Secondary Network Asset Temperature Behaviour Report

A description of analysis carried out within Phase 1 of the Celsius project

Draft Report for Electricity North West

Customer:

Electricity North West

Confidentiality, copyright & reproduction:

Contact:

Olivia Carpenter
Ricardo Energy & Environment
Gemini Building, Harwell, Didcot, OX11 0QR, United Kingdom

t: +44 (0) 1483 547013
e: olivia.carpenter@ricardo.com

Ricardo-AEA Ltd is certificated to ISO9001 and ISO14001

Author:

Olivia Carpenter, Michael Kelly, Mazin Tagelsir

Approved By:

Simon Terry, Sarah Carter

Date:

30 April 2018

Ricardo Energy & Environment reference:

Ref: ED62006- [Status]

Contents

1. Introduction

The Celsius project was awarded funding under Ofgem's 2016 Network Innovation Competition (NIC). It is being led by Electricity North West (ENW). Ricardo Energy & Environment are the technical consultant partners on the project.

The aim of Celsius is to increase the capacity of distribution substations, specifically the thermal capacity of transformers and cables, to support the uptake of low carbon technologies and reducing the need for network reinforcement. The project includes two phases:

- **Phase 1: Monitoring Trial (520 substations)** – Develops a detailed understanding of the operating temperature of assets with the aim of estimating the impact of a range of asset environment factors in asset rating, which is limited by the operating temperature of the transformer.
- **Phase 2: Cooling technology trial (100 substations)** – Identifies and demonstrates technologies that cool substation assets to release additional capacity.

This report describes the analysis and results for Phase 1 of the project, focusing on transformers only. As agreed with ENW, the project programme will consider cables later in 2018.

1.1 Monitoring Trial Data

The first phase of Celsius included 520 distribution substation sites, each of which has been fitted with at least one central communications hub and up to 30 sensors measuring transformer power and asset and ambient temperatures.

Celsius uses KeLVN monitoring equipment, provided by Ash Wireless, to provide substation data. Included in the suite installed in substation is:

- **Single temperature sensors**, which measure a single asset or environment temperature
- **Hex unit**, which can connect up to 6 sensors including current, voltage, and temperature. When voltage and current sensors are connected in the correct configuration, the Hex unit will calculate complex power from the data.
- **Internal transformer temperature monitoring**, which can also be connected through a Hex unit, and measures the temperature of the oil within a transformer.

This temperature and power monitoring equipment communicates half hourly readings (Temperature readings are instantaneous every half hour, while power readings are or the root mean square over a short period) to a central Hub installed within the substation. Data is transmitted from the Hubs, once per day over the mobile network – the data is sent in four groups, each containing 6 hours' worth of data for that substation. The Hub can receive commands for a short period once a day when it connects to the back end.

The KeLVN monitoring equipment was provided by Ash Wireless, and is described in more detail in a series of separate documents.

A mobile application was developed to support the installation of the equipment and recording of key installation and site data. A data management system was developed to receive data remotely from the monitoring equipment, and process, store, and provide access to it. Raw data is analysed and checked for validity and processed. This includes rounding the timestamps to the nearest half-hour in order to support validation. After processing, the data is stored. Importantly, both raw and processed data are retained for diagnostic and troubleshooting purposes. The installation app and data management system were both developed by Ricardo Energy & Environment.

In addition to the data available from the KeLVN equipment, Electricity North West have installed a number of transformers which have integrated fibre optic temperature monitoring throughout, including direct monitoring of the transformer hotspot. This data is stored through an ihost system separately to the Celsius data management system.

Weather data is also collected from freely available data sources.

1.2 Monitoring Trial Analysis Overview

The expected output of the monitoring trial analysis is a detailed understanding of the thermal behaviour of substation assets (in particular LV cables and transformers) under different asset environment conditions. The main focus is on sites that have higher loading compared to their rating. This will help to develop an understanding into how the sites behave when their load approaches their rating which is the most relevant operating conditions to predict and calculate ratings. The utilisation range for study was determined within the analysis. The site selection for the monitoring trial was carefully considered to include a representative mix of substation sites, taking into account the substation building structure, transformer rating, and transformer specification.

The methodology that has been used in the Monitoring Trial Analysis is as follows:

1. **Transformer Hotspot Estimation** – The majority of transformer temperature monitoring is placed on the surface of the transformer. However, the operating temperature which is most of interest to this analysis is the Hotspot temperature; the warmest point within the transformer. It was therefore necessary to start the analysis by estimating the temperature of this hotspot from externally measurable temperatures. **This analysis is reported in Section 2 of this report.**
2. **Temperature Factors Study** – A significant number of asset environment and loading factors were assessed to understand the impact on the operating temperature of the transformer. This detailed analysis investigated the relative influence and interaction between factors, as well as their impact on the temperature. **This analysis is reported in Section 3 of this report.**
3. **Thermal Rating Influences Analysis** – The conclusions of this analysis include the most influential factors, and a description of how they influence thermal behaviour. This is then used to draw conclusions about the influence of the asset environment factors on thermal rating. **This analysis is reported in Section 4 of this report.**

This document reflects the project findings up to April 2018. These findings will be built on and developed over the rest of the project, and therefore may evolve over time.

2. Transformer Hotspot Study

The transformer '*hotspot'* is the location of the hottest part of an operational transformer, and is located within the transformer itself, towards the centre. It is this temperature that is of interest within the Celsius analysis, but it is not possible to have retrofitted monitoring installed to measure the temperature directly. The aim of the Transformer Hotspot Study was to develop a method to estimate the transformer hotspot temperature from surface temperature measurements and other easily accessible information, for all ground mounted transformers within the Celsius trials. This calculated hotspot temperature was then used within the temperature factors analysis to analyse the impact of asset environment and loading factors on operating temperature.

The success criteria for this work has been determined by Ricardo Energy & Environment is for the model to predict measured data to within + or - 5°C. The approach taken to determine this was to derive the best correlation coefficient between multiple variables. The correlation coefficient R^2, which has a value between 0 and 1, indicates how good a model is for predicting the desired output term. As an indicator, an R^2 of 0.75 or 75% or above is considered to be the criteria for strong correlation which typically produces an error within + or – 5 °C.

This section describes the methodology, analysis and results for this study. There are several key limitations to the transformer hotspot study and its results, which are described in Section 2.4. The methodology and results described here have been updated since the previous report; 'Understanding and Calculating Maximum Operating Temperature of LV Cables and Transformers' delivered in December 2017.

2.1 Hotspot Methodology

The data available for the Transformer hotspot study was as follows:

- **Surface and ambient temperatures** – Over 500 substations were fitted with sensors that recorded transformer surface and ambient temperatures. The site selection included substations across a range of location and building types, specifications, mass, rating and age profiles.
- **Internal oil temperatures** – Out of the 500 transformers described above, 17 also had retrofitted internal top and bottom temperature sensors installed. The selection of these transformers aimed to include a representative sample of the GB ground mounted LV transformer portfolio, including a range of ages, specifications, and sizes. Due to safety and security reasons, all the transformers selected were indoor.
- **Hotspot temperatures** – 5 'smart' transformers which have integrated fibre optic temperature monitoring throughout, including direct monitoring of the transformer hotspot. These substations also have the oil, surface and ambient temperature sensors described above. These transformers are modern and of the same specification.

The approach to this study was as follows:

- **Step 1: Link between hotspot temperature and internal oil temperatures** – Step 1 involved investigating the relationship between the measured hotspot and internal oil temperatures of the smart transformers. Only those that were loaded above 25% average utilisation were considered in this analysis. Other data such as ambient temperature, transformer specification, rating, and mass was investigated to determine whether this improved the performance of the model. A generic model was developed for one of the smart transformers and will be validated with the other smart transformers to determine whether the model can be applied over a range of sites.
- **Step 2: Link between internal oil temperatures and external surface temperatures** – Step 2 involved using the internal oil measurements from the 17 transformer sites to establish a relationship between the internal oil temperatures and their corresponding external surface temperatures, installed as close as possible to the internal sensor positions. As in Step 1, other information was considered for inclusion in this method. A generic model was developed using 75% of the available data for each site, and the other 25% was used for validation.
- **Step 3: Hotspot temperature calculated from external surface temperature –** The results from step 1 and 2 were combined to develop an expression which can estimate the hotspot temperature from external surface temperature and other features.

Due to the volumes of data, and the complexity of the analysis required, a statistical software tool was required to analyse the data and to produce a hotspot estimation model for heavily loaded sites. Excel was not the best tool to use to carry out the analysis as it was not able to manipulate large enough data volumes. The statistical software package called Rstudio was chosen for the work in Celsius because:

- It is capable of accessing data from the Celsius servers;
- It performs data validation and manages data gaps effectively;
- It can manipulate large volumes of data; and
- There are standard statistical models and tools that can be used to perform complex analysis efficiently

2.2 Step 1: Link between hotspot temperature and internal oil temperatures

2.2.1 Analysis of Smart Transformer Data

The smart transformers with adequate loading at peak values of 75% or over were as follows:

Smart Transformer Site	Rating (KVA)	Transformer Specification
Kincardine Rd	1000	ESI 35-1
Clarendon Rd Whalley Range (Transformer 1)	1000	ESI 35-1
Clarendon Rd Whalley Range (Transformer 2)	1000	ESI 35-1

Table 1. Smart Transformers included in the analysis

For this stage, the Kincardine Rd site was selected to build the generic model that calculates hotspot from the internal temperature at the top of the tank while Clarendon Rd, which has two transformers, was used to validate this model. After doing an initial test, it was concluded that there is a direct correlation between the internal temperature at the top and bottom of the tank. This meant that only one internal temperature at either the top or bottom of the tank is needed to predict the hotspot temperature because they are closely related to each other. The internal temperature at the top was selected to be used for the analysis as this is the data point that is most closely related to the measured hotspot.

Due to the thermal mass of the transformer, there is a time delay between the temperature peak at the hotspot and the internal oil temperature to react. The calculation of this time lag is not an important output of the Celsius analysis as the focus is on the shape and magnitude of the temperature of the hotspot temperature. However, it needs to be considered in order to build the model and validate against measured data.

Scatter plots were used to investigate the relationships between two variables (for example, temperature of hotspot and top oil temperature) to predict the best mathematical model to calculate hotspot temperature. This was used to establish whether different variables have a strong influence on the temperature of the hotspot.

Scatter plots of the internal top oil temperature and the corresponding hotspot accounting for different time lags are shown below. The data is for Kincardine Rd, and runs from the start of July 2017 to the end of February 2018.

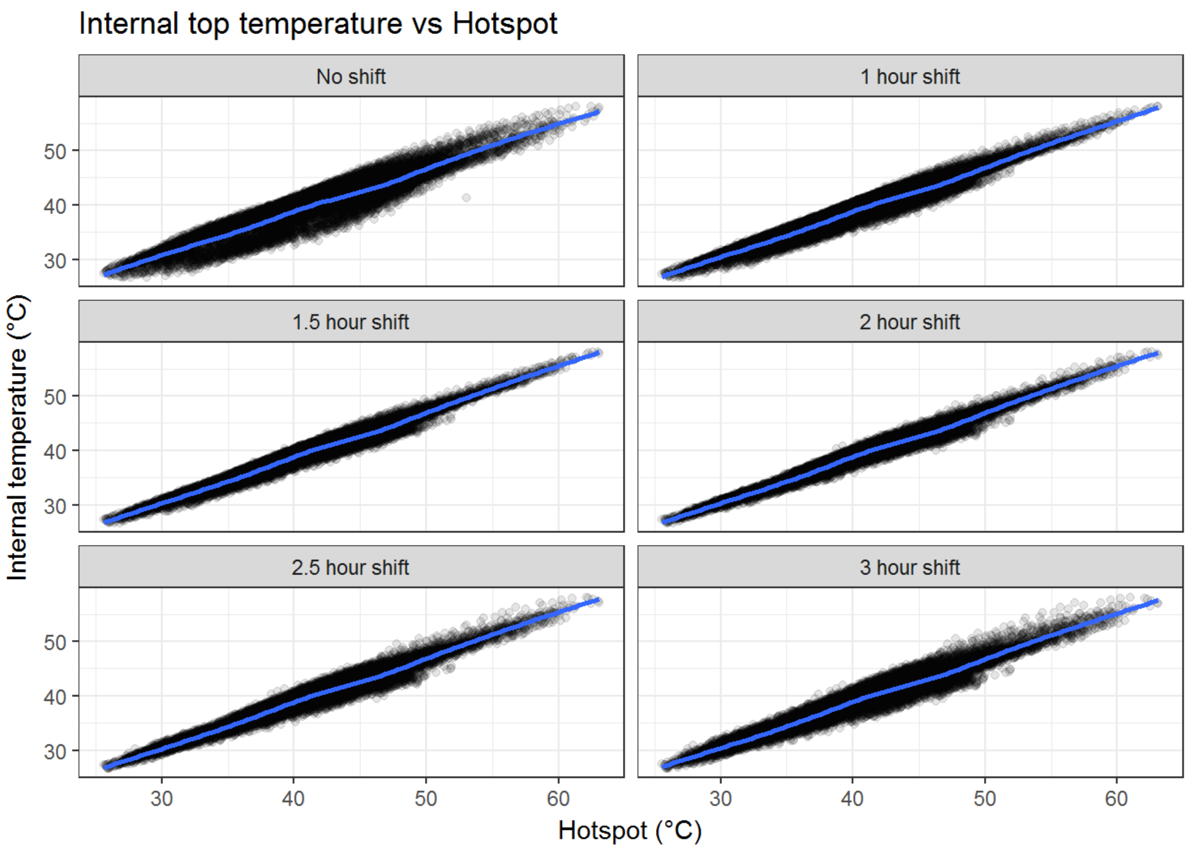

Figure 1. Correlation between hotspot temperature and internal top oil for different time lags

As can be seen from figure 1, there is clear correlation between internal temperature at the top of the tank and the measured hotspot data for all time lags. In each case the internal temperature data at the top of the tank has been time shifted backwards. This helps to bring both peak temperatures in line with each other. This time shifting mechanism for the data was repeated until the highest correlation coefficient was obtained. The table below shows the R^2 calculation for each time shift.

Hotspot vs Internal Top Temperature with a time shift	Correlation Coefficient (R^2)
No time shift	0.952
1 hour time shift	0.952
1.5 hour time shift	0.969
2 hour time shift	0.973
2.5 hour time shift	0.963
3 hour time shift	0.940

Table 2. Correlation Coefficient Results

The results show that shifting the internal data at the top of the tank backwards by 2 hours produces the highest R^2 value, which indicates that the thermal time lag between a peak in hotspot temperature, and the corresponding oil temperature peak is approximately 2 hours.

From this a regression analysis was run in Rstudio to produce an equation that calculated the measured hotspot based on the internal temperature data at the top of the tank which was shifted backwards by 2 hours. The expression was as follows:

$$HS = (1.169 * I_T) - 5.069$$

Where,

HS is the hotspot temperature in °C

I_T is the internal temperature at the top of the tank in °C

The residuals are the difference between the calculated and measured hotspot, and are therefore an indication of the error in the calculated verses the measured values. These were plotted against hotspot temperature (figure 2), and against date (figure 3).

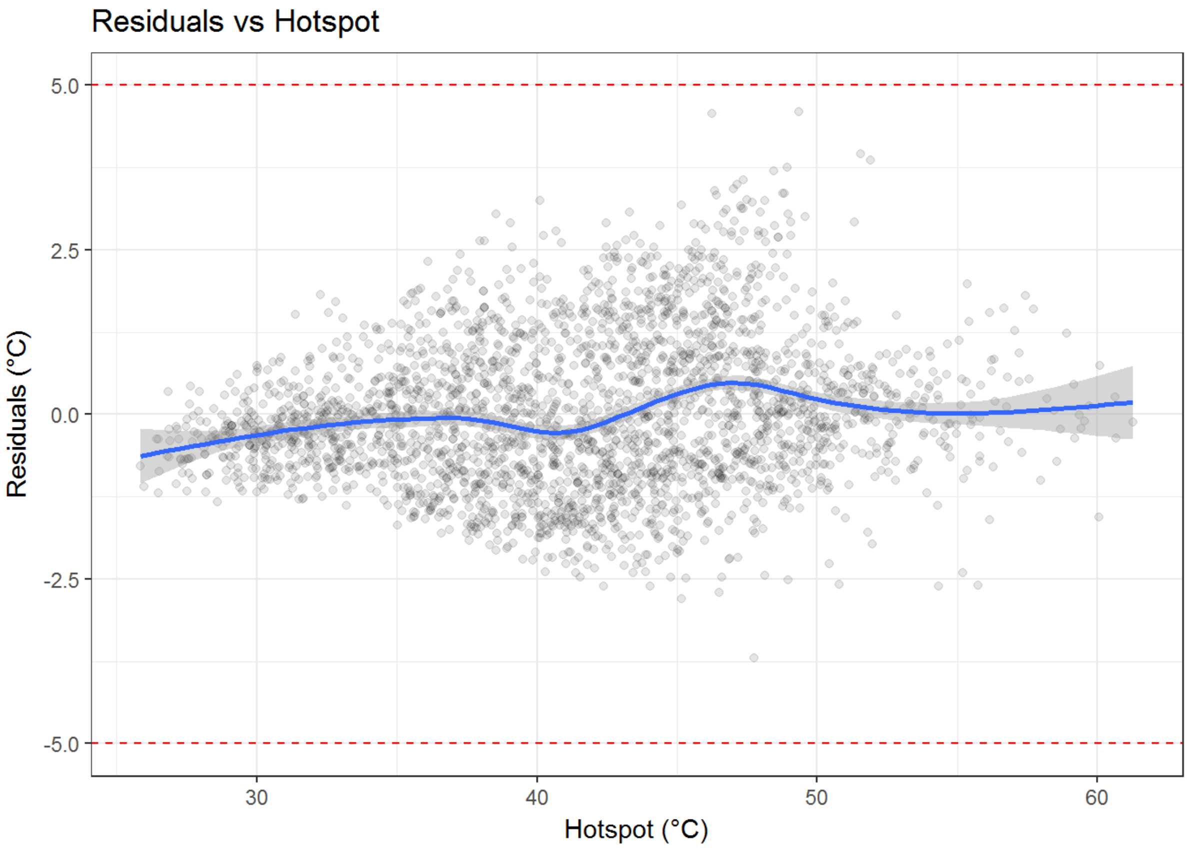

Figure 2. Residual between the calculated and measured hotspot sing a calculation based on surface temperature, plotted against hotspot

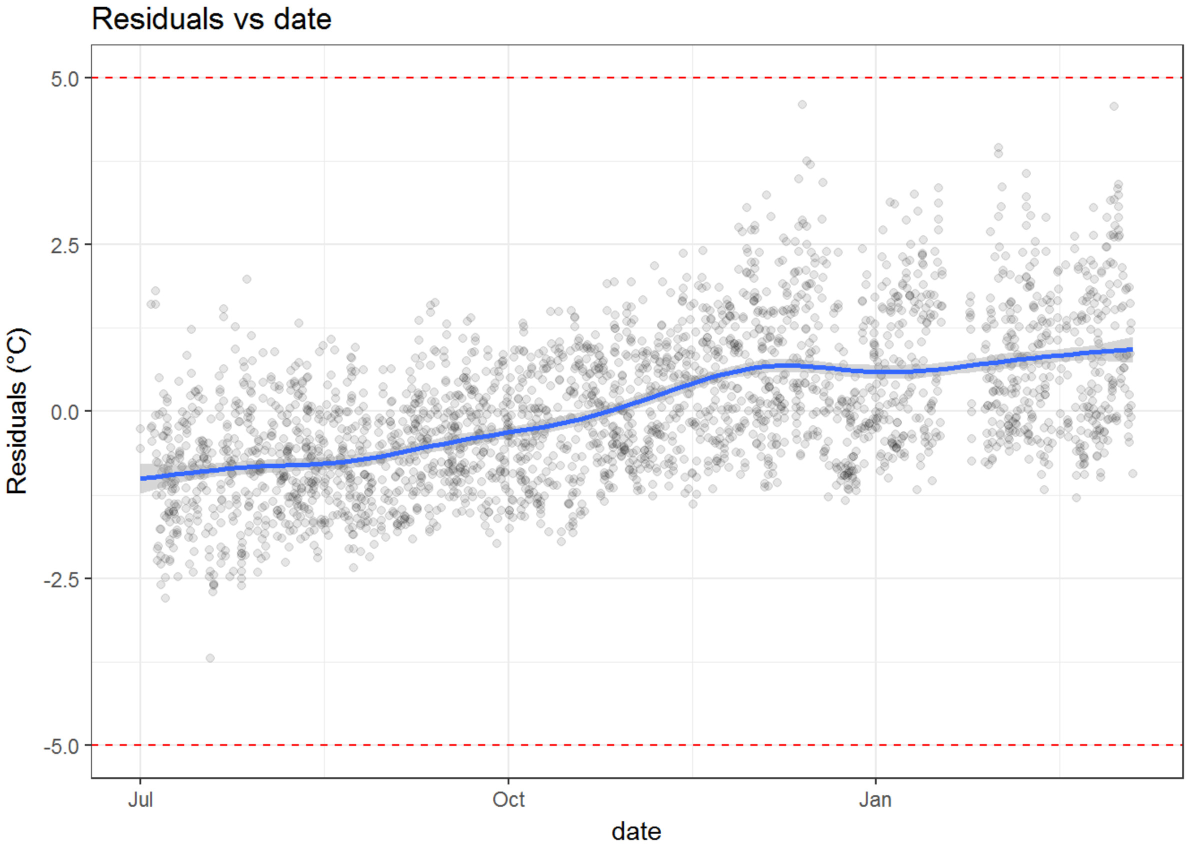

Figure 3. Residual between the calculated and measured hotspot sing a calculation based on surface temperature, plotted against measurement time

As can be seen from figure 2 and 3, the residuals are between -5°C and +5°C, indicating that the model is never more than 5°C from the measured result. Therefore, this is a good model, but there is room for improvement. Figure 2 indicates that there is little or no overall trend impacting this residual relating to hotspot data, as the average line shown in blue fits largely along the 0 axis. However, Figure 3 shows a trend in residuals over time, where they are lower in the summer months, and higher in winter. For this reason, a second variable of ambient temperature was selected for inclusion into the model.

By running a second regression that predicts the hotspot temperature based on the internal top oil temperature along with ambient temperature meant that the correlation coefficient increased from 0.97 to 0.99. By factoring in ambient temperature at high levels means that a more accurate model for calculated hotspot temperature was achieved. The new equation was as follows:

$$HS = (1.3828 * I_T) - (0.2594 * A_T) - 8.9964$$

Where,

HS is the hotspot temperature in °C

I_T is the internal temperature at the top of the tank in °C

A_T is the ambient temperature at high levels in °C

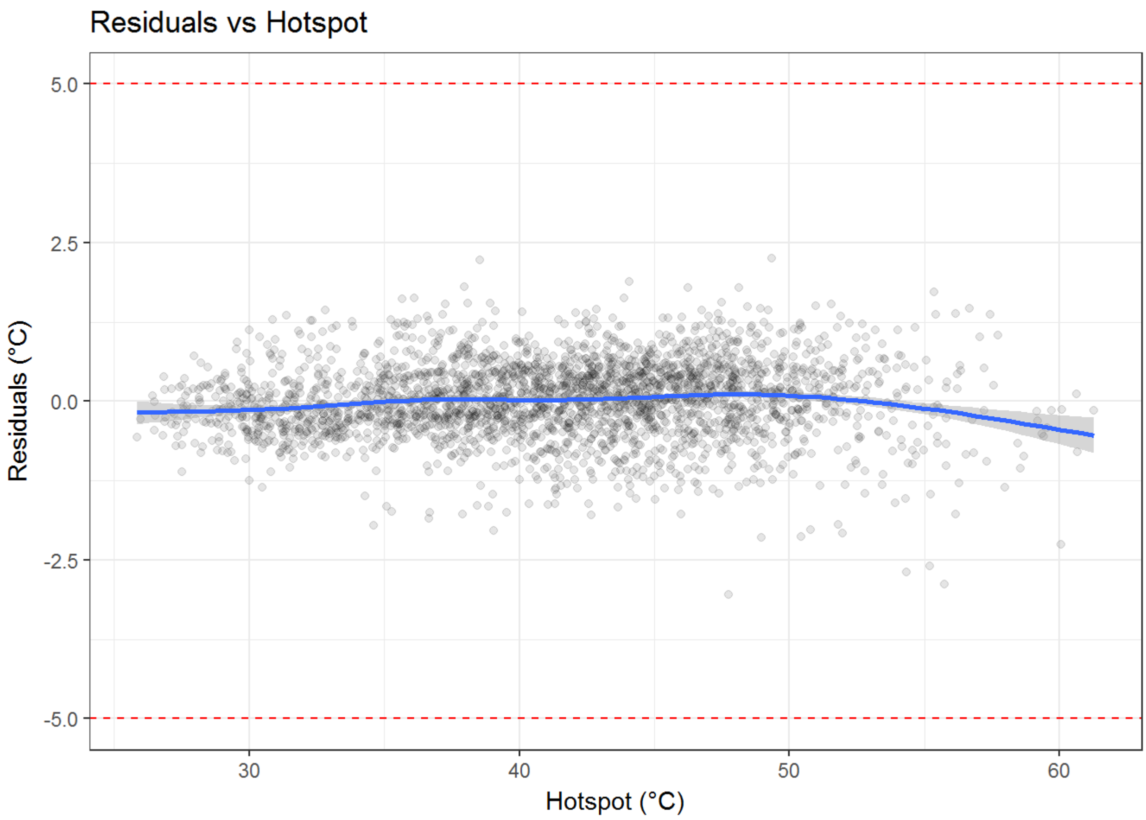

Figure 4. Residual between the calculated and measured hotspot sing a calculation based on surface and ambient temperature, plotted against Hotspot

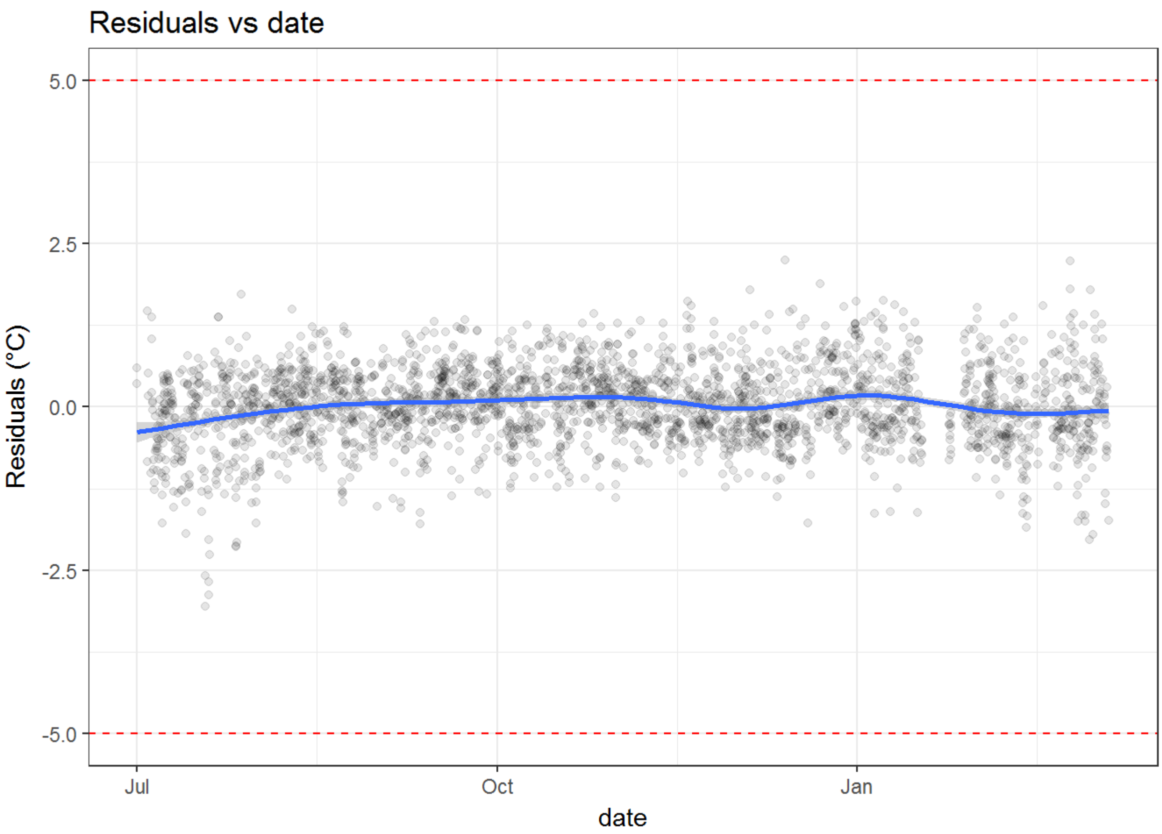

Figure 5. Residual between the calculated and measured hotspot sing a calculation based on surface and ambient temperature, plotted against measurement time

The results in figure 4 and 5 shows that residual between calculated and measured hotspot are even less than in figure 2 and 3, and are generally between -2.5°C and +2.5°C. This means that by taking ambient temperature into account has made the model better at predicating the hotspot temperature.

This model was validated with data from Clarendon Rd which has two transformers. The data used was from February 2018 only, as before this time there were serious data quality and completeness issues, and the site was too lightly loaded to provide useful validation data. Electricity North West increased loading and data quality issues were fixed so that validation data could be gained. Since February, the site loading has dropped again and therefore the useful loading data is not present. Electricity North West is considering this issue. The table below shows the loading levels of the two transformers at Clarendon Road.

Clarendon Rd	Average Utilisation %	Maximum Utilisation %
Transformer 1	44%	70%
Transformer 2	43%	72%

Table 3. Loading at Clarendon Rd

The generic model produced from Kincardine Rd was tested for this site and the results are as follows:

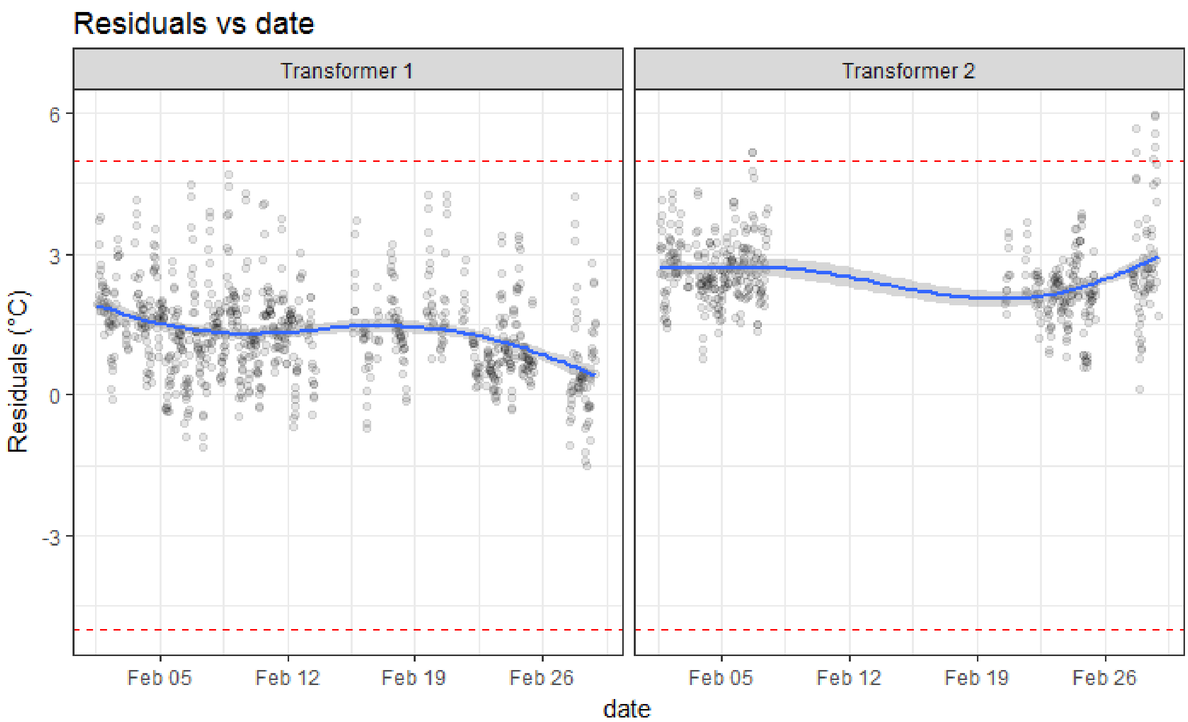

Figure 6. Validation test of hotspot prediction using data from Clarendon Rd

The residuals for both transformers are generally within -5°C and +5°C which suggest that a generic model that calculates hotspot temperature has been achieved. This will need to be tested for a much longer period when the site is heavily loaded.

2.2.2 Summary of results

Overall the generic model that predicts the hotspot temperature from the internal temperature at the top of the tank works reasonably well for Kincardine Rd and Clarendon Rd. The major limitation for this step is the lack of more example transformers with hotspot monitoring, that are adequately loaded for inclusion in the analysis. Clarendon Rd can only be used to validate the results from Kincardine Rd for a period when there are less data gaps, as this has reduced the robustness of the validation, and during a period when the site is heavily loaded. More data at heavier loading is required to test this model over a longer period.

2.3 Step 2: Link between internal top oil temperature and external surface temperatures

2.3.1 Analysis of internal monitoring sites

As mentioned previously, the next stage was to use all the sites that have internal monitoring installed at their transformers to develop a model that can calculate the internal top oil temperature from known information and data that can be gained easily from the outside of the transformer. The aim is to develop a model that can represent the range of sites, which includes transformers of different specifications and sizes, and different site characteristics, and to cover the higher loaded situations. The table below shows details for the 17 sites with internal monitoring installed. All internal monitoring sites shown in table 4 represents data that was downloaded from early July 2017 to the end of February 2018.

Transformer Site	Transformer Specification	Rating (KVA)	Average Utilisation %	Maximum Utilisation %
ABC Ardwick	ESI 35-1	1000	39.8%	77.1%
Clarendon Rd T1	ESI 35-1	1000	31.6%	72%
Clarendon Rd T2	ESI 35-1	1000	19.2%	73%
Dudley ST	T1	750	22.9%	52.6%
Helsby Way	ESI 35-1	800	15.9%	59.7%
Jenny LN	ESI 35-1	500	25.5%	60.3%
Jessel CL	T1	500	24.4%	51.4%
Kincardine Rd	ESI 35-1	1000	45.9%	88.7%
Lancaster Ave	T1	300	30.8%	89%
Leicester Ave	T1	500	26.1%	54.4%
Mount ST RMU	T1	750	23.2%	51.1%
Offerton DR	T1	500	34.2%	83%
Shackleton CT	ESI 35-1	500	25.3%	77%
Victoria RD Eccles	T1	750	45.2%	85.9%
Whitefriars	T1	750	42.5%	77.5%
Windsor Rd Prestwich	T1	1000	19.8%	68.6%
Woodend LN	ESI 35-1	315	30.4%	70.3%
Wordsworth Rd	ESI 35-1 UNIT	500	44.7%	97%

Table 4. Internal Monitoring sites with their load

The information used in the investigation included transformer and site characteristics as well as measured site data:

- **Site Characteristics**:
  - Site latitude
  - Site longitude
  - Site layout (including number of transformers, unit substation)
  - Substation building type (for example, brick, glass reinforces plastic, outdoor fenced enclosure)
- **Transformer Characteristics**:

- Oil mass
- Age of the transformer
- Transformer specification (T1, ESI 35-1)
- Transformer rating

- **Measured Data from Site**:
  - Surface temperature at the top and bottom of the tank
  - Ambient temperature at high levels
  - Other measured data from Celsius trial

All of these characteristics were considered in the model. The relative influence of each variable on the internal top oil temperature was compared, and the most influential were included in the model in turn to compare the fitness of the resulting model. As in Step 1, correlation coefficient R^2 is used as a measure of the success of the model to compare the results. This analysis approach considers each variable in turn. When an influential factor is included in the model, then where variables are closely related to this (for example, surface temperatures at different points on the transformer), they can often then be disregarded as they do not add additional information to the model.

Figure 7 below compares the R^2 for models containing the most influential characteristics. In this figure, the following variables are shown:

- **Surface_temp** – the temperature measured on the surface of the transformer at the top oil level
- **Tx.Sp** – the transformer specification. Three specifications were included; T1 (older transformer), ESI 35-1 (recent transformers conforming to the European standard), and ESI 35-1 UNIT (unit transformer)
- **Amb_temp** – the ambient temperature within the substation, measured away from the transformer at a high level off the ground.
- **Bld_type** – the substation building type, which for this subset of sites includes only stone/brick, and substations that are part of a larger building.
- **Age** – this is the age of the transformer, taken from the manufacture year.

All other factors that were included were found to have an insignificant impact on the success of the model, and so are not shown.

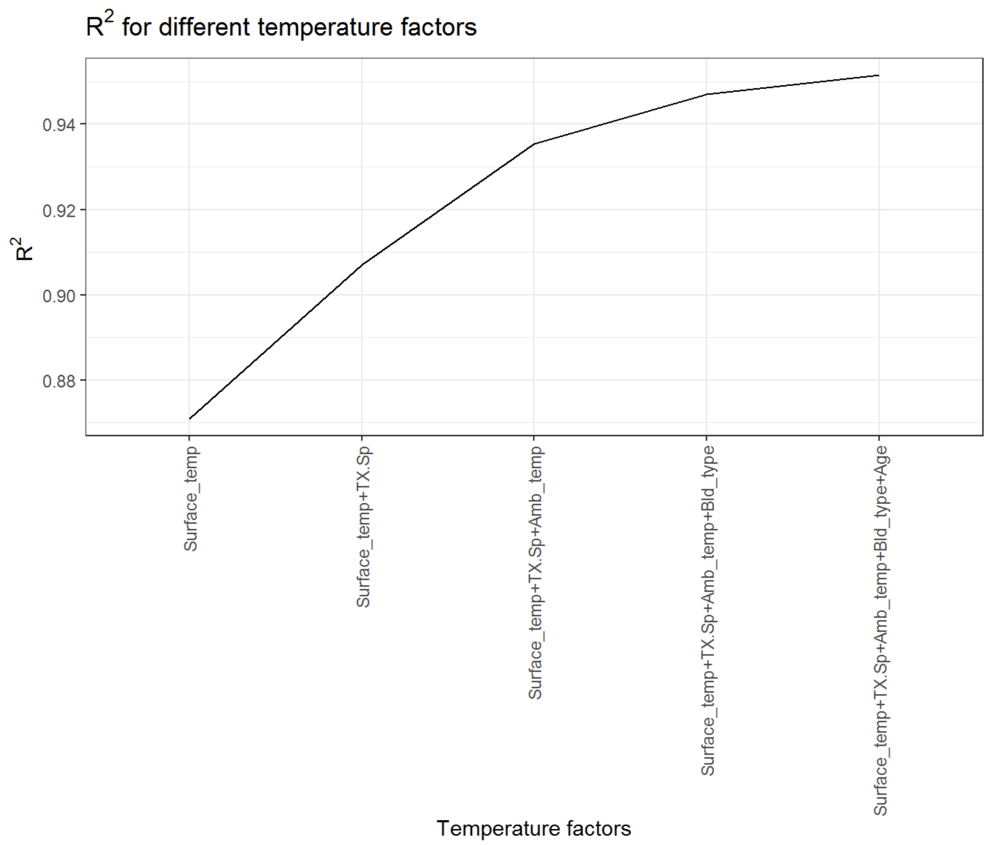

Figure 7. Measured internal top oil temperature correlated with model results for a range of temperature factors.

From figure 7, it can be seen that a model based on only surface temperature at the top oil level produces a strong model, with an R^2 of over 0.87. Adding the transformer specification followed by ambient temperature, building type, and age improves the model and the correlation coefficient variable increases.

There is a balance to be struck between fit of the model and its usability and complexity. The age of the transformer and the building type have a lesser effect on the model than surface and ambient temperatures and transformer specification, which is why they are not used to build the generic model.

The final model will predict the internal temperature at the top of the tank based on surface temperature, ambient temperature and for what transformer specification category the sites fall into. Note that there is only one example of the unit transformer specification, which will have an impact of the accuracy of this model.

T1: $I_T = (1.415 * S_T) - (0.393 * A_T) - 0.211$

ESI 35-1 UNIT: $I_T = (1.415 * S_T) - (0.393 * A_T) + 4.984$

ESI 35-1: $I_T = (1.415 * S_T) - (0.393 * A_T) + 1.245$

Where,

I_T is the internal temperature at the top of the tank in °C

S_T is the external surface temperature at the top of the tank in °C

A_T is the ambient temperature at the high level in °C

The three models were used to calculate the internal top oil temperature for each of the 17 sites with internal measurement, and this was compared with measured results to obtain the residuals, which represents the error in degrees Celsius between the calculated and measured internal temperature. Three plots are shown below which represents the residuals against the time of the measurements for all internal monitoring sites for each transformer specification category. The red dash line represents + and – 5 degrees Celsius error.

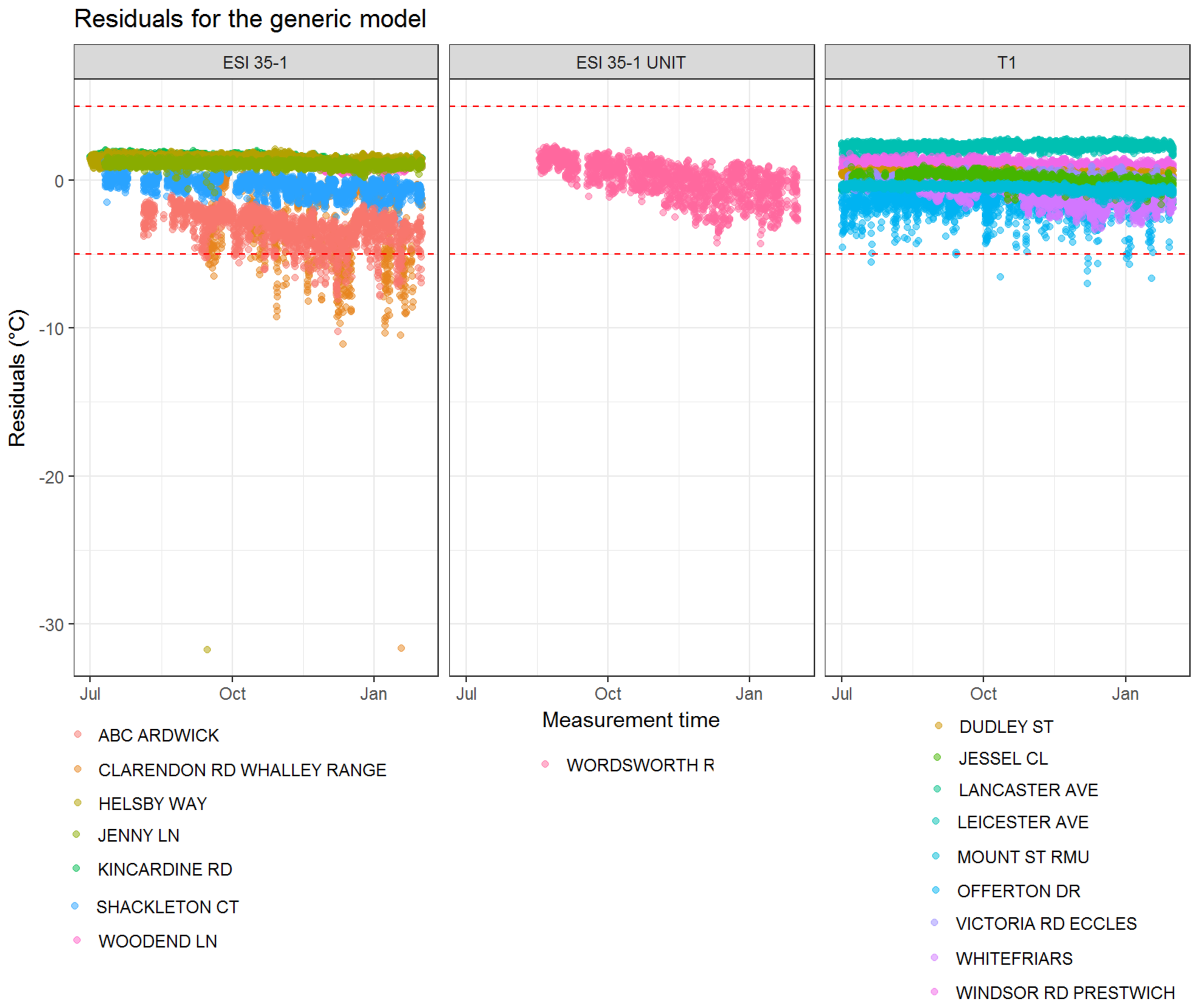

Figure 7. Residuals (the error in degrees Celsius between the calculated and measured internal temperature) for all internal monitoring sites

From the results, the three generic models work reasonably well across all 17 internal monitoring sites and has an error range that is within + or – 5 degrees Celsius. For the ESI 35-1 transformer spec category, the only two sites that are greater than 5 degrees Celsius are Clarendon Rd and ABC Ardwick. Clarendon Rd has two transformers within one site and the interaction between the transformers could explain why the assets are a higher temperature than expected. Also, both Clarendon Road and ABC Ardwick transformers are lightly loaded throughout the period from early July 2017 to February 2018, with transformer 2 at Clarendon road having an average utilisation of 19% over this period. It is intended to further investigate the discrepancies with these sites.

2.3.2 Summary of results

A generic model has been developed that can calculate the internal temperature at the top of the tank based on surface temperature, ambient temperature and from the transformer specification category.

As the level of loading over the 17 sites varies from heavy to light loading, the generic model to predict hotspot can be applied to all sites even when their load varies.

The main limitation from this section is the restrictions in the site selection, as the selection could not cover all building types and specifications. There are also issues with the loading of some of the lightly loaded sites; Mount St, Leicester Ave and Jessel cl. Helsby way cannot be more heavily loaded, and there are limited examples of extremely high loading amongst the sites. For these reasons, there are limitations to the degree to which the results can be applied across the full portfolio of sites and to explore the behaviour of very heavily loaded transformers.

2.4 Hotspot temperature calculated from external surface temperature

Using the results from section 2.2 and 2.3 a generic model can be developed which calculates the hotspot temperature from external surface temperature, ambient temperature and from which transformer specification the sites falls into. The results are as follows:

T1: $HS = (1.957 * S_T) - (0.802 * A_T) - 9.29$

ESI 35-1 UNT: $HS = (1.957 * S_T) - (0.802 * A_T) - 2.105$

ESI 35-1: $HS = (1.957 * S_T) - (0.802 * A_T) - 7.27$

Where,

HS is the hotspot temperature in °C

S_T is the surface temperature at the top of the tank in °C shifted backwards in time to compensate for time lag due to thermal mass. The appropriate time lag can be calculated for each site

A_T is the ambient temperature at high levels in °C

These three models will be used to estimate the hotspot temperature for all sites at the temperature factor stage which will now be discussed in section 3.

Note that there are several key limitations to the transformer hotspot study and its results, including:

- **Inadequate smart transformer data** – In many cases, the smart transformer data is significantly flawed, for reasons including the sites being extremely lowly loaded, the data communications from the smart transformers failing, or the KeLVN data being compromised causing significant data gaps. For these reasons, there is only one smart transformer site which has produced consistent, high quality data. A second site, Clarendon Road, has had data issues fixed and has been loaded to provide some validation data, but this is limited in volume. **Future work could be undertaken to validate and build on the results of this study using more numerous and varied examples of heavily loaded transformers where internal monitoring is available.**
- **Limitations to the selection of internal monitoring transformers** – The transformers with internal oil monitoring were selected to represent the breadth of site and transformer characteristics across the portfolio. However, due to safety and practical reasons, it was not possible to include outdoor substations or full unit Glass Reinforced Plastic enclosed substations. The result is that while it is possible to make reasonable assumptions about the application of the findings to the wider substation group, some of these applications are not tested. The selection also only includes one example of the unit transformer specification, which will have an impact of the accuracy of this model. **Future work could be undertaken on additional sites covering different building types, outdoor substations, and unit substation specifications.**
- **Limited examples of heavy loading** – The sites used in the hotspot study have a range of loading levels, from lightly loaded, to some examples of up to 90% loaded. However, as the focus of the Celsius project is about heavily loaded and high temperature sites, examples of transformers being very heavily loaded are required to ensure the analysis is accurate and representative at these levels, and there are only limited examples of sites at these loading levels within this study. **As demand increases over time, data from heavily loaded sites could be used to further validate the results of this study.**

These limitations together have an impact on the confidence with which the outcome of this study, the transformer hotspot temperature calculation methodology, can be used for the remaining analysis within Celsius, including the Transformer Temperature Factors work described in Section 3. The methodologies developed provide significant insight to the thermal behaviour of transformers, and the conclusions of this analysis are considered strong enough to be suitable for use in the Temperature Factors analysis to develop a robust methodology and draw initial conclusions. It is however, recommended that this analysis be updated when the required additional data described in the bullets above are available.

3. Transformer Temperature Factors Study

3.1 Approach

As mentioned in the previous section, the goal of the Transformer Temperature Factors Study was to understand what environmental and loading conditions have an influence on the maximum operating temperature of LV transformers. The purpose of carrying out this analysis was to develop a better and more accurate way of estimating the hotspot temperature for transformers under a range of conditions, and to draw conclusions about the influence of the asset environment factors on thermal rating. As asset ratings are considered conservative in most situations, more informed ratings can increase the capacity of the sites and allow more load to be applied which will help to ease the transition towards connecting low carbon technologies to the network and to cope with the rising demand.

This analysis focuses on ground mounted substations. Within Celsius, there are 520 ground mounted substations included in the monitoring trial. The monitoring configuration varies from site to site, including 128 sites with three phase voltage and current monitoring, and 336 sites with single phase monitoring. This was a decision that was made to reduce the cost to the project of monitoring the sites.

The work follows on from the Hotspot analysis work which makes use of the expression for predicting the hotspot temperature from surface temperature, ambient temperature and transformer specification type as described in Section 2. This is a generic expression which can be applied to all sites, noting the limitations described in section 2.4.

The approach taken for the temperature factors work is as follows:

- Identifying the major temperature factors – Identification of the major environmental factors that could have an influence on the hotspot temperature. General engineering principles about heat generation and flow in substation equipment were used to identify potential factors. Another important consideration was the availability of data for each of the factors.
- Investigating the influence of temperature factors – Determine what effect the environmental factors have on the maximum operating temperature. A model in Rstudio was created which identified which environmental factors have a strong influence on the hotspot temperature. These steps are described further in the sections below.

3.2 Identifying the major temperature factors

The first stage in the analysis was to identify all the major environmental factors that could have an influence on the hotspot temperature and narrow them down to the most important ones. Transformer loading also had to be considered, as this is a known driving factor of temperature.

The factors that were considered in this study fall under three categories; load conditions, transformer characteristics, and site characteristics. The following lists all the factors taken into account during the course of this study.

- **Load Conditions**:
  - Load (apparent power)
  - Load Profile
  - Utilisation
  - Harmonics
- **Site Characteristics**:
  - Site location (latitude and longitude)
  - Substation building type (for example, brick, glass reinforces plastic, outdoor fenced enclosure)
  - Site layout (including number of transformers, unit substation)
- **Transformer Characteristics**:
  - Age of the transformer
  - Transformer specification (T1, ESI 35-1)
  - The following sections of this report outline how the different sites were categorised when considering each temperature factor as well as how the required information was obtained.

3.2.1 Load

The total apparent power is calculated by summing the product of V and I for all three phases. Where there is only single-phase monitoring, the total load of the site is estimated by multiplying the single phase apparent power by three. From engineering principles, the heat generated is proportional to the square of current ($P_{heat} = I^2R$), and current is proportional to the apparent power (S = VI). Consequently, the heat generated is proportional to the square of the apparent power ($P_{heat} = S^{2(}R/V^2)$). Therefore, the temperature factor being considered here is the square of apparent power.

Load is considered an important temperature factor as its directly proportional to the heat produced by the asset. While the impact of load on hotspot temperature is not the focus of the Celsius investigations, it must be compensated for to investigate the impact of other factors, and therefore is an important part of the study.

3.2.2 Load Profile

Distribution substations will be exposed to different load's depending on location, weather, time of day or year. A substation based in the central business district will have a commercial load profile during the week and flat profile during the weekend. A substation supplying a neighbourhood will have a residential load that peaks in the morning and evening.

The Celsius analysis categorised load profiles using a normalising and clustering method, as described in the following sections.

3.2.2.1 Data Extraction and Preparation

Load data from Celsius sites included voltage and current, in some cases 3 phase, and in others single phase, which is used to calculate apparent power. For this study, only those sites with 3 phase power monitoring were used, which was a total of 128 sites. In order to categorise load profile, the data was considered in chunks of a single day, from across all Celsius sites. A database was generated with the data laid out in this form.

There were a number of issues with the data, including instances of:

- Data gaps
- Invalid data,
- Substations with a low load profile,

Days containing significant data gaps were removed as they could create inaccurate load profiles. This was addressed by filtering out days with more than 5 hours' data gap.

The sensors would, at times, report data that was clearly invalid. To make sure these measurements were not considered, all days containing measurements outside of the acceptable voltage range (220 V to 280 V) were excluded. Also, data with a low maximum load are not considered significant to the investigation. All days containing measurements outside of the acceptable current range (>15A) were excluded.

3.2.2.2 Normalisation

The data was normalised within each day's readings. All readings have been scaled using the maximum reading on the day set to 1. Thus, all hourly readings are in the range 0-1. The effect of this normalisation is to focus on the shape of the usage pattern and not on the total usage.

Two households with a similar shape but with differing total usages (e.g. if one household is much larger than the other) may have the same normalised load profile once scaled. As a consequence, the clustering algorithm can compare the consumption habits from two households or commercial areas of different sizes.

3.2.2.3 The Clustering Method

Research undertaken about clustering algorithms identified the K-means method as an appropriate solution for clustering the normalised load profiles. This method of load profile classification was used in the UK Power Networks Distribution Network Visibility Project.

The K-means method requires a number of clusters as an input parameter and works by randomly selecting an initial n locations for the centres of the clusters. Each data point is then assigned to one of the centre locations by selecting the centre that is nearest to that data point. Once all the data points are assigned, each collection of points is considered. A new centre of the allocated points is calculated and the centre for that cluster is reassigned. The K-means method uses Euclidean distance calculated for centre c = (c_1; c_2; ...; c_n) and point p = (p_1; p_2; ...; p_n) as

$$distance = \sqrt{\sum_i (c_i - p_i)^2}$$

The points are then reallocated to their new nearest centre and the algorithm continues as before until no changes are made to the allocations of points between iterations. The K-means algorithm returns the n centres of the clusters (called centroids), and the classification of each point in the cluster it belongs to.

The method is highly dependent on the initial random allocation of centres. To avoid this issue which gives rise to different results every time the clustering method is executed, the program was run 25 times with random centroids and the one with the least Mean Square Error was chosen.

3.2.2.4 Optimisation

As seen above, the K-means method requires a number of clusters as an input parameter. As we expect, the more clusters we choose, the more accurate the output will be, as the points will be closer to the centroid of the cluster they belong to. The extreme case is a number of clusters n equal to the number of points; in this case each point will be in its own cluster.

The optimal number of clusters is ambiguous and depends on the shapes and the number of the points on which the algorithm is applied. A good clustering scheme will create clusters where the members of a particular cluster are closely grouped, but where the different clusters are well separated. The method chosen to determine the optimal number of clusters uses a measure assessing the quality of the clusters generated: the Mean Index Adequacy (MIA):

$$MIA = \sqrt{\frac{1}{K}\sum_{k=1}^{K} d^2(r^{(k)}, C^{(k)})}$$

Where K is the number of clusters defined, $r^{(k)}$is one of the load profiles assigned to cluster number k, and $C^{(k)}$is the centroid of the cluster k. The optimal number of clusters should be chosen so that adding another cluster does not give a much better modelling of the data. Figure 10 shows the MIA decreases a lot as the number of clusters increases from 1 to 4, and progressively the marginal gains drop by adding a cluster. Therefore, 4 clusters were used for the optimisation.

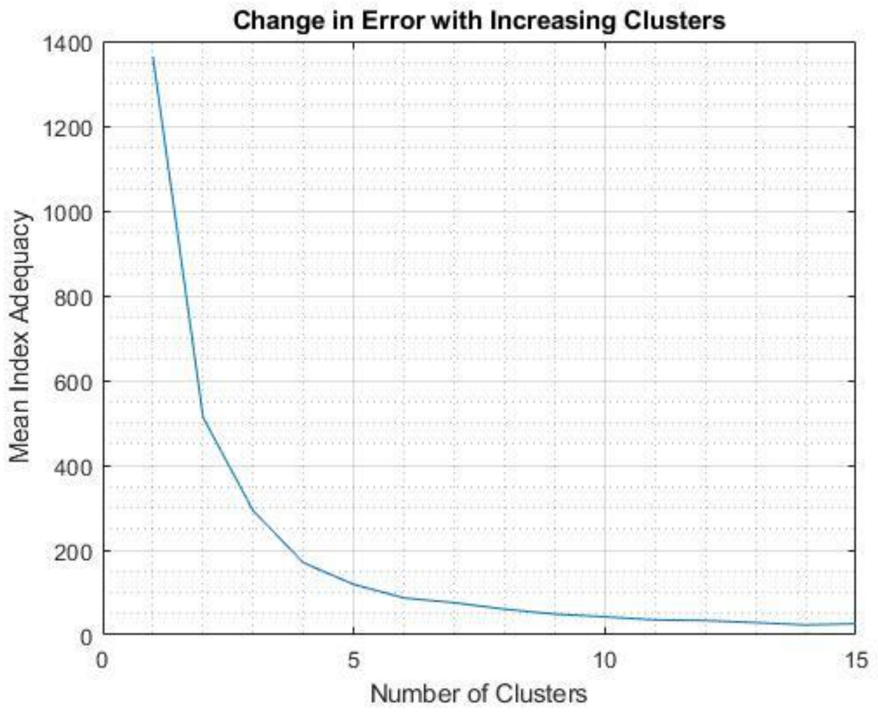

Figure 8. Change in Error with increasing Clusters

3.2.2.5 Results

Consequently, the following graph shows the four load profiles that represent 1 year's data from all type 1 and 1a sites. The clusters are named c1 to c4.

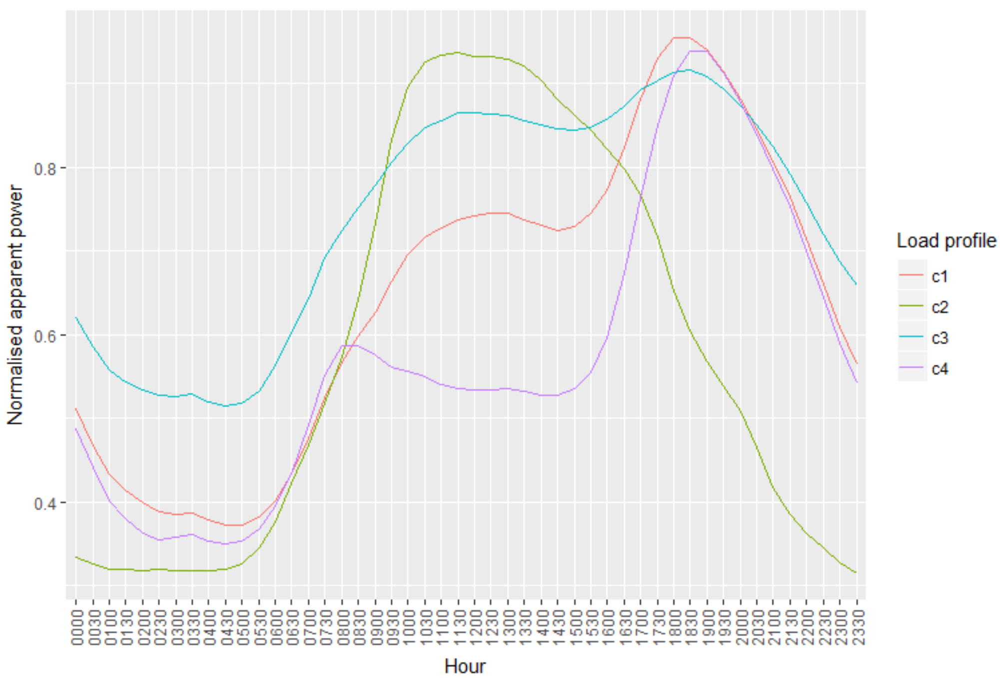

Figure 9. Number of clusters used for load profile

The 4 clusters are general representations of the typical load profile on ENW's distribution transformer.

These conclusions were applied across the whole Celsius trial portfolio by determining the load profile cluster for each site and for each day. For the sites with single phase monitoring, the total load of the site is estimated by multiplying the single phase apparent power by three.

3.2.3 Utilisation

Another load related factor that could have significant influence on the maximum operating temperature of the transformer is how much its loaded. The load profile factor mentioned above takes into account the change in load throughout the day. However, the actual value of load through the day is significant. The maximum utilisation for the day of a transformer can be used to add scaling information to the load profile, and is calculated for each day for each site using the equation below;

$$Max\ Utilisation = \frac{Max\ Apparent\ Power}{Transformer\ Rating},$$

for each day and site, the maximum utilisation values were calculated. Again, for the sites with single phase monitoring, the total load of the site is estimated by multiplying the single phase apparent power by three. Utilisation was used as a temperature factor because it gives an indication of how well loaded the asset is, which is likely to be an indication of the heat produced by the asset and hence the hotspot temperature.

3.2.4 Harmonics

Voltage and current harmonics are measured for every site. Harmonics are expressed as a measure of total harmonic distortion (THD). This is defined and calculated by the following equation (for example, for voltage):

$$THD = \frac{\sqrt{{V_2}^2 + {V_3}^2 + \cdots + {V_{25}}^2}}{V_1}$$

Where V_n is the RMS voltage of the n^{th} harmonic and n=1 is the fundamental frequency. At each site, the hex unit calculates the total harmonic distortion by measuring the harmonics up to the 25^{th} order.

As mentioned previously, there are sites with three phase monitoring, and sites with single phase monitoring within the Celsius trial. In order to analyse these sites alongside each other, The three phase THD for the sites with three phase monitoring were averaged and combined with the single voltage and current THD phases for the sites with single phase monitoring.

3.2.5 Building Type

The sites were each categorised into the following building types;

- Fenced Enclosure (outdoor),
- Glass Reinforced Plastic,
- Part of larger building,
- Stone/brick,

Table 6 illustrates some of the building types above.

Building Types	
Fenced Enclosure	Glass Reinforced Plastic
	
Stone/Brick	Part of a Larger Building
	

Table 5. Different Building Types

The building type could have a significant impact on the ambient temperature of the asset's operating environment. An outdoor transformer will experience more wind than an enclosed transformer. The increase of the convection heat transfer apparatus in this situation could result in higher operational limits for transformers. The opposite could be said during the summer when the outdoor transformer is exposed to the sun while the enclosed transformer is protected. For this reason, it has been determined that the building type could be an important temperature factor to investigate.

3.2.6 Location

The location is defined with 2 categories in the temperature factor investigation; the site's longitude and latitude coordinates. Locations of each site were found using GPS and stored on the Celsius server. We expect that location can make a difference on the operational limit of assets, for example, substations closer to coastal areas might be exposed to more wind and lower temperature allowing them to operate at higher ratings.

3.2.7 Transformer Type

There are 3 transformer types that every site that is of interest in the Celsius project can categorised into. These are as follows:

- T1
- ESI 35-1
- ESI 35-1 Unit

The T1 type is an older British Standard specification, and represents transformers that were manufactured up until the 1970s. The ESI 35-1 types are newer European standards that have been installed since the 1970s. Finally, ESI 35-1 Unit are unit transformers where both the LV transformer, LV switchboard and RMU are combined into one full unit.

3.2.8 Site Layout

The sites that are being investigated have a range of site layouts which could potentially have an impact on the hotspot temperature. For example, where sites have two transformers, A separate layout where the transformer, lv board and RMU are all separated, or a full unit where the transformer, lv board and RMU are combined into one full unit.

3.2.9 Age

The age of an asset can have significant effect on its operational efficiency. Assets can suffer from wear and tear, material and oil degrading and eventually malfunction after a certain period. Clearly, there are external factors that can accelerate the process of ageing such as its loading conditions, environment, and the number of times it has been maintained.

The data available for the Celsius project is the manufacture year, which is available from transformer nameplates.

3.3 Temperature Factor Investigation

Using the same technique that was used in figure 8 in section 2, linear regression was carried out and a R^2 value was obtained between the hotspot temperature and the models containing the most influential factors. In this figure, the following variables are shown:

- **S** – Apparent power through the transformer. The factor uses S^2, it is expected that the hotspot temperature will be related to the square of the current.
- **Site_ building type** – The substation building type, which for this subset of sites includes only stone/brick, and substations that are part of a larger building.
- **Age** – The age of the transformer, taken from the manufacture year.
- **Util** – The daily maximum utilisation of the transformer
- **Average_Voltage_THD** – The voltage Total Harmonic Distortion, calculated as shown in Section 3.2.4.
- **Load_cat** – The load profile category, determined as described in Section 3.2.2.
- **Layout_Type** – The site layout as described in Section 3.2.8
- **Average_Current_THD** – The current Total Harmonic Distortion, calculated as shown in Section 3.2.4.
- **Site_Longitude** – The longitude of the site
- **Site_Latitude** – The latitude of the site

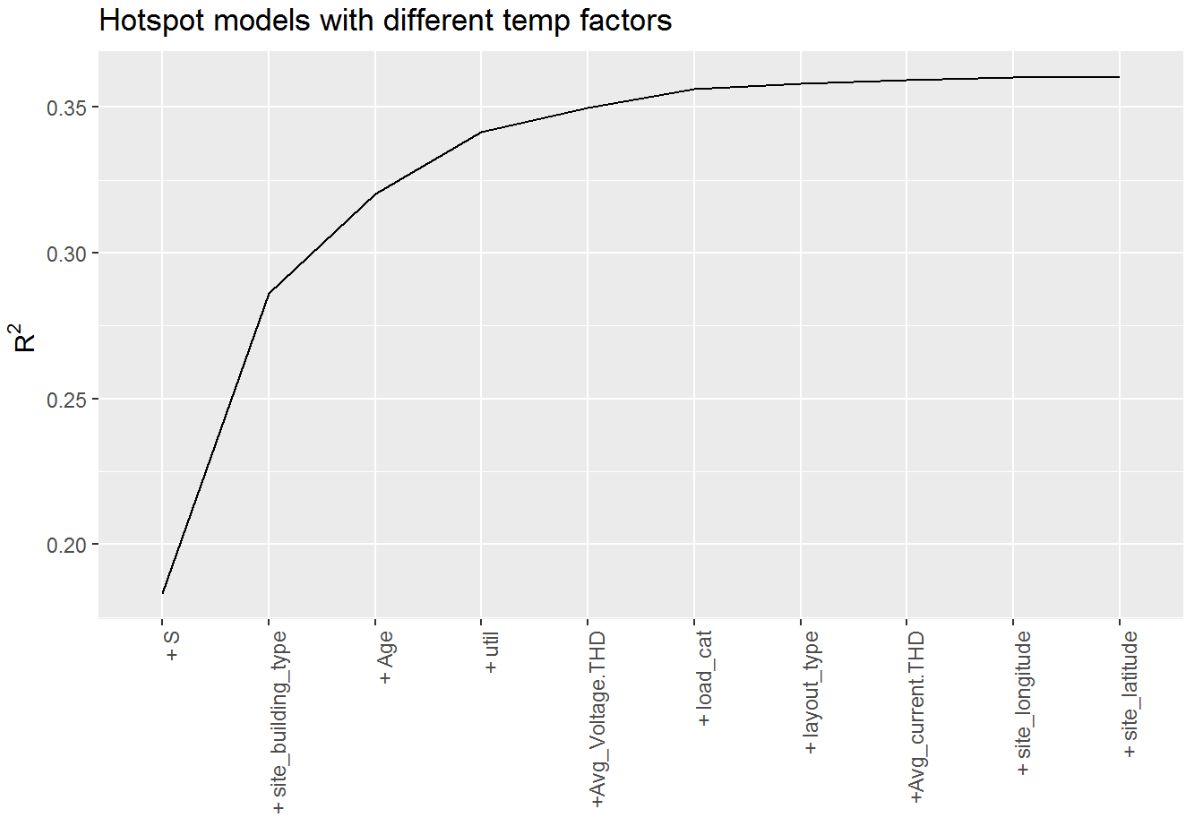

Figure 10: A comparison of models including different temperature factors and the resulting correlation with hotspot temperature. Note that each model builds on the last.

From figure 10, the total apparent power along with building type and age of the transformer have the strongest influence on the hotspot temperature for all sites. The total voltage harmonic distortion, percentage utilisation and load factors have a less significant effect, each improving the R^2 value by only a few percentage points. The layout type, the total current harmonic distortion and the site location (longitude and latitude) have no almost effect on the accuracy of calculating the hotspot temperature.

As mentioned in section 2, the R^2 value needs to be above 75% to be deemed acceptable. Therefore, from Figure 10, these factors do not produce an accurate model that can be used to predict the hotspot temperature.

A possible explanation for this is because the sites were lightly loaded, resulting in a less than expected correlation with site load. A similar graph was produced for sites that have over 50% utilisation and the results are as follows.

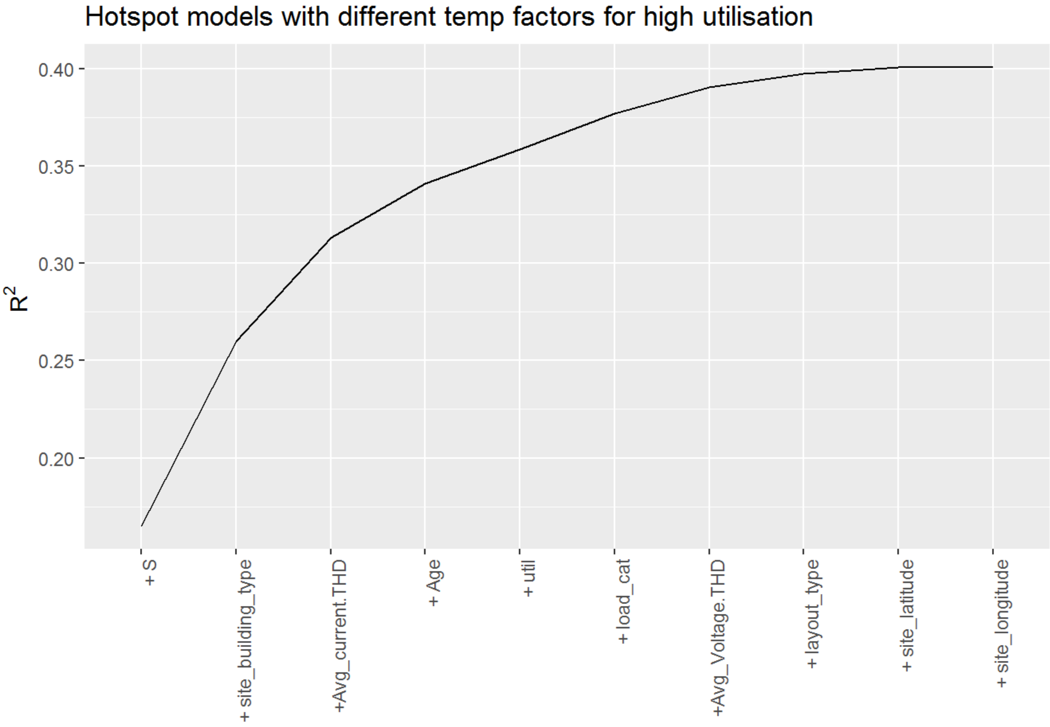

Figure 11. Hotspot from temperature factors for higher loading

The initial value of R^2 between the hotspot temperature and total apparent power drops from 0.16 in Figure 10 to 0.13 in Figure 11, suggesting that the correlation with load is still very low, and has actually decreased slightly. The final value of R^2 has increased due to the higher utilisation. However, overall the accuracy of the model has not improved significantly.

A further investigation on the effects of the most influential temperature factors, as seen in figure 10, was carried out, and the results are described in the following sections.

3.3.1 Temperature Factor Investigation: Load

First, the correlation between hotspot temperature and the square of apparent power can be seen in Figure 15. It shows a correlation that produces an R^2 value of 0.15. This is a low value but it does show a general increase in hotspot temperature as the apparent power increases.

These results can be used to conclude that, in general, the hotspot temperature increases for high loading values, but the relationship is not strong enough to provide a definitive equation that applies in all cases. This reinforces the conclusion that that there are other important factors that influence the hotspot temperature, and therefore load alone cannot be used as an accurate indication of hotspot temperature.

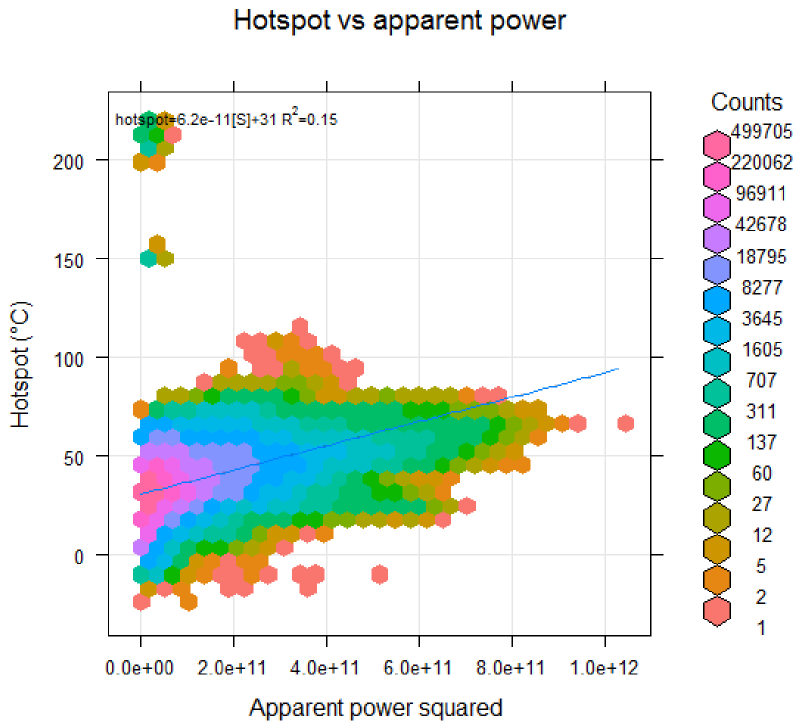

Figure 12: Effect on Hotspot temperature as load increases

To investigate the low correlation observed between load and hotspot temperature, the R^2 value was obtained between the hotspot temperature and the load squared for each site. Figure 13 shows these R^2 values ordered from largest to smallest. Each point in the graph represents a calculated R^2 value for a site.

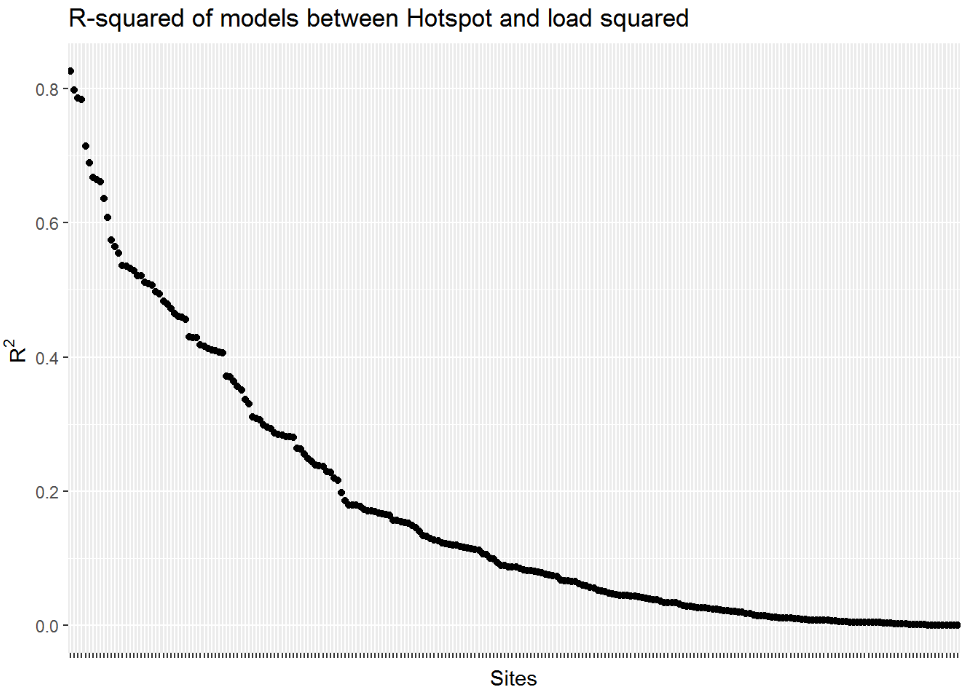

Figure 13: Correlation between hotspot and load for all sites

Approximately 98% of the sites have an R^2 value that is less than 0.7. This is poor correlation between hotspot and load for most of the sites. An explanation for this is that most of these sites were lightly loaded and thus the operating temperature is less impacted by the load and more influenced by environmental factors. The graph below shows the percentage utilisation for each site, with the black vertical line representing the range of utilisation, and the red dot showing the average. The order of sites along the x axis has been kept the same as in figure 13.

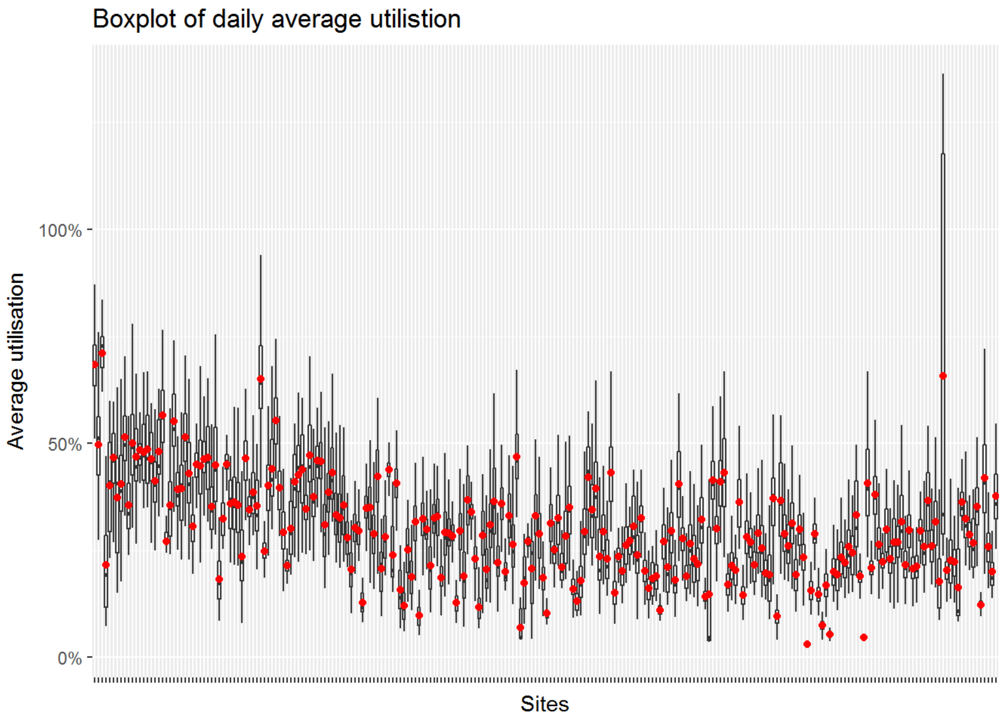

Figure 14: Boxplot of average utilisation (note that the outlier to the right of the figure has been investigated and appears to be due to erroneous data which has now been resolved)

Figure 14 shows some trend in the utilisation, with the left part of the graph having generally higher average utilisation, corresponding to higher R^2 than the right. However, this trend is not conclusive, with a wide band of variation and a number of outliers. This means that the high R^2 value observed in figure 13 could be due to higher loading, but further investigation is required.

3.3.2 Temperature Factor Investigation: Building Type

The effect of individual building type categories on hotspot temperature has also been investigated.

The trial includes the following building types:

- Fenced Enclosure (outdoor)
- Glass Reinforced Plastic
- Part of larger building
- Stone/brick

The graph below illustrates the average hotspot temperature for each building type (shown in grey) and the range of hotspot temperatures (shown in red). Generally, there is no clear difference in hotspot temperature for ‘Stone/brick’ or ‘Part of a Larger Building’ substations. However, the average hotspot temperature decreases significantly for ‘Glass Reinforced Plastic’ and even more for ‘Fenced Enclosures’.

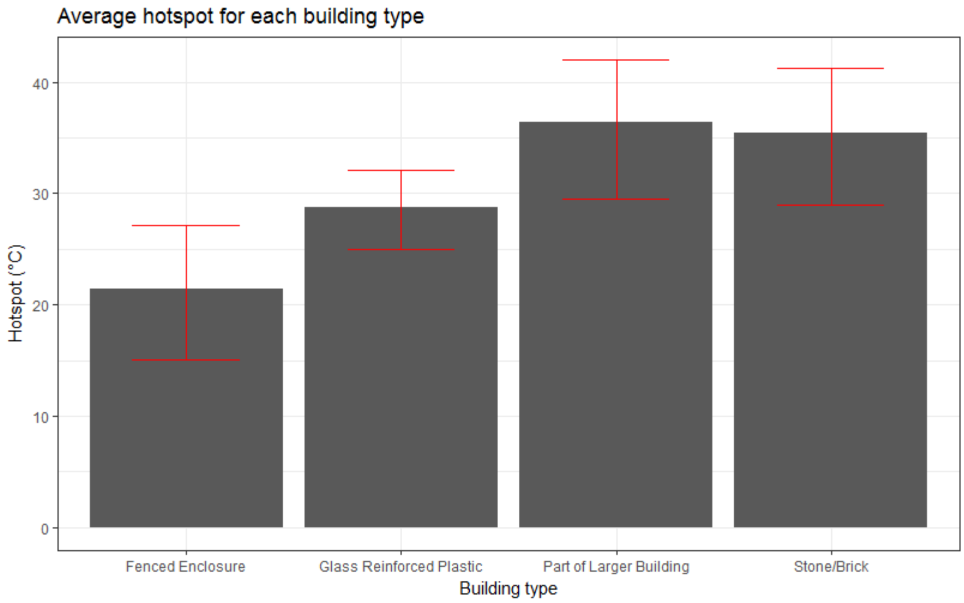

Figure 15: Effect of Hotspot Temperature with Different Building Types

Figure 16 compares data from the summer and the winter.

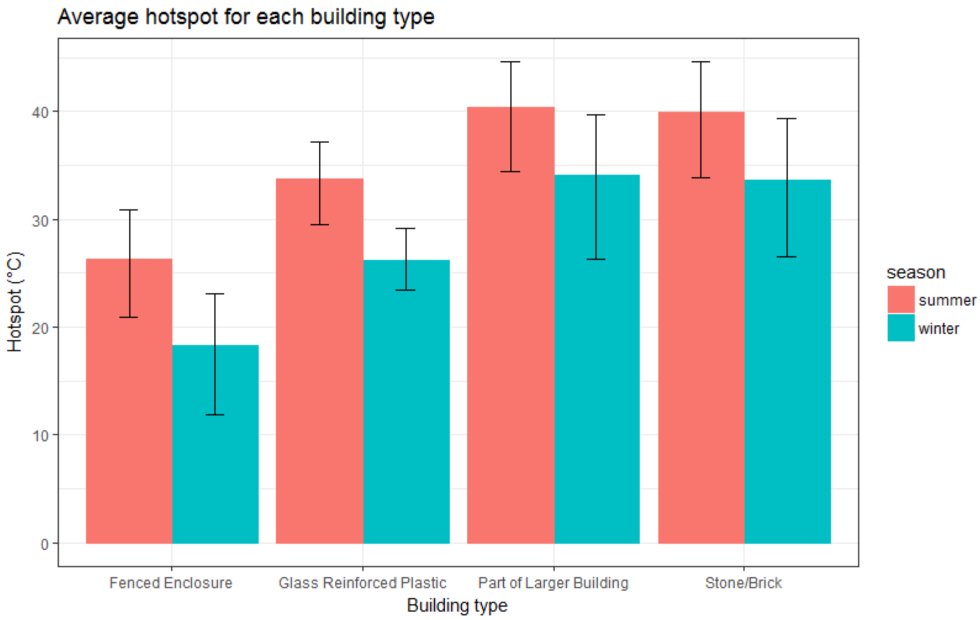

Figure 16: Effect of Hotspot Temperature in Different Season

Figure 16 shows that the relative performance of the building types does not change between winter and summer i.e. fenced enclosure is the coolest building type followed by GRP, part of larger building and stone/brick in both seasons. However, the difference between summer and winter temperatures are more pronounced in fenced enclosure (outdoor) sites and glass reinforced plastic substations, than compared with stone/brick buildings and substations that are part of a larger building.

It is important to take into account the loading of the sites within each building type. This is shown in Figure 17 below:

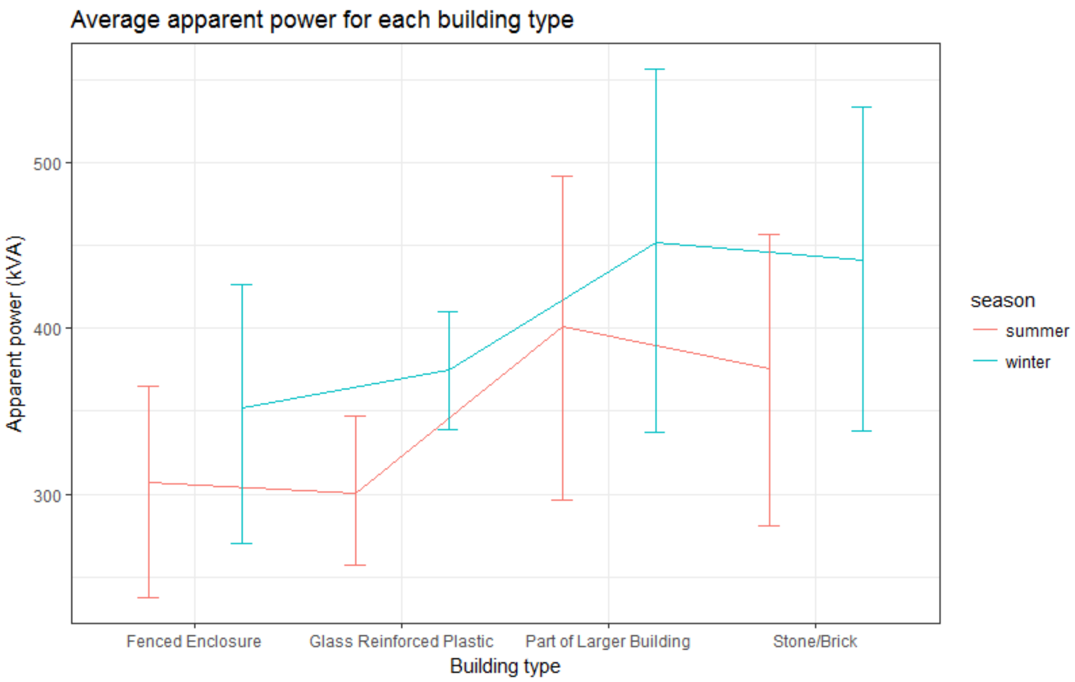

Figure 17: Average load for building type

The average apparent power is different for different building types. The glass reinforced plastic and fenced enclosure outdoor sites are loaded the least which explains their low average hotspot temperature. Although fenced enclosure sites have similar loads to glass reinforced plastic sites, they have a significantly lower average hotspot temperature.

From the investigations into building type, it can be concluded that the relationship between hotspot temperature and building type is not strong enough to conclude a definitive trend that applies in all cases. However, some insights can be drawn, including:

- Stone/Brick buildings and substations that are part of a larger building behave in a similar way, with relatively similar average overall temperatures at similar loading levels, and about 6°C difference in summer and winter temperatures. This is expected as they have a significant thermal mass, meaning that they respond to ambient temperature is a similar way.
- On average, Glass reinforced plastic substations and Fenced Enclosure are both much cooler than Stone/Brick buildings and substations that are part of a larger building, which is likely to be due to differing loading levels.
- With comparable loading levels, Fenced Enclosure buildings are cooler than Glass Reinforced Plastic substations. This may be the increased ventilation and convention possible with outdoor substations. The difference between summer and winter average temperatures at Glass reinforced plastic substations and Fenced Enclosure sites is greater than the Stone/Brick buildings and substations that are part of a larger building, at about 8°C. This is as expected, as these sites are more exposed to the ambient weather conditions; outdoor fenced enclosure transformers can be directly heated by the sun and cooled by the wind. Glass reinforced plastic substations have very little thermal mass, and are therefore also much more directly affected by the weather conditions.

3.3.3 Temperature Factor Investigation: Transformer Age

The next most influential factor was the transformer age. The correlation between age of the transformer and hotspot temperature has been investigated. The results show a relatively flat line of best fit with a very poor R^2. This suggests that the relationship between transformer age and hotspot temperature is weak enough that no significant conclusions can be drawn.

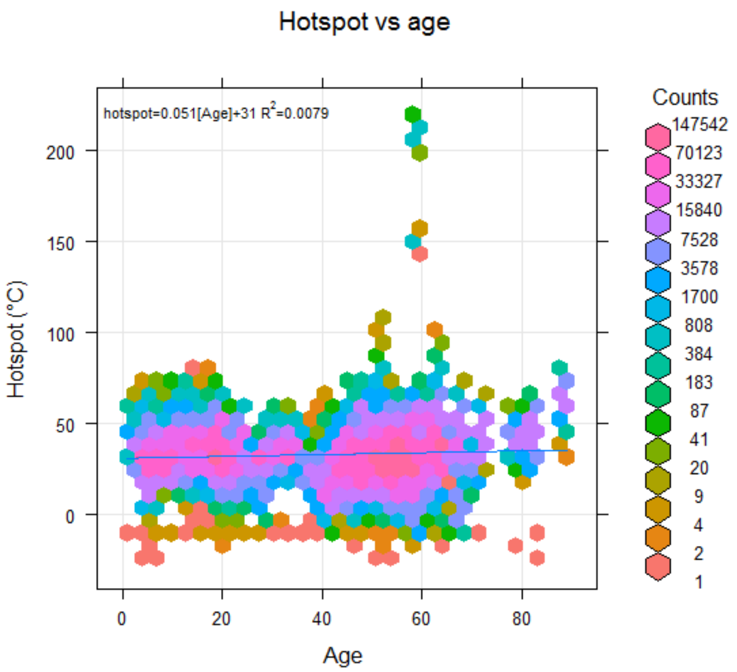

Figure 18: Effect of age on hotspot temperature

As all other temperature factor have a lower correlation as indicated in figure 13, it is not considered beneficial to carry out a similar detail study for the other factors.

3.4 Discussions and Limitations

The temperature factor study showed that, with the current data, it would be very difficult to predict hotspot temperature from the proposed temperature factors, even when considering loading.

The study showed that the load and hotspot temperature are very poorly correlated, which is unexpected. This could be due to the limitations in the data as detailed below. The study showed that there are temperature factors that are relatively more important than others. Transformer load and building type were the most influential factors, with building type impacting the way that the transformer temperature is influenced by loading and by the ambient conditions.

The results produced in this investigation, although informative, are not conclusive. There are key limitations to the results produced in this section.

- **Incomplete Hotspot Study work –** as described in Section 2.4, the conclusions of the Hotspot Study were not conclusive. As the temperature Factors work builds directly on the equations from the Hotspot Study, there are limitations to the conclusions made in this analysis**. If relevant data becomes available during the project the Hotspot Study work can be reviewed and updated results could be incorporated into the Temperature Factors Study methodology.**
- **Inadequate data quality** – It has been observed that site and sensor data have data gaps, and can report erroneous data. Data gaps and erroneous data can affect the analysis, for example when correlations and taking time shifts into account. The combination of missing data along with erroneous data can significantly distort the data. **It is intended to investigate methods to 'clean' and validate the data before further analysis is carried out and revisit the correlation assessment.**
- **Limited examples of heavy loading** – As the focus of the Celsius project is about heavily loaded and high temperature sites, examples of transformers being very heavily loaded are required to carry out correlations between hotspot temperature and load. However, only 5% of the data was for sites loaded at more than 80% utilisation, and less than 1% of the data was loaded at more than 90% (see figure 19, which shows the amount of data as a percentage of the total amount of data where the loading was larger than a given utilisation). This not only means the analysis is less applicable at heavily loaded conditions, but the light loaded sites may be distorting the results and decreasing correlation with temperature factors. The increase in load could also uncover other temperature factors that are more significant at higher loads. **As demand increases over time, data from heavily loaded sites could be used to further validate the results of this study.**

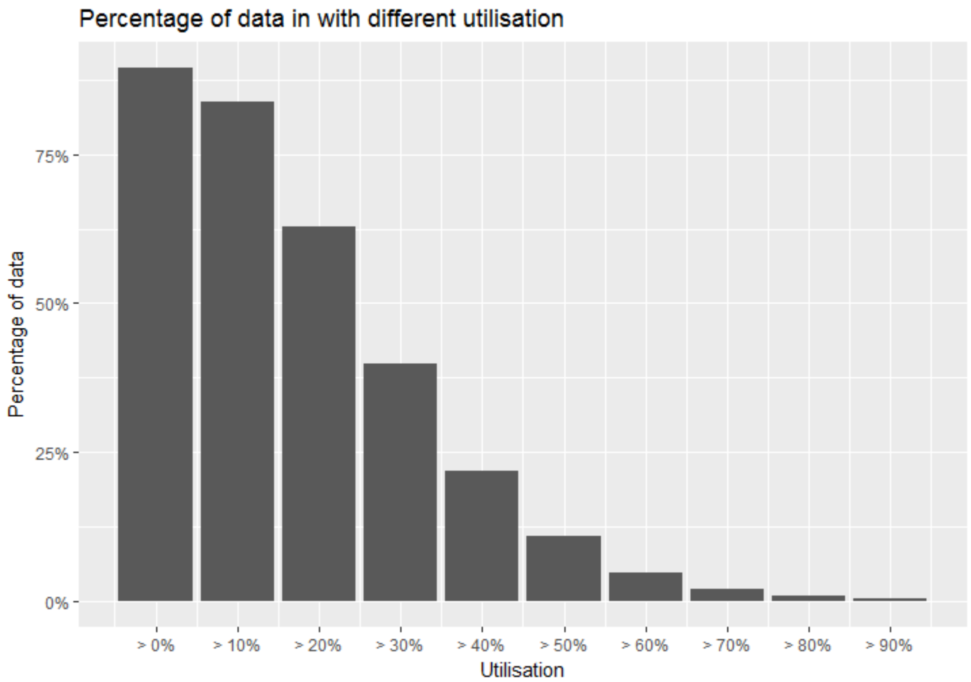

Figure 19: Amount of data available as the utilisation is increased

These limitations together have influenced the results of this study, potentially preventing more conclusive results to be found. However, the methodologies developed can provide significant insight to the thermal behaviour of transformers, and this analysis can be updated when more appropriate data becomes available, for example by completing the further work described in the bullet points above as far as possible within the Celsius project.

4. Thermal Rating Influences Analysis

The overall goal of the first phase of Celsius was to provide improved rating values for all sites that were informed by the most appropriate temperature factors. As ratings are considered to be conservative for most applications, more informed ratings would provide an increased capacity in distribution substations, enabling connection of low carbon technologies and avoid the costly alternative of reinforcing the substations and the network.

Following the conclusions from the Temperature Factors Study detailed in Section 3 of this report, it is not possible at this time to draw out definitive improved ratings. However, a potential approach to completing this work is described below, for use when the Temperature Factors Study has been updated and if more definitive conclusions can be drawn.

4.1 Converting Temperature Profile to Ratings

This section outlines the process that the project intended to take to convert the temperature profiles for the maximum operating temperature into ratings in KVA. ENW's code of practice provides maximum hotspot temperature for a transformer depending on the load profile, defined simply as continuous, cyclic, and emergency. There are seven load types, but the analysis will only consider continuous and normal cyclic load types as they cover the non-emergency situations.

Load Type	Temperature Criteria		Pre-fault loading
	Ambient temperature (°C)	Maximum hot spot temperature	
Continuous	Weighted average	98°C	-
Normal Cyclic	Weighted average	105°C	-
Long-Time Emergency Cyclic	Weighted average	115°C	-
3 day Emergency Cyclic	Absolute peak	125°C	-
3 hour Emergency	Absolute peak	125°C	Continuous rating
30 minute Emergency	Absolute peak	130°C	Long-term emergency rating
3 minute Auto Switching*	-	-	Any

Table 6. Maximum hotspot temperature from the load type, taken from the Electricity North West Code of Practice 382; Transformer Ratings

ENW's code of practice also provides information about the impact of ambient temperature on the assumed rating. Table 7 shows the percentage of the nameplate rating that can be assumed at different ambient temperatures. The combination of load type and weighted average ambient temperature will allow determination of ENW % Rating for each transformer (see table 7).

Load Type	Ratings (% of ONAN rating)			
	30°C	20°C	15°C	0°C
Continuous	90	100	105	120
Normal Cyclic	100	105	110	125
Long-Time Emergency Cyclic	105	115	120	135
3 day Emergency Cyclic	115	125	130	140
3 hour Emergency	120	130	135	150
30 minute Emergency	130	135	140	150
3 minute auto switching	150	150	150	150

Table 7: Transformer % rating by load type and ambient temperature, taken from the Electricity North West Code of Practice 382; Transformer Ratings

The thermal rating of a transformer is limited by the maximum operating temperature. The true rating is obtained by finding the value of load at maximum hotspot temperature. The improvement in transformer rating can be obtained by observing how far over ENW's code of practice rating, the true rating is. An illustration of this can be seen in the graph below.

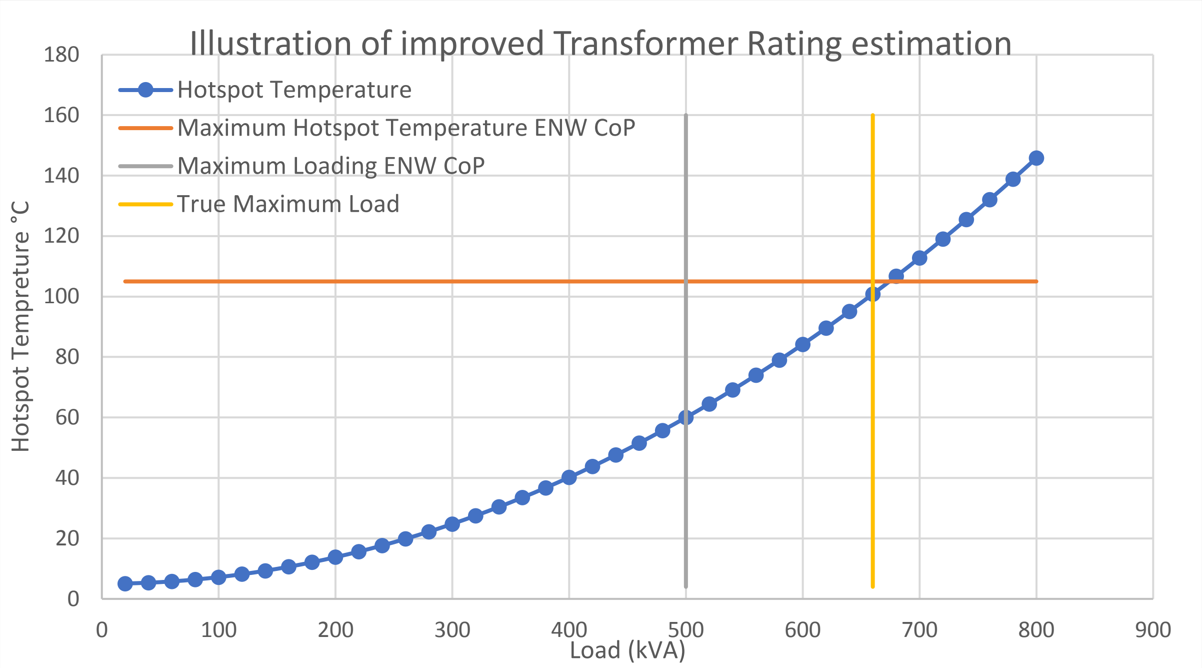

Figure 20:Improved Transformer Rating Calculation

Figure 20 shows an illustration of a Hotspot Temperature relationship with load. The shape and position of this profile would be influenced by the asset environment, for example, being less steep for outdoor substations in winter reflecting the increase in natural cooling of that environment.

The maximum hotspot temperature is given in the ENW code of practice. The profile can be used to determine the estimated rating, which is the estimated load at which this temperature is reached. On figure 20, this is represented by the yellow line.

4.2 Conclusions

Unfortunately, from the results obtained in section 3, it is not possible at this stage to provide recommend improved thermal ratings for all the sites selected for Celsius. The analysis shows that the load on the asset and the building types are major environmental factors that have an impact on the hotspot temperature (see figure 13 and 14). However, it is not possible to build a model that incorporates these major factors to accurately predict the hotspot temperature in different environments and locations and hence to refine site ratings.

As described in Sections 2 and 3, there are significant insights that have been developed through the analysis of the Celsius data. A key next step is to discuss the findings of this work with Electricity North West and agree the way forward for the completion of the Phase 1 analysis.