WESTERN POWER DISTRIBUTION

INNOVATION

NEXT GENERATION NETWORKS

LCT Detection Project

Webinar

1 May 2019

Welcome and introductions

Ricky Duke, Western Power Distribution

Agenda – LCT Detection project

- Fast facts and the LV network challenge
- The LCT Detection project
- Key business insights, lessons learned & quick wins
- Into business as usual
- Q&A

Fast facts and the LV network challenge

Gill Nowell, ElectraLink

Fast Facts

- Funded through the Network Innovation Allowance - £311k
- Hosted by WPD, delivered by ElectraLink and supported by IBM
- Project duration: November 2018 - March 2019
- Deliverables: Proof of concept models to identify unregistered electric vehicles, solar (other LCTs in the future), on its Low Voltage networks, to support network planning decisions for management of LCT uptake

The LCT challenge for LV networks

Images courtesy of Electric Nation and @EVclicks

The Data-based Solution

The project has established a mechanism for identifying EVs and other LCTs on the network:

- ElectraLink extracted data sent across the DTS regarding consumption and export relating to WPD's network.
- This data was analysed by IBM's cognitive analytics and where appropriate combined with third party datasets, to identify previously unknown LCTs on the network.
- This improved WPD information can be used to support a reporting framework to enable WPD to forecast future requirements for network monitoring and potential sites for active network management.

Image courtesy @EVclicks

The LCT detection project method and modelling

Max Hudson - IBM

Summary of our approach

Aim: Understand whether we can detect the presence of unknown low carbon technologies within WPD's network using the DTS dataset, through machine learning

The Crown Dataset

WPD Asset Records including known LCT installations

The DTS Dataset

6 years' worth of data sent across the DTN: Settlement; Switching and Metering.
ElectraLink has governance to provide this data to DNOs

The Machine Learning Capabilities

IBM Watson Studio: machine learning capabilities to analyse DTS and Crown Datasets

Summary of our approach

Machine learning uses algorithms to “learn” information directly from data without relying on a predetermined equations – approaching data analytics through examples, not instructions

Machine Learning found meaningful information from the Crown and DTS dataset and applied the learnings across new datasets for prediction, description and compression…

Key messages

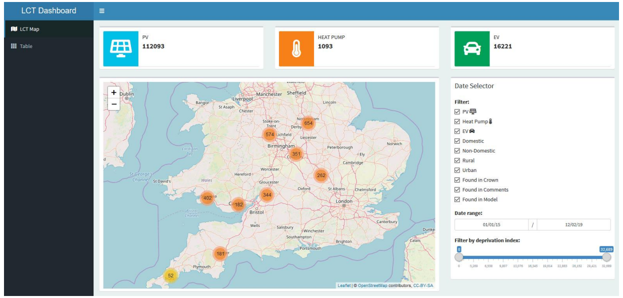
- The LCT Detection project has successfully developed Proof of Concept models that can identify thousands of likely low carbon technologies connected to the LV network which were previously unknown / unregistered
- On both structured and unstructured data…

This is a first!

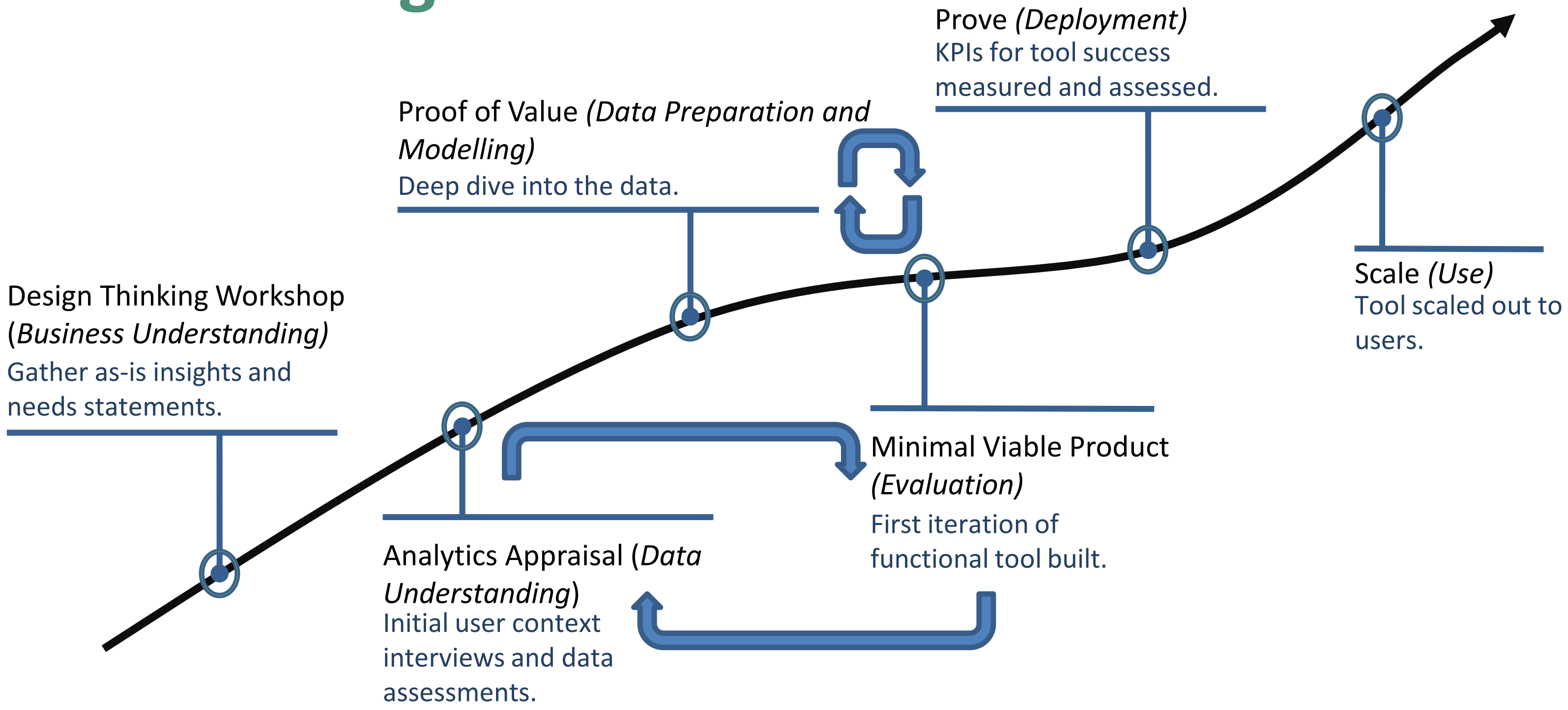
- The project has generated fantastic learning that could inform a second phase
  - The existing models are built on consumption data only
  - Other data, such as socio-economic data, will greatly enhance the model
  - As will use of a negative data set and more granular data
  - Opportunity to investigate other use cases

Data Tools

Visualisation Tools



CRISP DM & Agile



Hypotheses

WESTERN POWER DISTRIBUTION
INNOVATION

Hypotheses

Unstructured comment data from engineers on ground contains evidence of LCT proliferation

Keyword Search &
Natural Language Models

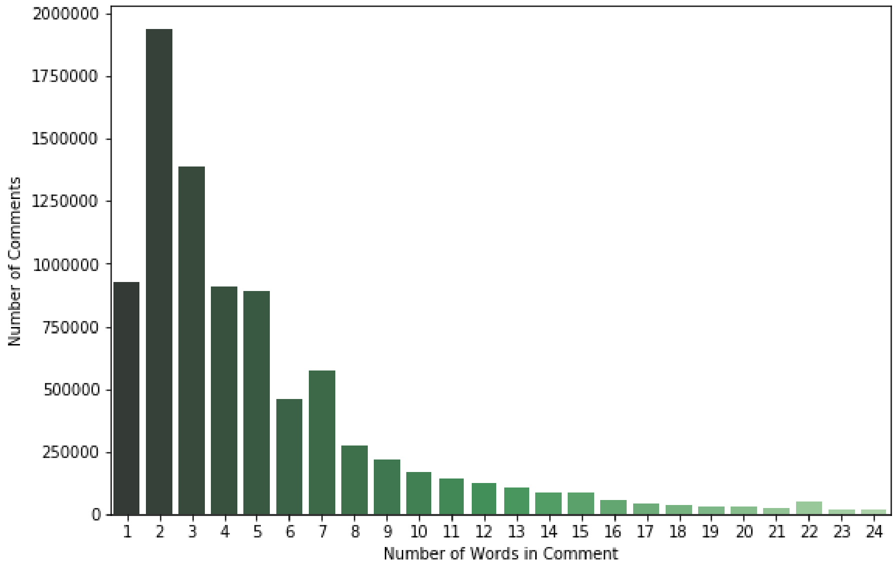
Changes in consumption behaviour indicate when MPANs have LCTs installed

Consumption Models to Identify
EV & PV

Unstructured data models

Unstructured Data Models – Initial Analysis

- The DTS dataset includes comments made by engineers on call-outs written to help engineers on future visits
- The comments are:
  - Usually short sentences with a few words
  - Often repeated



Comment	# of times present	% of all comments
no access	387,518	4.38%
<blank>	338,927	3.83%
k	320,652	3.63%
standard service standard service levels	294,845	3.33%
white door	156,424	1.77%
nrq	142,644	1.61%
no answer	141,066	1.60%
unable to obtain a remote meter reading	117,175	1.33%
no more info	103,716	1.17%
brown door	49,481	0.56%

Unstructured Data Models – Keyword Search

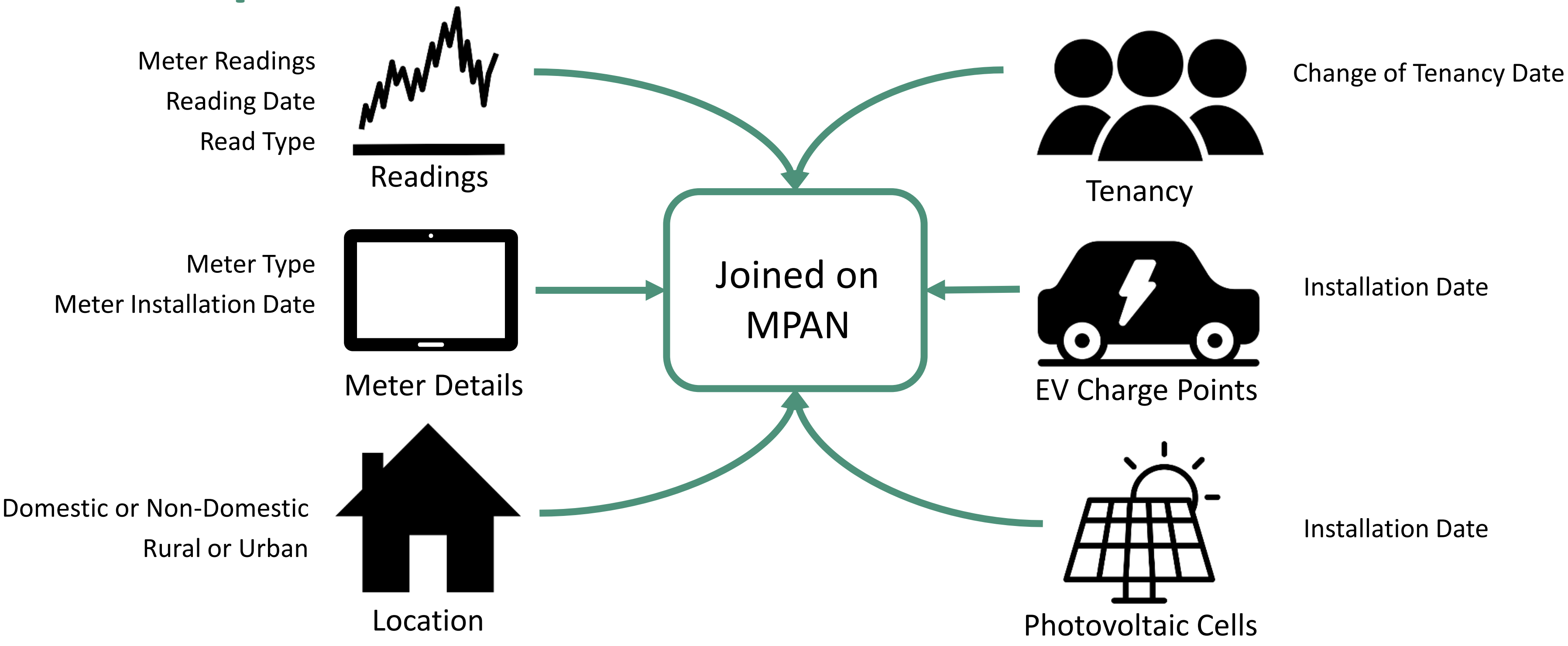
- Keyword searches related to LCT were undertaken to classify any MPANs with comments that contain these words
- Evidence was found of meters running backwards caused by presence of photovoltaic cells
- Most positive matches for the keywords were related to PV. This is likely because of engineers being called out for unusual readings caused by the running backwards phenomenon

Category	Keywords
EV	'charge point', 'electric vehicle', 'charging point', 'hybrid car', 'range extender', 'tesla', 'model x', 'model s', 'i3 ', 'i8 ', 'outlander', 'car plugged in'
PV	'voltaic', 'solar', 'pv ', ' pv'
Heat Pump	'heat pump', 'ground source'

Category	# of MPANs in tagged by keyword search	# of MPANs tagged by keyword search that are not in known dataset
EV	39	37
PV	3,042	1,924
Heat Pumps	4	4

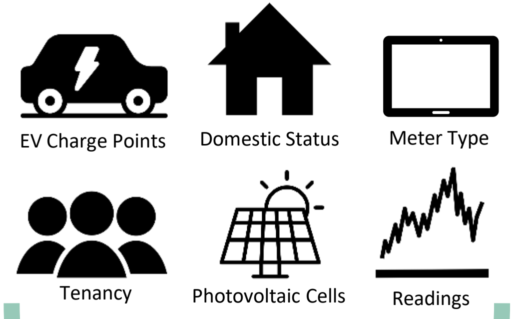
Consumption Data

Consumption Data



Data Processing

MPAN	Reading	Read Date	Meter Type	LCT Installed?	Installation Date	Deprivation Index	Rural / Urban
MPAN 1	3628	01/03/2016	S1	PV	24/08/2016	312	Rural
MPAN 1	3927	10/04/2016	S1	PV	24/08/2016	312	Rural
MPAN 1	5109	16/06/2016	S1	PV	24/08/2016	312	Rural



Remove:

- Anomalous spikes and drops in readings
- Duplicated readings
- Readings on same day

Data Processing

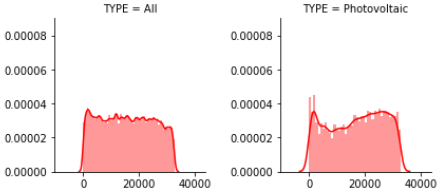
MPAN	Reading	Read Date	Meter Type	LCT Installed?	Installation Date	Deprivation Index	Rural / Urban
MPAN 1	3628	01/03/2016	S1	PV	24/08/2016	312	Rural
MPAN 1	3927	10/04/2016	S1	PV	24/08/2016	312	Rural
MPAN 1	5109	16/06/2016	S1	PV	24/08/2016	312	Rural

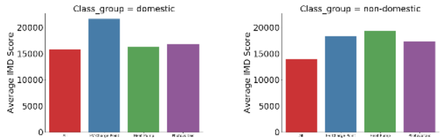
Feature Engineering

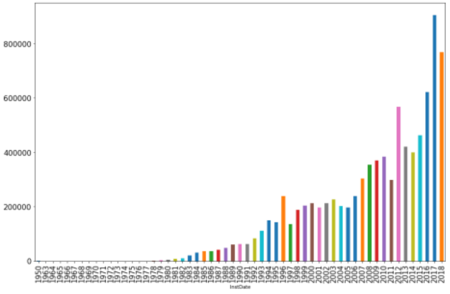
- Change since last reading
- Days between readings
- Average daily energy usage
- Change since same period last year
- Was LCT installed when the reading was taken?
- Has LCT been installed in last 12 months?
- Tenancy end dates

Analyses

Analyses - Population-level analysis





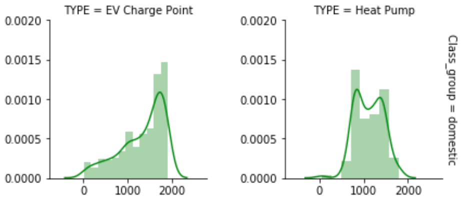


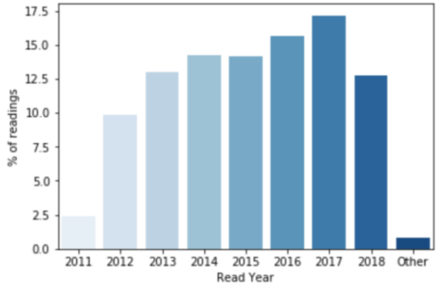
Exploring the data through analysis and visualisations helps us to understand the data, enabling us to begin to build our approach to modelling.

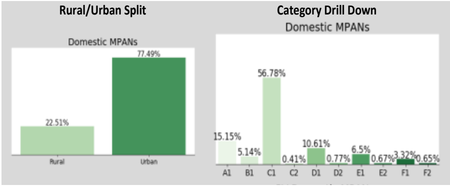
It also allows us to test some hypotheses:

EV is more likely in more affluent areas

LCT unlikely in properties with pre-payment meters



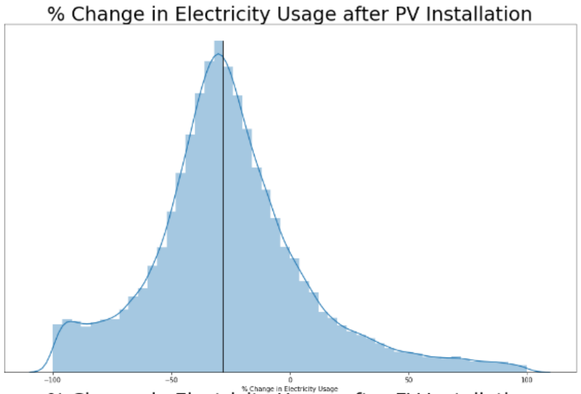




Analyses - Population-level consumption analysis

- When looking at multiple MPANs with LCT, the noise of individual MPANs’ consumption is averaged out and we can see behavioural changes once LCT has been installed.
- In the majority of cases for PV, there is a reduction in energy consumption once PV has been installed.
- In the majority of cases for EV, there is an increase in energy consumption once an EV Charge Point has been installed.

The results indicate that there are sufficient behavioural changes related to the installation of LCT that can be modelled using machine learning



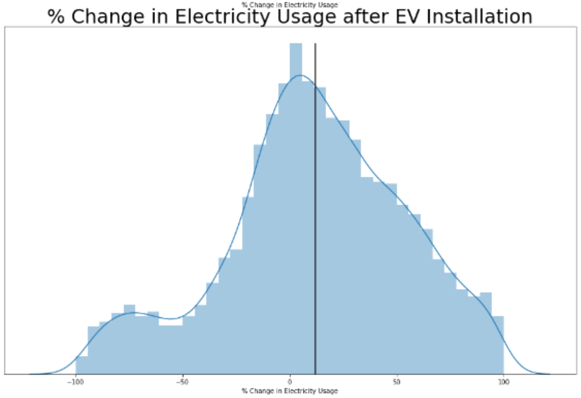
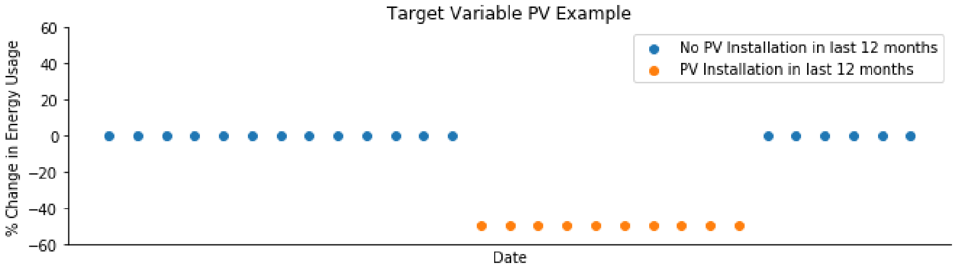


Image courtesy @EVclicks

Consumption models

Consumption Model – Target Variable

- Lack of a negative dataset (i.e. MPANs known not to have LCT) means:
    - Demographic data cannot be used
    - Cannot model on an MPAN level
- Consumption is therefore the input for the model
    - Due to different meter reading frequencies average daily consumption was calculated, and due different levels of consumption, change in average daily consumption for a given MPAN (controlled for meter, register and tenant) was calculated
    - To account for seasonal effects, change in average daily consumption compared to the same period the previous year was calculated for each MPAN
- Therefore, the target variable the model is trying to predict are the readings in the 12 months after LCT installation



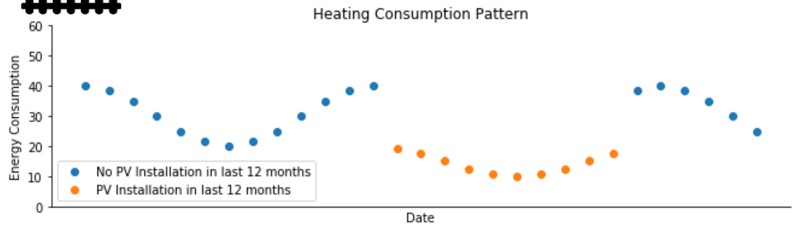
Initial Approach

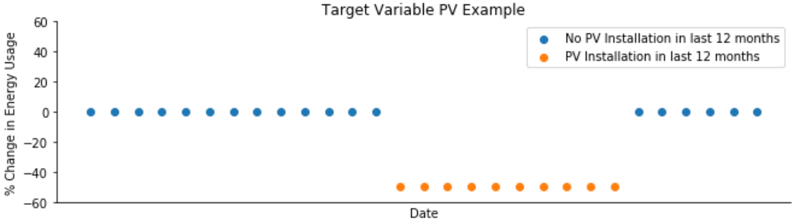
- The initial approach took all of the MPANs with the known LCT and built a model to predict LCT, however this was unsuccessful
- This was because:
    1. Differences in the frequency of readings for different meter types, as well as difference between domestic and non-domestic. Therefore, the model was restricted to domestic smart meters.

MType_group	avg(DaysSinceLastReading)
Smart	37.02326341921678
Others	35.18899326711983
Prepayment	63.17931551869645
Non-Smart	98.52526720696177

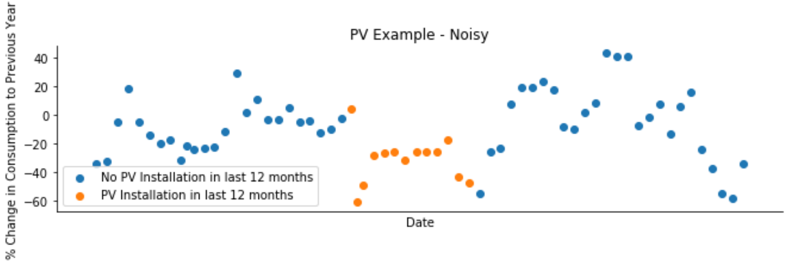
    2. A portion of the MPANs have a lot of noise from other unknown fluctuations in consumption

Seasonal Heating Energy Usage





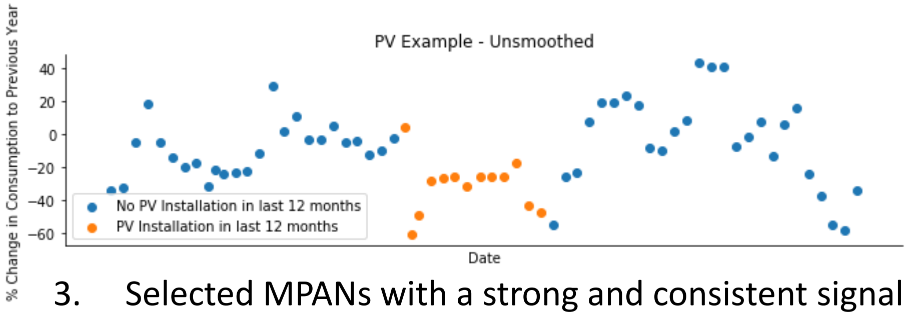
Combined Family Energy Usage

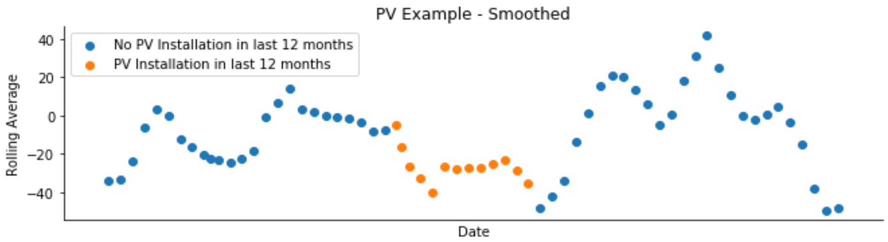


Developing New Approach

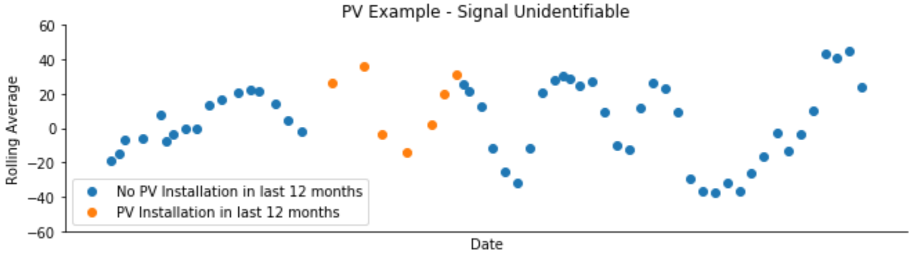
1. MPANs in use were filtered to include only domestic smart meters
2. Changed input to rolling average of consumption change compared to the same period last year, to smooth out the data

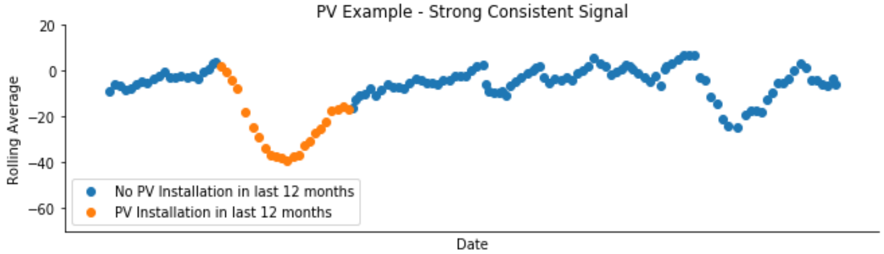
$$rolling\ average = \sum_{i=0}^{i=90}(\Delta\ Consumption) \Big/ number\ of\ readings$$





3. Selected MPANs with a strong and consistent signal from the known LCTs to build the model on





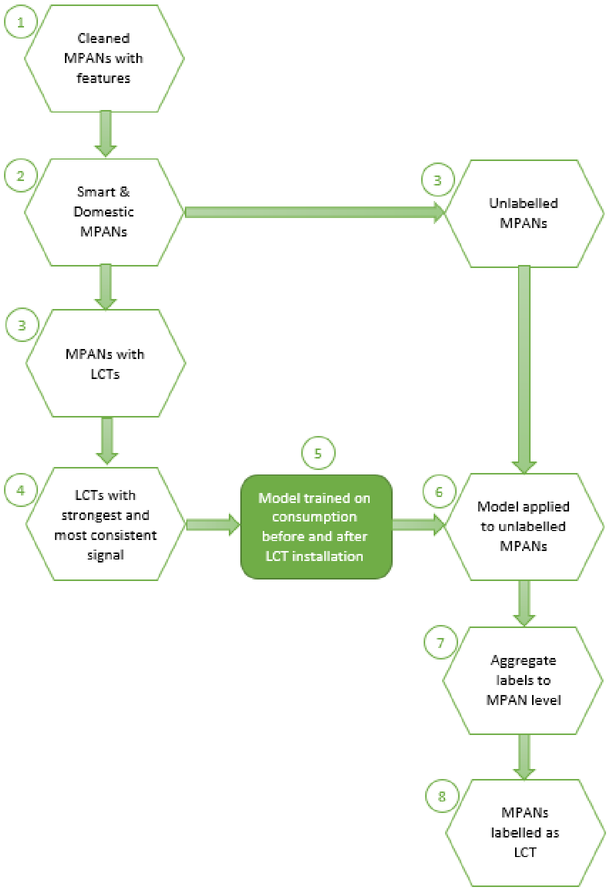
Final Approach

1. Selecting MPANs for modelling
   Domestic MPANs with smart meters known to have LCT are ranked by:
   1. Change in average daily consumption in the year before installing LCT compared to the year after
   2. Standard deviation of change in average daily consumption

   Once these MPANs are ranked, a proportion of the MPANs with the largest change in consumption and the lowest standard deviation are used to train the model.
2. Selecting MPANs as predicted to have LCT
   Two metrics were created to measure whether the aggregated predictions indicate whether an MPAN has LCT:
   1. Number of predicted periods of LCT / average number of readings per year
   2. Timespan between the first and last predicted period

   MPANs meeting the following criteria are deemed to indicate LCT:

Metric	Thresholds
Predicted Readings / Average No. Readings per Year	0.9 < x < 1.1
Timespan between first and last predicted period (days)	250 < x < 400



Results

MPANs meeting the criteria in the final approach are deemed to show similar enough consumption behaviour to the MPANs with known LCT and a strong consistent signal to indicate they may also have LCT present.

The number of MPANs identified for EV and PV are:

LCT Type	No. MPANs Predicted to have LCT
EV Charge Point	5,863
Photovoltaic	8,104

Lessons learned, key business insights & quick wins

Gill Nowell, ElectraLink – Max Hudson, IBM

Lessons learned – understanding WPD's business needs

Design Thinking Workshop

- At project start to provide a shared understanding of the end-users of the POC model and current processes relating to LCT proliferation.
- Captured business-led hypotheses around LCT proliferation and identified data to identify them.
- Key WPD personnel invited across innovation, network planning and policy business areas.

Business Values Report

- An additional project deliverable
- Supplementary to Sprint reports
- Ensured continuing alignment with WPD business needs

Recommendation for future innovation projects: adopt a workshop approach at start of project and establish a Business Values Report deliverable mid-project to ensure project is meeting WPD requirements.

Lessons learned – identification of LCTs in freeform text

Freeform (unstructured) text

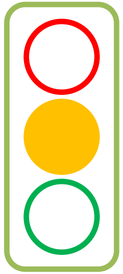
- Using machine learning, the project has identified instances of EV and PV from unstructured data i.e. the forms that field engineers fill out when going on site visits
- This is a valuable first!

Recommendation: it would add value to the data if engineers going out to properties recorded 'EV', 'PV, 'heat pumps' as a matter of course, in all cases, not just where issues are encountered

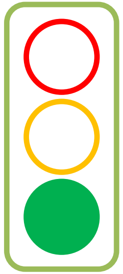
Interdependency of the data: generation of a negative dataset precluded us using other data

Did we learn what we hoped to?

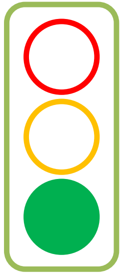
To what degree can a model accurately predict the presence of LCTs?



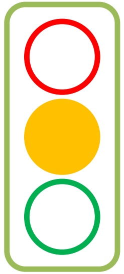
How is the performance of the model affected by the data sets available to it?



What are the options to validate the presence of different LCTs identified by the model?



What steps are needed to scale the model across the industry?



Transition to Business as Usual

Ricky Duke, Western Power Distribution

What we need as a DNO

- Better visibility of Low Carbon Technology connected to our network.

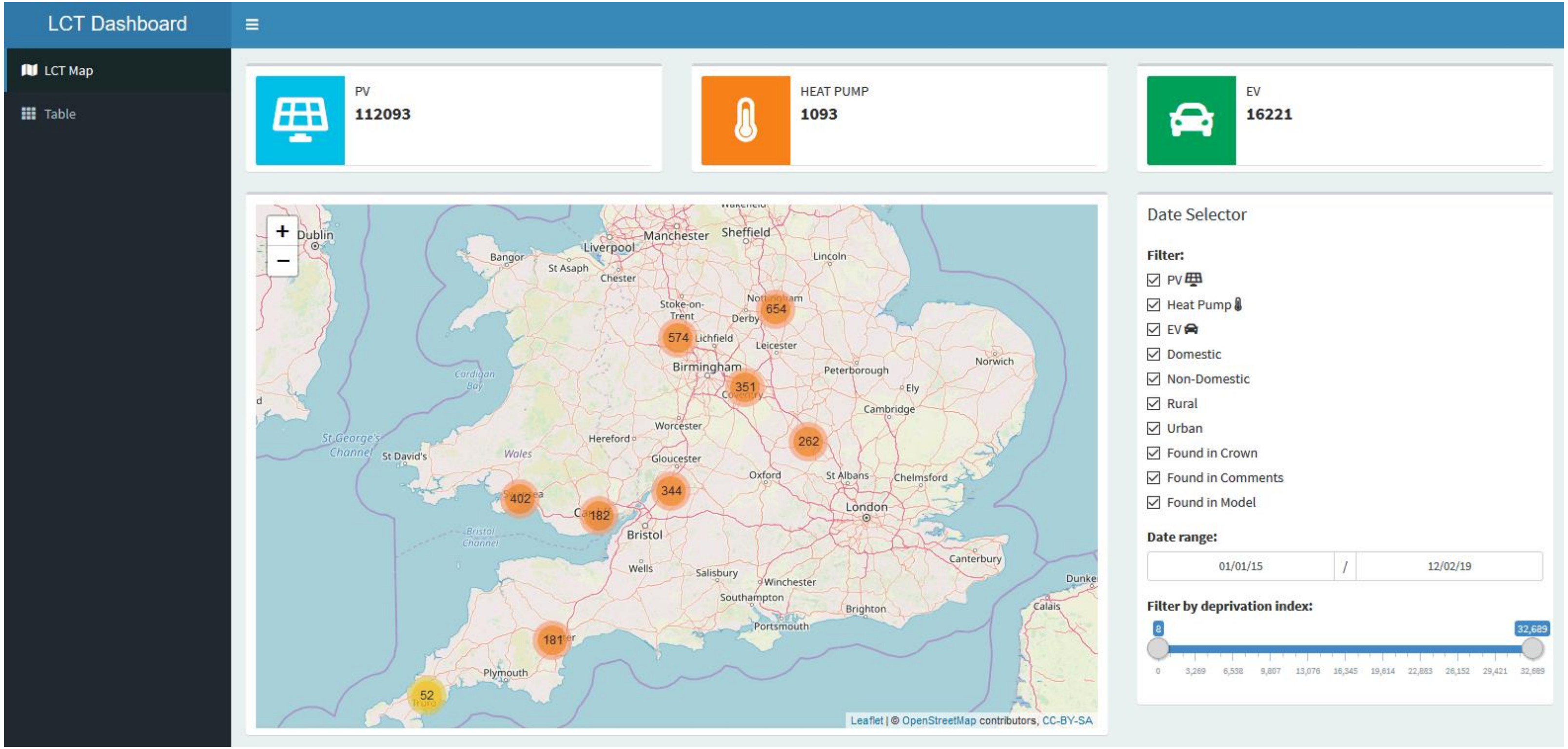
- This will enable network planners to plan the network more efficiently and allow for the effect of LCT on the network.

- Enable us to forecast better into the future, the more we know about our network, the more accurate we can forecast re-enforcement costs.

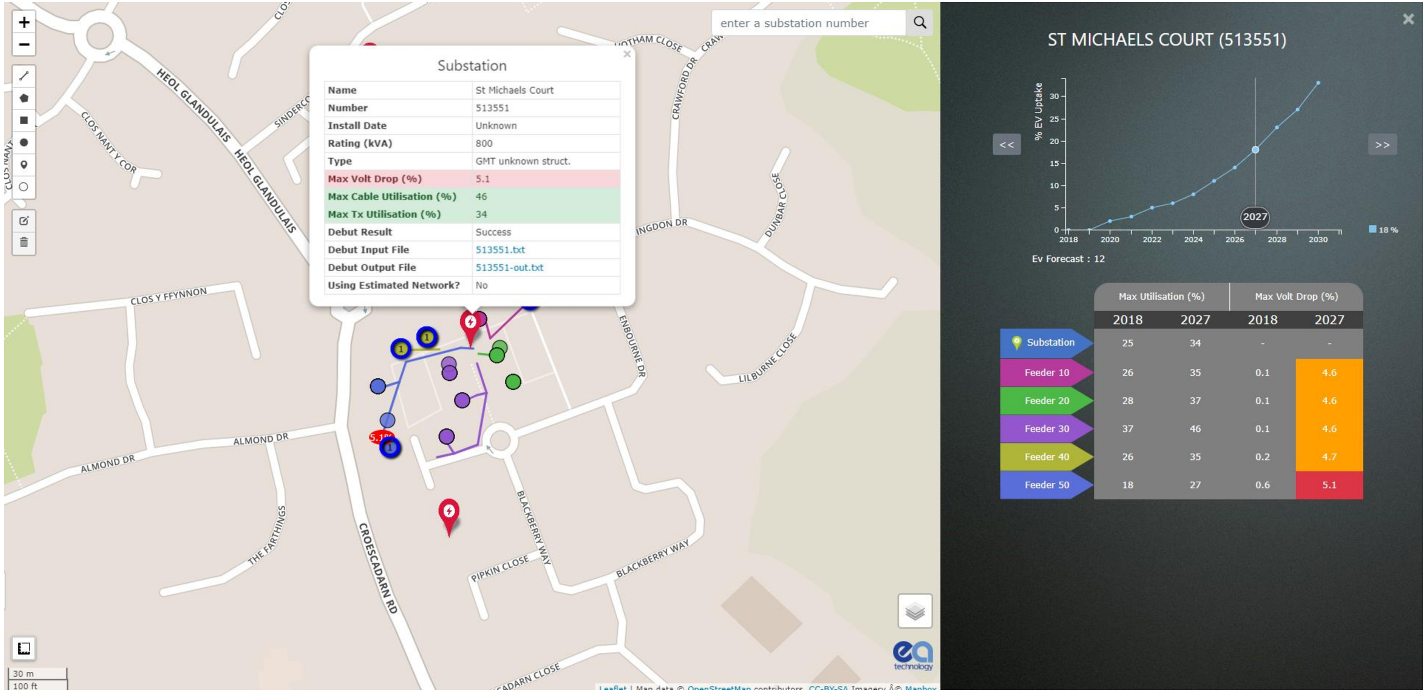
Integrating project outputs into BAU

- Standalone Dashboard/tool for network planners.
- Integration into a Network Assessment Type tool.
- Buying a Service i.e. 6 monthly report.

Standalone Dashboard



Network Assessment Tool Integration



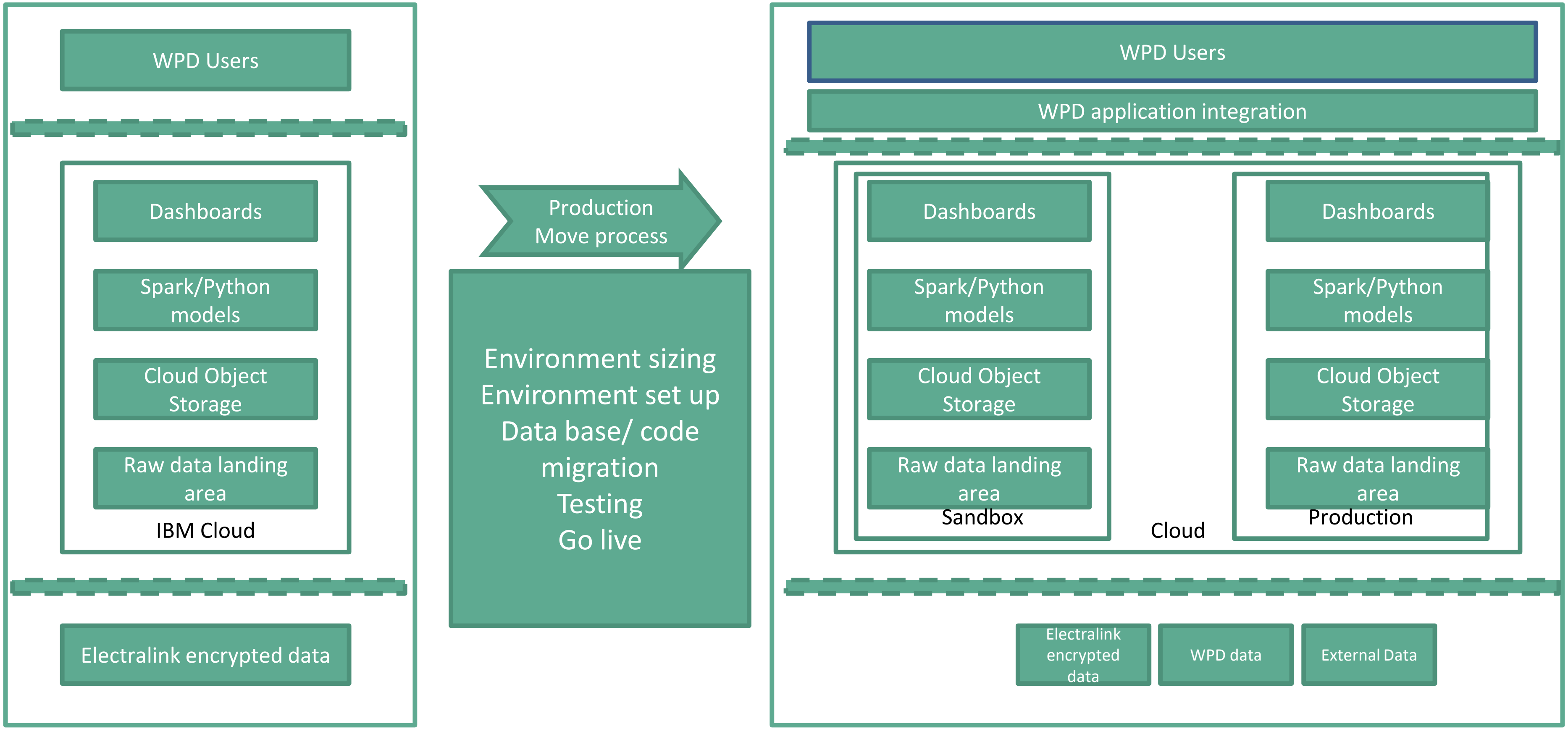
Buying of a service

- Going forward it is essential that WPD has a clear view of LCT hotspots or ‘Clusters’, so we can trigger re-enforcement or ANM before a fault occurs.
- Whilst LCT take-up is relatively low, a six monthly sweep of the data will suffice.
- As take-up increases, then this may become more regular, particularly with more accurate smart meter data as take-up increases.
- Localised searches on request, in response to overloads or voltage complaints.

WESTERN POWER DISTRIBUTION INNOVATION

Move to production - Dedicated WPD environment

LCT Detection



Q&A

THANKS FOR LISTENING

WESTERN POWER DISTRIBUTION

Serving the Midlands, South West and Wales

Ricky Duke	rduke@westernpower.co.uk
Max Hudson	max.hudson@uk.ibm.com
Gill Nowell	gill.nowell@electralink.co.uk