

BACHELOR

Visualisation of on-campus energy consumption

Bouwman, Nora M.

Award date:
2023

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Visualisation of on-campus energy consumption

Final report
Semester B 2022-2023

Nora Bouwman *

Eindhoven University of Technology

1 INTRODUCTION AND PROBLEM STATEMENT

Energy consumption at the campus of the Eindhoven University of Technology (TU/e) is high and it is reaching the limits of what the university can use. On Christmas Eve (24/12/2022), a baseline measurement was performed. At baseline TU/e uses 40% of their contractual energy usage. The university has a problem for two reasons. First, the university already went over their contractual limit several times. If this does not change, they will be blocked from the power grid. Second, the university wants to expand which is not possible as they are already consuming most of their contractual energy usage.

There is a lot of energy consumption data available, but at this moment, there are hardly any insights into this data. The goal of this project is to make a system that can help the Real Estate Team (the department responsible for energy consumption control) find what causes these enormous amounts of energy usage. The system will serve two purposes. First, it will help gain insights in which devices/buildings/et cetera causes these peaks. Second, the tool will help explore where the base load comes from. Which results in the following research question:

How can we use visualisation to gain insights in hierarchical time series of power consumption data of the TU/e?

This report is organised as follows. First, we present related works on peak detection in hierarchical time series data and visualisation techniques on hierarchical time series. Then, our methodology is described. We provide our result and findings on the basis of a Use Case. Penultimately, those results and findings are discussed. Finally, we close the paper with our conclusions and ideas on how to extend this work in the future.

2 RELATED WORK

Various visualisations have been introduced in the energy domain considering hierarchical times series.

Cuenca et al. [1] suggested to use MultiStream, a streamgraph approach to explore hierarchical time series. In their solution, the controller can link the overview and the multi-resolution view. The notion that the controller can use graphical zooming, semantic zooming and have the interaction technique of "overview + detail" and "focus + context", will be taken into account further in this report. However, their streamgraph approach will not be scalable to all university's assets and will miss the detail for the investigation on the hierarchical time series data.

In their paper, Wijk et al. [5], applied clustering and calendar-based visualisation of time series data. With this approach, they achieved

a presentation that provides a quick insight into both standard and exceptional patterns. Clustering is helpful to gain overview, but since clustering is performed, peak detection of the energy consumption is not possible as the data has to be aggregated. Such a calendar-based visualisation, without clustering, could help Real Estate observe the days that the university was close to its contractual limit.

Li et al. [2] introduced a node-linked technology to show hierarchy structures, where all visualisations, including a line chart, are plotted for all children of one parent. The approach Li et al. take is useful for the comparison of the leaf nodes of the same parent over a selected time period. The disadvantage of their visualisation is that it is not possible to compare leaf nodes of different parents. Therefore, their approach will likely not be useful for peak detection but might help Real Estate gain some insights.

LiveRAC presented by McLachlan et al. uses a reorderable matrix of area-aware charts [3]. LiveRAC on the one hand might have useful futures of multiple views and colouring, but on the other hand might be overly complex as they have eleven monitored parameters, while the energy consumption data only has one. In addition, their solution only visualises one day of time-series. This might be an addition to our tool, for daily monitoring, but gives no insights in historical data for a period longer than a day.

The concept introduced in Wills' paper [6] can be useful for finding which assets on average (or another aggregation) use the most energy. In addition, in his paper he uses the population of the Earth as example, meaning that this approach is scalable to 195 countries. This notion will not be helpful in detecting peaks on certain times but can still be used to determine the large consumers over a longer time period.

To conclude, the calendar-based approach of Van Wijk et al. [5] is used to find the days in which peaks in energy consumption have occurred. The concentric circle from Wills [6] is used to find the assets that result in peaks in energy consumption. What the dashboard needs in addition is an exploration tool where the user can select different nodes for a certain period.

3 METHODOLOGY

In order to gain insights in the energy consumption, the following research methodology. This section starts with Data and Task Analysis according to Munzners book [4]. Secondly, the data preparation that was performed is described. Thirdly, a dashboard was created that helps detect peaks and help with exploration. Fourthly, the visual encodings and interaction idioms in the dashboard are discussed. Lastly, evaluation was performed on a Use Case. The Use Case was presented to the Real Estate team at the TU/e. The used methodology can be found in Figure 1.

* e-mail: n.m.bouwman@student.tue.nl

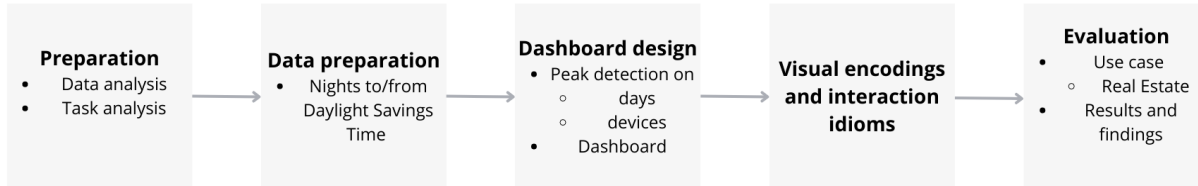


Figure 1: The research methodology. First, Data and Task Analysis was performed. Secondly, the data was prepared in order to start visualising the data. Thirdly, a dashboard was created that helps detect peaks and helps with exploration. Fourthly, evaluation of the visual encodings and interaction idioms in the dashboard. Lastly, evaluation on a Use Case and presentation to the Real Estate Team of TU/e

3.1 Data Analysis

The Data Analysis consists of the Domain Data Specification and the Data Abstraction.

3.1.1 Domain Data Specification

Two data sets were provided: the energy consumption data (main data) and weather data from the Royal Dutch Meteorological Institute (KNMI). The energy consumption data ranges from 31 December 2018 until 30 November 2022. The values of the assets are measured every 15 minutes. The data is hierarchical, i.e., a tree structure with the root node being a time series of the total energy consumption on campus. The root has three children nodes: Energy Station 1, Energy Station 2 and Enexis. Energy Station 1 provides the most energy. These energy stations have multiple loops, the loops go to a building and then to a device. In total, the data consists of 240 assets. A simplified version of this hierarchical structure can be found in Figure 2.

In addition to the main data set there is weather data available. This data is provided by the KNMI and has hourly measurements on various weather variables from 01 January 2011 until 18 January 2023. We kept the data for the dates 31 December 2018 until 30 November 2022, to fit the date range of the energy consumption data. For this project only the timestamp and the temperature variable from the weather data set were kept. We chose the temperature variable because we hypothesise that this is a key determinant of central heating energy consumption on campus.

3.1.2 Data Abstraction

For the Data Abstraction, we first considered the data and data set types. Then, we considered the attributes. Both were performed according to Muzners guidelines [4].

Data set

The main dataset can be seen as **atree**. A tree consists of **items (nodes)**, **links** and **attributes**. Every node is a **table** consisting of **items** (rows) and **attributes** (columns). This table is time series data, with a timestamp and energy consumption values (in kWh) as the two columns. The dataset availability is **dynamic**.

The weather dataset is a **table**, consisting of **items** and **attributes**. This table is time series data, with a timestamp and temperature (in °C) as the two columns. The dataset availability is **dynamic**.

Attributes

As concluded in the previous paragraph, the three attributes are timestamps, energy consumption value (in kWh) and temperature (in

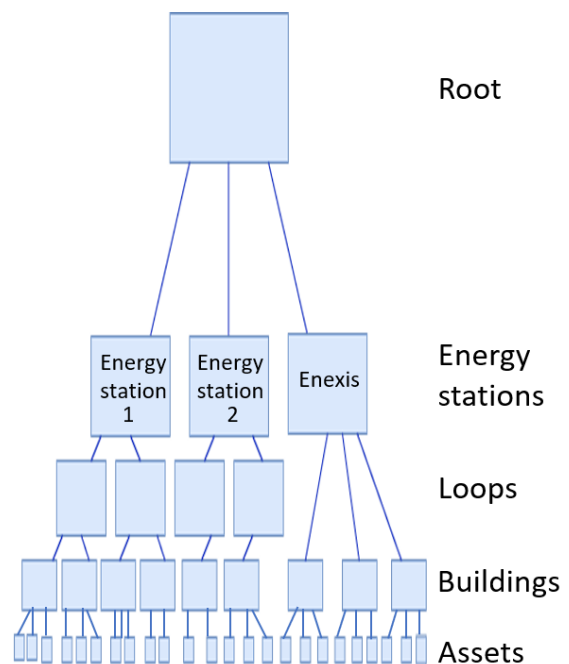


Figure 2: A simplified visualisation of the hierarchical structure of the main data set: the root with three energy stations; energy stations 1 and 2 having loops that go to buildings and then devices, and Enexis that immediately goes to buildings.

°C). The type of the timestamp attribute is **ordered, quantitative**. The ordering direction is **sequential**. The type of the energy consumption value (in kWh) attribute and the temperature (in °C) attribute are **ordered, quantitative** as well. The data abstraction of the attributes can be found in Table 1.

Attribute type	
Categorical	
Ordered	Timestamp, Power consumption (in kWh), Temperature (in °C)
Ordering direction	
Sequential	Timestamp
Diverging	
Cyclic	

Table 1: Data abstraction. The attribute timestamp is ordered and sequential. The attribute power consumption (in kWh) is ordered.

3.2 Task Analysis

The Task Analysis consists of the Domain Specific Task and the Task Abstraction.

3.2.1 Domain Specific Task

We investigated the research question via the following sub tasks and sub questions, dividing them into categories concerning peak detection and concerning data exploration.

Task concerning peak detection:

T1 Find the days in which peaks in energy consumption have occurred

Q1 On which days did peaks in energy consumption occur?

T2 Find the days in which the energy consumption per hour was close to reaching its limits. We defined the days when there was at least one occurrence of energy consumption per hour exceeding 7500 kWh (the contractual limit is 8000 kWh) per hour).

Q2 Which days was the University close to reaching its limits?

T3 Discover the correlation between peaks in energy consumption and the weather

Q3 Is there a correlation between the weather and the energy consumption?

T4 Find out if there is a pattern in the days in which peaks in energy consumption have occurred

Q4 What is the pattern in the days in which peaks in energy consumption have occurred?

T5 Find the assets that are large energy consumers.

Q5 Which assets on-campus result in peaks of energy consumption?

Task concerning data exploration:

T6 Select a time period

Q6 What happened in this period?

T7 Select an asset (energy station, loop, building, device)

Q7 What happened in this asset?

Tasks	Actions	Targets
Find the days in which peaks in energy consumption have occurred	Discover	Extremes
Find the days in which the energy consumption per hour was close to reaching its limits	Discover	Extremes
Discover the correlation between peaks in energy consumption and the weather	Discover	Correlation
Find out if there is a pattern in the days in which peaks in energy consumption have occurred	Discover	Features
Find the assets that are large consumers	Browse	Extremes
Select a time period	Lookup	-
Select an asset (energy station, loop, building, device)	Lookup	-

Table 2: Task abstraction. Every task has an action and a target, in this table all task are task are provided with an action and a target.

3.2.2 Task abstraction

The task abstraction is explained in Table 2.

3.3 Data preparation

The provided data required minimal additional preparation before it was utilised effectively. The only problem was the data during the nights in which times changed to and from Daylight Savings Time. The reverting to standard time in 2020 had the following affect, see Figure 3. As a solution, the data on the days there was a change in time was replaced by the data on the previous day for the hours 00:00 until 07:00.



Figure 3: Night in 2020 in which the time was reverted from Daylight Saving Time to standard time. This figure shows that the data does not follow its usual pattern.

3.4 Dashboard

The created dashboard uses three visualisation methods to create a transparent overview of the recorded energy consumption of all assets on the TU/e campus. It is organised in such a way to intuitively specify periods in time and assets of interest. The dashboard can be found in Figure 4. How the components work and relate to the tasks is explained in subsection 3.5.

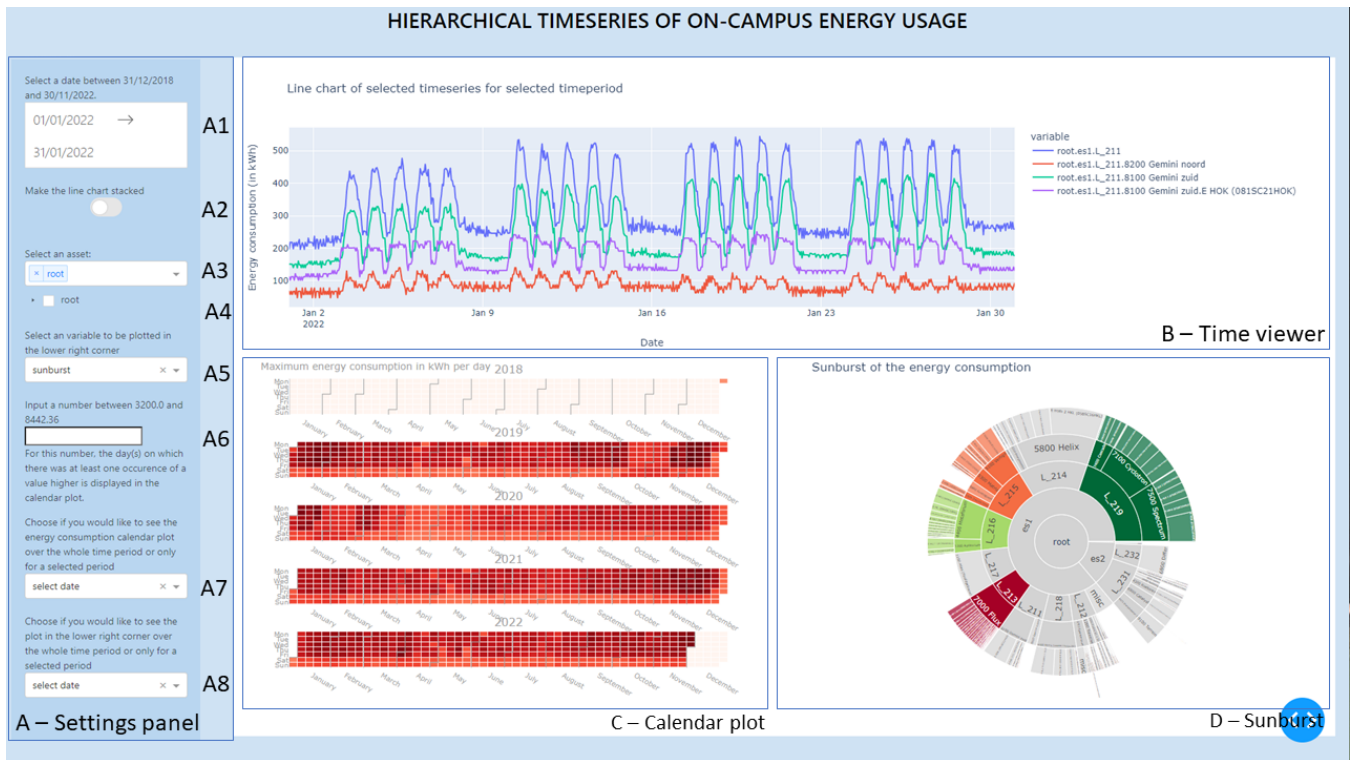


Figure 4: The dashboard. Component A enables the selection of options and input to the dashboard. It allows the user, among others, to select the date and select an asset. Component B plots the time series for a given date range and chosen asset(s). This lets the user explore the wanted data. Component C helps the user find the days in which peaks in energy consumption. If used in combination with component A6, it can also be used to find the days in which the energy consumption per hour was close to reaching its limits. Component D has two functionalities: (i) help the user find the assets that are large consumers, and (ii) summarise the weather data, similar to component C.

3.5 Visual encoding and interaction idiom

In this section, we contemplate the visual encodings that were used for the tasks abstracted in subsection 3.2, and explain the reasons these encodings are chosen. For the following tasks: "select a time period" and "Select an asset", interaction idioms are implemented.

3.5.1 A - Settings panel

The goal of this component is to let the user choose the settings and therefore control what they can see at the right side of the dashboard. The input panel has eight components. In A1, the dashboard enables the user to select a time range within the period (T6). In A2, the user can choose if they would like to see a line plot or an area plot. In A3 and A4 the dashboard allows the user to select one or more assets (T7) if they chose the line plot or area plot, respectively. The data plotted in component C, either energy consumption or weather data, can be chosen in A5. In component A6 the user can choose a number to find the days in which the energy consumption per hour was close to reaching its limits (T2), (T4) by giving a number. In A7 and A8, the user can choose if they would like to see the calendar (A7) and sunburst (A8) plot for the whole date range or only the selected date range. In A9, the user can choose a loop to input a threshold number for. Component A10 allows the user to input a threshold number for the chosen loop.

A1 For this component, we use a calendar in combination with input, such that the user can choose the date range. The used visualisation can be found in Figure 5.

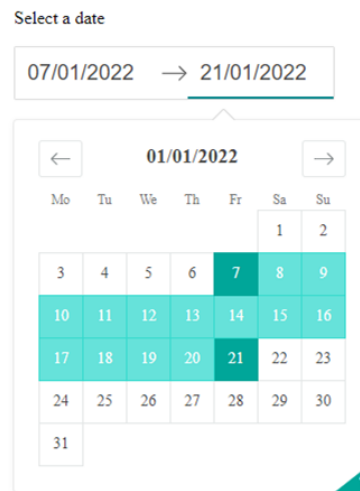


Figure 5: Selection method for task (T6) "select a time period". In this calendar the user can select a start and end date, for which they would like to see the visualisations.

A2 This component enables users to switch between a line chart and an area chart. The line chart presents cumulative asset values over a selected time period, providing a comprehensive overview. On the other hand, the area chart visually depicts the distribution of assets relative to a specified total, offering insights into asset allocation.

A3 and A4 We considered multiple possibilities for task (T7). The first visualisation that was considered was a tree and if an asset was selected, that file was plotted for the date range selected at A1. However, since there are 190 leaf nodes, this representation would not work as it is hard to scale to so many nodes and still be readable.

Therefore, the following solution was introduced. The user can scroll or type in a select box and sees all options in a drop-down menu, the interaction method can be found in Figure 6.

This option has a down-side however that the hierarchy is not



Figure 6: One selection method for the task "Select an asset". The desired asset(s) can be chosen in a drop-down menu, the chosen asset(s) will be plotted.

visible enough. For that reason, another solution was proposed. In this solution, see Figure 7, the user can click one or multiple assets to be plotted in component B. The disadvantage of this approach is the fact that selecting a parent result in all the children immediately being plotted. In the case where the root is selected all its 239 children and the root itself would be plotted, while this is not necessarily desirable.

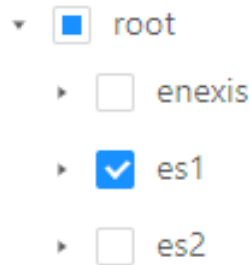


Figure 7: Another selection method for the task "Select an asset". The desired asset(s) can be chosen in a drop-down menu, the chosen asset(s) will be plotted.

At this point, the decision had been made to use Figure 6 for the line plot and Figure 7 for the area plot. For the line plot, you can select a node to be plotted without plotting its children. For the area plot, you can now see the hierarchy. In 3.5.2, an explanation in more detail why this decision was made follows.

A9 and A10 Figure 8 shows the bottom two options of A - setting panel. Besides the two options, the user gets a reference dictionary. The values of the reference dictionary contain the names of loop assets as keys and the values are the average consumption of that time period. The user can change this dictionary by selecting a key (A9) and changing its value (A10).

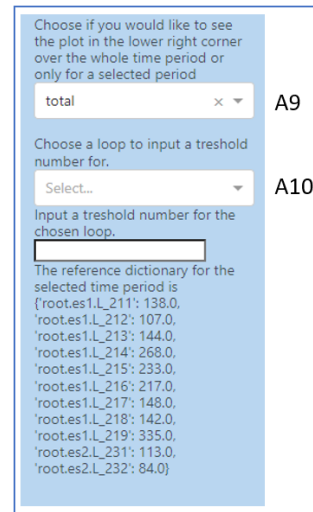


Figure 8: Bottom of A - settings panel. The last two components allow the user to select a loop and input a threshold number.

3.5.2 B - Time viewer

The goal of the time viewer is that the user can inspect the usage and patterns for a selected time period and the chosen assets. For this component, we employ a line/area plot based on the user's preference, which can be selected in component A2. The primary objective of the line chart is to facilitate a comparative analysis of asset values. Alternatively, by utilising an area chart, the utilisation of assets can be compared to the overall total.

The data given is per 15 minutes, while the university's limit is defined per hour. To find out which days were close to reaching 8000kWh we needed to re-sample the data to hourly intervals. Only when the user plots less than 3 days, the granularity of data should be every 15 minutes. Therefore, it was implemented that if the time difference of the dates selected at component A1 is bigger than two days, the hourly interval data is plotted, else the data with 15 minutes is plotted.

Figure 9 shows an example of a plot that can appear, based on the user input, in the Time viewer component. This plot has been made with a library such that the interaction idioms zooming, and panning have been automatically included. In Figure 9, it can be seen that all assets more or less follow the same pattern. The fact that in the first week of January there was a vacation can be seen when comparing it to the other weeks. There is a week/weekend and day/night pattern.

In addition to a line plot, we thought it would be in addition to help with task T5 if we could have an area plot. The disadvantage of an area plot is that it sums all the values. This is a problem because a parent is already the sum of his children. Therefore, we wrote a recursive algorithm to fix this problem. This algorithm considers two cases. If all the selected nodes are from the same ancestor, it plots the direct children of ancestor node. For example, if the root is selected its three children (the energy stations) will be plotted. The other case, where not all selected nodes have the same ancestor, the selected nodes will be plotted.

After the user has selected a date range in component A1, set component A2 to True, and one or more assets in component A4, the area plot for that date range and those assets will be plotted. If the user would choose 01 January 2022 until 31 January 2022

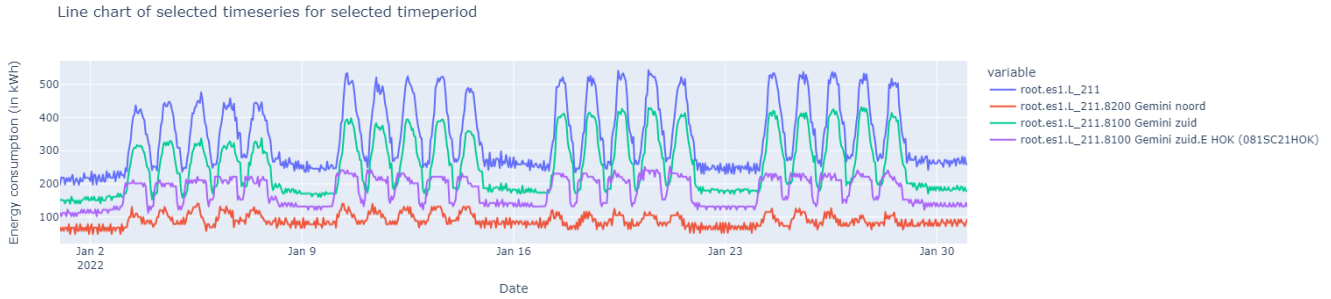


Figure 9: Plot to display the power consumption for a given date range and asset. Here the energy consumption of L_211 and two of its buildings (Gemini-Noord and Gemini-Zuid) and a device of a building are displayed on the y-axis for the month January 2022 (x-axis). In this figure, there are 5 peaks and then two days of flat lines, distinguishing working and weekend days. Besides this pattern, it can be seen that working days in the vacation (first week of January) have lower peaks than working days outside of the vacation (last three weeks of January).

in component A1 and would choose "root" in component A4, Figure 10 will be plotted in component B. This plot has been made with a library such that the interaction idioms zooming, and panning have been automatically included.

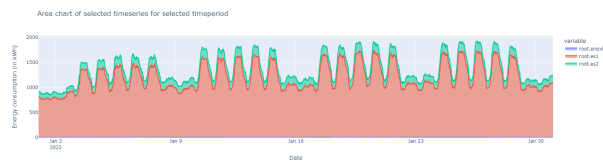


Figure 10: Plot to display the time series for a given date range and asset. Here the energy consumption of the roots children is displayed on the y-axis for the month January 2022 (x-axis).

In Figure 10, the vacation at the first week of January is again quite visible. Also, this plot also follows the weekday/weekend and day/night pattern as can be seen in Figure 9. In addition, it is visible that Energy Station 1 usage the most of the three energy stations.

3.5.3 C - Calendar plot

The goal of this visualisation is to find the days in which peaks in energy consumption have occurred (T1). To achieve this, we use a coloured calendar similar to Wijk et al. [5]. Since the power consumption is an ordered attribute, we use colour luminance or colour saturation as a channel. With days were peaks occurred having a higher luminance or saturation than days that had lower energy consumption. The disadvantage of this calendar is that the data should be aggregated to days, meaning details about the distribution during the day will be lost. In our case, it was decided to re-sample the data from every 15 minutes to hourly and use the maximum hour value of that day to be plotted. The resulted plot can be found in Figure 11.

If in A6 no value is chosen, and A7 is still set to total Figure 11 appears at component C.

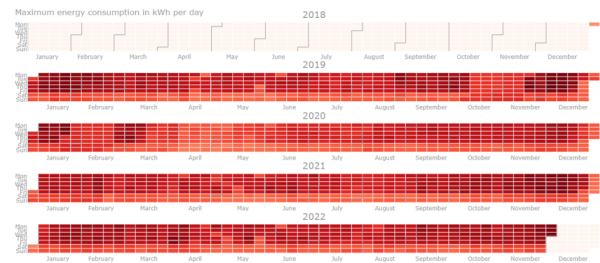


Figure 11: This figure shows the maximum hourly energy consumption per day of the last 4 years. This enables the user to find the days in which peaks in energy consumption have occurred (T1). Higher saturation indicates higher usage.

In component A6 the user can choose a number to find the days in which the energy consumption per hour was close to reaching its limits by giving a number. Again, the data is aggregated to hourly intervals. Then, it only keeps the rows where the root asset is higher than the given number. After, it takes the maximum value of that date. The limit of the University's energy consumption is 8000kWh, if a value of 7500 (kWh) is chosen in component A6, in component A5 of the dashboard the variable 'energy consumption' is chosen, and A7 is set to True Figure 12 is shown in component C.

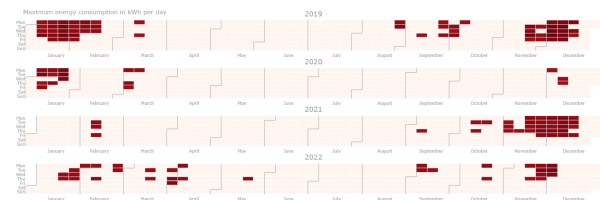


Figure 12: This figure shows the maximum hourly energy consumption per day for the days that had an hourly energy consumption higher than the number given in component A6. This figure tells the user that mostly during the weekdays in the winter months the university was close to reaching its limits.

For (T4), based on the information found in Figure 11 days found in Figure 12, we can find a pattern that helps answer the following questions:

- What is the difference in peaks on weekdays compared to the weekend?

- What is the difference in peaks in the summer compared to the winter?

In Figure 12, it can be seen that there are no peaks in the weekend as well as that most peaks occur in the winter months (January, February, March, April, September, November, December).

3.5.4 D - Sunburst

In component D either the sunburst plot or the weather calendar plot can be chosen. It was decided to plot only one of these at a time for spatial reasons. We deemed it more important for the user to see both plots with enough detail, than to see both at the same time. Additionally, there is no task where you need to compare information from these visualisations.

Sunburst

The goals of this visualisation are to find the assets that are large consumers (**T5**) and detect when a loop is over or under consuming. For this task, we use sunburst plots. These can be explained as "concentric circles, with radial depth, spatial position, and size showing tree depth and radial angular spatial position showing link relationships and sibling order" [4]. This is chosen because it is scalable to the 240 assets and in one glance it can be seen which assets are large consumers.

For this visualisation, every asset can only have one value, so the data needs to be aggregated. It was decided to take the sum of all assets, such that you can see which values cumulatively used most energy over the time period. In component A8, the user can choose between `total`, in that case the values are summed over the whole data period and the sunburst will be blue. In the case of `select date` the values are summed only over the period selected in component A1 and are coloured based on their usage.

At the core of the sunburst, you will find the root (all assets combined), the layer around the core are the energy stations, then a layer for the loops, the layer after for the buildings and in the outer two layers the buildings are visualised. A larger radial angular spatial position indicates higher usage for that asset.

Based on the input from A9 and A10, a dictionary with values arises. We consider five cases. In the following equations, Y is the average consumption for the time period selected at A1, X is the average consumption for the whole time period and σ is the standard deviation of the whole time period. The first case results in a dark red loop and the second case results in an orange loop. In the third case, everything is normal ranges and is grey. When an assets usage less than it normally does (*iv*) it becomes light green, when it uses much less then it normally does (*v*) it becomes dark green.

$$(i) \quad Y > X + 2\sigma$$

$$(ii) \quad X + \sigma < Y < X + 2\sigma$$

$$(iii) \quad X - \sigma < Y < X + \sigma$$

$$(iv) \quad X - 2\sigma < Y < X - \sigma$$

$$(v) \quad Y < X + 2\sigma$$

Since X ranges over almost four years and there is a lot of fluctuation in the data, other anomaly detection techniques to find the loops that over- and under consume should be considered in the future as it is outside the scope of this project.

Sunburst of the energy consumption



Figure 13: Visualisation for "Find the assets that are large consumers". This sunburst plot shows the assets with the highest consumption, using the radial angular spatial position. The colour is based on how much energy the loop consumes compared to the average. Red colours indicate over consumption, while green colours indicate under consumption.

In Figure 13, it can be seen that when summed over the whole time period Energy Station 1 is the largest energy station. L.219 is the biggest loop and Helix is the largest building in terms of energy consumption.

Calendar plot with weather variable

To discover the correlation between energy consumption and weather data, we made a plot similar to Figure 11. Figure 14 will appear at component D - Sunburst if in component A5 of the dashboard the variable 'weather' is chosen and A7 is set to `total`. Figure 14.

As temperature is an ordered attribute, we should use colour luminance or colour saturation as a channel. With days with higher temperature having a higher luminance or saturation than days with lower temperatures. For this visualisation, every day can only have one value, so the data needs to be aggregated. As there can be quite some variances in temperature over the day, it was decided to take the median temperature for every day. The resulted plot can be found in Figure 14.

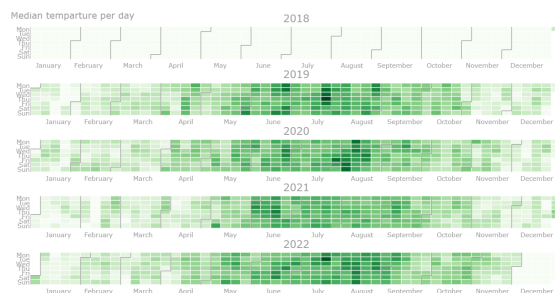


Figure 14: Visualisation for task **T3** "Discover the correlation between peaks in energy consumption and the weather". This figure shows the medium hourly temperature per day for the whole data period. The dark area in the middle represents the summer period.

Based on the input at component A8 the user can plot Figure 11, Figure 12, Figure 14 for the whole time period or only for the period selected at component A1.

In Figure 14, it can be seen that the winter months have lower luminance, meaning it is colder in those months and higher

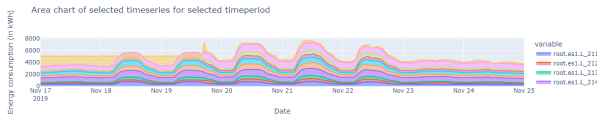


Figure 15: Area plot for Energy Station 1 between 17 and 25 November 2019. In this plot it can be seen that until the 19 November the miscellaneous category had high usages.

luminance, meaning it was warmer in these months.

4 USE CASE

In the following section, we demonstrate the capabilities of the dashboard. In addition, we answer the questions from subsection 3.2 and report noteworthy discoveries that arose from this dashboard. We started with broad data analysis. To do so, we plotted the root for the whole time period, we found that on 19/11/2019 around 18:00 a peak occurred. We then switched to the area plot to see which energy station was responsible for the peak. After further inspection, we found that the miscellaneous categories of the Energy Station 1 were responsible for that peak. Meaning that it is still not clear from which exact node was responsible as the miscellaneous category is the difference between the measured value of Energy Station 1 and all measured nodes from Energy Station 1.

Using the combination of the calendar view and the time viewer, the days in which peaks in energy consumption have occurred can be found, thus fulfilling (T1).

With the calendar view as in Figure 12, we completed (T2). It can be seen that most peaks occur in the winter months on weekdays. There was not a single occurrence of energy consumption per hour of the root being higher than 7500 kWh in the weekend. Over the whole time period there are only three occurrences in the summer months (May, June, July, August). Two of them are 26 and 27 August 2019. When looking at the calendar plot, we saw those day were indeed hot (maximum temperature of 32.2 °C and 33.8 °C respectively). However, in week 29 of 2019 it had been even hotter (over 34 °C for 4 days straight) and that did not result in peaks. The other one was 19 May 2022. There was a high increase

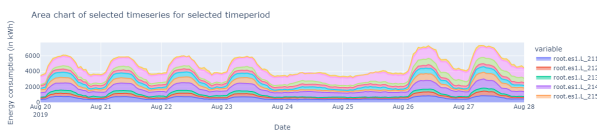


Figure 16: Area plot for Energy Station 1 in between 20 and 28 August 2019. This plot shows the last four days of a vacation, two weekend days, and the start of the academic year. It can be seen that the peaks of the vacation are much lower than those at the beginning of the academic year.

in the afternoon with the peak of the day being around 14:00. To determine what were the large consumers at this point, we first inspected the division between energy stations and saw that most of the consumption came from Energy Station 1, when plotting the loops. We saw that loop 219 used 1367 kWh at 14:00 and the miscellaneous category was at 1316. Therefore, it is not clear why there was a peak at this moment.

When considering the correlation between weather and energy consumption (T3), the impression is drawn that when the median temperature is low (light saturation in Figure 14), the energy

consumption is high (dark saturation in Figure 11 and presences of squares in Figure 12).

Regarding T4, we conclude that there is a pattern in the peaks of the energy consumption. The pattern is that peaks occur in the winter months during the weekdays.

Sunburst of the energy consumption

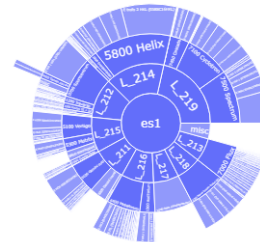


Figure 17: This sunburst shows the assets from Energy Station 1 for the whole time period. L_214 and L_219 are the biggest loops.

For T5, the sunburst was created. In combination with the area plot, so we can see the division of the total energy consumption over time. If we inspect Figure 13, we see that Energy station 1 is a large consumer. To, inspect that further, we solely focus on Energy Station 1, see Figure 17. There L_219 is the biggest loop, with 7500 Spectrum being its biggest child and then the device "E HOK (075SC19HOK)" being the largest child (using 3.8% of the total). However, Helix (from L_214) used more than Spectrum (24.582.436 and 16.973.235 kWh respectively) and Helix' biggest child "E trafo 2 HKL (059SC16HKL)" used 6.6% of the total.

As mentioned before, (T5) and (T6), are completed by component A1, A3 and A4 in the settings panel.

5 DISCUSSION

The process, visualisations and results will be discussed in this section.

Regarding the process, The Real Estate Team was asked its opinion on the current status on the dashboard two times. The first time we spoke with the team (T2) and (T3) were introduced. In addition, we considered different plots for (T5), namely: a sunburst, a tree map, a icicle cart, and a clickable stacked bar chart. When asked their opinion, they indicated that they prefer a sunburst.

When starting to experiment and documenting findings, we noticed a few bugs in a previous version of the dashboard. Previously, either the energy consumption variable or the weather variable could be plotted in C - Calendar plot. With this placement of calendar plots, it was not possible to compare the energy consumption and weather variable easily, making T3 not completed. Therefore, it was decided to leave the energy consumption calendar plot at component C - Calendar plot and letting the user choose if they want to see the sunburst plot or the calendar plot in component D - Sunburst.

Another missing part was the fact that the weather now only has the median and it would also be useful to have extra information, for example mean, minimum and maximum in the label. Thus, this was added.

Lastly, we noticed that the energy consumption calendar plot is given per hour and the values in the line/area plot are per 15 minutes. To remove inconsistency, we now plot Time viewer per hour unless less than two days plotted, in which case we still plot it per 15 minutes.

After these changes, we spoke with the Real Estate Team again. They mainly had positive feedback to the version of the dashboard that was presented. Although it was asked if the dashboard could have an additionally input field where the user can select a number for every loop. Then, for every loop it is checked how often they exceed their number. After, a gradient colour scale was implemented. The colour scale ranges from light yellow to dark red, where the loops that in comparison to the other loops did not exceed their number often are light yellow and loops that in comparison to other loops often exceeded their number are dark red. A disadvantage of this dashboard is that it does not facilitate comparing two events from the same plot at the same time. This means that the user often need to change the settings (or make screen shots, et cetera). As changing this did not fit our time frame it was included in 6.1.

In section 4, it can be seen that now peaks come from the miscellaneous category. This can mean two things: Energy Station 1 or its children are not precisely measured or there are assets on Energy Station 1 that currently have no meter. First, it should be determined which is the case. Then, based on those findings, they should either fix their measurements systems or add meters to the assets that currently do not have one.

Currently, only large devices are measured. To gain more insights in the data, it would be useful to have more devices metered. However, for this process one should consider the legal and ethical aspects of what it would mean if for example every students laptop charger would be metered.

6 CONCLUSION AND FUTURE WORK

We developed a tool to gain insights in hierarchical time series of power consumption data of the Eindhoven University of Technology. We identified user tasks from which appropriate visualisations and input options were selected. Next, we implemented these to a dashboard. The ability to compare different views of the data enables the user to find insights in the data. The settings panel equips the user to create the desired view. The time view enables users to detect patterns and trends over time. The calendar plot allows the user to obtain an overview of the maximum power consumption in the specified time period. The sunburst facilitates the user to find the assets that are large consumers. To conclude, by using a settings panel, a time viewer, a calendar plot and a sunburst, insights in hierarchical time series power consumption data of the TU/e can and are gained.

6.1 Future work

In the future, all components except for component A1 should be moved to their relevant visualisation as they only control one of the other components. Furthermore, it would be an addition if the user could compare two moments with each other. In addition, it would be nice to have brushing and linking between the Time viewer and the Sunburst. One could also consider dividing every square in the calendar plot to have the energy consumption in one half and weather in another half making comparison possibly simpler. Furthermore, it could be researched if the peaks often happen around the same time, and if so what time. Lastly, other anomaly detection techniques for the colour of the sunburst should be explored.

REFERENCES

- [1] E. Cuenca, A. Sallaberry, F. Y. Wang, and P. Poncelet. MultiStream: A Multiresolution Streamgraph Approach to Explore Hierarchical Time Series. *IEEE Transactions on Visualization and Computer Graphics*, 24(12):3160–3173, 12 2018. doi: 10.1109/tvcg.2018.2796591
- [2] N. Li, Z. Jiang, Z. Liu, and X. Meng. A method of hierarchical time-series data visualization. *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction*, 8 2013. doi: 10.1145/2493102.2493121
- [3] P. McLachlan, T. Munzner, E. Koutsofios, and S. North. LiveRAC. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 4 2008. doi: 10.1145/1357054.1357286
- [4] T. Munzner. *Visualization Analysis and Design (AK Peters Visualization Series)*. A K Peters/CRC Press, 1 ed., 12 2014.
- [5] J. Van Wijk and E. Van Selow. Cluster and calendar based visualization of time series data. *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99)*, 1999. doi: 10.1109/infvis.1999.801851
- [6] G. Wills. *Visualizing Hierarchical Data*, 2009.