UTRECHT UNIVERSITY

Department of Information and Computing Sciences

Applied Data Science master thesis

Finding human behaviour in hybrid heat pump power usage

Using predictive modelling

First examiner:

Laurens Stoop

Second examiner:

Ad Feelders

Candidate:

Ruben de Groot

In cooperation with:

Inversable BV

Intergas BV

July 9, 2023

Abstract

Climate change is causing countries to look for change and improvement in different sectors, including the heating industry. In the Netherlands, this is manifested, among other things, in a transition from the use of gas to electricity. Heat pumps could be a part of the solution to this problem, and the Dutch government has therefore initiated subsidies. However, not every house is suitable for a heat pump. For this reason, this thesis investigated the performance of hybrid heat pumps by searching for human behavioural patterns. The study demonstrates that by making use of predictive modelling, it is possible to forecast power usage trends. By doing so, we produced residuals with reduced influence of temperature, thereby enabling us to use Fourier analysis to look for human behavioural patterns. Although we found patterns that yielded indications, we could not give conclusive evidence that we found human behavioural patterns. However, we observed differences in how different heat pumps behaved under similar conditions. These findings suggest that hybrid heat pumps may not be suitable for every house included in this study. Further investigation is necessary to prove if our findings were signs of human behavioural patterns and how the differences in behaviour effects hybrid heat pump performance.

Preface

This master's thesis was undertaken as a part of the Applied Data Science program at Utrecht University. Two teams of master students, from Applied Data Science, collaborated with Inversable BV and Intergas Verwarming BV, which provided the team with the dataset collected for the "Demoproject Hybride". This project, initiated in November 2021, aims to gain valuable insights into the practical performance, savings, and applicability of hybrid heat pumps.

The thesis is based on the combined work of Sahar Pourahmad and Ruben de Groot. The overall aim was to find human influence within the temporal component of hybrid heat pump power usage. The methods applied by both students are interwoven and built on top of each other. This specific thesis focuses on the creation of a predictive model. The model removed the temperature influence in the time series so that Sahar could apply the second method in her thesis, which resolves around Fourier analysis. The combination of methods gives an insight into the possibilities of detecting patterns influenced by human behaviour and how this relates to differences in heat pump power usage between different devices.

We express our gratitude to Inversable BV and Intergas Verwarming BV for providing us with access to the dataset, which formed the cornerstone of our investigation. Furthermore, we extend our appreciation to our supervisors and mentors who offered their guidance and expertise throughout this research endeavour.

Jordi Beunk, Johanna Lems, Abdulhakim Özcan, Sahar Pourahmad, and Ruben de Groot collectively dedicated their time, skills, and efforts to this study, and we are proud to present the findings of our research in this master's thesis.

Special thanks

I would like to express my special thanks to Sahar Pourahmad and Laurens Stoop. Laurens for being a perfect guide throughout this research and Sahar for the wonderful collaboration. Sahar made invaluable contributions to the research, including providing the necessary research on Fourier analysis, contributing to the introduction, conducting research on outliers, and providing insights on assumptions for our methods used. This thesis would have been nowhere without the help of these two special individuals.

Contents

1. Introduction

1.1 Climate change

It is undeniable that we are all confronted with environmental challenges. In the past century, human activities have produced an artificial increase in the concentration of greenhouse gases in the atmosphere, resulting in the retention of the sun's energy within the Earth's system. The average surface temperature of Earth is estimated to rise between 2 °C and 6 °C by the end of the 21st century, and the rate of global warming has nearly doubled in the last 50 years [1]. The impact of global warming is far greater than just increasing temperatures. It has disrupted the natural water cycle, resulting in more intense rainfall, flooding, and drought in various regions. Global warming also affects rainfall patterns and causes rising sea levels, coastal erosion, and shifts in infectious disease ranges. The impact of rising greenhouse gas emissions on climate change is already evident and requires urgent action [1], [2].

By 2019, the concentrations of atmospheric carbon dioxide (CO_2) had reached levels higher than those observed in at least the past 2 million years [2]. As a result, international measures are being implemented to tackle climate change and its negative impacts. In 2015, world leaders of 194 parties, including the European Union, joined the United Nations Paris Agreement, committing to work together towards a net-zero emissions world [3].

1.2 Energy transition in the Netherlands

Among the countries of the European Union, the Netherlands has made notable progress on its transition to a carbon-neutral economy. The country aims for a swift transition to a low-carbon economy and has integrated greenhouse gas reduction targets into its energy and climate policy. In 2019, a climate agreement package was developed by the business community and civil society organizations [4]. The Dutch Ministry of Economics Agriculture and Innovation is aiming to reduce CO_2 emissions by 49% in 2030 compared to 1990 [5]. One of the key measures outlined in the Climate Agreement focuses on enhancing the energy efficiency of homes, and transition away from natural gas heating for new buildings, while also urging improvements in existing buildings to enable fossil-free heating methods [6].

A recent report from the Dutch Heating Industry (NVI), suggests that the $\mathrm{CO_2}$ reduction target in the built environment by 2030 can be accomplished by installing 1.7 million hybrid heating systems. This report emphasizes the potential of hybrid heat pumps, as a suitable and sustainable solution for millions of homes [7]. In May 2023, the government stated that there would be stricter requirements regarding the efficiency of heating installations, and heat pumps will become the minimum standard starting from 2026 [8].

1.3 Types of heat pumps

Moving away from fossil fuels towards renewable alternatives comes with an increase in electric-driven systems. One of these systems, designed to reduce gas usage by households, is the heat pump. There are multiple variants of the heat pump, with each having its own benefits and shortcomings. There are heat pumps with an outdoor unit that will extract heat from the outside air, while others will extract heat from below the ground [9]. This second option is generally more expensive since a part of the ground must be removed to install all the pipework that is needed to extract heat; therefore, most people will install an outside unit. A third option consists of the use of photovoltaic thermal collectors, typically abbreviated as PVT. These are panels that are laid down on the top of a house to collect heat; they are also used as solar panels.

1.4 Hybrid heat pump

The commonality between these different versions of the heat pump is that a house should be better isolated than what is seen as acceptable for a house that uses gas. This creates a problem for all households that want to save on gas and electricity but do not have all the prerequisites to install a heat pump [9]. Therefore, the Dutch government is experimenting with subsidizing hybrid heat pumps.

1.4.1 Hybrid heat pump components

A hybrid heat pump consists of two components, a smaller heat pump that runs on electricity and a smaller boiler that uses gas. As a result, the combined devices should have the benefits of both a heat pump and a boiler, namely stability and less energy use for heating a room, but also jump-start the heating of a room when it is cold, and the heat pump does not have enough power to heat the room quickly. This means that even if a house cools down more than a typical house that is isolated for the use of a heat pump, it can still benefit from a hybrid heat pump.

1.4.2 Savings

By implementing a hybrid heat pump, a reduction of around 20% in (CO_2) emissions from heating and hot water can be achieved, together with an approximate decrease of 80% in natural gas consumption [10]. A second advantage is that installation is generally not a big operation because most houses already have a boiler and so the needed pipes and connectors are already in place.

To investigate if hybrid heat pumps do indeed help to save power and reduce CO_2 emissions, the Dutch government needs more information about the performance of hybrid heat pumps in real-life situations.

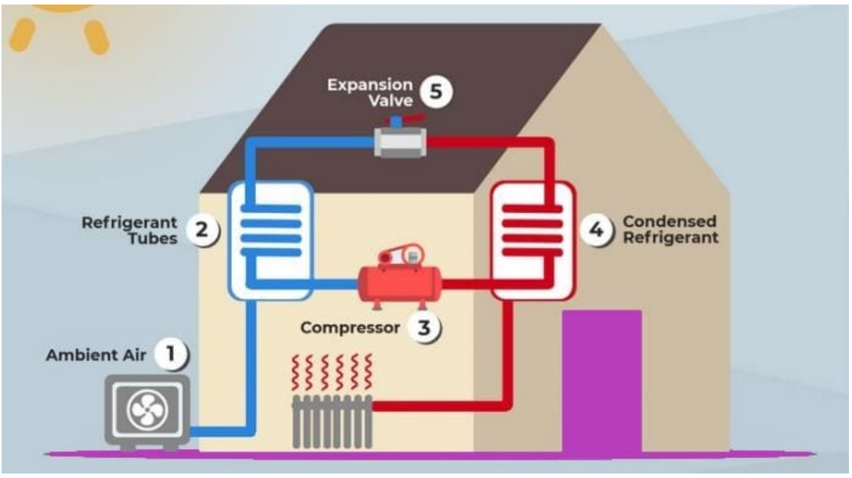


Figure 1.1: Illustration showing the components and sequential operation of a heat pump. Figure obtained from [11].

1.5 The heat pump explained

A heat pump works as follows (illustrated in fig. 1.1):

1. Heat is obtained from the air outdoors. It is directed or circulated over the heat exchange surface of the outer component of the heat pump.
2. The heat is sufficiently warm to make the liquid refrigerant inside the heat pump evaporate and transform into a gas.
3. This gas is then transported through a compressor, which boosts its pressure, resulting in an increase in temperature.
4. The heated gas is directed over the internal heat exchange surface. This heat can be either blown throughout the interior of the house or transferred to a central heating or hot water system.
5. As the heat is transferred into the house, the gas cools down, causing it to revert to a liquid state.
6. This cycle of reverse refrigeration repeats until the desired temperature is reached in your home, as set on your thermostat.

1.6 Previous research

There have been earlier projects about monitoring the performance of heat pumps in the Netherlands.

1.6.1 Installatiemonitor

In 2019 the project "Installatiemonitor" started. The goal of this project was to collect information about the real-life performance of both heat pumps and hybrid heat pumps. The project was a partnership between Enpuls, Gasterra, Gasunie, Liander, N-Tra, RVO, Stedin and Techniek Nederland and was carried out by consultancy firm BDH.

During this project 800 heat pumps were monitored until the 30th of June 2021, from which 450 were eventually analysed. They concluded that the release temperature of the heating system and surface of energy loss positively correlated with the energy usage of the heat pump. Additionally, a hybrid heat pump was found to significantly reduces CO_2 emissions and offers a financially attractive heating option [12].

1.6.2 Demo Project Hybride

This research is built upon the Inversable Demo Project Hybrid. The Inversable Demo Project Hybrid involves several organizations and institutions, including the Dutch Heating Industry (NVI), Technology Netherlands, the Ministry of the Interior and Kingdom Relations (BZK), the Ministry of Economic Affairs and Climate Policy (EZK), and the Netherlands Enterprise Agency (RVO). Utrecht University is also participating in this project. The current manufacturers participating in this project are Atag, Ferroli, Intergas, Nefit Bosch, Remeha, and Vaillant.

The project includes 200 participants who have either had a hybrid heat pump installed or will have one soon, starting from November 2022. Since these hybrid heat pumps differ from the ones previously installed in the Netherlands, mechanics underwent training to familiarize themselves with this new type of heating. Additionally, locating participants, despite the heat pumps being subsidized, has taken some time. Currently, Inversable has identified all 200 participants and continues to expand the data collection as heat pumps are installed and connected to a database [10].

1.7 Research question

Inversable suggested that it could be interesting to look at patterns of heat pump performance over time. Are there certain patterns to be found that influence the power usage of the heat pump? The "Installatiemonitor" research did not cover the temporal component of heat pump usage longer than a few hours, and the aspect of human behaviour was also not included. In this research, those two aspects are the core of the methodology.

In the end, a report about the findings of the performance of hybrid heat pumps will be produced for the Dutch government to see if it was worth the investment. If the demo was a success, more hybrid heat pump installations will follow.

We therefore chose the following research question: To what extent are human behavioural patterns present in hybrid heat pump performance?

1.8 Thesis structure

The rest of this thesis is structured as follows. In section 2 we will discuss what data we used and how we prepared it. In section 3 we discuss what methods we used, both for the predictive model as well as for analysing the residuals. In section 4 the results are presented. And finally, in section 5 we talk about our conclusions, start a discussion and highlight potential areas for future research.

2. Data

This chapter includes a description of the data followed by data wrangling, data exploring, data cleaning, and feature creation. Information about the data extraction can be found in the appendix A.2.

2.1 Data description

The time-series data consists of real-time measurements recorded in 169 houses. This could be 200 in the future, but as mentioned earlier, although all the 200 participants are found, not all of them have their heat pump installed yet. On average, there are 7 months of measured data per house, ranging from 20 days for the shortest duration to 15 months for the longest. During this period, various sensors were deployed in the houses, including the heating system sensor, heat pump sensor, boiler sensor, smart meter, and indoor climate sensor. In addition to data from the aforementioned sensors, local weather data was collected.

2.1.1 Sensors

A summary of these sensors is provided in fig. A.1. The 'time resolution' column in fig. A.1 indicates the frequency at which new measurements were recorded for each sensor. However, these time resolutions are more detailed than necessary for our analysis. Therefore, the data was aggregated on a daily basis. The aggregation techniques, indicated in the column 'Aggregation method' in fig. A.1, varied depending on the type of measurement. Here we provide a description of the aforementioned sensors,

The heating system sensor: Is responsible for monitoring the heating system within the house. It measures various parameters such as flow rate, and the supply and return temperature into the heating system. The energy and power of the thermal system are derived from this flow rate and supply temperature. The heating system sensor collects measurements every 5 seconds, providing detailed insights into the heating system's performance.

The heat pump sensor: Is positioned at the heat pump unit. It specifically measures the energy and power supplied to the heat pump. However, it does not directly measure the amount of energy converted by the heat pump and transferred to the heating system.

The boiler sensor: Monitors how much energy and electricity the boiler uses. It should be noted that a small portion of this energy could be used to heat water for showers or for other purposes besides heating systems.

The smart meter: Is responsible for reading the smart meter readings for the house. It transmits a "telegram" every 60 seconds for power readings and every 10 seconds for energy and gas readings. This enables real-time monitoring and analysis of energy consumption patterns. Please note that the power consumed and delivered values represent the combined total for all three phases of the network. Similarly, the energy consumed and delivered accounts for both high and low tariffs. Energy can be delivered by a house when it generates energy itself, such as through the use of solar panels.

The indoor climate sensor: Is typically placed in the living room and is primarily used to measure the indoor temperature and humidity. It provides valuable data for understanding the indoor climate conditions within the house.

Local weather data: Is collected via The Royal Netherlands Meteorological Institute (KNMI) and linked with each house by Inversable.

Although metadata of each house was also provided, this was not used in this research and therefore not described.

2.2 Data wrangling

A list of active devices given by Inversable was used to see what devices were usable. Devices are not activated at the same time, meaning that they

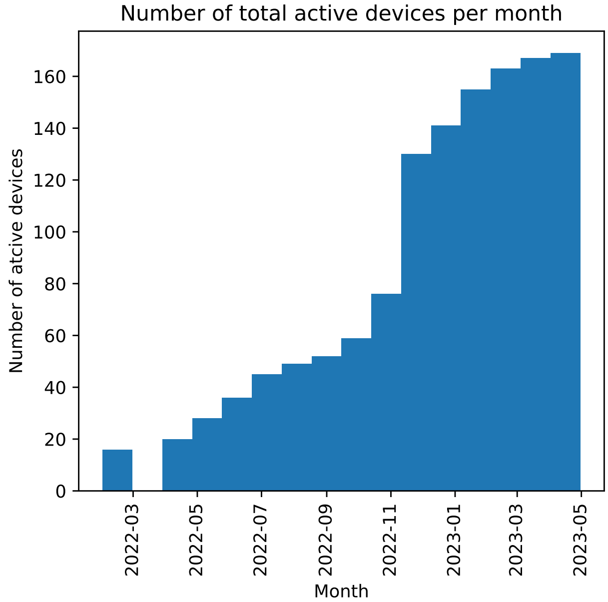


Figure 2.1: Number of active devices per month. The figure shows the gradual increase of the total active devices from February 2022 until May 2023

are not well comparable. If one device was active during the winter of 2022 and the other one was not, it has a different average heat pump power usage over the whole period. This statistic is important because it also influences the normalized heat pump power if a Z-score normalization was applied. In fig. 2.1, you can see how the number of active devices increases every month.

2.2.1 Data aggregation

Within the query procedure, the data was aggregated to daily values on the device level. In addition, only the period the device was active was extracted, further reducing the query size. The variables that are used in this thesis are aggregated as follows. Heat pump power is summed per day and for the in-/outside temperature the daily mean was used. The aggregation of other features can be found in fig. A.1.

2.2.2 Data selection

By analysing the autumn to spring period, we only take into account the period where heating is mostly used. Because we also had to take the number of active devices into account, we selected data from the start of October 2022 (October 6th) till April 2023 (April the first) resulting in 59 active devices.

For the local weather data, besides aggregation, the values were multiplied by 10 because the KNMI recorded the data in a scaled format, where the value ′0.1′ corresponded to a magnitude of 1, and the value ′2′ represented a magnitude of 20 °C.

If we want to compare all the devices equally, the heat pump power should be normalized. We chose Z-score normalization, a technique in which the mean of each value is subtracted and then divided by its SD. The Z-score normalization is applied to each device separately so that each has about the same range of values.

2.3 Data exploration

In this section, both specific traits of devices are investigated, as well as the distribution of some features.

2.3.1 Individual devices

By considering each device separately, we can examine the relationship between outdoor temperature and heat pump power in greater detail. As we plot them, we can see that devices can behave very differently depending on the temperature.

Without removing the relation with temperature, it is hard to say if this is caused by human behaviour. What we can see is that with sub-zero temperatures, some devices seem to use less power, while others use more. This happens with multiple devices on several occasions. Furthermore, some heat pump power trends are spikier than others. An example of the differences between devices can be found in fig. 2.2.

Are the users of the heat pump turning it off and on, does it have to do with poor isolation, is the boiler taking over, or is it a defect? To see exactly how common one behaviour is, all devices should be compared.

2.3.2 Distributions

Plotting all the values of heat pump power values per device per day and the average outside temperature on that day gives an insight into the distribution of heat pump power per temperature (fig. 2.3).

As mentioned before, multiple devices have below-average heat pump power usage with freezing temperatures, while others have the opposite amount. There are also a few outliers at higher temperatures, but most devices seem to behave quite similarly at those temperatures. If we try to plot a simple linear regression on outside temperature and heat pump power (blue line), we notice that it seems off. From the right it still looks right however, if we end up at sub-zero temperatures, we expect the curve to at least flatten. This is solved by creating a cubic function (orange line). This line does flatten and goes even down below zero, meaning that eventually, the majority of devices have below-average power usage.

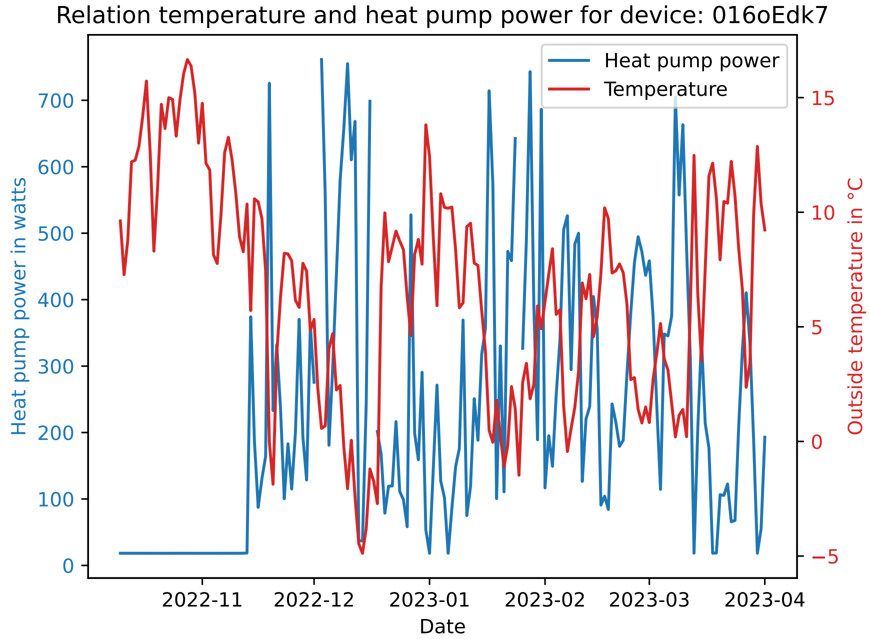
We expect that the difference in behaviour between the devices shown in fig. 2.2a and fig. 2.2b are also visible in the distribution plot (fig. 2.3). We can prove this by plotting these individual devices with the same regression and polynomial lines. It is visible that one device follows the regression line while the other follows the polynomial line more closely during sub-zero temperatures (fig. 2.4).

2.3.3 Polynomials

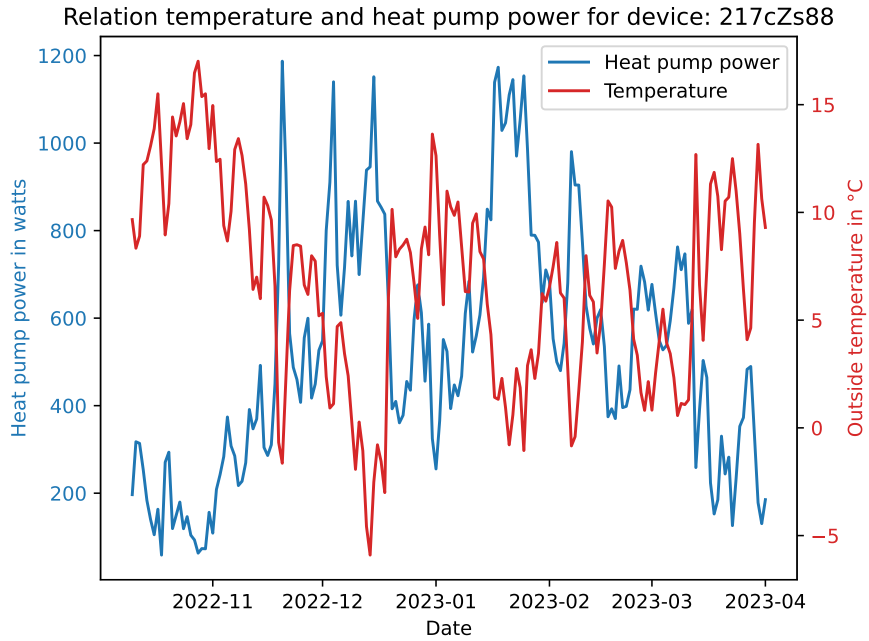
The cubic function is formed by incorporating a cubic polynomial into the outside temperature. This process leads to the creation of new features. A cubic polynomial is referred to as a third-degree polynomial because the expression's highest degree is 3, or equivalently, the power of the leading term is 3.

2.3.4 Day of the week

Another interesting property of heat pump usage to investigate is the difference between normal weekdays and weekends. To see if there actually is



(a) Device with a less pronounced correlation to the outside temperature.



(b) Device with stronger pronounced correlation to the outside temperature.

Figure 2.2: Time series depicting the relationship between heat pump power and outside temperature on a daily basis for individual devices with on the left y-axis heat pump power and the right y-axis outside temperature. The time series spans from October 2022 to April 2023. The blue line represents heat pump power, while the red line represents outside temperature.

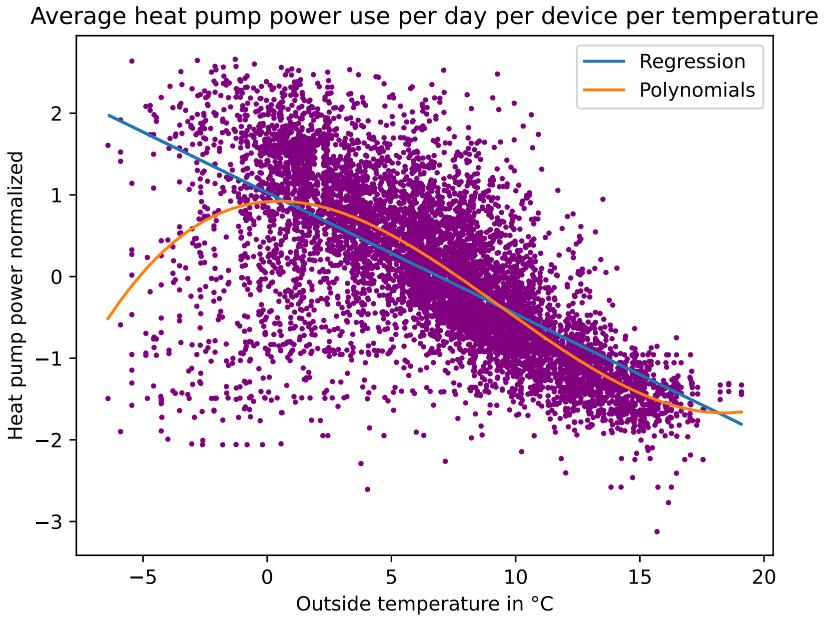


Figure 2.3: Scatter plot with regression and 3rd degree polynomials. It shows heat pump power vs outside temperature for each day and 45 devices during the period November 2022 - March 2023

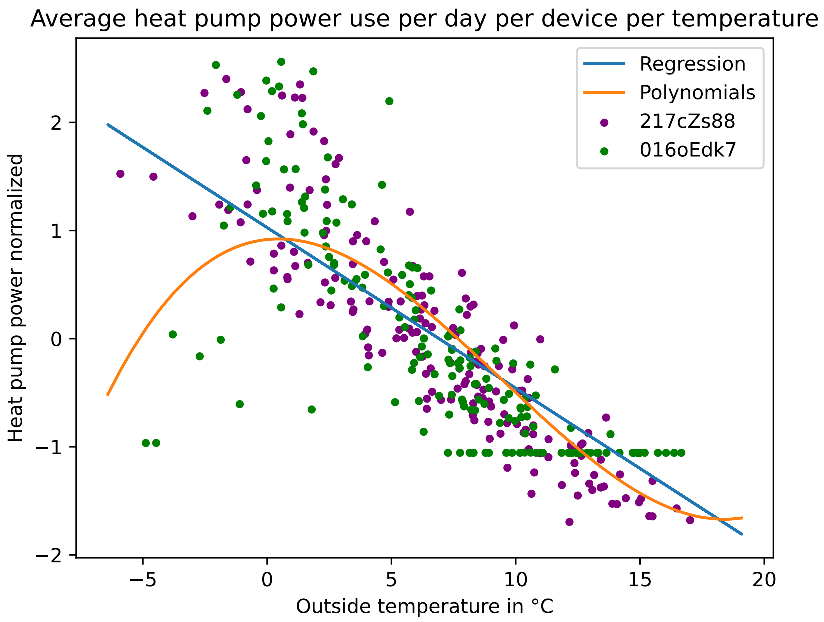


Figure 2.4: Scatter plot with regression and 3rd degree polynomials. It shows heat pump power vs outside temperature for each day and 2 devices during the period November 2022 - March 2023

a difference, we looked at both the means of heat pump usage on working days and that of the weekends.

The average heat pump power usage on working days is 495.83, and that on weekends is 495.93. It seems like there is close to no difference between these two, but what if we took the weekend days apart? For Saturday, the average power usage is 488.96, while on Sunday this is 497.01. Previous work shows that the day of the week does have an influence on temporal trends, which is why we still included it [13].

Based on this information, four features were derived, considering the day of the week as a factor (Weekend, Weekday, Saturday, Sunday).

2.3.5 Chosen features

While some features besides outside temperature had a relatively high correlation with heat pump power, these were all closely related to each other and the heat pump itself (multicollinearity). Although the actual method will be explained later (3), the idea is to create a model that can take away the trend based on temperature and show us human behaviour in what is left. Correlation is therefore important to find. Because of that, only the heat pump power, outside temperate and inside temperature was selected to be the basis of our future model.

2.4 Data cleaning

For data cleaning we looked at missing values, and outliers, removing unconventional or peculiar behaving devices and imputing the missing data.

2.4.1 Missing values

As you can see in fig. 2.5, some devices have in more than 10% of their rows missing data. While it is possible to impute all of it, the number of days that we can analyse is already quite small and imputing in a way that can pick up on both the temperature trend and human behaviour with this little amount of data is not without risk. Therefore, we made the decision to drop

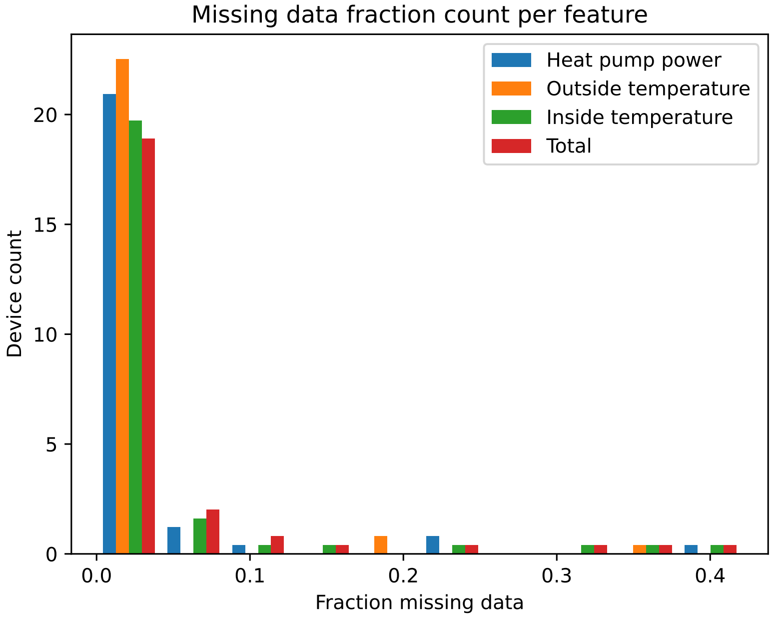


Figure 2.5: Histogram displaying the distribution of missing data per feature, shown as the count of percentages. Total depicts the percentage of rows where at least one missing is present.

6 devices out of our analysis, leaving us with 53 devices.

2.4.2 Outliers

For outliers, we looked at both individual data points and devices that are outliers themselves compared to other devices.

2.4.2.1 Outlier definition

An outlier is a data object that significantly deviates from the majority of the dataset. In contrast to noisy data, which represents random errors or variance, outliers are suspected of being generated by different mechanisms compared to the rest of the data. When an outlier is suspected to be a result of data collection or recording error, one possible approach is to remove the observation [14], [15].

2.4.2.2 Finding outliers

Finding outliers in data that might be heavily influenced by human behaviour brings some challenges with it. If it is not a recurrent pattern it can still be human behaviour because it happened on a holiday for example. In spite of that, behaviour that happens occasionally, once or twice a year, is not recognizable in data that describes only about 8 months of heat pump usage.

One option is to create a linear regression model, predict the heat pump power and use the residuals to determine if there are outliers. This way, it is possible to determine whether a given amount of heat pump power corresponds to the outside temperature, and if it does not, whether this behaviour occurs repeatedly and thus is not an outlier. The disadvantage of this approach is that the linear regression is trained on outliers, among other things.

Therefore, we chose to detect outliers directly on the original data, using 2.5 times the standard deviation. Choosing the right number of standard deviations was based on trial and error. With a standard deviation of 3, almost no values were seen as outliers, although we were sure there were some. With 2, too many of the values were seen as outliers, although they were probably just part of the human behaviour we are looking for or spikes because of temperature change.

In the end, only a few data points per device were identified as outliers, almost all of which used more heat pump power, as can be seen in fig. 2.6. This was to be expected, because the beginning of autumn has very low heat pump power values, close to zero, meaning that heat pump power values on the lower end are almost never marked as an outlier. After finding the positions of the outliers, their values were transformed into nan values.

2.4.2.3 Peculiar behaving devices

In some cases, whole devices were dropped for having strange values altogether. An example of one of these devices is visible in fig. 2.7. The average heat pump power usage of this device is too low compared to most other

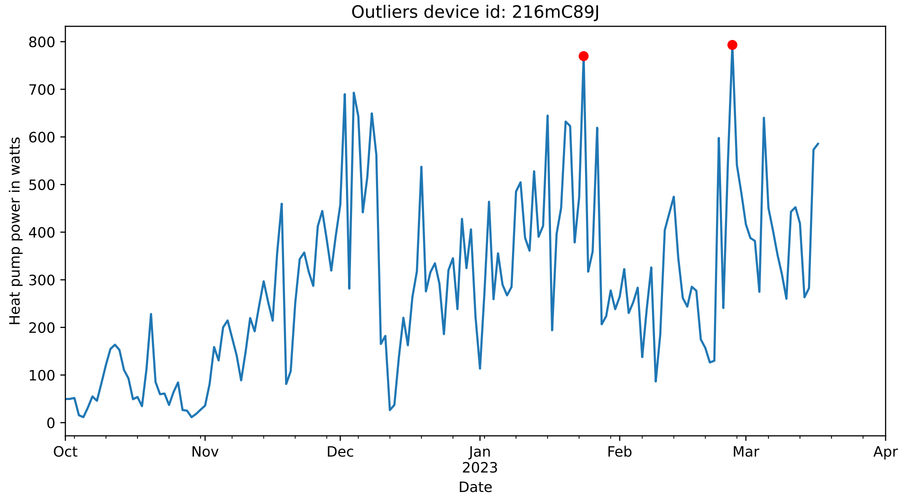


Figure 2.6: Time series of daily average heat pump power from a device during the period October 2022 – April 2023 with outliers, indicated by red dots.

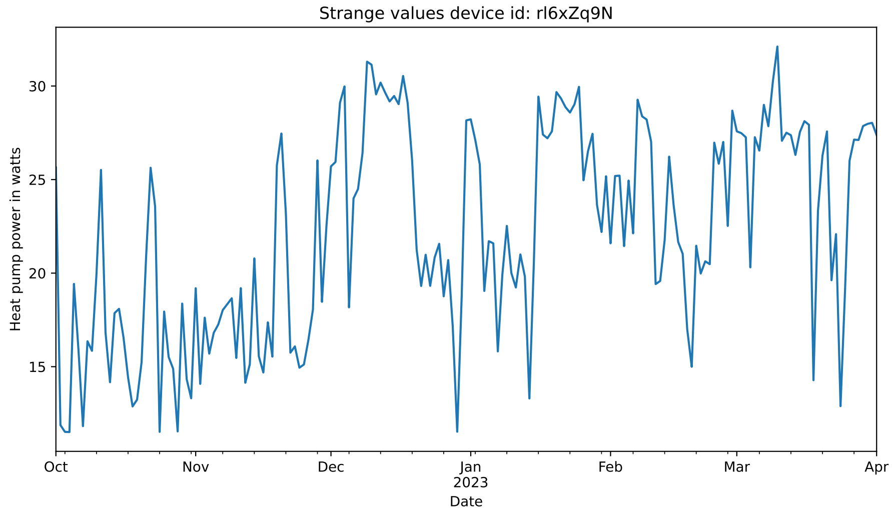


Figure 2.7: Time series of daily average heat pump power from a peculiar device during the period October 2022 – April 2023 with exceptionally low values.

houses and suggests that another form of heating is used. Besides looking at means and eyeballing, Inversable added a list of devices that exhibited similar odd behaviour. Combined with our own research, this led to the deletion of 8 devices, leaving us with 45 devices. The inside temperature was also checked for outliers.

As mentioned earlier, some devices show different behaviour during freezing temperatures than others. As a second cleaning stage for the creation of the final model, 15 of these devices were removed. Therefore, the data set only contains devices that use more heat pump power on average during freezing temperatures, instead of less like some other devices.

2.4.3 Imputing missing values

For imputing, we applied linear interpolation. A linear relation between temperature and heat pump power was found. However, the sub-zero temperatures showed large deviations from the estimated relation. Using polynomials or even splines resulted in some unpredictable behaviour with heat pump power values going below zero. Multi-imputation would also have been an option, like many others, but because of a limited amount of time, a relatively safe choice was made with linear interpolation. El-Nesr [16] wrote an article about interpolation methods on time series and made a list based on effectiveness expressed in explained variance (R2). It shows that on data sets that are not too complex (small number of features) linear interpolation performs well. Although, it is necessary to indicate that the dataset used in this example is more predictable by time than ours. The imputation was applied to each device individually to take the specific behaviour of that device into account.

After removing the odd devices, and the outliers and imputing the data, a few devices still had missing values. Those were the devices with missing values at the beginning of the extracted period. Therefore, we made the decision to shorten the period by moving the start date a few days forward, changing it from October 6th to October 10th. By changing the data, we also changed the statistics, as seen in figure table 2.1.

Table 2.1: Statistics of the data before and after all the cleaning steps.

Feature	Mean	New mean	SD	New SD
Heat pump power	414.11	495.86	360.50	352.94
Temperature inside	19.66	19.64	1.63	1.59
Temperature outside	7.04	6.92	4.63	4.61

Instantly notable is the standard deviation (SD) of heat pump power, it is almost as big as the average heat pump power usage. This was to be expected knowing that there is a difference in both house sizes, the number of inhabitants, and isolation measures.

2.4.4 Heating degree days

In the end, an extra feature was created named heating degree days. This feature is created by subtracting the actual temperature from 15 °C, unless the temperature is higher than 15 °C, then it will be set to zero. The reasoning behind this is that a heat pump generally only operates when the outside temperature is below 15 °C [17].

3. Method

Multiple models were created before the final model was used to make the predictions on which Fourier analysis is applied. We used Linear regression, polynomial regression and random forest models with any combinations of the selected features. Each of the models was scored with explained variance (R^2) to eventually create the final model, which is based on linear regression.

3.1 Prediction models

Every model was created using a train set that contains 66.67% of the data and a test set that contains 33.33% of the data. After every feature that we introduced in this thesis was tested and no further improvement was gained, the next model was tested.

3.1.1 Linear regression

The first models that were created were based on linear regression.

Linear regression is a fundamental and widely used kind of predictive analysis. The overall goal of regression is to investigate two things: (1) How well does a collection of predictor factors predict an outcome (dependent) variable? (2) Which factors, in particular, are significant predictors of the outcome variable, and how do they influence it (as indicated by the size and sign of the beta estimates)? These regression estimations are used to describe the link between one or more independent variables and one dependent variable [18].

3.1.2 Polynomials

We also implemented polynomial regression.

Polynomial Regression is a type of regression analysis in which the relationship between the independent and dependent variables is represented by an nth-degree polynomial. In our case, a 3rd-degree polynomial with outside temperature as the dependent variable

A basic linear regression model works only when the data relationship is linear. However, if we have non-linear data, linear regression will be unable to create a best-fit line. In such cases, simple regression analysis fails. As we have seen in data exploring (2.3.2), not all data follows a linear relationship between temperature and heat pump, the sub-zero temperatures showed large deviations from the estimated relation [19].

3.1.3 Random forest regression

The last model type we tried to implement was random forest regression.

Random Forest is an ensemble technique that can handle both regression and classification problems by combining many decision trees using a technique known as Bootstrap and Aggregation, or bagging. The core idea is to use numerous decision trees to determine the final output, rather than depending on individual decision trees [20].

3.1.4 Final model

Our final model is based on linear regression with just the in-/outside temperature since adding extra features did not improve the model. As explained during data cleaning (2.4.2.3), our final model used the dataset that was cleaned an extra time with the removal of the devices that behaved differently during the sub-zero temperatures. It was trained on 20 devices and tested on 10. This model outperformed all other models, even though polynomial regression had the best performance on the original data set. The results of this model will be explained in the results section.

3.2 Linear assumptions

When using linear regression, some assumptions are taken into account that need to be verified. The first assumption is that there exists a linear relationship between the predictors and the response variable. We have already shown that this linear relationship is there (2.3.2).

Another assumption is that the error terms are uncorrelated. The computation of standard errors for the estimated regression coefficients or fitted values relies on this assumption of uncorrelated errors. However, if there is actually correlation among the error terms, the estimated standard errors will tend to underestimate the true standard errors. Consequently, confidence and prediction intervals will be narrower than they should be. Time series data often exhibit correlated errors, whereas adjacent observations have positively correlated errors. To investigate this, we can plot the residuals against time. If there is no noticeable pattern, the errors are likely uncorrelated. However, if adjacent residuals show tracking or similarity, it suggests error term correlation [15]. To test this, an autocorrelation plot will be made to test for correlation between the residuals, this will be done in the results section 4.5.

The last assumption is that the errors are independent, have equal variance and are normally distributed [21]. To assess the normality assumption, the histogram of the residuals can be plotted. If the residuals follow a roughly symmetric bell-shaped distribution, it suggests that the assumption of normality is met. This will also be done in the results section (4.4).

3.3 Validation

Multiple validation methods were used. For comparing models, cross-validation and explained variance. For the final model, a histogram of the residuals was plotted in combination with an autocorrelation plot.

3.3.1 Cross validation

Cross-validation is a validation method that splits the data into a different train and test set for multiple iterations. Because some train/test splits will be more favourable for models than others, cross-validation gives a more complete insight into the model's actual performance. Each cross-validation was executed with 100 iterations.

3.3.2 Evaluation metric

For evaluating the models, we chose explained variance (R^2). This metric is not only useful for comparing models but also gives more information on how good an individual model is. For example, if a model has R^2 equal to 0.05 then it is definitely poor.

3.3.3 Autocorrelation

With the creation of the autocorrelation plot on the residuals, it is visible how well the model predicts the heat pump power on different time intervals and if the assumption of uncorrelated error terms is met. Autocorrelation looks at the correlation of a feature (in our case, heat pump power residuals) with itself on a later time step. All features start relatively high on correlation, this makes sense because they have a high correlation with themselves. Therefore, minor differences over one or two days may not significantly impact the correlation. It is like the weather, it is predictable for a few days, because it stays relatively the same, but over 20 or more days, normally not so much.

4. Results

4.1 Overview of the results

The results are split between the predictive model results and those of the Fourier analysis that was implemented on top of it by Sahar Pourahmad.

4.2 Data properties

The properties of the data that was used for the prediction model is different from the one used in the previous models. As explained in the data cleaning section, extra devices were removed. The statistics of the data are shown in table 4.1.

Table 4.1: Statistics for data used in final model

Feature	Mean	SD
Heat pump power	534.11	380.51
Temperature inside	19.73	1.69
Temperature outside	6.94	4.62

4.3 Final base model

Our final base model had a score of 0.644. This is an increase of 22% compared to our best previous model. To see how well this model predicts on individual devices, both best (fig. 4.1) and worst predictions are shown (fig. 4.2). The difference in score between these individual devices is relatively big and suggests there are still unexplained factors influencing individual devices. Despite the significant difference in scores, the line of the less accurately predicted device deviates less than the score suggests. While the overall trend is still well followed, the predicted values are slightly overestimated or underestimated.

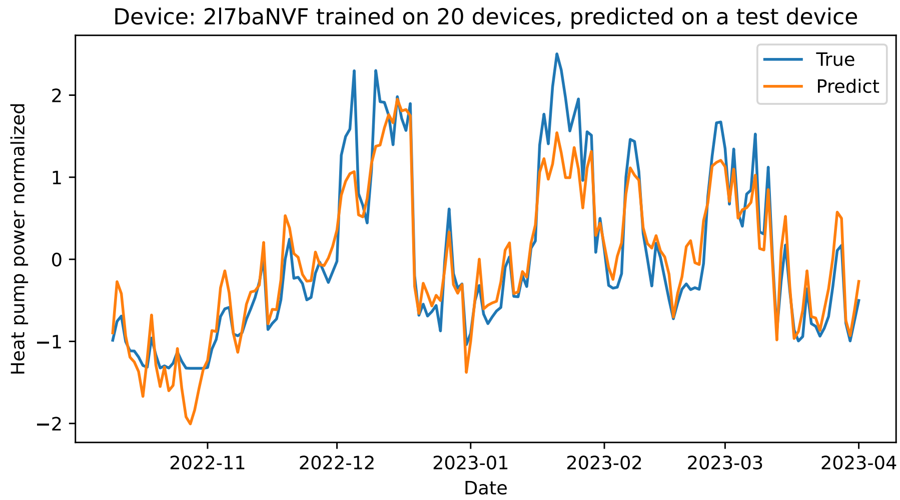


Figure 4.1: Time series depicting the relationship between the true and predicted normalized heat pump power on a daily basis for a well-predicted device. The time series spans from October 2022 to April 2023. The blue line represents the true heat pump power, while the orange line represents the prediction. R^2 score of 0.855.

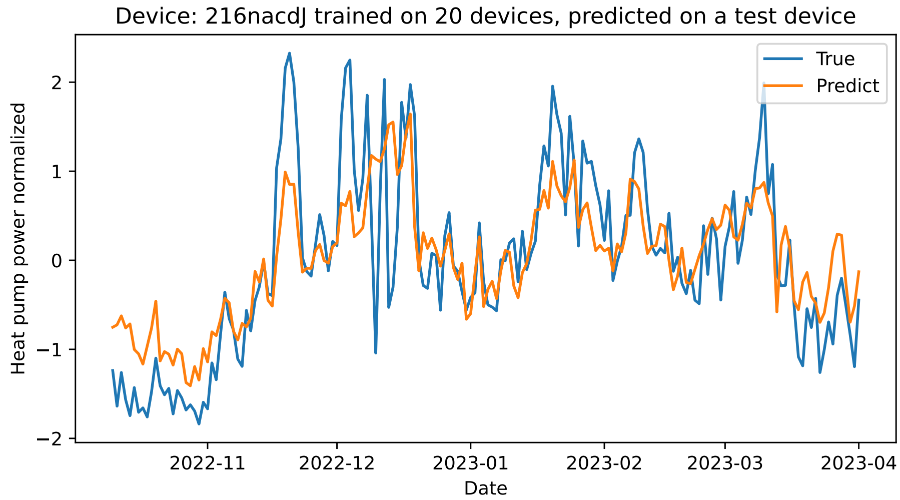


Figure 4.2: Time series depicting the relationship between the true and predicted normalized heat pump power on a daily basis for a poorly predicted device. The time series spans from October 2022 to April 2023. The blue line represents the true heat pump power, while the orange line represents the prediction. R^2 score of 0.184.

4.4 Residuals

As mentioned in the linear regression assumptions, residuals from the prediction of our model should follow a normal distribution. Plotting the residuals of our model in a histogram shows how well the normalized heat pump power of the test devices is predicted and proves that the assumption has been met (fig. 4.3).

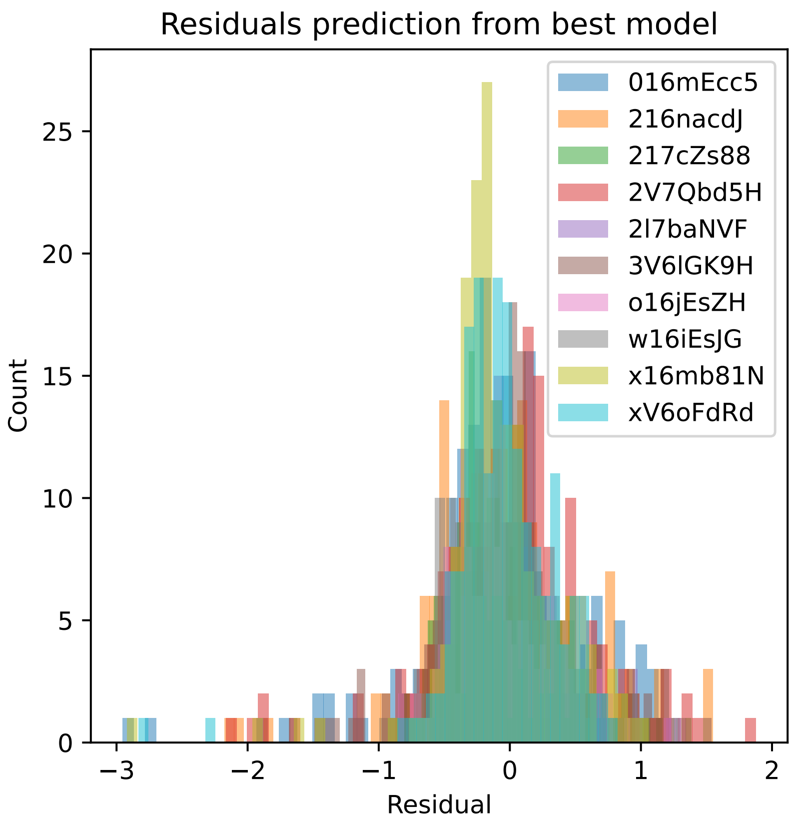


Figure 4.3: Histogram displaying the residuals count from the predictions of our best model on the test devices. The residuals are normalized heat pump power and therefore the number of standard deviations from the mean.

4.5 Autocorrelation

In figure fig. 4.4 you can see the autocorrelation graph. We observe that the model encounters challenges in making accurate predictions between 1 and 20 days. This changes after 20 days, and most devices stay between the boundary lines, meaning that the model can predict heat pump power on periods longer than 20 days quite well. There are a few devices that are interesting to discuss.

We observe a negative correlation around 20 days for 3V6IGK9H, meaning that when heat pump power is high it will be low 20 days afterwards and vice versa. Similarly, at 40 days, a high correlation is observed for this device. These could be signs of bad isolation, as it does not seem very stable.

Looking closer at this specific device reveals that the house was built between 1900-1945 and is detached. It has taken all possible isolation measures: Roof, walls, floor and double-glazed windows. It was given energy label C.

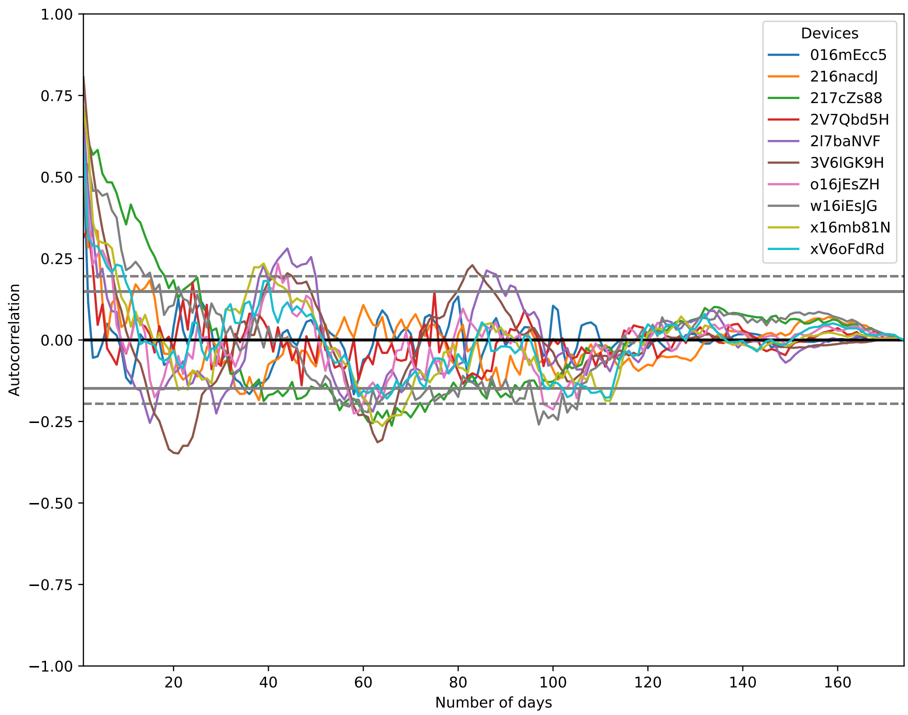


Figure 4.4: Autocorrelation graph displaying the correlation of residuals obtained from the predictions made by our best model on the test devices over a duration of 174 days.

The dark green line, device 217cZs88, has a slow descent in correlation, meaning that its heat pump power is relatively stable. When it uses a certain amount of heat pump power, it will roughly use the same amount the days after.

Looking closer at this specific device reveals that the house was built between 1980-2000 and is semi-detached. It has taken all possible isolation measures: Roof, walls, floor and double-glazed windows. It was given energy label B.

As you can see, even though both houses took the same isolation measures, the autocorrelation shows that their heat pumps do behave differently. One logical explanation is that the age of the house plays an important role, together with being detached or semi-detached. Another one could be found in the way the owner uses the heat pump. Although it is advised to keep it running, some people might turn it on and off.

4.6 Fourier analysis

By using Fourier analysis is it possible to find trends hidden in the data. To compare the difference between Fourier on pure heat pump power (fig. 4.5) and Fourier on residuals (fig. 4.6), both results are plotted.

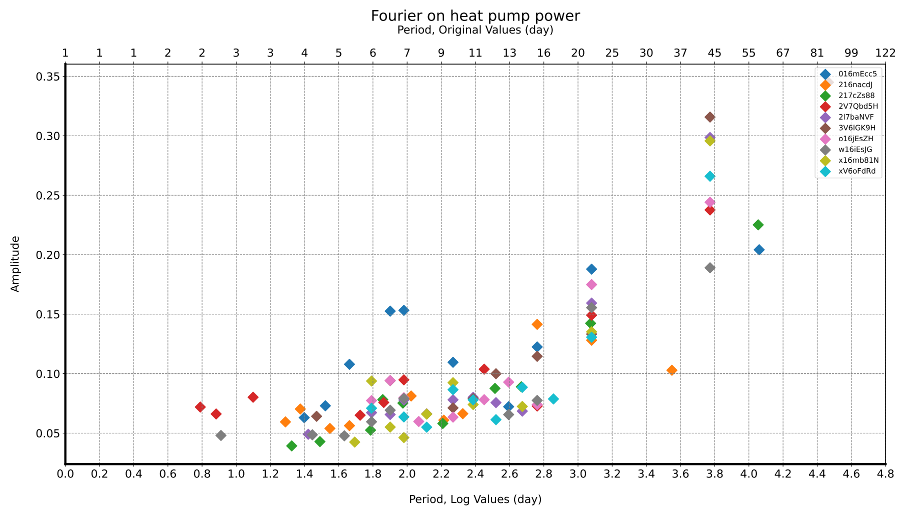


Figure 4.5: Scatter plot displaying the top 10 frequencies per device from the Fourier transform on heat pump power. With the logarithm of the period on the bottom x-axis and the real value of the period on the top x-axis. On the y-axis, the corresponding amplitudes, obtained from the Fourier transform.

By comparing the Fourier results showing patterns in the heat pump

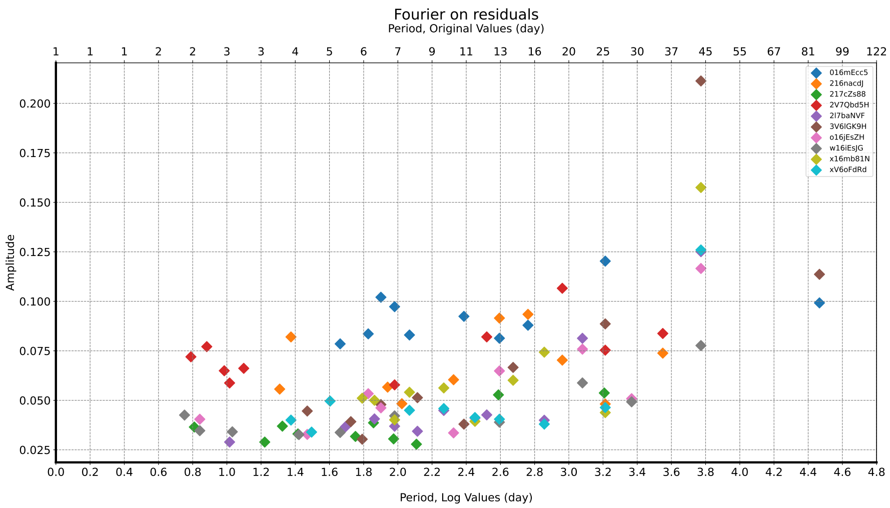


Figure 4.6: Scatter plot displaying the top 10 frequencies per device from the Fourier transform on the residuals. With the logarithm of the period on the bottom x-axis and the real value of the period on the top x-axis. On the y-axis, the corresponding amplitudes, obtained from the Fourier transform.

power (fig. 4.5) and the corresponding residuals for the same 10 houses (fig. 4.6), several observations can be made. The original heat pump power patterns associated with 6 days, 9–11 days, 11–13 days, 16 days, and 20–25 days are largely diminished in the residuals.

However, certain houses still exhibit residual values that demonstrate recurrent patterns at approximately 6-7, 13, 25, and 45 days. This suggests that even after removing the influence of temperature on the heat pump power, there are still periodic changes in power usage that are visible on the Fourier analysis of residuals. These recurring patterns may indicate behavioural patterns, such as weekly or biweekly activities that cause people to be away from their homes, resulting in changes in power usage.

5. Conclusion

The conclusion is that, based on the results, we have no conclusive evidence that we found human behavioural patterns. We did find certain patterns, but we were not able to really specify human behaviour by using the described methods. It does seem that some heat pumps show strange behaviour in ways that are not good for the overall performance. This could have implications for people who have hybrid heat pumps installed, even though their situation is not optimal for one, leading to higher power usage and cost than necessary. Besides power usage and cost, it could also harm the reputation of hybrid heat pumps.

5.1 Discussion

Different causes could be given for the difference in the performance of heat pumps during freezing temperatures. According to Inversable, it could be due to one of the following reasons. Firstly, it's possible that people set up the heat pump in a way that when it uses more power than desired, it stops and the boiler takes over. Secondly, it could be due to less optimal placement of the outside unit of the heat pump, such as being exposed to direct wind or being surrounded by snow or ice build-up, which can decrease its efficiency. Thirdly, it could be a faulty sensor. Additionally, it could be the difference between brands of heat pumps, one may be able to perform better in freezing conditions than others.

Furthermore, We would have liked to apply various imputation techniques. In this thesis, only linear imputation was applied, even though more complicated imputation techniques like MICE have proved to be very successful.

Besides imputation, our outlier detection method did not take the low temperature into account which creates some outlier peaks, and if nan values were present, outliers were not detected for that specific device.

Additionally, even though research showed the possible importance of the extra features that were created, they did, however, not improve the model. The reasons for this are unclear, and further investigation is required.

Lastly, this research was conducted within a relatively short time limit of two months. In addition, it is important to note that the project faced server issues at the beginning, resulting in a significant reduction of our available time by at least a few weeks.

5.2 Future research

For future research, more data over a longer period of time is required to make a more advanced analysis. In this thesis, the data covered only about 7 months of data.

Secondly, the difference in behaviour could be further analysed to find out if the power usage trends are influenced by measurement faults, defects, bad isolation, or maybe human behaviour after all. The actual impact of the behaviour of these devices is not measured, but it would be insightful to see what the impact on performance is and if measures have to be taken.

Lastly, clustering techniques could be applied to cluster devices that behave similarly. By doing so, it can be investigated whether these clustered devices also possess other overlapping characteristics.

Bibliography

[1] NASA Earth Observatory. "Global warming." (Jun. 3, 2010), [Online]. Available: https://earthobservatory.nasa.gov/features/GlobalWarming (visited on 06/26/2023).

[2] UN News. "Ipcc report: 'code red' for human driven global heating, warns un chief." (Dec. 14, 2021), [Online]. Available: https://news.un.org/en/story/2021/08/1097362 (visited on 06/26/2023).

[3] United Nations. "The paris agreement | united nations." (Aug. 9, 2021), [Online]. Available: https://www.un.org/en/climatechange/paris-agreement (visited on 06/26/2023).

[4] International Energy Agency. "The netherlands 2020 – analysis - iea." (Sep. 2020), [Online]. Available: https://www.iea.org/reports/the-netherlands-2020 (visited on 06/26/2023).

[5] Ministerie van Economische Zaken, Landbouw en Innovatie. "Wat is het klimaatakkoord?" (Jan. 22, 2020), [Online]. Available: https://www.rijksoverheid.nl/onderwerpen/klimaatverandering/klimaatakkoord/wat-is-het-klimaatakkoord (visited on 06/26/2023).

[6] Ministerie van Algemene Zaken. "Measures to reduce greenhouse gas emissions." (Jun. 2, 2021), [Online]. Available: https://www.government.nl/topics/climate-change/national-measures.

[7] Stichting Nederlandse Verduurzamingsindustrie. "Hybride warmtepompen, haalbaar en betaalbaar." (Nov. 19, 2021), [Online]. Available: https://www.nvi-go.nl/publicaties/hybride-warmtepompen-haalbaar-en-betaalbaar/ (visited on 06/26/2023).

[8] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. "Warmtepomp de norm vanaf 2026: Goed voor klimaat en de energierekening." (May 2, 2023), [Online]. Available: https://www.rijksoverheid.nl/actueel/nieuws/2023/05/01/warmtepomp-de-norm-vanaf-2026-goed-voor-klimaat-en-de-energierekening.

[9] Milieu Centraal. "Warmtepomp: Tips en uitleg." (Jun. 25, 2023), [Online]. Available: https://www.milieucentraal.nl/energie-besparen/duurzaam-verwarmen-en-koelen/warmtepomp-duurzaam-elektrisch-verwarmen/ (visited on 06/26/2023).

[10] Stichting Smart Energy Foundation. "Demoprojecthybride." (Jun. 26, 2023), [Online]. Available: https://www.demoprojecthybride.nl/ (visited on 06/26/2023).

[11] 1clickheat, *How-does-a-heat-pump-work*, Mar. 2022. [Online]. Available: https://1clickheat.com/wp-content/uploads/2022/03/How-Does-a-Heat-Pump-Work.jpg.

[12] M. Hommelberg, G. Janssen, and P. Friedel, “Publieke eindrapportage februari 2022,” Feb. 15, 2022. [Online]. Available: https://www.installatiemonitor.nl/eindrapportage-installatiemonitor-2/ (visited on 06/27/2023).

[13] H. Bloomfield, D. Brayshaw, L. Shaffrey, P. Coker, and H. Thornton, “The changing sensitivity of power systems to meteorological drivers: A case study of great britain,” *Environmental Research Letters*, vol. 13, no. 5, p. 054028, May 1, 2018. DOI: 10.1088/1748-9326/aabff9. [Online]. Available: https://doi.org/10.1088/1748-9326/aabff9.

[14] J. Han and M. Kamber, *Data Mining, Concepts and Techniques*. Jan. 1, 2012.

[15] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning, with Applications in R*. Springer, Sep. 1, 2021.

[16] M. El-Nesr, “Imputing the Time-Series Using Python - Dr Mohammad El-Nesr - Medium,” Mar. 2022. [Online]. Available: https://drnesr.medium.com/filling-gaps-of-a-time-series-using-python-d4bfddd8c460.

[17] E. E. Agency, *Heat and cold — mean air temperature*, Jan. 2022. [Online]. Available: https://www.eea.europa.eu/publications/europes-changing-climate-hazards-1/heat-and-cold/heat-and-cold-2014-mean.

[18] Statistics Solutions. “What is linear regression? - statistics solutions.” (Aug. 10, 2021), [Online]. Available: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/what-is-linear-regression/.

[19] A., “Understanding polynomial regression!!! - analytics vidhya - medium,” Dec. 15, 2021. [Online]. Available: https://medium.com/analytics-vidhya/understanding-polynomial-regression-5ac25b970e18.

[20] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. L. Sheridan, and B. P. Feuston, “Random forest: A classification and regression tool for compound classification and qsar modeling,” *Journal of Chemical Information and Computer Sciences*, vol. 43, no. 6, pp. 1947–1958, Nov. 1, 2003. DOI: 10.1021/ci034160g. [Online]. Available: https://doi.org/10.1021/ci034160g.

[21] J. J. Faraway, *Linear Models with R*. Aug. 12, 2004. DOI: 10.4324/9780203507278. [Online]. Available: https://doi.org/10.4324/9780203507278.

Appendices

A. Data

A.1 Data description

Description of the data can be found in fig. A.1.

A.2 Data extraction

The data that was used in this thesis only came from the heat pump sensor, the indoor climate sensor, and the local weather data. From the heat pump sensor, power in watts; from the climate sensor, the inside temperature of the house in degrees Celsius; and from the local weather data, the outside temperature in degrees Celsius.

The data was extracted by writing queries against the Influx database of Inversable. By setting the output of the influxDB towards Pandas data frames, it was possible to extract every table as a data frame. Due to server stability issues, it was not possible to extract the raw data. At the same time, raw data was not required for the method used.

General	Quantity	Min	Max	Unit	Time resolution [s]	Aggregation
Kampstrup Meter system heating	T_supply System supply temperature	-50	150	C°	5	Mean
	T_return System return temperature	-50	150	C°	5	Mean
	E_thermal_system kWh heat sum	-1000	1000000	kWh	5	Cum (max)
	P_thermal_system kW heat instantaneous (power)	-100000	100000	W	5	Mean
	F_system Water flow running throught the system	-100	100	liter/min	5	Mean
Heatpump	Energy kWh electricity sum heat pump	0	1000000	kWh	60	Cum(max)
	Power (W electricity at a point heat pump)	0	10000	W	5	Mean
Boiler	Energy kWh energy sum boiler	0	1000000	kWh	60	Cum(max)
	Power amount of W at a point for boiler	0	10000	W	5	Mean
DSMR Smart meter Dutch Smart Meter Requirements	p_consumed_L1: kW_laag fase 1 (kW low phase 1)	-50000	50000	W	10	Mean
	P_consumed_L2: kW_laag fase 2	-50000	50000	W	10	Mean
	P_consumed_L3: kW_laag fase 3	-50000	50000	W	10	Mean
	P_delivered_L1: kW_hoog fase 1 (kW high phase 1)	-50000	50000	W	10	Mean
	P_deliverd_L2: kW_hoog fase 2	-50000	50000	W	10	Mean
	P_delivered_L3: kW_hoog fase 3	-50000	50000	W	10	Mean
	E_consumed_low: kWh_laag fase 1	-1000000	1000000	kWh	60	Cum(max)
	E_consumed_high: kWh_hoog fase 2	-1000000	1000000	kWh	60	Cum(max)
	E_delivered_low: kWh_laag_terug (kWh low return)	-1000000	1000000	kWh	60	Cum(max)
	E_delivered_high: kWh_hoog_terug	-1000000	1000000	kWh	60	Cum(max)
	v_gas_consumed: Gasverbruik totaal (total gas usage)	0	1000000	m3	60	Cum(max
Sonoff	Kamertemperatuur (room temprature)	-30	100	C°	30	Mean
	Luchtvochtigheid (humidity)	0	100	%	30	Mean
KNMI	T (outdoor_temp)		Max(Min) = 20.9 Max(Max) = 38.9 Max(mean) =30.4913 (all after aggregation)	0.1 C°		Mean, min and max / 10
	SQ (sun_hours)		15.7 (after aggregation)	0.1 hour		Sum / 10
	FH (wind_speed)		17.209 (after aggregation)	0.1/ms		Mean / 10
	RH (rain)		56.4 (after aggregation)	0.1mm		Sum / 10
	Y (ice_form)	0	1	boolean		mean
	U (outdoor_humidity)	0	99.583 (after aggregation)	%		mean

Figure A.1: Data description

B. Trained models

B.0.1 Linear Regression 1

A linear regression model trained on all devices with a test set and cross-validation, using the feature's in-/outside temperature. Resulting in a R^2 score of 0.468.

It is roughly the same as when we tried to overfit. This means that we weren't actually overfitting and probably need more features that could explain the unexplained variance.

B.0.2 Linear Regression 2, extra features

A linear regression model trained on all devices with test set and cross-validation, using different combinations of the extra features created but without polynomials. Resulting in a R^2 score of also 0.468.

Adding the extra features brought no improvements.

B.0.3 Linear Regression 3, 3rd degree polynomial

A linear regression model trained on all devices with a test set and cross-validation, using the features inside temperature and 3rd-degree polynomials for the outside temperature. Resulting in a R^2 score of 0.526.

The highest score so far when using all devices.

B.0.4 Random forest regression

To see if other models besides linear regression would score better, we tried to implement a random forest regression. The random forest trained on all devices with a test set and cross-validation, using the feature's in-/outside temperature. Resulting in a R^2 score of 0.428.

This means that the random forest regression probably is not the ideal method for this data.