Electricity data analytics: Opportunities and Challenges

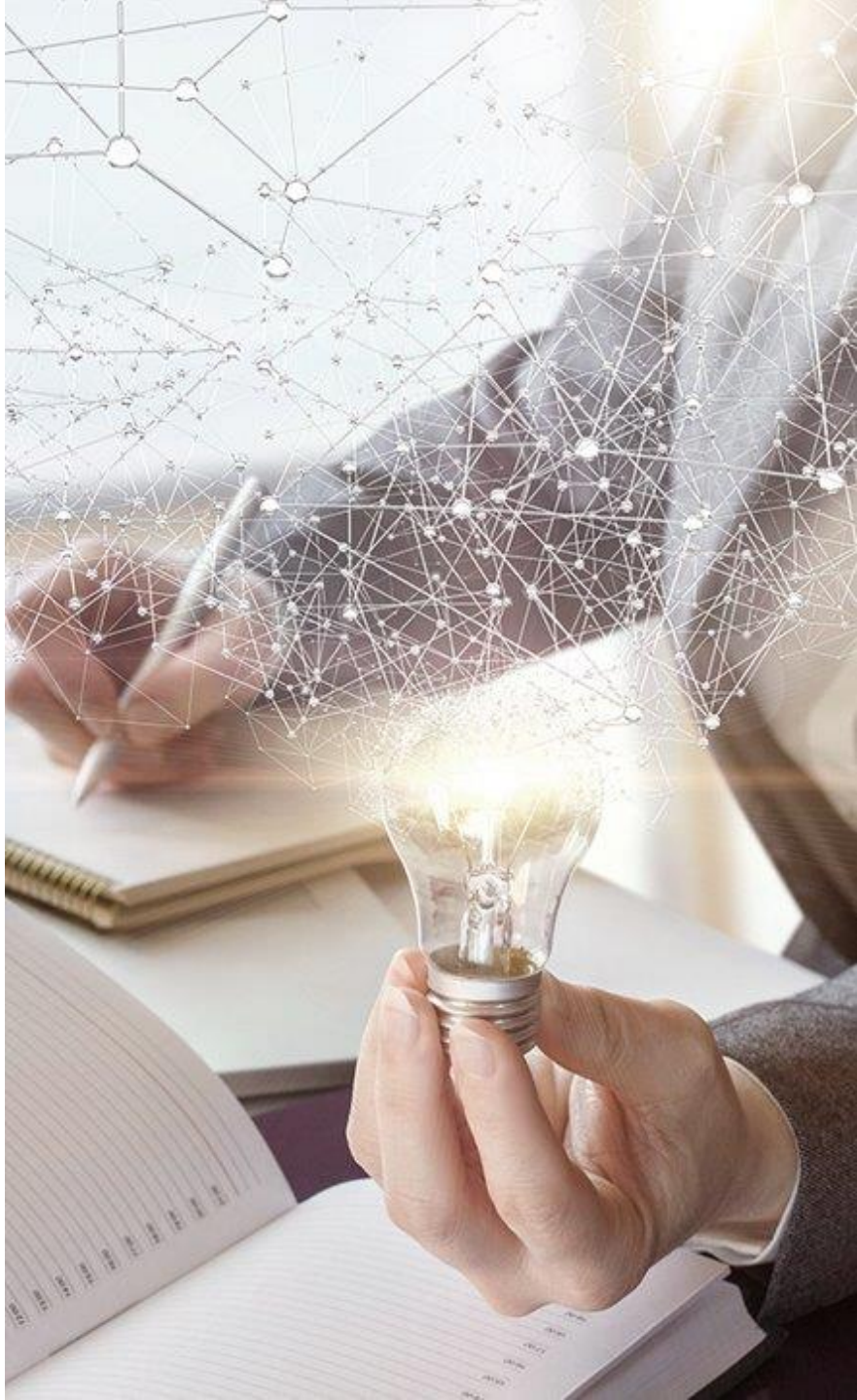
Dr Stephen Haben

Digital and Data Consultant– Energy Systems Catapult
Visiting Research Fellow – University of Oxford

20th October 2021

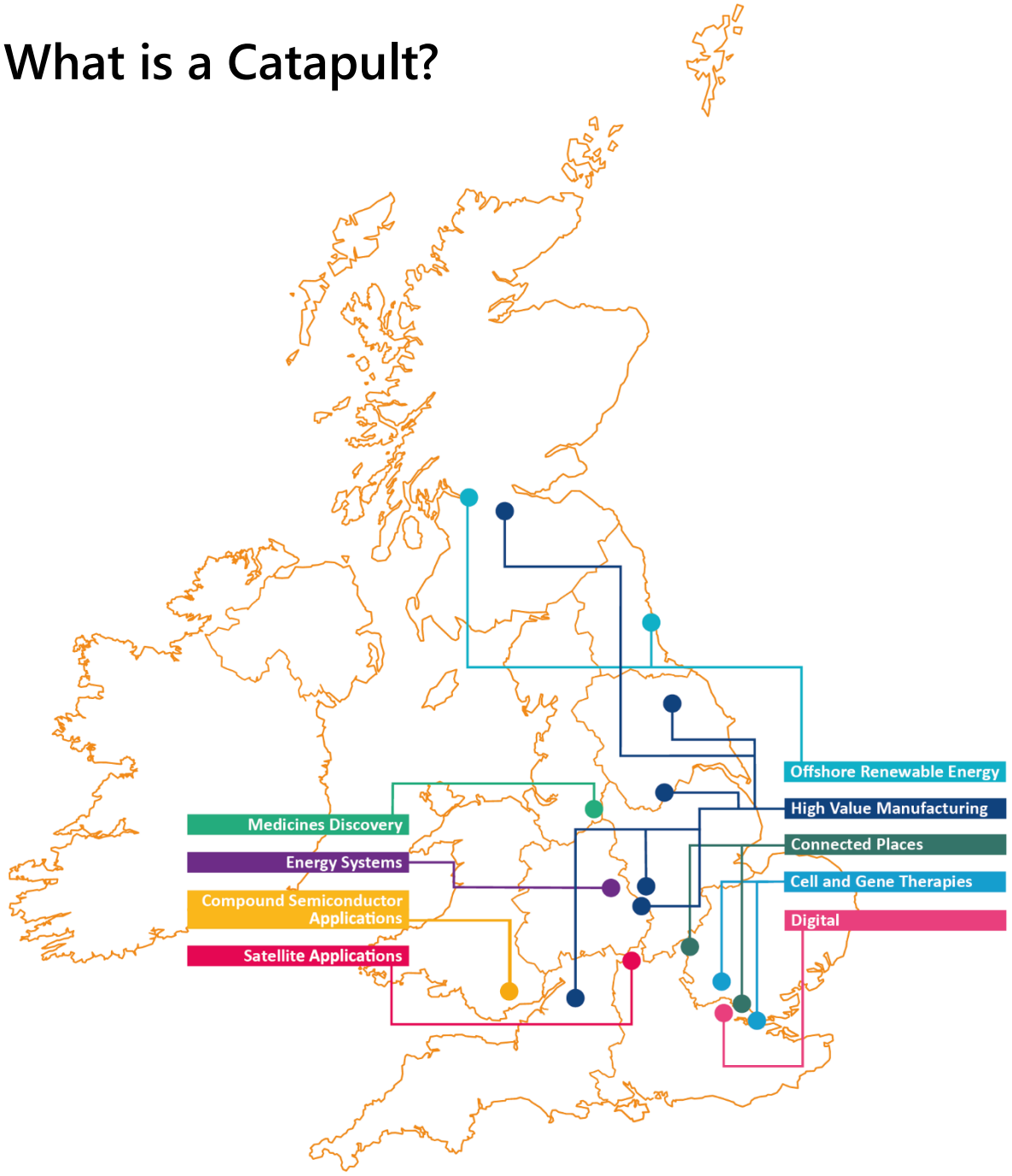
@EnergySysCat

Contents



Introduction to Energy Systems Catapult

What is a Catapult?

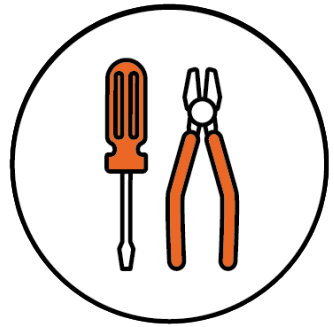


CATAPULT
Energy Systems

We work with Innovate UK

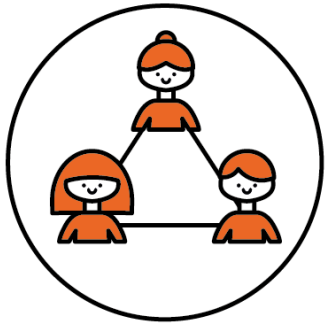
Established and overseen by Innovate UK

9 catapults



Technical capabilities, equipment, and other resources

Solve key problems and develop new products and services



Bridge the gap between stakeholders in the sector

Open up opportunities for innovators, in the UK and globally

About Energy Systems Catapult

Mission: **Unleash innovation and open new markets to capture the clean growth opportunity**

200 Innovation experts

Hubs in Birmingham and Derby

We work with Innovate UK

Established, overseen and part-funded by Innovate UK. Independent from Government. Not for profit

Bridge the gap between stakeholders in the sector

Supporting innovators

Research

Trials

Systems engineering

Digital

Modelling and simulation

Our Capabilities and Assets

Modelling

National Energy System Modelling
Local Area Energy Planning and Modelling
Building Energy System Modelling

Energy System Modelling Environment™
EnergyPath Networks™
Home Energy Dynamics
Storage and Flexibility Model

Digital and Data

Data Science
Data Systems

Living Lab
Energy Knowledge eXchange™

Consumer Insight

Research
Design
Trials

People Lab
Home Truths®

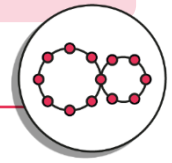
Markets, Policy and Regulation

Policy and Regulatory Knowledge
Economic Appraisal

Systems Integration

Systems Engineering and Integration
Dynamic Energy System Simulation
Dynamic Energy System Architecting
Business Model Innovation
Energy System Integration Guides

EnergyPath Operations™

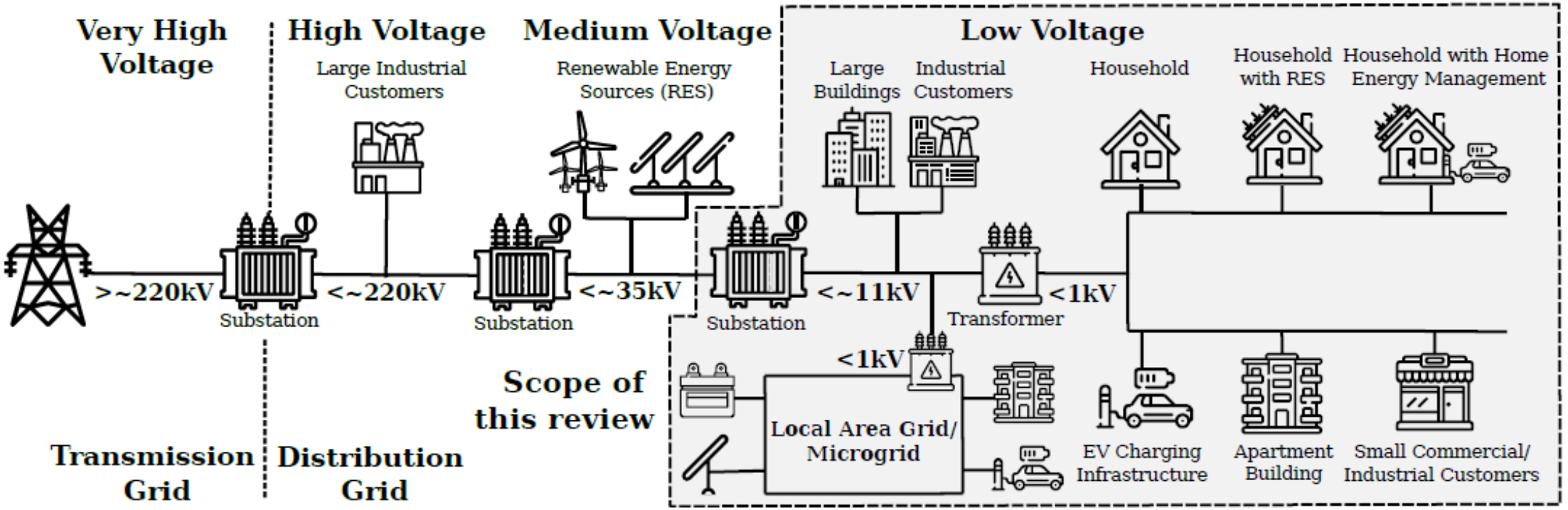


Infrastructure and Engineering

Networks and Energy Storage
Renewables
Transport
Nuclear
Carbon Capture and Storage, Industry and Hydrogen
Bioenergy

Challenges: LV Forecasting Case Study

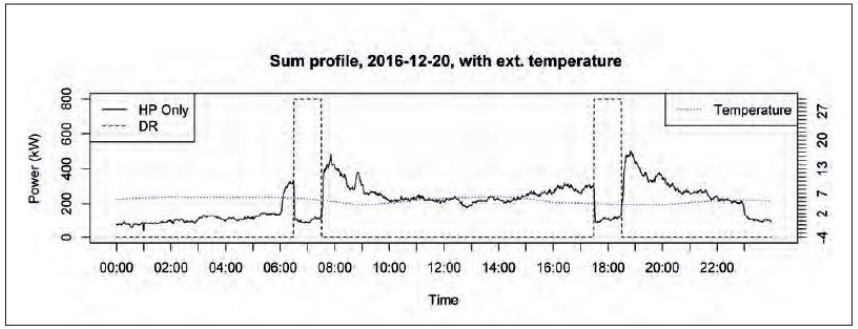
What do we mean by LV?



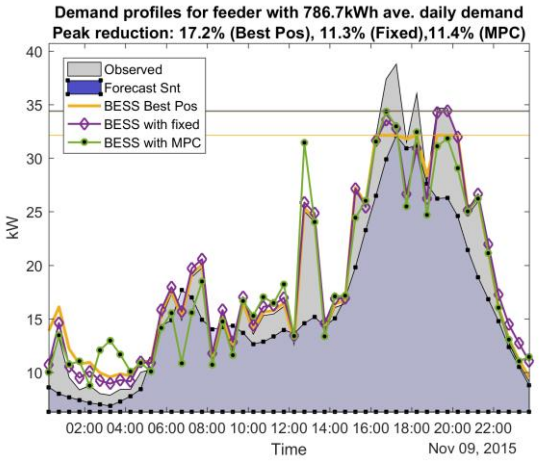
Picture from Review of low voltage load forecasting: Methods, applications, and recommendations, Haben et al. Applied Energy, 2021. https://doi.org/10.1016/j.apenergy.2021.117798

Forecasting Applications

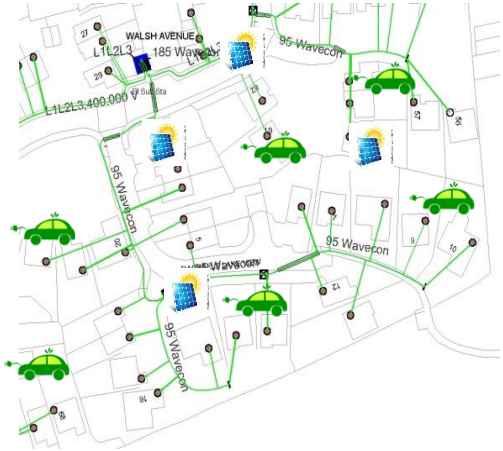
Heat pump demand side response



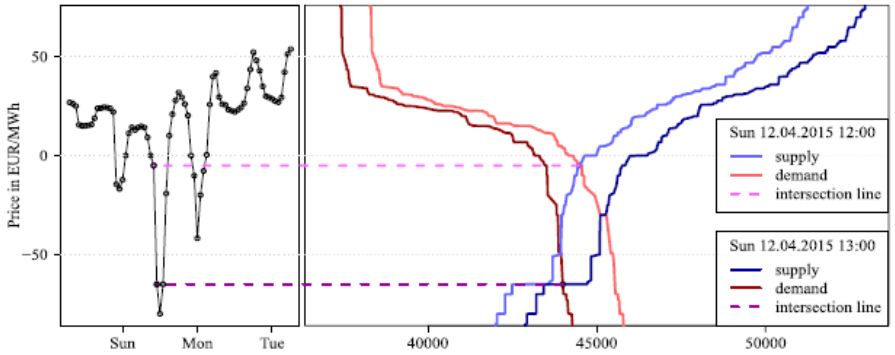
Smart management systems



Long term scenario planning



Day ahead wholesale electricity markets



- **Network Design and Planning** – location and sizing of substations, location sectionalising switches, storage location.
- **Network Operations and Control** – grid management, storage, feed-in limits, minimise curtailment losses, cost reductions, maximise PV hosting capacity, Voltage control …
- **Anomaly Detection –** theft detection systems, malicious attacks, early warning systems, …
- **Trading** – Peer-to-peer trading, feed in to market responses, energy trading algorithms, ….
- **Simulating Inputs, Missing data, Privacy Protection** - imputing missing values, generating pseudo observations for state-estimation, differential privacy, …

Pictures clockwise from top left:

http://media.ontheplatform.org.uk/sites/default/files/GMCA%20NEDO%20Smart%20Communities%20Exec%20Report%20FINAL.pdf

Evaluating the effectiveness of storage control in reducing peak demand on low-voltage feeders, T. Yunusov, S. Haben, T. Lee, F. Ziel, W. Holderbaum, B. Potter, Proceedings CIRED 2017

Long term individual load forecast under different electrical vehicles uptake scenarios, A. Poghosyan, D. V. Greetham, S. Haben and T. Lee, Applied Energy, vol. 157, pp. 699--709, 2015

Electricity price forecasting using sale and purchase curves: The X-Model, Florian Ziel, Rick Steinert, Energy Economics, Volume 59, 2016,

Narrative Context this Talk: Storage Control for LV Networks

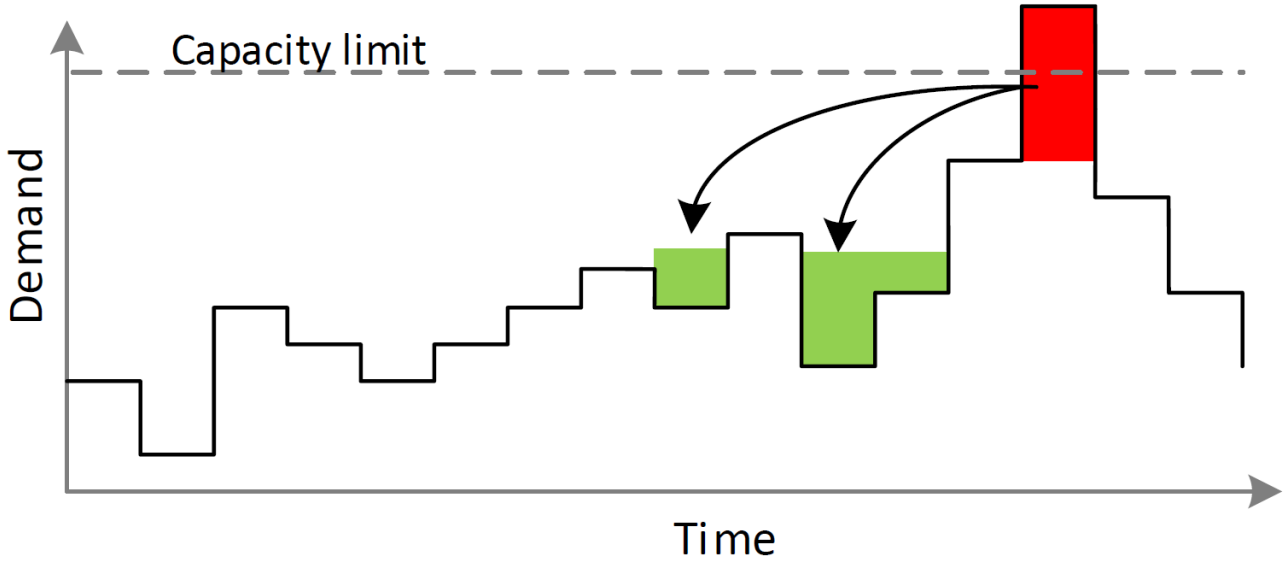


Photo courtesy Timur Yunusov

Storage Control LV Residential Feeders

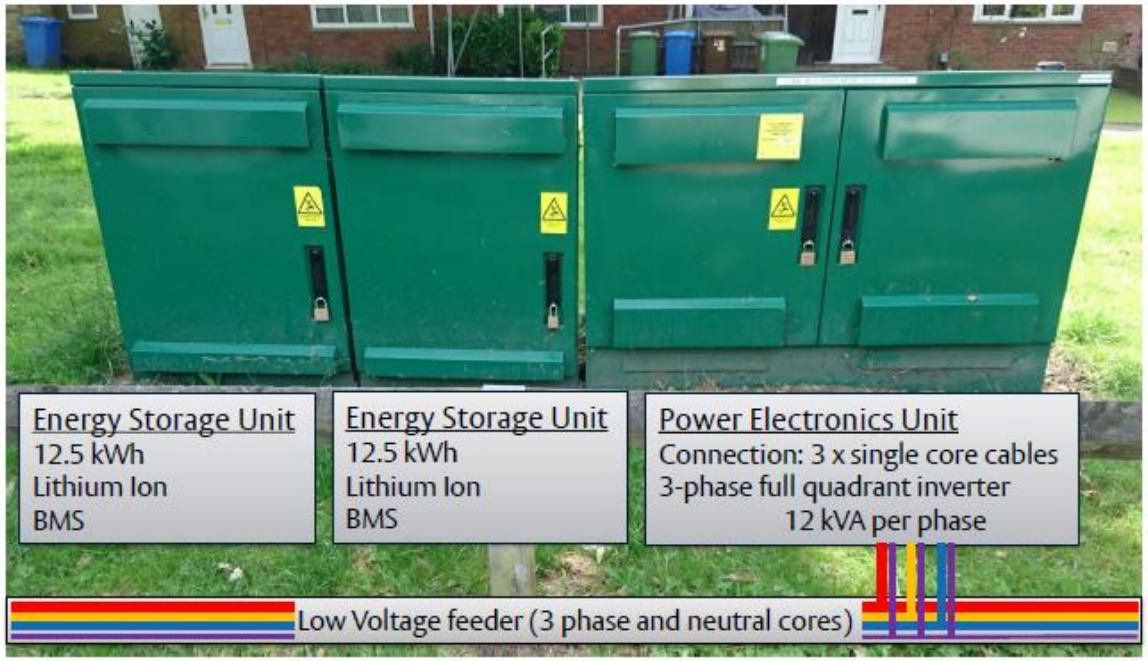
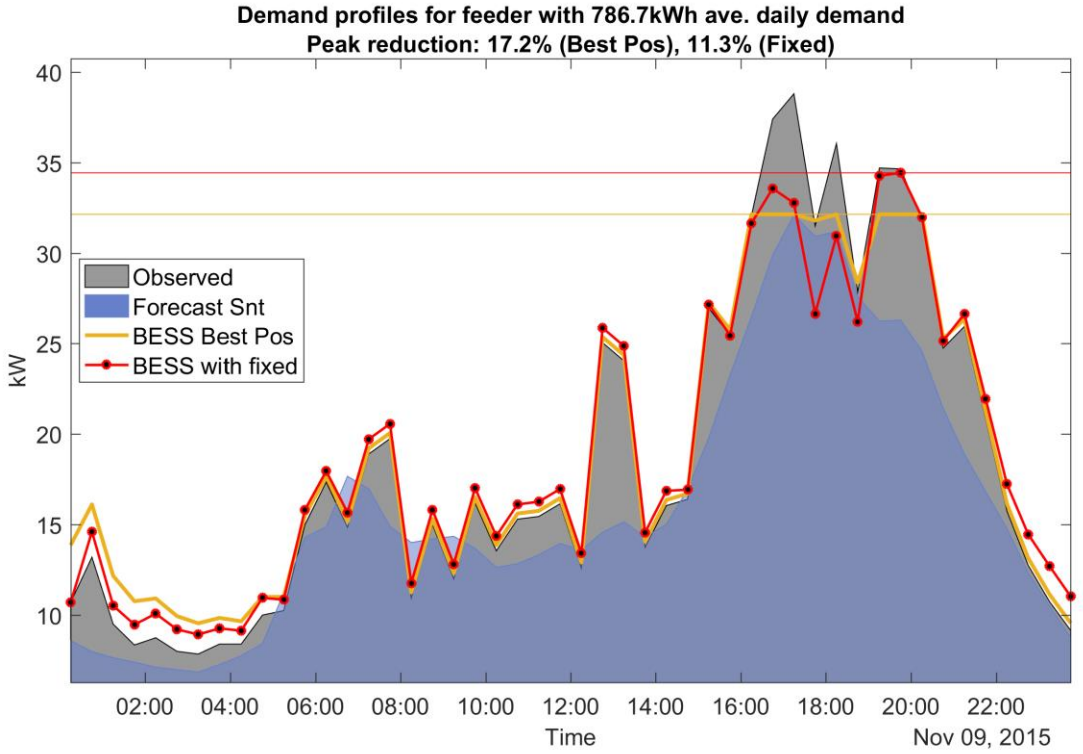
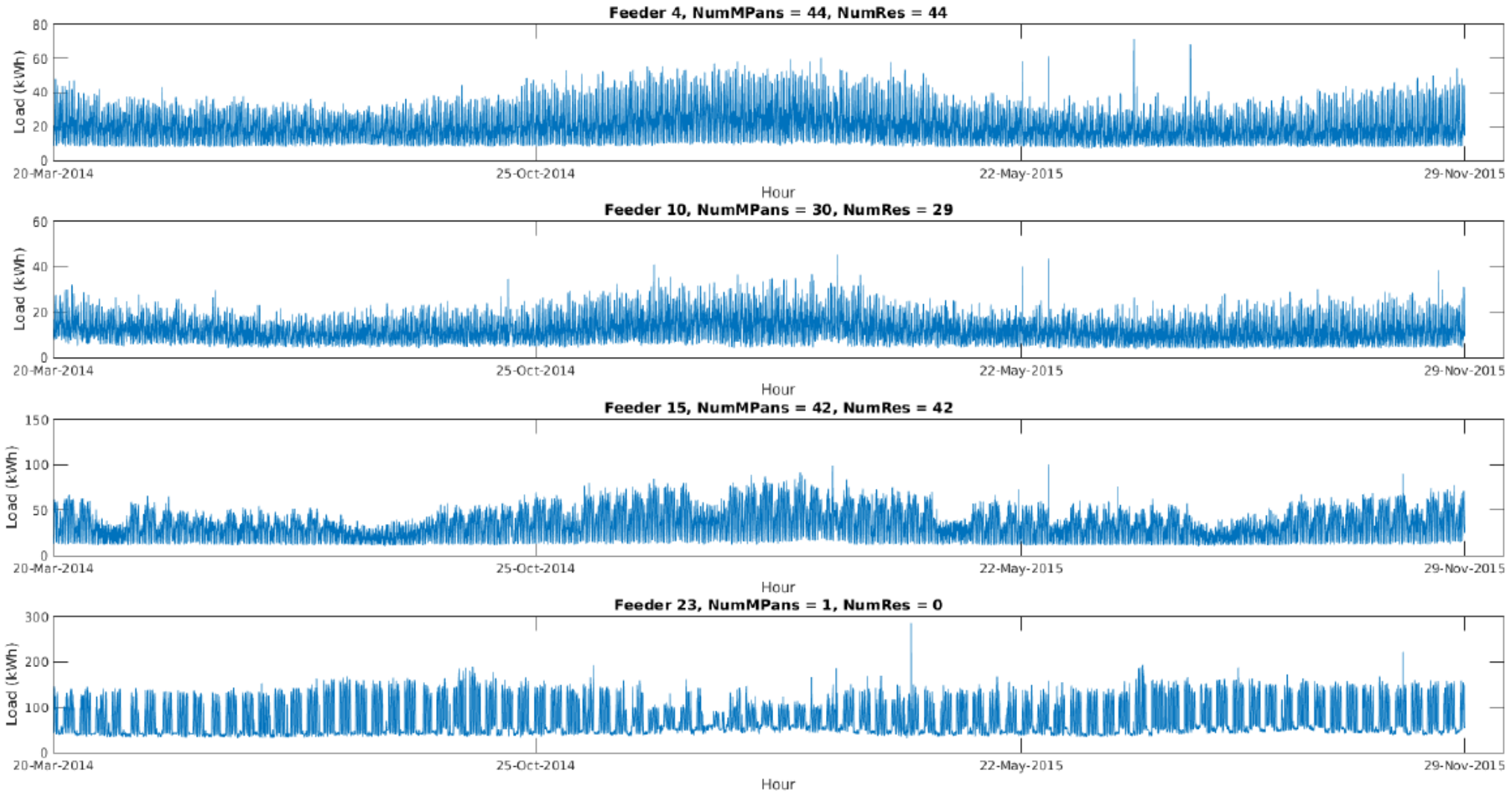


Photo courtesy Timur Yunusov

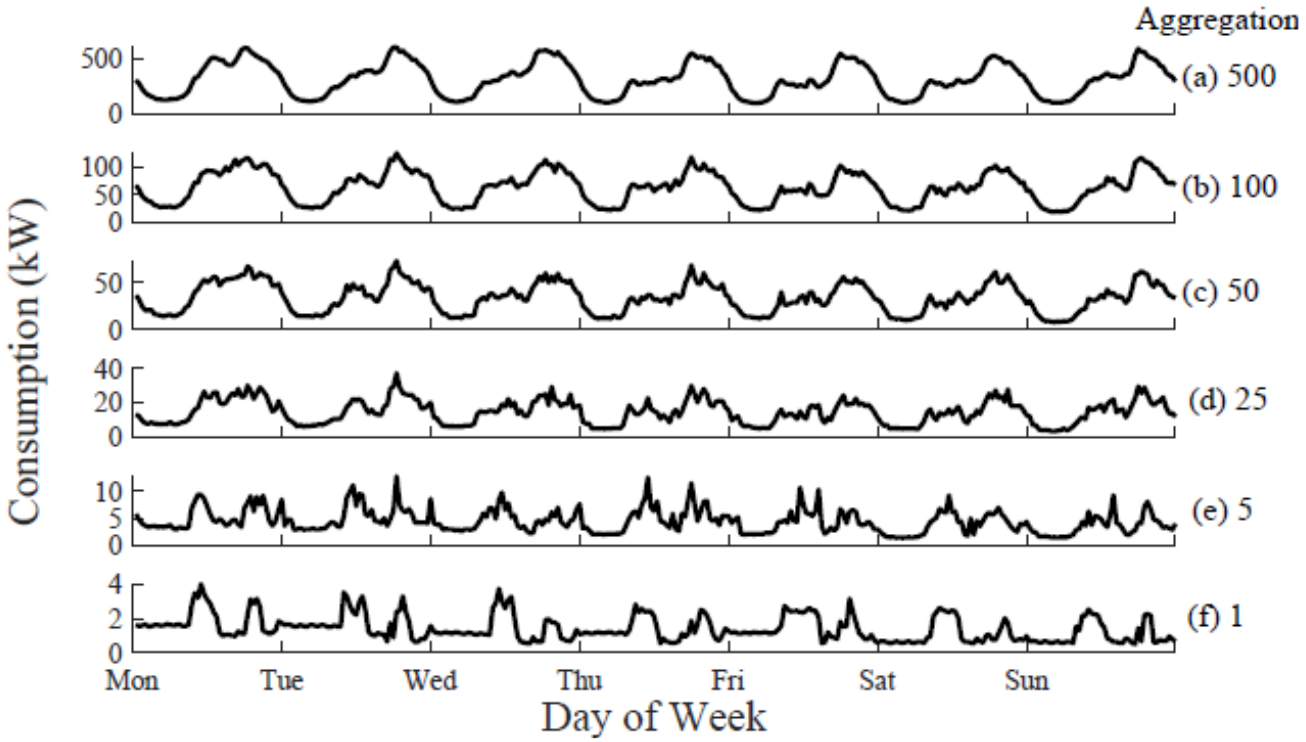


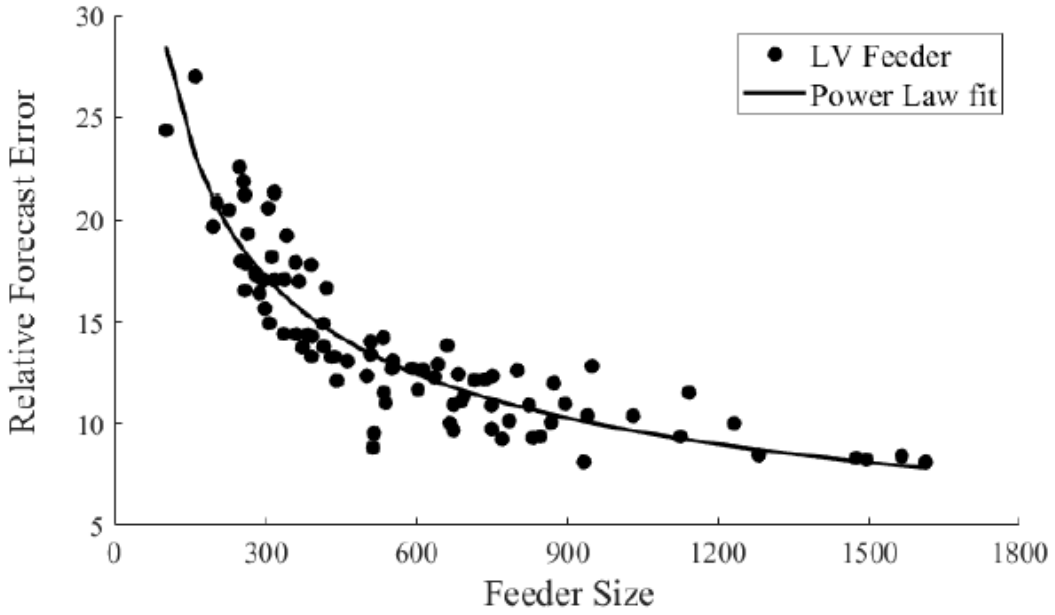
Picture: Evaluating the effectiveness of storage control in reducing peak demand on low-voltage feeders, T. Yunusov, S. Haben, T. Lee, F. Ziel, W. Holderbaum, B. Potter, Proceedings CIRED 2017

LV Distribution Demand



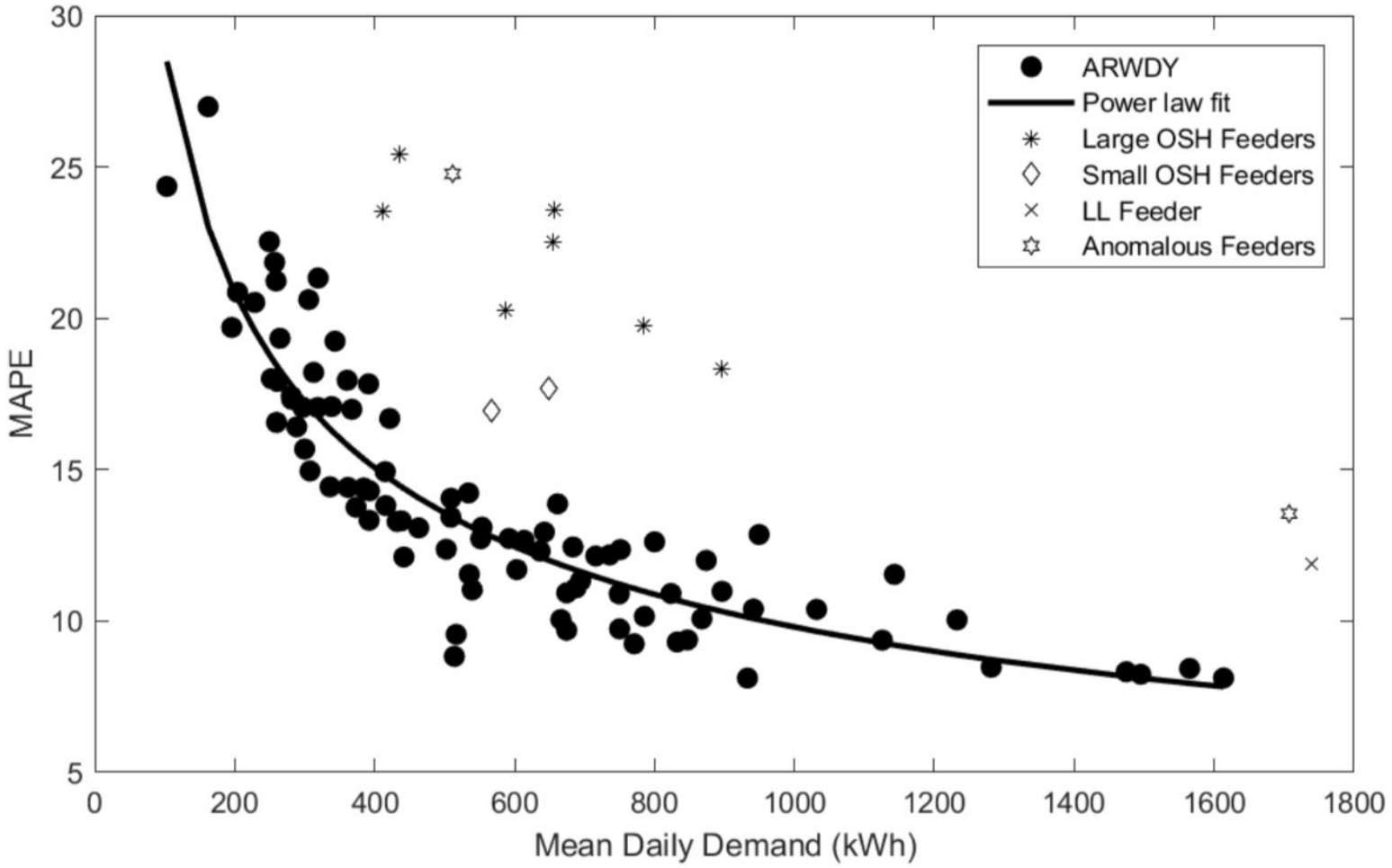
LV feeder Size and Forecast Accuracy

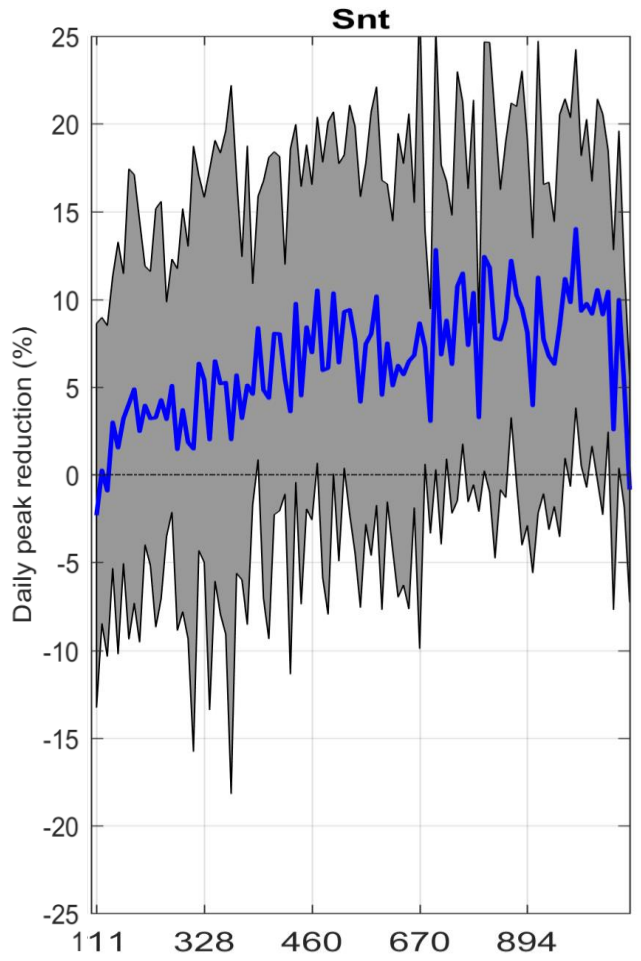




Pictures from Review of low voltage load forecasting: Methods, applications, and recommendations, Haben et al. Applied Energy, 2021. https://doi.org/10.1016/j.apenergy.2021.117798

Residential LV networks – just aggregate of standard households?

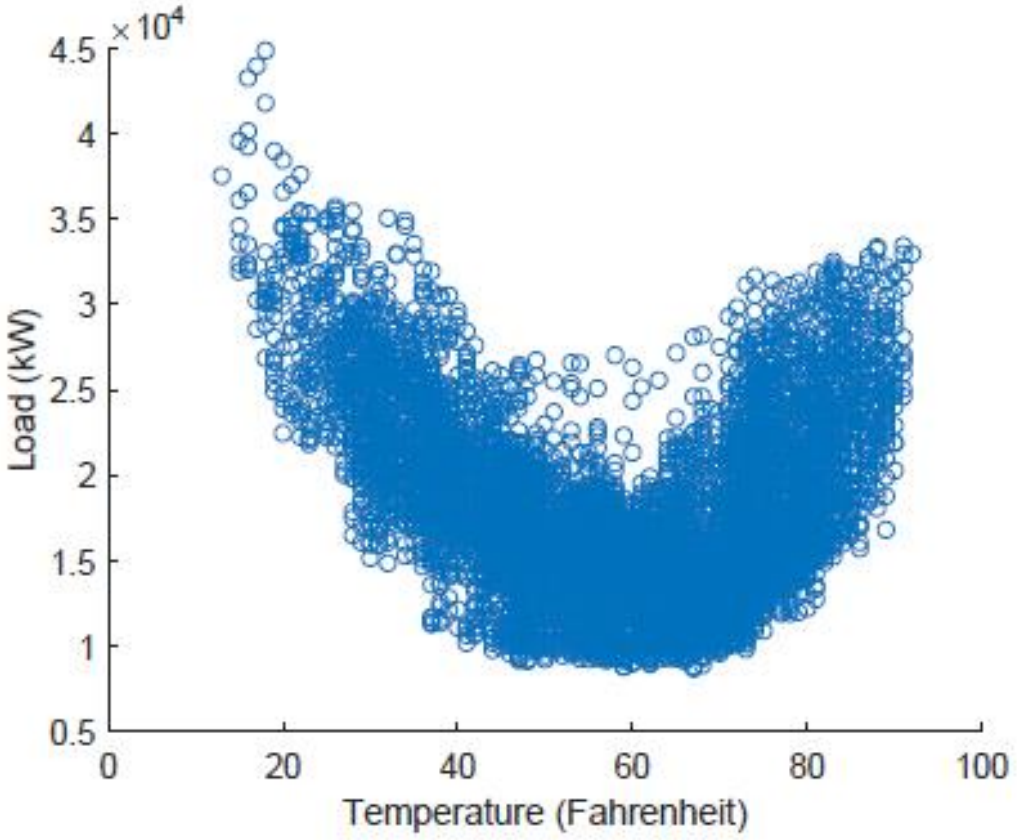




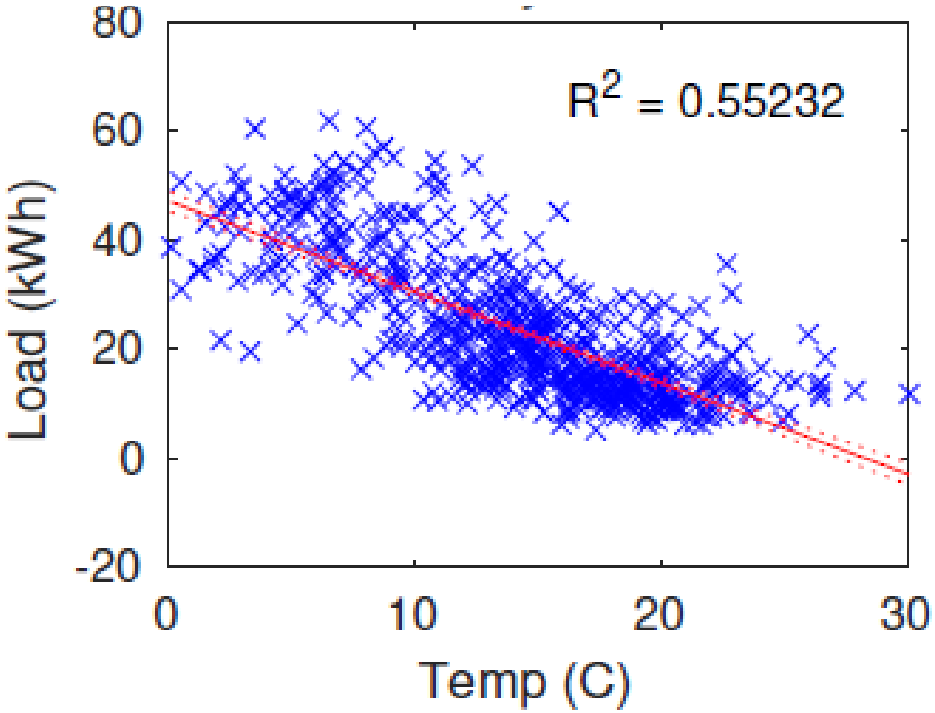
Sources: Short Term Load Forecasts of Low Voltage Demand and the Effects of Temperature, S. Haben, G. Giasemidis, F. Ziel and S. Arora, International Journal of Forecasting, 2019.
Evaluating the effectiveness of storage control in reducing peak demand on low voltage feeders, T Yunusov, S Haben, T Lee, F Ziel, W Holderbaum, B Potter, 24th International Conference & Exhibition on Electricity Distribution (CIRED), Glasgow, 2017.

Influence of Temperature?

GEFCOM 2014 – US Data

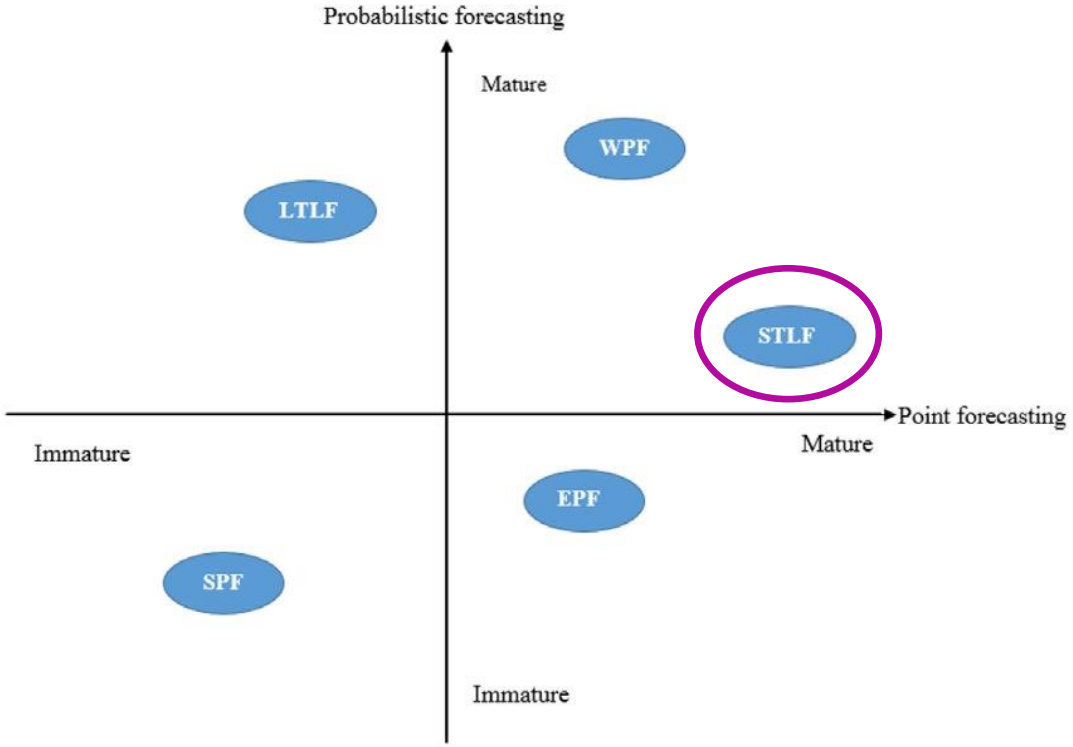


TVV Data 2016 – UK LV demand Data

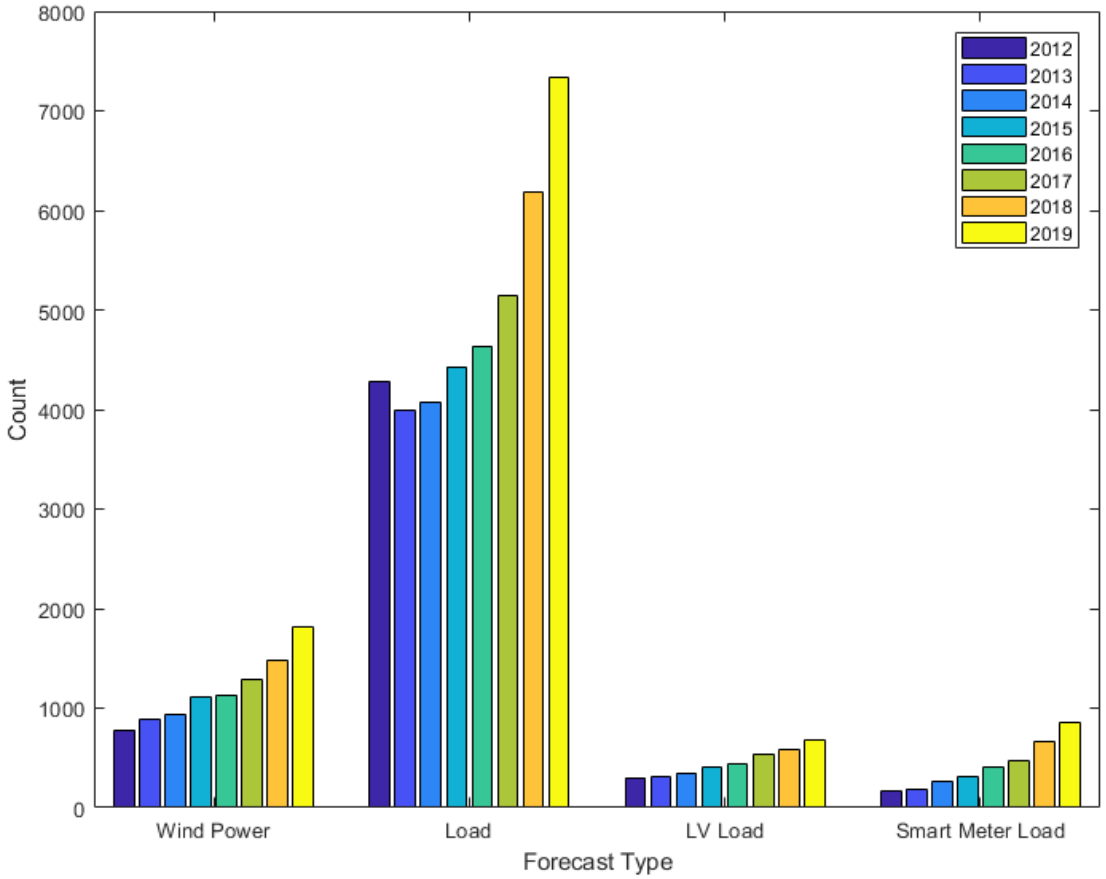


GEFCOM data from: Tao Hong, Pierre Pinson, Shu Fan, Hamidreza Zareipour, Alberto Troccoli and Rob J. Hyndman, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond", International Journal of Forecasting, vol.32, no.3, pp 896-913, July-September, 2016.

Current state of load forecasting



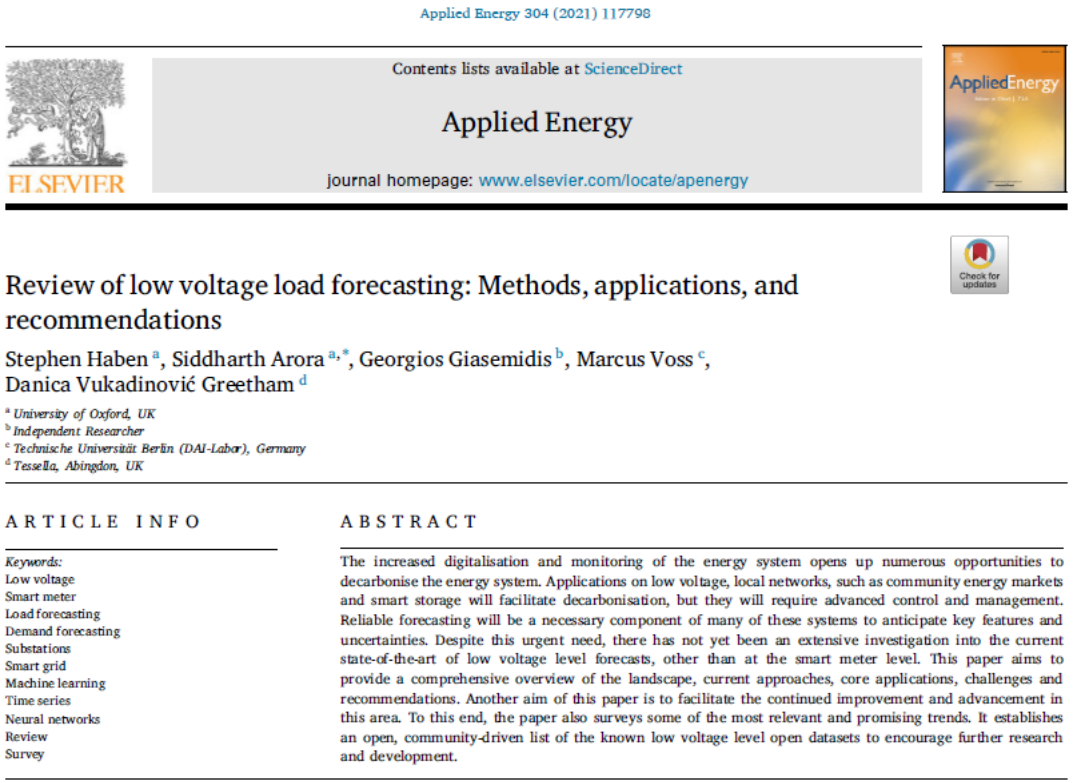
Picture credit: Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond, T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, R. J. Hyndman, International Journal of Forecasting, 2016.



Source data: Google scholar.

Issues: Review 221 Papers in LV forecasting

- Many papers use no benchmark, if they do it is non-competitive.
- Forecasts often ignored as part of an application
- Only 44 papers (<22%) utilised probabilistic forecasts.
- Methodologies unclear:
  - Size of Validation/Testing period
  - Resolution of data
  - Forecast horizon
- Resolution Issues:
  - 99 papers were at resolution of half hourly or hour.
  - Only 11 papers considered data of resolution of 1 minute or less
- Data Issues (Next Slides)



Applied Energy 304 (2021) 117798

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Review of low voltage load forecasting: Methods, applications, and recommendations

Stephen Haben [a], Siddharth Arora [a,*], Georgios Giasemidis [b], Marcus Voss [c], Danica Vukadinović Greetham [d]

[a] University of Oxford, UK
[b] Independent Researcher
[c] Technische Universität Berlin (DAI-Labor), Germany
[d] Tessella, Abingdon, UK

ARTICLE INFO

Keywords:
Low voltage
Smart meter
Load forecasting
Demand forecasting
Substations
Smart grid
Machine learning
Time series
Neural networks
Review
Survey

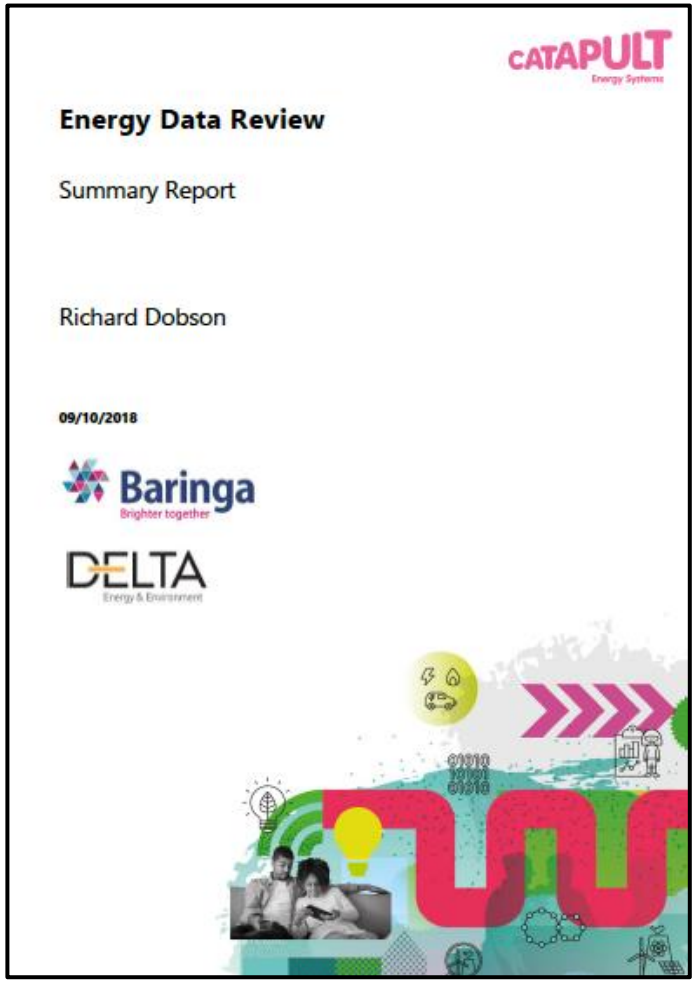
ABSTRACT

The increased digitalisation and monitoring of the energy system opens up numerous opportunities to decarbonise the energy system. Applications on low voltage, local networks, such as community energy markets and smart storage will facilitate decarbonisation, but they will require advanced control and management. Reliable forecasting will be a necessary component of many of these systems to anticipate key features and uncertainties. Despite this urgent need, there has not yet been an extensive investigation into the current state-of-the-art of low voltage level forecasts, other than at the smart meter level. This paper aims to provide a comprehensive overview of the landscape, current approaches, core applications, challenges and recommendations. Another aim of this paper is to facilitate the continued improvement and advancement in this area. To this end, the paper also surveys some of the most relevant and promising trends. It establishes an open, community-driven list of the known low voltage level open datasets to encourage further research and development.

Data Issues

Data: General Issues

- Not open or only shared with strict conditions.
- Not clear what data exists.
- Documentation – often unclear but sometimes doesn't even exist.
- Unclear or no Licensing attached – what are you allowed to do with the data?
- No uniform standards or meta-data formats.
- Not easy to find or easy to search formats.
- Much data not in Machine Readable Format.



Review 221 Papers in LV forecasting Revisited: Further Data Issues

- Only 52 papers used at least 1 openly available data sets. Of these
  - 22 (42%) used Irish Smart meter data.
  - 4 used UK Low Carbon London.
  - 4 used Ausgrid.
  - 3 the Umass dataset.

- Irish Data:
  - A decade old.
  - only 2 years worth of data.
  - Many in trial also subject to interventions.
  - Most households 3-4 bedrooms.

- List of LV data sets here: https://low-voltage-loadforecasting.github.io/

Applied Energy 304 (2021) 117798

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Review of low voltage load forecasting: Methods, applications, and recommendations

Stephen Haben [a], Siddharth Arora [a,*], Georgios Giasemidis [b], Marcus Voss [c], Danica Vukadinović Greetham [d]

[a] University of Oxford, UK
[b] Independent Researcher
[c] Technische Universität Berlin (DAI-Labor), Germany
[d] Tessella, Abingdon, UK

ARTICLE INFO

Keywords:
Low voltage
Smart meter
Load forecasting
Demand forecasting
Substations
Smart grid
Machine learning
Time series
Neural networks
Review
Survey

ABSTRACT

The increased digitalisation and monitoring of the energy system opens up numerous opportunities to decarbonise the energy system. Applications on low voltage, local networks, such as community energy markets and smart storage will facilitate decarbonisation, but they will require advanced control and management. Reliable forecasting will be a necessary component of many of these systems to anticipate key features and uncertainties. Despite this urgent need, there has not yet been an extensive investigation into the current state-of-the-art of low voltage level forecasts, other than at the smart meter level. This paper aims to provide a comprehensive overview of the landscape, current approaches, core applications, challenges and recommendations. Another aim of this paper is to facilitate the continued improvement and advancement in this area. To this end, the paper also surveys some of the most relevant and promising trends. It establishes an open, community-driven list of the known low voltage level open datasets to encourage further research and development.

Energy Data Taskforce

A Strategy for a Modern Digitalised Energy System

- In October 2018 the Energy Data Taskforce was established to provide Government, Ofgem and Industry with a set of recommendations on how data can assist with unlocking the opportunities provided by a modern, decarbonised and decentralised Energy System at the best value to consumers.

- In June 2019 the Energy Data Taskforce published a report entitled

  A Strategy for a Modern Digitalised Energy System

  which presents five key recommendations that will modernise the UK energy system and drive it towards a net zero carbon future through an integrated data and digital strategy throughout the sector.

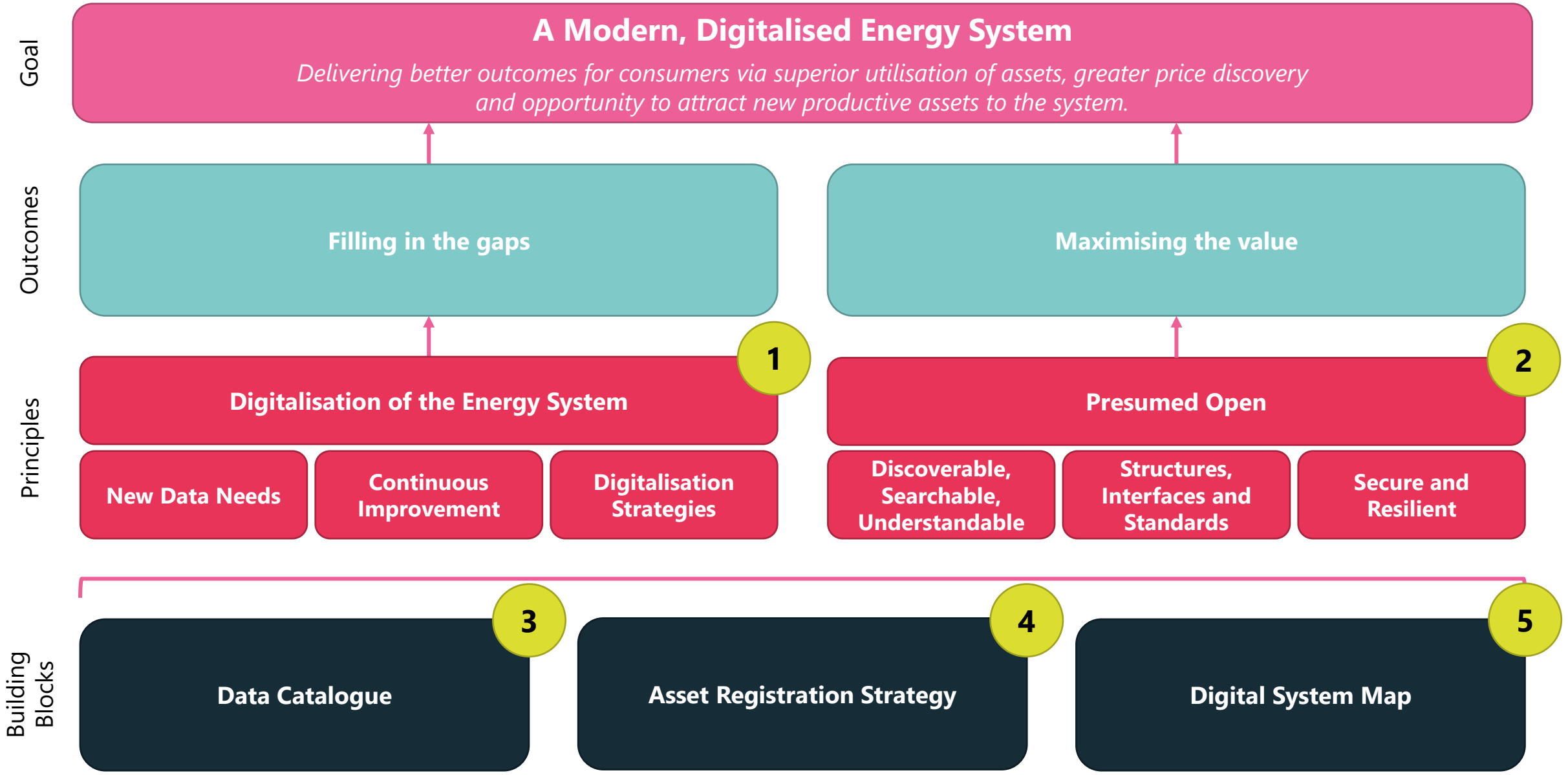
Commissioned by:

ofgem Making a positive difference for energy consumers

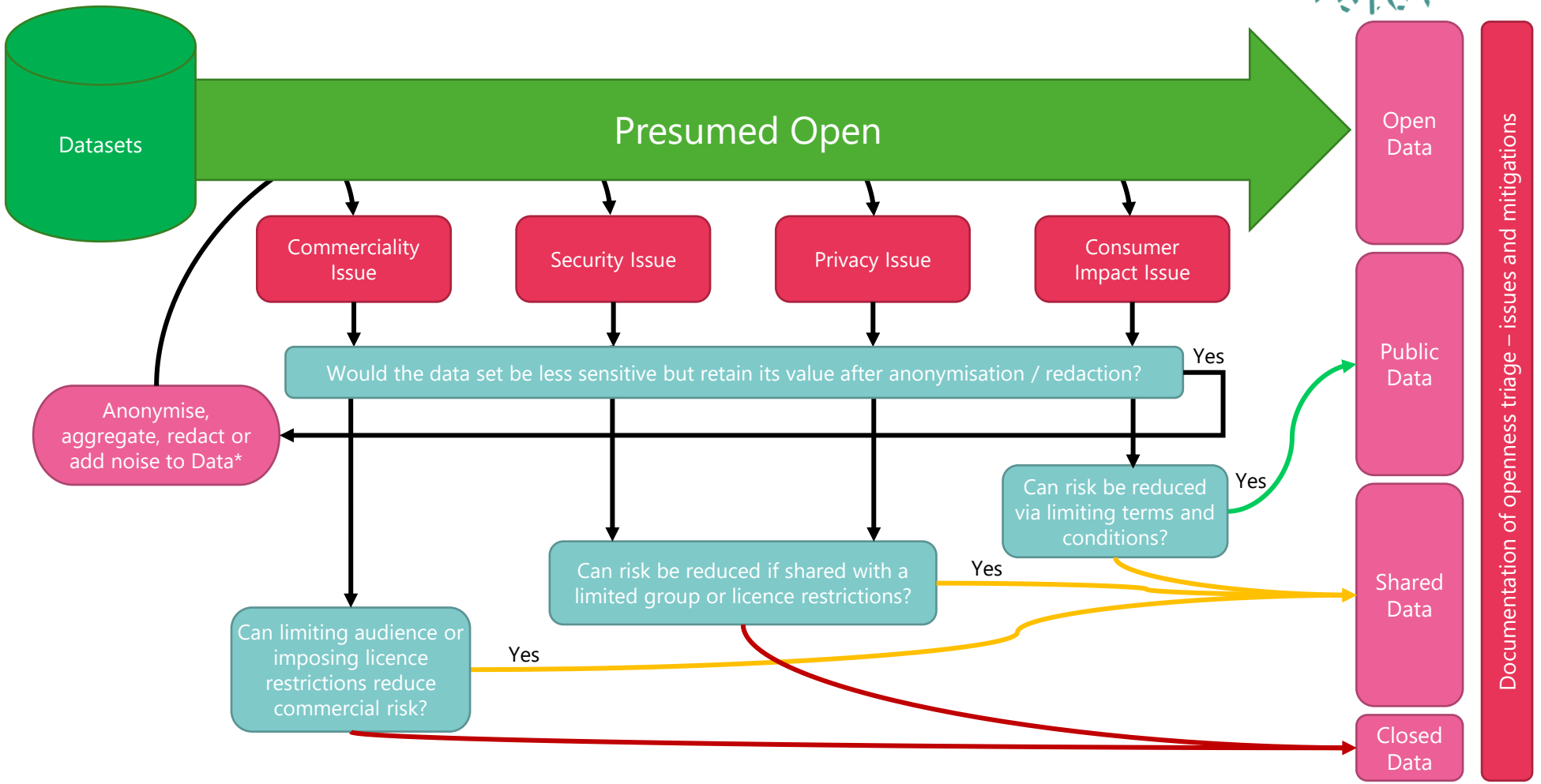
Innovate UK

Department for Business, Energy & Industrial Strategy

Energy Data Task Force



Triaging Openness

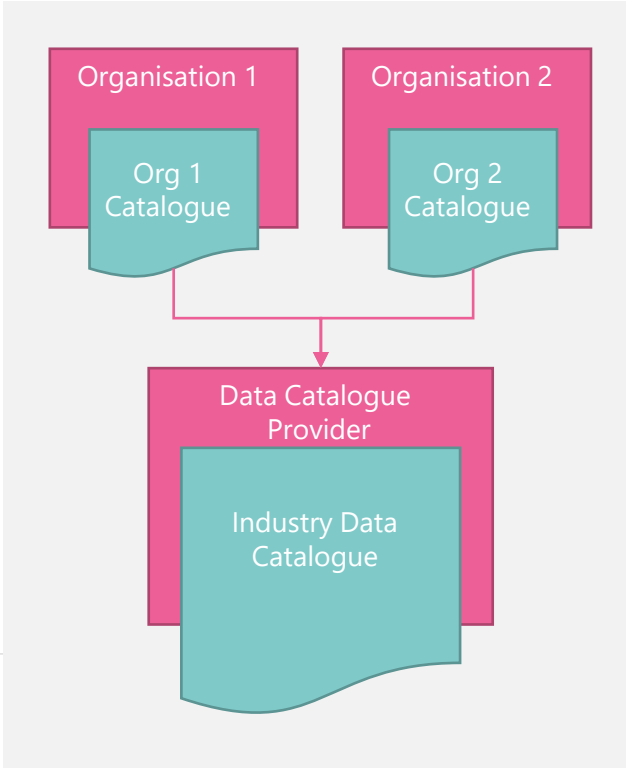


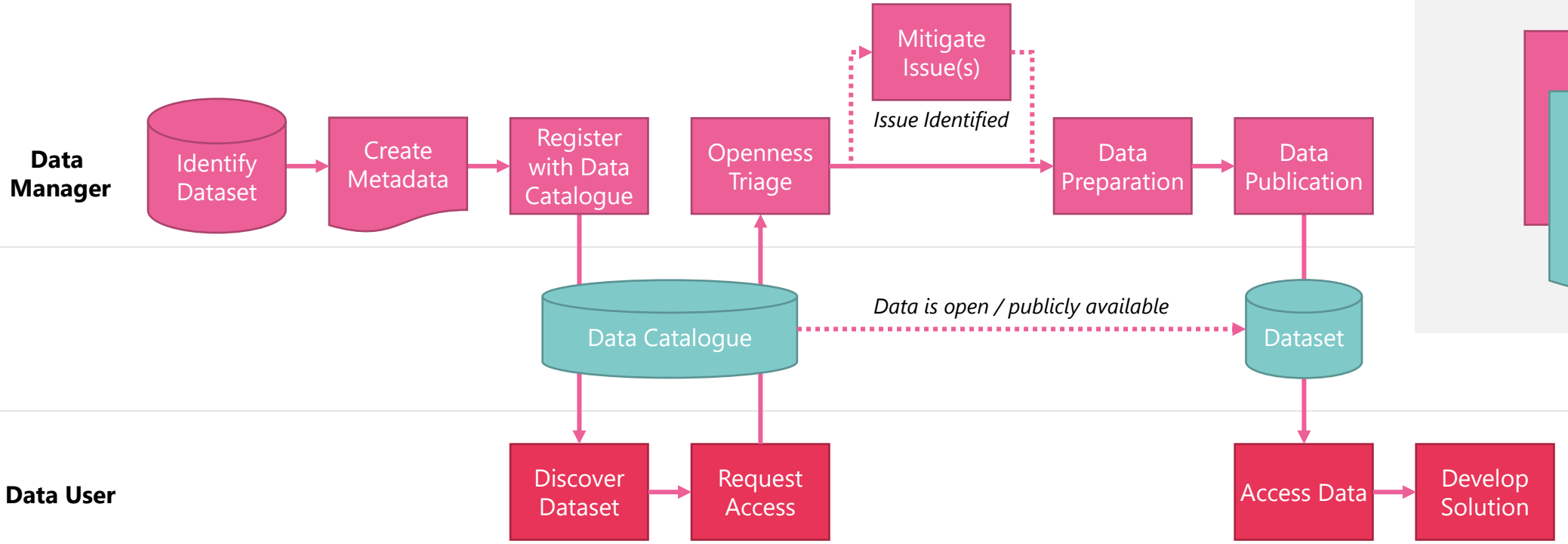
*Multiple stages of anonymisation / redaction may be required to address different issues (e.g. privacy and security) but repeated application should be limited

Recommendation 3:
Data Catalogue

Recommendation 3: Visibility of Data

A Data Catalogue should be established to provide visibility through standardised metadata of Energy System Datasets across Government, the regulator and industry. Government and Ofgem should mandate industry participation though regulatory and policy frameworks.





Sector Engagement and Reaction

"Digitalisation and data are essential to managing the energy system efficiently and securely through the energy transition. We welcome this industry-wide digitalisation strategy and we look forward to closely supporting it's delivery."

Fintan Slye, Director – National Grid Electricity System Operator

"We will help initiate the implementation of the report's recommendations by seeking where necessary, to facilitate a change to the Balancing and Settlement Code this year to further embed the principle of Open Data within the BSC."

Mark Bygraves, Chief Executive - Elexon

"Data is vital to creating a smarter, cleaner and more flexible energy system that will unlock a whole range of new benefits for Britain. This report recognises that energy networks are the nerve centre of that system."

David Smith, CEO - Energy Networks Association

"This report sets out a positive vision for a more transparent and accessible energy system that will benefit market participants and customers alike by enabling decarbonisation and increasing innovation across the system. It is vital we capitalise on the gathering momentum in this area.

Lawrence Slade FEI, Chief Executive – Energy UK

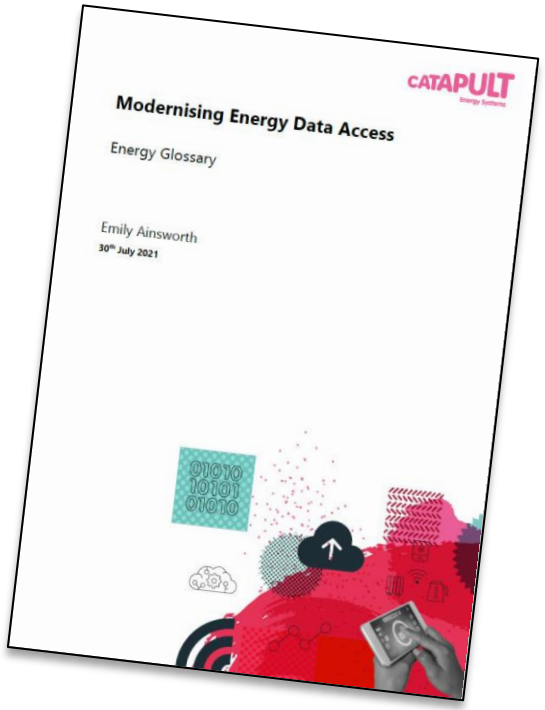
"The EDTF has managed to galvanise usually very disparate parts of the energy sector into agreeing on tangible steps that can be taken now and will be transformative in the long run…"

Sian Jones, CEO - Xoserve

Improving Energy Data Visibility

- Icebreaker one:
  - Open Energy platform, Part of Modernising Energy Data Access (MEDA) program: data sharing ecosystem focused on governance, licensing and authentication.
  - Energy Data Visibility Project, generating a data catalogue and search engine for open data sets. https://openenergy.org.uk/

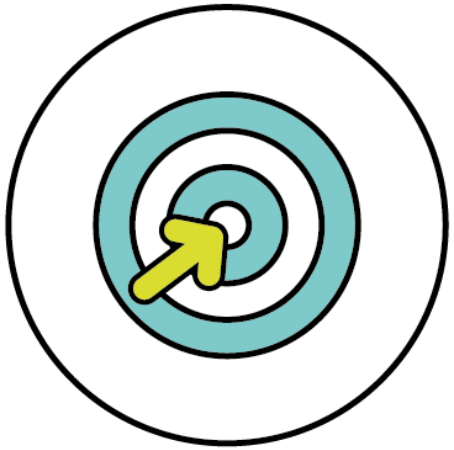
- Energy Systems Catapult through MEDA
  - Energy Data Glossaries
  - Energy Data Use Cases
  - Data Interfaces

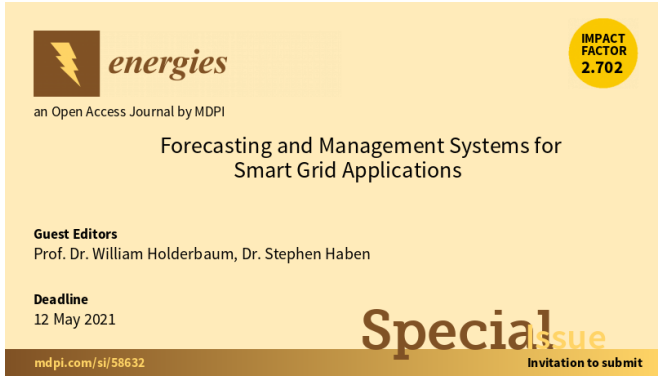


Value of Data:
Data Science Challenges

Challenge Aims

- Presumed Open Data Project – an NIA project led by Western Power Distribution
- Two main aims:
  - Maximise the Visibility of Data.
  - **Maximise the Value of Data.**
- Aim of the data science challenge was to demonstrate the value in increasing data accessibility:
  - What techniques provide the good performance?
  - What datasets demonstrate the most value?
  - What value can be engineered from the data?
  - How much data is needed?
  - What other datasets may be required?
- Challenge assessed over four weekly tasks plus one practice challenge.
- Prizes: Ideas Pitch and Publication in Energies Journal





Challenge Scenario

- A 6MWh/2.5MW battery is connected to a primary distribution substation and a 5MW solar farm in Devon, southwest England.
- Design the control of a storage device to support the distribution network to:
  - Maximise the daily evening peak reduction.
  - Using as much solar photovoltaic energy as possible.
- This will be done for each day for the week following the current challenge date.
- In other words it is ***a constrained optimisation/control problem under uncertainty.***
- There will be **four** assessed weeks as part of this challenge.

Scoring

For each day ($d = 1, \dots, 7$) of the current task week a score is calculate given by:

$$S_d = R_{d,peak}(p_{d,1}C_1 + p_{d,2}C_2)$$

Peak reduction (percentage) during the evening period on day d

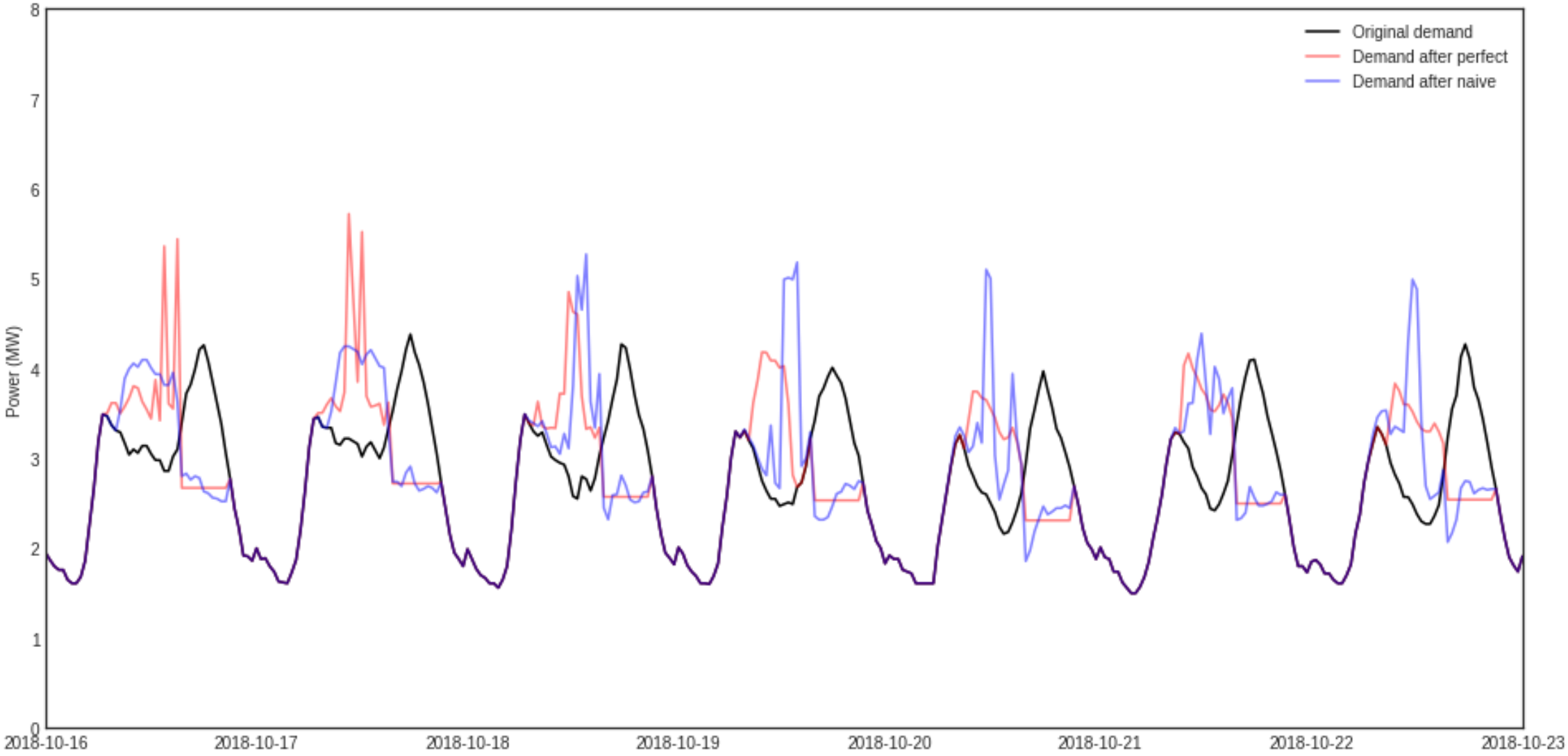
Proportion of energy stored in the battery from solar energy and from the grid on day d.

- $C_1 = 3$, $C_2 = 1$ are weights for the solar and grid energy, respectively. These weights are based on the relatively lifetime GHG emissions intensity of solar and electricity from the grid.

The final score for the current task week, is simply the average over each day of the week:

$$S_{final} = \frac{\sum_{d=1}^{7} S_d}{7}.$$

Benchmarks for Comparison



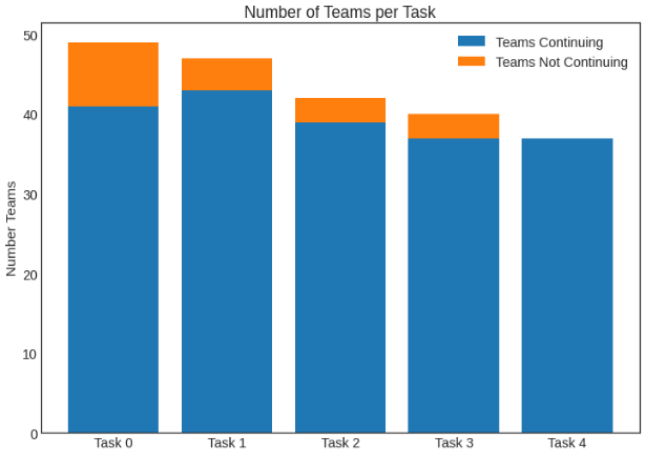
Presumed Open Data Project: Data Science Challenge

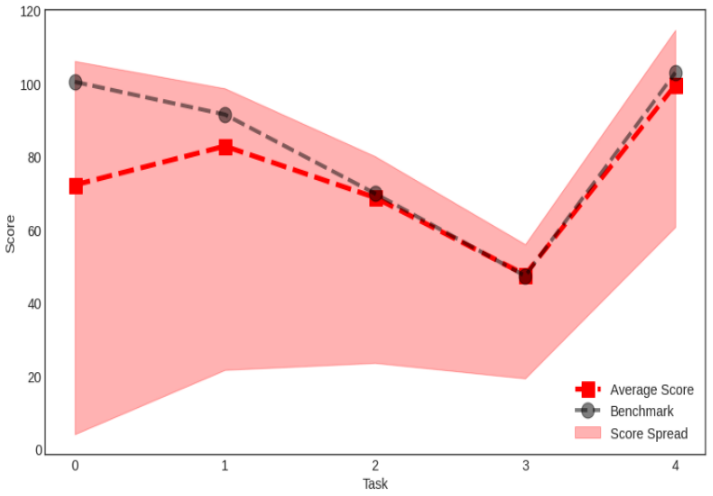
Participation:

- 365 Registrations for the kick-off event.
- 55 teams - a total of 142 individuals – participated in a least one round.
- 15 different countries
- Individuals come from 72 different organisations/institutions (30 universities)

Improvement

- From 26% beating benchmark (Practice Challenge) to 65% (Task 4)





Main Points

- Participation – there is a lot of **interest and value** in making these sorts of dataset available.
- Variety – rare opportunity to see and compare variety of methods, approaches, ways of utilising data on the same problem.
- Source for benchmarks and state of the art.
- Several paths to achieve similar and good results. No “perfect” method or approach.
- New Challenges starting soon! Get in contact if interested!

Final Points

Final Summary

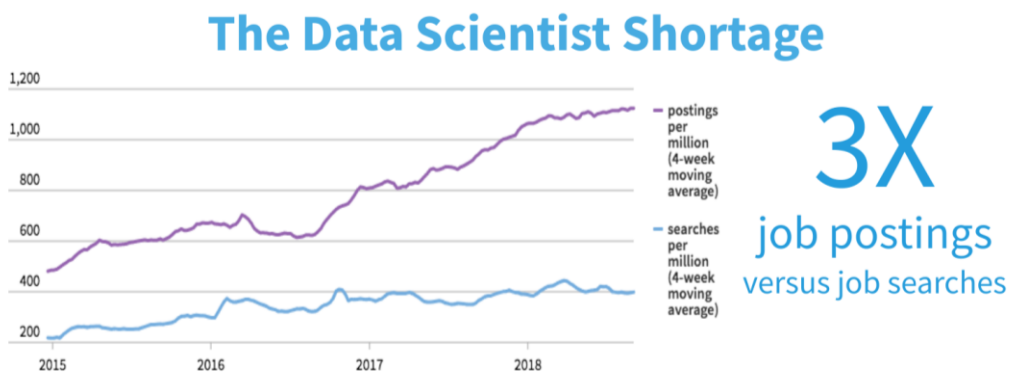
- Data essential to drive innovation. Needs cross sector support and action.
- Investigate methods for privacy protection in LV level datasets.
- Validate research against multiple (hopefully open) datasets.
- Compare against existing methodologies in the data.
- Move towards probabilistic methods.
- For applications investigate how important the forecast accuracy is for performance.

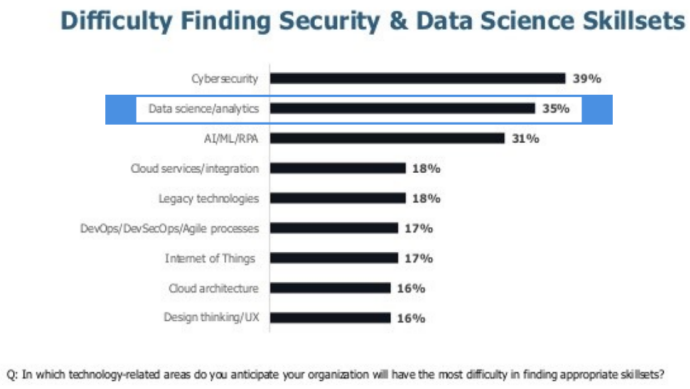
Data Science Skills

Some challenges for data science and forecasting

"The labour market and skills gaps – "in the EU, there are 6 million data workers who collect, store, manage or analyse data as their primary activity, representing about 3% of the workforce. " International Data Corporation.

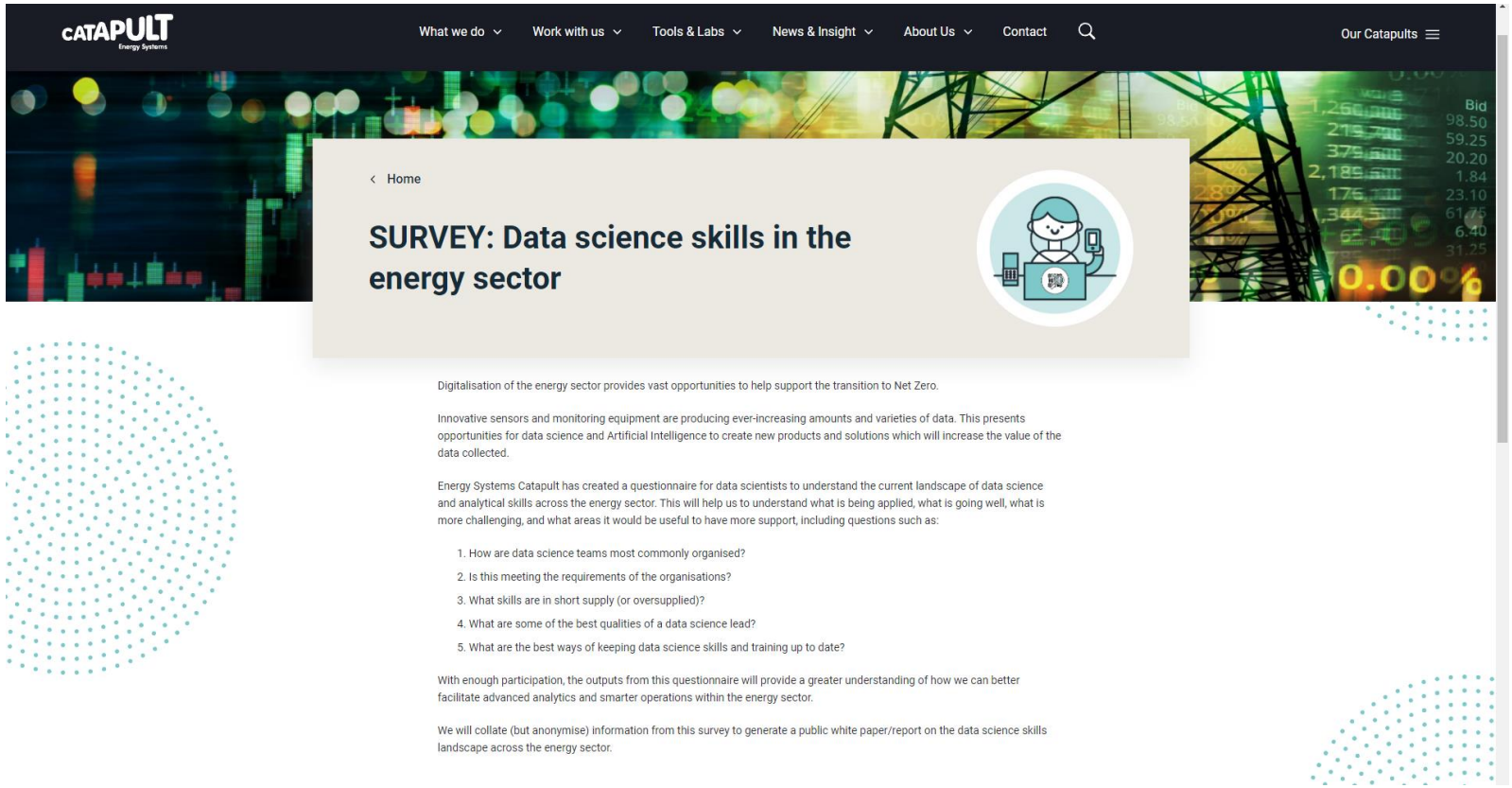
- "**Two-thirds of datavores who tried to recruit analysts in the previous 12 months struggled to fill at least one vacancy**." Analytic Britain, NESTA, 2015.





Pictures: https://quanthub.com/data-scientist-shortage-2020/

Please Fill in Our Survey



https://es.catapult.org.uk/news/survey-data-science-skills-in-the-energy-sector/

Thank you!

CATAPULT
Energy Systems

Stephen Haben

Stephen.haben@es.catapult.org.uk