' above, the project developed a long list of use case themes. The summary table below captures the main themes, their priority and to what extent the eleven use cases progressed these themes.

Use case theme	Priority	Progress	Comments
Thermal limits violations	High	●	See use case 1.x
Voltage violations	High	●	See use case 2.x
New connections	High	●	See use cases 3.x
Network model maintenance	High	●	See use cases 4
Alternative supply arrangement	High	●	See use case 5
Strategic Network Modelling	High	●	See use cases 6
Understanding Customer Demand	High	●	Covered through use cases 1.x to 6
Network Reinforcement	High	●	Covered through use cases 1.x to 6
Losses assessment	High	●	Covered through use cases 1.x to 6
System Reliability	Medium	◑	Fault detection using smart meter data is out of scope, the real-time monitoring and management of system integrity has not been covered
Outage Management	Medium	◑	Outage planning has been covered but not the real-time outage management aspect since network operations are out-of-scope
Load Switching	Medium	◑	Network planning aspect associated with switching Customers Load is covered in use case 5 for the LV networks and in use case 1.2, 2.2 and 6. Network operations not covered as out-of-scope of the project
Phase Balancing	Medium	●	Covered through use case 1.1 and 2.1
Fault Response	Low	○	Not covered as it is mainly related to Network Operations and not Network Planning. The principles are covered off in Use Case 5
Customer Contact	Low	○	Out-of-scope
Load Index Reporting	Low	◑	Some the analysis will support the regulatory reporting of the load index, particularly that related to use case 1.x
DUoS Charging	Low	○	The novel analysis techniques may provide a better view of network use of system charges

Key:

● fully covered

◑ partially covered – foundations in place

○ not covered

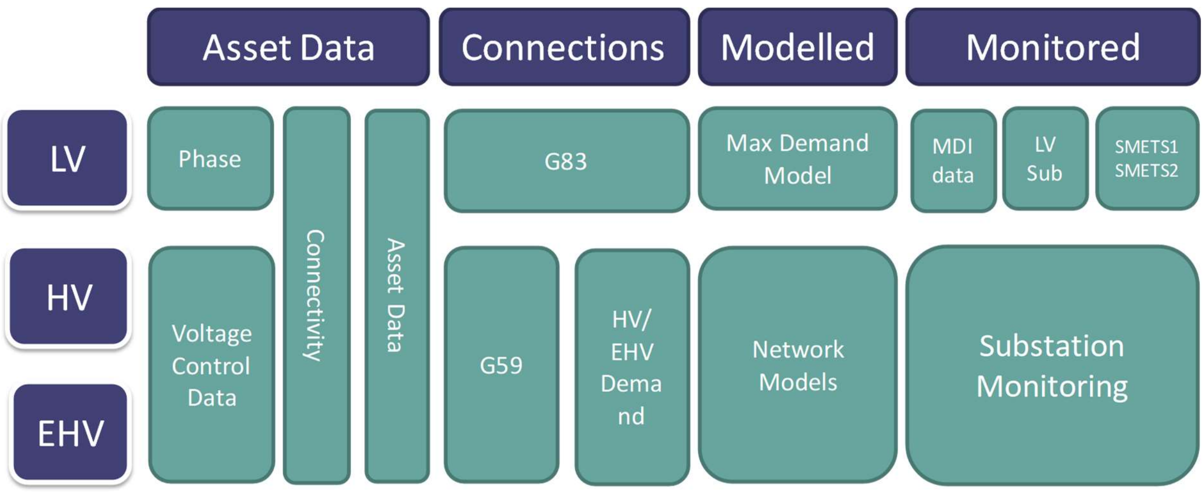
5.1.3 Data review

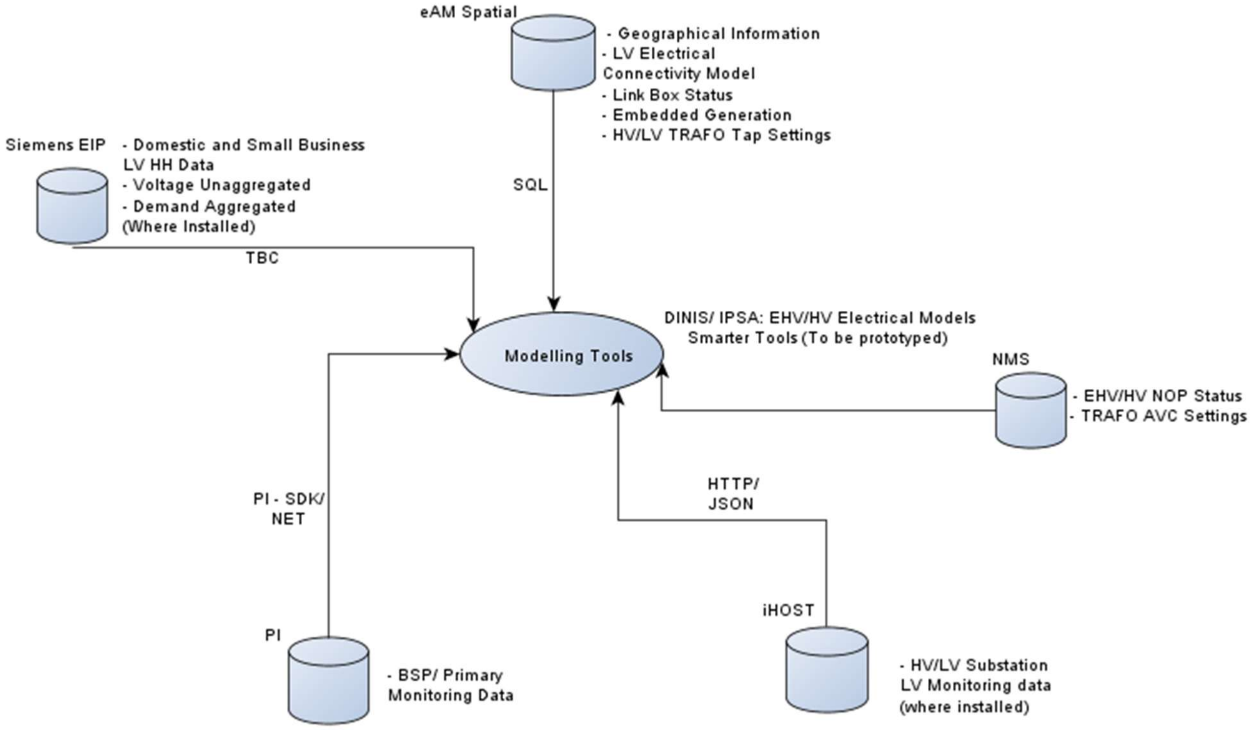
The purpose of the data review was to assess the quality and availability of the data which could be used by the project, to establish how that data could be retrieved, and to identify tools, code and scripts needed to retrieve the information. A summary of the data sources available, how they are accessed, and concerns regarding them are given here:

- Northeast EHV Network Models – IPSA. Can be used directly as the project will build test networks in IPSA.

- Yorkshire EHV, Yorkshire HV and Northeast HV Network Models – DINIS. Scripts are required to convert DINIS models to IPSA.
- LV Network Models – eAM Spatial. The electrical connectivity data can be accessed directly from the Oracle database (which eAM Spatial is based upon) using SQL. Scripts need developing to extract the data and convert to IPSA.
- EHV and Primary substation SCADA monitoring data – PI. Data can be extracted using Excel Macros which interrogate the PI database directly. PI can also be interrogated using Microsoft's .NET framework using programming languages such as C#.
- Secondary substation LV maximum demand indications – Amp2View. Data is available as a spreadsheet. However, there are significant issues with MDI data accuracy related to equipment malfunction and whether measurements taken in the field have been read and recorded reliably.
- LV network monitoring – Accessed over iHost platform. There is a low rollout of LV network monitoring. iHost data can be accessed manually via a web browser or automatically via an HTTP/JSON API.
- Embedded generation size and location – eAM Spatial. Location and type can be access using SQL from the Oracle database but there is no direct access to eAM where the generator kVA rating is stored (this data needs retrieving manually).
- Primary substation AVC target voltages – NMS PowerOn. Information needs retrieving manually from the PowerOn user interface.
- Live smart meter data – NPg's secure smart meter environment (Siemens EIP). Direct access to voltage data is currently not available. Data needs manually requesting and extracting on physical media.
- Historic smart meter data - CLNR data set. Historic smart meter data available as CSV files from CLNR project portal.

The diagrams below illustrate the data sources relevant to each voltage level, and the protocols which can be used to extract the data into modelling tools.





5.1.4 Test networks

We needed to identify representative test networks which we could model in the future work streams. The project scope was to develop two multi-voltage test networks, so we decided to choose one GSP in the NPg Northeast licence area feeding a predominately rural area and one GSP in the NPg Yorkshire area feeding a predominantly urban area.

A set of selection criteria were developed based on our project objectives and use cases and these were used to score candidate primary networks within each of these GSP groups.

The score summary sheets are shown below:

Northeast:

Criteria importance	Criteria	Darlington North - Rise Carr	Darlington North - Darlington West	Darlington North- Newton Aycliffe South	Leeming Bar /Bedale	Darlington North- Aycliffe Industrial	Bowesfield Field - Stokesley	Leeming Bar - Thirsk	North Tees (66kV) - North Tees T3/T4	North Tees (66kV) - North Tees T1/T2	Leeming Bar - Barden Friar
		6.6 kV	11 kV	11kV	11 kV	11 kV	11 kV	11 kV	11 kV	11 kV	20kV
		Urban	Urban	Urban	Semi-Rural	Urban	Semi-rural	Rural	Urban	Urban	Rural
High	Network representative of typical NPg urban or rural netwo	Medium	Medium	Medium	High	Medium	High	High	High	High	Low
High	Network representative of typical GB urban or rural networ	Medium	High	High	Medium	High	Medium	High	High	High	Low
High	Presence of low carbon technology at LV	Low	High	Low	Medium	Low	Low	High	Medium	Medium	Low
High	Presence of low carbon technology at HV/EHV	Medium	Medium	High	Medium	High	High	High	High	High	Low
Medium	Existing LV monitoring	High	Low	Low	Low	Low	Low	Low	Low	Low	Low
High	Potential to resolve known issues	High	High	High	High	High	High	High	High	Medium	Medium
Medium	Existing smart meters - SMETS1 or SMETS 2	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Medium	Current model maturity	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
	Score	42	47	44	44	44	44	53	50	47	26
	Score (%)	67%	75%	70%	70%	70%	70%	84%	79%	75%	41%

Yorkshire:

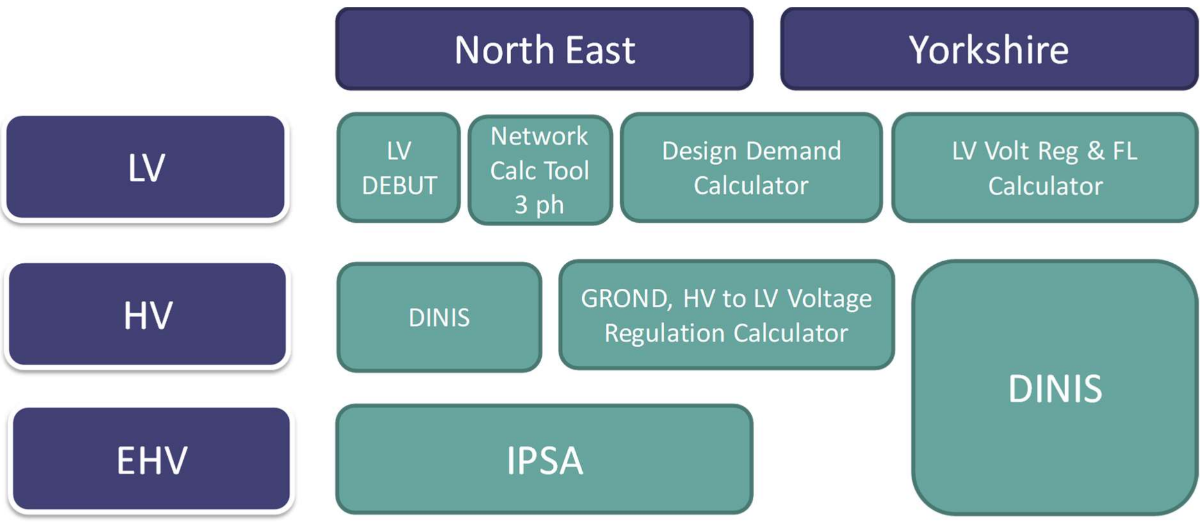
Criteria importance	Criteria	Beverly (132/33)-Southgate 5444	Beverley 132/33 - Southwood Road 5628	Driffield/Beverley-Kirkburn 6101	Hull West - Skillings Lane 8028	Beverly 132/33kV-ENDIKE LANE	Beverly 132/33kV-FIRST AVENUE	CORNWALL STREET 132/11kV	Hull South - CLARENDON STREET	Hull South - West Docks	Sculcoates A 132/11kV	SALTEND North-Hull east132/33-Westcott Street
		11 kV	11 kV	11 kV	11 kV	11 kV	11 kV	11 kV	11 kV	11 kV	11 kV	11 kV
		Rural	Mixed	Rural	Mixed	Urban	urban	Urban	Urban	Urban	Urban	Urban
High	Network representative of typical NPg urban or rural networks	Medium	Medium	Medium	Medium	High	High	High	High	High	High	High
High	Network representative of typical GB urban or rural networks	Low	Medium	High	Medium	High	High	High	High	High	High	High
High	Presence of low carbon technology at LV	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
High	Presence of low carbon technology at HV/EHV	High	High	High	High	Low	Medium	Low	Low	Low	Low	High
Medium	Existing LV monitoring	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
High	Potential to resolve known issues	High	High	High	High	High	High	High	High	High	High	High
Medium	Existing smart meters - SMETS1 or SMETS 2	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Medium	Current model maturity	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Medium
	Score	38	41	44	41	41	44	41	41	41	41	47
	Score (%)	60%	65%	70%	65%	65%	70%	65%	65%	65%	65%	75%

Based on these scores, the following multi-voltage level networks to be studied in the later project workstreams, were identified:

- Norton (GSP) - Leeming Bar (132/33kV) – Thirsk (33/11kV) – Sinderby (11kV/LV) – Sinderby (LV)
- Creyke Beck (GSP) – Beverly (132/33kV) – First Avenue (33/11kV) – Cranwood (11kV/LV) – Cranwood (LV)

5.1.5 Existing modelling tools

During the horizon scanning workstream we thought it pertinent to document the existing network planning and design modelling tools used within NPg. This allows us to consider the functionality required of new modelling tools, and which tools may become obsolete. The tools and their applications are summarised below:



Name	Use	Comments
DEBUT	LV design (Northeast)	DEBUT is an older version of WINDEBUT and is a graphical tool for LV network study. It performs voltage drop, fault current and loop impedance studies on the LV network. Also helps to identify the appropriate LV fuse required. It is not possible to model generation in DEBUT.
Network Calc Tool 3 ph (Excel Tool)	LV design (Northeast)	Also called as Northeast LV Design Spreadsheet. Primarily used for motors (starting studies), welders fusing & flicker study. Provides P28 limit, loop impedance, voltage fluctuation, fault current and fuse size.
LV Volt Reg & FL Calculator (Excel Tool)	LV design (Yorkshire)	In-house tool that calculates the voltage drop, earth loop impedance, fault current and required fuse size.
Design Demand Calculator (DDC) (Excel Tool)	LV Design and Planning (Northeast and Yorkshire)	This tool provides the Equivalent ADMD (kW) for any given LV feeder depending on the number of customers and the type of load connected. Apart from normal domestic load, it also takes in account of heat pumps and electric vehicle customers. Based on the Equivalent ADMD (kW) values, the DDC calculates the total design demand (i.e. total load) in kVA that any supplying transformer will have to supply. The DDC calculates the Equivalent ADMD (kW) values for three typical loads which are assumed to be the majority of the network load i.e. General Domestic (GD) with no electric heating, GD with Heat pumps and GD with Electric vehicles.
HV to LV Voltage Regulation Calculator (Excel Tool)	HV/LV Design and Planning (Northeast and Yorkshire)	This modelling tool helps It helps to identify the appropriate transformer tap of secondary substation (i.e. 20,000/433V, 11,000/433V, 6,000/433V) to keep the LV voltage within the statutory limits. It requires preliminary information such as the target operational voltage of the primary substation and minimum & maximum HV voltage drop on the feeder.
GROND	HV Design (Northeast and Yorkshire)	This is an HV fault level calculation and reliability study tool. It uses DINIS data for network details and requires definition of the source/upstream fault level.
DINIS Yorkshire	HV and EHV Design and Planning (Yorkshire)	DINIS is a power flow modelling software and is used to carry out load flow, fault level, voltage drop and X/R studies. The HV DINIS model is linked with the EHV DINIS model.
DINIS Northeast	HV Design and Planning (Northeast)	The EHV model is not present in the North East version of DINIS and is modelled as an equivalent grid infeed.

IPSA	HV and EHV Design and Planning (Northeast)	IPSA is a power flow modelling software and is used to carry out load flow, fault level, voltage drop and X/R studies. HV feeders from the primary are not modelled in the IPSA master network file apart from feeders with generation.

5.2 Workstream 2 – LV Network Model Methodology

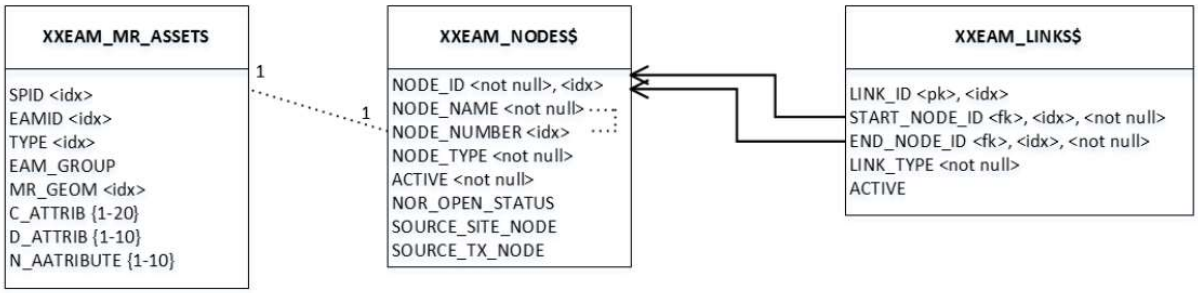
Start: June 2018, Finish: December 2018.

The objectives of this workstream were to:

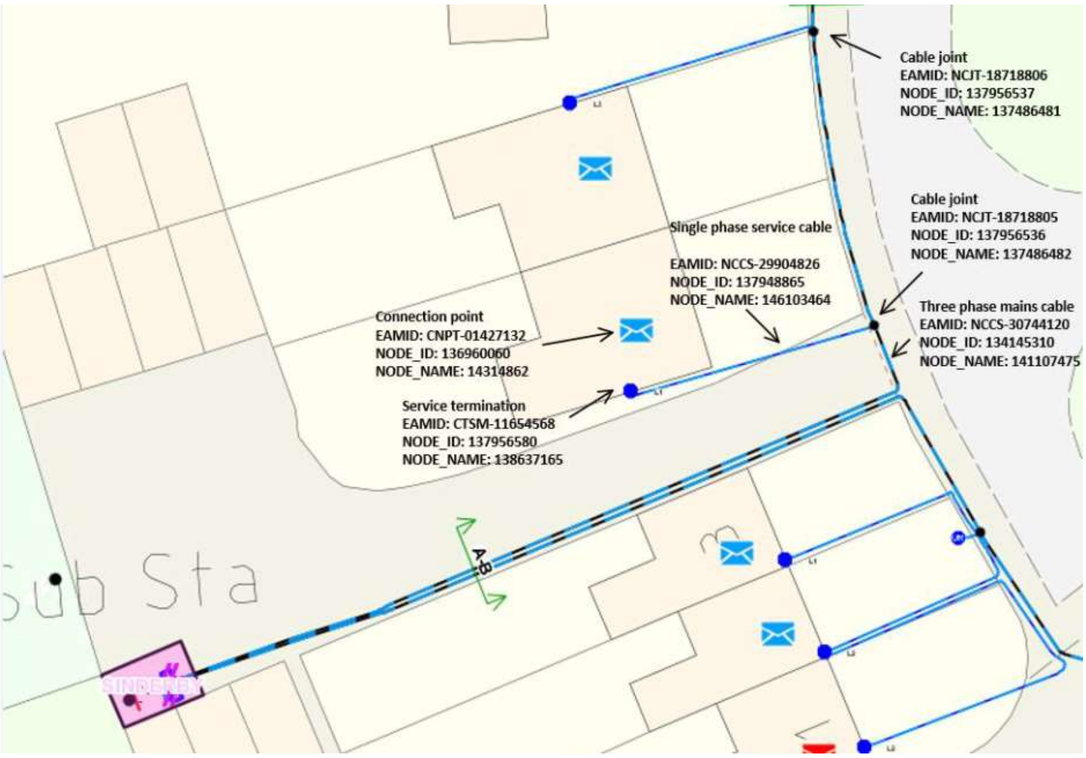
- develop and accurately represent two test LV distribution networks in power systems software (IPSA) through an automated process;
- to identify issues with the existing data sources; and
- to implement workarounds to these issues.

NPg has recently developed two databases to store asset information: Spatial and eAM. Spatial is used primarily to store asset location data and acts as a geographic information system along with the web application iSMART which is used to view the data. eAM is the central asset database. In the LV model build process we only accessed the Spatial database, as the Spatial database does contain some frequently used asset information which was sufficient for our purposes. The only exception is embedded generation kVA rating which we extracted manually from eAM.

Rather than storing data about electrical assets such as cables, overhead lines, transformers, customers etc, in separate tables, the design philosophy for the Spatial database is a meta-model, where all the data required to build an electrical connectivity model is stored in three tables:



A section of the Sinderby LV test network is shown below. Each asset such as a cable joint, cable, service termination etc is a separate node. A links table comprising start and end node IDs links each node together to form the electrical connectivity model. An assets table contains information such as cable types (300mm^2 Al waveform etc), and crucially the geometry information which allows the iSMART web application to plot circuits visually and for us allows us to extract the circuit lengths.



A recent copy of the data in the above three tables had been extracted from the NPg production database into a development database so we were able to develop and test scripts in the development environment. We developed SQL scripts using 'views' and 'table joins' to extract the electrical connectivity information for our test networks into CSV files. We developed scripts in Python to import these into IPSA. We created a separate data source containing cable types with their related impedance data as in IPSA each circuit is modelled explicitly with its impedance. This differs for other tools such as DIgSILENT Powerfactory, where circuits are modelled with their lengths and cable type, and the cable type impedances being modelled in the power systems tool itself. The scripts allow the creation of three phase balanced or single phase unbalanced networks depending upon whether customer phase connectivity information was available. Comprehensive internal documentation was produced to explain how to use these scripts, along with a detailed table of data quality issues and possible workarounds/solutions. Ideally, most of these would be fixed by cleansing the source data in Spatial. Some of these issues are presented below:

Issue	Temporary workaround	Long term fix
Embedded generation kVA rating not available in assets table	Entered manually once network build in IPSA was finalised.	Embedded generator kVA rating could be stored in assets table in Spatial in the future.
Phase connectivity information missing in many cases	Field staff identified customer phases manually at Sinderby using Hasys phase identification equipment. The information was entered manually.	Identify phase connectivity for all customers by going out into the field using Hasys units and/or use automatic methods based on smart meter data. When customer phase connectivity is identified it should be updated in Spatial and a record kept as to how this was determined.
Transformer start and end nodes often not stored correctly in Spatial. Some nodes especially those in substation sites are missing substation and transformer reference IDs	Information manually fixed in csv files.	Will require manual data cleansing in Spatial.

Connection Point nodes at customer premises are being entered into node-branch matrix as a branch since the Connection Point is linked to a Service Termination but without any physical line connecting them.	Connection Point nodes are ignored in the script and excluded in the IPSA model.	
Easting and northings co-ordinates for some nodes in the assets table do not match those in the circuit geometric data.	Use the geometric data as the correct information source.	
Cable types missing from some mains and service cables in Cranwood and Crandyke data set	Refer back to data sources used prior to Spatial such as Main Records.	Will require manual data cleansing in Spatial. Where information is not known, a conservative assumption should be made.
Several LV mains joints, service joints, cables and unmetered customers are marked as abandoned but were contained in the data export	Manually removed from CSV files	Extract process should ignore these assets.

5.3 Workstream 3 – Multi-Voltage Level Model Methodology

Start: July 2018, Finish: May 2019.

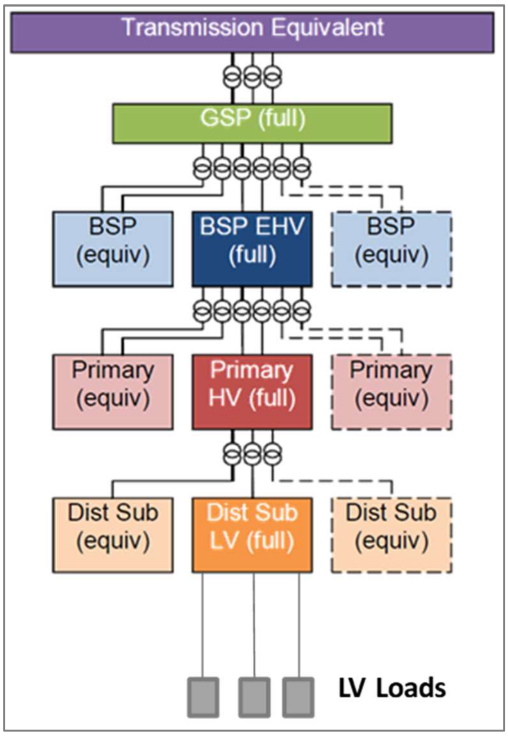
The objectives of this workstream were to:

- develop and test a methodology to build a multi-voltage level distribution network in power system analysis software using, where possible, semi-automated processes;
- develop a multi-voltage level network model which is suitable for holistic analysis of voltage behaviour across different voltage levels;
- identify issues with current DNO network data systems which pose challenges to developing a robust multi-voltage model, to learn lessons, and to provide recommendation on how to address these challenges in the future; and
- identify any data model discrepancies which need to be correct in the existing data systems.

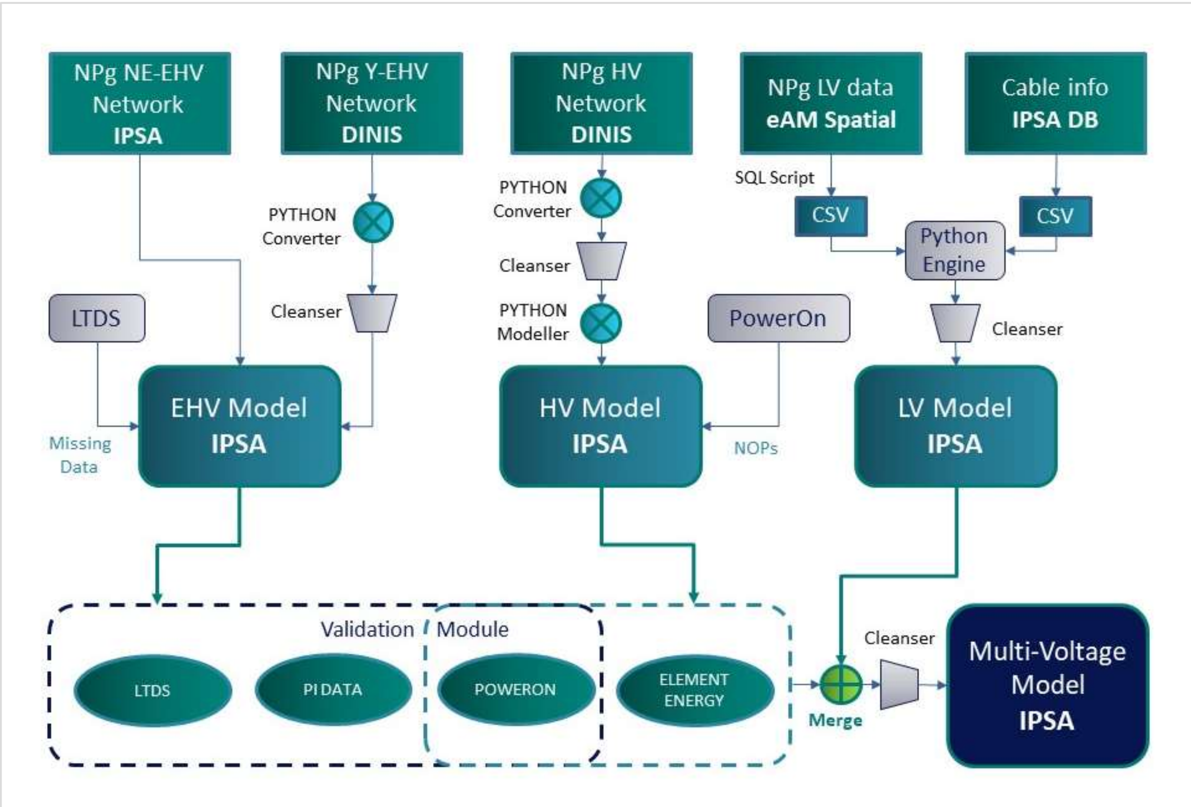
The Horizon Scanning workstream identified two multi-voltage networks which would be developed in workstream 3:

- A Northeast rural network: Norton (GSP) - Leeming Bar (132/33kV) – Thirsk (33/11kV) – Sinderby (11kV/LV) – Sinderby (LV)
- A Yorkshire urban network: Creyke Beck (GSP) – Beverly (132/33kV) – First Avenue (33/11kV) – Cranwood (11kV/LV) – Cranwood (LV)

The multi-voltage networks represent a thin slice through the network at each voltage level with parts of the network not fully modelled represented by equivalents:



Creating the multi-voltage networks involved a range of different data sources for input and validation. The network models at each of the voltage levels were developed from different sources and merged together. The figure below summarises the multi-voltage network model build process.



The end result of the process, was the production of two multi-voltage network models developed in IPSA. The key steps in the model production process are described below:

1 EHV model creation

- Conversion of Yorkshire DINIS model into IPSA. The first step was to generate a CSV export file from DINIS. Only the immediate and adjacent EHV networks to Creyke Beck/Beverley/First Avenue were of interest and required in the multi-voltage level models. It was important to carefully define the boundaries of networks of interest before the DINIS data extract process was started.
- The ‘dinisreader’ Python script, available in the IPSA installation files was then used to convert the DINIS CSV export file into IPSA.
- After the model was converted to IPSA, it was necessary to manually check and update the model:
    - It took some effort to rearrange the model elements into a Single Line Diagram as all the elements for a particular substation are stacked on top of one another.
    - The busbars have alphanumeric identifiers and these need to be changed to the busbar names.
    - The Long Term Development Statement (LTDS) was consulted to manually fix some parts of the network which branched off in the IPSA model but were not captures in the DINIS extraction process.
- For the Northeast IPSA model the part of the network of interest (Norton/Leeming Bar/Thirsk) was identified and the model reduced to only contain this immediate and adjacent networks. Interfaces were replaced by equivalent loads.
- For both the Northeast and Yorkshire models the equivalent loads were mapped to SCADA data provided by the PI system.
- A load flow was run on each network. IPSA generated warnings or did not converge for some parts of the network where circuit impedances had not been converted correctly. Some manual fixing was required at this stage.

2 HV model creation

- Both the Yorkshire and Northeast HV networks are modelled in DINIS, so the conversion process into IPSA was similar to that followed for the Northeast EHV networks.
- Secondary substation load data was taken from load growth models provided to NPg by Element Energy.
- HV circuit breakers, switches and Normally Open Points were not imported during the conversion process, and it was necessary for these to be added manually by viewing the PowerOn NMS data.

3 LV model creation

- The LV networks were created in Workstream 2.

4 Model Validation

- After the network model build, the models were verified and validated against available SCADA data, network diagrams, current running arrangements and distribution substation loading based on the Element Energy Load Growth model.

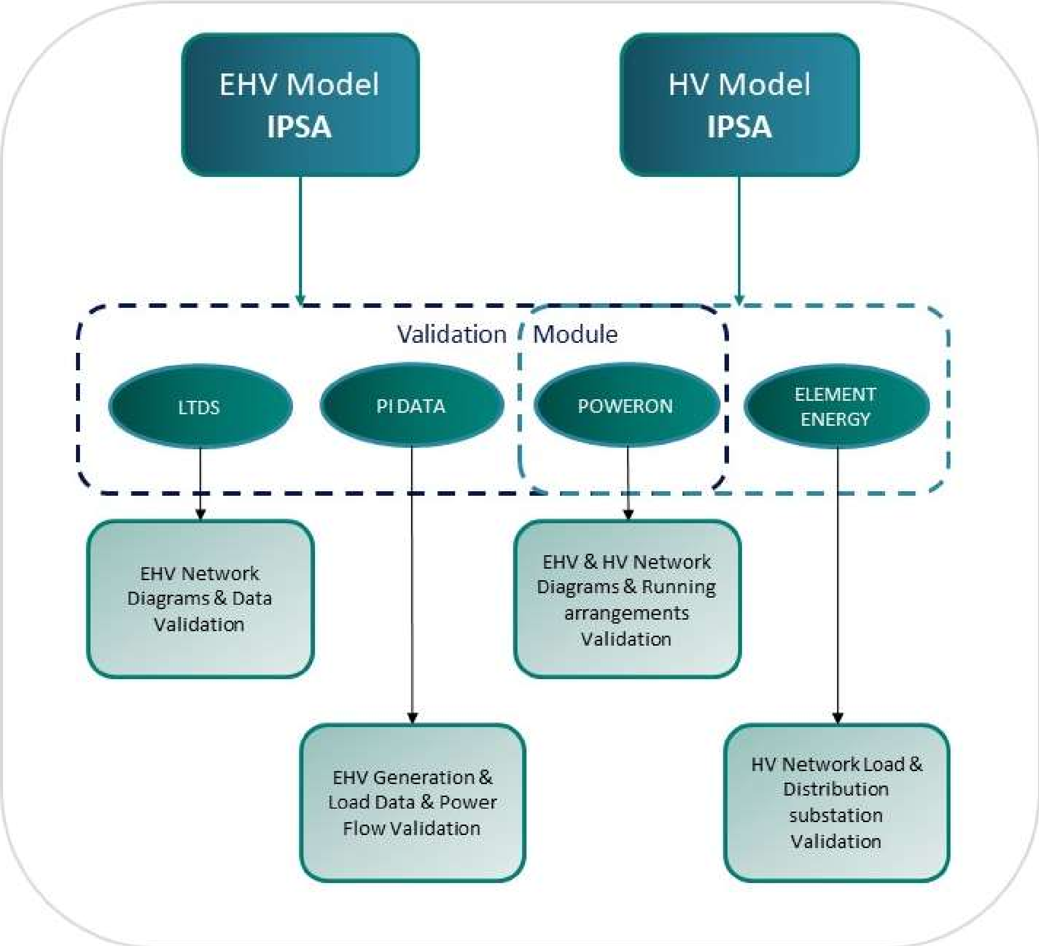
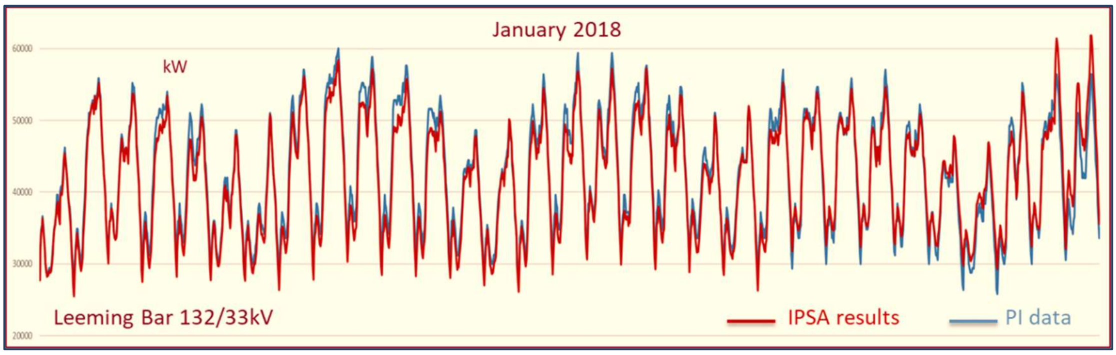
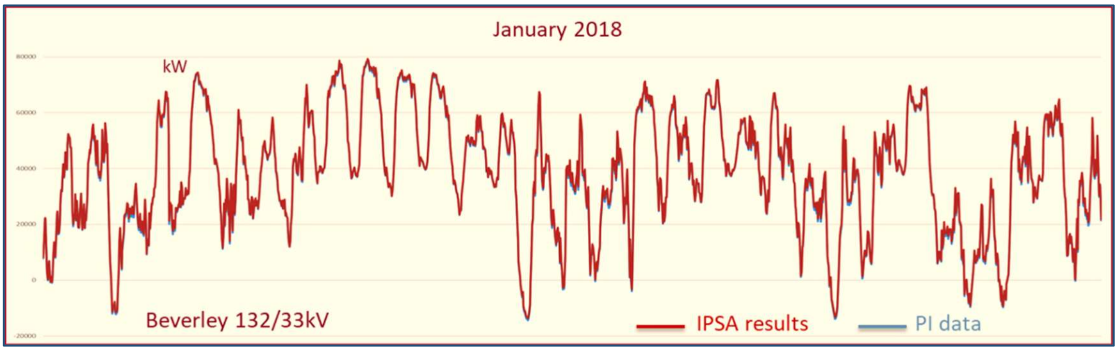


Figure 1 – Network Model Validation

- Validation using LTDS - The EHV network was validated against the network and asset data in the LTDS. Some parts of the Yorkshire network were updated using the LTDS data and then later validated against the DINIS data.
- Validation using PowerOn NMS - The EHV & HV Network Models were compared and validated against the PowerOn NMS manually to ensure that the running arrangements and the network diagrams were accurate and up to date. The normally open points (NOPs) in the HV network were manually updated. PowerOn NMS was also used to identify the PI tags associated with each feeder for the validation of power flow using SCADA data from PI.
- Validation using PI Data - The PI data of primary substation loads was imported into the EHV networks using Python scripts to simulate power flow through the network. The measured values at the 132/33kV transformers from the PI data and the modelled values from the load flow calculations in IPSA are compared below.



PI validation at Leeming Bar 132/33kV in the Northeast

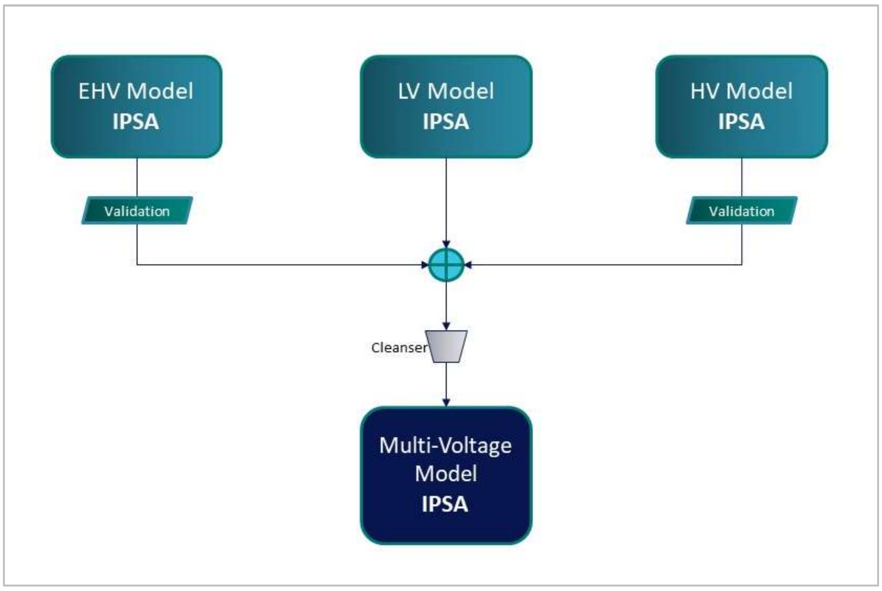


PI validation at Beverley 132/33kV in the Yorkshire

- The differences observed in Leeming Bar between the PI data and the IPSA results was mainly due to instances of PI data containing "Bad Data" or "Calculation failed" fields. Periods of missing or bad data were populated with the most recent data that was deemed to be good, leading to stepped PI curves. Ideally, such cases could have been scaled but owing to the number of such cases this method was not tested. It is to be noted that in Yorkshire where most of the data was available, the validation results are very good.
- Validation using the Element Energy Load Growth Data - The Element Energy Load Growth Data was used to compare the loads and the distribution substations in the HV networks and was used to initialise distribution substation loads to test the model connectivity with a load flow analysis. This exercise identified quite a few discrepancies in the network data obtained from DINIS including NOPs not being imported/converted. This made the network meshed and a load flow was not able to converge. Once the NOPs were correctly defined based on PowerOn data, the load flow converged.

5 Model merge

- The relevant EHV, HV and LV models were manually merged together to create a multi-voltage network model.



- A load flow was run to ensure that the network connectivity is not lost and the results are compared and validated against the individual models.

Issues, Solutions and Lessons Learnt

The process of producing the multi-voltage networks was more complex than initially envisaged and a number of learning points emerged, including:

- The PI data showed negative load at two primary substations where there was insufficient generation present that could produce this profile. The negative load data was converted to positive load data.
- Some primary substation transformers were missing taps after model conversion. This was due to a problem in the formatting of the DINIS export CSV file. In these cases the taps were added manually.
- Zero impedance circuits (circuit breakers/switches) at primary substations were removed from the model, as these could cause problems with load flow convergence.
- Line impedance values less than 1×10^{-4} were set to 0.001, so as not to cause problems with load flow convergence.
- The primary target voltage for all Northeast substations on the DINIS HV model needed to be updated. Ideally, they should have been updated following the target operational voltage changes in 2012.
- Some circuits in the Yorkshire EHV model were observed to be modelled at the base of 1MVA rather than 100MVA. This was identified by load flow warnings in IPSA as the voltage drop was high. The data was verified, and it was identified that these inconsistencies existed in the DINIS model and did not occur during the conversion process. The circuit data was compared against that in the LTDS and corrected.

- The primary transformer maximum step size should be sourced from the EHV IPSA primary substation data which should be more up to date than the DINIS model data.
- PI data can contain “Bad Data” or “Calculation failed” fields. These data items were populated with the last known good value.
- Some reactive power kVAr data was identified as being outside reasonable values; this was partly due to incorrect PI tags. The kVAr data in PI is generally not good enough to carry out the same checks that are used to validate active power flows.
- The Element Energy Load Growth Data identified that some loads were not present in the DINIS network models. These were added to the multi-voltage model and the issue raised internally.
- Secondary transformer impedance tap range and tap position were not always available. It was possible to make assumptions but ultimately the nameplate data needs reading on site and the asset management database updating.
- In the Northeast, the EHV IPSA models use a per-unit base voltage of 11.5kV at primary substations, which matches the secondary winding no-load voltage of the primary transformers. However, in the HV DINIS models the per-unit base voltage is 11kV. Care needs to be taken to ensure that the correct per-unit base voltage is used in the merged model, and that per-unit parameters are converted correctly where different per-unit bases are used. In the merged model an 11kV base voltage is used.
- Distribution transformers have a no-load secondary voltage of 433V. Transformer tapping ranges needed converting when these transformers were modelled in the multi-voltage model with an LV base of 400V.
- DINIS does not use transformer tapchanger relay bandwidth in the load flow calculations so this had to be added manually to the IPSA model. This could be scripted in the future.
- In the Yorkshire DINIS model, the EHV/HV transformer resistance and reactance values are separately applied to primary and secondary winding whereas in the Northeast IPSA EHV/HV transformer resistance and reactance values are only allocated on the primary side i.e. EHV side. The DINIS conversion script addressed this by calculating the total resistance and reactance value.
- The DINIS csv file is not a properly tabled set of data but rather disparate data not arranged in a script friendly format. For example, the node name is found in the 4th column whereas for some it is in the 50th. This is because all the element data are lumped together in one sheet. This leads to issues in the conversion process such as the loss of busbar names and circuit breaker data. It is necessary to check for these individual details manually and to build python scripts for specific purposes. The DINIS csv file had to be formatted manually to extract specific data such as transformer and circuit details. This formatting also had to be done as part of the process to extract the busbar names as these were initially imported as alphanumeric identifiers in the new IPSA file.
- The Northeast EHV models do not include any circuit susceptance value, in contrast to the Yorkshire EHV models. A sensitivity study was carried out by removing circuit susceptance values from the Yorkshire EHV model to ascertain the impact; voltage differences of greater than 1% were

seen on some 33kV circuits. This is material and the inclusion of circuit susceptance in the EHV models would improve system modelling.

5.4 Workstream 4 – Smart Meter Data Analytics

Start: April 2018, Finish: February 2019.

The smart meter data analytics report is available at:

https://www.northernpowergrid.com/asset/0/document/4803.pdf

The workstream focused on three areas:

1. The problems associated with smart meter consumption data aggregation and disaggregation;

2. Determining whether the phase connectivity of customers on LV networks can be determined from smart meter data; and

3. The generation of probability distributions of loads at a given time for customers based on a suitable range of demand characteristics.

When the workstream started in April 2018, there were 20 SMETS2 smart meters deployed in NPg's two licence areas, rising to 339 in September 2018. Due to this low deployment, the fact the smart meters were not clustered on any particular LV networks, and issues with accessing the smart meter data, it was decided to use smart meter data collected during the CLNR project. Accessing SMETS1 smart meters was also ruled out as there was no mechanism for NPg to interrogate these meters. The CLNR smart meter data set, specifically the TC1a data, includes approximately 8,000 half-hourly consumption time series, covering two and a half years and includes two winters, collected between January 2011 and December 2013. After filtering out monitored customers with poor data quality, the number of series / customers fell to around 5,000. As a result of the trial start and end dates, for each customer we essentially had data for two winters and three summers. The analytics works was carried out in the statistical programming tool R Studio.

5.4.2 Aggregation

Distribution Standard Licence Condition 10A (SLC 10A) requires that DNOs cannot access electricity consumption data from a domestic customer relating to a period less than one month. We developed and tested techniques to dis-aggregate aggregated smart meter data to recreate load profiles for individual customers. The aggregated 30-minute interval consumption profiles were multiplied by weighting factors determined for each customer to produce dis-aggregated consumption profiles. Two different metrics were used to calculate the relevant weighting factors to apply to the aggregated consumption data. These two metrics were: 1) the total monthly consumption for each customer and 2) the peak monthly consumption for each customer. Weighting factors were calculated by summing the monthly consumption or peak consumption for all customers whose load profiles had been aggregated. Next the proportion of this total that each customer's consumption represented was calculated by dividing each individual customer's peak or total consumption (dependent on the method being used) by the total for all customers whose data had been aggregated. This was carried out for customer aggregations levels of 2, 5 and 10 customers.

Both disaggregation methods investigated had a high level of error:

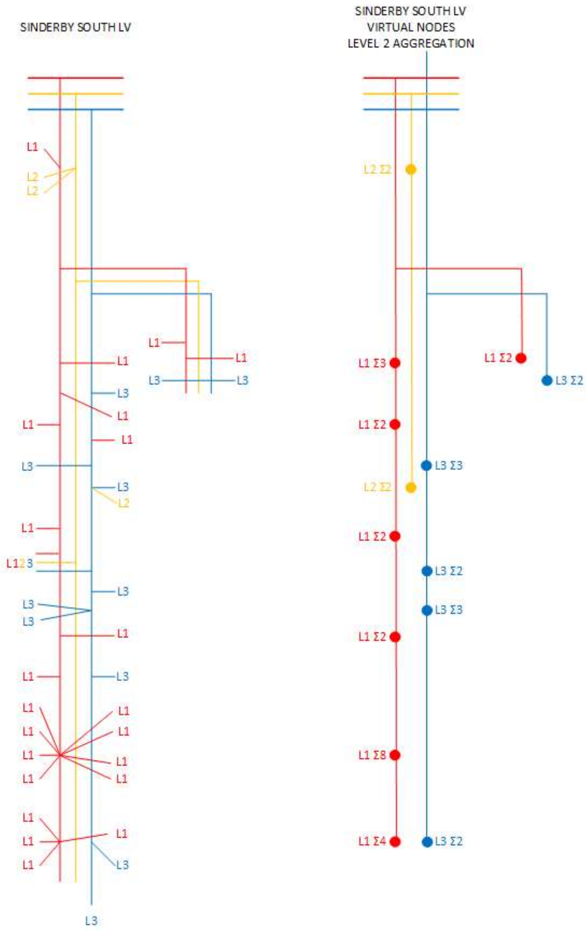
Aggregation Level	Total Consumption Method Average Error	Peak Consumption Method Average Error
2	±33.9%	±39.5%
5	±49.5%	±58.0%
10	±57%	±66%

The aggregation and disaggregation process caused a high degree of error to be introduced. The level of information lost, particularly the shape and height of the consumption peaks raised the question of the usefulness of disaggregated data.

As reported in the Smart Meter Aggregation Assessment Report[1] these results supported the findings that an aggregation level of 2 sufficiently anonymises customer consumption data on days of both high and low consumption; any higher aggregation level would introduce far higher levels of error with no practical benefit. NPg's smart meter data privacy plan is based on an aggregation level of 2.

One aspect we briefly investigated to overcome some of the data aggregation issues was the use of "virtual nodes" in an LV network model. These are points in the network where downstream consumption is aggregated to produce an equivalent load model at that virtual node. This allows aggregated consumption data to be used directly and removes the need to disaggregate it. We looked at methods for creating these virtual nodes automatically for different aggregation levels. The figure below, shows an approach to this, based on an LV feeder from one of our test networks.

[1] Smart Meter Aggregation Assessment Final Report, Smart Meter Aggregation Assessment Final Report – Benefits Reduction, Energy Networks Association (ENA), 2015



The general approach to create the virtual nodes is as follows:

- A trace is started up the mains cable from its end, towards the source substation. A count is kept of the number of properties connected to each phase;
- When the count meets the required aggregation level on that phase, a virtual node is dropped into the model at the service cable / mains cable joint;
- Virtual nodes are created separately on each phase;
- Where there are more customers at the service cable / mains cable joint than the aggregation level, the customers are grouped together at an aggregation level equal to the number of customers; and
- If the source substation is reached and there are not enough customers in the phase count to meet the aggregation level, these remaining customers are grouped with the last created virtual node on that phase.

By visual inspection of the figure above, we see that for an aggregation level of two, a reasonable model of the LV cable can be built. However, for greater aggregation levels (five or ten in the example), the virtual nodes became too far spread out and the ability to model power flows towards the ends of the feeder become more and more limited. There are also only four customers connected to the yellow phase L2, so it is not possible to aggregate these up to five or ten.

The Novel Analysis Techniques workstream ran in parallel with this workstream, and in it a methodology was developed combining network characterisation of circuit sections with exceedance expectation functions of downstream demand to determine the risk of circuit sections becoming overloaded or outside voltage limits. This side stepped the need for virtual nodes, as time series consumption data is aggregated across different customers to develop a 'what happened' view of the consumption of customer connected to circuit sections. The novel analysis methodology focused on producing building probability density functions and exceedance expectation functions of 'what could have happened' or 'what might happen'.

5.4.3 Phase connectivity

The literature review identified several methods that could be used to determine phase connectivity based on smart meter measurements. These can be divided into the following classes:

- Voltage correlation. This method relies on the fact that neighbouring properties on the same phase will have similar voltage profiles.
- Power flow summation. These generally rely on a full rollout of smart meters on a network, and require power flow measurements on LV feeder ways at secondary distribution substations. Rollout of smart meters in Great Britain is progressing at a slower rate than previously envisaged and is driven by customer demand and hence is effectively random; there are very few, if any areas of the NPg network with clusters of accessible smart meters. The majority of secondary distribution substations do not have measurement equipment installed on feeder ways. The methods may also be compromised in Great Britain due to the requirements for aggregation of consumption data under Standard Licence Condition 10A.
- Other. Several academic papers looked at using signals that are not available from SMETS2 meters such as harmonic measurements or phase angles.

Therefore, we explored the application of voltage correlation algorithms. As we had no voltage measurements from smart meters on our test networks, we adopted an approach whereby dummy half-hour consumption data (CLNR), was applied to our test networks, and unbalanced load flows run in IPSA, to produce synthesised voltage data.

Two test networks were used. The first, Sinderby, is a rural network comprising 58 customers on two feeders. The second, Cranwood, is an urban network with 675 customers on six feeders. Sinderby is in NPg's Northeast licence area, where phase connectivity information has not been recorded as standard. Field measurements were taken at each property using a Hasys phase identification unit, to identify the phase connection of each property. Cranwood is in NPg's Yorkshire licence area where the majority of properties have their phase connectivity recorded in NPg's asset database.

We developed K-means clustering algorithms in the R programming language to attempt to identify the phase connectivity of customers using simulated voltages at each customer property, and compared these against the actual phase connectivity records to determine how successful the algorithm was.

K-means clustering is a type of unsupervised learning, which aims to group data based on the similarity of the specified variables, the number of groups the data is segmented into is set by the value K. The algorithm works iteratively to assign each individual row of data to one of K groups based on the values of each variable, clustering those with the most similar values together.

These principles have been applied to phase identification of single phase customers. Utilising the hypothesis that customers on the same phase will have similar voltage characteristics, the clustering algorithm groups customers based upon these characteristics of their voltage profiles. For the method to be considered to be successful, customers connected to the same phase would be grouped together.. The value of K was set to three as we want to segment the customers into three groups, one for each phase. Six different sets of statistical markers of the voltage data in various combinations were used as inputs to the K-means algorithms:

1. Mean and Median
2. Mean, Mode and Median
3. Mean, Mode, Median and Standard Deviation
4. Mean, Mode, Median, Standard Deviation, Maximum, Minimum, and Variance
5. First 50 values in the voltage data set of each customer
6. First 50 differences in the voltage data set for each customer

Although these K-means methods were successful in the academic papers we reviewed, our results did not produce satisfactory results. One of the reasons for this could be that the input consumption data from CLNR was half-hourly. Previous research used much smaller time granularity, and one paper showed accuracy became significantly worse when the time steps were increased from 10 seconds to 5 minutes. The default granularity for recording average voltage measurements on GB SMETS2 meters is 30 minutes, although this can be set to a lower value. Another reason could be that in the IPSA model the voltages across the three phases on the slack bus (located at the 11kV busbar on the 11kV/LV transformer) are set to the same value. Any natural variation in the voltage across the phases would help in the clustering process, but this feature wasn't incorporated in the modelling.

We originally intended to examine the effects of smart meter voltage measurement accuracy on the phase identification methods. This was not pursued as the methods were unsuccessful with no error assigned.

During the course of the work, we also became aware of a problem with the SMETS2 specification in that the time period over which the time step is defined does not have to be clock synchronised, unlike energy consumption. This means that one smart meter could be recording an average between say 11:06 and 11:36, whilst a neighbouring smart meter could be recording between 11:23 and 11:53. This problem has been observed on the small amount of SMETS2 meters deployed on NPg's network. However, this observation would not have affected the findings of this assessment as the voltage profiles used were calculated.

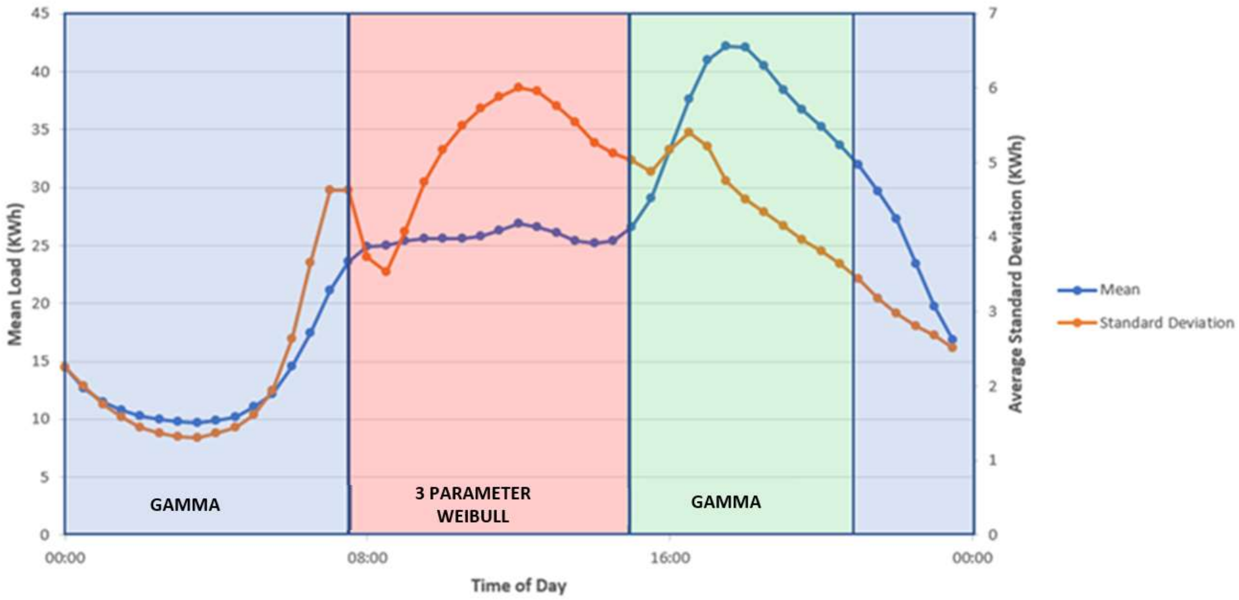
In the context of phase identification using SMETS2 data in Great Britain, we recommend future work in this area considering:

- The time of the year / day when demand is highest and the voltage drop on the network is higher, so that voltage variations along a feeder are more pronounced, hence it should be easier to detect any patterns.
- Analysing LV feeders where large number of SMETS2 meters are deployed.
- The voltage variations on individual feeders rather than a combination of feeders supplied from the same substation. Where the LV feeder comprises branches then each branch could be analysed individually.
- The voltage measurement accuracy of smart meters and the effect of this on the efficacy of the phase identification methods assessed.
- A way to resolve the clock synchronisation problems for voltage measurements on SMETS2 meters.

5.4.4 Probability distributions of customer load

The aim here was to produce probability density functions (PDFs) and cumulative density functions (CDFs) that allow the estimation of the most likely values of the shape parameters used to define the parametric distributions that model the consumption characteristics of groups of customers for a given time of day and season. The outputs of this work would feed directly into the Novel Analysis Techniques workstream.

Firstly, the CLNR data set was segmented by season (Winter, Spring, Summer and Autumn). The next step was to understand the mean consumption and variance per half hour for each of the seasons for groups of differing numbers of customers (5, 10, 50, and 100). To generate these mean and variance values, plots of the mean aggregated consumption for each half-hour segment of the day, 1000 scenarios were run for each group size and season. An example plot for the mean half hourly consumption for 100 customers in winter is shown below:



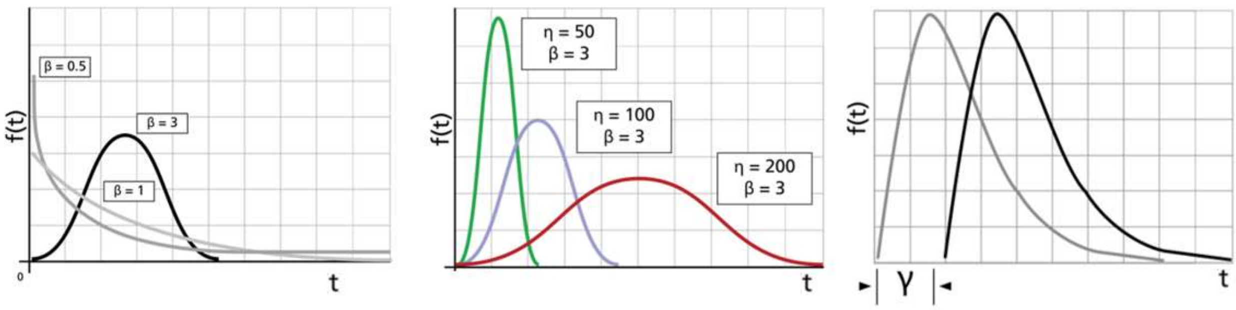
From the plot it can be seen that average consumption throughout the day is variable, therefore matching a parametric distribution with a constant mean to the entire day would not accurately model consumption. To account for this, two steps were taken: 1) the day was segmented into periods with similar mean consumption and 2) distributions were selected that can be scaled either by addition or multiplication to adjust the mean to fit the distribution for the whole period and then scaled back to produce half-hourly distributions.

Two distributions that can be scaled are the gamma distribution and the three parameter Weibull distribution. The gamma distribution can be scaled by multiplication; therefore, it is the best distribution to model segments with a changeable mean and changeable standard deviation. The three parameter Weibull distribution is scaled by addition; therefore, it is the best distribution to model segments with a constant mean and changeable standard deviation. These distributions are far more suitable than the normal distributions used in ACE49 as they can represent the skewness inherent in the PDF shapes.

The PDF of the three parameter Weibull distribution is described by the equation:

$$f(t) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}$$

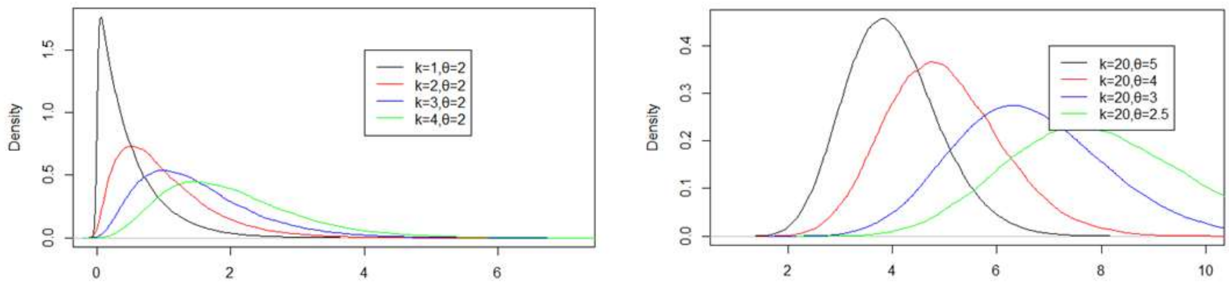
Where β is the shape parameter, η is the scale parameter and γ is the threshold parameter. The effect of changing these parameters can be seen below:



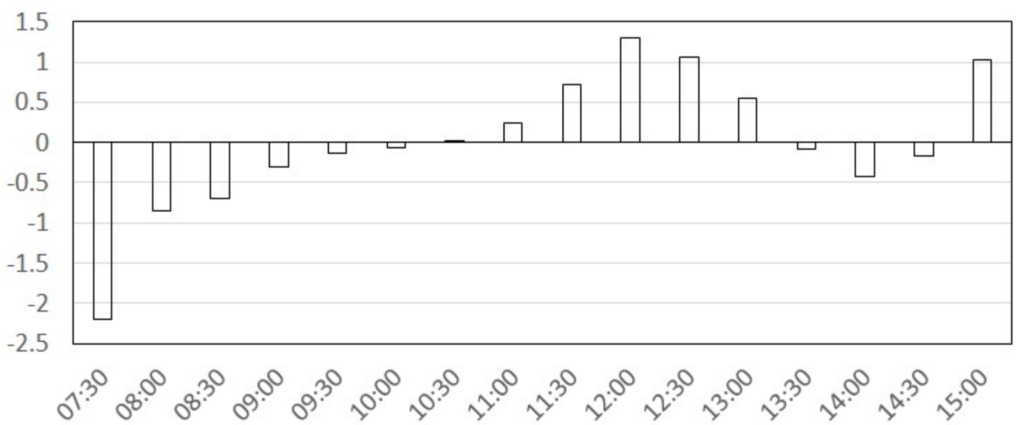
The probability density function of the Gamma distribution is described by the equation:

$$f(t) = \frac{1}{\Gamma(k)\theta^k} t^{k-1} e^{-\frac{t}{\theta}}$$

Where k is the shape parameter and θ is the scale parameter. A scaling factor is calculated to normalise the mean for a particular half-hour in the day and this is multiplied by the shape parameter. The effect of changing these parameters can be seen below.

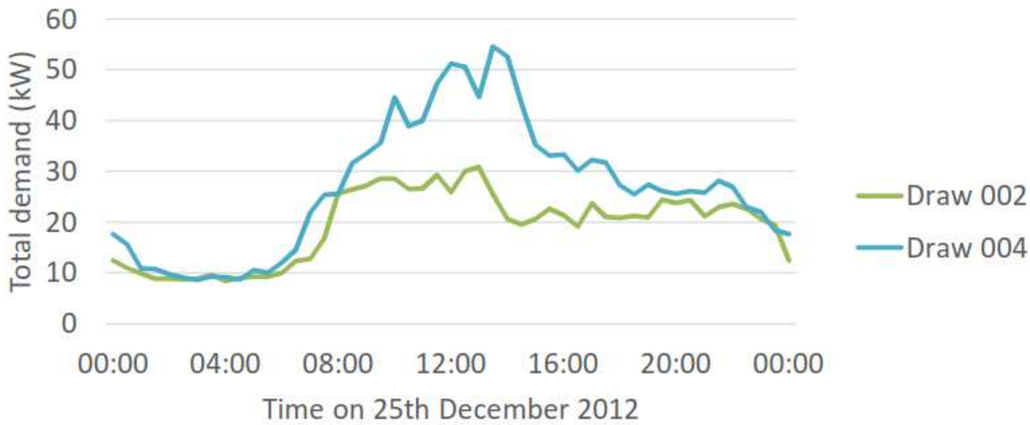


We have now created a method to determine for any grouping of customers, within one of the three segmented time periods in the day, and for a particular season, a PDF. Next, we generate scaling factors that can be used scale the PDF for the particular half-hour, for example:

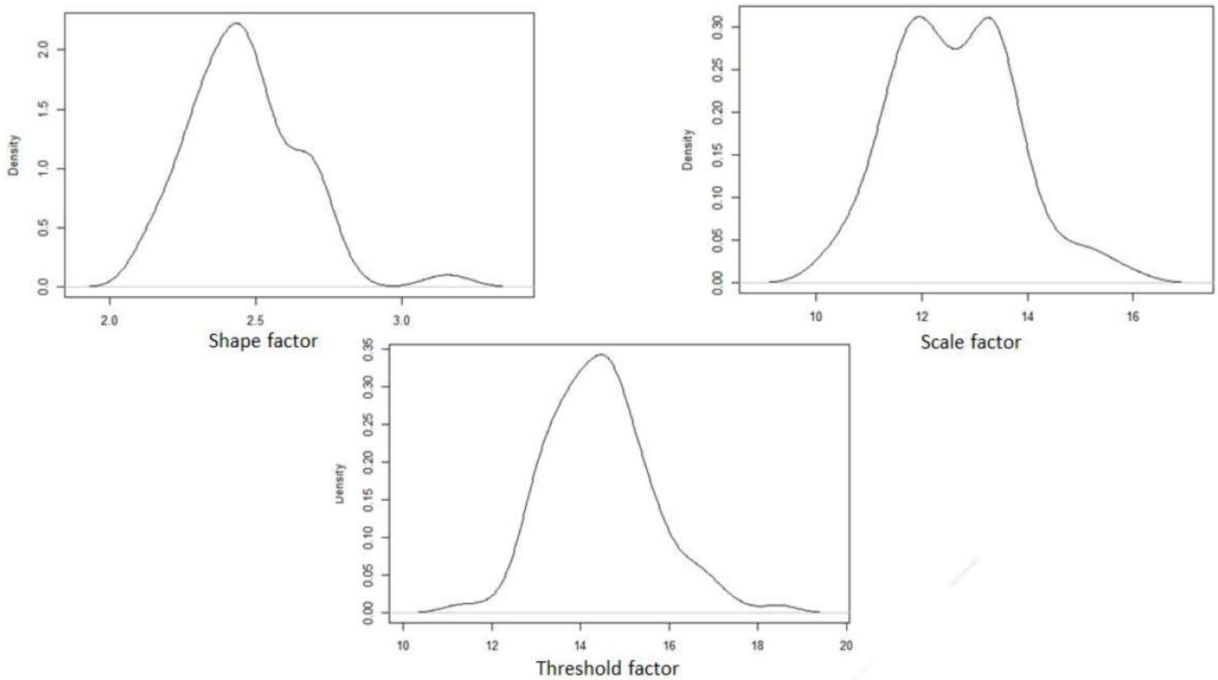


For the Weibull distribution, the scaling factor is calculated to normalise the mean for the particular half-hour in the day and this is added to the threshold parameter. For the Gamma distribution, it is multiplied to the shape parameter.

So that we had a suitably large set of observations from the CLNR data set, the sampling algorithms were run 1000s of times across randomly chosen groups of 5, 10, 50 and 100 customers. Each “draw” resulted in a different consumption profile for the selected day:



Each draw had its own PDF and set of parameters, so there was uncertainty with the PDFs themselves. This is represented below by the probability density plots for the parameters in a three parameter Weibull distribution for a particular season and segment of the day. We did not go as far as fitting PDFs to these distributions themselves but in a full Bayesian statistical model this could be done, with the parameters of these PDFs being known as ‘hyper-parameters’.



With this information we can understand the uncertainty associated with our demand exceedance predictions, and formalise statements such as: “we know with 90% confidence that there is a chance that a particular demand level will be exceeded once in every ten years”.

5.5 Workstream 5 – Novel Analysis Techniques

Start: June 2018, Finish: December 2019.

5.5.1 LV network modelling – existing approaches and limitations

With the exception of designing new connections, there has historically been little need to monitor or model LV networks, which means that existing models and the data they require are reasonably simple. Like other DNOs in Great Britain. NPg use one of two existing methods:

- The **ACE 49** method, developed in the 1970s and 1980s, which describes a simple statistical model for understanding the demand for groups of customers types that could be supplied from an LV network. There are several strong assumptions made in order to derive this model, including the definition of a 90th percentile level of risk. In addition, the detail and intent of some important aspects of the method aren't completely clear.
- The **After Diversity Maximum Demand (ADMD) method**, which generally refers to empirically determined values of the per customer demand, for a group of N customers. By definition, ADMDs don't relate to a defined level of network risk.

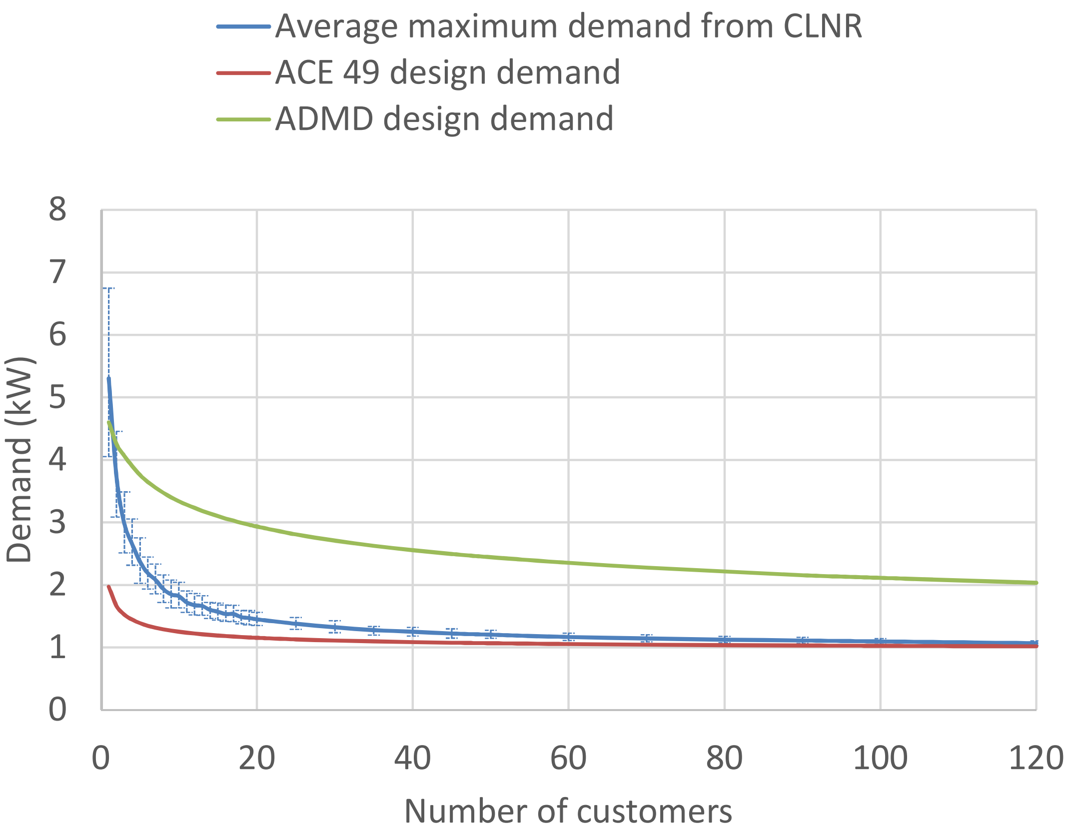
Both methods have shortcomings. One weakness of both methods is that, in their typical use, they assume that all customers of a specific type exhibit the "average" behaviour for that type of customer. Trial data shows that this is not the case when dealing with small groups of customers. Another weakness is that power flows on the network are often not studied, or are studied in a simplified way.

We conducted an experiment whereby we simulated the aggregated annual half hourly demand profiles of 200 trial groups each with N domestic customers by sampling from the CLNR TC1a data set (general domestic customers with no electric heating). We did this for groups of varying sizes ranging from N=1 to N=120 customers. The calculation procedure was as follows:

- 200 trial groups each containing 120 customer annual half-hourly demand profiles were generated by randomly selecting customer profiles from the TC1a data set;
- For N=1, one customer profile was randomly chosen from each of the 200 groups. For each customer profile their half-hourly demand data was scanned to determine the maximum half-hourly demand for that customer. From the 200 maximum half-hourly demand values the average was calculated;
- For N=2, two customer profiles were randomly chosen from each of the 200 groups. For each of the two customer profiles the half-hourly demand data was scanned to determine the half-hourly period which resulted in the maximum combined total demand for that customer. This combined total demand was divided by two, to give an average maximum demand for those two customers within that group. This was repeated for each of the 200 groups to give 200 maximum demand values from which the average was calculated;
- The process was continued until N=120.

The results are shown by the blue line in the figure below, with the 10th and 90th percentile shown by the vertical bars.

Alongside this, we calculated the design demand which would be provided for the same values of N by the ACE49 and ADMD methods as described in accordance with the NPg LV Design Code of Practice. The results are shown by the red and green lines in the figure below.



The following observations are made based on these trials:

- The ACE49 method consistently underestimates the demand for smaller values of N, whereas the ADMD method consistently overestimates it for all values of N.
- The ACE49 method's underestimation is more pronounced for smaller customer group sizes, N. For larger values of N, the trial sample and the ACE49 design demand converge at approximately 1 kW per customer. The ADMD method, on the other, converges to a figure of approximately 2.1 kW per customer.
- There is significant variation in the peak demand produced by different groups of customers, especially for small values of N. Neither the ACE49 method nor the ADMD method account for this variation.

There is the potential for rapid increase of demand on LV networks, particularly due to customers adopting Low Carbon Technologies (LCTs) in support of Government objectives for decarbonisation of heat and transport. In the future, both the ACE49 and the ADMD methods may cease to be suitable for designing efficient and secure LV networks.

The roll-out of smart metering and LV monitoring could provide DNOs with a rich source of data to improve their LV design practices. Our motivation for the workstream was to develop enhanced methodologies which can incorporate this new information to continue to design economically efficient networks.

5.5.2 Novel analysis techniques at LV

The novel analysis techniques at low voltage report is available at:

https://www.northernpowergrid.com/asset/0/document/4918.pdf

The novel analysis techniques we developed are rooted in the use of a statistical model to reflect both the *variability* and *uncertainty* in demand on LV networks. More specifically, we proposed the use of a *Bayesian statistical model* for representing demand:

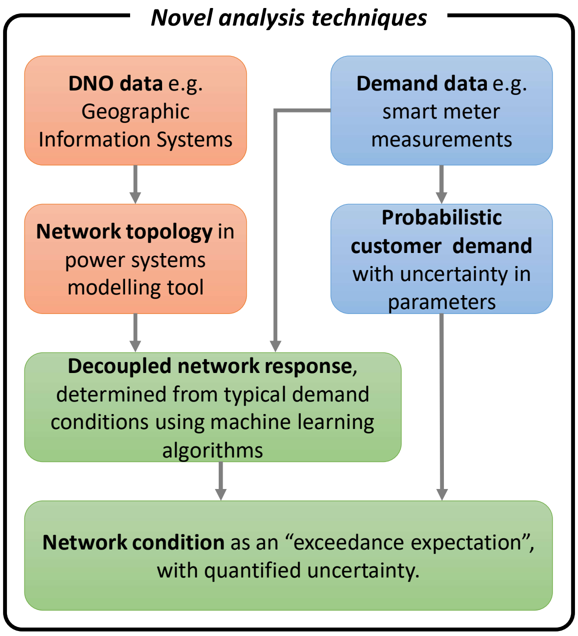
- In Bayesian statistics, probabilities are viewed as representing subjective beliefs, rather than the long-run frequency of some measured phenomenon. This is important when dealing with problems for which there is not much data. Initial beliefs are formalised mathematically as *'prior probability distributions'*.
- Our method proposes that prior probability distributions should be formed based on existing data sets from projects such as CLNR. Bayesian statistics allows for the initial prior beliefs to reflect the uncertainty which exists when trying to understand the demand of customers supplied at LV level, without specific local data.
- When data becomes available from smart meters or LV monitoring or both, it can be used to update the 'prior', according to a procedure known as Bayesian updating, to form a *'posterior probability distribution'*. It is expected that this will reduce the uncertainty in the estimate.
- This procedure can be repeated indefinitely every time new data becomes available. Eventually, the initial prior belief will have very little influence on the demand estimate.

Our method also captures the impacts which these demands will have on the network, in terms of thermal utilisation and voltage excursion, to inform network planning and new connection designs, based on detailed AC power flow modelling. However, rather than attempting to run thousands of AC power flow simulations using Monte-Carlo sampling every time a study is run, we proposed a method which decouples the AC power flow modelling from the modelling of demand.

This is achieved by running large sets of aggregated customer demands through a network model, storing the outputs, and then regressing the outputs of the AC load flow against the demand inputs. This requires considerably fewer samples than a conventional Monte-Carlo AC load flow. This takes advantage of the fact that, even though AC load-flow is non-linear, by observing the results for thousands of combinations of demand, it should always be possible to estimate the load flow results for all credible demand levels.

The figure below provides an overview of the methodology developed. The production of detailed LV power system models "network topology" is covered in the previously described Workstream 2 (LV Network Model Methodology).

Modelling	Challenge	Novel analysis technique
1 Network topology	Network operators don't typically have detailed power systems models of their LV networks	**Automated LV model build** of a power systems model based on GIS data
2 Customer demand	Customer demand is subject to lots of uncertainty, even for a known mix of customer types.	A **Bayesian customer demand model**, to account for uncertainty in customer demand.
3 Network condition	AC load flow is computational expensive, making "Monte-Carlo" load flow unattractive.	**Decoupled network response** - analysis of power flow and voltage for varied demand conditions.



Bayesian customer demand model

The Bayesian approach to statistics describes nothing other than people's state of knowledge about the phenomenon or quantity of interest. The main consequence of this is that everything which is uncertain is treated as a random variable. This means that random quantities such as customer demand are modelled as having probability distributions belonging to established parametric 'families', and characterised by a set of parameters. However, the distribution parameters are themselves modelled as random variables, characterised by their own set of hyper-parameters. A process to create these probability distribution functions (PDFs) and associated hyper-parameters was covered in Workstream 4 (Smart Meter Data Analytics) as previously described.

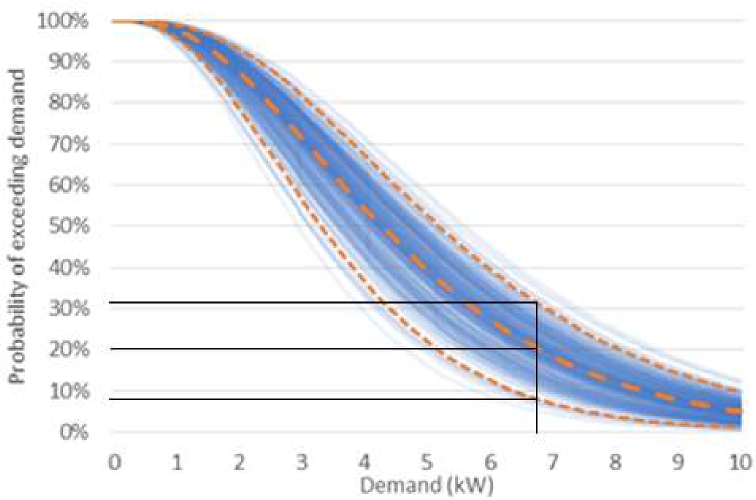
As this process defines variables together with their inherent uncertainties, it is natural to explicitly quantify the uncertainty within a model – such as the uncertainty associated with modelling the demand of a specific group of customers rather than a typical group, or the uncertainties associated with the demand from electric vehicles or heat pumps.

Bayesian statistics is unique in treating probability distributions as subjective, so that the statistical modeller's prior beliefs constitute a valid and necessary part of their model, even when no specific data is available. As a result, the approach is well suited to problems where there is limited data, or data from multiple sources.

The statistical modeller's prior beliefs are translated into mathematical form through the construction of prior probability distributions. For example, if the modeller believes that the quantity of interest is likely to lie within a certain range, but knows nothing else, then a suitable prior distribution would be to assume a constant probability density within that range, and zero density outside of the range.

In our methodology, we started with the CLNR dataset as the basis for our initial prior belief of typical customer consumption. As smart meter data along with other sources such as annual consumption data and LV network monitoring are received, Baysian updating is used to produce a posterior probability distribution. As further data is received in the future this posterior become the new prior and is again updated to produce a new posterior probability distribution. The process is continuous and indefinite and allows us to incorporate new sources of data when they become available.

The figure below demonstrates the inherent uncertainties with Baysian distributions. There are multiple exceedance expectation functions (represented by the faint and deep blue lines) representing plausible demand patterns on a particular network. The central orange line shows the average of these, whilst the outer orange lines include 95% of plausible customer behaviours. In the illustrative example of the figure, the black lines help to demonstrate that the expected demand that is exceeded 20% of the time is 6.8kW, and 95% of samples have the probability of occurrence for this demand within the range 8% to 31%.

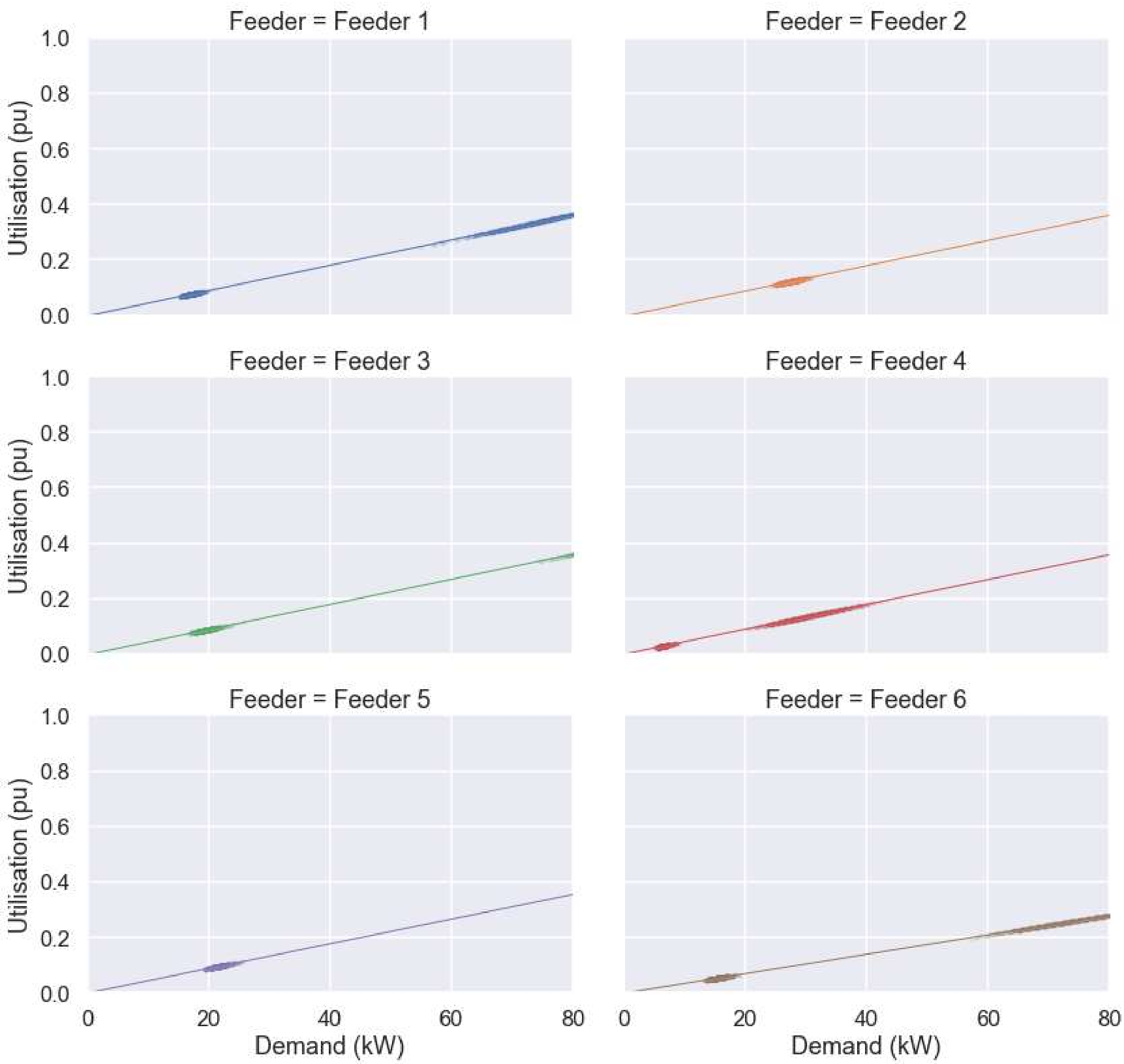


In the long-run, as new data is incorporated within the model, the uncertainty around the parameters in the figure will decrease, with the range of possible values "tightening" around the average. However, the process will not always be one-way, i.e. sudden changes in demand patterns following the uptake of new technologies, may temporarily increase uncertainty.

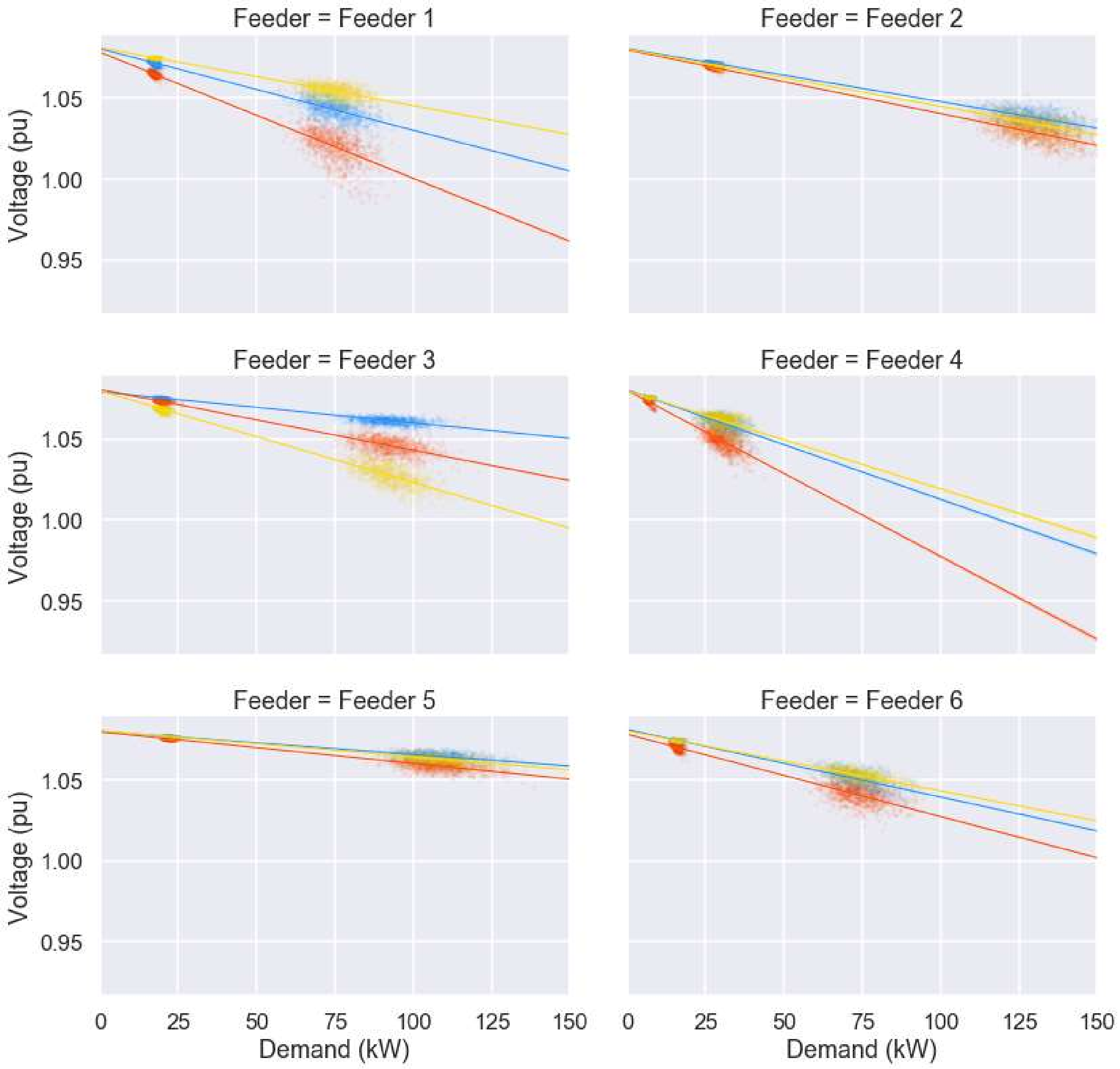
Decoupled network response

We have proposed a method which decouples the AC power flow modelling from the modelling of demand. This has allowed us to find simple equations that relate thermal utilisation and network voltage to the feeder demand. This method speeds up the time it takes to calculate thermal utilisation and voltage exceedance risks for particular network design use cases.

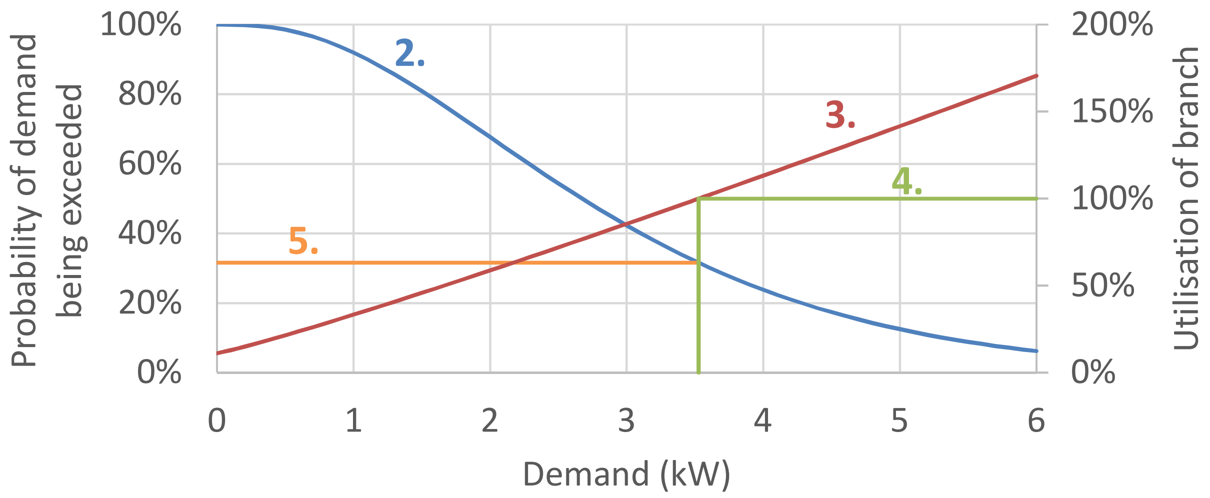
We illustrate how this works using our Cranwood test network which has six feeders. The figure below shows the thermal utilisation for each of the six feeder ways at the substation as the customer demand increases. The individual data points represent the results from one sample of customer demands, with a best fit line drawn through the results.



The figure below shows how the voltage falls on the ends of each of the six feeders as the customer demand increases. There are three lines for each feeder representing each of the three phases.



A stylised step by step example of how to use the decoupled network response modelling combined with the Bayesian customer demand level, to establish the probability that a given utilisation threshold will be breached, is presented below:



Step 1: Produce PDFs of the aggregate customer demand for each half-hour of a day and season, using the methods described previously (in Workstream 4).

Step 2: Use these PDFs to generate a set of "exceedance functions" i.e. the probability that demands exceed the input value. This is one minus the cumulative probability of any given level of demand. (Line 2 in the diagram above).

Step 3: Based on the network response trials completed in the IPSA AC power flow model, an equation can be found which describes the utilisation (or voltage) of a feeder section as a function of the relevant aggregated demand. In this illustrative case, this is represented reasonably well by a linear function, which depends on only one variable (e.g. aggregate downstream demand). (Line 3 in the diagram above).

Step 4: The network planner can determine what level of aggregate customer demand leads to an utilisation of 100% (or other utilisation), from Step 3. In the example we see that an aggregated downstream demand of 3.5kW leads to a utilisation of 100%. (Line 4 in the diagram above).

Step 5: By comparing this to the demand model exceedance function (Line 2), the probability that demand will be high enough such that the feeder section has a utilisation of 100% or higher can be estimated. In this stylised example, there is a 32% chance of the feeder section utilisation exceeding 100% utilisation. (Line 4 in the diagram above). This would be mean that for a given year, there are 5,606 half-hour periods (0.32 x 24h/day x 365 days x 2) during which half-hourly aggregate customer demand is likely to cause the feeder section utilisation to exceed 100%.

5.5.3 A functional specification for a novel LV modelling tool ('Smart LV Design')

The functional specification for a novel LV analysis tool ('Smart LV Design') is available at:

https://www.northernpowergrid.com/asset/0/document/5218.pdf

The functional specification describes a minimum viable product (MVP) that can be used to implement some of the concepts developed during the project. The tool would be integrated with Northern Powergrid design modelling tools software and IT systems, and would interact with Siemen's secure Energy IP environment, in which smart meter data will be held, and the eAM Spatial system. It would be expected that, if the tool were developed, the MVP would be tested with a limited set of users and their experiences fed into the development of the second version of the tool. The MVP does not rely on large quantities of smart meter consumption meter data being available, so development could progress immediately.

The goal of the tool is to offer an alternative to the DEBUT and the ADMD spreadsheet tools. The replacement of legacy tools would be justified on account of the new tool's ability to flexibly leverage new data sources, to produce more accurate predictions of thermal and voltage violations on LV networks.

The tool's architecture is conceived in a modular way such that the core functionality is separated from DNO specific interfaces such as to smart meter systems or GIS databases. Within the architecture is a Common Information Model (CIM) compliant database holding the electrical network models. This database could be used with other DNO tools if required.

The user interface to the tool would be a web browser with the LV networks represented as LV skeletons on top of Ordnance Survey mapping data.

The tool would incorporate the following features:

- Sophisticated data-driven statistical modelling;
- Risk based analysis – identifying demands that exceed asset capacities and statutory voltage limits, their impact and their frequency;
- Extendible to capture the statistical properties of multiple, correlated demands - for situations where the state of a network component cannot be determined with sufficient accuracy by a single aggregated demand. This would ensure that the high and low extremes observed over multiple years are accurately modelled;
- Ability to carry out analysis in a true power system modelling environment, used economically to manage computational resource requirements; and
- Dynamic and flexible, will update to incorporate new data, such as increasing availability of smart meter data, and learning about the usage patterns of new technologies and their uptake.

5.5.4 Multi-voltage level novel analysis

The multi-voltage level novel analysis report is available at:

https://www.northernpowergrid.com/asset/0/document/5323.pdf

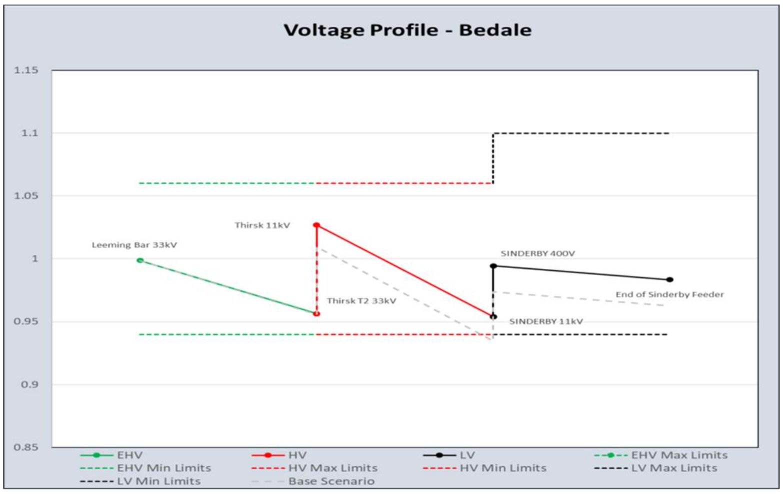
The following two multi-voltage level (MVL) models were developed in Workstream 3:

- A Northeast rural network: Norton (GSP) - Leeming Bar (132/33kV) – Thirsk (33/11kV) – Sinderby (11kV/LV) – Sinderby (LV).
- A Yorkshire urban network: Creyke Beck (GSP) – Beverly (132/33kV) – First Avenue (33/11kV) – Cranwood (11kV/LV) – Cranwood (LV).

In the analysis we used these two MVL models to explore holistic voltage behaviour under a range of interdependent network load states and topologies including:

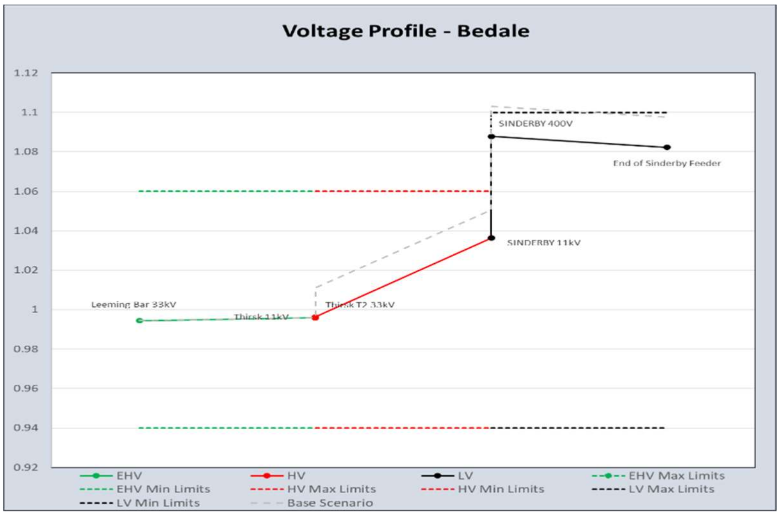
- Voltage variation across voltage levels under credible but challenging First Circuit Outage (FCO) conditions at both EHV and HV at times of i) peak demand and ii) summer minimum demand/peak generation.
- Voltage variation across voltage levels under increasing demand / generation scenarios that result in voltage violations beyond statutory limits.

We then explored the effectiveness and application of voltage management solutions to address the voltage violations modelled. As an example, the following case study presents a winter peak demand, zero HV generation scenario on the Northeast rural network. The figure below shows how we can represent the voltage profile along with the statutory voltage limits across the multi-voltage levels and provides a useful visual aid to a network designer. The grey dashed line shows that the voltage when the network supplies the winter peak demand drops below the HV statutory limits at Sinderby 11kV substation. The red and black lines show that application of load drop compensation (LDC). After application of LDC the 11kV target voltage increases at Thirsk Primary substation from 1.01pu to 1.03pu, and this brings the 11kV voltage at Sinderby 11kV substation to within statutory limits.



We then look at the summer minimum demand, maximum generation scenario:

The grey dashed line shows that the voltage when the network supplies the summer minimum demand and the HV connected generation is at a maximum, the LV voltage on the feeder way at Sinderby 400V busbar goes above statutory limits. The red and black lines show that application of load drop compensation (LDC). The application of LDC brings the LV busbar voltage at the secondary substation to within limits by lowering the 11kV target voltage at Thirsk Primary substation from 1.01pu to 1.00pu. The MVL model is thus allowing us to see the effect of applying LDC at HV on the LV system as well as on the HV system.



Based on analysis of the representative multi-voltage networks under a range of scenarios and voltage management solutions, we concluded the following learning outcomes and recommendations:

Outcomes

- The rural Thirsk network has comparatively less voltage headroom and legroom as feeders are much longer, consistent with rural network topology. The relatively urban First Avenue network is more resilient to the impact of connecting new demand and generation in terms of voltage; thermal constraints are likely to occur first. With heat and transport electrification, rural networks are likely to be more affected by voltage issues.
- A number of existing and novel solutions can be implemented to unlock network voltage capacity on both the First Avenue and Thirsk networks under increasing demand and generation scenarios.
- The present NPg voltage policy is be suitable for existing and future loading conditions under credible operational regimes. The allocation of voltage drop across the HV and LV voltage levels is broadly appropriate for both urban and rural networks. However, our analysis has illustrated that in some cases the voltage at the 400V busbar is less than 1pu and is not aligned with the HV and LV design principles, yet voltages at the end of the LV feeder are within limits. Our findings support the recent NPg policy to purchase new transformers with taps settings of -2.5% (voltage boost) to 7.5% (voltage buck) as it enables an improved balance of voltage headroom and legroom.
- Changes to the target voltage at EHV substations makes very little difference to downstream voltage as the primary substations are set to manage the voltage to a target set point, with sufficient tap range to generally accommodate a wide range of loading conditions. For the representative networks analysed, the primary transformer tap changers are lying towards the higher end of the tapping range, potentially reducing the amount of generation headroom.
- Assessment of the two representative networks indicates that transformers installed in both rural and urban networks are able to meet the requirements set out in Engineering Recommendation P10.
- A voltage reduction solution in conjunction with a LDC appears to be the most suitable to implement at primary substation level if high voltage constraints occur. LDC can help manage voltages in various future demand and generation conditions at HV and LV although this will vary from network to network and the loading patterns of feeders connected to those networks would need to be fairly similar, including under various contingency conditions, in order for LDC to be

applied successfully. Inline voltage regulators can be implemented to address feeder specific voltage issues.

- Alternatively, target voltage settings on new AVC relays fitted with remote communications could be changed remotely on a seasonal basis, for example, by Network Operations, resulting in a similar voltage change to LDC.
- Solutions may also enable new outage topologies to be considered. For example, in the Thirsk network, backfeeding the Thirsk circuits from Bedale can potentially be considered with the deployment of an in-line voltage regulator. This would provide the operational engineers with more options when there is an outage as well as enabling the connection of some further demand or generation.
- For localised LV voltage issues, manually changing the HV/LV substation tap settings may be a suitable short-term solution for demand increase. However, settings would then need to be readjusted seasonally if there was a moderate to high uptake of PV; in this scenario the solution would not be a feasible long-term solution. In addition to the resources required to change the tap position, as the transformers would need to be switched off, there would be an increase in Customer Interruptions and Customer Minutes Lost. A wider scale rollout of distribution transformers equipped with OLTCs may provide a more flexible future proof solution however, this has implications for capital expenditure.
- The MVL modelling approach enables cross-network voltage issues to be better studied and solutions tested to resolve issues with specific networks as well as provide strategic guidance on voltage policy. For example, the connection of generation at HV combined with high uptake of PV at LV. The implications of a voltage management solution applied at primary substation, to connected generation or an in-line voltage regulator along a feeder can also be assessed across a range of voltage levels. By bounding the model to a thin slice and relevant voltage levels, it is possible to build quickly and efficiently. Most value is delivered for HV and LV multi-voltage level models, the addition of EHV models does not provide significant additional value.

Recommendations

Our recommendations on network modelling and application of voltage control solutions and management techniques, both at design and strategic level, are provided below.

- This innovative and efficient MVL modelling approach developed here should be applied to NPg networks where voltage is specifically an issue across HV and LV networks. This could occur due to connection of material generation and/or demand both at HV and on LV feeders supplied from the HV feeder for example. It should also be used to advise policy on strategic voltage solutions, for example on the use of target voltage changes, on their own or in conjunction with LDC across a wider range of network types beyond the two networks studied here.
- The modelling of NPg networks could be improved to better capture voltage behaviour e.g. by including susceptance values for network models.
- Rural networks are more likely to suffer from voltage issues and should be considered for a more detailed assessment where there is significant embedded generation or high electric vehicle/heat pump uptake. There is potentially a case for treating rural and urban networks differently in voltage policy however, this can lead to lack of clarity for networks that are difficult to classify. This should be explored further.
- The present NPg voltage policy should be retained for now but reviewed as loading conditions change materially. Transformer tap ranges and specifications, for example, appear to be broadly suitable for urban and rural network types based on the two representative networks studied, with

transformers generally not reaching the end of their tapping range. However, the two networks studied are not likely to be fully representative of the wide and varied range of network characteristics across NPg so we would recommend further verification of these recommendations on other networks. This can be efficiently carried out based on the methodology developed.

- The tap change step and range should be sufficient to enable future connection of demand and generation, as low carbon technology uptake increases and distributed energy resources are increasingly deployed. For the representative networks analysed, the primary transformer tap changers are lying towards the higher end of the tapping range, potentially reducing the amount of generation headroom.
- In some cases, it may be found that NPg HV and LV design principles are overly limiting and compliance with voltage statutory limits can still be achieved although the design principles are not. These cases should be tracked to support assessment of wider applicability and thresholds to ensure that the balance of HV and LV voltage drop is appropriate. The risk associated with the treatment of AVC deadband to voltage non-compliance should also be considered in more detail.
- Use of revised target voltage, with LDC at primary substations to manage the impact of new demand and generation on voltage can be quite flexible. However, the application depends on the loading patterns of HV feeders being relatively similar including under various contingency conditions. Alternatively, target voltages could potentially be broadly set for rural and urban networks and adjusted seasonally. Strategic application of LDC across a range of network types should be explored further, which could be achieved using this modelling framework. An improved understanding of when to deploy target voltage changes and/or LDC, how these might interact including during contingency conditions, and how this could best be reflected in both the planning and design process and voltage policy is required.
- Change of generator voltage / reactive power characteristics can provide some material voltage support when the generator is in service. This solution should be studied in more detail under a wider range of conditions to understand the behaviour of generation operating in different modes e.g. PV, PQ, QV, (as defined in Engineering Recommendation G99) and how the benefits could be realised in practice.
- Where there is a requirement to replace a HV/LV transformer due to asset condition or load, the deployment of a transformer equipped with an OLTC should be considered to both provide future proof voltage support.
- There is expected to be an increase in the volume of devices that are connected to the network as the penetration of LCTs increases. Some of these devices, e.g. a primary transformer AVC and a secondary transformer AVC, are DNO assets and others will be customer assets e.g. V2G and domestic batteries which may be able to provide reactive power and voltage control services. Further work should be done to understand the potential interactions, implications and complexities.

6 Performance compared to the original project scope, objectives and success criteria

6.1 Project performance against scope

The project scope, as specified in the NIA PEA (Project Eligibility Assessment) document is divided across the five work streams:

Workstream 1 – Horizon Scanning

Required Outputs:

1) Summary of relevant learning from other innovation projects.; and
2) Identified Use Cases and Test Networks.

The required outputs have been delivered.

Literature Review: https://www.northernpowergrid.com/asset/0/document/4772.pdf

Use Cases: https://www.northernpowergrid.com/asset/0/document/4773.pdf

Data Review: An internal NPg report was produced. A summary is given in section 5.1.3.

Test Networks: Two multi-voltage level networks were identified as described in section 5.1.4.

An additional output produced was a review of existing modelling tools as described in section 5.1.5.

Workstream 2 – LV Network Model Methodology

Required Outputs:

1) Methodology for efficient LV network model build and analysis;
2) Two LV test networks; and
3) Methodology for dealing with data quality and phase connectivity issues.

The objectives of this workstream have been achieved. A detailed internal NPg report has been produced which describes an automated process for building LV network models in IPSA from NPg's eAM Spatial system; this is described briefly in section 5.2. Two LV networks were built in IPSA using this process, a rural network (Sinderby), and an urban network (Crandyke). Whilst building these networks a number of data quality issues were found in eAM Spatial. Temporary workarounds were used to address these and long term fixes have been suggested (as summarised in section 5.2). We were hoping that one of the outputs from Workstream 4 would be a method for identifying customer phase connectivity based on smart meter data, and that the output from this could be fed into Workstream 2. As this was unsuccessful we deployed field staff to Sinderby with a Phase Identification Unit manufactured by Hasys[2]. This is an easy and quick to use device that allows the phase a customer is connected to, to be identified, without entering the customer's property. This will become one of our preferred methods for identifying customer phase connectivity where this information has not been previously recorded in our systems.

Workstream 3 – Multi-Voltage Level Model Methodology

Required Outputs:

1) Methodology for multi-voltage level network model build and analysis;
2) Two multi-voltage level test networks; and
3) Recommendations on voltage control solutions.

The objectives of this workstream have been achieved. A detailed internal NPg report has been produced which describes a methodology for building multi-voltage level networks; this is described in section 5.3.

[2] https://www.haysys.co.uk/haysys_PIU_Main_page.aspx

Two multi-voltage level networks have been developed in IPSA, a Northeast rural network, and a Yorkshire urban network.

Recommendations on voltage control solutions are described in the report:

https://www.northernpowergrid.com/asset/0/document/5323.pdf

Workstream 4 – Smart Meter Data Analytics

Required Outputs:

1) Recommendations on use of smart meter data analytics for LV and multi-voltage level network modelling.

This workstream was partially successful. The use cases developed during the Horizon Scanning workstream and the smart meter data analytic problem statements developed as a result of these prioritised data analytics at LV rather than at HV or EHV. The methods we tested for phase identification of LV customers using voltage correlation were unsuccessful, primarily to the lack of available good quality data. We found that aggregating consumption data of two or more customers was enough to sufficiently anonymise it, as required by SLC10A. The most successful part of the data analytics work was a method for generating three parameter Wiebull and two parameter Gamma probability distributions for customer loads from smart meter data; this work was fed into Workstream 5.

A detailed report for this workstream is available at:

https://www.northernpowergrid.com/asset/0/document/4803.pdf

Workstream 5 – Novel Analysis Techniques

Required Outputs:

1) Prototype script with novel analysis technique algorithm; and
2) User requirements as input to functional specification for a future network modelling tool.

The objectives of this workstream were achieved.

A novel analysis technique was developed for modelling LV networks and is described in the detailed project report:

https://www.northernpowergrid.com/asset/0/document/4918.pdf

Detailed mathematical algorithms are presented in the appendices of this report. A number of prototype scripts were developed in the R programming language and the results of studies carried out on the Sinderby and Cranwood LV test networks are shown in this report.

The two mutli-voltage level models developed in Workstream 3 were used to carry out novel studies exploring holistic voltage behaviour under a range of network load states and topologies. This is described in detail in the project report:

https://www.northernpowergrid.com/asset/0/document/5323.pdf

Processes were developed to allow different scenarios to be run and processes developed to display voltage profiles visually across different voltage levels in Excel.

User requirements were developed and a functional specification produced for a future LV modelling tool:

https://www.northernpowergrid.com/asset/0/document/5218.pdf

6.2 Project performance against objectives

Five project objectives were defined in the NIA PEA document:

1. Deliver recommendations on improved network EHV/HV/LV network build and holistic network analysis under a range of conditions

This objective has been achieved in Workstream 3 and Workstream 5.

- Workstream 3 – A methodology has been developed to build multi-voltage EHV/HV/LV networks. These use existing power system models built in IPSA and DINIS, SCADA data held in the PI system, and load growth models for secondary substations provided by Element Energy. Two multi-voltage test networks have been developed. These test networks were used to study voltage behaviour holistically across the voltage levels under first circuit outage conditions, and a range of demand and generation conditions; and
- Workstream 5 – These test networks were used to study voltage behaviour holistically across the voltage levels under first circuit outage conditions, and a range of demand and generation conditions.

2. Provide recommendations to improve network planning and design solutions

This objective has been achieved in Workstream 5, by development of i) LV Novel Analysis Techniques, and ii) Multi-Voltage Level Novel Analysis:

i) <u>LV Novel Analysis</u>

Methodologies have been developed which improve on the existing ACE 49 and ADMD methods for analysing thermal utilisation and voltage levels on LV networks. The methodologies describe a risk based approach to LV design, using existing LV network models held in NPg's eAM Spatial system, and using smart meter consumption data to model customer demand.

ii) <u>Multi-Voltage Level Novel Analysis</u>

We demonstrated through the use of multi-voltage level models how cross network voltage issues can be better studied and solutions tested. For example:

- Ability to assess the effects of the connection of generation at HV combines with high uptake of PV at LV; and
- The implications of a voltage management solutions applied at a primary substation can be assessed at HV and LV, such as LDC or seasonal target voltage set points.

For the test networks studied we found that NPg's existing voltage management policy is suitable for existing loading conditions.

The two multi-voltage level network models are available to the business for planners to use in the future, whether for strategic level studies or to test design solutions. A methodology is described to build further multi-voltage level networks to model other parts of the network if required.

3. Provide recommendations on how to deal with challenges for smart meter data utilisation

This objective has been achieved in Workstream 4 and Workstream 5. To summarise:

- Analysis has shown that aggregation levels of two or greater are sufficient to anonymise customer consumption half-hourly data;
- A methodology has been developed to create customer demand models based on smart meter data. The methodology uses a base model using on CLNR data. Through a process of Bayesian updating these can be updated with smart meter data for specific LV networks as this data becomes available to the DNO. The methodology also allows the models to be continually refreshed and updated as more recent smart meter data becomes available; and
- Although we were unsuccessful in being able to demonstrate a method to identify LV customer phase connectivity using voltage correlation techniques, we believe that the method should be re-examined if in the future when more voltage data from smart meters becomes available, an LV network with a large amount of phase imbalance and high rollout of smart meters is identified.

4. Validation of equipment specifications

This objective has been partially achieved through the multi-voltage level analysis work. The two networks studied are not likely to be fully representative of the wide and varied range of network characteristics across NPg so we recommend further verification of these observations and recommendations following application of the MVL on other networks. However, to summarise, based on the test networks we did study:

- The tap range of primary transformers can accommodate a wide range of loading conditions although for the representative networks analysed, the primary transformer tap changers are typically operating towards the higher end of the tapping range, potentially reducing the amount of generation headroom.
- The assumptions related to the modelling of tapchanger deadband can be material in those situations where network voltages are approaching their upper or lower limits. The risk associated with tapchange relays spending a significant amount of time at the extremes of the deadband range should be considered in more detail.
- The specification of and the loading on supply point transformers is such that the requirements of Engineering Recommendation P10 are satisfied.
- Load drop compensation could be implemented to cater for various future demand and generation conditions at HV and LV although the application would be network specific and would need to consider the loading patterns of all feeders connected to those networks – LDC is more effective where loading patterns are fairly similar under normal and under contingency conditions.
- Target voltage settings on new AVC relays fitted with remote communications could be changed remotely on a seasonal basis. The MVL modelling approach can be deployed to develop an improved understanding of when to deploy target voltage changes and/or LDC, how application of these solutions might interact including during contingency conditions, and how application of these solutions could best be reflected in both the planning and design process.
- Advantages were seen in being able to seasonally tap secondary HV/LV transformers where there is a high uptake of solar PV generation. As these transformers have off-load tap changers, this could justify the replacement of the transformer with one with an on-load tap changer.

5. Creation of a set of requirements for future functional specifications for new power systems software

As described previously in Workstream 5, this objective has been fully achieved. A functional specification which describes a minimum viable product for a new LV design tool has been developed.

Through development of the multi-voltage level models, knowledge has been gained on how to combine EHV, HV and LV models held in different software packages and how to develop appropriate scripts to manage this process. This is informing our requirements for our suite of new DSO analysis tools.

6.3 Project performance against success criteria

Three success criteria were defined in the NIA PEA document:

1. The network methodologies developed actually propose potentially acceptable solutions for addressing smart metering challenges which will significantly improve the design and planning assumptions especially at LV

This criterion has been met through the development of the LV novel analysis techniques and the functional specification for a new LV design tool ('Smart LV Design').

2. The network methodologies allow the modelling of more innovative solutions due to the improved knowledge and visibility of the holistic operation of the combined networks

This criterion has been met. The multi-voltage level models have been developed in the IPSA power systems modelling tool. This tool is capable of modelling traditional and innovative solutions. In our work we prioritised modelling of seasonal target voltage changing on primary transformer AVC relays, application of LDC, on-load HV/LV tap changers at secondary distribution substations, interconnection with adjacent networks and load shifting, using generator voltage / reactive power control functionality. The tool can also model STATCOMs and capacitor banks. It would also be possible to model Demand Side Response, possibly through use of the Python scripting functionality of IPSA.

3. The network methodologies assist in achieving the potential £5M of network reinforcement benefit due to use of smart metering data

The LV methodologies developed provide a way to model thermal utilisation and voltage more accurately than existing methods allowing the selection of conductor size for new connections and reinforcement to be more precise, however the benefits cannot be considered in isolation from economic considerations such as optimisation to reduce losses.

The functional specification for a new LV design tool 'Smart LV Design' could either be taken forward as a standalone product to be developed or components used as part of a broader suite of LV design tools. A separate NPg innovation project, AutoDesign (a LV connections Self Service Tool) has a user front-end that would have a very similar graphical interface to Smart LV Design. There is the potential to use the AutoDesign front-end with alternative back end calculation engines such as the one described in the Smart LV Design functional specification.

7 Required modifications to the planned approach during the course of the project

The following modifications were required to the planned approach:

- The project did not include a budget for LV network monitoring. However, there is a programme to deploy LV network monitoring at certain secondary distribution substations. Once the test networks were identified Sinderby and Cranwood were prioritised for deployment within that programme. However, there were delays to the deployment of equipment within that programme and by the time site surveys had been carried out our there would have been no time left in our project to be able to record sufficient data and incorporate it into our analysis. Therefore we did not include LV network monitoring data into our study cases.
- Rather than collecting and analysing SMETS2 smart meter data, the CLNR data set was used as a substitute. There were several reasons for this:
    - **Low deployment of SMETS 2 smart meters** - When Workstream 4 started in April 2018, there were only 20 SMETS2 smart meters deployed in NPg's two licence areas, rising to 339 in September 2018. The CLNR smart meter data set, specifically the TC1a data, includes approximately 8,000 half-hourly consumption time series, covering two and a half years and includes two winters.
    - **Data access and security** – NPg's Siemens EIP head end system had only recently gone live when this project started, and is deployed within its own secure environment with limited authorised personnel being able to access it, none of who whom were on the project team.
    - **Data Privacy Plan not approved** - NPg's data privacy plan concerning the collection and use of smart meter data consumption data had not been submitted to nor approved by Ofgem.
- During the Horizon Scanning Workstream, within the project team there was a lack of understanding of the existing software tools and data sources available across NPg departments and at the different voltage levels. A side activity was initiated which documented the existing NPg tools and data sources as well as the design processes used at EHV, HV and LV.
- Scripted IPSA did not have the feature to produce output voltage results from unbalanced load flows into external files. This introduced an unexpected delay into Workstream 4, but was managed by delaying the LV phase identification work until this feature was incorporated into IPSA by TNEI.

8 Lessons learnt for future projects

The main lessons learnt during the project were:

- In order to build LV electrical network connectivity models suitable for power system analysis from records held in Geographic Information Systems, a large amount of data cleansing is required and the amount of manual effort required should not be underestimated;
- The project assumed that a few months' data would be available from LV network monitoring on the test networks. By the time the test networks were identified during the horizon scanning workstream and equipment deployed, too much time had passed for it to be of use. The time period required to capture data and the time required for procuring any monitoring equipment needs considering for future projects. Improved alignment in the timescales of relevant related projects should be considered in future; and
- Future projects relying on the use of a smart meter data should confirm whether the required volume and access to that data is available within the timescales of the project. They should also

examine in detail the accuracy, resolution and period of time over which voltage and consumption data is captured and stored in GB smart meters. Many international academic papers rely on data sets from a small set of smart meter data with access to data at accuracy levels and time steps which are not realistic in the context of GB smart meters.

9 The outcomes of the project

Prototype software scripts in the R programming language are available for:

- Disaggregating smart meter consumption data;
- Determining phase connectivity of LV customers using voltage correlation techniques based on K-means clustering;
- Creating probability distributions of load at a given time based on demand characteristics.

We have developed novel analysis techniques at LV which describe a risk-based approach utilising new data sources to be enable LV design which improves on existing methods. This includes the production of a functional specification for a minimum viable product.

SQL and Python scripts to extract LV networks from NPg's Spatial database into IPSA have been developed.

A methodology and process (including Python scripts) has been developed to create multi-voltage level network models in IPSA from various sources including: Spatial, DINIS, PI, Element Energy Load Growth models and SCADA.

Two working multi-voltage models in IPSA have been developed, one for the Norton/Leeming Bar/Thirsk network and another for the Creyke Beck/ Beverly/First Avenue network. Design studies and the testing of voltage control solutions have been carried out on these networks. We have been able to demonstrate that multi-voltage level models allow a more holistic approach to voltage control and management to be made.

10 Planned implementation

The LV novel analysis work will feed into national reviews currently underway on Engineering Recommendation P5 ('Design methods for LV underground networks for new housing developments') to assess the case of establishing a more consistent approach to the design of LV networks. If there is a business case to adopt the novel LV methodologies in NPg, it is likely that these would be developed alongside the AutoDesign tool, as this would reduce the development effort required to build a graphical front-end for the tool. It should also be recognised that LV network monitoring and smart metering should provide the default understanding of demand on LV networks, with modelling being used to fill in the gaps, a "measure what you can, model what you can't" approach.

Valuable knowledge has been gained from building the multi-voltage network models including extraction of data from SCADA, converting EHV/HV and LV models stored in different tools into a single model, and the how scripting to assist in automating these processes. These lessons learned will be used to inform the requirements for the set of DSO analysis tool suite we are planning to procure.

11 Project budget

The project budget was £399,510 which rose to £411,910 by the end of the project. The budget increase was due to the cost of expenses not being factored into one of the contractor's original estimate. This was within the project contingency of 10% of the project budget.

12 Technology Readiness Level

The TRL at the start of the project was three ('Experimental proof of concept'). The target TRL of four ('Technology validated in lab') was achieved.

13 Learning dissemination

The project has been disseminated via the following channels:

- A paper was accepted and published at CIRED 2019 - "Novel Analysis Techniques for LV Network Planning using Smart Meter Data";
- Gordan McFadzean (TNEI) gave a presentation on the novel LV analysis work at CIRED 2019;
- The project was exhibited by NPg at LCNI 2019. Over the course of the two day conference, the project manager was available at a dedicated stand to present the project and answer any queries;
- A dedicated micro-site has been set up for the project on NPg's innovation portal: https://www.northernpowergrid.com/innovation/projects/smart-network-design-methodologies

  Extensive project reports are available for download;
- TNEI hosted a CIGRE UK webinar on the project in November 2019;
- Lunch and learn sessions in NPg's Northeast and Yorkshire areas took place in February and March 2020.

Later in the year TNEI plan to:

- Present the project at the ENA Energy Innovation Forum;
- Publish a paper at Cigre 2020 titled "Quantifying risk in low voltage network planning using smart meters and probabilistic modelling".