Raising energy awareness: outlines for concrete techniques for analyzing the energy usage of residential and business consumers based on smart meter data
## Contents

1 **Introduction** 10

2 **Conventions & notation** 14

3 **Background** 17
   3.1 The Dutch energy market 17
   3.2 The host organizations 20
      3.2.1 Huismerk Energie 20
      3.2.2 De Groene Stroomfabriek 21
   3.3 Smart meters 21
      3.3.1 Smart metering systems 21
      3.3.2 P1 and P4 data 22
      3.3.3 Deployment and penetration 24
      3.3.4 Benefits of smart meters 24
      3.3.5 Privacy concerns and legislation in the Netherlands 25
   3.4 Smart meter feedback 26
      3.4.1 Transtheoretical model of change 26
      3.4.2 Feedback 28
      3.4.3 Existing energy management systems 31
   3.5 Rebound effects 33
      3.5.1 Definitions 34
      3.5.2 Typologies and a classification 35
      3.5.3 Beneficial effects 39
      3.5.4 Rebound estimates 39
      3.5.5 Avoiding rebound 41

4 **Methodology** 42

5 **Signal restoration** 46
   5.1 Transient spike & noise reduction 46
   5.2 PV generation 48
      5.2.1 Smart meter limitations 48
      5.2.2 Intrusive methods for retrieving PV generation 50
      5.2.3 Modelling approaches 50
      5.2.4 Applications 56
   5.3 Energy data normalization 56
      5.3.1 Influences on consumption: a conceptual model 57
      5.3.2 Consumption profile fractions 62
      5.3.3 Cycles 65
      5.3.4 Normalization 65
   5.4 Conclusions 73
6 Smart meter solutions
6.1 Conversion techniques
  6.1.1 Consumption-to-cost conversion
  6.1.2 Consumption-to-emission conversion
  6.1.3 CO₂ emission visualization
6.2 Non-intrusive load monitoring
  6.2.1 Problem formulation
  6.2.2 A state-based classification of appliance load signals
  6.2.3 General NILM framework
  6.2.4 Common NILM techniques
  6.2.5 Overview of results
  6.2.6 Applications
6.3 Consumption extremes
  6.3.1 Discussion of typology
  6.3.2 Peak load
  6.3.3 Peak consumption
  6.3.4 Base load
  6.3.5 Idle consumption
6.4 Change detection
  6.4.1 Edge clustering
  6.4.2 Peak consumption saturation
  6.4.3 Long-term detection
  6.4.4 Hybrid mode
6.5 Benchmarking & comparison
  6.5.1 Comparative feedback
  6.5.2 Historic feedback
  6.5.3 Normative feedback
  6.5.4 Energy saving tips
  6.5.5 Manufacturing energy analysis

7 Discussion

8 Conclusions & recommendations
  8.1 Solution roadmap for Huismek Energie
  8.2 Solution roadmap for De Groene Stroomfabriek

Appendices

References

Glossary
List of Figures

1. A scheme of how the various techniques are designed to support the journey from smart meter data to output on any graphical user interface (GUI). ........................................ 13
2. Graphical depiction of the interdependence of $t$, $M$ and $C$. ........................................ 16
3. The distribution of electricity and gas consumption among households in the Netherlands. Adopted from (Gerdes, Marbus, & Boelhouwer, 2017). ................................. 18
4. The Huismerk Energie and De Groene Stroomfabriek logos. ........................................ 21
5. The components and communication portals of a smart meter, adapted from Hoenkamp, Huitema, and de Moor-van Vugt (2011); AlAbdulkarim and Lukszo (2008). ........................................ 22
6. Schematic depiction of the adapted Transtheoretical Model of Change as described in Riche, Dodge, and Metoyer (2010). ........................................ 27
7. Purchase prices and any available solar and smart plug expansion pack prices among the examined devices........................................ 33
8. Schematic representation of the rebound effect, adapted from Santarius (2012). ........ 33
9. Adaptation of a classification of rebound effects originally devised by Santarius (2012). The focus of this study will be on financial, psychological and material effects at the level of individuals and firms........................................ 36
10. Structure of the description of each of the proposed solutions. ...................................... 44
11. Load signal from the refrigerator from House 1 of the REDD database (Kolter & Johnson, 2011). ........................................ 46
12. Two copies of the refrigerator sample from Figure 11 processed with median filters with different window ranges: $R = 2$ (top) and $R = 10$ (bottom). ...................................... 47
13. Example depiction of consumption, generation and net flow as well as the measuring areas of counters $M_I$ and $M_{II}$. ........................................ 49
14. Common steps in many PV performance models. ........................................ 51
15. Regional differences in received hours of sunshine on 4th July, 2017. ............................ 53
16. Triangulation method for estimating irradiation $\gamma$ at $x$ based on irradiation figures $y_1$, $y_2$ and $y_3$ from nearest stations $s_1$, $s_2$ and $s_3$. ........................................ 54
17. Factors influencing household energy usage, based on the discussion in Kowsari and Zerriffi (2011). ........................................ 58
18. Self-made flowchart model for describing how energy consumption of a household is shaped by endogenous and exogenous factors. ........................................ 62
19. Normalized annual lapse of electricity (E1A, E3C) and gas (G1A, G2B) consumption profiles. Multi-year average temperatures were inserted as temperature coefficients to determine the gas profiles. ........................................ 64
20. Average electricity consumption for the E1A and E3C profiles in an average summer and winter week. ........................................ 65
21. Day- and quarter-hour-based self-normalization of NEDU E1A profile fractions from 1st March, 2017 (left) and a time-shifted copy (right). ........................................ 67
22. Annual and seasonal mean values and standard deviations of untouched and normalized weekly electricity consumption figures from three P4 signals. ........ 69
23. Annual and seasonal mean values and standard deviations of untouched and normalized weekly gas consumption figures from two P4 signals. .................. 70
24. The conversion steps from smart meter measurements towards CO$_2$ visualizations. .... 75
25 Splitting of a general energy bill into its various components. The bold entries
are influenced by consumption. ........................................ 76
26 Screenshot from a visualization by Carbon Visuals of total daily GHG emissions
in New York. ................................................................. 80
27 Depiction of the NILM problem for \( n = 6 \) devices, loosely based on a similar
picture from Hart (1992). .................................................. 82
28 State-based categorization of appliances. ................................ 82
29 Common operation scheme for many NILM techniques. .......... 83
30 States, transients and some related appliance features depicted for an example
load signal. ................................................................. 84
31 A factorial hidden Markov model applied to the disaggregation problem for \( n \)
appliances, adapted from Huss (2015). ................................ 87
32 Architecture of a feed-forward ANN (applied to the energy disaggregation prob-
lem) including a close-up of an artificial neuron, adapted from Ruzzelli, Nicolas,
Schoofs, and O’Hare (2010); Jain, Mao, and Mohiuddin (1996). ....... 89
33 Schematical visualization of the DTW alignment for two appliance signals, based
on a similar picture in Keogh and Ratanamahatana (2005). .......... 91
34 Schematic overview of several of the notions discussed above. ........ 95
35 Schematic depiction of the \( k \)th iterative step in the flat level base load estimation
algorithm. ................................................................. 101
36 Results from a test of the flat level method on a sample from House 2 from the
REDD database. ............................................................. 104
37 Test results of the recurrent value method used on a household with PV panels.
Shows two separate runs featuring \( k = 10 \) and \( k = 20 \) clusters. .......... 107
38 Test results from the flat level method applied to a five-day P4 data sample from
a medium-sized company. ............................................... 110
39 An example depiction of two-sigma intervals for \( L \) clusters on the power spectrum
based on the probability distribution for each cluster. ................. 112
40 Schemetical depiction of the peak consumption saturation method. ....... 115
41 Test results of the peak saturation technique applied to three unprocessed con-
sumption signals A,B and C previously used in Figure 22. ............... 116
42 A plot of the arctangent function (with \( x \)-axis values in radians). .......... 119
43 Schematic example depiction of the long-term change detection method. .... 119
44 Test results from the long-term method for three consumption signals A, B and C.121
45 Example depiction of the modelling ranking approach for comparative feedback
showing the probability density function of a log-normal distribution. .... 125
46 Example depiction of a possible target progress visualization for a one-month
period. ........................................................................ 131
47 Two examples of simple regression. ......................................... 133
48 Outline of the proposed regression technique and its applications. .......... 134
49 Possible results when correlation is weak (left) or when there is little variation in
production (right). ........................................................ 135
50 Hypothetical situation in which start-up lag in production output creates a bunch
of outliers in the \((v,C)\)-plane which may lead to bad model choice decisions. ... 137
51 Extrapolation efforts going south due to bad model choices. .......... 137
Example of segmented regression performed on a step-wise consumption-production relation. If production exceeds the threshold $T$, an additional machine is switched on.

List of Tables

1. Dutch grid connection classification system based on capacity and operational time (NEDU, 2017a).
2. Benchmark tariffs for 1 kWh of electricity. All amounts include government levies and VAT.
3. Some characteristics of P1 and P4 data.
4. Smart meter counters and what they record from consumption ($C$) and generation ($G$) under different circumstances.
6. Alternative classification scheme of energy-influencing factors.
7. Example situations for all possible types of indirect factors.
9. Average GHG emission intensities from various electricity generation assets.
10. Average GHG emission intensities from gas combustion for some types of gas.
11. Appliance features determinable from 1 Hz data. *Reactive power is not captured by all smart meters.
12. The electricity consumption and associated costs (based on the tariffs in Chapter 2) of refrigerators. The energy consumption of a new refrigerator is based on an A+++ model.
13. Summary of the advantages and disadvantages of each of the detection techniques discussed in this section.
14. Ranking of the solutions relevant for Huismerk Energie based on the above six variables.
15. Ranking of the solutions relevant for De Groene Stroomfabriek based on the above six variables.
Abstract

As of 2018, anthropogenic climate change has been acknowledged as a major challenge that needs rapid mitigation. Apart from large investments in sources of renewable energy worldwide, more frugal use of energy at the end of the consumer is also considered an important piece of the puzzle. A frequently cited way to achieve such reductions is to encourage consumers to change their behaviour by sending them personalized energy feedback. Smart metering systems being deployed everywhere in the EU seem to be an ideal facilitator of such feedback. In this text, a large number of concrete methods are described for the supporting companies to implement, which help sculpture raw smart meter energy data into various different feedback outputs that are promising in terms of their anticipated energy reductions. Each of these comes with a discussion of both technical aspects and possible energy savings. The final product of this thesis is a ranking of the individual applications, based on their technical feasibility and projected energy savings.
Foreword

This thesis was written at the end of the Science, Management & Innovation specialization of the master’s degree in mathematics at the Radboud University, Nijmegen. The text revolves around the utilization of smart meter data for energy management systems, a subject which unites aspects from the worlds of social science and applied mathematics, thus gluing my seemingly disjoint master specialization directions together.

I am most grateful to my primary supervisor Gerwin Hament for his ever-enthusiastic support, wide-ranging expertise on many relevant aspects of the energy market, and thorough and invaluable feedback on numerous draft versions throughout the project. Secondly, I thank my mother for her moral support which greatly helped pulling me through the less-sensational final phases of the project. Additionally, I thank Florian Knobloch for his support and feedback during the second half of this thesis. I am also grateful to Sue, Jelmer, Arianne, Teun, Luud, Xanne, Diederik and Frank for their suggestions and for correcting numerous typos and other oversights in the draft pieces I sent them. Finally, I thank the people at Huismerk Energie for facilitating my thesis with a quality workplace and lots of assistance, which has greatly helped me gain some insight into the energy market.
1 Introduction

By now, mitigation of anthropogenic climate change has firmly occupied a high position on the political agendas of many countries. The 2015 Paris Agreement requires all ratifying countries set themselves a Intended Nationally Determined Contribution, which is a non-binding national target for greenhouse gas emission reductions (Rogelj et al., 2016). There are numerous possibilities for a country to engage in, in order to achieve these carbon reductions, e.g. investing in renewable energy, afforestation or less-conventional means of carbon sequestration, or stimulating a circular economy. This thesis focuses on yet another possibility: namely that of accomplishing energy end-use reductions in buildings. Residential and commercial buildings accoun for about 40% of final energy use and 60% of electricity use worldwide (Faustine, Mvungi, Kaijage, & Michael, 2017; L. Yang, Yan, & Lam, 2014). In the Netherlands, over 20% of electricity and about 25% natural gas consumption is attributed to households (Papachristos, 2015; Gerdes et al., 2017). Thus, significant carbon reductions can be achieved upon realizing a notable decrease in building energy consumption (Kneifel, 2010).

A large step towards achieving energy reductions in Dutch buildings was taken in May 2017 when an energy covenant titled 10 PJ energy reductions in the built environment was agreed on by a club of government institutions, grid operators, energy suppliers and trade associations. It consists of a three-stage plan to achieve a 10 PJ reduction of final energy consumption in the Netherlands (H. Kamp et al., 2017). One component of this covenant has the participating energy suppliers (including Huismerk Energie and De Groene Stroomfabriek) committed to help their customers save energy by offering them enhanced feedback on their energy consumption. The aim of this thesis is to get the supportive companies up and running with respect to such endeavours:

| Main aim: | Provide Huismerk Energie and De Groene Stroomfabriek with a broad, comprehensible introduction to the opportunities and limitations of smart meter data and address the feedback-induced energy savings associated with specific applications based on this data. |

To this end, this text is built up as follows. Chapter 3 is an interdisciplinary background overview in which numerous topics relevant to this study are introduced. It starts off with a brief introduction to the Dutch energy market and its most important players. The host companies that facilitated this thesis are then introduced briefly in order for external readers to familiarize themselves with the companies’ visions. A number of the examples provided in later chapters relate to the context of Huismerk Energie and their customers. The remainder of Chapter 3 features a number of smart meter-specific elements, including their build characteristics, data output and deployment progress in the Netherlands, as well as the envisaged benefits of smart meters for different market players, and a discussion of relevant privacy legislation regarding the utilization of the data (Section 3.3). Additionally, an overview of factors that influence the effectiveness of energy feedback (Section 3.4.2) and an assessment of existing smart meter data applications (Section 3.4.3) are provided here. The chapter closes with a discussion of rebound effects (1) to help the reader place projected savings into perspective (Section 3.5). In chapter 4 the methodology of this research project is explained.

---

(1) Erosion of projected energy savings due to secondary consequences of either the feedback or the energy savings incited by the feedback.
However, the body of this text is formed by the solutions described in Chapters 5 and 6. The latter includes a discussion on the eventual applications, namely appliance recognition methods (Section 6.2), peak and base consumption values for both households and organizations (Section 6.3), various methods to automatically take notice of changes in a consumption signal (Section 6.4) and a number of comparative techniques that allow consumers to compare their energy consumption to their own historical data and to data of other consumers (Section 6.5). This last section also features (personalized) saving tips and goal-setting designs, and an additional firm-specific application that allows organizations to create a more detailed inventory of how energy consumption is related to their day-to-day operations. The output of all of these methods is given in terms of energy use (in kWh or m$^3$), which can be converted to a monetary unit (Section 6.1.1) or associated CO$_2$ emissions (Section 6.1.2). Potential energy savings are estimated for each of the solutions individually, based on empirical evidence from scientific literature.

The applications from Chapter 6 transform smart metering data into useful and personalized energy statistics that are meant to increase energy awareness among consumers and trigger consumption reductions. The performance of these solutions in terms of energy savings is expected to be higher when the input data used is a pure mirror image of the target consumer’s behaviour. As such, it is desirable to sift external interferences from the input signal as much as possible, which is the topic of Chapter 5. To this end, a conceptual model is drawn up in Section 5.3.1, which describes the effect various household-specific and external influences have on household energy consumption. One specific class of external effects, i.e. the influence of physical factors such as temperature and seasonal variations, is then discussed in more detail in Sections 5.3.2 - 5.3.4. Two additional cleaning endeavors of influences more technical in nature are discussed in Sections 5.1 and 5.2. The former discusses a filtering method to smoothen noisy high-resolution data series, while Section 5.2 describes means to estimate the generation figures from any present PV installations. Such estimations are necessary because smart meters do not register PV-generated amounts of electricity separately, but rather present a net figure from which real consumption can no longer be accurately derived. Figure 1 shows how the individual methods from Chapters 5 and 6 are interconnected.

There are large differences between individual applications described in this text in terms of complexity, projected energy savings, anticipated accuracy and their dependency on the corrective methods from Chapter 5. Therefore, two rankings are presented in Chapter 8, so that the implementing agencies can work their way through the solutions, starting with those that yield the highest savings and require the least effort. It may be that by realizing the easier solutions, the company’s programmers get new insights with respect to tackling one of the remaining problems that are perceived as difficult here. It may also be that problems described in this text have been solved externally by the time the earlier solutions have been built, for instance if a policy is installed that obliges solar panel yield to be automatically communicated to the grid operator as a separate data source.

This text is intended to provide Huismerk Energie and De Groene Stroomfabriek with an broad starting point for smart meter data applications. To serve a wide audience, the author has attempted to handle a light approach, with many examples and supporting images in favour of rigorous, complex theoretical discussions. Nevertheless, the subject is data-driven and readers with no affinity with numbers and basic arithmetic whatsoever may get lost in the more technically demanding sections later on. These technical details can be skipped though without losing
sight of the bigger picture. Readers with some programming experience will find a number of uncomplicated algorithm schematics with detailed discussions on the roles of various parameters, which should help kick-start the implementation and testing phases of any chosen solution. Additionally, some ready-made Matlab code, used to quick-test a number of the solutions, has been included in the Appendix.

The focus of the smart meter data applications presented in this thesis will be mostly on electricity consumption in the residential market. Therefore, a number of examples are in the context of Huismerk Energie and their customers. Some solutions can also be applied to business consumers and some are even business-only. In such cases a separate subsection will be devoted to this particular setting and the viewpoint will tacitly be changed to that of de Groene Stroomfabriek’s customer base. Although most solutions are related to electricity usage, there will be a few digressions on similar techniques for gas consumption, whenever relevant.
FIGURE 1: A scheme of how the various techniques are designed to support the journey from smart meter data to output on any graphical user interface (GUI).
2 Conventions & notation

User profiles
The Dutch organization for energy data exchange (NEDU) has designed a classification of energy consumers based on their consumption volumes and patterns. The various categories and their classification criteria are shown in Table 1 below. A more detailed presentation of the consumption patterns for some of the categories is given in Section 5.3.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Capacity</th>
<th>Other classifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small consumers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1a</td>
<td>( C \leq 3 \times 25 \text{ A} )</td>
<td>single tariff</td>
</tr>
<tr>
<td>E1b</td>
<td>( C \leq 3 \times 25 \text{ A} )</td>
<td>night tariff (07:00 AM - 11:00 PM)</td>
</tr>
<tr>
<td>E1c</td>
<td>( C \leq 3 \times 25 \text{ A} )</td>
<td>evening tariff (07:00 AM - 09:00 PM)</td>
</tr>
<tr>
<td>Large consumers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2a</td>
<td>( 3 \times 25 \text{ A} &lt; C \leq 3 \times 80 \text{ A} )</td>
<td>single tariff</td>
</tr>
<tr>
<td>E2b</td>
<td>( 3 \times 25 \text{ A} &lt; C \leq 3 \times 80 \text{ A} )</td>
<td>double tariff</td>
</tr>
<tr>
<td>E3a</td>
<td>( 3 \times 80 \text{ A} &gt; C ) \quad T \leq 2,000 \text{ h} )</td>
<td></td>
</tr>
<tr>
<td>E3b</td>
<td>( 3 \times 80 \text{ A} &gt; C ) \quad 2,000 \text{ h} &lt; T \leq 3,000 \text{ h} )</td>
<td></td>
</tr>
<tr>
<td>E3c</td>
<td>( 3 \times 80 \text{ A} &gt; C ) \quad 3,000 \text{ h} &lt; T \leq 5,000 \text{ h} )</td>
<td></td>
</tr>
<tr>
<td>E3d</td>
<td>( 3 \times 80 \text{ A} &gt; C ) \quad 5,000 \text{ h} \leq T )</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4a</td>
<td>( C &lt; 100 \text{ kW} )</td>
<td></td>
</tr>
</tbody>
</table>

| Profile          | Capacity                                      |                                    |
|------------------|-----------------------------------------------|                                    |
| Small consumers  |                                               |                                    |
| G1a              | \( C < 5,000 \text{ m}^{3} \)                 |                                    |
| Large consumers  |                                               |                                    |
| G2a              | \( 5,000 \text{ m}^{3} \leq C \leq 170,000 \text{ m}^{3} \) \quad T < 750 \text{ h} \) |                                |
| G2b              | \( 5,000 \text{ m}^{3} \leq C \leq 170,000 \text{ m}^{3} \) \quad 750 \text{ h} \leq T < 1,500 \text{ h} \) |                                |
| G2c              | \( 5,000 \text{ m}^{3} \leq C \leq 170,000 \text{ m}^{3} \) \quad 1,500 \text{ h} \leq T \) |                                |

Table 1: Dutch grid connection classification system based on capacity and operational time (NEDU, 2017a).

Electricity prices
As a benchmark for electricity prices, the tariffs for electricity from (Huismerk Energie energy prices, 2017) will be used in this text, see Table 2. Some customers utilize a double tariff pricing structure, meaning they pay somewhat more for their electricity during peak hours (Mo-Fr, between 07:00 AM and 11:00 PM) and get a reduction during the remaining off-peak hours.(2) All solutions discussed in this text work equally well for single and double tariff clientele, yet diligent differentiation between peak and off-peak tariffs is a tiresome effort which unnecessarily complicates notation. Therefore, all techniques will be discussed in the context of a customer with a single tariff contract. The double tariff case will be elaborated on when there is reason to. A more detailed discussion of energy prices is given in Section 6.1.1.

<table>
<thead>
<tr>
<th>Price per kWh</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Single tariff</td>
<td>€0.18718</td>
</tr>
<tr>
<td>Double tariff</td>
<td></td>
</tr>
<tr>
<td>(07:00 AM - 11:00 PM)</td>
<td>€0.19202</td>
</tr>
<tr>
<td>(11:00 PM - 07:00 AM)</td>
<td>€0.17992</td>
</tr>
</tbody>
</table>

Table 2: Benchmark tariffs for 1 kWh of electricity. All amounts include government levies and VAT.

(2)In the counties of Limburg and Noord-Brabant and in some other municipalities, peak hours run from 07:00 AM to 21:00 PM on Mo-Fr.
Sequences
Some notation will be fixed throughout this text. A sequence of \( N \) numbers will be denoted as
\[
x = (x_1, x_2, \ldots, x_N).
\]
The entries \( x_i \) are called the components of the sequence and the subscript \( i \) is called the index of a component. The notations \( x \) and \( (x_1, \ldots, x_N) \) are used interchangeably to describe the same object: the bold letter is invoked to encapsulate the entire sequence when there is no need to specify its components or its length. The starting index of a sequence need not be 1, sometimes it is more convenient to take 0 as starting point.

2.1 Convention: The following sequences are fixed throughout this text:

- \( t \) will consistently be used to denote a time series. Time series will typically start at index 0 to give rise to \( N \) intervals, (the periods between two subsequent instants from \( t \)). It is tacitly assumed the duration of each time frame is equal, i.e. \( t_i - t_{i-1} \) is the same for all \( i \). This duration is often aligned with the sampling rate of a dataset, but longer intervals are sometimes used as well.

- \( M \) will be used to denote readings from a (smart) meter which show cumulative consumption. Meter readings will always be used in conjunction with a time series, i.e. \( M_i \) will be the meter reading at time stamp \( t_i \).

- \( C \) is used to denote power consumption between two time stamps. If \( t = (t_0, \ldots, t_N) \) is a time series and \( (M_0, \ldots, M_N) \) denote the corresponding meter readings, then
\[
C_i = M_i - M_{i-1},
\]
for \( 1 \leq i \leq N \), so \( C = (C_1, \ldots, C_N) \) is sequence of length \( N \). To denote consumption between two non-subsequent time stamps \( t_i \) and \( t_j \), the notation \( C_{i,j} \) will be used. The standard units for energy consumption used in this text are kilowatt hours (kWh) for electricity and cubic metre (m\(^3\)) for gas.

- When using electricity data at a frequency of about 1 Hz, electric power, i.e. the rate at which electrical energy is delivered in an electric circuit (Jewett & Serway, 2008), becomes a quantity of interest. In this text, electric power will be denoted by \( P \) and more or less represents an instantaneous version of consumption \( C \). It is measured in Watts (W).

Figure 2 below shows \( t, M \) and \( C \) in a picture. Throughout this text, schematic visuals of meter readings and power sequences will be represented by smooth solid lines rather than by individual points to increase clarity.
Figure 2: Graphical depiction of the interdependence of $t$, $M$ and $C$. 
3 Background

This background section contains introductions on various aspects that touch upon the main topic of this thesis. These include the energy market, the host organizations that supported this thesis, smart meters and their data, privacy legislation regarding the utilization of this data, energy management systems, energy feedback studies and rebound effects.

3.1 The Dutch energy market

This thesis being about energy conservation in the Netherlands, it is pertinent to take a brief look at gas and energy consumption patterns and at the underlying system of production and distribution. This section features some introductory material on domestic energy use and on the most important players on the Dutch energy market.

Energy consumption in the Netherlands

Energy usage can be divided into direct energy consumption, i.e. on site electricity usage and combustion of gas and other fuels by the consuming party, and indirect energy consumption, which relates to all energy embodied in products and services consumed or used, e.g. food, furniture, waste treatment, etc. (Vringer & Blok, 1995). A smart meter only measures gas and electricity usage within a building. Hence, the reader should continuously keep in mind that the energy reductions discussed in this text represent only a portion of total energy usage: all indirect energy usage and all direct fuel combustion from (non-electric) transportation and less-conventional heating assets such as wood stoves will not be considered.

Household energy consumption

In a household setting it has been estimated indirect energy usage constitutes about half of total household energy usage in the Netherlands (Vringer & Blok, 1995; Nijdam, Wilting, Goedkoop, & Madsen, 2005; Steg, 2008). Of the remaining half if direct usage, about two-thirds is gas and electricity consumption (Gerdes et al., 2017). Thus, any energy reductions in a household context listed in this text, relate to about one-third of total domestic energy usage. Of this, about 37% is from electricity usage and 63% is from gas combustion. Refrigerators (13%) and lighting (12%) have the largest share in household electricity usage, whereas gas-fired combi-boilers are responsible for the large majority of gas consumption and therefore are the overall number one energy users (Gerdes et al., 2017).

Naturally, the exact amount of gas and electricity consumed differs between individual households, yet most households use about 1500 – 3500 kWh of electricity and 1000 – 1800 m³ of gas annually. The distribution of electricity and gas consumption among different households is shown in Figure 3 below.(3) A notable aspect from these graphs is that the bar corresponding to the households with the smallest gas usage is very large. A possible reason for this category being so well-represented is that some modern dwelling types come equipped with electrical heat pumps and boilers for heating and electric stoves for cooking, thus have (almost) no gas demand.

---

(3) Some more details on the exact distribution of electricity consumption among households are given in Section 6.5.1. A discussion of the distribution of energy consumption over different periods can be found in Sections 5.3.2 and 5.3.3.
Organizational energy consumption

There are great differences between the energy usage of different firms. The energy consumption of an organization depends, amongst other things, on the sector in which it operates and on the size of the firm (Statistics Netherlands, 2018; Fix, 2017; Gerdes et al., 2017; Wilting, 1996). A small service provider will have a very different energy profile compared to some major industrial party operating numerous large pieces of machinery. Therefore, it is difficult to make any general statements about organizational energy consumption. As a result, it will also be hard to determine the applicability of smart meter solutions and any corresponding energy reductions for organizations in general: it is very well possible some of the techniques discussed later in this text yield great results for one firm, while failing completely for the other. As such, organizations will require some individual attention when implementing any of the data applications discussed in this text.

The energy market

The Dutch energy market is a modern liberalized one as many others in the European Union
following a large privatization process during the late 90s and the early 00s (Serena, 2014). A brief overview of the most important players within the Dutch gas and electricity markets is presented below.

- **Producers:** all energy-producing agents. With regard to electricity this includes the exploitants of coal- and gas-fired power plants, wind parks and solar farms, but also a growing number of (semi-)self-sufficient consumers with their own solar panel installations. Examples of large electricity producers in the Netherlands include Essent, Electrabel and NUON (Gerdes et al., 2017). As for natural gas, the Nederlandse Aardolie Maatschappij (NAM) supplies about 75% of all natural gas used in the Netherlands (Nederlandse Aardolie Maatschappij, 2018). Some smaller parties also supply some biogas: in 2016, 0.9% of all gas consumed in the Netherlands was biogas (Gerdes et al., 2017).

- **Consumers:** the end-users of energy, which can be households or companies. Consumers buy their energy from their energy supplier of choice either at a fixed or flexible rate.

- **Grid operators:** the utility companies responsible for operating and maintaining power grids and gas networks in the Netherlands (Gerdes et al., 2017). The companies carrying these responsibilities on a national level (i.e. the high-voltage power grid and the gas transmission network) are TenneT (electricity) (TenneT Holding B.V., 2018) and Gasunie Transport Services (gas) (Nederlandse Gasunie N.V., 2018). By law, the Dutch government is the sole shareholder of both companies (Electricity law, 1998; Gas law, 2000). The smaller energy networks are managed by regional grid operators of which there are seven in total, the largest being Enexis and Liander (Energieleveranciers.nl, 2018; Gerdes et al., 2017). Grid operators should always be independent from suppliers and producers (Electricity law, 1998; Gas law, 2000).

- **Program responsible parties:** the parties responsible for providing daily forecasts of electricity consumption by (groups of) customers for the upcoming day at a quarter-hourly resolution (TenneT Holding B.V., 2014). Everyone with a connection to the electricity grid is obliged to provide such projections in order for the transmission system operator to be able to maintain the balance between production and consumption. For small consumers (i.e. the E1A, E1B and E1C profiles from Table 1 on page 14) the burden of program responsibility lies at their respective energy suppliers, who can either arrange this themselves or outsource the task to a third party (TenneT Holding B.V., 2014).

- **Transmission system operator (TSO):** the entities responsible for the task of congestion management, i.e. the maintaining the balance between electricity consumption and production. In the Netherlands, this task is carried out by national grid operator TenneT (TenneT Holding B.V., 2018). Congestion management is vital for the functioning of the electrical grid. As for now, there is no economic way to store electric power on a large scale, thus production and consumption need to be equal at any point in time, otherwise a power outage may occur. At the moment, TenneT achieves this balance by dispatching generators, based on the consumption forecasts provided by program responsible parties (Sagwal & Kumar, 2016; Pillay, Karthikeyan, & Kothari, 2015).\(^{(4)}\) It is therefore essential the predictions provided by the program responsible parties are accurate. The TSO

\(^{(4)}\)In the not-so-distant future, the increasing share of renewable assets (which are hard to steer due to the volatility of the weather) in the energy mix are expected to induce an overhaul of the current system in which demand is aligned with the available supply (Hartmann, Thomsen, & Wanapinit, 2018).
charges any anomalies between the forecasted and actual demand at the imbalance tariff (TenneT Holding B.V., 2014).

- **Metering organization**: the entity responsible for recording (smart) meter readings and passing these on to the grid operator. The metering organization owns all gas and electricity meters and often also organize any maintenance if necessary. Small consumers pay their metering organization through their energy supplier, whereas large consumers receive a separate bill (Gerdes et al., 2017). There are a number of certified metering organizations in the Netherlands. Grid operators carry the metering responsibility for households (Authority for Consumers and Markets, 2017).

- **Suppliers**: the intermediary parties between end consumers and the other market players. Suppliers buy their energy from a program responsible party and sell it to their customers. They also provide service to their customers concerning most energy-related matters (Gerdes et al., 2017). Small consumers pay all their energy costs through their energy supplier who then redistributes this money over all beneficiaries (see Section 6.1.1).

- **Supervisory body**: an independent organization monitoring whether all of the above market parties comply with the enacted regulations. In the Netherlands this task is carried out by the Authority for Consumers and Markets (Authority for Consumers and Markets, 2018a).

There may be some overlap between these parties. For example, many suppliers also produce some electricity and often have their own program responsibility, while grid operators often also do the metering work (Gerdes et al., 2017). There is however a strict separation between commercial parties (suppliers, program responsible parties and producers) and the grid operators. The energy market is expected to undergo a number of profound changes in the coming decades in order to be able to handle larger shares of renewables in the energy mix. A number of these may affect the techniques discussed in this text and will be briefly mentioned in Section 5.3.4.

### 3.2 The host organizations

This project was supported by energy suppliers Huismerk Energie and De Groene Stroomfabriek, which together with The Energy Trading Company\(^{(5)}\) jointly make up a tripartite energy company called the Energy Transition Group, situated in Nijmegen, the Netherlands. To provide some background context for this thesis, both companies and what they stand for will be briefly introduced.

#### 3.2.1 Huismerk Energie

Huismerk Energie is a Dutch supplier of gas and electricity active on the residential market. The company was founded in 2013 by Peter den Biesen and Loek Otters. It profiles itself as an innovative energy company which supplies 100% renewable energy from local sources only. Wind energy is their main power source, generated at several wind farms spread across the country, but their electricity portfolio also includes a biomass power plant, a small hydro facility, a number of solar farms and numerous smaller PV installations lying on the roofs of their customers. All in all, 88% of the electricity supplied by Huismerk Energie is from wind sources, 10% is from the biomass plant and 2% is solar generation (Huismerk Energie, 2018).

\(^{(5)}\)The program responsible party for both energy suppliers
Green gas (CO\textsubscript{2} neutral biogas) is not yet produced on scales large enough for widespread adoption (Huismerk Energie, 2018). As an interim solution, Huismerk Energie supplies CO\textsubscript{2}-compensated gas, which is ordinary natural gas from fossil origin, the greenhouse gas emissions of which are compensated through investments in projects that help reduce carbon emissions elsewhere. CO\textsubscript{2}-compensated gas is not renewable (it is natural gas after all, the reserves of which will be depleted at some point), however its harmful greenhouse component is made up for (PIANOo, 2017).

### 3.2.2 De Groene Stroomfabriek

De Groene Stroomfabriek is the business market oriented-equivalent of Huismerk Energie. They provide tailor-made energy solutions to all business customers willing to switch to renewable energy. Their energy mix has a larger share of biomass (90.1%), complemented by some wind (9.8%) and solar (0.1%) assets. Just like Huismerk Energie, De Groene Stroomfabriek supplies CO\textsubscript{2}-compensated natural gas (De Groene Stroomfabriek, 2018).

![Huismerk Energie and De Groene Stroomfabriek logos.](image)

**Figure 4:** The Huismerk Energie and De Groene Stroomfabriek logos.

### 3.3 Smart meters

Smart meters are expected to form a cornerstone in the future energy infrastructure (Alahakoon & Yu, 2016; Ebeid, Heick, & Jacobsen, 2017; Yan, Qian, Sharif, & Tipper, 2013). As a result, smart meter rollout procedures have been set into motion in many countries. This section provides an overview of various aspects of smart meters, including the progress and prospects of deployment in different countries, technical aspects, envisioned advantages over traditional metering systems for different stakeholders, and privacy issues. Following the latter, the Dutch privacy legislation related to smart meters will be discussed.

#### 3.3.1 Smart metering systems

A smart metering system consists of the following components:

1. A smart meter, the metering device which measures several variables related to energy consumption in the building.

2. The Central Access Servers where the measured data is stored.

3. Four communication ports through which system data is exchanged with other systems and authorized parties.
The four ports are referred to as the P1, P2, P3 and P4 ports (AlAbdulkarim & Lukszo, 2008). The P1 port is the consumer portal, a read-only plug reserved for local applications, such as a graphical user interface (GUI). The P2 port is used for communication with other metering instruments (e.g. gas and water meters in the building) (Schellebeek, 2012) and for grid operator equipment. The P3 port sends quarter-hourly and hourly consumption data to the Central Access Server from which the information is passed on through the P4 portal to network operators and authorized independent service providers and utility companies (Hoenkamp et al., 2011; Netbeheer Nederland, 2014). Figure 5 shows the various components and data exchange possibilities of smart metering systems.

3.3.2 P1 and P4 data

There are two ways to acquire smart metering data: it can either be extracted directly from the smart meter by plugging a cable into the P1 port, or it can be retrieved from the CAS via the P4 port. These types of data are commonly referred to as P1 and P4 data respectively and have different properties. Some characteristics of both data types are shown in Table 3.
**Table 3:** Some characteristics of P1 and P4 data.

<table>
<thead>
<tr>
<th></th>
<th>P1 data</th>
<th>P4 data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sampling interval</strong></td>
<td>(Electricity)</td>
<td>1-10 sec</td>
</tr>
<tr>
<td></td>
<td>(Gas)</td>
<td>5 min - 1 hour</td>
</tr>
<tr>
<td><strong>Meter readings</strong></td>
<td>(Electricity)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>(Gas)</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Active Power</strong></td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td><strong>Reactive Power</strong></td>
<td>Some</td>
<td>×</td>
</tr>
<tr>
<td><strong>Access to data</strong></td>
<td>Real-time</td>
<td>Delay of ±1 day</td>
</tr>
<tr>
<td><strong>Requirements</strong></td>
<td>Additional equipment</td>
<td>ODA certification</td>
</tr>
</tbody>
</table>

P1 data
The P1 port is the intended platform for future in-home energy solutions (van Bijnen, 2010). P1 data is real time comes at a sampling rate of about 1-10 seconds for electricity and 5 minutes to 1 hour for gas (van Wylick, 2016). P1 data contains active power measurements (van Bijnen, 2010) and newer types of smart meters also support measuring reactive power (see below) (Zeifman & Roth, 2011). Policy makers have high hopes for the P1 port as a basis for future home energy solutions, but the adoption rate of P1-based applications has been disappointing so far on both the supply and demands sides (Vringer & Dassen, 2016; van Elburg & Uitzinger, 2014). According to van Wylick (2016), the costs pose the main barrier on both ends of the spectrum: service providers fear high investment costs of P1 solutions and consumers are reluctant to pay large amounts of money for a solution which only provides them with some additional details about their energy consumption.\(^{(6)}\)

3.1 Remark: In alternating current (AC) circuits two types of electric power are often distinguished. Active power is the ‘useful power’ dissipated by a load, measured in Watt (W). It is what is measured by smart meters and the quantity used for the majority of P1 data applications in this text. Reactive power on the other hand does not do any work and is not consumed either: it is transferred back and forth between the source and the load. It is measured in Volt-Ampere reactive (VAR) (Alexander & Sadiku, 2013).

Running electrical appliances always consume a positive amount of active power. Some appliances also have a reactive component, which can be positive (e.g. in electric motors, transformers) or negative (e.g. in many computer electronics, radio circuits). Other appliances (e.g. a light bulb or heat pump) do not have any reactive component whatsoever (Alexander & Sadiku, 2013). Reactive power is charged to larger industries, but not to households (Authority for Consumers and Markets, 2018b). Nevertheless, it will have a number of uses regarding the applications in this thesis.

P4 data
At the moment, P4 data is requested by the energy supplier every two months for a consumption summary and once for the annual bill (van Elburg & Uitzinger, 2014). In light of the energy covenant signed earlier this year, the frequency at which the consumption summaries are provided is expected to double, leading to a total of thirteen standard read-outs of the P4

\(^{(6)}\)See Section 3.4.3 for prices and technical specifications of some existing P1 solutions.
port for each household (H. Kamp et al., 2017). Customers can sign a so-called certification overige diensten aanbieder (ODA), by which they authorize their energy supplier to acquire P4 data much more often for energy insight and more extensive billing purposes (van Bijnen, 2010). In practice, this is the road many utility companies and independent third parties have taken, mainly due to its approachability. However, the possibilities for utilizing P4 data are much more restrained, as the sampling rate is very low and the data comes with a delay of approximately one day (van Wylick, 2016). A number of the solutions proposed in this study\(^7\) are expected to underperform when used as a P4-based feature.

### 3.3.3 Deployment and penetration

**Rollout in other countries**

Smart meters have been widely deployed in several world regions and countries, e.g. Ontario (Canada), Victoria (Australia), Japan, South Korea and parts of the US (Hoenkamp et al., 2011). In Europe, Sweden, Italy and Finland were the first to adopt and smart meter penetration rates in these countries were already at 95-100% by 2014 (Zhou & Brown, 2017).

**Rollout in the Netherlands**

In the Netherlands, a national roll out of smart meters entered the political agenda in 2004, two years prior to a European Electricity Directive (Directive 2006/32/EC, 2006) obligating member states to initiate such a large-scale deployment (Hoenkamp et al., 2011). In 2008, two bills were submitted to the Dutch House of Parliament, in which technical standards were set and infringement penalties were proposed for consumers who refused the installation of their new meter (Zhou & Brown, 2017). After some alterations, both bills passed the House of Parliament in July 2008. However, at around the same time various civil organizations began raising questions about privacy and data security issues surrounding smart meters (Cuijpers & Koops, 2008). Several highly regarded institutions including the Committee for the Protection of Personal Data\(^8\), applied research institute TNO and the Consumers’ Association criticized various aspects of the proposed acts, the Senate ultimately declined to approve both bills in 2009 (Hoenkamp et al., 2011; Cuijpers & Koops, 2013). Eventually, an amended version of the bill (in which accepting the smart meter is no longer mandatory) was adopted in 2011, paving the way for the currently ongoing large-scale rollout (Van Elburg, 2014).

### 3.3.4 Benefits of smart meters

Smart meters provide several advantages over their traditional equivalent for many stakeholders. A few of the envisioned benefits which originally drove decision-makers to initiate the large-scale deployment of smart metering systems are given for different groups in the energy supply chain.

- **Energy suppliers** can assess consumption more accurately, reduce their operational costs associated with manual meter reading efforts and may mitigate the administrative burden of customer switching (Dimitropoulos, 2007). Also, smart meters facilitate experiments with different pricing schemes (Deconinck & Decroix, 2009; Hu, Kim, Wang, & Byrne, 2015) and the development of tailor-made energy solutions which may reduce customer churn rates (Rogai, 2006).

---

\(^7\)The peak load estimation from Section 6.3.2 and the edge clustering technique from Section 6.4.1 for example.

\(^8\)The College Bescherming Persoonsgegevens in Dutch, renamed later as the Autoriteit Persoonsgegevens (Personal Data Protection Authority).
• *Transmission system-* and *grid operators* expect to benefit from peak consumption reductions and an overall increase in demand-side flexibility, enabling them to achieve greater penetrations of renewable technologies (Strbac, 2008). The option to remotely disconnect and reconnect power inlets in a safe way is an important feature in future smart grid systems (Zheng, Gao, & Lin, 2013). Moreover, enhanced metering can aid power theft and outage detection endeavours, provide more detailed power quality data and improve network investment decisions (Dimitropoulos, 2007).

• *Governments* hope smart meters live up to their expectations regarding energy efficiency improvements (See Section 3.4), which may help bridge the gap towards achieving carbon reduction targets (McKenna, Richardson, & Thomson, 2012).

• *End consumers* hope to reduce their energy costs when personalized smart meter-based feedback hands them pointers for easy ways to save energy (Darby, 2006; Dimitropoulos, 2007).

3.3.5 Privacy concerns and legislation in the Netherlands

Opposed to the various advantages smart meters provide over traditional metering devices, there are also several negative voices, mainly regarding consumer privacy. The general opinion on the privacy aspects of smart metering data utilization differs greatly between different countries. In the Netherlands and Germany there is deep public concern about smart meter privacy and data security, while in Scandinavian countries there is hardly any resistance against unbridled utilization energy data (Zhou & Brown, 2017).

Smart meter data can potentially be used for remote monitoring of individual consumption habits. Many different spying opportunities involving multiple actors have been identified in literature (e.g. Lisovich, Mulligan, & Wicker, 2010; Quinn, 2009; Cavoukian, Polonetsky, & Wolf, 2010). A few example concerns have been bundled by McKenna et al. (2012) such as:

• Burglars finding out when homes are unoccupied.

• Commercial organizations targeting specific individuals based on their energy consumption behaviour.

• Insurance companies performing personalized risk assessments based on energy consumption data to adjust their insurance policies.

• Law enforcement agencies using energy data to detect illegal activities (e.g. drug labs).

• Partners investigating each other’s behaviour.

In the Netherlands the utilization of personal data is currently regulated under the *Personal Data Protection Act* (Wet bescherming persoonsgegevens, 2001) and the compliance with this law is monitored by the Personal Data Protection Authority. From 25th May, 2018 on however, the European General Data Protection Regulation (Directive 95/46/EC, 2016) will take effect, replacing all national data privacy regulations including the Personal Data Protection Act above. Nevertheless, the Personal Data Protection Authority remains responsible for the relevant legal compliance in the Netherlands and maintains its powers to impose fines in case of a violation.
In order to assess the relevant privacy aspects of the smart meter-based techniques described in this thesis, an interview with Huismerk Energie’s consulting solicitor Karine Koster was conducted (personal communication, 22nd January, 2018). Ms. Koster named two key aspects of the privacy side of energy data utilization in general, based on the upcoming European act. First and foremost, the implementing organization must always have the customer’s unequivocal permission to use their personal energy data. An ODA certification covers this aspect, i.e. upon signing the certification, customers authorize their energy supplier to utilize their consumption data in order to generate more detailed feedback. Secondly, the implementing organization should communicate to its consumers in a transparent way, exactly how and for what purposes the data is used. Most organizations include this in their privacy statement.

With respect to the division of customers into subgroups based on additional characteristics such as their dwelling type, Ms. Koster remarked such endeavours are allowed as long as no data can be traced back to individual customers. When using cluster-wide figures (such as a cluster’s mean value) in comparative efforts, a rule of thumb is that each group should have at least three members to ensure the requisite anonymity.

3.4 Smart meter feedback

Potential energy savings have been one of the main drivers for large scale rollouts of smart meters in many European countries (Hierzinger et al., 2012; Wilson, 2015). In pre-rollout effectiveness assessments in the Netherlands, it was stated a nationwide introduction of smart meters in all buildings would yield energy savings within the range of 3.5-6% (Gerwen, Koenis, Schrijner, & Widdershoven, 2010). By now, people have come to see that a smart meter by itself is nothing more than an advanced metering system which does not yield any meaningful energy savings unless the data it generates is sculptured into something consumers understand. This is referred to in literature as smart meter feedback.

3.4.1 Transtheoretical model of change

Smart meter feedback is part of the process of achieving energy efficiency through behaviour change, an attempt to affect people’s habits or their choices of electric appliances in order to reduce end-consumption. A number of studies have shown energy use in identical homes may differ by a factor two to three (Parker, Hoak, Meier, & Brown, 2006; Winett, Neale, & Grier, 1979; Socolow, 1978), which is why behavioural aspects of energy efficiency have been addressed in many studies (e.g. Froehlich, 2009; Fischer, 2008; Darby, 2006). A model which has been cited in the context of energy efficiency by several authors(9) is the Transtheoretical Model of Change (TMC) from Prochaska and DiClemente (1986), which describes the ‘stages of change’ an individual goes through when engaging in the process of intentional behaviour change. A three-stage version distilled from the original TMC is presented by Riche et al. (2010) will be adopted here, as this adaptation is tailored to the context of domestic energy use. The three stages of the adapted model are raising awareness, informing complex changes and maintaining sustainable routines, shown in Figure 6 below. As with the original TMC, none of the steps can be skipped, but completed stages can be revisited in case of a relapse (Riche et al., 2010; Prochaska, DiClemente, & Norcross, 1992). Riche et al. underline that any energy feedback system should support these three steps.

(9)See e.g. H. A. He, Greenberg, and Huang (2010); Riche et al. (2010); Scherling (2017); R. A. Howell (2014).
A recent situation in the Netherlands indicates many customers have not yet completed the awareness stage when it comes to energy consumption. Over 3000 citizens claimed having a faulty meter after Dutch TV show Tros Radar reported on a study which showed electromagnetic signals from household appliances can interfere with smart meters, resulting in abnormally high consumption measurements (Leferink, Keyer, & Melentjev, 2016). However, follow-up research conducted by Netbeheer Nederland revealed none of the callers’ discrepancies were caused by electromagnetic interference. In most cases it turned out consumers had just been confronted with their actual consumption for the first time. Consumption was especially higher than expected when the new meter had been placed in winter, just after the birth of a child or after the installment of a new appliance (Netbeheer Nederland, 2017). A series of qualitative studies surveyed by Froehlich (2009) endorse the conclusion that, as of yet, consumer awareness about personal energy usage is generally unsatisfactory. Most respondents were unable to properly estimate their total energy consumption or disaggregate their energy use per appliance. Also, many people (incorrectly) believed it to be better to leave the heating on overnight because re-heating the home in the morning would lead to overall higher energy usage. The main observation made in many text on energy conservation is that energy consumption is something abstract which goes without notice (Fitzpatrick & Smith, 2009; Twyman, Smith, & Arnall, 2015; Whitmarsh, Seyfang, & O’Neill, 2011; Dowd, Itaoka, Ashworth, Saito, & de Best-Waldhober, 2014). Another important observation is that (at least in an OECD country such as the Netherlands) one does not run out of energy, unlike most goods, which gradually decrease in quantity or quality upon using them (Froehlich, 2009). There are still good reasons to conserve energy (e.g. energy consumption has financial and environmental implications), but all of these are indirect.

In view of the above empirical results and the TMC model’s left-to-right structure (i.e. the requirement that one needs to complete a stage in order to proceed to the next one), one may conclude that the awareness stage requires some more attention if one wants to accomplish persistant energy savings. Therefore, this study will be mostly ‘left-oriented’ in terms of the TMC: the majority of applications discussed in this thesis are designed to inform consumers on specific aspects of their consumption (thus raising awareness), few are meant to directly induce

\[\text{(10)}\] As Crabb (1992) formulated it: “People do not use energy, they use devices and products.”
changes (e.g. the refrigerator payback time and the insulation quality assessments from Sections 6.2 and 5.3.4), while attention will be paid to long-term effects only sporadically.

### 3.4.2 Feedback

Increased awareness and possible subsequent energy savings can be achieved through the provision of feedback, defined by Karlin, Zinger, and Ford (2015) as “the process of giving people information about their behaviour that can be used to reinforce and/or modify future actions”. In the smart meter setting, there are many different ways to set this up. At the moment, the only smart meter feedback mandatory in the Netherlands relates to enhanced billing, i.e. sending consumers a bimonthly digest of their energy costs (instead of one single bill at the end of the year) (H. Kamp et al., 2017). This by itself is not a very effective feedback method, as is also reflected in the poor initial smart meter-related savings achieved in the Netherlands (less than 1%) (Uitzinger & Uitdenbogerd, 2014).

There is a large body of literature on the effectiveness of different types of smart meter feedback. A large number of field experiments have been conducted towards the effectiveness of different applications and GUI output designs regarding energy conservation. One particularly extensive study towards this topic which bundles numerous different aspects together has been conducted by Froehlich (2009) and will provide the main source for this section. Froehlich (2009) has identified ten “design dimensions” which affect the effectiveness of energy feedback. The author himself acknowledges some of these dimensions overlap with each other, which is why some related dimensions have been grouped here. What remains are six categories: granularity, output, availability, persistence, design and level of interaction, each of which will be discussed in a smart meter context, with links to some of the applications discussed later in this text. In this section, there will be limited attention to the potential energy savings associated with specific solutions though, as this will be done for each solution individually in Chapter 6.

**Granularity**

The first category relates to the degree of precision of the feedback regarding a number of different variables. This concerns mostly time-related aspects, such as the possibility for the user to ‘zoom in’ on his or her data, which is of course limited by the frequency of the measuring installation. In the smart meter setting, P1 data allows for very detailed figures containing information on sudden changes, whereas P4-based feedback is bounded by 15-minute windows and gives a more generic overview of any events. Although there are little empirical results on the best data resolution of energy feedback, real-time feedback has been found very effective (see Availability below) and inherently uses a higher data frequency.

Precision aspects of feedback are not limited to time-related elements though. One can for instance also zoom in on the building that is measured and show room- or appliance-specific feedback. In most cases, this is achieved by deploying additional submetering components which measure every room/device separately. However, the emerging field of non-intrusive load monitoring is occupied with deducing the energy consumption by individual appliances from a main load signal without using any additional measuring equipment, as will be discussed in Section 6.2. Appliance-specific feedback has been found to yield very high energy savings in the

---

(11) After this clustering, there is still some overlap between the categories though.
range of 7-12% (Armel, Gupta, Shrimali, & Albert, 2013; Chakravarty & Gupta, 2013; Ueno, Inada, Saeki, & Tsuji, 2005).

**OUTPUT**
The second aspect that determines the effectiveness of energy feedback is the extent to which users can relate to the feedback output that is presented. In general, it pays off to convert certain amounts of energy to a different physical quantity that comes with a more perceptible measurement unit, as this gives users different object to compare their energy use to (Wood & Newborough, 2007). Apart from the ordinary energy units (mostly kWh for electricity and m$^3$ for gas) two possible categories are often found in literature, namely monetary units (energy consumption converted to their associated costs) and environmental units (usually CO$_2$ emissions associated with a certain amount of consumed energy). These will be discussed in Sections 6.1.1 - 6.1.3. Apart from adjusting the measurement unit, feedback output can be presented as a relative amount, compared to one’s own past consumption data for instance, or that of others. These topics will be discussed in Sections 6.5.1 - 6.5.3.

**Availability**
The availability of feedback is another important determinant of the effectiveness of feedback. Availability can have a number of different meanings in a feedback context. Darby (2006) distinguishes between direct feedback, i.e. unprocessed energy information which is provided right away, and indirect feedback, which is processed and comes with a certain delay. To place these concepts in a smart meter context: many P1 applications count as direct feedback as P1 data is extracted from the meter instantaneously, whereas all P4 applications fall in the category of indirect feedback, due to the one-day lag inherent to such measurements. As such, P4-based feedback is less available from a time-perspective than P1-based feedback, as one can only request data that is over a day old.

Another important aspect of this is the presentation medium used to display the feedback and the location of this medium. According to (Froehlich, 2009), electronic media are preferred over traditional paper, because the latter “is non-interactive and can only display static information.” However, there numerous different possibilities for digitally displaying smart meter feedback, i.e. via e-mail, a web page or a smart phone app, or by means of an additional in-home display specifically installed for the purpose of providing feedback. The latter seem to be especially effective: a survey of twelve different studies towards the effectiveness of in-home displays in terms of energy conservation showed savings of 2-18% (Faruqui, Sergici, & Sharif, 2010). However, it should be noted that all of the systems provided various types of direct (real-time) feedback, which by itself is already quite effective (Darby, 2006; Vringer & Dassen, 2016), making it a bit inconclusive as to what part exactly is linked to the medium on which the feedback is displayed and what part is induced by the feedback itself. Smart phones have also been quoted by many as ideal platforms for energy feedback, due to their portability and connectivity (Coşkun & Erbüğ, 2016). Moreover, they are the preferred feedback medium among consumers (M. Weiss, Loock, Staake, Mattern, & Fleisch, 2010).

Scientists have also been experimenting with decentralized appliance-specific feedback, by attaching a display to one or more appliances. McCalley and Midden (2002) reported savings of as much as 21% on washing-related energy use after installing such screens to the washing machines of participant. Consumers on the other hand prefer one central feedback system over separate displays for each device (Fitzpatrick & Smith, 2009).
Persistence
Related to the availability is the persistence of a feedback system. Systems can either be passively waiting for the user to log on and see the feedback, or actively send (parts of) the feedback to the user via e-mail or any other notification system. Studies show light active pushes every now and then can be very effective, but overdoing it or sending irrelevant information too often can lead to counter-effective results (Froehlich, 2009). On a similar note, the extent to which the system attempts to have the user engage in specific action increases the success rate of feedback. The human brain tends to prioritize information that is concrete and personalized (Borgida & Nisbett, 1977). From smart meter data one may derive a lot of information that can be used to create personalized feedback, which will be the topic of Section 6.5.4.

Design
The graphic design of a feedback system also affects the success rate of its applications, yet few studies have addressed this dimension. One very extensive empirical study on this topic however is the one by Chiang (2015), who has assessed the influence of various design aspects of energy managements system displays on the effectiveness of such systems. Three different designs for communicating changes in energy consumption were tested in that study, namely:

1. A numerical design, in which changes are communicated as raw numbers, displayed in different colours: ordinary black for displaying average consumption figures, green to address a substandard consumption value, and orange to display phases of relatively high energy use.

2. An analogue design, in which the numbers were replaced by an analogue meter (such as the traditional speedometers in cars). Rather than showing numbers, the scale was divided in differently coloured areas, using the same colours (fulfilling the same role) as in the numerical case.

3. A faces design, where consumption was shown by means of smiley faces, showing a happy face if less energy was consumed and a unhappy one if consumption was high. Additionally, the smileys took on different colours, with again green being used to address frugal use and orange assigned to display high energy consumption.

The numerical design turned out to be most effective for conveying energy information, whereas the faces design was least effective in this aspect. However, in terms of drawing user’s attention, there was strong indication that the faces design worked better and the faces design also realized the highest savings of the three design types. In terms of user preferences, the numerical design was highly liked (over 80% of the participants liked this design), while the faces design was strongly disliked by almost 80% of its users.

In another survey conducted by (Rodgers & Bartram, 2011), a number of more abstract artistic design choices were implemented. It was found that many respondents are sympathetic to such alternative representations of energy use, different from traditional graphs or depictive graphics. From the above studies, one can conclude that the design of an energy feedback system might indeed affect the effectiveness of such systems and that it may be worthwhile to experiment with different design choices.

Level of interaction
Finally, Froehlich argues adding an option to share one’s energy saving-achievements on social
media may lead to peer pressure phenomena and trigger others users to also behave more frugally. (Petkov, Köbler, Foth, & Krcmar, 2011) wonder why energy savings “still remain a lone activity in our interconnected world.” Nevertheless, estimates of the effects of social media options in scientific studies are rather limited and do not provide a unified answer to the question of the effect social media features have on consumption (Bull, Lemon, Everitt, & Stuart, 2015).

A second dimension on interaction sometimes discussed in literature is that of gamification, i.e. having consumers compete with others to find out who can achieve most reductions. In some ways, gamification is an extension of the comparative feedback and goal-setting features that will be discussed in Sections 6.5.1 and 6.5.3. Orland et al. (2014) found that a users really drawn into a game decreased their consumption by 13%. Noteworthy was the large difference between the effect on work days (7%) and non-work days (23%). Interested readers may find a number of game designs in Fijnheer and van Oostendorp (2015).

3.4.3 Existing energy management systems

Energy management systems are beginning to flourish, profiting from the fast-growing coverage of smart meters. To get an impression of existing marketable technologies, a total of 52 energy management systems have been surveyed in this thesis. The information originates from several company websites and from (Milieu Centraal, 2018a), an initiative of Dutch energy organizations Milieu Centraal, Netbeheer Nederland and Energie Nederland. The majority of these systems are available in the Netherlands and most use smart meter data. A total of 24 systems use the P1 port, 15 rely on P4 data, 12 on external measuring clamps or pulse meters and 1 on manual input by the user. All in all, 36 devices are able to provide realtime, direct feedback on a smart phone, web page or separate display. The majority of systems measure both electricity and gas (some additionally also measure water consumption), but 11 devices focus only on electricity use and 2 are gas-specific.

Functionalities
The available features differ greatly from one device to another. The majority of the applications examined are able to convert energy consumption to costs. In most cases, this requires the user to manually insert their energy tariffs, but some supplier-coupled devices do this automatically based on the energy contract. Only four of the devices surveyed have a function for converting energy consumption into associated CO₂ emissions.

Almost all of the applications feature a ‘historic feedback’ option, although the exact possibilities for this function vary from just having access to old data to more comprehensive statistics. A total of 23 products comes equipped with a ‘comparative feedback’ functionality, allowing users to compare their energy use over different periods to that of other users. Besides, a goal-setting feature was found in 20 products and 24 systems provide energy saving tips (either personal or generic).

By default, only one product measures solar generation from PV panels (Cohere Energy Solutions’s Maxem, which is designed specifically for households with solar panels and electric vehicles (Cohere Energy Solutions, 2018)). A total of 28 systems can be expanded with additional clamps or wiring for solar generation assessments, however these expansion packs require

(12) Allowing a consumer to compare his energy usage between different periods.
an additional investment by the consumer (see the Cost section below). Eighteen systems are
also able to measure the consumption of individual appliances (and in some cases steer them
too), however these smart plugs have to be purchased separately as well. Six systems come
equipped with a few smart plugs. Five devices are able to assess the energy usage of individual
appliances from the aggregate signal (this is called non-intrusive load monitoring, which will be
discussed in Section 6.2). Besides, a number of less common functions were found among the
systems examined. These include:

- A standby power assessment feature, which measures the total energy consumption from
  all appliances in standby or continuous-power mode (e.g. Your2Power’s BeeClear device
  (Your2Power, 2017) and Greenchoice’s BOKS Live (Greenchoice, 2018)).
- A standby killer option, which automatically shuts down smart plug-enabled appliances
  whenever they are in standby mode too long (E.ON, 2016).
- A peak consumption estimation functionality (2Wire’s MEMo, (2Wire, 2018)).
- The ability to operate smart home appliances (e.g. Smappee (Smappee N.V., 2018)),
electric vehicle charging processes (the aforementioned Maxem device (Cohere Energy
  Solutions, 2018)) and/or a smart thermostat (Eneco’s Toon for instance (Eneco Beheer
  N.V., 2015)).
- An ‘unusual consumption’ detector (Delta’s Delta Comfort Wijzer has this feature (Delta
  N.V., 2013)).
- A limiting option, which sends a notification whenever energy consumption exceeds a
  preset threshold (e.g. GoGreen’s Totaal Verbruiksmeter (GoGreen B.V., 2012)).
- Weather data comparison functionalities (e.g. the ECOhome and Mindergas.nl applications
  (ECOhome, 2018; Mindergas.nl, 2018)).
- Energy cost (Umeter, 2018) and PV generation prediction features (Qurrent’s Qbox (Qurrent,
  2016)).

Costs
Prices vary greatly among the systems surveyed here. Of the 52 devices, 16 come with no
acquisition costs, although most of these oblige the user to enter into an agreement of some sort
with monthly charges of its own. A total of 10 devices are (semi-)exclusive to customers of an
energy supplier. Otherwise, all price ranges are roughly equally well-represented, see Figure 7.

Figure 7 also shows the costs of any available solar and smart plug expansion packs. Most
smart plugs cost about 25 euros a piece, but there are outliers both upwards and downwards.
Smappee’s smart plugs cost €11.- a piece and two are provided with the main product itself
(which in turn is one of the more expensive ones among the devices examined, at a cost lying
between €229.- (basic product) and €599.- (Smappee Plus)), while 2Wire’s MEMo smart plugs
cost €99.- each. Solar expansion packs are available at prices mostly in between €25.- and
€100.-, but many of them need to be installed by a technician and the above prices do not
include such installation fees.
Figure 7: Purchase prices and any available solar and smart plug expansion pack prices among the examined devices.

3.5 Rebound effects

Suppose a new type of car fuel is developed, allowing car owners to drive more kilometres per litre than before. Assuming the costs of the new fuel per litre are equal to those for traditional gasoline, the fuel costs per kilometre fall, thus car owners will save on their mobility expenses. If they spend this money on anything requiring energy, this reduces the net energy savings from the fuel improvement. This is an important phenomenon in the context of energy efficiency, generally referred to as the rebound effect or take-back effect, a schematic depiction of which is given in Figure 8 below.

Figure 8: Schematic representation of the rebound effect, adapted from Santarius (2012).

The rebound effect was first described in 1865 by Jevons, who observed technical improvements on steam engines resulted in an increase of coal consumption rather than a decrease (Jevons, 1906). Although the subject initially did not receive much attention, Khazzoom (1980) revived the topic and it has been subject to a large number of studies ever since. However, these studies are rather divergent in terms of definitions, methodological approaches and data sources (Abdessalem & Labidi, 2016). Also, though general consensus on the existence of the effect has
been reached by now, the nature and size of the phenomenon are still under debate (Greening, Greene, & Difiglio, 2000; Sorrell, Dimitropoulos, & Sommerville, 2009).

As for the relation to this text, rebound effects may occur following feedback-induced energy savings too. Consider for example a household that manages to cut their energy consumption by 5% upon receiving more detailed feedback on their consumption. Such energy savings induce substantial financial savings. Whenever the family decides to spend some of these savings on a luxurious espresso machine, part of the original reduction will be offset: the espresso machine probably uses more electricity than the family’s old coffee maker and some energy is embedded in the espresso machine itself (related to the manufacturing process and the materials used).

The rebound discussion presented here will be approached mostly from a descriptive rather than a quantitative angle and is primarily meant to shed a different light on potential feedback-induced energy conservation and temper any excessive optimism regarding such savings. Additionally, a number of possible ‘anti-rebound’ side effects, i.e. effects that boost energy savings, are discussed in Section 3.5.3. Some empirical estimates will be presented in Section 3.5.4 to give the reader a feeling for potential sizes of different effects, but the outcomes of different studies vary considerably, making it difficult to pinpoint an exact range. Because of this, and the fact that it hard to say what types of feedback discussed in this text are most likely to trigger which rebound effects, rebound energy will not be included as a determinant for the final ranking of applications in Chapter 8. However, a number of measures to help reduce rebound impacts and foster beneficial side-effects will be discussed in Section 3.5.5.

3.5.1 Definitions

Definitions of the rebound effect found in literature vary widely. In their frequently cited text, Greening et al. (2000) describe the effect as follows: “Gains in the efficiency of energy consumption will result in an effective reduction in the per unit price of energy services. As a result, consumption of energy services should increase, partially offsetting the impact of the efficiency gain.” While this definition may be a good representation of the more classical rebound viewpoint, it is too restrictive for the intended applications in this section as it explicitly implies the effect is triggered by a financial mechanism, which is not uncommon to include in the definition (e.g. Khazzoom, 1980; Maxwell et al., 2011), but excludes a number of more recently identified types of effects (see Section 3.5.2).

A more flexible definition is given by Van den Bergh (2011) who defines the rebound effect as “the phenomenon that greater energy efficiency, or plain energy conservation through changes in behavior or choices (by firms or consumers), triggers additional energy use so that the net effect on total energy use over time becomes uncertain”. This description does not commit to a certain type of effect: any erosion of projected energy savings counts as rebound effect. Also, it captures a number of relevant aspects such as the fact that there may be a difference between the reactions of organizations and residential consumers when it comes to energy conservation.

The importance of making such a distinction is underlined by Sanne (2000), who argues the perception of individuals may be influenced by things such as leisure, social recognition and beauty, whereas firms primarily seek to increase their profits. A second element that can be noted from Van den Bergh’s definition is that rebound effects may be variable over time: an attempt to change one’s consumption behaviour through feedback may have the desired effect on a short notice, but the attention to the feedback may get ‘backgrounded’ on the long run.
A distinction between short- and long-term rebound effects is made explicitly in Jägerbrand, Dickinson, Mellin, Viklund, and Dahlberg (2014).

### 3.5.2 Typologies and a classification

Many different types of rebound effects can be found in literature. One particular distinction made in the majority of texts is that between *direct* and *indirect effects*. When an efficiency improvement of a certain energy service results in an increase in demand of that same service, we speak of a *direct rebound effect*, whereas an increase in demand of any other service or good that needs energy to be produced is referred to as an *indirect rebound effect* (Freire-González, 2017). In relation to the car fuel example given earlier, a direct effect would be driving more, while spending the saved money on eating more meat would be an indirect effect.

Macro-economic studies focus on *economy-wide rebound effects*. Greening et al. (2000) for instance, remark that “A fall in the real price of energy services may reduce the price of intermediate and final goods throughout the economy, leading to a series of price and quantity adjustments, with energy-intensive goods and sectors likely to gain at the expense of less-energy-intensive ones.” In some studies the term *transformational rebound effects* is used to describe the effects technological improvements have on overall consumer preferences and indirectly on energy consumption (e.g. Greening et al., 2000; Borenstein, 2013). Macro-economic rebound types however, will not be discussed here, since the smart meter solutions devised in this text primarily affect the energy consumption of individual consumers and are not expected to trigger any significant shifts on larger scales.

Building on the work of Van den Bergh (2011); Jenkins, Nordhaus, and Shellenberger (2011) and Paech (2011), Santarius (2012) devised a classification that partitions thirteen different rebound mechanisms into four different categories, namely *financial effects*, *material effects*, *psychological effects* and *cross-factor effects* (see Figure 9). Systematic divisions of such scale are rare and become outdated relatively quickly as new types of rebound mechanisms are identified every now and then. It would therefore be rash to claim Santarius’ classification captures all elements of rebound theory. However, for the moment it is the most extensive division known to the author and will therefore be used as a basis for the discussion.

Some modifications have been made regarding the terminology used in Santarius’ model. The term “cross-factor effects” has been replaced by *process-efficiency effects*, which arguably captures the content of these mechanisms better.\(^{(13)}\) Also, because psychological effects are a relatively novel and less-polished aspect of the rebound discussion, these will discussed based on a few more recent studies. Specifically, the terminology regarding psychological effects has been updated to a later work by the same author Santarius and Soland (2016).

As mentioned already, the focus here will be mainly on financial, psychological and material mechanisms that act on a micro-economic level, as these are closest to the scope of this text (smart meter solutions targeting the behaviour of individuals and firms). All remaining effects will be mentioned briefly, but will not be discussed in detail.

\(^{(13)}\)Accordingly, the terminologies “cross-factor effect”, “material cross-factor effect” and “multiple cross-factor effect” have been replaced by “labour efficiency effect”, “mechanization efficiency effect” and “economy-wide efficiency effect” respectively.

---

35
Financial and Psychological Mechanisms

**Financial mechanisms**

The classical perspective on the rebound effect involves a change of cash flows somewhere in the process. Santarius (2012) distinguishes between *income effects* and *reinvestment effects*, which actually refer to the same mechanism: energy savings leading to a real income gain which results in additional spending on something that involves additional energy usage. The difference between the two is that the income effect refers to the reaction of individuals on energy-induced money savings, whereas the reinvestment effect relates to investment decisions made by organizations. Both types also have a positive counterpart, i.e. the freed-up capital can also be spent on something that reduces consumption even further. This will be discussed in Section 3.5.3. As all energy-savings from smart meter feedback directly lead to monetary savings, both the income and reinvestment effect may occur for all smart meter applications discussed in this text.

**Psychological mechanisms**

Psychological or *mental* rebound effects (Girod & De Haan, 2009) are a relatively novel aspect of the rebound discussion, set in motion by changes in the symbolic meaning of a certain technology as perceived by customers (Santarius, 2012). The first type up for discussion is the *attenuated consequences effect*. The underlying theory used to explain this effect draws upon the theory of *mental accounting* from social psychology, defined by Thaler (1999) as the set of cognitive operations used by individuals and households to organize, evaluate and keep track of financial activities. This is done from *mental budgets*, described by Krishnamurthy and Prokopec (2009) as “self-specified allowances for behaviors”. A mental budget can be seen as an imaginary bank account running in parallel to the real one, which influences spending decisions made by individuals. According to Heath and Soll (1996), the fact that budgets cannot perfectly anticipate consumption opportunities, people may earmark too much or too little money for a
particular category, leading them to overconsume or underconsume in that category. Girod and De Haan (2009) argue similar bookkeeping mechanisms exist for how an individual perceives the environmental impact of their daily operations. In other words, people earmark a certain amount of environmental damage to individual activities and ‘spend’ from their environmental mental budget when they carry out these activities. The theory of environmental mental accounting is supported by Gruener and Hirschauer (2016); Parag and Strickland (2009).

The theory of mental accounting can be used to explain the attenuated consequences effect. This effect arises when “an efficient improvement on a certain technology leads to a re-evaluation of personal monetary, social or emotional consequences of using that technology, increasing the motivation to use it” (Santarius & Soland, 2016). In other words, the perceived financial or environmental costs of using a certain technology drop, which lowers the barriers to use it. Empirical evidence of this effect can be derived from a study by Ohta and Fujii (2011), who found that users who adopted a more environmentally friendly hybrid Prius car drove 1.6 times more than they did with their previous cars. As for an example in the setting of smart meter feedback, suppose Huismerk Energie decides to send its customers information on their energy-related CO₂ emissions. As Huismerk Energie supplies green energy only, customers may get the feeling their energy consumption does not have a large impact on their carbon footprint, i.e. the perceived environmental costs of using energy supplied by Huismerk Energie drop, which means consuming the same amount of energy now uses up less of the environmental budget. The attenuated consequences effect manifests itself when a customer starts to ‘spend’ the remaining budget by using more energy.

The other two effects discussed occur “via a decrease in feelings of responsibility to refrain from environmentally damaging behaviour” following a comparison of individual responsibility against either the personal moral balance (‘good deeds vs. bad deeds’) or that of others (Santarius & Soland, 2016). The moral licensing effect turns up when a ‘good deed’ licenses a customer to engage in environmentally damaging behaviour in another area. Santarius (2012) provides some examples, e.g. the purchase of an energy efficient car licensing a holiday overseas (requiring air travel). The diffusion of responsibility effect occurs when an individual compares his or her personal moral balance to that of others. Santarius and Soland (2016) describe this as follows: “Due to the purchase or use of an efficiency-improved technology, perceived responsibility for protecting the environment through frugal usage of that technology diffuses to other agents (e.g. engineers, policy makers, other consumers as potential adopters of efficient technologies), which leads to a decreased motivation to frugally use that technology.” In a feedback context this may occur when consumers at some point have made efforts to reduce their energy usage and get the feeling they have “played [their] part and that it is now up to others to do the same” (Santarius & Soland, 2016). A special case of this effect in a feedback setting is the boomerang effect discussed by Vine, Buys, and Morris (2013) in the setting where energy consumption is compared within a group of consumers (i.e. comparative feedback, see Section 6.5.1). While users with a relatively high consumption may indeed lower their consumption following such feedback, users who originally used less than the group’s average sometimes increase their consumption after having seen they are behaving ‘overly frugal’. Empirical evidence of this effect is found in (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007), who interestingly enough also found

(14) Considering the limited understanding people generally have of the consequences of their behaviour for the environment (see Section 3.4.2), the gap between their mental budget and their actual spendings (i.e. their environmental footprint) is likely to be fuzzier than in the original financial setting.
that households who consumed less than the average and received an encouraging message (in the form of a smiley face) maintained their low consumption, opposed to the group of below-average consuming households who did not receive such messages.

Material effects
Material effects correspond to a range of equipment-related rebound mechanisms. The embodied energy effect occurs when a certain energy efficiency improvement requires a material energy investment. The emission figures from various renewable technologies used in Table 9 which are allocated to the production of solar panels, wind turbines etc. are examples of this. The consumption accumulation effect, occurs when a certain asset is replaced by a more efficient one and the old item remains operative somehow. This happens for example, when a new, efficient car is bought, and the old one is sold online or passed on to a family member. This effect may also occur as a consequence of the refrigerator payback time application from Section 6.2. If a consumer invests in a new refrigerator because of long-term financial benefits, it may be that they decide to decrease the payback time somewhat further by selling the old model, in which case it will continue to drain energy elsewhere.

Other mechanisms
The remaining effects that bear too little relevance to this text will now be briefly touched upon. This includes a number of macro-economic effects, as well as the entire class of process efficiency effects.

Economy-wide effects
As pointed out already, some of the effects act on a macro-economic level. The market price effect for instance occurs when an energy efficiency improvement causes an economy-wide price drop for an energy(-related) product, increasing demand for that product. Another example is the new market effect, which relates to a certain innovation giving birth to a new market, requiring a whole new production chain to be set up (which naturally comes with energy implications). The ongoing evolution of the industry of electric vehicles is an example of this effect. Both of these effects lie outside the scope of this text though.

Process efficiency effects
When an efficiency improvement allows a certain procedure to be carried out more efficiently, demand for energy often increases (Santarius, 2012). One mechanism, discussed in e.g. Bin- swanger (2001); Jalas (2002), which may explain this relation is the situation where an efficiency improvement yields time savings on a process, while the energy requirements remain unchanged. If so, the same amount of energy is used up in a shorter time frame, thus if any energy is consumed during the time that has become available, this means a net increase in energy use. This may also occur if the energy requirements of the process drop due to an efficiency improvement, but the time savings of an efficiency improvement are greater than the energy savings.

Process efficiency effects can occur on the production side (in labour and mechanization improvements), but also on the consumption side (for instance if a consumer gains a higher internet connection speed more data will be exchanged within the same time frame, inducing higher energy use at data storage facilities) and on a macro-economic level (if a higher internet connection speed is realized on a national scale for instance). The smart meter-based solutions proposed in this text most likely do not yield process efficiency improvements of any kind. Therefore, these types of rebound effects will receive no further attention.
3.5.3 Beneficial effects

As opposed to rebound effects, energy efficiency improvements can also induce beneficial side effects\(^{(15)}\) leading to additional energy conservation on top of the predicted consumption reductions incited by the technological improvement. Santarius and Soland (2016) describe two such effects, namely improved responsibility effects and improved control over frugal use effects. The improved responsibility effect is described as the positive counterpart of both the moral licensing and diffusion of responsibility effects, i.e. “the key process in the improved responsibility effect is an increase in felt responsibility (responsibility compared to others or compared to the individual moral balance)” (Santarius & Soland, 2016). The improved control over frugal use effects on the other hand occurs via “a re-evaluation of the behavioural control over using [a] technology in frugal ways, which decreases the motivation to purchase/use that technology.”, however Santarius and Soland (2016) were unable to provide empirical support for this effect.

The improved responsibility effect partially overlaps with what is termed in many other studies as pro environmental spillover, i.e. “[the] effect of an intervention on subsequent behaviors not targeted by the intervention” (Truelove, Carrico, Weber, Raimi, & Vandenbergh, 2014). Spillover of pro-environmental behaviour has been the subject of numerous different studies for decades (e.g. Frey, 1993; Thøgersen & Ölander, 2003; Margetts & Kashima, 2017). There may be pointers for stimulating spillover by means of careful framing of the feedback, as will be discussed in Section 3.5.5. Finally, some authors mention the prebound-effect, often discussed in the context of home-heating, describing the phenomenon that occupants of poorly isolated homes tend to behave more economically with respect to their space heating (Sunikka-Blank & Galvin, 2012; Teli et al., 2016). This effect may have indirect implications for the current study however, in the sense that it seems unlikely this effect can be encouraged with energy feedback, while a loss of this prebound effect counts as a rebound effect and reduces potential savings. More specifically, if consumers who have been behaving overly-economical according to this prebound supposition decide to upgrade the insulation quality of their dwelling based on the insulation assessment discussed in Section 5.3.4, they may lose part of their frugality and start using more energy.

3.5.4 Rebound estimates

The above exposition naturally raises the question of the numerical size of different rebound effects. The literature does not provide a unanimous answer though: estimates of the amount of rebound energy vary widely between different studies (Borenstein, 2013). Most studies do use the same measure however, defined in terms of the lost part of energy conservation initiated by the energy efficiency improvement:

\[
RE = \frac{\Delta P - \Delta A}{\Delta P} \cdot 100\%,
\]

where \(\Delta P\) refers to the projected savings and \(\Delta A\) to the savings actually achieved. A rebound of 10% thus means 10% of the theoretical savings from the improvement are offset by an increase

\(^{(15)}\) Some authors refer to these as “negative rebound effects” which is logical in view of the formula that is often used to calculate rebound energy (see Section 3.5.4), but may be somewhat confusing. Here the term ‘beneficial effects’ will be used, as is done in Santarius and Soland (2016).
in consumption. A rebound effect of more than 100% means the energy efficiency improvement has caused a net increase in energy consumption (often referred to as backfire, e.g. Saunders (2000)), which in the above formula occurs when $\Delta_A$ turns out negative. The formula can also be used to measure beneficial effects: whenever actual savings exceed the anticipated ones, the above formula will return a negative percentage.\(^{(16)}\)

Even with a seemingly simple equation as the one above, it is inherently complex to give a quantitative estimate of the size of different types of rebound effects for a number of reasons. First of all, there is not always consensus about the potential energy savings $\Delta_P$ from a certain technological advance. When a diesel car is replaced by a hybrid car for instance, savings depend on multiple factors such as the average driving distance and speed. As such, one can only estimate the size of the rebound effect, given that the original estimation of energy savings was accurate. More specifically, if actual savings deviate from the expected ones this may indeed be due to rebound or beneficial effects, but it can just as well be that the projected savings were off.

Another issue with giving quantitative rebound estimates is that identifying the exact cause of the phenomenon requires a qualitative assessment. Even though it may be relevant to identify different types of mechanisms, it is hard to separate different effects in terms of their impact, because one cannot derive an individual’s motivation for making certain energy-related decisions without asking them personally. One may hypothesize disappointing energy savings are caused by a psychological rebound effect (for instance because of a lack of a price incentive), yet without any additional context it cannot be determined whether consumers used more energy than expected because of a drop in feelings of responsibility or because of a re-evaluation of consequences.

Additionally, the exact size of effects depends on a large number socio-economic factors. Rebound effects tend to be larger in low income groups for example (Maxwell et al., 2011). Also, the size of effects can differ between different types of energy, e.g. Chitnis, Sorrell, Druckman, Firth, and Jackson (2014) found that the take-back is relatively small for measures affecting domestic energy use compared to measures affecting fuel use or food waste. Additionally, a EU-wide study of micro-economic rebound effects also found that the effect differs greatly per country (Freire-González, 2017).

Furthermore, there are different views on interdependencies between different types of rebound effects. Some studies have shown low direct rebound effects initiated large indirect effects (Freire-González, 2017; Chitnis et al., 2014), while others claim a mitigating influence (Gillingham, Kotchen, Rapson, & Wagner, 2013; Santarius, 2016). This makes it hard to compare findings from different studies since some focus on direct effects only (e.g. Sorrell et al., 2009), others include also indirect effects, but exclude economy-wide effects (e.g. Freire-González, 2017; Borenstein, 2013), while yet others study country-specific macro-economic rebounds (e.g. Shao, Huang, & Yang, 2014; Broberg, Berg, & Samakovlis, 2015) or global effects (e.g. Wei & Liu, 2017). In view of the above it is only natural different studies on the rebound effect come up with diverse outcomes in terms of empirical estimates. In the setting of household energy consumption in the Netherlands, Berkhout, Muskens, and Veltuijsen (2000) reported direct effects of about 15%, whereas Freire-González (2017) (who targeted both direct and indirect effects)

\(^{(16)}\)Which is why the term “negative rebound effect” is used to describe these phenomena in some texts.
found effects as large as 53%. This underlines the viewpoint of e.g. Chitnis et al. (2014); Maxwell et al. (2011) that indirect effects are not to be neglected. A longer list of estimations collected from different studies conducted in different countries can be found in (Freire-González, 2017).

3.5.5 Avoiding rebound

Research on mitigation of rebound effects is typically conducted from a public policy perspective (e.g. Van den Bergh, 2011; Vivanco, Kemp, & van der Voet, 2016; Chan & Gillingham, 2014), making it difficult to extract measures that are implementable by a commercial organization such as Huismerk Energie. Van den Bergh (2011) gives five policy directions for rebound mitigation, most of which involve governmental instruments like taxes and subsidies. An energy supplier could mimic these instruments by implementing a pricing system which rewards consumers who behave economically, but such practice carries financial risks. Other than that, the only remedy in Van den Bergh’s taxonomy that is applicable for a commercial organization is that of information provision: adding information about rebound effects to the feedback sent to consumers to make them aware of the existence of such effects and stimulate voluntary action to avoid them. There may also be potential for clever framing of feedback messages to counteract, but according to Karlin, Sanguinetti, et al. (2015), such directions need yet to be examined.

Opposed to mitigating rebound effects, one may also attempt to stimulate beneficial effects and there seem to be more useful pointers in literature that are executable by a commercial organization. In a large survey among German households Steinhorst, Klöckner, and Matthies (2015) found that a large emphasis on the financial benefits of saving electricity nullified the potential for environmentally friendly behaviour in other aspects, whereas a focus on the environmental benefits from electricity savings did ignite significant spillover effects. In a resource-based study, Margetts and Kashima (2017) found spillover from one behaviour to another is more likely to occur when the resources (time, money, effort) required to perform both behaviours are similar, because it is then more likely people perceive these behaviours as similar. Thus it may be fruitful to think about grouping different feedback aspects that draw upon the same pool of resources.
4 Methodology

The bulk of this thesis consists of two parts spread over chapters 5 and 6. The main aim are the smart meter data applications discussed in Chapter 6. These have come about based on the assessment of existing energy management systems (Section 3.4.3), results and recommendations from smart meter feedback studies (Section 3.4), wishes from employees of Huismerk Energie and De Groene Stroomfabriek (personal communication, September 2017), and own creativity.

In relation to the model displayed in Figure 6, the focus of this study is mostly ‘left-oriented’: the majority of applications discussed in this thesis are designed to inform consumers on specific aspects of their consumption (thus raising awareness), few are meant to directly induce changes (e.g. the refrigerator payback time and the insulation quality assessments from Sections 6.2 and 5.3.4), while little attention will be payed to long-term effects.

The applications from Chapter 6 transform smart metering data into useful and personalized energy statistics that are meant to increase energy awareness among consumers and trigger consumption reductions. Yet, raw smart meter data may in many cases be ‘contaminated’ with additional information, leading to algorithms failing to fulfill their primary objectives. Three different directions in which this may occur are identified in Chapter 5. The first of these concerns the presence of measurement noise and so-called ‘transient spikes’ in P1 data. These appear at rather random moments and may offset the mechanics of the data applications. The median filtering technique described in Section 5.1 is expected to be a simple and adequate tool for removing such noise without causing serious damage to the signal itself.

As for the other two directions, which concern (a) private electricity generation by consumers and (b) external weather effects, it may be less obvious as to why these form a problem to begin with. In many cases, the applications will actually work fine without taking into consideration any of the external effects discussed in Chapter 5, yet, higher levels of awareness and greater energy savings may be achieved if one does. This will be explained for both types, illustrated by means of a practical example.

(a) Starting off with the first type, i.e. the situation where a consumer produces its own solar energy. As will be explained in Section 5.2, smart meter data is no longer a measure of total consumption for such households, but rather a measure of the net flow of electricity, i.e. the difference between consumption and generation. As a result, the smart meter output of a partly self-sufficient consumer is no longer a reflection of the inhabitants activities alone: an additional factor, influenced primarily by external factors such as the positioning of the sun, is mixed into the signal. It will be hard to draw any conclusions about the consumption aspects when using this data, especially since solar generation is variable in time.

As for a practical example of this, consider the historic feedback application, which is discussed in Section 6.5.2 and allows consumers to compare different periods of their own data. This application can be run for a consumer with a private solar installation, in which case net figures will be compared. However, if there is a difference in the net figures between two subsequent days, one cannot determine whether this originates from a change in consumption, generation or both. Because both consumption and generation
are compared simultaneously, very little can be concluded of either aspect individually. In Section 5.2, methods will be discussed to estimate the amount of generated solar energy by a set of panels. Combined with the net figures measured by the smart meter, this enables one to retrieve total consumption.

(b) In Section 5.3 the influence of weather effects (and more generally, of ‘exogenous factors’) on consumption is discussed. If one is able to isolate these from a smart meter signal, one may enlarge the aspects of consumption that are determined by the behaviour of inhabitants. To illustrate this, consider the change detection feature, which is discussed in Section 6.4. Leaving aside all of the technical details, this solution is designed to take notice of any sudden increases in electricity usage that may lead to higher consumption over time. When applied to the raw smart meter signal of a typical Dutch household, the technique will start detecting changes that occur every year, due to weather influences and changes in solar influx. Also, it may fail to detect any changes during the part of the year where consumption is low, because its reference point are higher standards, set during the opposite season. If however the change detection feature is applied to a dataset that has been subdued to a seasonal correction (which is one of the topics of Chapter 5), then it is no longer triggered by external seasonal differences and has a better chance at finding anomalies in consumption that originate from an household-specific, internal development.

The above types of externalities are discussed in Chapter 5, including a number of generic solutions that can be applied to boost the end applications. As of yet, these solutions may fail to achieve a desirable level of accuracy though, and implementing them introduces a level of uncertainty into the value chain. As such, each of the applications from Chapter 6 comes with a discussion of its vulnerability to problems caused by private electricity generation and weather effects. In some cases, the effects are small enough to ignore, whereas others may turn up adverse results when the application is run on raw smart meter data. For some applications there are additional workarounds through which the generic solutions can be avoided entirely. For example, if for a household equipped with solar panels one wants to retrieve information about the lower levels of consumption (constituted by appliances in standby- or low-power modes), one can try to do so during the night, when solar panels do not generate any electricity. To the author’s surprise, PV estimation and weather normalization of consumption data is only rarely discussed in literature on energy feedback (Froehlich (2009) and (Brandon & Lewis, 1999) mention weather normalization in the context of historic feedback).

As already touched upon in Section 3.1, it is anticipated the Dutch energy market will undergo a number of possibly profound transformations in the not-so-distant future. Though these topics are not within the scope of this text, a number of the projected changes may affect the functioning of one or more of the techniques discussed in this text and will therefore be briefly discussed in this context in Section 5.3.4. The digression has been kept concise though, as it is a bit uncertain how exactly the market will develop.

**Solution discussion scheme**

The discussion of each of the technical solutions from Chapters 5 and 6 will roughly follow the same five-step framework shown in Figure 10. The steps include:

1. **Introduction:** A brief opening, in which the ultimate aim of the technique is stated and some motivation is given as to its relevance.
2. **Setting**: A description of the assumed setting and input data. For some techniques a framework is delineated more specifically through a set of assumptions.

3. **Method**: An exposition of the proposed technique used to meet the goal. Depending on whether the solution is already described in an other text, the degree of detail varies greatly: some techniques are mentioned only briefly, whereas others feature a step-by-step outline.

4. **Technical discussion**: An assortment of various aspects of the expected technical performance of the solution in practical situations. May or may not include a discussion of any
   - adjustable parameters,
   - assumptions postulated upfront,
   - data correction methods required (see Chapter 5),
   - test results,
   - expected performance and possible improvements,
   - applications.

5. **Energy conservation**: A discussion of potential energy savings associated with the application, based on evidence from literature. In some cases, this will be done for several different approaches at once. For example, Section 6.4 features a number of different change detection techniques. The discussion of energy savings associated with these techniques will be presented on change detection in general and is therefore presented after all of the different technical aspects have been assessed.

![Figure 10: Structure of the description of each of the proposed solutions.](image)

Although the above framework will form the main line in the discussion of each of the techniques, the prominence of each individual step differs between individual solutions. When the discussion of a technique is lengthy, relevant stages are marked by a subheading. If however a solution does not require a comprehensive treatment (for instance because it is very simple or already well-documented elsewhere), the steps will be run through more quickly and a clear partitioning using subheadings is tacitly left out to avoid unnecessary disruption of the narrative.

**Test results**

Most of the self-conceived solutions have been tested on consumer data, the results of which are included in their respective technical discussions. The amount of usable test data has been limited though and as such, the test results do not pass as a proper performance evaluations of the applications. Still, the test results give a good first impression of each of the techniques.
in practical situations and also help identify potential issues that can be investigated in a more expansive series of follow-up tests. None of the applications have been implemented by Huismerk Energie so far, so there are not test results on any consumer reactions as of yet.

**Averaging**
Throughout technical Chapters 5 and 6, there will be a lot of situations where an average needs to be taken over a set of data. There are various different ways to take an average, of which the arithmetic mean and median are probably most commonly used. In this text, the median will be the standard tool for averaging, because it is unaffected by outliers (opposed to the ordinary mean). If one would like to give outlying values a small vote as well, taking a weighted average may be a solution. To give the reader an idea for how this works, a weighted averaging approach has been implemented in Section 6.4.3. If preferred, one may replace the median by a similar approach in other sections as well.

**Energy savings**
In Section 3.4.1, an adaptation of the Transtheoretical Model of Change has been introduced. As hypothesized, consumers go through three stages when adopting a new, more energy efficient habit, and each stage needs to be completed in order to proceed to the next. As mentioned, the energy feedback outputs presented in this text are meant to conduct the first two stages of this process. Awareness however, is somewhat difficult to measure as it requires a qualitative assessment. When it comes to the effect of individual feedback designs, there are little such qualitative results available in literature. There are, however, some quantitative results on the energy savings associated with a number of different types of feedback. Therefore, this text will focus on the effects of different types of energy feedback in terms of energy savings reported in these studies. Given the hypothesis that each stage of the Transtheoretical Model of Change needs to be completed in order to move on to the next one, it will be assumed awareness has been raised sufficiently by a certain feedback type whenever significant savings have been reported.
5 Signal restoration

Energy consumption is influenced by numerous factors. Several of the techniques discussed in this thesis focus on the influence of household-specific factors on total consumption. This necessitates ways to distinguish the effect of individual influences from the rest. A conceptual model will be drawn up in which different factors which influence energy consumption are assessed. Three externalities will be discussed in more detail, as well as ways to isolate them from the data. First of all, a filtering technique will be discussed for smoothing noisy P1 signals. Secondly, methods for retrieving at-home solar generation will be treated, to be able to separate consumption and generation from a smart meter signal. Finally, weather influences and seasonal patterns are examined and techniques to compensate for such effects will be treated.

5.1 Transient spike & noise reduction

Certain electrical appliances\(^{(17)}\) draw a brief but intense \textit{inrush current} upon first starting, which is visible in power data as a short-lived \textit{transient current spike} (Kryter & Haynes, 1989). Such spikes may reach up to 6 kW (Barsim, Streubel, & Yang, 2012), but typically last only 5-500 µs (Key, Mansoor, & Martzloff, 1996) and therefore have a negligible effect on consumption. Because of their short duration, spikes can only be detected reliably by energy meters which sample at a frequency in the range of kHz or higher (Hart, 1992).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{load_signal.png}
\caption{Load signal from the refrigerator from House 1 of the REDD database (Kolter & Johnson, 2011).}
\end{figure}

\(^{(17)}\)Especially the ones with a built-in electric motor, such as a washing machine or vacuum cleaner.
Transient spikes are occasionally captured by smart meters (e.g. Figure 11), but their 1 Hz sampling rate is too coarse to reliably pinpoint them in general, which makes spikes unpredictable disturbances to the signal. Apart from transient spikes, metering data can contain *measurement noise*, as can be seen in Figure 11. The random appearance of noise and large transient spikes in data reduces the performance of some of the techniques discussed in this text. It is therefore desirable to create a processed copy of each P1 dataset in which imperfections have been smoothened.

A simple but effective means of noise reduction widely used in image processing is a *median filter*. A standard 1D median filter removes outliers in a one-dimensional dataset by replacing each entry by the median of its neighbouring data points (J.-H. Wang & Lin, 1997). Explicitly, if \( \mathbf{P} = (P_1, \ldots, P_N) \) is a sequence of P1 power measurements at subsequent time stamps \( t_1, \ldots, t_N \), then (starting at \( P_1 \)) each \( P_i \) is replaced with

\[
\text{median}(P_{i-R}, \ldots, P_{i-1}, P_{i+1}, \ldots, P_{i+R}).
\]

The range parameter \( R \) determines the *window size* of the filter. A larger windows allows for

\[\text{Figure 12: Two copies of the refrigerator sample from Figure 11 processed with median filters with different window ranges: } R = 2 \text{ (top) and } R = 10 \text{ (bottom).}\]
greater noise reduction, but also creates a more smeared picture (see e.g. Ko & Lee, 1991). Figure 12 above shows two median-filter processed copies of the refrigerator signal from Figure 11 with different window settings.

Both median filters succeed at removing the voltage spike, but the median filter with a range parameter of $R = 10$ performs a lot more better at noise reduction than the $R = 2$ filter. However, some information has been lost as well: the smaller “transient bumps” arising when the refrigerator switches on, are captured well at a $1$ Hz frequency (top graph) and can be useful in energy disaggregation (see Section 6.2), but did not survive a median filter with an $R = 10$ window range (bottom graph). The right value for $R$ thus is method-dependent and can be difficult to determine. In some cases, multiple, differently filtered copies of the data can be used for different applications. Alternatively, $R$ could be set adaptively, see e.g. Juneja and Mohana (2009); J.-H. Wang and Lin (1997).

5.2 PV generation

Private energy production from solar photovoltaic (PV) installations has grown explosively over the past few years in the Netherlands (Gerdes et al., 2017; H. G. J. Kamp, 2017a). Although smart meters are designed to capture all the necessary data to comply with current legislation, they do not measure the amount of electricity actually generated. This undermines the functioning of many of the techniques discussed later. The solutions discussed in this section are meant to restore, rather than utilize, smart meter signals. Nevertheless, some possible applications will be discussed which add relevance to PV yield retrieval efforts from an energy awareness perspective.

5.2.1 Smart meter limitations

When a household invests in a set of solar panels, they become partly self-sufficient in their electricity supply. Under the current Dutch net metering legislation, such *prosumers* (consumers who partly produce their own energy) enjoy fiscal benefits: only the net flow of electricity (i.e. the difference between consumption and PV generation) is taxed. Moreover, electricity may be fed back into the grid whenever the generated electricity is not entirely consumed in the residence. This surplus is deducted from overall consumption on the electricity bill (H. G. J. Kamp, 2017b).

PV modules and all running electrical appliances in a house thus jointly determine a bidirectional net flow of electricity. Energy meters only measure these net flows. Traditional meters have a single counter which runs forwards or backwards depending on whether there is a surplus in consumption or production, while smart meters have separate counters for both situations: instants of overproduction are registered on a second counter to keep track of feed-in (Authority for Consumers and Markets, 2017). Many smart meters are designed to differentiate between peak and off-peak hours as well and therefore come with four different counters. Table 4 below shows the operating range of each of the counters.

Whenever a four-counter meter is used for a customer with a single-tariff contract, the measurements from $M_{III}$ and $M_{IV}$ are added to $M_I$ and $M_{II}$ respectively. This practice will be adopted in this text to keep things orderly, so from now on two-counter metering systems will be considered, one measuring net import and the other net export.
Table 4: Smart meter counters and what they record from consumption \((C)\) and generation \((G)\) under different circumstances.

<table>
<thead>
<tr>
<th>Counter</th>
<th>Records</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M^I)</td>
<td>(C - G)</td>
<td>(C &gt; G), peak hours</td>
</tr>
<tr>
<td>(M^{II})</td>
<td>(G - C)</td>
<td>(C &lt; G), peak hours</td>
</tr>
<tr>
<td>(M^{III})</td>
<td>(C - G)</td>
<td>(C &gt; G), off-peak hours</td>
</tr>
<tr>
<td>(M^{IV})</td>
<td>(G - C)</td>
<td>(C &lt; G), off-peak hours</td>
</tr>
</tbody>
</table>

However, (smart) energy meters generally do not register actual at-home generation separately (van Elburg, 2017). Hence, actual electricity consumption can also no longer be determined when the panels are operational, only the net amount of electricity supplied by the utility or fed back into the grid are known, see Figure 13 below.

Figure 13: Example depiction of consumption, generation and net flow as well as the measuring areas of counters \(M^I\) and \(M^{II}\).

This is problematic for many of the smart meter data-based solutions discussed later, both from a technical angle (as the consumption signal is distorted) and from an energy awareness viewpoint (because the consumption signal shows values which are lower than the actual consumption). However, if generation over a certain period can be retrieved somehow, consumption can be re-determined as well. Explicitly, using the notation from Chapter 2, if between two time stamps \(t_0\) and \(t_1\) an amount \(G\) of electricity has been generated, then gross consumption \(C\) during that time equals:

\[
C = (M^I_1 - M^I_0) - (M^{II}_1 - M^{II}_0) + G.
\]

It should be noted however, that consumption at any intermediate moment remains unknown. In other words, the sampling rate of the PV generation data determines how fine-grained the
restored information about consumption will be. Some of the solutions discussed in this study require a higher level of precision than others. A number of techniques for establishing generation will be treated in this section, all with their individual pros and cons. These methods only concern private generation from solar PV modules, other types of at-home generation will not be considered.

5.2.2 Intrusive methods for retrieving PV generation

For starters, PV data can be acquired via so-called *intrusive methods*, that is by directly measuring inside a residence. Two ways to set this up will be briefly touched upon.

**Manual entry**

Often, a display of some sort showing generation is integrated in the inverter that comes with the PV panels upon installation. A tool could be integrated in Huismerk Energie’s web portal, allowing customers to manually input generation values displayed on their inverter. However, such an unautomated method requires a lot of effort at the consumer end, is prone to human error and loses its effect completely once the manual input ceases. Also, irregular insertion of the data might cause confusion about whether consumption figures relate to updated data or not.

**Additional metering**

PV generation can also be retrieved by installing an additional meter which directly measures production at the source and sends this data to a server via an integrated transmitter. Such a submetering approach yields 100% accurate knowledge of generation at high frequencies if desired and therefore is the preferred method to support the other smart meter data solutions, especially those that rely on precise consumption data. It however requires additional hardware at the consumer’s residence (raising the financial burden for the customer in addition to the costs already incurred in the initial investment) and obliges Huismerk Energie to set up the necessary infrastructure for secure transmission and storage of the data. Also, it might lead to disagreements about who is responsible for maintenance and repair of the components.

This *submetering* approach is more attractive for business customers, to whom a bunch of metering components generally mean a minor investment compared to a large energy deal. According to Gerwin Hament, former director at De Groene Stroomfabriek, business customers value the ability to continuously monitor the functioning and yield of their PV installation (personal communication, 22nd November, 2017). De Groene Stroomfabriek could design a private lease option for a PV meter in which various DBFOM(18) responsibilities are recorded. In return, firms get a fully reliable experience, not only from their PV generation figures, but also with respect to the smart meter data solutions discussed in this text. For organizations, this is an important aspect of software products: a comparison of five software quality assessments widely used by organizations found ‘reliability’ was the only component all models had in common (Jamwal, 2010).

5.2.3 Modelling approaches

Recently, a novel feature has been implemented in the Huismerk Energie’s web portal which enables customers to indicate whether they possess a PV installation, and if so, the number

(18) DBFOM stands for Design Build Finance Operate and Maintenance.
of panels they installed. This information enables non-intrusive estimation of generation from private PV modules. Such endeavours are sometimes referred to as nowcasting and are a special case of the larger topic of photovoltaic power forecasting. A lot of research has been conducted on the latter topic, since reliable predictions of power generation from renewable assets are key in securing the practicability of future congestion management tasks.

An extensive review of various PV forecasting approaches is provided by Antonanzas et al. (2016), who distinguish two methodologies for such techniques: PV performance models and statistical models. Both will be elaborated on, but from a rather general perspective so that the advantages and drawbacks of both can be assessed in a nowcasting context. For an extensive overview of the specific techniques used, the reader is referred to the original text. Deeper knowledge about the mathematical aspects of each of the techniques can be found in e.g. Inman, Pedro, and Coimbra (2013).

**PV performance models**

PV performance models utilize a deterministic and intuitive method by following the natural path of energy from extraterrestrial sun ray to panel output. Using weather forecasts or satellite imagery, solar irradiance and other weather factors at the target location are predicted, which in turn are converted to expected power generation using a model of a PV module. An overview of the steps of a general approach is given in Figure 14 below. PV performance models in general achieve lower scores in terms of forecasting accuracy than statistical models, and these deviations largely originate from the weather prediction component (Dolara, Grimaccia, Leva, Mussetta, & Ogliari, 2015; Brabec et al., 2011). However, in a nowcasting setting, PV performance methods might be more viable, since actual weather data is more accurate than a forecast.

![Figure 14: Common steps in many PV performance models.](image)

As PV performance models attempt to simulate the many circumstances under which electricity is generated, it is fundamental to start the discussion by assessing the parameters that influence photovoltaic generation processes. The yield of a PV module depends on internal characteristics of the panels, such as PV array size and peak power, inverter efficiency and maximum power point tracking losses, and on external influences, most notably solar irradiation and cell temperature (Notton, Lazarov, & Stoyanov, 2010; Lyden, Haque, Gargoom, & Negnevitsky, 2012).

**Solar irradiance**

Global horizontal irradiance (GHI) is the total amount of shortwave radiation received from above by a surface horizontal to the ground, both directly and indirectly through diffusion (Velds & Hoeven, 1992). It is a measure commonly used in relation to describe the potential for photovoltaic installations in a certain area. Two types\(^{(19)}\) of influences will be distinguished in this text:

\(^{(19)}\)This distinction is adopted from Lyden et al. (2012), although originally no names were assigned to the
• **Macro-climatic factors** relate to large-scale processes and influences that determine the baseline GHI. These include latitude, season, time of day and cloud cover (Lyden et al., 2012).

• **Micro-climatic factors** are the parameters which govern transient radiation on site, e.g. temperature, humidity, albedo, panel orientation and tilt angle, wind speed, dust covering panels and shading from nearby objects (Lyden et al., 2012; Bergin, Ghoroi, Dixit, Schauer, & Shindell, 2017; Mani & Pillai, 2010).

Intuitively, the macro-climatic factors can be understood as the influences that are equal on a regional scale, whereas micro-climatic factors offset this baseline irradiation received by a certain set of panels from neighbouring installations. The influence of many micro-climatic factors is virtually indeterminable unless a 3D model of the surroundings is created. However, one of the predominant determinants is the baseline GHI (Lyden et al., 2012), which can be estimated relatively accurately as long as the geographical location is known.

The amount of solar radiation varies greatly between regions. Table 5 shows the extremes of average sunshine hours and irradiation in the Netherlands over the period 1981-2010, retrieved from the Dutch meteorological institute KNMI (KNMI, 2011). In De Kooij, Noord-Holland 15 per cent more hours of sunshine are registered on average than in Deelen, Gelderland. Total received irradiation not only depends on sunshine hours, but also on latitude. In the end, Vlissingen, Zeeland is the station receiving the most radiation: about 10 per cent more than the station in Eelde, Drenthe. The difference in irradiation thus is smaller than the difference in sunshine hours, mainly because clouds only partly block incoming radiation (Velds & Hoeven, 1992).

<table>
<thead>
<tr>
<th>Location</th>
<th>Hours of sunshine (h)</th>
<th>Solar irradiation (J/cm(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deelen</td>
<td>1522.2</td>
<td>351,713</td>
</tr>
<tr>
<td>Eelde</td>
<td>1550.0</td>
<td>351,086</td>
</tr>
<tr>
<td>De Kooij</td>
<td>1751.4</td>
<td>383,260</td>
</tr>
<tr>
<td>Vlissingen</td>
<td>1733.1</td>
<td>386,126</td>
</tr>
</tbody>
</table>

Table 5: Multi-year (1981-2010) average sunshine data from four KNMI weather stations in the Netherlands.

Apart from annual differences, sunshine hours can also show large and unpredictable regional disparities throughout the day. This is captured in Figure 15, which shows a map of registered sunshine hours on 4th July 2017 for a number of KNMI stations (KNMI, 2017).

The terminology used here agrees with the classification used by winemakers for categorizing scale differences of climate and weather influences on their plantations (Robinson & Harding, 2015).

\(^{(20)}\)Albedo refers to the reflectivity of the surrounding area.
Therefore, for PV performance models to be reasonably accurate, the location of the panels should be known to some degree. Knowledge of the full address is expected to yield the best results, but if customers are reluctant to authorize the use of this information, their city or municipality already yields a fairly high accuracy (Saint-Drenan, Bofinger, Ernst, Landgraf, & Rohrig, 2011) and could be an acceptable alternative.

In solar forecasting studies, baseline irradiance is usually modelled using (predictions of) the macro-climatic parameters above by using weather models or satellite imagery. However for nowcasting applications, it is more cost-effective in terms of computational efforts to interpolate irradiance at a location based on data from the nearest weather stations. Saint-Drenan et al.
(2011) have proposed such a technique using inverse distance weighting interpolation, which works as follows. If \( x \) is the customer’s location, let \( s_1, \ldots, s_N \) be the KNMI weather stations within a certain radius. For each weather station \( s_i \), let \( d_i \) be the distance between \( x \) and the station and let \( y_i \) be the solar irradiation measured at that station during a certain time frame. The situation has been sketched for three stations in Figure 16 below.

![Figure 16: Triangulation method for estimating irradiation \( \bar{y} \) at \( x \) based on irradiation figures \( y_1, y_2 \) and \( y_3 \) from nearest stations \( s_1, s_2 \) and \( s_3 \).](image)

To estimate the irradiance \( \bar{y} \) at \( x \) a weighted average of \( y_1, \ldots, y_N \) is taken, i.e. a weight \( w_i \) is assigned to every weather station \( s_i \) and \( \bar{y} \) is set as

\[
\bar{y} = \frac{\sum_{i=1}^{N} w_i y_i}{\sum_{i=1}^{N} w_i}.
\]

Inverse distance weighting means setting \( w_i \) as a negative power of the distance between \( s_i \) and \( x \), i.e. picking \( w_i = d_i^{-p} \), where \( p \) is a positive number chosen heuristically. The philosophy behind this technique is that \( \bar{y} \) is estimated as an average which is biased towards closer stations. Higher values of \( p \) create a stronger bias. To increase the accuracy at locations alongside the border, measurements from German and Belgian stations could be added. For remote areas, the method could additionally be reinforced by using data from satellite models or by including data from trustworthy amateur weather stations.

**Modelling micro-climatic factors**

Even when the location is known, the baseline GHI can still deviate significantly from the amount of radiation captured by the PV installation due to micro-climatic factors. Some of these can to some extent be implemented in the model. Local weather variations are partially induced by differences in soil type and the level of urbanization (Mahfouf, Richard, & Mascart, 1987; Tapper, 1990; D. Lee, 1979), which can be embedded in the model. There might also be opportunities for dust modelling. Dust slowly accumulates under dry-weather conditions with high levels of particulates, and partially washes off during rainfall (Kimber, Mitchell, Nogradi, & Wenger, 2006). Using air quality and precipitation measurements, such oscillations can be built into the general model to improve accuracy. Similarly, snow covering the panels leads to large yield losses in winter (but small annual losses) (Thevenard & Pelland, 2013). Still, snow cover could be modelled based on weather data. Dust and snow models should be manually overruleable in case the owner cleans the panels. This could be realized by implementing a button into the portal, which can also act as a reminder to point customers in the right direction.
PHOTOVOLTAIC CELL MODELLING

Once the incoming radiation has been estimated, a panel model is often used to estimate the amount of electricity generated under the predicted conditions (e.g. Farivar & Asaei, 2011; Villalva, Gazoli, & Ruppert Filho, 2009; A. Chatterjee, Keyhani, & Kapoor, 2011). However, these models all require intricate knowledge about the specifics of a panel. Villalva et al. (2009) name cell temperature, series and shunt resistances, the leakage current of the diodes and the number of individual cells connected in series and in parallel as such. Additionally, the efficiency of the inverter needs to be taken into account. These parameters can only be known if the customer provides them, or if there are precise measurements of the amount of electricity generated under different levels of irradiance. Both rely on the willingness of customers to invest time in improving the model.

Instead of modelling each panel separately, it might be more practical to use a standard model, which converts irradiance to panel output based on an average panel and inverter. In view of the fact that solar energy technology has seen large performance improvements over the past decades (Green, Emery, Hishikawa, Warta, & Dunlop, 2015), such an estimation could be improved if the manufacturing date is known. This also enables the inclusion of an age correction in the model: the yield from a PV module is estimated to fall by about 0.5% each year (Jordan & Kurtz, 2013).

CORRECTIONS AND CALIBRATION

Some micro-climatic parameters cannot be modelled sufficiently accurate, or not at all (Thevenard & Pelland, 2013). The lack of such knowledge and the implementation of the rather coarse panel conversion technique can be partly compensated for through calibration. This is done by comparing the yield estimated by the model $G_{\text{est}}$ to an actual measurement $G_{\text{act}}$ supplied by the PV owner. From this, the deviation of the estimation can be determined, i.e.

$$d = \frac{G_{\text{act}}}{G_{\text{est}}}.$$  

This number $d$ can then be used to modify all previous and subsequent estimations by multiplying them with $d$. This can be applied both globally (i.e. adjusting all estimations) and on the level of the individual customer. Global scaling can be applied in the testing phase of the technique, to neutralize structural under- or overestimating. Individual calibration can be used to compensate for missing knowledge about micro-climatic influences and panel specs. If enough customers calibrate the model, statistical analysis can be performed on the total collection of (anonymized) calibration numbers to assess whether there are regional or seasonal differences in deviation. If this is the case, seasonal differences can be accounted for by varying $d$ over the year and regional differences can be corrected for by adjusting the yield estimation of all customers in the same municipality using the inverse distance weighting technique discussed earlier.

However, the calibration procedure can make things worse when faulty measurements are inserted or when the calibration is performed in circumstances where the estimation deviates more than on average. The deviation of the method is expected to vary from day to day (whimsical weather conditions for instance are likely to cause larger divergence than stable conditions) and between seasons (lower objects might obstruct direct sunlight in winter but not in summer for instance). As about 75% of solar energy is generated between March and September (Velds
& Hoeven, 1992), it is best to either calibrate during this period or to discriminate between seasons. To avoid worsening the situation too much, a reduced correction could be applied when the customer calibrates for the first time, gradually building towards a full scaling after a few entries.

**Statistical models**
The second genre of forecasting techniques applies methods from statistics or artificial intelligence (mostly artificial neural networks (Antonanzas et al., 2016), which are discussed in Section 6.2.4) to predict generation based on historical data (Antonanzas et al., 2016). This eliminates the need for real-time meteorological data and also works its way around the indeterminable nature of some micro-climatic parameters. Although such models have proven to be superior to PV performance models in terms of forecasting accuracy (Graditi, Ferlito, & Adinolfi, 2016), their drawback lies in the fact that they require a large set of historical meteorological and power measurements acquired over a longer period. Each PV installation needs its own training data, and the quality of this data largely determines the model’s precision (Antonanzas et al., 2016). This decreases the usability of such modelling approaches for instant PV estimations in residential cases, because each individual PV module needs to be studied, thus requiring the deployment of additional metering devices after all.

### 5.2.4 Applications
Apart from restoring the consumption signal, knowledge about PV generation can be utilized for different aims. A few suggestions will be discussed briefly, for which it is assumed a solar energy retrieving method has been chosen which achieves a reasonable accuracy.

As a simple application, PV generation can be shown to the customer in relation to their consumption. Instead of showing net values only, generation and consumption can be visualized side-by-side to show what part of consumption is covered by the PV installation. Using the conversion techniques which will be discussed in Section 6.1.1, abated electricity expenses can be estimated. If generation over some period is known, future generation can be roughly estimated through extrapolation. By comparing the total hours of sunshine for a period to the multi-year average, past generation can be scaled towards normal conditions. The trend found can be extrapolated towards future generation based on a variety of scenarios, i.e. extrapolating with different levels of sunniness, panel degradation and dust accumulation rates, etc. For each of these scenarios, and estimate on the payback time for the panels can be given, although this largely depends on developments in net metering legislation and electricity prices.

### 5.3 Energy data normalization
Energy usage is subject to a large number of influences. Some of these are native to a specific household, whereas others affect larger groups of consumers. When assessing the behavioural aspects of energy consumption of one specific household, it is desirable to have a method for separating the parts of energy consumption that are determined by actions of the inhabitants from those shaped by externalities. In this section, a model will be presented to aid in such endeavours. Also, an attempt will be made at isolating the effect of one specific class of external influences, namely that of weather conditions and seasonal differences. Normalization techniques will be discussed to remove these effects from a generic smart meter signal in order to enlarge the sought-after consumer-driven elements.
5.3.1 Influences on consumption: a conceptual model

Energy consumption differs between individual households and firms and is generally not uniformly distributed in time (see Figure 19, page 64). The amount of final energy used by a consumer is subject to a large number of influences. In this first section, a classification of energy-influencing factors is presented and interactions between factors are assessed. The discussion is partially based on an analysis and classification of factors that determine household energy choices in developing countries presented by Kowsari and Zerriffi (2011). Because a large number of these factors overlap with those in studies towards influences on domestic energy consumption in OECD countries such as the Netherlands, Kowsari and Zerriffi’s model will be a good starting point for the discussion here.

Kowsari and Zerriffi (2011) classify influences on household energy consumption as either exogenous (external) or endogenous (household-specific). Yet, they do not specify what marks the delineation between these two categories. The discussion here will therefore commence with an attempt to define a somewhat more rigid separation.

- **Exogenous factors**: those variables that affect the energy consumption of a significant part of a regional population.
- **Endogenous factors**: all other variables that significantly affect the energy consumption of one household.

The above distinction requires some clarification with respect to what is meant by “significant” and by “regional population”. By a region will be referred to a coherent area the size of at least a considerable town (with no upper limit, i.e. all of the Netherlands is allowed). If for example a small-sized thunderstorm puts a municipality in darkness, causing people to massively turn on their lighting earlier than they normally would have, the thunderstorm-induced darkness counts as an exogenous factor, whereas if a block of houses share the same building characteristics, these will count as individual endogenous factors that overlap by coincidence. A factor is marked as significant if its impact on the mean energy figures from the respective region triggers a statistically significant deviation from the expected course. This definition is somewhat unwieldy as it assumes a perfect understanding of the effect all other factors individually exert on regional energy consumption. Also, it requires an agreement on what counts as statistically significant.

Rather than fine-tuning the exact course of the border between endogenous and exogenous influences at a microscopic level, the discussion here will focus on those factors for which it is clear to what category they belong. Inevitably, this will leave a somewhat gray area of factors for which it is unclear whether they are exogenous or endogenous. For the sake of comprehension, a large number of examples will be given, without worrying about the statistical significance of effects or whether any presumed consumer responses are ‘natural’. This is intended to help the reader grasp a deeper understanding of the complex mixture of different factors that shape energy usage.

Within the two categories of factors, Kowsari and Zerriffi (2011) discern a number of subclasses. These can for a large part be adopted when categorizing influences on household energy use in the Netherlands. Two categories which made an appearance in the original model as exogenous...
factors have been removed due to not being relevant in the current setting. These comprised of energy supply factors and energy device characteristics.\(^{(21)}\)

Two new categories have been added. The first being an exogenous one consisting of certain widespread practices such as ‘sleeping at night’ and ‘9-to-5 work days’, which are responsible for daily patterns in household energy use depicted in Figure 21 (page 67). The second addition is an endogenous category entitled presence & activity, which consists of ‘on-off’ determinants, i.e. householders being at home or not, and ‘intensity measures’, related to the energy intensity of different activities performed by those who are at home. Figure 17 shows an overview of the categories. Each of the individual categories will be discussed in some more detail.

![Figure 17: Factors influencing household energy usage, based on the discussion in Kowsari and Zerriffi (2011).](image)

The eventual aim of this text is to achieve energy conservation through enhanced smart meter feedback. To this end, it is desirable to somehow isolate the endogenous elements of consumption from a generic smart meter signal. If successful, this will allow not only to investigate and target the purely behavioural aspects of energy consumption, but also the extent to which different consumers are able to conquer the effects of exogenous influences (e.g. differences in temperature-sensitivity of gas consumption, see Section 5.3.4). However, as the reader will learn shortly,\(^{(22)}\) there are a large number of endogenous and exogenous factors the effects of which are strongly intertwined, making it virtually impossible to precisely determine what part of consumption is determined by either one of them for an arbitrary consumption signal. Nevertheless, exogenous factors were defined as those factors that are notable from mean consumption figures, which provides some leads for roughly assessing their impact.

\(^{(21)}\)The first refers to “the affordability, availability, accessibility and reliability of energy supplies [which] are found to influence household fuel choice.” Though Kowsari and Zerriffi (2011) remark “the affordability of a fuel is determined by its price, which may also affect the quantity of fuel consumed”, this will be considered an economic household characteristic in this text. The availability, accessibility and reliability of energy supplies are hardly relevant in the context of this thesis: the focus of this text is on the clientele of an energy supplier, so the availability and accessibility are guaranteed, and with the average availability of electricity in the Netherlands over the past fourteen years at 99.995% (Gerdes et al., 2017), effects on consumption from blackouts will be neglected (although arguably, a major blackout may have a notable impact on domestic energy consumption). The second category of energy device characteristics refers to “high capital costs associated with using modern energy conversion technologies” as a barrier to switching to modern energy systems, which is also not relevant in the present context, although it may become so in the future if the transition towards renewable energy takes shape.

\(^{(22)}\)Impatient readers may turn to Figure 18 on page 62
Exogenous factors

Physical environment

A number of geographic factors impact energy usage and the way it is distributed throughout the year. In the Netherlands, both average gas and electricity demand are higher in winter than in summer (Hekkenberg, Benders, Moll, & Uiterkamp, 2009). In general, these differences are greater in residential and commercial markets than in industrial sectors (Elkhafif, 1996), possibly caused by the fact that heating-related energy has a larger share in total consumption of households than it in commercial energy use (Opstelten, Bakker, Kester, Borsboom, & Elkhuisen, 2007). Gas consumption is very temperature-sensitive (Mirasgedis et al., 2007; Sailor & Muñoz, 1997), while seasonal patterns in electricity usage result primarily from the fluctuating influx of solar radiation and varying economic activity throughout the year (Pardo, Meneu, & Valor, 2002). In fact, Tol, Dorland, Lise, Olsthoorn, and Spaninks (2000) found no significant relation between temperature and electricity consumption in the Netherlands whatsoever. In a more recent study by Hekkenberg et al. (2009) however, a small temperature-dependence of electricity demand in fall and in late spring was established, which was found to have become stronger over the past decades. In the end, whether there is a relation between outdoor temperatures and the electricity demand for a specific household or not, depends primarily on the appliances owned, see Section 5.3.4.

Economic climate

A second class of factors is that of the larger national economic circumstances. The relation between economic growth and energy consumption is a well-studied topic within the field of energy economics (Acaravci & Ozturk, 2010; Soytas & Sari, 2003). Economic growth goes hand in hand with increasing energy consumption figures, whereas the 2008 financial crisis lead to undersized consumption figures compared to earlier years (Peters et al., 2012).

Policies and regulations

A number of government policies and regulations may also affect household energy use patterns. Most of these influences encompass some price mechanism, e.g. the effects from peak and off-peak electricity tariffs, energy taxes and energy-related subsidies and regulations (such as the current ones for electric vehicles). Some policies may induce non-monetary effects, e.g. most shops not being allowed to be open on Christmas Day. More generally, some public holidays have notable impact on consumption patterns of both households and firms nationwide (NEDU, 2017a, 2017b).

Widespread practices

The final category of exogenous factors is that of widespread practices, which relate to a number of customs a majority adheres to. Some examples were already mentioned above, namely sleep and working routines. These are examples of nationally embedded practices. Effects may also be regional however. Consider for instance the yearly Carnival festivities which are celebrated especially exuberantly in the southern parts of the Netherlands and are likely to affect household consumption patterns in this region, as many people leave their homes to watch the local parade.

Endogenous factors

Economic household characteristics

The category of economic household characteristics is more or less a miniature version of the exogenous economic climate discussed above. It consists mainly of the factors that relate to the
financial situation of specific households, of which income is probably the one most encountered in literature. A large number of studies support a strong correlation between income and the amount of final energy consumed (e.g. C.-C. Lee, 2006; Elias & Victor, 2005; Fitzgerald, Barnes, & McGranahan, 1990). Energy usage is often thought to follow an inverse-U trajectory\(^{(23)}\) with increasing incomes, the underlying philosophy being that once basic energy demands of a consumer are met, a part of their income will flow towards more energy efficient assets, setting in motion a series of energy reductions (Acaravci & Ozturk, 2010). This hypothesis has been the subject of a great deal of criticism due to the limited amount of statistical support (e.g. Stern, 2004; Dasgupta, Laplante, Wang, & Wheeler, 2002).

The class of economic household characteristics is however not limited to financial mechanisms. An example of a non-monetary factor would be the employment status of residents. Households with one or more unemployed members on average have a flatter daily profile than those whose members are at work during the day (B. Anderson, Lin, Newing, Bahaj, & James, 2017). Naturally, employment status is likely to affect the income of the household as a whole as well, thus also incorporates an indirect, financial effect.

**Non-economic household characteristics**

There are also a number of non-economic household characteristics that have been found to influence energy usage. Examples are household size, gender, age and level of education of the residents (Wier, Lenzen, Munksgaard, & Smed, 2001; Kowsari & Zerriffi, 2011; Heltberg, Kojima, & Bacon, 2003), and a number of build characteristics of their accommodation such as dwelling size and insulation quality (Brounen, Kok, & Quigley, 2012).

**Cultural & behavioural characteristics**

Beside the materialistic and demographic factors mentioned above, a number of lifestyle adaptations, habits and preferences (Kowsari & Zerriffi, 2011) may also shape energy usage patterns of a household. Possible examples are one’s food tastes (i.e. a person who regularly engages in multihour slow-cooking efforts will have higher energy usage from cooking equipment than someone who dines outdoors most of the time) and preferred living room temperature. Also, aspects of one’s cultural background and beliefs may influence energy usage. A pro-environmentalist for instance is likely to make different decisions when it comes to appliance choices than someone to whom energy conservation has less priority.

**Presence & activity**

The final class of endogenous factors relates to whether and how many residents are at home or not and what they are doing when they are. The latter includes whether people are asleep or awake. This category can have a rather dramatic effect on the influence of other factors on energy consumption. For example, it may be –20°C outside, but if consumers are not at home at that moment, their energy consumption will probably not be affected by the wintery circumstances.

**INTERRELATIONS BETWEEN FACTORS**

In the end, it is always the consumer who determines the final energy consumption. The above two categories of factors affect all energy-related decisions by a consumer, rather than energy consumption itself. In this text, non-decisions due to inhabitants being asleep or away from

\(^{(23)}\)Known as the environmental Kuznets curve hypothesis.

60
home are also counted as decision-making. Also, the relation between decision-making and energy consumption is expected to be instantaneous and without any further influences from externalities.

As already remarked by Kowsari and Zerriffi (2011), there is a lot of interplay between individual factors and the way they affect energy consumption. Apart from the above distinction between endogenous and exogenous factors and the various subcategories, all influences will be subsumed to a parallel classification. Within this second scheme, direct factors are those variables that impact energy-related decision-making by a consumer straight away, whereas indirect factors are the ones that affect another factor or the impact of another factor. Indirect factors are further sub-classified as either primary factors, i.e. the ones that influence the impact of another factor on consumption, or secondary factors, which affect other factors themselves. For example, the insulation quality of a building affects the impact of outdoor temperatures on a consumer’s feeling of comfort and thus counts as a primary indirect factor. A consumer’s employment status partly determines a household’s total income and therefore is a secondary factor.

With respect to the categories of endogenous and exogenous influences, indirect factors can also be labelled as internal or external, depending on whether they influence a factor or effect from the same class (an endogenous factor influencing (the effect of) another endogenous factor for instance) or from the opposite class (e.g. an endogenous factor influencing (the effect of) an exogenous one). Applied to the examples given just now, the insulation quality would be an external factor and the employment status would be internal. One should note that the above classification is not disjoint: one factor can be of multiple types. Table 6 summarizes the different types of indirect effects.

<table>
<thead>
<tr>
<th>Indirect factors</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td>Influences the effect of a direct factor from the same class.</td>
<td>Influences the effect of a direct factor from the opposite class.</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td>Influences a direct factor from the same class.</td>
<td>Influences a direct factor from the opposite class.</td>
</tr>
</tbody>
</table>

Table 6: Alternative classification scheme of energy-influencing factors.

In the remainder of this section, the possibilities of various factor cross-classification type combinations will be assessed. Starting with the class of direct factors: there are both exogenous and endogenous factors that directly impact energy decisions. Outdoor temperature is an example of an exogenous factor and a consumer getting hungry is an endogenous one. As for indirect factors, there are eight possible combinations in total, two of which have already been covered by the insulation (endo → exo, primary) and employment (endo → endo, secondary) examples above. Examples of the remaining categories are given in Table 7 below. It is assumed individual consumers are unable to alter exogenous factors themselves (one cannot change the weather). In other words, the category (endo → exo, secondary) is presumed empty.
Example situation | Type of influence
---|---
Residents dining outdoors. | endo → endo primary
Employment status affecting income. | endo → endo secondary
Insulation quality mitigating influence of outdoor temperatures. | endo → exo primary
- | endo → exo secondary
Sunlight waking a consumer earlier than usual. | exo → endo primary
Government policy increasing customer spending power. | exo → endo secondary
Thunderstorm blocking sunlight. | exo → exo primary
Worsening economy leading to new policy. | exo → exo secondary

Table 7: Example situations for all possible types of indirect factors

The above interrelations are summarized in Figure 18.

![Figure 18: Self-made flowchart model for describing how energy consumption of a household is shaped by endogenous and exogenous factors.]

Over the course of the next few sections, this will be tried for exogenous factors that relate to the physical environment. The main reason for this is that these are largely predictable in the sense that seasonal fluctuations are an every-year phenomenon. As a result, the effect of annual temperature and influx oscillations on energy consumption is fairly well-understood. Therefore, this class of exogenous effects is expected to be the easiest one to handle with respect to normalization efforts. Analogous endeavours for the other external factors are left for further research.

The remainder of this chapter is organized as follows. In section 5.3.2, an energy dataset consisting of nation-wide averages is introduced which will be used for studying recurrent exogenous influences on domestic energy consumption (Section 5.3.3) and for naive normalization efforts (Section 5.3.4), which use the national averages to correct the data of individual customers. This method is then expanded into a more sophisticated version, and an additional application is described in the form of an insulation quality assessment.

5.3.2 Consumption profile fractions

Each year, the Dutch organization for energy data exchange (NEDU) provides standard annual consumption figures for all user profiles distinguished in Table 1 (page 14). These consumption figures model the distribution of energy consumption of profile members in time, based on past
consumption data, climate records and dates of public holidays (NEDU, 2017a, 2017b). Each profile consists of a sequence of profile fractions, which replicate the relative energy consumption of profile members for a series of time slots. The time slots jointly cover an entire year and their duration corresponds to the sampling rate of smart meter P4 data: fifteen minutes for electricity and an hour for gas (NEDU, 2017a, 2017b).

The electricity profile fractions are normalized, i.e. the annual sum of all fractions is 1. For gas, the fractions are temperature-dependent: a fraction becomes higher when the average temperature during the corresponding time slot is below the climate normal and vice versa. The formula for a single profile fraction uses a temperature-independent consumption figure $\text{TOP}_i$, a heating temperature $\text{TST}_i$ (which is a threshold value for the outdoor temperature $T_i$ below which consumers are assumed to turn on the heating), and a regression coefficient $\text{RER}_i$ which describes the relation between outside temperatures and heating-related gas consumption. The formula for each profile fraction is as follows:

$$
\text{PF}_i = \begin{cases} 
(T\text{ST}_i - T_i) \cdot \text{RER}_i + \text{TOP}_i & \text{if } T\text{ST}_i > T_i \\
\text{TOP}_i & \text{if } T\text{ST}_i \leq T_i
\end{cases}
$$

When multi-year mean temperatures are inserted, the annual sum of all gas profile fractions equals 1, otherwise it need not be.

The profile fractions can be used to assess the global distribution of annual electricity and gas consumption over different periods. For instance, to estimate the portion of annual electricity consumption taking place in March, sum all profile fractions for March and multiply by the annual consumption. For comparative endeavors though, it is sometimes more convenient to use the deviation from the annual average. To this end, the deviation factor $f_i$ of each profile fraction $\text{PF}_i$ will be defined as the quotient of the profile fraction for that sample $\text{PF}_i$ and the average taken over all profile fractions $\overline{\text{PF}}$:

$$
f_i = \frac{\text{PF}_i}{\overline{\text{PF}}} \quad (1)
$$

Deviation factors sway around 1: if $f_i > 1$, this means $\text{PF}_i$ is above the annual average, while an $f_i < 1$ corresponds to a sub-standard profile fraction. Profile fractions and deviation numbers can be downsampling to visualize coarser patterns. Ergo, $\text{PF}^{\text{day}}_i$, $\text{PF}^{\text{week}}_i$ and $\text{PF}^{\text{month}}_i$ will denote daily, weekly (24) and monthly profile fractions, which can be acquired by summing all profile fractions in the selected period. By dividing by the average periodical profile fraction in equation (1) above, the deviation numbers $f^{\text{day}}_i$, $f^{\text{week}}_i$ and $f^{\text{month}}_i$ can be computed for each period.

A high resolution annual lapse of daily deviation numbers for electricity and gas consumption for residential (E1A and G1A) and large industrial (E3C and G2B) profiles is shown in Figure (24). When weekly time scales are considered in this text, the first week will be initiated at 1st January. If this happens to be a Wednesday, all weeks will then run from Wednesday up to the next Tuesday. When doing this, week 53 will consist of only one or two days, which can be problematic for some applications (e.g. normalization using deviation numbers, see below). There are various ways to resolve this. One may for instance multiply $\text{PF}^{\text{week}}_{53}$ by 7 divided by the number of days in the last week to model a full week. Alternatively, the last week can be completed by adding a few days from the new year.
Figure 19: Normalized annual lapse of electricity (E1A, E3C) and gas (G1A, G2B) consumption profiles. Multi-year average temperatures were inserted as temperature coefficients to determine the gas profiles.
Average weekly profile fractions in summer and winter for electricity profiles E1A and E3C are shown in Figure 20.

5.3.3 Cycles

From the NEDU profiles depicted in Figures 19 and 20, three recurring cycles may be identified:

1. Daily cycles: patterns in daily energy consumption, which in the residential case shows a modest peak in the morning, a more pronounced peak in the evening and a clear dip during the night. The industrial profiles display a more simplistic course, with one larger bump during office hours on weekdays and less-defined patterns in the weekends.

2. Weekly cycles: the division of energy usage on a weekly basis, which shows differences between weekdays and weekends, especially in commercial and industrial sectors.

3. Yearly cycles: the global distribution of energy consumption throughout the year, shaped by seasonal differences in weather and solar irradiation.

Quite some information can be derived from Figures 19 and 20. First of all, the fact that (on average), the Netherlands is a winter-peaking market: both electricity and gas consumption are higher in winter than in summer. As noted already in Section 5.3.1, residential energy use shows greater seasonal fluctuations than industrial consumption. Also, the figures confirm the findings discussed in Section 5.3.1 that gas demand to a large extent follows outdoor temperatures, while electricity use seems to be independent from temperature fluctuations.

5.3.4 Normalization

Now that the generic patterns on different time scales have been identified, means to negate weather influences in consumption data can be assessed.

Method description

The method presented in this section uses a general model of annual energy usage (in this case,
the NEDU profiles) to correct an individual consumer’s energy data: a correction is applied to
a consumer’s data based on what is expected from the national average. Whenever a NEDU
profile fraction is high, the corresponding real data sample is lowered and if a profile fraction
is lower, the corresponding data sample is boosted. This is done using the deviation numbers
introduced earlier. The procedure is as follows: the \( i \)th normalized entry \( \hat{C}_i \) is obtained by
dividing the \( i \)th consumption measurement \( C_i \) by the deviation number \( f_i \) carrying the same
time stamp:

\[
\hat{C}_i = \frac{C_i}{f_i}.
\]  

(2)

For gas, actual temperatures\(^{(25)}\) should be used to determine the profile fractions and derived
deviation numbers. The above procedure can be executed over single (quarter-hourly or hourly)
consumption measurements, but also at coarser resolutions. To perform day-based normalization
for instance, \( f_i \) is replaced by \( f_{\text{day}}^i \) in equation (2) above, where the deviation number used
corresponds to the day on which \( C_i \) was recorded.\(^{(26)}\) Normalizing over larger time frames
creates a more subtle effect and can be used to bypass the cycles discussed in Section 5.3.3,
avoiding errors due to misalignment of patterns as will become apparent soon. In this text,
three time window possibilities for the normalization procedure will be discussed:

1. *Quarter-hourly (electricity) or hourly (gas) resolution.* The highest correction resolution,
which means that each consumption value is normalized by its a separate deviation number.
This means the scaling data will be very detailed, in the sense that it follows all of the
cycles discussed in Section 5.3.3. Hence, such a sample-based normalization procedure
will only be successful if the patterns in the actual dataset match those from the NEDU
profile used for the correction. If the patterns do not match, the normalization may lead
to divergence phenomena. This is illustrated in Figure 21 below.

2. *Daily resolution.* When downsampling to daily deviation numbers \( f_{\text{day}}^i \), the scaling data
no longer contains information about energy usage throughout the day, which means daily
patterns are no longer affected by the normalization procedure. Weekly patterns such
as the ratio between consumption on weekdays and weekend are still altered by day-
based corrections. Thus, problems may arise for households with unconventional weekly
schedules. For instance, if one family member works from home on Mondays, consumption
on Mondays will most likely be higher than anticipated and a day-based normalization
procedure may lead to divergence phenomena every Monday for this particular household.

3. *Weekly resolution.* When using weekly corrections \( f_{\text{week}}^i \), both daily and weekly cycles
are bypassed. For instance, the influence of a family member working from home on
Mondays is no longer a problem, as it occurs every week and therefore hardly affects the
proportion between different weekly consumption values. However, at such low resolutions,
the step size between subsequent deviation numbers is rather large, which may therefore
introduce unwanted step-like patterns in corrected data as well. These can be smoothed
by interpolating if desired.

\(^{(25)}\) Ideally, regional temperatures are inserted, based on the customer’s location.
\(^{(26)}\) In particular, all individual consumption values from one day are normalized by the same amount.
Which resolution to pick depends on a number of factors. If a household’s electricity consumption indeed solely depends on slow-changing factors such as solar influx (see Section 5.3.1), then a slower daily or weekly normalization resolution would be fitting: picking a higher resolution makes it likely one also compensates for non-seasonal influences. Gas consumption however is temperature-dependent, and since temperatures may vary greatly throughout the day, a finer resolutions is required in order to model these changes.

However, using a (quarter-)hourly resolutions for normalization may go wrong when a consumer’s timetable differs from that of the average consumer. To illustrate this, the different effects of the above normalization resolutions are shown in Figure 21 below, where the NEDU E1A electricity profile fractions from 1st March, 2017 are self-normalized at quarter-hourly and daily resolutions.\(^{(27)}\) The left-side graph shows the effects when the signals are perfectly aligned. This is the ideal setting for sample-based normalization, which is why the result of the quarter-hourly self-normalization process is a flat line hovering at a constant value equal to the annual average profile fraction. The day-based correction only lowered the entire signal by a small amount (as daily consumption on the first week of March is anticipated to be slightly above the annual average), leaving its lapse more or less intact.\(^{(28)}\)

![Figure 21: Day- and quarter-hour-based self-normalization of NEDU E1A profile fractions from 1st March, 2017 (left) and a time-shifted copy (right).](image)

The graph on the right shows what may happen when a signal is not in sync with the NEDU profile’s course. The original sample has been time-shifted by four hours. This causes the sample-based correction to create a large, unwanted peak during the night, while the day-based run was hardly affected. Although a complete four hour time-shift of a sample may not be

\(^{(27)}\)The results from a week-based correction were almost indistinguishable from the day-based run, hence the former has been omitted from the figure.

\(^{(28)}\)The result of the daily self-normalization is that the total daily amount after the correction equals the daily consumption average.
too realistic, a household’s energy demand always perfectly following the anticipated NEDU course is also unlikely. Even if local divergences similar to the one shown in Figure 21 only happen occasionally, the normalization procedure altogether may damage a signal and reduce the performance of the envisaged applications instead of improving it. Hence, from a theoretical point of view, day- and week-based corrections are recommended as the safer option.

**TECHNICAL DISCUSSION**

**Test results**

The deviation number normalization procedure was tested at all three resolutions on the electricity and gas data from three of Huismerk Energie’s customers.\(^{(29)}\) All tests for electricity were performed on the quarter-hourly recordings between 00:15 AM on 1st January, 2016 and 00:00 AM on 1st January, 2017, i.e. all samples from 2016.\(^{(30)}\) All tests for gas were performed on hourly recordings between 01:00 AM on 1st January, 2016 and 00:00 AM on 1st January, 2017. The NEDU electricity and gas profile fractions from 2016 were used for the corrections. For the gas normalizations, hourly temperature records from KNMI stations were inserted as temperature coefficients (obtained from KNMI, 2018). As the two data samples used were both from inhabitants of Nijmegen, Gelderland, the weighted average (with squared inverse distances as weights, as discussed in Section 5.2) of temperature recordings from the three nearest KNMI stations, Volkel, Noord-Brabant (23 km), Deelen, Gelderland (29 km) and Arcen, Limburg (45 km) was used.

Corrections were applied on a quarter-hourly/hourly, daily and weekly basis, after which the original sample as well as all corrected samples were downsampled (through summation) to a weekly resolution for comparison.\(^{(31)}\) The year was subdivided into seasons based on a whole-week classification shown in Table 8 below.

<table>
<thead>
<tr>
<th>Season</th>
<th>Week numbers</th>
<th>Starting date</th>
<th>End date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>1-9</td>
<td>Friday, 1st January</td>
<td>Thursday, 3rd March</td>
</tr>
<tr>
<td>Spring</td>
<td>10-22</td>
<td>Friday, 4th March</td>
<td>Thursday, 2nd June</td>
</tr>
<tr>
<td>Summer</td>
<td>23-35</td>
<td>Friday, 3rd June</td>
<td>Thursday, 1st September</td>
</tr>
<tr>
<td>Autumn</td>
<td>36-48</td>
<td>Friday, 2nd September</td>
<td>Thursday, 1st December</td>
</tr>
<tr>
<td>Winter</td>
<td>49-53</td>
<td>Friday, 2nd December</td>
<td>Saturday, 31th December</td>
</tr>
</tbody>
</table>

**Table 8:** Week-based seasonal classification for 2016.

Of each original and corrected sample, the mean value and standard deviation were computed, both annually and seasonally. The results are shown in Figures 22 (electricity) and 23 (gas). These statistics can be used as a measure of success for the corrections, the philosophy being that an efficacious correction homogenizes weekly consumption values throughout the year (i.e. it reduces the standard deviation), without affecting the annual mean too much (i.e. there is a balance between the total portions of the data that are increased and decreased).

\(^{(29)}\) The gas recordings of customer C were incomplete and have been omitted.

\(^{(30)}\) The time stamp of a sample represents the end point of the recording period, i.e. the sample ending at 00:15 AM on 1st January encompasses the electricity consumption between 00:00 AM and 00:15 AM on that day.

\(^{(31)}\) Since the 1st January, 2017 was on a Friday, all weeks considered here run from Friday to Thursday. The 53rd week consisted of only two days, which was compensated for by a multiplying total consumption from that week by 7/2.
Figure 22: Annual and seasonal mean values and standard deviations of untouched and normalized weekly electricity consumption figures from three P4 signals.
Mean weekly consumption (m$^3$)
Weekly consumption standard deviation (m$^3$)

50
25
20
10

year
winter
spring
summer
autumn
year
winter
spring
summer
autumn

Figure 23: Annual and seasonal mean values and standard deviations of untouched and normalized weekly gas consumption figures from two P4 signals.

It should be noted upfront that the set of testing material is too small to draw any broad universal conclusions, yet some observations can be made from the above figures. Starting with the electricity data, one can note there is no significant difference in any of the test cases between day- and week-based corrections in terms of performance, neither on an annual nor on a seasonal basis. Also, it can be seen all of the corrections had a detrimental effect on sample A. This is
likely to be caused by the fact the annual course of consumption in household A does not match the anticipated pattern from the NEDU profile. Indeed, the seasonal mean values from the original data of sample A in Figure 22 show a peak in autumn and a drop in spring, which signifies a phase-shift in comparison to the anticipated winter-high, summer-low lapse.

Besides, the coarser day- and week normalizations seem to attain a higher degree of stability regarding conserving the annual mean than their 15 minute counterpart, at least based on the three test cases. The all-year mean value has not been altered by any of the day- and week-based corrections, whereas the annual average of the quarter-hourly corrected copies of samples A and C is significantly higher than the mean values from the original samples. This implies the quarter-hourly normalization also increased total consumption in these cases, which is undesirable. The 15-minute normalization performed better on sample B, but worse for sample C than the coarser variants. However, while the day- and week-based corrections improved both B and C, the 15 minute run had a negative effect on sample C. A larger test run should determine which of the methods performs best in general.

As for the gas corrections, one may note that there is little spreading between the outcomes of different resolutions, although (contrary to what was expected beforehand) the daily and weekly normalizations yielded better performance than the hourly one. The daily and weekly normalization efforts performed especially well on sample B: the seasonal mean values are very close to one another and the spreading has been greatly reduced in all seasons except summer. Sample A has also been normalized fairly well, but the corrections should have been even weightier; there is still some of the original fluctuations left after the normalization.

**Expected performance & possible improvements**

One issue with seasonal normalization of energy data in general is that it is difficult to check whether a normalization has been successful, mostly because it is very hard to determine exactly what part of a consumption signal is season-dependent, without any additional knowledge about the target residence. The energy consumption of some processes, such as lighting, heating and air conditioning most likely shows a clear lapse throughout the year, whereas others are mainly determined by the behaviour of the consumer. A smart meter signal being the sum of all these processes, becomes a complex and typically volatile signal from which it may be difficult to determine the underlying seasonal component.

In the above, the mean and standard deviation were used to evaluate the outcome. When engaging in larger test runs, it may be preferable to use the median absolute deviation as a measure of variability. This is a median-version of the standard deviation, insusceptible to outliers (D. C. Howell, 2014). Additionally, the median could be used instead of the mean to make the entire assessment process more robust. Also, feedback mechanisms could be implemented to control the intensity of the normalization. For instance, one could cap the maximum amount that is adjusted for each individual sample.

A second problem with the above procedure is that the seasonal course depicted in Figure 19 need not be the same for everyone. In some Southern European countries, the annual distribution of electricity consumption is different from that in the Netherlands, with relatively high consumption in summer and a more significant degree of temperature-dependence (Mirasgedis

---

(32) In the words of Benders, Kok, Moll, Wiersma, and Noorman (2006): “the average household does not exist.”
et al., 2006). These observations have been attributed to a higher penetration and usage of air conditioning units in these countries (Bessec & Fouquau, 2008). Therefore, if a Dutch household has an air conditioner installed in their home, their summer electricity demand might show some similarities with that of a consumer in a Mediterranean country. Applying the generic normalization procedure described above to such households is likely to lead to counter-productive results regarding electricity use in summer.

A mismatch may also occur when a residence has a higher degree of electrification than most, for instance when it is equipped with an electric heat pump instead of a traditional gas-fired heating installation. For such households, electricity consumption will likely take over a large part of the role usually fulfilled by gas consumption. In particular, electricity consumption will become a much more temperature-dependent quantity. Rashly applying normalization based on the standard figures to the consumption data of such households is likely to lead to large divergence phenomena for both electricity and gas in winter. At the moment, 1.8% of the Dutch households use electricity for heating, however this number is likely to increase in the near future in view of the cabinet’s intentions to phase out natural gas as the standard heating resource for all new-builds (VVD, CDA, D66 & Christenunie, 2017).

Instead of normalizing all consumers’ energy data using the same generic model, it may be better to create customized profiles for groups of customers. One could look into clustering various electricity and gas signals based on their annual course. To make signals comparable, each entry of a consumer’s signal should be divided by their total yearly consumption (so that the individual consumption measurements of each consumer signal sum up to 1 on an annual basis). Alternatively, one could look into sine-regression, i.e. fitting a wave function onto a one-year collection of daily measurements. This should reduce the possibility of a misalignment when applying normalization in a later stage. However, a drawback is that it can only be applied to customers of whom over a year of data is available.

**Applications**

Apart from attempting to remove exogenous weather influences from a consumption signal as described above, one can also try to examine exactly what impact weather effects have on the consumption of a household. From this, one may derive some knowledge on for instance the insulation quality of a building. According to (Milieu Centraal, 2018b), three out of four Dutch households is not yet insulated properly, while well-insulated houses may use up to 80% less heating energy than poorly-isolated ones. For the technique discussed here to work it is assumed the dominant heating source of a household is known, that is, one has derived from a consumer’s electricity and gas data whether the household’s heating system runs on gas or electricity as described above. In the remainder of this section, the term ‘energy’ will refer to this dominant heating source.

One way to approach this is through sine regression which was already mentioned above. A wave function with a period of one year is expected to mimic the greater underlying seasonal changes quite well. The amplitude of the resulting sine wave carries information about the difference between a consumer’s summer and winter consumption. By normalizing each customer’s data (by dividing each measurement by the total annual amount), one can compare the sine waves and amplitudes of different customers. The larger the amplitude, the more energy is consumed.

---

(33) See Section 6.5.5 for a discussion on regression.
in winter. This may be a sign an overly large part of energy goes to heating, which could indicate the quality of the insulation is not too good. The above procedure can be executed for each of the subcategories distinguished in the normalization procedure above. One may for instance compare the fluctuations in energy use for all inhabitants of terraced houses. Consumers with very large seasonal fluctuations in their consumption may then be sent a message to point them at the fact that renewing their insulation may be profitable on the long run.

The sine wave procedure is expected to work best for consumers who heat their house using gas, although it can also be applied to customers with electrical heat pumps. Thing become tricky though when a customer’s energy consumption does not follow a simple wave pattern. For instance, if a household has both an electrical heat pump and an air conditioning unit, there may be two peaks, one in winter and one in summer. In this case it will be difficult to extract any information about the quality of the insulation.

5.4 Conclusions

Three directions in which smart meter data signals may need pre-processing have been examined. Median filtering is expected a strong tool for cleaning up messy signals, flexible enough to achieve the right amount of smoothing while maintaining the important characteristics of the signal. The seasonal correction and PV estimation methods on the other hand are prone to errors and may not meet the desired accuracy for their envisioned applications in some situations. The difficulty in both situations is that it is impossible to model the habitat of one specific consumer: one can only create a generic model and attempt to refine it here and there. A properly conducted series of test runs should determine whether the correction methods described here are adequately precise.

Each of the techniques discussed in Chapter 6 will come with a short discussion of the extent to which they require any of the above signal restoration. In general, applying any kind of modification to the signal introduces a degree of uncertainty in the chain from data to application. The PV estimation and seasonal correction being especially precarious, those types of signal restoration will be avoided as much as possible and alternative approaches will be suggested if possible. For some techniques, the corrections are inevitable, so some extra attention should be devoted to these methods in the testing phase.

Finally, the author acknowledges that a number of anticipated technological paradigm shifts may induce a complete overhaul of the current energy system and affect the functioning of a number of the techniques described in this thesis. One such is the anticipated transition towards smart grids, i.e. “automated, widely distributed energy delivery network[s], [...] characterized by a two-way flow of electricity and information [...] capable of monitoring everything from power plants to customer preferences to individual appliances” (U.S. Department of Energy, 2009). Although several pilots have already been completed and it may not be long until smart grids are widely adopted, their precise architecture is uncertain as of yet (Y. Zhang, Chen, & Gao, 2017; Internation Energy Agency, 2015). As such, it is unclear in which aspects smart grids will influence the results of this thesis.

Another projected transition is the expected decarbonization of heating and transportation sectors by means of electrification, coupled with a significant reduction in the carbon intensity of the power supply mix (Afman & Rooijers, 2017; J. Weiss, Hledik, Hagerty, & Gorman, 2017).
The impact of electric heat pumps on the distribution of domestic energy consumption has already been touched upon in the discussion above. Nonetheless, the transportation sector is also expected to undergo large-scale electrification, both industrially and domestically (Cazzola, Gorner, Munuera, Schuitmaker, & Maroney, 2017). The extent to which the shift towards e-mobility will affect the techniques discussed in this text, greatly depends on the role of electric vehicles in future passenger traffic. One possibility is that electric vehicles replace conventional cars as privately owned assets. In such a scenario, their batteries are likely to play a major role in the demand-response aspect of future smart grids (Tan, Ramachandaramurthy, & Yong, 2016) and in- and outflows of electricity for individual households may change drastically. In opposition, others foresee the simultaneous emergence of autonomous cars will set in motion a transition towards a ‘taxi economy’ in which hardly anyone owns a car (Arbib & Seba, 2017). If so, at-home energy consumption may stick more closely to the situation as it is now.

All in all, the exact way in which the above paradigm shifts will unfold are very uncertain at this point in time. For that reason, no general scheme has been developed to cope with possible consequences associated with these future developments for the techniques proposed in this text. Nevertheless, some of the solutions will include a short digression on possible ramifications for electric vehicle or heat pump owners, since such assets have already been adopted by a small segment of the population.
6 Smart meter solutions

In this chapter, a number of based feedback applications based on smart meter data are proposed. Most of them relate to electricity consumption in the residential market, but some of them can also be applied to gas consumption as well, and there are some sections on organizational energy usage. Each of the solutions comes with an extensive assessment of technical challenges, possible improvements and any dependencies on the data corrections discussed in the previous chapter; the techniques that have been realized based on own creativity will feature a more extensive technical discussion than those already well-documented in literature. All of the applications can be implemented on any platform of choice, such as a web page, a smartphone app or a local computer screen.

6.1 Conversion techniques

One of the reasons why people in general have difficulties understanding the impact of their consumption, is that entities like climate change, energy consumption and CO\textsubscript{2} emissions are invisible and intangible for consumers, see Section 3.4.2. Converting abstract consumption units like ‘kilowatt hours’ (electricity) and ‘cubic metres’ (gas) into variables people understand may help bridge this gap (Hargreaves, Nye, & Burgess, 2010). Two such alternative units will be presented in this section, namely costs (Section 6.1.1) and CO\textsubscript{2} emissions (Section 6.1.2). Since the latter still carries a somewhat abstract flavour, CO\textsubscript{2} visualizations will be discussed in Section 6.1.3. Figure 24 provides an outline of the conversion steps that will be covered.

<table>
<thead>
<tr>
<th>INPUT DATA</th>
<th>CONVERSION</th>
<th>COMPARISON</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity consumption (kWh)</td>
<td>Costs (€)</td>
<td>Emissions from everyday activities</td>
<td>GUI</td>
</tr>
<tr>
<td>Gas consumption (Nm\textsuperscript{3})</td>
<td>GHG emissions (kg CO\textsubscript{2}e)</td>
<td>Contents of familiar objects</td>
<td></td>
</tr>
<tr>
<td>GHG emissions (m\textsuperscript{3} CO\textsubscript{2}e)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 24:** The conversion steps from smart meter measurements towards CO\textsubscript{2} visualizations.

### 6.1.1 Consumption-to-cost conversion

In Section 3.1 the various players in the energy supply chain have been identified, all of which need to be payed when consuming energy. Dutch consumers transfer all their energy-related...
payables to their energy supplier, who in turn distributes the proper components of the total amount to the grid operator, transmission system operator and to the government. The exact layout of the energy bill may vary between different energy suppliers, but there are guidelines, which are monitored by the Dutch Authority for Consumers and Markets (Authority for Consumers and Markets, 2018c). The Dutch consumer association Consumentenbond has attempted to provide a general split-up of energy costs into four categories, shown in Figure 25 below (Consumentenbond, 2016).

<table>
<thead>
<tr>
<th><strong>Supplier fees</strong></th>
<th><strong>Grid costs</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy costs</td>
<td>Network costs</td>
</tr>
<tr>
<td>Fixed supplier fee</td>
<td>Capacity tariff</td>
</tr>
<tr>
<td>Regional supplement (G)</td>
<td>System services fee</td>
</tr>
<tr>
<td>Correction factor (G)</td>
<td>Meter costs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Taxes &amp; government levies</strong></th>
<th><strong>Other costs</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy tax</td>
<td>Other costs</td>
</tr>
<tr>
<td>Energy tax reduction</td>
<td>Offsetting of monthly payments</td>
</tr>
<tr>
<td>Levy for sustainable energy</td>
<td></td>
</tr>
<tr>
<td>VAT</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 25:** Splitting of a general energy bill into its various components. The bold entries are influenced by consumption.

The general supplier fees include the energy costs (based on the tariff per unit of electricity/gas consumed) and a fixed administrative supplier fee. The costs for gas consumption are furthermore subdued to a regional supplement (which depends on the distance between consumer’s location and the natural gas field in Slochteren, Groningen), as well as a caloric correction factor determined by the caloric value of the gas supplied (Consumentenbond, 2016).

The grid costs include a series of fixed costs, including network costs (a fee payed to the grid operator), a capacity tariff (based on the capacity of the connection, see Table 1 for the relevant classification), a system services fee (a fee for the transmission system operator) and meter-costs, i.e. rent for the energy meter installed in the residence. The category of other costs is a residual group containing for instance monthly installments, either already payed by the consumer or not. This category has mainly an administrative purpose and does not affect the relation between consumption and energy costs.

The government charges an energy tax on the amount of electricity and gas consumed. A fixed amount of this is returned to the consumer, as a warranty for a certain standard of living. The
bill also includes a levy for sustainable energy called the *Opslag Duurzame Energie* (ODE) in Dutch, the profits of which are used to subsidize the generation of renewable energy. Over the entire bill, 21% VAT is charged.

**Conversion**

Computing the costs per unit of energy consumed is not very complex per se. Grouping the consumption-dependent costs as $Y_{\text{con}}$, the fixed costs as $Y_{\text{fix}}$ (independently for gas and electricity) and using a factor 1.21 to account for the VAT charge, the total costs $Y_{\text{tot}}$ as a function of consumption is given by

$$Y_{\text{tot}}(C) = 1.21 \cdot (C \cdot Y_{\text{con}} + Y_{\text{fix}}).$$

However, the above formula is not suitable for calculating the costs of 1 kWh of electricity, as the total costs will almost entirely be determined by the fixed segments in this case. Therefore, for the applications used in this text, only the consumption-dependent factor $Y_{\text{con}}$ will be used for the conversion and $Y_{\text{fix}}$ is set at zero.\(^{(34)}\) When exercising this in practice, it is imperative the customer is aware of the fact their final energy will include fixed costs and will therefore be significantly higher than intermediate estimates which only show the consumption-related costs.

**Energy conservation**

Showing the monetary consequences of one’s energy consumption is a common way to convert energy usage into a more tangible variable. Yet, it can be misleading when energy prices fluctuate heavily and may be less persuasive for households where energy costs are low in proportion to their income (Froehlich, 2009). Additionally, too strong an emphasis on the monetary aspects of energy use may limit pro-environmental spillover effects (Steinhorst et al., 2015; Buchanan, Russo, & Anderson, 2015), as as been discussed already in Section 3.5.3. Although monetary information on energy consumption is a natural feedback mechanism which can be easily implemented (see Section 6.1.1) and is wanted by users (Vassileva, Odlaire, Wallin, & Dahlquist, 2012), it does not yield significant energy savings by itself (Schultz, Estrada, Schmitt, Sokoloski, & Silva-Send, 2015).

**6.1.2 Consumption-to-emission conversion**

Apart from the financial impact of energy consumption for the consumer, environmental effects can also be displayed. In this section, conversion figures for transforming smart meter consumption data into $\text{CO}_2$ emissions will be presented as a starting point for so-called carbon visualizations discussed in Section 6.1.3. This includes a calculation of the average amount of greenhouse gas (GHG) emissions from electricity generation and from the combustion of natural gas in the Netherlands.

The various conversion figures used in this section originate from CO2emissiefactoren.nl, an initiative from several organizations and the Dutch Ministry of Infrastructure and the Environment. Their aim is to standardize one set of emission figures to prevent an uncontrolled sprawl of different benchmark numbers for $\text{CO}_2$ assessments (Milieu Centraal, 2017a). The importance of such figures is underlined by Padgett, Steinemann, Clarke, and Vandenbergh (2008) who argue the success of carbon calculators regarding raising energy awareness, hinges on such calculators producing a unified output.

\(^{(34)}\)The electricity prices displayed in Table 2 comply with this convention.
6.1 Remark: Although CO\textsubscript{2} is the most common and well-known greenhouse gas, there are others which are stronger in terms of their heat absorption capabilities. Climate scientists use one measure for all GHGs to enable comparison of different mixtures of gases, based on the notion of *global warming potential* (GWP). The GWP of a greenhouse gas is a measure of the amount of heat trapped in the Earth’s atmosphere by 1 mass-unit of that gas, relative to the amount of heat trapped by the same mass-unit of CO\textsubscript{2} over a given period of time (a period of 100 years is commonly used, sometimes GWP\textsubscript{100} is used to specify the time interval). Global warming potential is measured in *CO\textsubscript{2} equivalents* (CO\textsubscript{2}e). Methane (CH\textsubscript{4}) for instance has a GWP\textsubscript{100} of 28 CO\textsubscript{2}e, i.e. 1 kg of methane contributes equally much to global warming as 28 kg of carbon dioxide over a period of 100 years (Myhre et al., 2013).

**Emissions from electricity generation**

Table 9 below shows GHG emission intensities of the most common electricity generation assets. The fossil mix is the share-based weighted average of greenhouse gas emissions from coal-, natural gas- and nuclear-fired power plants in the Netherlands. The emissions from renewable sources are from manufacturing of the generating units (solar panels, wind turbines, etc.) (Milieu Centraal, 2017c).

<table>
<thead>
<tr>
<th>Source</th>
<th>GHG emission intensity (kg CO\textsubscript{2}e/kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fossil mix</td>
<td>0.526</td>
</tr>
<tr>
<td>Wind</td>
<td>0.012</td>
</tr>
<tr>
<td>Hydro</td>
<td>0.004</td>
</tr>
<tr>
<td>Solar</td>
<td>0.070</td>
</tr>
<tr>
<td>Biomass</td>
<td>0.189</td>
</tr>
</tbody>
</table>

Table 9: Average GHG emission intensities from various electricity generation assets.

The emission figure for biomass requires some attention. CO\textsubscript{2} is released into the air when biomass is combusted, but it is common practice to assume an equal amount of carbon dioxide is absorbed in the process of growing the biomass feedstock (Koruba, Piotrowski, & Latosińska, 2017). With such a line of reasoning, one arrives at the conclusion bio-based energy has no impact on climate change. However, taking into account the many side effects of the entire biomass cycle, the climate neutrality of bio-based energy becomes questionable. In general, there are emissions from growing, harvesting, processing and transportation of the feedstock as well as *land use change* emissions, which are allocated when a biologically productive crop (such as a forest) in which much carbon is stored is replaced by a relatively less productive crop used for biofuel production (Zelm et al., 2015; Valin et al., 2015; Lange, 2011).

The biomass emission figure shown in Table 9 is an estimation of the average for the Netherlands, but the authors warn about the considerable differences between different bio-based power plants and advice to backtrace and use the estimated GHG emission intensity for each different bio-based power plant separately (Milieu Centraal, 2017c; Otten & Afman, 2015). The bio-based power plant in Huismerk Energie’s energy portfolio for example, runs on residual pruning materials from its nearby surroundings. With such a local waste-based setup come no cultivation or land use change emissions, and the impact of transportation of the feedstock is also smaller. The emission intensity from this particular plant are therefore likely to be lower than the national biomass average.

78
EMISSIONS FROM THE COMBUSTION OF NATURAL GAS

Table 10 shows emissions from the consumption of various types of gas (Milieu Centraal, 2017b). The large majority of gas consumption (99.1%) still concerns natural gas, the other 0.9% is biogas (Gerdes et al., 2017). Huismerk Energie supplies natural gas which is CO$_2$ compensated, meaning that CO$_2$ emissions from consumption are compensated for through investments in sustainable energy projects or afforestation. For lack of a standard benchmark figure, CO$_2$ compensated gas is assumed to be carbon neutral in this text.

<table>
<thead>
<tr>
<th>Source</th>
<th>GHG emission intensity (kg CO$_2$/Nm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural gas</td>
<td>1.887</td>
</tr>
<tr>
<td>Natural gas (CO$_2$ compensated)</td>
<td>0.000</td>
</tr>
<tr>
<td>Biogas (landfill gas)</td>
<td>0.398</td>
</tr>
<tr>
<td>Biogas (co-fermentation)</td>
<td>1.260</td>
</tr>
</tbody>
</table>

Table 10: Average GHG emission intensities from gas combustion for some types of gas.

MASS-TO-VOLUME CONVERSION

While GHG emissions are usually expressed based on their mass, some visualizations (such as the one shown in Figure 26 on page 80) are volume-based. Gay-Lussac’s ideal gas law describes the relation between volume and mass for any gas. For more details about this relation the reader can turn to any book on basic chemistry, e.g. (Timberlake & Timberlake, 2017). For now, it is only important to know the volume of a fixed amount of gas depends on temperature and atmospheric pressure. The relation between mass and volume for CO$_2$ (or any other gas) at a temperature of 0°C and a pressure of 1,013.25 hPa is given by

$$1 \text{ kg CO}_2 \leftrightarrow 0.509 \text{ Nm}^3 \text{ CO}_2.$$ 

Nm$^3$ stands for normal cubic metres, which is the unit for the amount of gas which has a volume of one cubic metre under the above temperature and pressure conditions.$^{(35)}$ Note that in case of a mix of GHGs, the above method converts a mass-unit of CO$_2$ equivalents into a volume-unit of CO$_2$ equivalents. Thus, any graphic based on the above conversion will show the volume of an amount of CO$_2$ molecules with the same global warming potential as the original mix of greenhouse gases.

6.1.3 CO$_2$ emission visualization

Computer-based visualization has been proposed as a countermeasure against the abstract flavour of climate change many times (Wibeck, Neset, & Linnér, 2013; Ballantyne, Wibeck, & Neset, 2016; Grainger, Mao, & Buytaert, 2016). Yet, scientific literature on visualizing household-specific CO$_2$ emissions is limited. Whilst various studies discussing smart meter feedback recommend including an estimation of energy-related CO$_2$ emissions to highlight the environmental impact of private energy consumption (Darby, 2006; Karlin, Zinger, & Ford, 2015; Zvingilaite & Togeby, 2015), none of those talk about putting those emissions into perspective. This is an unfortunate omission, as carbon pollution expressed as a plain number still make little sense to the average consumer (Whitmash et al., 2011).

$^{(35)}$In this section normal cubic metres will be used for computational efforts. However, many sources of literature do not specify temperature and pressure conditions when providing figures on volumes of natural gas.
Despite the lack of personalized visual representations of the environmental impact of at-home energy consumption, some organizations have created larger-scale CO$_2$ visualizations. An example is Carbon Visuals, who display total carbon dioxide emissions in New York through large blue balloons and show how those accumulate over a period of an hour, day or year, as shown in Figure 26 (CarbonVisuals, 2012). Although this may not be a very realistic depiction of reality (emissions from a city do not emerge at one spot and spread throughout the atmosphere rather than piling up at one location), it can help people to get a grasp of the amount of pollution from one city. A comparable example is Carbon Quilt, a tool which converts the CO$_2$ emissions from a city, country or energy source over a period of time into an area that could theoretically be covered by a blanket consisting of the carbon dioxide the thickness of a sheet of paper (CarbonVisuals, 2017). Both examples are forms of volume-based emission visualizations. In general, volume comparisons can be performed in many different ways, using all kinds of objects.

As for mass-based options, a simple but effective means of comparison is to relate GHG emissions from household gas and electricity consumption to the emissions from different transportation methods. Specifically, using the average emissions per kilometre from a medium-sized car, or any public transportation asset,\(^{(36)}\) a portion of carbon emissions can be expressed in terms of a distance by the chosen mode of transport. According to (Whitmarsh et al., 2011), “there is particular value in providing such comparative information so that individuals understand the relative contribution of different activities.” Indeed, when showing the environmental impact of private and public transportation methods simultaneously, consumers are pointed at the large environmental impact of personal car use compared to public transport alternatives.

Alternatively, the environmental impact of consumption can be related to the net CO$_2$ exchange rate of an area of forest over time. For example, one can think of showing the number of hectares of forest required to offset a consumer’s energy usage. General estimations of the necessary conversion figures can be found in a number of texts (e.g. Anthoni et al., 2004; Wofsy et al., 1993). However, carbon sequestration by forests or other terrestrial ecosystems is a complex topic. The total CO$_2$ exchange depends on a number of aspects and processes such as the latitude, age, soil type and tree compound of the forest (Anthoni et al., 2004). Also, though forests are generally considered to be a net sink of carbon, the majority of the stored carbon is released back into the air when wood eventually degrades (Schlesinger & Lichter, 2001). Due

\(^{(36)}\)These emission figures too, are available from CO2emissiefactoren.nl (Milieu Centraal, 2017d).
to this complexity, it may prove difficult to select a right number that connects to a customer’s understanding. A possibility would be to use a benchmark number for the uptake of an average forest in the Netherlands. This can be computed by using the total sequestration numbers for the country as a whole (which can be found in Lof et al. (2017)) and then divide by the total forested area (see e.g. Stichting Probos (2013)). One could then even relate a customer’s energy usage to a forested area on the map close to a customer’s home.

**Energy conservation**
Energy consumption expressed in terms of carbon emissions may raise awareness for the environmental side of consumption (Froehlich, 2009) and is more likely to stimulate trigger spillover to other environmentally friendly behaviours (Steinhorst et al., 2015). However, when shown to customers of an energy company supplying 100% renewable energy (such as Huismerk Energie), associated CO\(_2\) emissions will be very low and may give a skewed image of the actual environmental impact of energy consumption. Indeed, energy-related CO\(_2\) emissions do not contain any information on the scarcity of renewable energy: one’s personal energy-related CO\(_2\) emissions may be very low when using green energy, but the environmental consequences of using more renewable energy lie in the fact that someone else will need to use more fossil energy. Such indirect carbon emissions may be difficult to explain to a customer though.

### 6.2 Non-intrusive load monitoring

Numerous studies have reported particularly large energy savings can be achieved by presenting real-time information on consumption at an appliance-specific level (e.g. Batra, Singh, & Whitehouse, 2016; Neenan, Robinson, & Boisvert, 2009; Fischer, 2007). This naturally requires knowledge about the electricity use of each device in a building. Such data can be acquired directly by deploying a number of smart plugs, each measuring the power consumption of one specific device. This however requires a great amount of sensing equipment which can be costly (Zoha, Ghulak, Imran, & Rajasegarar, 2012). An alternative approach is to infer the consumption of each appliance by analyzing the aggregate load data monitored by a single meter. This technique is referred to as **non-intrusive load monitoring** (NILM). The recent large-scale deployment of smart meters in many countries has rekindled the interest in developing effective NILM methods. In this section, the basic principles of NILM techniques are presented. For a more detailed discussion of individual techniques, the reader can turn to extensive reviews of modern research on the topic, e.g. Tabatabaei, Dick, and Xu (2017); Faustine et al. (2017).

#### 6.2.1 Problem formulation

The load disaggregation problem has been defined as follows by e.g. Faustine et al. (2017). If \(a^1, \ldots, a^n\) are the electrical appliances in a building and \(P = (P_1, \ldots, P_N)\) is the aggregate load signal as measured by a (smart) meter at time stamps \(t = (t_1, \ldots, t_N)\), then the task is to find the individual power consumption \(p^j = (p^j_1, \ldots, p^j_N)\) of each appliance \(a^j\) such that

\[
P_i = \sum_{j=1}^{n} p^j_i + \eta_i,
\]

for all \(1 \leq i \leq N\). Here \(\eta_i\) is a (small) term of measurement noise. The problem has been depicted for \(n = 6\) appliances in Figure 27.
6.2.2 A state-based classification of appliance load signals

By a state, a steady operating condition of an appliance in which it draws a constant amount of power will be meant. In literature appliances are often classified based on the number of states they can attain (e.g. Zeifman & Roth, 2011; Zoha et al., 2012), see Figure 28.

1. **Single-state appliances** are either ON or OFF, e.g. light bulbs, toasters.

2. **Multi-state appliances** are devices which can attain a number of constant loads. Washing machines and clothes dryers are often named as examples (Zeifman & Roth, 2011; Liao, Elafoudi, Stankovic, & Stankovic, 2014), but in fact a lot of appliances have several modes which correspond to different levels of energy intensity.

3. **Continuously-varying appliances** are devices of which the load at some point fails to remain at a stable level. Refrigerator signals for instance tend to suffer from a substantial decrease in load between switching ON and OFF, and therefore fall in this category, see e.g. Aiad and Lee (2016), or Figure 11.
4. Permanent devices are those assets which continuously draw approximately the same amount of power. These include hard-wired smoke alarms, WiFi routers, alarm clocks and several appliances in standby mode (Zeifman & Roth, 2011). Two or more of such appliances cannot be distinguished from one another from an aggregate signal. Therefore, this category will be excluded from further examination in this section. However, power drawn by the class of all permanent and standby power devices as a whole can be estimated, see Section 6.3.4.

6.2.3 General NILM framework

While many different approaches to the NILM problem have been proposed, most share a few common steps. A general four-step framework for NILM techniques discussed by Zoha et al. (2012) will be adopted here, see Figure 29.

**Figure 29**: Common operation scheme for many NILM techniques.

### Data acquisition

First of all, an aggregate input signal needs to be acquired for developing the algorithm. For the testing and verification phase, it is convenient to have a sub-metered disaggregated copy of the same sample as well. Some techniques rely on such data in the training phase (see the discussion on inference and learning below).

The required sampling rate of the aggregate input data depends on the technique used. The majority of techniques are designed to be compatible with the future smart metering infrastructure and therefore rely on 1 Hz data (Bonfigli, Squartini, Fagiani, & Piazza, 2015). A few studies have been devoted to techniques based on data with even lower sampling rates (e.g. Powers, Margossian, & Smith, 1991; Batra et al., 2016; Al Abassi, Johnsson, Karlsson, Schön, & Wågberg, 2015), but these are a minority. Data of a higher sampling frequency (in the order of kHz or MHz) allows the development of more sophisticated algorithms, but requires additional sensing equipment. Examples of such methods are Bilski and Winiecki (2016); Bouhouras et al. (2017); Duarte, Delmar, Goossen, Barner, and Gomez-Luna (2012). This text being centered around smart meter-based solutions, the focus will mainly be on 1 Hz techniques here. Data need not be weather corrected for this method, but noise reduction is often applied (Mostafavi & Cox, 2017; Long, Chen, & Li, 2016). Whether PV generation needs to be accounted for depends on the NILM technique used.

### Feature extraction

The general strategy of most NILM algorithms is to recognize and separate devices based on one or more specific appliance features. Zoha et al. (2012) devised a classification of many appliance features found in NILM literature.
1. *Steady-state features* are aspects of an appliance signal when running steadily, see Figure 30. Examples of some features used in literature include active power (e.g. Parson, Ghosh, Weal, & Rogers, 2012; Liao et al., 2014), reactive power (e.g. Zeifman & Roth, 2011), duration (e.g. H. Kim, Marwah, Arlitt, Lyon, & Han, 2011), current (e.g. Makonin, Popowich, Bartram, Gill, & Bajic, 2013) and current and voltage waveforms (e.g. Barsim, Mauch, & Yang, 2016).

2. *Transient features* are characteristics of appliance state transitions. Examples of transient features used include the shape (e.g. Mostafavi & Cox, 2017), magnitude and sign (e.g. Zhao, Stankovic, & Stankovic, 2016), duration (e.g. Meziane, Ravier, Lamarque, Le Bunetel, & Raingeaud, 2017) and harmonics (e.g. S.-m. Wang & Bo, 2017) of such transients.

3. *Non-traditional features* are the remaining features which do not belong to either of the above categories. Some techniques for instance match the entire structure of an appliance signal from the instant it is switched on to the moment it is shut down (e.g. B. Liu, Luan, & Yu, 2017). Others include external factors in their method, such as the average consumption of neighbours (Batra, Singh, & Whitehouse, 2015).

NILM techniques are often classified as state-based or event-based depending on whether they use state- or transient features from appliances. However, a few techniques use neither (e.g. Batra et al., 2016) whereas some hybrid methods rely on both (Huss, 2015). State-based techniques usually invoke a modelling approach such as hidden Markov models or artificial neural networks (see Section 6.2.4), while event-based methods often rely on a feature clustering algorithm (Section 6.2.4). Again, there are exceptions: Le, Vrigneau, and Sentieys (2015) for instance have proposed a deterministic state-based approach.

Most techniques attempt to extract one or more of the above appliance features from the aggregate input signal and store these in a database. State features such as current waveforms generally require a sampling rate that is a multiple of 50 Hz (Zeifman & Roth, 2011), which is the fundamental mains frequency used in Europe (Neidhofer, 2011). For example, to determine a signal waveform, about 20 data points are required per fundamental frequency, so the sampling rate should be at least 1 kHz (Matthews, Soibelman, Berges, & Goldman, 2008). The features that can be captured from smart meter P1 data are shown in Table 11.

---

(37)i.e. whether a transient concerns a rising or a falling edge, see Figure 30.
### Table 11: Appliance features determinable from 1 Hz data. *Reactive power is not captured by all smart meters.

<table>
<thead>
<tr>
<th>Steady-state features</th>
<th>Transient features</th>
<th>Non-traditional features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real/reactive power</td>
<td>Event timestamp</td>
<td>Appliance load pattern</td>
</tr>
<tr>
<td>State duration</td>
<td>Real/reactive power edge magnitude</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Edge sign</td>
<td>Rough transient shape</td>
</tr>
</tbody>
</table>

Events, i.e. instants at which one or more appliances go through a state transition (Liao et al., 2014), can be located using an event detection algorithm. In case of 1 Hz data, this is usually performed by comparing the difference between power readings at subsequent time stamps \( t_i \) and \( t_{i-1} \) against a chosen detection threshold \( T_{\text{det}} \), \(^{(38)}\) i.e. an increment is detected if

\[
P_i - P_{i-1} > T_{\text{det}}.\]

Some authors have proposed a method of adaptive thresholding, that is, \( T_{\text{det}} \) is automatically adjusted to different situations (e.g. Zhao et al., 2016). At a sampling rate of 1 Hz, only macroscopic features such as the timestamp, magnitude and the sign of a transient can be extracted. In some cases, rough aspects of the shape of an edge such as its steepness and a transient bump can be captured, see Figures 30 and 12. To derive other features, higher sampling rates are required, as well as more sophisticated event detection algorithms. Examples of such can be found in Meziane et al. (2017); S.-m. Wang and Bo (2017).

### Inference and Learning

Every residence has its own unique mix of electrical appliances and usage patterns (Bouhouras et al., 2017). Therefore a NILM algorithm needs to be adjusted to a specific situation, which is done during a training phase. The exact procedure of this learning period depends on the NILM method used, but two strategies may be identified in general:

1. **Feature clustering:** Extracted features are considered ‘signatures’ of appliances. Similar appliance features are grouped and assumed to originated from the same device.

2. **Parameter setting:** Various parameters of a general model are calibrated to adjust to a specific building.

Regardless of the method used, a method’s training phase requires a certain degree of supervision. **Supervised methods** require labelled training data: sub-metered power readings from each appliance in the target building. The database is thus acquired intrusively, after which the algorithm takes over and performs disaggregation non-intrusively. Supervised methods generally achieve higher levels of accuracy than unsupervised methods (Tabatabaei et al., 2017), but their dependency on ground-truth data for each specific building decreases their attractiveness for large-scale application. Moreover, the sub-metering procedure must be followed again each time a new device is added to the building. **Unsupervised methods** build their database of appliance signals by discovering and storing patterns from actual data. Such techniques train ‘on the go’ and in general are more easily implementable but less accurate than supervised methods (Tabatabaei et al., 2017). Some modelling approaches use statistical methods to estimate parameter values (Parson et al., 2012). **Semi-supervised methods** use a small amount of generic

\(^{(38)}\)There will always be some variation in the load, the threshold is used to avoid detection of such ‘noise’.
labelled training data to improve the pattern recognition functionality applied by unsupervised methods. As such, semi-supervised methods manage to grab a bit of the benefits of both methods (e.g. Li, Sawyer, & Dick, 2015).

**Testing and Verification**

The final stage of the NILM method development process involves the testing of the algorithm. Apart from testing the general functioning and deciding on a safe training period of the algorithm, this should include an examination of the technical limits of the method. The energy disaggregation problem becomes increasingly hard to solve as the number of devices in a building grows. Also, if appliance usage overlaps their individual power consumption is indistinguishable. It is important to determine the technical limits of a developed NILM algorithm prior to developing a marketable product out of it.

### 6.2.4 Common NILM techniques

In the enormous number of studies performed on NILM methods, a few techniques are used most often. While the list of techniques described here is not exhaustive with respect to NILM techniques used in literature, many studies draw upon at least one of them. At the end of this section some results of each of the techniques will be presented.

**Hidden Markov models**

Appliance modelling with *hidden Markov models* (HMMs) is the most commonly used approach in appliance disaggregation (Huss, 2015). An HMM models an appliance as a sequence $s = (s_1, \ldots, s_N)$ of *states* and a corresponding sequence $p = (p_1, \ldots, p_N)$ of *observations* at time stamps $t = (t_1, \ldots, t_N)$. These observations represent the appliance’s load in its current state, i.e. $p_i$ is the power consumption of the appliance in state $s_i$. The central assumption is that the sequence $s$ is assumed to have the *Markov property*: each entry $s_i$ in the sequence only depends on its predecessor $s_{i-1}$ (Huss, 2015).

An HMM only models one appliance, so a factorial hidden Markov model (FHMM) is needed to model all appliances in a house. An FHMM is nothing but a series of HMMs run in parallel, see Figure 31. Each appliance $a_j$ is modelled by a separate HMM so that all observations $p_1^j, \ldots, p_n^j$ at time stamp $t_i$ add up to the measured total consumption $P_i$, i.e.

$$P_i \approx \sum_{j=1}^{n} p_i^j$$

at every instant $t_i$. The idea is to somehow infer the probability of every appliance shifting from one state to another, so that whenever an event takes place, the model calculates which device is most likely to have changed state (H. Kim et al., 2011). Many authors also model the probability distribution of each observation $p_i^j$ (e.g. Agyeman, Han, & Han, 2015; Bonfigli et al., 2015; Parson et al., 2012). The probabilities are usually assessed using labelled training data (e.g. Bonfigli et al., 2017; Kolter, Batra, & Ng, 2010), but can also be estimated in an unsupervised matter, for instance by using statistical methods (e.g. H. Kim et al., 2011).

---

(39) For instance if several devices are plugged into the same socket box which is turned on and off with a switch.
A drawback of the use of FHMMs is that they suffer from computational complexity issues, meaning that the model encompasses so many computational steps, an average computer cannot calculate the output fast enough. In general, the complexity of FHMMs grows exponential with the number of appliances (Long et al., 2016). To give the reader a feeling for what this entails, consider modelling a house with 10 single-state appliances for a one-day period with a resolution of 1 second. This already implies there are $2^{10 \cdot 24 \cdot 60 \cdot 60}$ possible state transitions to model.

To overcome complexity issues, so-called Monte Carlo simulations are often invoked, which are analogous to the technique of sampling in population research. The output of such a method is no longer a numerical value, but rather a probability distribution describing a range in which the output is most likely to be; the certainty of the output can be increased by increasing the size of the sample, but this increases the complexity (Robert, 2004). The aim of a Monte Carlo simulation can thus be thought of as finding the optimal balance between accuracy and speed. A Monte Carlo method often encountered in the NILM literature is Gibbs sampling (see e.g M. J. Johnson & Willsky, 2013; H. Kim et al., 2011).

**Artificial neural Networks**

*Deep learning* is a common machine learning technique used in many NILM methods and more generally to tackle pattern classification, clustering, function approximation, forecasting and optimization tasks (Jain et al., 1996). It uses *artificial neural networks* (ANNs), a type of model
inspired by the structure of the human brain. It has been shown ANNs are able to approximate any given function with arbitrary precision (Leshno, Lin, Pinkus, & Schocken, 1993).

A basic ANN is what is referred to in mathematics as a graph: a web of nodes connected by edges. The nodes represent (artificial) neurons and the edges allow information to pass from one neuron to the next. Neural nets come in various shapes and sizes (see e.g. Basheer & Hajmeer, 2000) but a large number them have their neurons arranged in layers (Poole & Mackworth, 2010). Each neuron in one layer is edge-connected to every neuron in the next, yet there are no interconnections between neurons within the same layer. The first and last layers are called the input and output layers respectively and those in between are referred to as hidden layers, some of which may serve a specific purpose within the network (see for instance W. He & Chai, 2016; C. Zhang, 2016). Many neural networks are feed-forward, meaning information flows from the input layer through each of the hidden layers towards the output layer, without any feedback loops (Poole & Mackworth, 2010).

Each neuron carries a set of weights \((w_1, \ldots, w_n)\), a bias \(b\) and an activation function \(\sigma\). The weighted sum of the \(n\) inputs from the previous layer and the \(n\) weights is computed, and the bias is added. The resulting number is inserted into the activation function and the resulting number is “fired” as output to all nodes in the subsequent layer (Basheer & Hajmeer, 2000). The activation function is typically a simple operation to make sure the output is bounded (Tomlinson Jr., Walker, & Sivilotti, 1990). In case of the energy disaggregation problem, the input layer consists of only one node in which the aggregate load signal measured by the (smart) meter is inserted. The output layer consists of a number of nodes equal to the number of appliances in the target building (C. Zhang, 2016). The number of intermediate nodes and layers depends on the technique used, but larger neural nets can have several thousands or even millions of nodes (Shazeer et al., 2017). Figure 32 below shows the schematics of a basic ANN applied to the energy disaggregation problem.

To apply a neural network in a particular situation it needs training. The training procedure of an ANN consists of calibrating its parameters to align the model’s output with ground-truth data. This thus requires labelled training data consisting of an aggregate load signal and \(n\) individual appliance load signals measured over the same period of time. The aggregate load signal is inserted into the input layer and the model produces an output. This output is compared to the ground-truth data of the individual appliances and the weights, biases and thresholds are updated to match the generated output with what it should be. There are various ways to update the parameters, an overview of these techniques can be found in e.g. Forouzanfar, Dajani, Groza, Bolic, and Rajan (2010). The data examples can be inserted either one-by-one or in batches, both methods have their pros and cons (see Basheer & Hajmeer, 2000).

**Feature clustering techniques**

Many non-modelling NILM techniques described in literature are based on matching appliance features. Various classification approaches have been proposed, both supervised e.g. Meehan, Mc Ardle, and Daniels (2014); Ardeleanu and Donciu (2012); Chou, Chuang, and Chang (2012) and unsupervised e.g. Zhao et al. (2016); T. Liu, Ding, and Gu (2015); K. D. Anderson (2014); Gonzalez, Debusschère, and Bacha (2012). Examples of clustering techniques found in NILM literature include \(k\)-nearest neighbours (e.g. C. C. Yang, Soh, & Yap, 2017; Chahine et al.,

\[\text{An ANN with loops is called a recurrent neural network (RNN) (Kelly & Knottenbelt, 2015).}\]
For smart meter P1 data, attention has been devoted especially to edge-clustering techniques, i.e. matching rising and falling edges in the active power load diagram (Tabatabaei et al., 2017). Such methods generally work very well for single-state devices, but sometimes fail when applied to multi-state and continuously varying appliances (Faustine et al., 2017). In general, two issues may be identified:
1. The magnitude of rising and falling edges from the same appliance does not coincide.

2. The number of rising and falling edges in an appliance signal is not equal.

The former situation can occur for both multi-state and continuously varying appliances (see Figure 28), whereas the latter is inherent to multi-state devices (see for instance Figure 30). Both situations can lead to mismatches when there are other appliances in the same power range, which may set-off a chain reaction of mismatches.

Edge-based techniques in general have difficulties working with devices with similar outputs. For instance, two appliances which produce rising edges of about 200 W are difficult to distinguish from each other. To overcome this, many techniques use multiple appliance features to increase the clustering. One such feature commonly used is reactive power, found in e.g. Kelly and Knottenbelt (2015); Barsim, Streubel, and Yang (2014); Jazizadeh, Becerik-Gerber, Berges, and Soibelman (2014). If one of the two 200 W-appliances above uses 50 VAR of reactive power while the other does not have any reactive component, the two appliances can still be distinguished. This subject will be elaborated on in more detail in Section 6.4.1.

Shape matching techniques
A few NILM studies have focused on matching the shape of an appliance output signal to those stored in a database. One technique for such efforts is dynamic time warping (DTW), a time-elastic similarity measure which has seen successful applications in voice recognition software (e.g. Muda, Begam, & Elamvazuthi, 2010). Its strength lies in comparing two similar samples, one of which has been shifted or stretched in time (B. Liu et al., 2017). This is especially convenient for matching e.g. a refrigerator load signal, as the running duration of a refrigerator can be variable (see Figure 11). The output of the DTW procedure is an optimal warping path between two signals, as shown in Figure 33 below. The sum of all individual warping distances defines a total “DTW distance” between the two signals, and can be used as a measure for the similarity between the shapes of both signals. See B. Liu et al. (2017) or Liao et al. (2014) for a complete description of the DTW technique itself and how it applies to NILM.

It can be challenging to isolate each signal from the aggregate load. Liao et al. (2014) have proposed a way to achieve this by subtracting each appliance signal from the main load after it has been matched, however this might turn out to be complicated in practice. Also, the DTW procedure suffers from a high complexity, which is why Monte Carlo simulation techniques are often implemented (e.g. Niels & Vuurpijl, 2005). To reduce the amount of computations, there might be possibilities for a co-op method with an edge-detection feature: once a rising edge has been detected, only those signals in the database which kick-off with a similar rising edge are considered, instead of performing DTW on every signal.

6.2.5 Overview of results

Performance metrics
Comparing the results of individual techniques with each other is difficult. In the first place because researchers use different metrics to measure the performance of their techniques. Faustine et al. (2017) have identified as much as 11 different metrics, various combinations of which are used for comparison efforts in NILM literature. Many techniques have their own strong points and limitations, which cannot be captured by a single metric. A few different domains
in which techniques may or may not stand out include performance per appliance type, adaptability to various situations and increasing number of appliances, computational speed, degree of self-sufficiency, learning time, and complexity (Tabatabaei et al., 2017; Faustine et al., 2017; Klemenjak & Goldsborough, 2016). Some NILM publications attempt to bundle some of these metrics for a broader comparison: Bonfigli et al. (2017); Cominola, Giuliani, Piga, Castelletti, and Rizzoli (2017) for instance included radar charts to compare the appliance-specific performance of their technique to that of other methods.

A common performance measure is the accuracy, which is defined as the ratio between the number of correct matches $M_C$ and the total number of matches that could have been detected in theory, i.e. sum of the number of correct matches and the number of false matches $M_F$:

$$\text{Acc} = \frac{M_C}{M_C + M_F}.$$  

However, Faustine et al. (2017) remark this metric is unsuitable for appliances which are rarely used: if a TV is on about 10% of the time, any algorithm predicting it is never used already scores an accuracy of 90%. Therefore, more sophisticated metrics have been designed, but not everyone uses these (Faustine et al., 2017).

Also, measuring the performance of NILM algorithms has been constrained by the limited number of datasets available online (Faustine et al., 2017). Notably house 2 from the REDD database is a popular test case on which many empirical evaluations are based (e.g. Cominola et al., 2017; Aiad & Lee, 2016; Liao et al., 2014). On the one hand this is a positive thing, as the differences in the disaggregating performance of various techniques are laid bare. However, this particular residence of the REDD database is the one with the smallest number of appliances (with a
total of 10 appliance groups (Kolter & Johnson, 2011)) and therefore a relatively easy setting to disaggregate. Therefore, this house is insufficient regarding testing an approach’s to deal with complex situations. In general the limited amount of test data limits verification of each technique’s ability to adapt to different situations. In particular, the author has not encountered any experimental setups with more than one refrigerator.

Comparison
Despite the large amount of modern studies towards developing NILM methods (41), it seems an efficient and accurate solution has yet to be found. In general the best NILM algorithms using 1 Hz frequency data encountered by the author attain reported accuracy scores of around 80%. Precision scores of some techniques based on higher frequency data reach up to 90%, but these are not compatible with the current smart metering infrastructure.

Although in view of the above it is difficult to rank different NILM approaches, a number of studies have attempted to assess the performance of different approaches to the NILM problem. A few comparative studies reported deep learning approaches outperformed the HMM-based methods they examined (C. Zhang, Zhong, Wang, Goddard, & Sutton, 2016; Bonfigli, Fagiani, Squartini, & Piazza, 2016). A larger cross-method comparison by Bonfigli et al. (2015) gathered results from 11 different studies using various approaches. Their comparison is based on the original test results of each of the techniques, which are rather diverse in terms of the settings making it difficult to draw any hard conclusions. Generally, the tested methods showed an accuracy of around 80% for simple houses, dropping to 70% for more complex testing environments.

Additionally, a few energy monitoring devices available have implemented a non-intrusive disaggregation feature. Large market players include Smappee (Smappee N.V., 2018), Neurio (Neurio, 2018) and Sense (Sense, 2018). Unfortunately there seems to be no study comparing the technical performance of these systems. However, Rosberg (2016) interviewed some Smappee users, who reported numerous issues, including the mixing up of different appliance signals and difficulties when dealing with multiple refrigerators.

6.2.6 Applications

Important applications of NILM lie in locating inefficiencies in specific devices (Chakravarty & Gupta, 2013) and in split billing: providing customers with a break down of their energy costs per appliance, using the conversion methods discussed in Section 6.1.1. However such applications can only be realized if the underlying NILM structures are reasonably precise, which is doubtful for the moment. However, even if the NILM theory developed up until this point turns out to be too restrained to be implemented in a marketable product, some aspects may still be used for smaller applications. The refrigerator payback time assessment feature discussed below is an example of such practice.

Refrigerator Payback Time Assessment

As discussed in Section 3.1, refrigerators are the largest power drainers in the domestic setting, constituting about 13% of an average household’s consumption in the Netherlands (Gerdes et al., 2017). The past few years, considerable headway has been made in improving the energy

---

(41) If any indication, Google Scholar reports 1436 citations of Hart’s initial paper on the topic (Hart, 1992), of which over 1330 from after 2009.
efficiency of refrigerators. Yet, many old models are still in circulation, often used as a so-called “beer fridge” (Young, 2008). Milieu Centraal (2017e), has provided a list of the energy consumption of older models refrigerators shown in Table 12. Of course the exact electricity use also depends on other factors, such as size, door-open time, etc.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Average annual consumption (kWh)</th>
<th>Annual costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>160</td>
<td>€29.95</td>
</tr>
<tr>
<td>8</td>
<td>350</td>
<td>€65.51</td>
</tr>
<tr>
<td>10</td>
<td>380</td>
<td>€71.13</td>
</tr>
<tr>
<td>15</td>
<td>390</td>
<td>€73.00</td>
</tr>
<tr>
<td>20</td>
<td>450</td>
<td>€84.23</td>
</tr>
<tr>
<td>25</td>
<td>500</td>
<td>€93.59</td>
</tr>
</tbody>
</table>

Table 12: The electricity consumption and associated costs (based on the tariffs in Chapter 2) of refrigerators. The energy consumption of a new refrigerator is based on an A+++ model.

It can be seen that it often pays off to replace an old refrigerator by an energy-efficient one on the long term. In general, the payback time $t_{\text{payback}}$ (in years) of such an investment can be estimated, based on the annual energy consumption of the old and new model $C_{\text{old}}$ and $C_{\text{new}}$ and the price $Y$ of the new refrigerator:

$$t_{\text{payback}} = \frac{Y}{C_{\text{old}} - C_{\text{new}}}.$$

The price and consumption of the new model can either be based on an average number or on values provided by the customer. The consumption of the old model can be based on the age (given the customer provides this information), or it can be estimated from a smart meter signal using NILM. Especially HMM-based NILM techniques, have shown reliable performances when it comes to detection of refrigerators (Faustine et al., 2017; Liao et al., 2014).

Alternatively, there might be possibilities to extract the refrigerator signal heuristically, using a hybrid method relying on a number of techniques discussed in this thesis. Using the base power figure found in Section 6.3.4, one can determine resting hours in the target house. During these hours, the task of energy disaggregation is relatively easy, as there are very few active appliances. Using a generic database of refrigerator signals, a DTW algorithm can determine which signal(s) is or are most likely to originate from a refrigerator.

**Energy conservation**

As mentioned already on numerous occasions in this text, appliance-specific feedback has been found to be a very effective way to achieve energy reductions with savings in the range of 7-12%. According to (Karjalainen, 2011), a disaggregation by appliance provides delivers valuable information to help users get a feeling for the relevance of individual actions. As has been discussed in Section 3.4.2, knowledge among consumers about their own total energy use and the way individual appliances contribute to this amount is inadequate as of yet and disaggregated feedback may help fill this gap. For this to work however, the author believes it is invaluable the disaggregated information is correct, since an inaccurate allocation of energy use to certain appliances may lead to consumers undertaking rigorous action they believe is significant, while
actual savings are disappointing. Based on the above technical discussion and the results from the best-working algorithms, it is highly uncertain if NILM is a good approach for achieving energy savings.

Compared to NILM in general, the refrigerator payback time assessment described above seems to be more reliable from a technical perspective as of now. Using the figures from Table 12 and the fact that refrigerators make up about 13% of household energy use on average in the Netherlands, one can calculate that replacing a 15 year old refrigerator by a new one would induce energy savings of about 7.7%. However, this is under the assumption that users indeed decide to replace their refrigerator upon receiving the feedback, dispose of their old one and refrain from selling or giving it to anyone else. Also, there will be no effect at all for consumers who already own an efficient model. All in all actual savings are most likely significantly lower than the 7.7% from the naive estimation above.

6.3 Consumption extremes

In this section, the upper and lower ends of energy consumption are examined. After a brief introduction of some notions, several techniques will be discussed to identify various types of extreme values in both P1 and P4 signals.

6.3.1 Discussion of typology

When it comes to electricity demand and consumption, a number of different ‘highs’ and ‘lows’ can be observed. The upper and lower regions of the main load will be disassembled into different components. Most of these components have also been identified in one or more other studies, but there is little uniformity in the terminology used. To avoid any mix-ups, one set of notions is fixed in this section, which will be used in the remainder of this text. Most of the definitions given are depicted in Figure 34 below.

- The peak load will refer to the highest power demand of a household or firm within a certain period (Gönen, 1986). The absolute load maximum is often attained during a very short-lived transient current spike. Transient spikes are often caused by inrush currents drawn by an appliance switching on (Waegeneers & Jacobs, 2014), but can also appear at random other instants (see Figure 36 on page 104). For some applications however, the highest load value irrespective of such spikes is more interesting, as will be discussed in Section 6.3.2. Therefore, in this text two different peak loads will be distinguished: the absolute load refers to the maximal load value measured (possibly during a transient), whereas the non-transient peak load will be the maximal load value after all transient peaks have been shaved off in the data sample.

- In this text peak time corresponds to the daily 15-minute window during which consumption is maximal. Peak consumption equals the total consumption during this window.

- Standby power refers the electricity used by appliances that are not performing their primary purpose but are ready to do so (Siderius, 1998). Examples of notorious standby

\[^{(42)}\text{From a rebound perspective this is especially bad, as in such case the consumer will get an unwarranted feeling of having done the right thing, which may instigate a number of psychological rebound mechanisms. Due to the actual savings being so low, this might even lead to backfire phenomena in such situations.}\]

94
users include numerous multimedia appliances (TVs, PCs, VCRs, etc.), as well as kitchen devices such as coffeemakers, ovens and microwaves (Sigurjónsdóttir, 2013).

- Some appliances such as and WiFi routers, clocks, alarms and monitoring systems are often considered to be ‘on’ rather than in standby (Harrington, Siderius, & Ellis, 2008). These devices generate a flat load and are sometimes referred to as always-on appliances or continuous appliances (Bijker, Xia, & Zhang, 2009; Firth, Lomas, Wright, & Wall, 2008).

- The individual categories of continuous and standby mode appliances are indistinguishable from an aggregate smart meter signal. Collectively, such devices define the base load, defined by Camilleri, Isaacs, and French (2006) as “the typical lowest power consumption of the entire house when there is no active occupant demand and all cycling appliances (e.g. refrigeration) are in off-cycle.” Such cycling or periodic devices turn on automatically every once in a while, and the power drawn by this class of devices will be referred to as intermittent loads. Apart from refrigerators, examples of periodic devices creating intermittent loads are freezers, air conditioners, furnaces, aquarium heaters, and (de-)humidifiers (Delforge, Schmidt, & Schmidt, 2015).

- Total electricity consumption associated with the base load over a specified time frame is referred to as the base consumption.

- Base power and consumption may in some cases be difficult to determine from a smart meter P4 signal as will be discussed in Section 6.3.4. A generally more easily computable type of ‘low’, is the mean idle load, which refers to the average power measured in a residence or office when there is no human activity. This encompasses the base load and the mean value of all intermittent loads. Total consumption during those instances is referred to as idle consumption.

![Figure 34: Schematic overview of several of the notions discussed above.](image)

\[\text{Figure 34: Schematic overview of several of the notions discussed above.}\]

\(\text{(43)}\) Figure 11 shows a picture of a load signal from a refrigerator which may help give the reader a feeling for what consumption patterns of such appliances may look like.
There is no general consensus regarding the various notions distinguished above. In particular, the terms ‘peak load’, ‘base load’ and ‘standby power’ are used to describe different quantities. In energy demand management, peak and base load refer to the highest and lowest values in overall energy demand (see e.g. Sijm et al., 2017). Some authors (e.g. Yohanis, Mondol, Wright, & Norton, 2008) use the term base load to refer to what is defined here as the mean idle load.

Also, the distinction between standby power and always-on consumption is not always made. Some authors define the standby or lower-power mode of an appliance as “the lowest level of consumption that cannot be switched off” (e.g. Nipkow & Bush, 2003). This definition tacitly marks the operational mode of many always-on appliances as a standby mode, since this often is the only possible state for such devices. Harrington et al. (2008) on the other hand does discriminate between always-on and standby consumption, but continues by stating the former category “can easily and appropriately be included within a broad standby policy approach as their design and main function resembles the secondary functions for many other products of concern.”

Even though this might be true, it is worth considering there are differences between the two. One such is the fact that appliances in low-power mode can often be shut down without greatly affecting the overall well-being of the user (e.g. turning off a TV in standby mode only increases the booting time by a few settings), whereas unplugging a continuous device often has more severe consequences (upon shutting down the WiFi router for instance, the internet signal is lost). Thus even though both types of could be targeted by the same policy measures, the resulting energy reductions for both categories may be very different. From a technical point of view, standby and always-on consumption are virtually indistinguishable as separate categories from a smart meter signal, unless any additional knowledge about the appliance mixture in the residence is available. Therefore, in the discussion on smart meter data solutions below, only the base load will be considered.

### 6.3.2 Peak load

Peak load is a notion mainly important for grid operators and transmission system operators, but in some cases, it can make a difference for organizations as well. A consumer’s peak load determines the required capacity of their connection to the grid and therefore the network costs charged by the grid operator (van Wezel, 2015). An overview of some connection capacities is given in Table 1 in Chapter 2. In case a firm’s peak load is close to the margins of its capacity, the company could look into reducing their peak load in order to remain in or enter a lower price class.

The required capacity of a grid connection is determined by the absolute peak load, which may or may not correspond to a transient spike (van der Hoeven, 2008). Fuses and circuit breakers on the other hand typically have an in-built delay mechanism allowing short pulses of up to twenty times the normal tripping current to pass through (ABB SACE, 2007). The required threshold value of such safety components therefore depends on the non-transient peak load. Hence both peak load quantities can be relevant to individual consumers, although in different ways.

Peak load values may further become an increasingly important aspect of local grid management in the not too distant future. The total capacity of a local grid is mainly determined by non-transient peak loads of individual households connected to it, assuming inrush currents generally
J. Liu (2013) found that at-home charging of an electric vehicle increases the daily (non-transient) peak load of an average household during peak times by 71.4%. The expected rise in popularity of electric vehicles will lead to more instances of overloaded local power grids in the near future (Sijm et al., 2017).

**Determining the peak load**

As noted in Section 5.1, transient spikes can only be captured reliably by energy meters which sample at a frequency of 1 kHz or higher. Although they are occasionally registered in P1 data of a smart meter (see Figure 36), additional metering is required for an accurate estimation of the absolute peak load. The non-transient peak load can however be determined from smart meter P1 data, since transient peaks can be shaved off using a median filter. When executing this, it is recommended to use a filter with a very high range parameter, in order to make sure all peaks are tackled by the filter.

**6.3.3 Peak consumption**

In the residential case, the highest electricity consumption is usually attained during the evening, as shown by Figure 20 (p. 65), but the situation may be different for individual consumers. Firms attain their highest consumption somewhere during office hours, the exact time of which depends on the type of company. Understanding of peak consumption can help the customer place their total consumption into perspective. Also, it is used in the peak saturation change detection technique discussed in Section 6.4.2.

**Determining peak consumption**

The peak consumption can be extracted from P4 data as the maximal entry among daily consumption recordings. In case of P1 data, additional options are available. In general, at a sampling rate of 1 Hz, the peak consumption corresponds to the maximal string of 900 measurements. However, one need not stick to the clock-aligned time stamps standardized for P4 data. Also, for some applications (such as the one in Section 6.4.2) time windows of duration shorter than fifteen minutes may sometimes be preferable.

**Energy conservation**

As of now, peak load and peak consumption do not seem to be very relevant in terms of energy savings. Some feedback studies report on peak demand reductions too (e.g. Darby, 2001), but this seems to be a consequence of more general savings rather than the focus of these studies. Still, ‘peak shaving’ may become an important aspect of energy feedback in the future, when electricity grids start to suffer from higher peak consumption due to electrification, as the test conducted by Liander in Lochem has shown. As such, one could just as well start with feedback on peak energy use now. However, if peak energy reductions are achieved by a displacement of energy consumption to off-peak hours, possibly stimulated by an off-peak pricing mechanism (Holland & Mansur, 2008), then there will be no energy savings.

---

Footnotes:

(44) The probability of multiple households starting an appliance at exactly the same instant is very small, apart from the situation when power is restored after a blackout.

(45) A stress test conducted by Liander in Lochem some years ago confirmed local grids are currently unable to handle large e-mobility related increases in peak electricity demand. (Alliander N.V., 2016).
6.3.4 Base load

There are two factors that contribute to the electricity consumption of an appliance, namely the power it draws and the total time it is used. Some devices, such as a washing machine, draw a huge amount of power when on, but are generally off most of the time while other appliances (semi-)continuously run at a modest wattage. The latter class of devices may seem innocent, but can induce significant amounts of electricity consumption figures on a yearly basis. An imaginary 1W-appliance running non-stop, uses 8.76 kWh of electricity a year, which (based on the electricity tariffs from Chapter 2) costs about €1.64, i.e.

| 1W always on | $8.76$ kWh/yr | $€1.64$/yr |

Literature on estimations of the impact of base consumption is rare, but the subclass of standby power-related consumption has been subject of several studies. Standby power as a phenomenon has been identified in the 1980s (Harrington et al., 2008) and deemed a challenge by Sandberg (1993), who drew attention to the increasing number of appliances with standby modes and the associated ‘energy leakage’ (Gram-Hanssen, 2010). Several studies have estimated the impact of standby power consumption. A. K. Meier (2001) showed standby power varies greatly between different countries, ranging from about 30W in China to over 100W in New Zealand and the United States, although the exact numbers may have changed by now. Ross and Meier (2001) pinpointed a range from 14-169W (5-26% of consumption) with an average of 67W in California, USA. A more recent study in the EU-27 found home appliance standby consumption contributed for about 5.9% of total consumption, 47.5 TWh each year (Bertoldi, Hirl, & Labanca, 2012). Estimates from the Netherlands are relatively rare, Siderius (1998) showed it accounts for about 10% of national residential electricity use, a number acknowledged by a few more recent studies (e.g. Papachristos, 2015; Gram-Hanssen, 2010). Based on an average household electricity consumption of 2966 kWh per year (Gerdes et al., 2017), this means a little less than 300 kWh annually, which is equivalent to about 34 W of continuous standby power. This is more or less in line with the amount of 39.8 W found by De Almeida, Fonseca, Schlomann, and Feilberg (2011) based on the average standby power drawn by certain appliances.

Determining the base load

In contrast to the peak load, the base load is a rather complex quantity which can be difficult to determine from a smart metering signal. Two techniques targeting base load values in load and consumption data will be proposed. The first method is a P1-specific option which attempts to determine the base load as a flat level during so-called base load hot spots i.e. instances at which there is no human activity in the building and no periodic appliances are running, see Figure 35 on page 101. The second method attempts to estimate the base load based on recurring values in the data and has a higher P4-compatibility. While the methods are very different, both assume a certain level of stability to the base load quantity. Also, both estimation methods are discussed in a household setting. The mean idle load discussed in Section 6.3.5 is unfolded from a business viewpoint, as it is expected to be a more reliably determinable and more interesting quantity for organizations than the base load.

Flat level method

The flat-level method uses the base load’s assumed stability and the higher resolution of smart meter P1 data to assess the base load as a stable level of relatively low electricity usage. It is a flexible technique, adjustable to various different situations.
**Setting**

The input of the proposed technique is a time series \( t = (t_1, \ldots, t_N) \) and corresponding power readings \( P = (P_1, \ldots, P_N) \). The following conditions are presumed:

### 6.2 Assumption:

1. A household’s base load is a steady quantity.
2. The base load dominates total consumption in a residence on a regular basis, for a duration covering several consecutive data samples.

The assumptions jointly guarantee the visibility of one or more base load hot spots in the data. The above assumptions translate to the following: there exist at least one sequence of consecutive load samples \( P_i, P_{i+1}, \ldots, P_{i+n} \) in \( P \) that are almost equal to the average base load \( P_{\text{base}} \), i.e.

\[
|P_{i+j} - P_{\text{base}}| < T_{\text{tol}}
\]

for \( j = 0, 1, \ldots, n \) and some small tolerance \( T_{\text{tol}} \). The extent to which the above statements are expected to hold true in practical situations will be assessed in the theoretical discussion below.

**Method description**

The flat level method attempts to locate the consecutive values \( P_i, P_{i+1}, \ldots, P_{i+n} \) which meet the condition in (3) above. However, since \( P_{\text{base}} \) is unknown, the technique resorts to finding contiguous series of \( P_i \)'s swaying around some low value, while avoiding detecting any spike values resulting from faulty measurements. The data should be pre-processed using a median filter to remove such outliers, however it cannot be assumed the median filter targets them all \(^{46}\) unless a very drastic filter is applied. The proposed algorithm is designed to be able to deal with some remaining noise. Also, it features a number of scaling parameters which add a degree of flexibility to the technique regarding its ability to cope with different situations. The role each of the parameters has within the algorithm will be discussed in more detail after the technique itself has been described.

Only the most recent measurements will be used, so in the first step of the method an integer number \( M \leq N \) is chosen and the \( M \) last values \( P_{N-M+1}, \ldots, P_N \) and \( t_{N-M+1}, \ldots, t_N \) are pulled from \( P \) and \( t \). As the other samples will not be used any further, the extracted samples may be relabeled as \( P = (P_1, \ldots, P_M) \) and \( t = (t_1, \ldots, t_M) \) respecting the original order. Choosing another integer \( L < M \) the \( L \) lowest load values \( p = (p_1, \ldots, p_L) \) are isolated from \( P \). Thus each \( p_i \) corresponds to some \( P_j \) in \( P \) and all remaining values in \( P \) are at least as large as the largest one in \( p \).

The next step is to investigate whether (the majority of) the \( L \) low points \( p_1, \ldots, p_L \) found indeed relate to base load hot spots and not to occasional noise values which survived the median filter. This is achieved by assessing the stability in load among the neighbouring entries of each \( p_i \) in \( P \). For this an *adaptive thresholding* method is deployed: an iterative technique for determining

\(^{46}\)see for instance Figure 36 below which shows numerous spikes, even though a relatively drastic median filter was applied
the ‘perfect threshold’, by consecutively increasing a small initial threshold $T_0$ by tiny amounts $T_{\text{raise}}$ until a certain predetermined condition is met.

This is carried out as follows. Suppose $p_i$ corresponded to $P_j$ in the original sequence $\mathbf{P}$. The difference between $P_j$ and $P_{j+1}$ is compared to the initial threshold $T_0$. If this difference is smaller than or equal to $T_0$, the algorithm moves on to the next sample, i.e. it checks whether $|P_j - P_{j+2}| \leq T_0$ and so on. This procedure is continued until for some $P_{j+r_{j,0}}$, the difference is larger than $T_0$, in which case the number $r_{j,0}$ is stored. Similarly, the algorithm checks back in time from $P_j$ to assess whether the distance between $P_j$ and consecutively $P_{j-1}, P_{j-2}, \ldots$ is smaller than $T_0$. The number $s_{j,0}$ corresponds to the first entry $P_{j-s_{j,0}}$ which differs more than $T_0$ from $P_j$. This concludes the first iterative step, and the result is a maximal interval of consecutive samples

$$(P_{j-s_{j,0}+1}, \ldots, P_{j-1}, P_j, P_{j+1}, \ldots, P_{j+r_{j,0}-1}),$$

which are less than $T_0$ away from $P_j$. The length $l_{j,0}$ of this interval equals $r_{j,0} + s_{j,0} - 1$, and is at least 1 (since the interval always contains $P_j$). The interval can be of short length though: in the most extreme case it only contains $P_j$, either because $P_j$ corresponds to an isolated low value (i.e. both $P_{j-1}$ and $P_{j+1}$ are much larger) or because the $T_0$ was just too small to cope with load fluctuations.

This procedure is carried out for each $p_i$. The majority of resulting intervals should be sufficiently lengthy to guarantee they correspond to stable lows. In other words, at least $\alpha\%$ of the interval lengths should exceed a predetermined tolerance parameter $T_{\text{tol}}$. If this is not the case, threshold $T_0$ is increased by $T_{\text{raise}}$ and the intervals are recalculated. With this higher threshold, these new intervals are at least as long as their predecessors.

The above process is re-run for a number of iterative steps, with increasingly lenient conditions, until the majority of intervals is sufficiently long. In the $k$th iterative step the threshold is set as $T_k = T_0 + k \cdot T_{\text{raise}}$ and $r_{j,k}, s_{j,k}$ are the smallest positive numbers for which

$$(P_{j-s_{j,k}}, \ldots, P_{j-1}, P_j, P_{j+1}, \ldots, P_{j+r_{j,k}-1}),$$

$|P_j - P_{j-s_{j,k}}| > T_k,$

$|P_j - P_{j+r_{j,k}}| > T_k.$

Figure 35 depicts the above procedure for the $k$th iterative step.

After each iteration it is checked whether any of the interval lengths $l_{j,k}r_{j,k} + s_{j,k} - 1$ exceeds $T_{\text{tol}}$. If this is the case, the interval is stored and excluded in any further iterations to avoid unnecessary computations. When after some $K$ steps, some $\alpha\%$ of the intervals is indeed larger than $T_{\text{tol}}$ the iteration process ends. At this point, all intervals with $l_{j,K} < T_{\text{tol}}$ are discarded. Of each of the remaining intervals, the median of its power values is computed. The base load can be set as the minimum over all these interval medians. Alternatively, multiple quantities can be returned if there are multiple stable low levels (as is the case in Figure 37 on page 107).

**Discussion of assumptions**

Although the author has not been able to find any direct evidence of the first statement of
Assumption 6.2 in literature, there are numerous indications to support the claim there is a certain level of stability in the base load in general. For starters, fluctuations in standby power found by Camilleri et al. (2006) and Nipkow and Bush (2003) were generally very small, as were load fluctuations from common Hi-Fi systems (Horowitz et al., 2013) and the smoke alarm in Houses 3 and 4 from the REDD database (Kolter & Johnson, 2011). The data examples used for testing in this text also support the stability assumption, although in one case there were multiple steady levels (see Figure 37). Base load levels may in general attain different stable values over sustained periods of time, for instance when people leave on different combinations of lights over night, or occasionally turn off their Hi-Fi equipment. Both base load estimation techniques discussed here can return multiple outputs corresponding to these different levels if desired. Ultimately, the stability assumption is no tough requirement: some smaller fluctuations in load are permissible, as the proposed adaptive thresholding feature is designed to deal with those.

The second condition from Assumption 6.2 on the other hand is imperative, in the sense that the above method will fail if the base load is persistently masked by loads from the alternating use of other appliances. The extent to which the base load is visible in the data is influenced by a number of factors. While it may not be unrealistic to assume a period during which no electricity is actively consumed regularly occurs in the majority of residences (e.g. when all occupants are away or asleep), this hypothesis may fail to hold for dwellings housing larger numbers of people (fraternity homes, shelters, etc.). Also, the amount of base load quality time may be reduced if householders structurally make use of off-peak tariffs by running washing machine and dishwasher programs at night. Multiple intermittent loads from different periodic devices can cover large parts of the base load signal as well. The extent to which (a combination of) the above events completely conceal the base load, needs to be determined through a series of test runs. In the end, with the proper parameter configuration, the flat level technique only requires a string of base load values covering a few samples, which (considering a P1 signal contains 86,400 samples a day) is not expected to be a major constraint in most domestic situations.

**Data corrections**

As mentioned already in the outline of the method, a median filter should be applied to get rid
of as much noise as possible. As the aim of the algorithm is a steady level which is determined by taking averages anyway, there might be opportunities for setting filter’s range parameter quite high in order to remove large amounts of noise.

When customers have solar panels, the above technique can still be applied to the hours between sunset and sunrise when there is no generation. This obviates the need for PV estimation and is not expected to give problems in most cases, since for the majority of households, the base load is attained at night (Nelson & Hydro, 2008). Nevertheless, if a household has one or more night owl members, or structurally exploits night tariffs by running washing machine and dishwasher programs at night, a restriction to night time may cause the technique to fail.

When running the method in night-only mode, it is recommended to use a longer data window, to increase the chance of capturing a base load hot spot. This is wise especially in summer, when the time between sunset and sunrise is less than 8 hours.

Choice of parameters
Overall, as much as six parameters can be set for this technique. These are $M$, the window length over which the base load is estimated; $L$, the number of lowest values used from this window; $T_0$, the initial threshold; $T_{\text{raise}}$, the amount to increase the threshold by in each iterative step; $T_{\text{tol}}$, the minimum average interval length and $\alpha$, the number of intervals which should meet the interval length condition. Each of these will now be briefly discussed regarding its impact on the performance of the technique.

The window size $M$ is meant to restrict the algorithm to recent consumption data only, allowing the output to be updated regularly. This window should be long enough to contain at least one base load hot-spot, or the algorithm will wrongly label a higher power level as the base load, thus longer window setting are recommended if a base load hot spots occur only rarely. Selecting a large window will however delay the detection of any increase in the base load. Because the algorithm is based on local minima in the window, such increases (e.g. following the event of a new multimedia device being added to the circuit) are only detected once the majority of minimal entries in the window correspond to the new situation, i.e. when the window has shifted almost entirely by its own length. Thus, $M$ more or less determines the time it takes to detect an increase in the base load. The detection of a base load decrease on the other hand does not directly depend on $M$, because such a situation will insert lower minimal values into the window, which immediately get recognized at the expense of the old ones.

The number of smallest values used, $L$ can also be varied. The optimal value for $L$ depends on the desired output. If the aim is to find the minimal base load level (i.e. the minimal interval median), then only one interval is required, so a small value of $L$ suffices. If one attempts to return multiple base load levels however, then $L$ should be set high enough so that for each of the target intervals at least one of the $p_i$’s lies within that interval. The best value for $L$ may be difficult to determine in this case. As for upper and lower bounds; the majority of all $L$ smallest load values from the window should originate from base load hot spots for the algorithm to perform as intended. When set too high, non-base load values will be considered, whereas too low a value for $L$ may result in the majority of $p_1, \ldots, p_L$ being all spike values. A well-fitting value for $L$ is best determined through a series of experiments.

(47) The recurrent value method has been tested on a household with PV panels, see Figure 37 on page 107.
The adaptive thresholding step is an important feature of the method. The sizes of the initial and raising thresholds determine the number of iterative steps up to reaching the final threshold value. Smaller values increase precision, but also increase the clock speed of the algorithm. The tolerance parameter for the desired interval length $T_{tol}$ should be set at at least the chosen median filter’s window range, since smaller strings of noise values have already been tackled in the pre-processing step. Larger interval settings allow the method to estimate the median value of a base load hot spot with higher precision, but a shorter interval length requirement is more suited in situations where base load hot spots are rare. Pointers for setting $T_{tol}$ adaptively are given in the performance improvement section below.

**Test results**

The above method was implemented in Matlab (see Appendix 8.2) and tested on a sample from House 2 from the REDD database (Kolter & Johnson, 2011). Half a day of data was used, running from 12:00:00 PM on 18th April, 2011 until 11:59:59 PM the same day. The data was preprocessed with a median filter with range parameter $R = 25$. The window size $M$ was chosen to cover the entire sample and $L$ was set as $0.01 \cdot M$ rounded above. Thresholds were set at $T_0 = 5$ W and $T_{raise} = 2.5$ W and 70% of the intervals was required to meet the desired interval length which was set at 30 consecutive samples. The algorithm returned a base load value of 22.23 W after one iteration. Figure 36 on the next page shows a plot of the results.

**Expected performance & possible improvements**

Various improvements can be implemented to the above method to increase both its speed and performance. For instance, if after a number of iterations two intervals corresponding to minima $p_{i1}$ and $p_{i2}$ overlap, they could be merged into one instead of running all further computations twice. The final interval could be given a weight corresponding to the number of $p$’s it has absorbed. Another speed increase could be realized by removing some of the higher load values from the data prior to starting the iteration procedure, since the algorithm eventually restricts to the lower-valued regions of the data anyway.

Also additional auto-adjusting with respect to some of the method’s parameters could be realized by setting a cap on the threshold $T_{k}$. A relatively high initial interval requirement $T_{tol}$ could be set for example, which is reduced each time $T_{k}$ exceeds a certain value. This has two advantages: it allows the algorithm to adjust to different situations by itself, and it prevents the threshold to get to undesirably high values. However, it also increases the number of iterations and therefore reduces the speed.

Lastly, some safety measures could be built in, ordering the algorithm to re-run the estimation with different parameter settings if a previous run came up with an unlikely value. One way to achieve this is by setting up a range in which the base load value is likely to be, based on the consumer’s average consumption. Another possibility is to constantly monitor for any large changes in the base load quantity. For instance, if the estimated base load suddenly doubles, it is likely the chosen parameter settings did not fit to the situation.

**Recurrent value method**

The second technique estimates the base load as the lowest regularly recurrent value in the chain of all load or consumption values. It is discussed here as a P4-based method, but can naturally also be used in conjunction with P1 data as an alternative or supplement to the flat
**Figure 36:** Results from a test of the flat level method on a sample from House 2 from the REDD database.
level technique. The philosophy is that each time the base load dominates total consumption for an entire fifteen-minute metering window, the consumption recorded for that window will be close to the same constant value.

**SETTING**
The input of the proposed technique is a time series $t = (t_1, \ldots, t_N)$ and corresponding consumption data $C = (C_1, \ldots, C_N)$. The proposed method for estimating standby power presumes the following conditions:

<table>
<thead>
<tr>
<th>6.3 Assumption:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A household’s base load is a steady quantity.</td>
</tr>
<tr>
<td>2. The base load dominates total 15-minute consumption in a residence on a regular basis.</td>
</tr>
</tbody>
</table>

The reader should note that Assumption 6.3 is a less stringent version of Assumption 6.2. However, when applying the recurrent value method to P4 data, the above assumption is more stringent than Assumption 6.2 was for P1 data. It then implies base load hot spots covering an entire 15-minute metering window regularly recur, i.e.: there exist at least a collection of load samples $C_{i_1}, C_{i_2}, \ldots, C_{i_n}$ in $C$ which are all almost equal to the base load-associated consumption $C_{\text{base}}$, i.e.

$$|C_{ij} - C_{\text{base}}| < T_{\text{tol}}$$  (4)

for $j = 0, 1, \ldots, n$ and some small tolerance $T_{\text{tol}}$.

**METHOD DESCRIPTION**
As with the flat level method, the technique starts of by extracting the $M$ most recent measurements from $C$ and $t$, where $M$ is a chosen window parameter. The extracted consumption values are grouped by size using a clustering algorithm, see Section 6.2.4. As many different combinations of appliances are possible and the exact composition may vary within a 15 minute window, there will probably be a large number of unique values, which make up small-sized clusters. However, consumption measurements corresponding to a 15 minute window in which only the base load drives up consumption are expected to recur regularly in the setting determined by Assumption 6.3. Discarding all clusters with less elements than a certain tolerance $T_{\text{tol}}$, the base load is estimated as the lowest mean or median cluster value among all remaining clusters.

**TECHNICAL DISCUSSION**
A large part of the technical discussion on the flat level method is applicable to the recurrent value technique, most notably the remarks on assumptions, data corrections and safety measures. The reader should however keep in mind the setting sketched for the recurrent value method is more restrictive than the one given in Assumption 6.2, thus it is advised to stack some insurance mechanisms when deploying the recurrent value method in P4 mode.

**Choice of parameters**
Two parameters can be set for this method, namely the window size $M$ and the tolerance $T_{\text{tol}}$. 

105
for the minimal cluster size. Additionally, some clustering techniques (such as the k-means algorithm deployed in the test case below) require the number of clusters $k$ to be pre-defined. The considerations regarding the choice of $M$ are similar to those for the flat level method, so the reader is redirected to the discussion above. The choice for $T_{tol}$ is somewhat dependent on the number of clusters. A larger number of clusters generally means there are less elements in each cluster, thus a smaller value for $T_{tol}$ is recommended for higher values of $k$. In turn, more clusters can help capturing different base load values in case there are more than one. The test run discussed hereafter shows the results for two different choices of $k$.

**Test results**
The method has been tested on a P4 dataset from a four-person household with solar panels. One week of data running from 00:00 AM February 5th 2016 until 00:00 AM February 12th 2016 was extracted (672 measurements). During this week, sunrise took place between 8:00 - 8:10 AM and sunset between 17:31 - 17:42 PM. Therefore, any measurements between 8:00 AM and 17:45 PM were disabled in all calculations to make sure no consumption was masked by any electricity generation from the PV installation. On the remaining data points, Matlab’s $k$-means clustering algorithm was run twice, once with $k = 10$ and once with $k = 20$ clusters. The lowest cluster median values were 39.3 Wh for the $k = 10$ run and 22.2 Wh for the $k = 20$ run. Both runs had additional clusters with median values below 75 Wh: for the $k = 10$ run there was one with a median value of 62.1 Wh and for the $k = 20$ version there were three, with median values of 33.3 Wh, 44.2 Wh and 59.8 Wh. Figure 37 (page 107) shows what these numbers represent regarding the last three days of the data.

Figure 37 shows the method performs reasonably well in terms of finding low steady values. However, in the current example there seem to be multiple steady levels during idle hours, of which a few seemingly include intermittent load values from cyclic appliances. Although the $k = 20$ run did manage to capture them all, it will be difficult in practice to automatically decide what value to return as output.

**Energy conservation**
As established above, electricity consumption by appliances in standby and low-power mode has a significant share in total electricity use: about 10% on average. As already mentioned however, it is often unclear what is included in this category and what is not, making it difficult to estimate the effect of reductions within this category on total energy use. A study by Parker et al. (2006) showed the implementation of a feedback system which measured energy usage of the category of standby mode appliances yielded a 90 W consumption reduction within this category. If indeed these Watts were being drawn continuously, all year long prior to the feedback, then this would mean a staggering 788 kWh consumption reduction on an annual basis (most households consume 1500-3500 kWh of electricity annually, see Section 3.1. Despite from the large uncertainty revolving around the effect of base load reductions, it has been affirmed by many that a large part of standby and low-power mode consumption can be easily avoided without affecting the quality of life (e.g. Parker et al., 2006; A. Meier, Lin, Liu, & Li, 2004).

**6.3.5 Idle consumption**
Even if the base load cannot be estimated anyhow, the mean idle load can still be determined in many cases. Since the mean idle load seems to be more valuable for organizations than for
Figure 37: Test results of the recurrent value method used on a household with PV panels. Shows two separate runs featuring $k = 10$ and $k = 20$ clusters.
most householders, it will be discussed from a business perspective. Nevertheless, most of the
discussion is applicable to residential situations as well.

Especially for organizations which are closed at least one day a week, the mean idle load provides
an estimate of the average electricity usage outside of office hours. Since such ‘weekend con-
sumption’ usually is not directly related to any labour conducted on behalf of the organization,
interesting cost reductions could be induced by avoiding a portion of it could theoretically be
avoided without affecting performance. Two methods will be discussed: one which requires the
customer to provide information of the general opening hours of the firm, and one automatized
technique.

**Manual method**
Computing the idle consumption is only a matter of summing the consumption measurements
during idle hours. The difficulty lies in delineating the moments which are considered idle hours.
One way to do this is by requesting the customer firm to provide a table with times at which
the office normally opens and closes. However, when blindly using these times, occasionally
overworking employees can throw the calculation off course. To overcome this, one could compute
the idle load for a sample of the data and then extrapolate to the remaining idle hours. If $C_1, \ldots, C_N$ are the consumption values outside of office hours, the sample idle load can be
assessed by choosing some window length parameter $M$, and setting the sample idle consumption
$\tilde{C}_{\text{idle}}$ as the minimum amongst all strings of consumption values of length $M$.

$$\tilde{C}_{\text{idle}} = \min\{C_{1,M}, \ldots, C_{N-M+1,N}\},$$

where in accordance with the notation from Chapter 2, $C_{i,j}$ equals consumption between time
stamps $t_i$ and $t_j$. When applying the above procedure, it is important to choose $M$ carefully.
When set too low (i.e. when the sample only contains part of one on-off cycle of most periodic
devices), the above minimum corresponds to a large number of individual cyclic devices being in
their off-cycle and will return an underestimate of the real idle consumption. As $M$ is increased,
samples of length $M$ start to include more periods and will be less determined by the state of
individual appliances. However, setting $M$ too high increases the probability of windows being
contaminated by consumption corresponding to activity in the firm.

**Automatized method**
As an automatized alternative to the above, one could deploy the flat level method from Section
6.3.4 with a different parameter configuration. The philosophy is that relative to an organiza-
tion’s energy intensive day-to-day operations, load fluctuations during idle hours are small:

**6.4 Assumption:** Compared to electricity usage during office hours, a firm’s idle hour
consumption is small and relatively constant.

Whether Assumption 6.4 holds depends primarily on the size and energy usage of the firm. In
‘zoom-out mode’, i.e. using coarser parameter settings, the flat level method can be used to find
the mean load as a stable level. This encompasses a long window, higher threshold settings and
a long interval requirement. Also, outliers are relevant to the mean idle load, so in this case the
mean rather than the median should be taken over each of the intervals.
TECHNICAL DISCUSSION

Data corrections
If a firm has its own PV installation, this will affect the idle consumption values as measured by the smart meter. In this case, one may attempt to retrieve the generated amount of electricity, or restrict to nighttime consumption values and then extrapolate the results to the missing daytime hours.

Test results
The flat load algorithm was tested on a five-day P4 sample from an medium-sized firm with a clear Monday-to-Friday work week structure. The window again covered the entire sample, but only $L = 3$ minimal values were considered. Thresholds were set at $T_0 = 20$ and $T_{raise} = 10$ and the minimal interval requirement $T_{tol}$ was set at 160 samples. After four iterations, Matlab returned a mean idle quarter-hourly consumption value of 35.59 kWh, see Figure 38 below.

Expected performance & possible improvements
The automatized method is expected to be a solid way of determining idle consumption for firms, as long as there are periods during which the firm is not operative. As already mentioned, the above technique can also be applied to households, which may prove useful if they do not meet the criteria for applying any of the base load estimation methods.

ENERGY CONSERVATION
The idle consumption is expected to be one of the most effective feedback mechanisms when it comes to organizations, since it concerns energy consumption that for a large part does not affect a firm’s active operations. In the end, many firms are not operational during a significant amount of time (weekends, nights, holidays), thus any energy reductions during these hours automatically induce large net energy savings especially for larger organizations. To demonstrate this, consider the above test case shown in Figure 38 and for the sake of argument say that this firm is operational 5/7th of the time. Based on the idle load test results, this firm consumes about 356 MWh of electricity annually during idle hours.\(^{48}\) A 10\% reduction of idle consumption would thus mean net savings of 35.6 MWh a year, which equals the consumption of almost 12 households consuming 3000 kWh of electricity on a yearly basis. Naturally, energy savings may differ greatly between individual firms, depending on their size, running time and efficiency during idle hours.

6.4 Change detection
When smart meter measurements are collected over a longer period, certain patterns become apparent. Using auto-comparison techniques, one can determine whether real-time data significantly deviates from found standards. Shifts in electricity consumption can arise for instance when a consumer adopts a new device or replaces an old one, but can also be due to changes in the family situation (e.g. when a child is born or when the employment status of a family member changes). A sudden increment may however also be an indication of electrical leakage caused by a defect somewhere in the circuit.

The unwanted occurrence of an electric potential between two objects is termed stray voltage and often originates from faulty wiring or insulation (Appleman & Gustafson, 1985). It poses a

\(^{48}\) The fact this seems to be a ten-thousand-fold of quarter-hourly consumption is a coincidence. The number has been computed by multiplying quarter-hourly consumption successively by $4 \cdot 24 \cdot 365 \cdot 2/7 \approx 10,011$.  

109
Figure 38: Test results from the flat level method applied to a five-day P4 data sample from a medium-sized company.
fire hazard (American Petroleum Institute, 1956), may cause dangerous electric shocks (Zipse, 1999), can damage sensitive assets, lead to power outages and result in serious wastage of energy (Brennan, 1993; Lefcourt, 1991). Additionally, numerous studies have shown stray voltage problems on farms affect the health and productivity of dairy animals (e.g. Appleman & Gustafson, 1985; Rigalma et al., 2010). Potential damage could be greatly reduced if the issue is detected in time, but this can be difficult in practice.

Three change detection techniques based on smart meter data are proposed in this section, two of which are specifically designed to detect increments. The cause of a detected shift cannot be traced with these methods, yet it may still be of interest to customers to be notified about any increases in consumption, to avoid any unpleasant surprises at the end of the month. All methods require a learning phase in order to determine consumption patterns used for real-time monitoring later. The techniques operate on different time spans and can be deployed simultaneously for maximal performance (see Section 6.4.4).

### 6.4.1 Edge clustering

The first method is based on the detection of rising edges, which is derived from the energy disaggregation problem discussed in Section 6.2. Of the three increment detection techniques treated here, this one has the fastest response time: once the learning period has been completed, it can locate abnormalities instantaneously. It requires real-time power measurements and therefore can only be operated using P1 data.

**Method description**

As introduced in Section 6.2, a **rising edge** refers to a sudden increase in the main load, which generally marks the event of an appliance turning on or shifting to a more energy-intensive state (Zhao et al., 2016). Detecting rising edges can be achieved by comparing the difference between power readings from a smart meter at subsequent time stamps \(t_i - t_{i-1}\) against a chosen detection threshold \(T_{det}\), i.e. an increment is detected if

\[
P_i - P_{i-1} > T_{det}.
\]

If a rising edge occurs, the difference \(E_i = P_i - P_{i-1}\) is stored. However, instead of matching the rising edges to falling edges as done in edge-based NILM techniques, the rising edges itself are clustered here. During the learning phase, a database \(E = (E_1, \ldots, E_J)\) consisting of rising edge magnitudes is formed. The entries in \(E\) can be partitioned into a number of clusters (which correspond to the appliances that triggered the increment) using a clustering algorithm.\(^{49}\)

After this step a number of clusters \(K_1, \ldots, K_L\) consisting of rising edges of similar magnitude are obtained. Once enough data has been collected, the distribution of samples within each cluster can be assessed. Based on these distributions, an area on the continuum can be assigned to each cluster which contains a certain portion of the cluster’s data points. For example, if the values in a cluster follow a normal distribution, about 95% of the points are expected to lie in so-called **two-sigma intervals**, that is within two standard deviation from the mean. The

\(^{49}\)In this situation, a so-called (normal) mixture model may be the preferred clustering tool, as will be explained in the technical discussion below.
procedure has been sketched in Figure 39, where edge magnitudes in each cluster have been assumed to follow a normal distribution.

![Figure 39: An example depiction of two-sigma intervals for L clusters on the power spectrum based on the probability distribution for each cluster.](image)

Once the database has been constructed, real-time operations of the method consist of checking whether newly incoming rising edges belong to any of the clusters, by verifying whether the power difference lies in one of the above intervals. If it is identified as a member of a cluster, the rising edge may either be added to the cluster or discarded (depending on the data storage capacity). If at some point an ‘unknown’ increments pops up, it could be an indication of a new appliance or an electrical leak.

**TECHNICAL DISCUSSION**

*Choice of parameters*

The edge-based detection method involves a threshold parameter $T_{\text{det}}$ which should be set high enough to avoid detection of measurement noise, but low enough to be able to capture increments of smaller magnitude. One could choose a higher value to have the method only target appliances that draw larger amounts of power, as this is expected to increase the method’s performance (see also the improvements section below). Additionally, the clustering algorithm and any probability models may introduce a number of parameters, but these depend on the method and will not be discussed.

*Data corrections*

As stated, the edge clustering method is a PI-based technique. The data should be pre-processed using a median filter to dispose of noise and voltage spikes (see Section 5.1). The technique does not require any seasonal corrections, but season-specific appliances (e.g. air conditioning units or a heat pumps) can be missed if the learning phase is executed in the opposite season. In this case, such devices may be detected as new appliances once they are frequently operated.

As for households with PV panels, it may not be necessary to back-trace generation amounts for this technique. As long as the amount of electricity generated by the PV module does not fluctuate heavily within a second, no significant falling edges\(^{(50)}\) are produced by the installation in the data. Under most weather conditions, the output of a PV module is not expected to regularly show any large shifts within a one-second window, unless maybe the installation itself

\(^{(50)}\)Note that a falling edge in electricity production produces a rising edge in the net flow.
is very large (in which case normal fluctuations may already be substantial). If this assumption turns out to be false however and large sudden drops in generation do occur often, the PV estimation methods discussed in Section 5.2 probably lack the precision to accurately pinpoint such fluctuations, thus additional sensing equipment would be required. Alternatively, one could try restricting to rising edges which recur regularly, as a falling edge produced by a PV module is expected to be rather random in magnitude.

**Expected performance & possible improvements**

The edge clustering technique is fast, but has a few drawbacks. The inrush power signal from a new appliance or an electrical leak can only be detected if it does not overlap with the ranges of an existing appliance. If a household’s collection of electrical appliances is large, or if certain appliances attain a variety of loads, the power spectrum can get cluttered by a proliferation of clusters or by clusters with a very wide span (i.e. clusters with a high standard deviation).\(^{(51)}\) In such cases, no new appliances can be detected anymore by the above method. To narrow the intervals in such cases, the thresholds could be set more stringent, or the accuracy of the intervals could be lowered, but this increases the probability of a false detection.

Alternatively, if reactive power is measured by a smart meter, the clustering can be performed in two dimensions. Using the time stamp of each measurement, real and reactive power differences can be linked and the intervals in Figure 39 are replaced by areas in the plane. This may greatly increase the “empty space” in between the clusters, and therefore is expected to yield a significant performance improvement.

Also, the performance of the method can be improved by restricting to appliances which are regularly used. By introducing another threshold \(T_{\text{tol}}\), clusters which at the end of the learning phase contain less elements than \(T_{\text{tol}}\), are removed. This reduces the number of clusters and therefore increases the probability of new events to be picked up. However, in such a case it is advised to apply the same tolerance requirement also to the detection side, i.e. a new event should only be reported after it has been detected numerous times. This might also be a solution to the issue of sudden large changes in solar PV output interfering with the method: assuming such PV-related edges are of random magnitude, they will not be clustered as an appliance, and therefore will be neglected by the algorithm. In relation to electrical failures, a continuous defect or appliance will no longer be detected (because it stabilizes after one rising edge), whereas (failures in) regularly recurrent devices have a higher chance of being recognized.

By extension, one could think of performing additional frequency analyses on the clusters. By checking the time stamps of each of the edges, one can derive how often a certain edge appears (which would relate to how often the corresponding appliance is used) and assess how this number evolves over time. This also holds some prospects for expanding the method to the detection of decreases in consumption.

Finally, one should keep in mind that the distribution of a cluster may be complex, even if it turns out the inrush signals of all appliances are neatly normally distributed. Suppose for instance, the rising edges produced by two appliances are normally distributed with mean values \(\mu_1 = 180\, \text{W}, \mu_2 = 195\, \text{W}\) and standard deviations \(\sigma_1 = \sigma_2 = 10\, \text{W}\). Then the two distributions are intertwined and it is likely that the rising edges from both appliances end up in the same

\(^{(51)}\) The reader could picture this as the power spectrum in Figure 39 being (almost) entirely covered by intervals.
cluster. If this happens, the cluster as a whole will follow a *Gaussian-mixture-distribution* which is more difficult to cope with than a normal distribution (Nasrabadi, 2007). A tool for executing the clustering step in such situations is to deploy a *mixture model*, which may be able to identify the underlying structures. There are a number of texts on this subject the reader can consult (e.g. Bilmes, 1998; Nasrabadi, 2007).

### 6.4.2 Peak consumption saturation

The second technique mimics an argument sometimes deployed by climate scientists to support the theory of global warming, namely that the observed increase in the number of heat records as of late is induced by globally rising temperatures (Rahmstorf & Coumou, 2011). Similar logic may hold true in the setting of domestic energy consumption: if a household’s electricity usage increases for some reason, it is likely their peak consumption (as defined in Section 6.3.2) attains higher peaks as well.

**Setting**

The technique builds around the following assertion:

| 6.5 Assumption: | The appliances in a building jointly define a *peak consumption saturation level*, a theoretical soft ceiling which is only exceeded in situations where implausible combinations of appliances are simultaneously used. |

This saturation level has not been found in literature nor has it been substantiated with rigorous empirical evidence, which makes its existence as well as the functioning of the technique below somewhat doubtful. Yet, it is not unlikely maximal consumption accumulates at some point because a household consists of a limited number of people, who each can only use a few appliances at the same time. Some indications of the existence of a peak saturation level are given in the technical discussion section below, see Figure 41. It should be noted the saturation level is not a hard limit: it can be exceeded, but it is assumed this rarely happens. This makes it a rather ill-defined concept from a mathematical point of view. One could go for a more rigorous notion by replacing the saturation level by a continuous saturation range, which describes the probability of consumption exceeding various levels. Such a range is difficult to determine though, and since the technique does not use any numerical features of the saturation aspect (only that maximal consumption stagnates to some extent), this will not be practiced here.

**Method description**

The technique uses time frames of a fixed duration $D$. During the training phase of the method, the consumption data is divided into such intervals, and the consumption $C_i$ during each interval is computed. The $N$ highest consumption values are stored as a top $N$ ranking $\mathbf{R} = (R_1, \ldots, R_N)$ where $R_i \geq R_{i+1}$ for all $1 \leq i \leq N$. The ranking is continuously updated, i.e. if a new consumption measurement is higher than $R_N$, it is included at the proper position in $\mathbf{R}$ at the cost of $R_N$. As the peak consumption saturation level is approached by more and more entries in $\mathbf{R}$, the event of new measurements infiltrating $\mathbf{R}$ should become increasingly rare. The training phase ends once this happens less often than a certain tolerance $T_{tol}$. If at any subsequent point in time the frequency at which peak measurements enter $\mathbf{R}$ suddenly ramps up (i.e. if a detection threshold
$T_{\text{det}}$ is exceeded), this could be an indication of a saturation shift caused by a new appliance or an electrical leak, see Figure 40.

![Schematical depiction of the peak consumption saturation method.](image)

**Figure 40:** Schematical depiction of the peak consumption saturation method.

**TECHNICAL DISCUSSION**

**Choice of parameters**

A number of parameters can be adjusted, namely the interval duration $D$, the length $N$ of the ranking, the tolerance threshold $T_{\text{tol}}$ and the detection threshold $T_{\text{det}}$. A short interval length allows for quick data acquisition and therefore shorter learning periods, and also has a higher chance of capturing instances at which the peak saturation level is approached (i.e. when there is a maximal overlapping of appliances). Some appliances such as a blender or a microwave, are generally used briefly. Hence, using shorter intervals, the likeliness of capturing the electricity consumption of those appliances is higher. Too short intervals pose the risk of the consumption values being influenced by noise and voltage surges (which sometimes survive filtering endeavours).

Higher values of $N$ lengthen the learning time. Peak values are only attained a few times a day, so data accumulation is not expected to go swiftly. Hence if $N$ is set high, it will take long to fill the ranking. Also, high values of $N$ can create ‘vulnerable’ entries in the lower sections of the ranking, i.e. values which are easily surpassed by others. As a result it may take long for the ranking to stabilize. On the other hand, too short rankings might get filled up with occasional extremes. At this point in time, it is difficult to give a good range for $N$, as this depends largely on the stability of the saturation level. The same goes for the tolerance and detection parameters, optimal values for which should be determined in a testing procedure.

This method can be used both with P1 and P4 data and the exact choice of parameters depends on the type of data used. When P1 data is inserted, short time frames can be used, which enable much faster data accumulation and a better approximation of the saturation level. The method can also be used with P4 data, but this necessitates shorter rankings, and even then the learning phase may take long (see Figure 41). In general, the technique is expected to perform better and faster with P1 data and provide greater flexibility with regard to the choices of $N, T_{\text{tol}}$ and $T_{\text{det}}$. However, when using P1 data, particular attention should be devoted to the detection phase. Whenever at some point in time a peak consumption value is higher than the lowest
value in $R$, this will most likely also be the case for some subsequent values as situations do not change every second. This may generate long strings of the same value, which may completely fill the ranking. Hence, either these subsequent values should be ignored, or a very long ranking should be used.

**Data corrections**

If the method is used with P1 data, the data should be pre-processed with a median filter, to avoid large spikes influencing peak consumption levels. Also, consumption shifts with the seasons, so it is likely the peak saturation level will also be influenced by seasonal differences. Therefore, it is recommended to apply some type of normalization, or otherwise the saturation level may increase during winter and the algorithm may fail to detect anything in summer. If the seasonal influence on the above technique turns out to be small compared to a ‘real’ shift in consumption, one could decide to leave the data and relax the detection threshold $T_{det}$ somewhat to have the technique only aim at larger increases. When it comes to PV generation, one may either attempt to account for the amount of electricity generated, or the method can be restricted to nighttime consumption (as was done in Section 6.3.4). This greatly limits the performance of the method during the summer half-year though.

**Test results**

The peak saturation technique has been tested on the P4 electricity data from 2016 of the three households A,B and C used previously for testing efforts in Section 5.3 (see Figure 22). The above method scheme was implemented in Matlab (see Appendix 8.2), the output of the program being a *cumulative update number*, a number initiated at 0, which is increased by 1 every time the ranking of highest quarter-hourly consumption entries is penetrated by a new value. The results of a test run with $R = 15$ are shown in Figure 41 below.

![Figure 41: Test results of the peak saturation technique applied to three unprocessed consumption signals A,B and C previously used in Figure 22.](image-url)
One may observe that after about half a year, the cumulative update number stabilizes for all three signals during the summer months. However, all three signals also show a slow reincrease during autumn, possibly caused by the fact that the data was not normalized. The developments for signal B in the final weeks of the year are particularly interesting as they show an increase in consumption somewhere mid-December. This could be a detected increment, possibly the result of inefficient Christmas lighting.

**Expected performance & possible improvements**

A benefit of this technique is that the ranking feature allows the detection of relatively subtle increments. There are some drawbacks as well though. First of all, the method is limited to detecting increments during peak time only. Thus, even if a new appliance is added, it will only be detected if it is used regularly at peak times. Nonetheless, any significant electrical leakage due to a failure in the general circuit or in an appliance which is regularly on during peak times,\(^{(52)}\) can be detected.

Furthermore it should be noted that if the algorithm is at some point successful and there is indeed a shift in the peak consumption saturation level, the algorithm needs to undergo re-training to adjust to the new situation. A re-learning program should also be initiated if the customer manages to cut his peak consumption, otherwise the algorithm might never be able to detect anything anymore. Detection of such reductions can be realized by long-term techniques in Section 6.4.3, which are able to measure decreases as well. Finally, the learning period should be proceeded with care. For instance if the customer goes on vacation, no more peaks will be recorded and the learning period might be wrongfully terminated. This too, can be organized by cooperating with long-term detection methods, but can also be coordinated with the customer.

### 6.4.3 Long-term detection

Shifts in consumption can also be recognized by looking at trends in smart metering measurements acquired over a longer period. In this section, a technique will be described for analyzing the long-term behaviour of a customer’s total daily consumption. With some minor adjustments, a similar solution can be developed for different aspects of a customer’s consumption data, such as their peak consumption, base load or idle consumption. Since the underlying framework is by and large the same for all of these cases, a detailed description of the technique will be given here for total daily consumption only. Pointers as to what is different in the other situations will be handed to the reader in the technical discussion afterwards.

**Method description**

The long-term method uses the average daily consumption as a benchmark to compare future consumption values to. There may be outliers towards both lower and higher consumption values and since the ‘outlying potential’ of the higher entries is much larger than for lower ones (i.e. there is hardly a cap on daily consumption, whereas the base load more or less makes up a lower bound), the median daily consumption is expected to be a better representation of the average than the mean. Hence, the sample median will be used in this section as benchmark value for the average and will be denoted by \(\mu\).

During the learning period, the \(n\)th cumulative median \(\mu_n\) is computed over all consumption values \(c_{day}^{1}, \ldots, c_{day}^{n}\) available at that point in time. Thus, with each day, the median is computed

\(^{(52)}\)The reader should note this applies to all continuous and most standby mode devices
over more consumption values. As \( n \) grows, \( \mu_n \) is expected to stabilize, and the learning period is terminated on day \( N \) if the difference between the highest and lowest cumulative median value over the last \( M \) days drops below a certain tolerance threshold \( T_{\text{tol}} \), i.e.

\[
\left( \max_{n=N-M+1,\ldots,N} \mu_n \right) - \left( \min_{n=N-M+1,\ldots,N} \mu_n \right) < T_{\text{tol}}.
\]

When this happens, the median over the entire learning period \( \mu = \mu_N \) is fixed as the benchmark average. From this point on, incoming daily consumption values are compared to this average to assess whether there is a structural increase or decrease in consumption. To do this, for each daily entry \( C_{\text{day}}^i \) the deviation

\[
\delta_i = C_{\text{day}}^i - \mu
\]

is computed. The sign of \( \delta_i \) indicates whether a value is above- (positive sign) or below-average (negative sign). By averaging\(^{(53)}\) over \( N \) subsequent \( \delta_i \)'s one can assess how consumption develops on the long term, i.e.

\[
\Delta_i^{(N)} = \frac{1}{N} \sum_{j=i-N+1}^{i} \delta_j
\]

is a measure of total deviation over a period of \( N \) days. However, in some situations this sum can be largely determined by one spike value, e.g. one large outlier value \( C_{\text{day}}^i \) may on its own compensate for a smaller structural difference in the opposite direction and ‘fool’ the algorithm into thinking there is no change going on. To reduce the influence of outliers one can replace the above formula for \( \Delta_i^{(N)} \) by

\[
\Delta_i^{(N)} = \frac{1}{N} \sum_{j=i-N+1}^{i} \sigma(\delta_j)
\]

where \( \sigma \) is a *levelling function* which leaves smaller entries untouched, while reducing the magnitude of higher \( \delta_i \)'s. A possible choice for \( \sigma \) could be (a scaled version of) the arctangent function measured in radians, as depicted in Figure 42\(^{(54)}\). The arctangent function fits the profile: it maps small values more or less to itself, yet its output stops growing when extremer values are inserted and asymptotically approaches \( \pm \pi/2 \approx \pm 1.57 \). Ergo, when applying this function to each \( \delta_i \), a very large entry will still add more weight than a smaller one, but the effect is reduced and a weight can never grow beyond \( \pm \pi/2 \). When doing this in practice, it is recommended

\(^{(53)}\) As announced already in the methodology chapter, in this section a weighted averaging technique will be illustrated in favour of the median that is used for taking averages in the rest of this text. When using this weighted averaging method, one gains control over the effect of outlier values on the average, as will be shown. Translated to the current setting, the depicted method will allow the user to choose the amount of influence a day of very high or very low consumption has on the measure of deviation \( \Delta_i^{(N)} \). Naturally, if one does not want to pay any special attention to outliers, the weighted average may be replaced by a median and all further digressions on weighted averaging in this section can be ignored.

\(^{(54)}\) S. Kim and Kim (2016) have applied a similar technique in a different setting.
the arctangent function is scaled to be compatible with a particular dataset, details for this are given in the technical discussion.

Figure 42: A plot of the arctangent function (with x-axis values in radians).

The $\Delta_i^{(N)}$'s can be used to assess changes in consumption at different scales. The idea is that the median was set subject to a certain stability condition. While there may be fluctuation among the individual $\delta_i$ values, averaging over larger numbers of $\delta_i$'s should yield a relatively small number, as long as nothing changes within the household. If for some reason consumption is increased or decreased, the balance between positive and negative $\delta_i$'s is disrupted, leading to $\Delta_i^{(N)}$ values leaning towards something significantly either positive or negative. A change is detected if

$$|\Delta_i^{(N)}| > T_{\text{det}},$$

where $T_{\text{det}}$ is a detection threshold. This threshold should be set based on the choice of $N$, see the technical discussion below. The method is sketched in Figure 43.

Figure 43: Schematical example depiction of the long-term change detection method.
TECHNICAL DISCUSSION

Choice of parameters
A number of parameters need to be set for this method. First of all, the combo of an \( M \)-day time frame and the tolerance threshold \( T_{\text{tol}} \) that jointly decide when to terminate the learning phase. Since it is of vital importance for this technique that the final median value is a good representation of the average consumption, these should be set rather stringent.

A second set of parameters, less explicitly mentioned in the above discussion is the scaling function \( \sigma \). While the arctangent function may be a solid pick for reducing the influence of outliers, the way it affects the data in its neutral form, may not be desirable an may moreover be very different when applied to different sets of data. Consider for the sake of the argument an example household whose energy consumption is rather fluctuating, say most \( \delta_i \)'s are about 3 kWh, with outliers towards 9 kWh. In this case, the arc tangent will map all \( \delta_i \)'s within a relatively small range somewhere betwee 1.25 and 1.50, creating a relatively uniform output in which deviations due to outliers are hardly separable from ‘normal’ deviations. When on the contrary deviations are very small, different problems occur, in the sense that the arctangent will hardly affect the \( \delta_i \)'s. One can overcome both types of problems and ensure a higher degree of uniformity regarding the effect of the arctangent function on different datasets, by scaling it. To do this, set

\[
\sigma(x) = \arctan(c \cdot x),
\]

where \( c > 0 \) is a scaling parameter. When adjusting \( c \), the curve depicted in Figure 42 will either be compressed \( (c > 1) \) or be stretched \( (c < 1) \) along the \( x \)-axis. This parameter can therefore be used to draw the majority of the \( \delta_i \)'s into the ‘sweet spot’ of the arctangent’s domain, where it affects precisely the desired amount of data. It can be set based on for instance the deviation in \( C_i \)'s, which may be determined during the learning period. Too large \( \delta_i \)'s are then corrected by setting \( c \) somewhat smaller than 1 and smaller deviations can be corrected by a \( c > 1 \). A proper scaling protocol is to be determined through a series of tests.

The final parameters to set are the time scale \( N \) over which \( \Delta_i \) is computed and the detection threshold. The first actually does not need to be chosen, one may run the technique for different values of \( N \) simultaneously. For smaller values of \( N \), the technique can be used for medium term change detection, whereas for larger values of \( N \), \( \Delta_i^{(N)} \) hold information about any possible long-term changes. The choice of \( T_{\text{det}} \) depends on \( N \). For smaller values of \( N \), \( \Delta_i^{(N)} \) is more heavily influenced by individual \( \delta_i \)'s and will to some extend follow the natural fluctuations in consumption. Ergo, a higher detection threshold is required in order to ensure only significant changes are detected. As \( N \) grows, natural fluctuations are expected to cancel out, so small deviations of \( \Delta_i^{(N)} \) away from 0 can already be marked as significant. For large values of \( N \), the algorithm will target the more structural changes. Again, a proper formula for setting \( T_{\text{det}} \) based on \( N \) should be determined in a number of test runs.

Data corrections
When applying this technique to the data of a customer with their own PV installation, one needs to estimate the amount of generated electricity, because the technique uses daily measurements as inputs (i.e. on cannot restrict to nighttime consumption values). However, a rather rough estimation of total daily generation may already suffice here.
As for seasonal influences on consumption, one can apply a correction to the $\Delta_i^{(N)}$ figures instead of the data itself. When doing this, one should take into account there is some time delay in what $\Delta_i^{(N)}$ expresses, namely the variation over the past $N$ days. Therefore, one should correct accordingly (for instance by using the average deviation number from the last $N$ days. When going for long-running applications such as annual deviation figures, seasonal correction is no longer necessary in most cases, since such figures already capture the entire cycle and seasonal differences cancel one another.

**Test results**
The above procedure was tried on the three 2016 datasets A, B and C used for testing in the previous section. Using a seven-day window and a 0.25 kWh range, the cumulative median stabilized after 69, 59 and 35 days for samples A, B and C respectively. In real time it would take some time to denote this stability, which is why the learning time was terminated after 100 days to be on the safe side. From two weeks past this point on, biweekly levelled deviation $\Delta_i^{(14)}$ was computed, with $\sigma$ the (unscaled) arctangent function described above. The results are shown in Figure 44 below. Additionally, the deviation over the entire 266-day period was computed for each signal.

![Figure 44: Test results from the long-term method for three consumption signals A, B and C.](image)

As mentioned above already, each $\Delta_i^{(14)}$ value in Figure 44 holds information about consumption in the two week prior to the time stamp, i.e. there is some delay regarding the time it takes for any changes to become visible. Also, due to the averaging effort, each individual deviation value depends to large extent on the same consumption data as its predecessor. Hence, it may take some time until certain changes become apparent.

Having said that, a number of observations can be made from Figure 44. For starters, the graph
for signal A shows above-average consumption for the majority of the year, whereas signal B inhibits substandard electricity usage. Signal C fluctuates around zero, with only moderate peaks and a pretty clean balance between periods of positive and negative terms. The same conclusions can be drawn from the $\Delta (266)$ figures: the numbers for customers A and B are quite significantly positive and negative (seeing that $\pm 1.57$ is the maximum score, which would require tremendous over- and under-consumption during the entire period), while customer C’s total deviation is almost zero.

Also, a significant dip in consumption is notable for customer A’s data. A look at customer A’s original data showed there was a two-week period in the second half of September in which consumption was at a stable and very low level. Possibly the customer was on vacation during these two weeks. If the fourteen days were removed from the data, customer A’s total deviation would have been $\Delta^{(252)} = 0.443$. Arguably, a customer’s deviation is captured better without any such periods of absence. One could try to automatically remove these periods from the data. This will be discussed in the improvement section below.

Finally, one could wonder if the ‘December increase’ in customer B’s data as found by the peak saturation method in the previous section, is also detected by the current method. Indeed a rather steep increase in signal B’s $\Delta^{(14)}$ values in December can be observed from the test results, however taking the delay-effect into account, this change must have started at the end of the first week of December, while a significant change was only noted by the peak saturation method after the second week of the month. The author has no idea as to what may cause this discrepancy, and recommends a more thorough study towards the symbiosis between the different change detection methods described in this text.

**Expected performance & possible improvements**

The long-term method is expected to be a robust way of detecting changes, which is able to note very subtle changes, especially when deployed in annual mode. Although the method generally has a slower detection speed than the other techniques discussed here, it is very flexible in terms of the choice of detection time: one may trade detection accuracy for speed. Obviously multiple time windows can be run simultaneously, and the detection algorithm can be run with increasingly long time windows as time passes. Just as with the peak saturation method, this method requires re-training upon detecting a change in consumption.

As seen in the test section just now, it may be desirable to remove any vacation-like periods from the data. This can be achieved by applying the flat level method from Section 6.3.4 to the customer’s daily consumption data with a different parameter configuration. It may require somewhat higher threshold settings and a multi-day interval length requirement, depending on the vacation length that one tries to uncover. After a longer idle period has been detected, the corresponding measurements can be removed from the data.

Besides, this technique can be applied not only to total daily consumption, but also to daily peak, base and idle load values. For peak consumption, there is some overlap with the peak saturation technique discussed in the previous section. It acts on a different time scale and can be used as an addition. The majority of the discussion above can be copied to each of the different cases, albeit with some minor modifications here and there. For instance, since peak, base and idle consumption is much lower than the total daily electricity usage, one will
need different scaling parameters for the arctangent function. The base load and idle load being relatively stable, it may be better to remove scaling function altogether and just remove any significant outliers from the data (since for the base load these are most likely the result of a wrongful estimation). An additional advantage of applying the method to the base load is that it may not be influenced by any seasonal patterns (see Section 5.3) and that it is often assessed at night, thus obviating the need for any PV corrections.

6.4.4 Hybrid mode

The data requirements, strengths and weaknesses of each individual change detection technique discussed so far are bundled in Table 13 below. It can be seen there is little overlap between the edge detection, peak saturation and long-term techniques. Therefore, the methods can be operated in hybrid mode, strengthening one another and delivering a better overall performance. As such, methods can verify or disprove each others results. For instance if the edge clustering method registers an unknown edge, the other techniques can be used to check whether they also record the change. Such back-to-back operations allow for less stringent thresholds and can therefore be used to detect smaller changes. Also, additional information may be derived in some cases. For example if a new edge is detected and the peak saturation level shows a shift as well, but no significant increase in consumption is found on the long term, it could be that the shift is caused by a new appliance which is regularly used during peak times, but does not consume a lot of electricity altogether.

A disadvantage of the hybrid mode is that its speed is determined by the slowest method, which in this case is the long-term detection method. In general, all three techniques suffer from the fact that they require a learning period and operate by comparing to the situation they trained on. Hence, any electrical failures already occurring in a building cannot be detected. Also, the above methods do not provide any pointers as to the cause of an increase in consumption. Various externalities can trigger an increment, e.g., following the sudden unemployment of a household member, a child moving back in with their parents, etc.

Energy conservation

The author has not been able to find any scientific literature discussing the consequences, in terms of consumer reactions and potential savings, of implementation a change detection feature. Assuming the feature works as intended, one should note however that for consumers who do not change anything significant enough to be notable from a smart meter signal, the feature does not do anything at all. Hence for this group of users, there is no customer reaction, nor will there be any energy savings. Secondly, there may also be a group of customers who consider a change detection feature as an invasion of their privacy and reject any feedback from such features. Finally, there are some consumers who indeed may profit from a change detection feature, because their equipment starts leaking energy, or because they purchased a new device that turns out to use much more electricity than they expected. For this last group of customers, potential savings can be very large. At this point in time, it is unknown how many consumers fall in each of the above categories, making it difficult to estimate the overall impact of a change detection feature on energy use. All in all, it seems more likely this feature lies within the less effective regions of the spectrum of feedback applications discussed in this text.
### Table 13: Summary of the advantages and disadvantages of each of the detection techniques discussed in this section.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Data requirements</th>
<th>Advantages &amp; disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Edge detection</strong></td>
<td>P1 data</td>
<td>⊕ Fast detection speed.</td>
</tr>
<tr>
<td></td>
<td>Noise filter</td>
<td>⊖ Cannot detect appliances with overlapping load.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Does not detect subtle changes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Decreased performance for large number of devices.</td>
</tr>
<tr>
<td><strong>Peak saturation</strong></td>
<td>Noise filter (P1 only)</td>
<td>⊕ Can detect subtle increments.</td>
</tr>
<tr>
<td></td>
<td>Seasonal peak normalization</td>
<td>⊕ Can only detect changes occurring during peak time.</td>
</tr>
<tr>
<td></td>
<td>PV compensation</td>
<td>⊖ Needs retraining after every change.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Learning period can derail in some situations.</td>
</tr>
<tr>
<td><strong>Long-term total</strong></td>
<td>Seasonal normalization</td>
<td>⊕ Can detect both increases and decreases.</td>
</tr>
<tr>
<td></td>
<td>PV compensation</td>
<td>⊕ Can detect subtle changes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Needs retraining after change.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Slow detection speed.</td>
</tr>
<tr>
<td><strong>Long-term peak</strong></td>
<td>Seasonal peak normalization</td>
<td>⊕ Can detect both increases and decreases.</td>
</tr>
<tr>
<td></td>
<td>PV compensation</td>
<td>⊕ Can detect subtle changes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Needs retraining after every change.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Can only detect changes occurring during peak time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Slow detection speed.</td>
</tr>
<tr>
<td><strong>Long-term base/idle consumption</strong></td>
<td></td>
<td>⊕ Can detect both increases and decreases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊕ Detects subtle changes in base power.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Needs retraining after every change.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Possible issues with misdetected base load values.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⊖ Slow detection speed.</td>
</tr>
</tbody>
</table>

---

### 6.5 Benchmarking & comparison

In this section, a number of comparative features will be discussed to help place energy usage into perspective to help people increase their understanding of their own consumption. Consumption data of individual households will consecutively be compared to that of users with a similar profile, and to their own historical data. Goal-setting and saving tips will be discussed as means for self-incited energy conservation at the consumer end. For firms, an additional “energy per unit” benchmarking method will be proposed, which comes with promising applications.

#### 6.5.1 Comparative feedback

*Comparative feedback* is a notion which has been described in many studies on energy conservation (e.g. Ayres, Raseman, & Shih, 2013; Fischer, 2008; Siero, Bakker, Dekker, & van den Burg, 1996) and has been implemented in several energy managing systems (see Section 3.4.3). It entails the option for customers to compare their own consumption to that of others, which can be done on different time scales (days, weeks, months, years) and within a specific subgroup of customers (e.g. customers occupying a similar dwelling type, customers with an electric vehicle, etc.).

**Method**

One way to approach cross-consumer consumption comparison is by pin-pointing one’s position
on the ranking of all consumers sorted by the amount of energy they consume. The corresponding GUI output for a consumer could read something like: “In March, 37% percent of customers with a similar profile used less electricity than you.” Such a ranking can be created from the collection of all (anonymized) individual household consumption figures. However, if consumers are to choose custom time frames, or subgroups for this comparison, the list needs to be recalculated for every specific choice. Thus, full datasets of every consumer must be available at all times and large calculations need to be carried out over and over again.

To reduce data storage requirements and computational efforts, one could resort to only using a sample of the consumer population. Alternatively, the average energy consumption used in the comparison could be modelled using a probability distribution. Kuusela, Norros, Weiss, and Sorasalmi (2015) found electricity consumption follows a log-normal distribution, see Figure 45 below. In order to adjust this distribution to a specific group of customers, only the mean and standard deviation of consumption among the group’s members are required. Once the distribution has been scaled, it can be used to estimate a specific customer’s position as has been depicted in Figure 45. The area under the graph left from a specific customer’s consumption value relates to the proportion of households consuming less than the target (the total area under the graph equals 1 since it concerns a probability distribution).

The NEDU profiles discussed in Section 5.3 are a model of how the mean value is expected to evolve throughout the year, based on past records and the distribution of public holidays. A similar prognosis can be created for the standard deviation, either for the entire population or for a certain subgroup of choice. Using both these forecasts, the consumption distribution model displayed in Figure 45 can be extended to any other period in the year, allowing for comparisons on different time scales. Instead of full datasets of a large number of consumers, this method only requires a few parameters and the NEDU profiles to create the full model. Being probabilistic in nature, the method is also less accurate.

(55) As demonstrated in Section 5.3 it is unwise to apply a generic model such as a NEDU profile to adjust a specific consumer’s signal, since the latter need not follow onto the average trend. However, in the current setting the NEDU profiles are used to model how the mean consumption computed over its profile members behaves over time, which is a more reliable application.
**Technical discussion**

*Data corrections*

To make for a ‘fair’ comparison, one should attempt to create an equal level playing field for all consumers. As long as the period over which energy consumption is compared is the same for all participants in the comparison, it may not be necessary to apply any of the seasonal normalization procedures from Section 5.3. As electricity demand depends primarily on solar influx (see Section 5.2), which can be assumed to be more or less the same for all inhabitants of the Netherlands most of the year\(^{(56)}\), in which case normalization can be omitted. Gas consumption on the other hand, being so heavily dependent on outdoor temperatures, may require some more micro-management. However, instead of applying temperature corrections to all consumer’s gas consumption figures, one could restrict to comparing gas consumption of customers living in the same region to ensure a sufficient degree of homogeneity regarding ambient temperatures.

If a household owns any PV panels, comparing their electricity consumption to that of people who do not generate any electricity gives skewed results. One can either use one of the techniques from Section 5.2 to assess the amount of electricity generated by any individual systems, or subdivide customers into a number of groups based on the number of solar panels they own. Comparison between customers is then only performed within these groups (e.g. the aggregate consumption and generation by customers with four solar panels is compared to the average within the group of customers owning four solar panels). To account for regional differences in solar irradiance, either a correction factor could be applied based on weather reports (see Section 5.2), or comparison could be narrowed down further to customers with the same number of panels in the vicinity. For the latter, it is not clear whether there is enough data to realize a sufficiently supported comparison in each subgroup. Also, it is impossible to account for any mutual micro-climatic differences between customers.

*Expected performance & possible improvements*

It may be desirable to show cross-period comparisons for gas and electricity simultaneously, because a customer’s choices in appliances might offset the balance between electricity and gas consumption. For several household activities such as cooking and heating, there are appliances running on gas and appliances using power. At the moment, the majority of Dutch households uses gas for heating and cooking (Gerdes et al., 2017), so if a certain consumer uses an electrical solution for either one of these processes, the balance between their electricity and gas consumption will be different from that of the average consumer. As a result, the above comparison may return a very poor score in the electricity ranking and a very high one on the gas chart. When comparative feedback figures for gas and electricity accompany one another, this may be of aid to the customer in increasing their overall awareness about their energy consumption. Additionally, an aggregate ranking could be realized by weighting gas and electricity consumption somehow, for instance based on their caloric value, or costs.

Still there could be other situations in which consumption figures deviate from standards. If for example a customer uses wood for heating, their gas usage is likely to be much lower than

\(^{(56)}\)In June and December, there is a 30-minute difference in daily daylight time between the utmost northern and southern parts of the Netherlands due to latitudinal differences (Velds & Hoeven, 1992). Also, differences in sunniness may influence consumption from day to day, although the author has not been able to find any evidence of this. To be on the safe side, one could restrict all cross-consumer comparisons to consumers within the same region.
the national average without any sign of compensation in the electricity consumption figures. As discussed in Section 5.3, there may be some potential for subdividing customers into various categories based on the balance between their gas and electricity usage.

**Generalizations**

Apart from total consumption, different aspects of consumption can also be compared to other consumers. Possible examples include the peak and base load quantities, refrigerator consumption (see Section 6.2) and the susceptibility to temperature influences (as discussed in Section 5.3.4). In all cases, one can either go for a direct approach by repeatedly taking random samples from the database of customers, or via a modelling approach, as described above.

**ENERGY CONSERVATION**

Though comparative feedback features have been installed in a considerable number of energy management systems and may seem to be a logical approach for achieving energy reductions, the actual reductions from such feedback are somewhat disappointing. Ayres et al. (2013) found savings of 1.2-2.3% in two field experiments, though it should be remarked the feedback was being provided only once a month. Nevertheless, past studies have shown customers are interested in comparisons (Egan, 1999b; Haakana, Sillanpää, & Talsi, 1997), though in some cases consumers have been suspicious about the validity of the comparisons (Darby, 2006). It may be these feelings were evoked by the fact that users were unable to verify the results, which is less likely to occur when comparing energy consumption to one's personal data (which will be discussed in the next section). Though savings from comparative feedback may be low at the moment, it may be that the technique can be improved by enhancing its competitive aspects through gamification (see Section 3.4.2.

**6.5.2 Historic feedback**

*Historic feedback* refers to comparing present consumption figures to those recorded in the past. According to Darby (2006), this type of indirect feedback has yielded higher energy savings than comparative and normative feedback. Cross-period comparison itself requires hardly any computations (only the usual weather and PV corrections, discussed in the technical discussion below). Consumption during a certain period can be compared to any other period, given there is data available. If there is a contrast between the two, it can be shown as a difference or as a proportion.

**TECHNICAL DISCUSSION**

*Data corrections*

Historic feedback requires some form of seasonal compensation when it concerns different periods (e.g. when comparing consumption in February and September). For electricity, normalization can be avoided by restricting to only comparing the same period (e.g. comparing June 2018 with June 2017) or periods which are within a few days of each other (e.g. comparing two subsequent Saturdays, or one week with the next). For gas consumption, some form of compensation for temperature differences should always be implemented, since the volatility of the weather can create very different circumstances upon otherwise similar periods. This can either be realized by means of normalization or by accompanying the raw numbers with a temperature graph to place the numbers into perspective.
If a consumer has a PV installation, a meaningful cross-period electricity consumption comparison should include an estimate of the amount of electricity generated in both target periods. As with temperatures, the solar influx is a very volatile variable which may vary greatly between different periods.

**Energy conservation**

Historic feedback is one of the more effective and more popular types of indirect feedback. Most energy management systems surveyed in Section 3.4.3 feature an option for users to compare their most recent energy data to their own past consumption data. Additionally, historic feedback is popular among consumers (Darby, 2006; Fitzpatrick & Smith, 2009). Estimates of the effect of historic feedback often present savings in the range of 1-5% (e.g. Raw & Ross, 2011), but some authors argue is rather context-dependent and may have very different effect on different audiences (McKerracher & Torriti, 2013). There are also researchers who challenge the belief that historic feedback is superior to comparative feedback. Egan (1999a) for instance states: “[while] it can be useful to detect anomalies in ones personal energy usage patterns, it is a poor indicator of fundamental problems in the energy consumption”, arguing that historic data should be presented in conjunction with a comparison to the data of others to put one’s data in perspective. Additionally, McKerracher and Torriti (2013) found historic feedback is more effective when provided in conjunction with real-time feedback.

### 6.5.3 Normative feedback

One of the fundamental techniques to support individuals in regulating their behaviour is to have them aim at a certain target. Findings from studies towards goal setting, have seen successful application in numerous field such as education, sport, health, social behaviours and the environment (Epton, Currie, & Armitage, 2017). Often deemed as the founders of modern goal-setting theory, Locke, Shaw, Saari, and Latham (1981) have defined a *goal* as “the object or aim of an action”. At the time they identified two conditions which largely influence the effectiveness of goal setting regarding achieving behaviour change, namely that:

1. the goal must be conscious and specific,
2. the goal must be sufficiently difficult.

In other words, a well-defined goal should be more than a general intention and should be a significant step past from what is usually achieved. A positive relationship between goal difficulty and the eventual achievement has been verified by several experimental studies in multiple disciplines (e.g. Epton et al., 2017; Latham & Brown, 2006), including that of energy conservation (Becker, 1978). The success rate may drop however once the goal difficulty exceeds the limits of one’s abilities (Erez & Zidon, 1984). Later studies have lead to a number of additions to the above list, which were bundled in Locke and Latham (2002, 2006). Specifically, the effectiveness of goal setting is particularly high if:

3. the goal is publicly visible (Hollenbeck, Williams, & Klein, 1989; Epton et al., 2017),
4. the subject is high in need for the achievement (Hollenbeck et al., 1989),
5. the subject’s locus of control\(^{(57)}\) is internal (Hollenbeck et al., 1989),
6. it concerns a *cooperative group goal*, i.e. a goal structure which ensures the target can only be achieved through a collaborative effort by several group members (D. W. Johnson, Maruyama, Johnson, Nelson, & Skon, 1981).
7. the goal is framed in a non-threatening fashion (Drach-Zahavy & Erez, 2002),
8. the subject is regularly provided with interim feedback on his or her progress towards reaching the goal (Karlin, Zinger, & Ford, 2015). Bandura and Schunk (1981) found that providing a goal and information about progress toward that goal could serve as a form of behaviour modification, much like providing a reward or punishment.

There are also some factors on which the literature does not provide an unambiguous conclusion as to the effect on the success of goal-setting efforts. An example from an energy conservation context is the question whether it matters if people get assigned a target or get to pick one themselves. In a study by Abrahamse, Steg, Vlek, and Rothengatter (2005) no significant difference in terms of the amount of energy saved between these two groups of participants was found, whereas McCalley and Midden (2002) found this depends on the subject’s social orientation, with “pro-self individuals saving more energy when allowed to self-set a goal and pro-social individuals saving more energy when assigned a goal.”

When designing a goal-setting feature for an energy management system based on smart meter data, a number of the above items can be taken into account. The goal’s difficulty for instance can be automatically set based on the consumer’s historical data (see ‘An automated target-setting method’ below) and could be clarified to the customer in terms of tangible examples (e.g. naming a set of specific actions which would secure the goal to be reached). As smart meters more or less continuously monitor the subject’s energy consumption, it is relatively easy to regularly update a customer on their progress (see ‘Interim feedback’ below). Also, in any digital environment an option for sharing one’s enrollment in a goal setting attempt on social media could make the effort publicly visible and even some kind of community structure could be designed in which customers can interact with one another and try to take on a common goal.

Some of the above aspects however, depend mainly on the personality of each individual consumer. Whether a customer feels a reduction of energy consumption has priority or not for instance, depends upon his or her stance towards climate change and possibly on their financial situation (in which case monetary savings trigger the need for energy conservation). Also, a customer’s locus of control depends on their personal circumstances such as their socio-economic status (Meyerhoff, 2004; Maqsud & Rouhani, 1991). Some of these aspects can be partially steered when the feature as a whole is well-framed, but there will always be differences at the level of the individual which will be difficult to control with a generic goal-setting feature.

**A goal-setting feature design**

An outline for a possible goal-setting design will now be sketched. Possibilities to integrate a

\(^{(57)}\)Referring to Julian Rotter’s *locus of control* theory, according to which an individual is somewhere on the continuum between ‘internals’, who believe that one has influence over outcomes through ability, effort, or skills, and ‘externals’ who believe that forces outside the control of the individual determine outcomes (Rotter, 1966).
number of the aspects in the above list will be given, such as an automatized target finder and a feedback functionality. However it is eventually up to the reader to decide upon how the details are filled in, such as the way the functionality as a whole is framed and whether any social media linkages or communal structures are implemented.

AN AUTOMATED TARGET-SETTING METHOD
In order to choose a customer-specific target which meets the conditions listed above, an automated method will be proposed where the target is based on any available historical consumption data. First of all, a period over which the target should run needs to be chosen. This can be a day, week, month, year or any customized period preferred by the customer. Once the time frame has been fixed, the target can be set as a percentage of the customer’s past consumption. Optionally, one can have the customer choose between various goal difficulties to which different percentages are assigned. Reasonable percentages should be determined through a series of experiments.

The process of determining a target based on a customer’s energy data can generally be carried out more accurately the more data is available. If the period over which the data is available stretches out for more than a year, a thorough analysis towards the average consumption as well as any seasonal fluctuations can be assessed, allowing for a well-picked season-dependent target (see the technical discussion below). Such efforts become increasingly surrounded with uncertainties the less data is available. When there are very little past records or none at all, a last resort for automated target-setting in such a case is to set the goal based on the consumption data of customers with a similar profile, e.g. those inhabiting a similar housing type.

INTERIM FEEDBACK
Apart from the eventual aim itself, the road towards it is very important (Karlin, Zinger, & Ford, 2015). It is thus important to regularly provide the customer with an update on their progress, especially when the effort concerns a longer period. A basic form of interim feedback could for instance be a comparison between the portion of the target amount which has been consumed and the time passed, see Figure 46. For longer-term aims, more sophisticated updates can be provided once a sufficient part of the time has passed. Using a regression algorithm, a trend line can be drawn in a consumer’s cumulative energy consumption. By extrapolating this line from the initial stages of the target period towards the remaining time, an estimate can be given to whether it is feasible the target will be met. Figure 46 shows a visualization of this procedure.

TECHNICAL DISCUSSION
Data corrections
Assuming there are enough records to assess any annual consumption patterns, these can be used to pick a season-dependent goal. This way situations in which a target is far too easy or incredibly difficult (both of which may be detrimental to the success of the whole goal-setting procedure) may be avoided. Instead of normalizing the data, one could set the target as a percentage of the customer’s average consumption and then apply the normalization correction to the goal instead.

However, extrapolation is tricky and may give a false impression if executed carelessly. This will be discussed in more detail in the next section.
When a customer has solar panels, the goal-setting practice can be applied to their net consumption or generation figures. This way the need for any generation estimation methods is circumvented, but it also introduce a degree of luck into achieving the objective, since PV generation depends on the number of sunshine hours that will be registered during the running period, which cannot be accurately determined beforehand and may significantly deviate from multi-year averages. If this is deemed undesirable, generation needs to be estimated somehow, allowing the goal-setting feature to be applied to pure consumption figures instead.

**Energy conservation**

A program involving non-binding self-set goals in an energy context was devised by Harding and Hsiaw (2014) and resulted in energy savings of 11% for those who had set the most realistic goals. The average savings in that study were 4% though and one may argued that when setting a goal based on past data, one is simultaneously engaged with a historic feedback feature which makes it difficult to determine what part of the savings can really be traced back to the goal-setting feature. In a Dutch study where several types of feedback, including a goal-setting feature were combined, participant reduced their energy consumption by 5.1% over a five month period (Abrahamse et al., 2005). Moreover, the authors found that most participants had adopted a number of energy-saving behaviours during the course of the study, increasing the likelihood of stable savings on the longer term (see Section 3.4.1).

### 6.5.4 Energy saving tips

A substantial number of the energy management systems examined in Section 3.4.3 send their users saving tips to hand them specific pointers for reducing their consumption. Most of these systems only show generic tips, that may fail to encourage individual users. However, a number of the devices examined send out personalized saving tips based on an analysis of a specific consumer’s energy data. Saving tips have proven to be an effective way of achieving consumption reductions (Fischer, 2008), especially when they have a personal touch (Benders et al., 2006).
A saving-tip feature can be developed in parallel with all of the other applications discussed in this text. Starting off with a number of generic pieces of advice (e.g. replacing old model refrigerators, substituting classic light bulbs by LEDs, improving insulation, etc.) supported by numeric savings if possible, tips can be gradually personalized as more sophisticated data applications are implemented. Examples of this already discussed in this text are the refrigerator payback time assessment from Section 6.2 and the insulation assessment from Section 5.3.4. Personal tips can also be based on the cross-consumer comparisons discussed in Section 6.5.1. For instance if a household’s base load turns out to be above-average compared to that of other users occupying a similar housing type, one may send out some tips that concerns reducing standby power.

Additionally, one could think of a ‘tip-of-the-day’ type of feature, which sends out daily suggestions for possible energy savings. Tips can be sorted based on their perceived difficulty and potential energy savings, allowing users to choose between either a lighter or a more challenging feedback approach. This however requires a substantial database of tips, all accompanied by a decent estimation of their impact on energy use and quality of life.

**Energy conservation**

As with many other types of feedback, saving tips are often provided in addition to other types of feedback, making it difficult to isolate the effect from saving tips on behaviour and on consumption. Abrahamse et al. (2005) cite one rather dated study by Kantola, Syme, and Campbell (1984), where several groups of customers where given different types of feedback. The groups which only received some generic saving tips did not do significantly better than the control group, which supports the point made by (Benders et al., 2006) that saving tips should be personalized. Nevertheless, well saving tips may bridge the gap between static numbers about one’s energy consumption and specific action and may as such act as a conductor for the information from other applications.

**6.5.5 Manufacturing energy analysis**

Cost reduction is an important aspect of firms’ efforts to increase their financial benefits (Rust, Moorman, & Dickson, 2002). Accurate product-cost information is crucial to competitive success (Cooper & Kaplan, 1988). In this section, a method will be proposed which allows companies to draw conclusions from the relation between their energy usage and a second quantity that measures the company’s productivity somehow. For the sake of comprehension, the technique will be discussed here based on the case of a *factory* producing *tangible goods* (i.e. production volume will be chosen as the second quantity). Generalizations to different settings are discussed in the technical discussion afterwards. The author has not been able to find this particular application in literature, although there is some overlap with the work by Ghazanfari (2015); Sharp (1996).

The primary aim of the technique discussed in this section is to express a manufacturer’s energy consumption $C$ as a function of the production volume $v_i$, i.e.

\[
C = f(v).
\]  

(5)

where $f$ is a suitable function. If such a manufacturing energy model can be established, valuable information about the firm’s present and future energy consumption may be derived from it.
To find the sought-after function, a statistical modelling technique called *regression analysis* is deployed. There is a vast amount of literature on this subject, covering a whole range of different techniques designed for various applications. However, writing up a rigorous mathematical introduction on regression is outside the scope of this text and will not be pursued. Rather, a somewhat sketchy picture will be presented, to give the reader a feeling for the general idea behind regression and the role it plays within the smart metering application discussed in this section, without requiring any expert mathematical knowledge. If the reader is interested in a more detailed discussion, there are plenty of great, practically oriented, introductory texts on the topic, e.g. Harrell (2001); Draper and Smith (2014); S. Chatterjee and Hadi (2015); Montgomery, Peck, and Vining (2012); Weisberg (2005).

Regression analysis is an umbrella term for a number of statistical processes for modelling relationships between one *dependent variable* $Y$ and one or more independent variables $X = (X_1, \ldots, X_N)$ (Soto, 2013). Commonly, the goal of a regression analysis is to craft a model which estimates the behaviour of the dependent variable under changing conditions (i.e. changing values of the independent variables). In elementary situations, such a model is often crafted by plotting a collection of measurements of the dependent and independent variables acquired over time. If this plot reveals a pattern of some kind, a suitable regression function is fit onto the data in a way that its graph mimics the identified pattern, see Figure 47 for two examples. If successful, one can to some extent predict the behaviour of $Y$ in situations that have not yet been encountered, through inter- and extrapolation of the regression function (Klosterman, 1990).

![Figure 47: Two examples of simple regression.](image)

**Setting**

In what follows, regression will be applied to the pair $(v, C)$, where $v$ represents a firm’s manufacturing output and $C$ their energy consumption. The procedure can only be carried out if these variables satisfy a number of conditions, listed in e.g. Berry (1993). Rather than going

---

Regression analysis applied to situations with one dependent and one independent variable is referred to as *simple regression*, whereas *multivariable regression* relates to all problems that feature multiple independent variables. It is also possible to have more than one dependent variable, in which case one speaks of *multivariate regression*, however this will not be discussed in this text.
into the details, it will be assumed these conditions are sufficiently met; the technical discussion features a number of practical examples showing what may happen if one of them is violated.

### 6.6 Assumption: The pair \((v, C)\) satisfies the requirements for a regression analysis.

The specific ingredients of the method are sequences of measurements of both a factory’s production figures (which need to be provided by the company) \(v = (v_1, \ldots, v_N)\) and energy consumption \(C = (C_1, \ldots, C_N)\) corresponding to a time series \(t = (t_0, \ldots, t_N)\), such that \(t_i - t_{i-1}\) is equal for every \(1 \leq i \leq N\). The duration of the time frames over which production and energy consumption are measured does influence the technique: the pros and cons of various window length decisions will be commented on in the technical discussion.

#### Method

When scatter-plotting measurements in the \((v, C)\)-plane, a regression algorithm can be used to fit a curve through the data points. This curve represents the function \(f\) described in equation (5) described at the beginning of this section. Once a well-fitting function has been found, valuable information about both present and future energy consumption can be derived from it. First off, one can predict energy consumption at various manufacturing outputs. For example, if all product orders for one day have been received, one can predict the corresponding manufacturing energy costs by evaluating \(f\) at the total number of ordered products \(v^*\), i.e. \(f(v^*)\) should provide a fairly accurate estimate of the corresponding energy consumption. This can be done not just within the range of the regression function, but also for production volumes that have not yet been encountered, by extrapolating \(f\). This inevitably introduces somewhat larger uncertainty margins, especially when it concerns values that lie further outside of the range of the original data (see the technical discussion below).

Extrapolation all the way to the left yields an estimation of all power consumption during office hours not directly related to the actual production process, i.e. \(f(0)\). This may be an interesting...
quantity for energy-saving endeavours that do not hinder manufacturing itself. Extrapolation to the right gives a prediction of power consumption for larger production volumes that have not yet be attained, but may be in the future. This may provide a bit of a direction of future energy costs for growing firms, although one should keep in mind extrapolation comes with uncertainties. Figure 48 depicts the above for a fictional example dataset.

Also, one can assess energy costs associated with the manufacturing of one single product or batch of products, by considering the following product function:

\[
g = \frac{f(v)}{v}. \tag{6}\]

This function can be useful in determining profit margins and how these develop if one manages to cut down on energy costs.

**TECHNICAL DISCUSSION**

*Discussion of assumptions*

As mentioned, the standard regression model uses a number of assumptions, which concern various aspects of the dependent variable, the independent variables and the error terms (the natural deviations of individual measurements of \( Y \) from their predicted mean values) (Berry, 1993). The assumptions themselves are somewhat technical, and a rigorous discussion does not fit within the practical approach adhered to in this text. The assumptions are important though, as will be shown through multiple practical examples spread throughout this section.

Two examples of possible outcomes when a condition is violated are depicted in Figure 49. In the graph on the left there is too little correlation between production and energy consumption. This can be the result of disturbing influences of external factors, examples of which will be given in the remainder of this section. In the graph on the right there is too little spreading among observations of the independent variable. Translated to the current setting, this happens when a factory’s production output is very steady. In both cases, it will be very difficult to fit a regression curve and draw any conclusions.

![Figure 49: Possible results when correlation is weak (left) or when there is little variation in production (right).](image-url)
Data corrections
The regression technique as described above will only succeed if energy consumption is not significantly influenced by anything external. Even though industrial energy consumption is less susceptible to weather influences than domestic energy usage (see Section 5.3), such influences may still pose too large a disturbance to the above method. Instead of normalizing the data, one may attempt to implement some meteorological conditions into the regression model as independent variables. For example, one could attempt to model how gas consumption develops under changing production and temperature conditions. A similar procedure could be followed when a firm has its own solar energy production unit. In this case, one could add solar irradiation measurements as an independent variable.

Choice of parameters
The above method features a number of parameters, but most of these are related to the type of regression model used, which is highly situational and will not be discussed here. The only non-regression parameter to set is the time frame over which the data is acquired. As mentioned already, the choice of window duration may affect the technique. In general, shorter time frames increase the rate at which data is acquired. However, in some cases using a short time frame may have undesirable side-effects, as illustrated by the following examples:

- When the firm in question manufactures its products in low quantities, the impact of individual products on the energy-product correlation may become too large. Consider for instance an automobile factory which completes about 4 cars every fifteen minutes. When using quarter-hourly windows for the data acquisition, the instance where a car is marked as manufactured mere seconds after one time window finishes creates a large outlier: the last car is not counted as a product for this window, while almost all of the energy used to manufacture it is. This problem does not occur when the number of products is larger. For example, when considering a firm producing 30,000 toothpaste tops each 15 minutes, it does not matter if one individual item falls within one window or the next.

- A possible second issue of using shorter time frames, is that conditions may vary between different windows. Consider for example a factory that is closed at night. When the production process is initiated in the morning, it may take a while until the first finished products are completed. Therefore, during the first few time windows very low production volumes are registered, while energy consumption is regular. This ‘start-up lag’ in production output can create weird results and may lead to bad model choice decisions, as shown in Figure 50. To circumvent this problem, one can try to include the effect in the regression model by adding an independent variable which simulates it somehow. Alternatively, one could downsample to a daily resolution, as then every sample includes the same start-up lag. This greatly reduces the data acquisition rate however.

Expected performance & possible improvements
An obvious disadvantage of the regression method is that it requires the cooperation of the company in providing the necessary manufacturing data. Also, a part of the procedure may need to be carried out manually, to tailor it to a specific company. On the plus side, most of the information derived concerns long-term knowledge, so the process as a whole only has to be carried out once (that is, until the production process is radically changed).
In absence of any test data, it is difficult to comment on the chances of success of the above method. The extrapolation aspect is probably the most susceptible to problems, as extrapolation of regression models is one of the hot potatoes in statistics (e.g. Loh, Chen, & Zheng, 2007; Bystritskaya, Pomerantsev, & Rodionova, 2000; Chiou-shuang, 1978; Snee, 1977). Extrapolation by itself is already prone to errors, but when applied as proposed in the above, an additional uncertainty arises in the correctness of the regression model. Two common types of bad modelling choices are shown in Figure 51. Overfitting occurs when one overly attempts to have the regression function go past every measurement. Usually, this can only be achieved by using a highly fluctuating function which does not model the situation well and generalizes badly to unseen situations (Sharma, Nori, & Aiken, 2014). The opposite can also occur; underfitting relates to situations where the chosen regression function is too simplistic to properly model a situation. Both types of misreads of a situation can have a detrimental effect on extrapolation efforts, as Figure 51 shows. Pointers to find the best trade-off between the two extremes can be found in Sharma et al. (2014).
**Generalizations**

Production processes may be somewhat more complicated than has been presumed so far. Consider for example a factory which manufactures a variety of products in different proportions. If the energy intensity for the manufacturing processes differs between these products, then total energy consumption depends not only on the number of manufactured products, but also on the proportion in which they are produced. In this case, one could keep track of the production volumes for each product and switch to multivariable regression to see if $C$ can be written as a function of the $u_i$'s. In this case the energy consumption per product may vary between different combinations of products, which could reveal interesting optimization opportunities in determining cost-efficient product configurations.

It can also be that increasing production volumes at some point necessitates switching on additional machinery, which may create a step-wise increase somewhere in the $(v, C)$ diagram. If a significant step size can be identified in the data, one can apply segmented regression: partitioning the data into a number of connected segments and fitting different regression functions onto each segment, see Figure 52. This creates a more detailed model of reality and can therefore increase the accuracy of its applications. This is also something to keep in mind when extrapolating, i.e. consider the productive capacity of different machines when estimating future energy consumption.

![Figure 52: Example of segmented regression performed on a step-wise consumption-production relation. If production exceeds the threshold $T$, an additional machine is switched on.](image)

Up until now, the regression method has been applied to factories, i.e. based on the relation between energy consumption and the number of products manufactured. However, this quantity does not always represent a firm’s state of operations, can sometimes be difficult to determine and may not meet the regression conditions. Fortunately, the above method is in no way limited to the production volume of a factory. In fact, many quantities that correlate to a firm’s energy consumption can be used to carry out the above procedure:
6.7 Assumption: There exists some quantity $X$ which captures an aspect of the firm’s productivity and is correlated to the firm’s energy consumption in a way that $(X, C)$ meets the conditions for a regression analysis.

Consider for instance a service company that provides a certain kind of services. Then, the number of products (services) is probably not a good measure for the firm’s activity, as there may be large differences in the duration of individual services unrelated to energy consumption. However, there are many substitutes, such as ‘the number of people logged onto their account in the office’, ‘the firm’s internet data usage’ or even more creative quantities, such as ‘the number of free parking spots surrounding the office building’ or ‘the amount of coffee consumed’. As long as it is a somewhat variable quantity which makes a good representation of the activity within the office building and complies with Assumption 6.7. To increase the accuracy of the estimations, one could try to run several approaches simultaneously, i.e. operating the procedure for different input variables of $X$ and observe whether the estimates agree.

Energy Conservation

For the above method it is difficult to say anything about the effect on energy use, since it has not yet been tested in any environment. However, a number of scenarios are imaginable. For example, it could be that based on the information derived from the above technique, a firm decides to adjust their production process to reduce their operational costs. It is possible though this leads to an increase in energy use, if realizing a scale-up turns out to be the favourable option for the firm. All in all the consequences of the above are best to be determined through a series of tests.
7 Discussion

This thesis has provided a more practical approach to topics that have been discussed in more theoretical settings on numerous occasions. While a great number of papers discussed the effectiveness of feedback on energy consumption in different settings, very few go into the details of how to realize such feedback in practice. Therefore, the main contribution of this thesis is that it may facilitate a more rapid implementation of different types of feedback for a much larger public.

An important aspect of this text that may have been overlooked in other studies is the continuous attention of weather effects and at-home PV generation on energy data applications. As mentioned already in Chapter 4, a small number of studies talk about weather corrections in relation to historic feedback, but the implications of external effects are much larger than that. The solutions discussed in this text to resolve the issues caused by these implications may not be satisfactory though. Developing better ways to correct the data of individual households regarding these effects could be an avenue for further research.

Another aspect of the existing literature to which this project may be of use, is the disorganized set of notions used to address low levels of household energy consumption. This makes it difficult to relate the results of different empirical studies on topics such as standby power to each other, since it is not always entirely clear which aspects are included in such assessments and which are not. Though the terminology presented in Section 6.3.1 is not new and may not be perfect, a similarly organized disassembling of the lower ends of household energy consumption is unknown to the author. As such, it may provide a starting point for the establishment of one unambiguous set of notions regarding a number of different layers in household energy use.

Limitations

This study has attempted to draw on both the technical and the behavioural aspects of feedback to provide a complete picture. However, for both these aspects one element has received too little attention, namely that of testing. The author is aware that this is a serious limitation of this project. On the technical side, many solutions have been designed based on mathematical intuition and on a very limited set of data recordings. Even though most of the technical solutions have been discussed in many different situations that could theoretically arise at some point, it may very well be that one or more solutions fail to function properly when exposed to a real-world setting overlooked here.

Similarly, none of the solutions have been implemented as of yet, so it is at this point uncertain how customers will react to them. Numerous empirical estimates from literature have been added to provide at least a bit of a direction as to what one may expect from each of the applications and how they compare to one another. Yet, the author does not recommend to overvalue these estimates, as the circumstances under which many experiments were conducted are not equal. There are so many secondary variables that influence consumption (climate, season, country, dwelling type, income, etc.) that it is hardly justifiable to copy-paste the results of some American or Chinese study conducted during an unknown season in an unknown town with unknown types of residences, to the situation here in the Netherlands. As such it is not unlikely the eventual impact of the solutions sketched here, is very different from the expectations based on the empirical estimates presented here.
8 Conclusions & recommendations

As this text features a fair number of possible applications, it is possible the reader may not know where to start when implementing. In this final section, the applications discussed throughout this text will be ranked based on a number of variables, which relate to a number of technical aspects and to anticipated energy savings. More specifically, the assessment is based on the following questions:

(A) \textit{P1 dependency}: Does the application depend on P1 data?

\begin{itemize}
\item 1 (yes) - fully depends on P1 data.
\item 3 (both) - can be deployed with P4 data, but may work better with P1 data.
\item 5 (no) - fully P4 compatible.
\end{itemize}

(B) \textit{Technical complexity}: How difficult will it be to build the solution?

\begin{itemize}
\item 1 (very high) - requires high level of expertise on mathematics and/or machine learning
\item 2 (high) - requires some knowledge on mathematics and/or machine learning
\item 3 (medium) - can still be tricky but does not require any specific skills other than general programming experience and some proficiency with basic arithmetic.
\item 4 (low) - not very demanding, technically.
\item 5 (very low) - no technical aspects involved.
\end{itemize}

(C) \textit{Technical performance}: How likely is it that the solution will work properly? (Without taking into account problems from private PV generation and weather effects).

\begin{itemize}
\item 1 (very low) - solution is most likely inoperable in most settings.
\item 2 (low) - it is likely the solution will fail in a significant number of situations.
\item 3 (medium) - there may be situations in which the solution may fail.
\item 4 (high) - solution is expected to almost always work as intended.
\item 5 (very high) - can hardly go wrong.
\end{itemize}

(D) \textit{Dependency on PV estimations}: to what extent does the technique still work when consumers own a PV installation and can such problems be bypassed easily?

\begin{itemize}
\item 1 (very high) - solution can only be operated when PV generation is retrieved.
\item 2 (high) - solution performs significantly below the intended level due to PV-related problems.
\item 3 (medium) - there is a way to deal with PV-related problems which works most of the time.
\item 4 (low) - there is a good workaround to avoid problems with PV generation.
\item 5 (very low) - not affected by PV generation.
\end{itemize}

(E) \textit{Dependency on weather corrections}: To what extent does the technique require weather-normalized data?

\begin{itemize}
\item 1 (very high) - solution can only be operated on weather-normalized data.
\item 2 (high) - solution performs significantly below the intended level due to weather effects.
\item 3 (medium) - there is a way to deal with weather-related problems which works most of the time.
\item 4 (low) - there is a good workaround to avoid problems with weather influences.
\item 5 (very low) - not affected by weather influences.
\end{itemize}

(F) \textit{Projected energy savings}: What energy savings are associated with the technique based on empirical evidence found in literature?
Points are awarded based on the discussions of each individual solution in Chapter 5 and 6. In some cases, one or more questions cannot be answered based on literature, in which points were awarded based on an educated guess. The ranking presented here can be used as a roadmap for the development phase of smart meter applications, allowing to start off with the easier ones and adding more sophisticated options in subsequent development stages. There will be separate roadmaps for households and business customers, since some solutions apply to only one of those settings and the added value per application may also differ between the two sectors.

### 8.1 Solution roadmap for Huismerk Energie

Table 14 shows a ranking of the solutions discussed in this text that are applicable to Huismerk Energie. In this ranking all of the 6 aspects are equally weighted, which may lead to an overemphasis on technical aspects. However a different weighting system can be applied or a few variables can be discarded if one would like to focus more on the aspect of energy savings, complexity, etc. It should be noted the ranking needs to be redetermined once the first P1 application has been realized, because from that point on P1 dependence of a solution poses less of a barrier.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Section</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>Total</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulation</td>
<td>5.3.4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>Cost</td>
<td>6.1.1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td></td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>$CO_2$</td>
<td>6.1.2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td></td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>NILM (HMM)</td>
<td>6.2.4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>NILM (ANN)</td>
<td>6.2.4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>NILM (Edge-based)</td>
<td>6.2.4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>NILM (Shape-based)</td>
<td>6.2.4</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>6.2.4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>Peak load (Flat-level)</td>
<td>6.3.4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>Peak load (Recurrent)</td>
<td>6.3.4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>Idle consumption</td>
<td>6.3.5</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>Change detection (Hybrid)</td>
<td>6.4.4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>Change detection (Edge clustering)</td>
<td>6.4.1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Change detection (Peak saturation)</td>
<td>6.4.2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>Change detection (Long-term)</td>
<td>6.4.3</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Comparative</td>
<td>6.5.1</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>Historic</td>
<td>6.5.2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>Normative</td>
<td>6.5.3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Saving tips</td>
<td>6.5.4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>26</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 14: Ranking of the solutions relevant for Huismerk Energie based on the above six variables.
As can be seen from the above table, the output conversion techniques from Section 6.1, the different types of indirect feedback discussed in Sections 6.5.1 - 6.5.4 and the recurrent value base load method attain the highest scores. This is mainly due to their low complexity, independence of P1 data and insusceptibility to weather effects and ramifications of private PV generation. This conclusion is in line with the findings from Section 3.4.3, where similar features were among the ones most commonly implemented in existing energy management systems.\(^{(60)}\) Once these features have been implemented, the refrigerator payback time and insulation assessment are interesting data applications, with higher rewards in terms of energy savings. Despite their high savings potential, NILM methods are not technically mature as of yet and appear not very high on the list.

### 8.2 Solution roadmap for De Groene Stroomfabriek

A number of the applications discussed in this text are hardly relevant for organizations and have been omitted (some of the remaining solutions may still not be relevant to individual firms). Again, one may rearrange the ranking by adding different weights to each variable. Also, additional points could be awarded to solutions that are relevant for both Huismerk Energie and De Groene Stroomfabriek and are similar in terms of technical specifications.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Section</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>Total</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>6.1.1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>6.1.2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>Peak load</td>
<td>6.3.2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>Idle consumption</td>
<td>6.3.5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td>Change detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Hybrid)</td>
<td>6.4.4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>(Peak saturation)</td>
<td>6.4.2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>(Long-term)</td>
<td>6.4.3</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>Historic</td>
<td>6.5.2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td>Normative</td>
<td>6.5.3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>Saving tips</td>
<td>6.5.4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>Manufacturing analysis</td>
<td>6.5.5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>15</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 15: Ranking of the solutions relevant for De Groene Stroomfabriek based on the above six variables.

For business customers too, output conversion and saving tips are on top of the list. Yet, idle consumption is also a very attractive quantity in this case. The manufacturing analysis feature, though developed specifically for the class of business customers, is not very high on the list. This is mainly due to the many technical uncertainties surrounding the technique.

\(^{(60)}\)Except for estimations of the base load.
Appendices

Matlab Code

Median filter

```matlab
function [Q] = median_filter(P,R)

% INPUT
P #Input power signal.
R #Range parameter.

% OUTPUT
Q #Smoothened power signal.

% ASSUMPTION
#It is assumed length(P) >> 2R.

% PROGRAMM

[I,~]=size(P); #Determines the length of P.
Q=zeros(I,1);

for i=1:R
    Q(i) = median(P(1:i+R)); #Replaces the first R values by the
                   median taken over the first i+R numbers.
end

for i=R+1:I-R
    Q(i) = median(P(i-R:i+R)); #Replaces all non-boundary values by
                   the median of itself and the neighbouring 2R numbers.
end

for i=I-R+1:I
    Q(i) = median(P(i-R:I)); #Replaces the last R values by the
                   median taken over the last i+R numbers.
end
```
Flat level method

```matlab
function \[Q] = flat\_level(P, M, L, T0, Traise, Ttol, a)

% INPUT
P ... track of the input power signal.
M ... Window size.
L ... Number of lowest values considered.
T0 ... Initial threshold.
Traise ... Threshold raise amount.
Ttol ... Interval length requirement.
a ... Number of intervals that should meet the length requirement.

% OUTPUT
Q ... An estimate of the base load.

% ASSUMPTION
It is assumed length(P) > M > L.

% PROGRAM
[I, ~] = size(P); % Determines the length of input vector P.
R = zeros(M, 1); % Initiates a vector of length M to contain the M last values from P.
for i = 1:M
    R(i) = P(I-M+i); % Copies the last M values from P to R.
end
r = R; % Initiates a temporary copy of R.
S = zeros(L,1); % Initiates a vector of length L to hold the L lowest values from R.
V = zeros(L,1); % Initiates an index vector of length L to keep track of the original positions of each S(j) on R.
for j = 1:L
    [S(j), V(j)] = min(r); % Copies the jth lowest value from r to S and the corresponding index to V.
    r(V(j)) = max(P); % Replaces the jth lowest value from r by something very high to allow for the detection of the (j+1)th lowest value in the next iteration step.
end
# Now S consists of the L lowest values (ordered by size) from R and V contains the corresponding indexes of these minima.
A = zeros(L, 3); % Initiates a matrix to keep track of the extent to which each interval stretches out the left (first column) and to
```

---

145
the right (second column). The third column is used to track the
original index of the value so that it can be located when values
are removed and the original order is scrambled.

b = 0; #Initiates counter to keep track of the number of intervals
that meet the Ttol requirement.
s = S;
v = V; #Initiates temporary copies of S and V from which values may
be removed.
T = T0;
while b < a
  for j = L:-1:1 #Runs backwards for loop to avoid problems when
    indexes are removed.
    c = 0;
    d = 0; #Initiates left and right side counts.
    while abs(s(j)-R(v(j)-c)) < T && v(j)-c >= 2 #checks to
      the left of R(v(j)) whether values deviate less than T
      from s(j). This procedure is continued until either
      the threshold is exceeded or there are no more samples
      in R.
      c = c+1;
    end

    while abs(s(j)-R(v(j)+d)) < T && v(j)+d <= M-1 #checks to
      the right of R(v(j)) whether values deviate less than T
      from s(j). This procedure is continued until either
      the threshold is exceeded or there are no more
      samples in R.
      d = d+1;
    end
  end

if c+d+1 > Ttol
  A(b+1,1) = c-1;
  A(b+1,2) = d-1; #Writes the final left- and right-side
  interval length counts onto the (b+1)th row of A (c-1
  and d-1 were the largest values for which the
  threshold was not exceeded on either side).
  A(b+1,3) = v(j);
  s(j) = [];
  v(j) = []; #Removes the jth entry from s and v.
  b = b+1;
end
\[ \text{end} \]

\[ [L, \sim] = \text{size}(s); \]  
#Updates the size to the new situation in which entries may have been removed from s.

\[ T = T + \text{Threshold}; \]  
#Raises the threshold for the next iteration.

\[ \text{end} \]

\[ A(\text{all}(\sim \text{any}(A),2),:) = []; \]  
#Removes all zero rows from A. The length of A equals \( b \) after this step.

\[ M = \text{zeros}(b,1); \]  
#Initiates a vector in which the median value for each interval will be stored.

\[ \text{for } j = 1:b \]

\[ B = \text{zeros}(A(j,1)+(A(j,2)+1),1); \]  
#Initiates a vector of length equal to the length of the \( j \)th interval.

\[ \text{for } k = 1:A(j,1)+A(j,2)+1 \]

\[ B(k) = R(A(j,3) - A(j,1) - 1 + k); \]  
#Copies the \( j \)th interval onto B.

\[ M(j) = \text{median}(B,1); \]  
#Takes the median of this interval.

\[ \text{end} \]

\[ B = []; \]  
#Resets B for the \((j+1)\)th interval.

\[ \text{end} \]

\[ Q = \text{min}(M); \]  
#The minimal interval median is returned as the base load.
Peak saturation

```matlab
function [Q]=peaksat(P,L)

% INPUTS
P #Input consumption signal.
L #Length of the ranking.

% OUTPUT
Q #The cumulative update number which keeps track of instances at which the record is updated.

% PROGRAM
[I,˜] = size(P); #Determines the length of input vector P.
R=zeros(L,1); #Initiates the default ranking.
B=zeros(L,1); #Initiates a temporary side list for determining the position of a new entry on R.
Q=zeros(I,1); #Initiates the cumulative update number count list.
for i=1:I
    if P(i) > R(L) #Triggered whenever a consumption measurement is higher than the lowest value on R.
        for k=1:L
            B(k)=R(k)-P(i); #The difference between P(i) and each R entry.
        end
        [j,˜] = size(B(B>0)); #j equals the last entry in R which is larger than P(i).
        j=j+1; #j+1 equals the envisioned position of P(i) on R.
        for m=L:j+1:-1 #Runs backwards for loop to avoid problems when indexes are removed.
            R(m)=R(m-1); #Moves all R(m)’s smaller than P(i) downwards by one position.
        end
        R(j)=P(i); #Adds P(i) to R at the proper position.
        Q(i)=max(Q)+1; #If P(i) entered R, Q(i)=Q(i-1)+1.
    else
        Q(i)=max(Q); #If P(i) did not enter R, Q(i)=Q(i-1).
    end
end
```
References


Chahine, K., Drissi, K. E. K., Pasquier, C., Kerroum, K., Faure, C., Jouannet, T., & Michou,


Deconinck, G., & Decroix, B. (2009). Smart metering tariff schemes combined with distributed energy resources. In *Proceedings of the fourth international conference on critical infrastructures (cris)*, 2009 (pp. 1–8).


---

*Energy efficiency, 1*(1), 79–104.

---

156


Jazizadeh, F., Becerik-Gerber, B., Berges, M., & Soibelman, L. (2014). Unsupervised clustering of residential electricity consumption measurements for facilitated user-centric non-


Kim, H., Marwah, M., Arlitt, M., Lyon, G., & Han, J. (2011). Unsupervised disaggregation
of low frequency power measurements. In *Proceedings of the 2011 SIAM International Conference on Data Mining* (pp. 747–758).


Policy, 39(5), 2373–2385.


phones as energy consumption feedback devices. In *International conference on mobile and ubiquitous systems: Computing, networking, and services* (pp. 63–77).


Zhang, C. (2016). *A Deep Learning Approach to Identifying Household Appliances from Elec-


Glossary

absolute load ................................................................. 94
active power ................................................................. 23
adaptive thresholding ..................................................... 99
always-on appliance ....................................................... 95
artificial neural network .................................................. 87
attenuated consequences effect ......................................... 37

base consumption ......................................................... 95
base load ................................................................. 95
base load hot spot ......................................................... 98
boomerang effect .......................................................... 37

caloric correction factor .................................................. 76
capacity tariff ............................................................... 76
comparative feedback ...................................................... 124
congestion management .................................................. 19
consumption accumulation effect ...................................... 38
continuous appliance ...................................................... 95
continuously-varying appliance ......................................... 82
cooperative group goal ................................................... 129
cumulative median .......................................................... 117
cumulative update number .................................................. 116

dependent variable ......................................................... 133
deviation factor ............................................................. 63
diffusion of responsibility effect ........................................... 37
direct energy consumption .................................................. 17
direct factor ................................................................. 61
direct feedback ............................................................. 29
direct rebound effect ...................................................... 35
dynamic time warping ....................................................... 90

economy-wide rebound effect ............................................ 35
electric power ............................................................... 15
embodied energy effect ..................................................... 38
endogenous factor ........................................................... 57
energy costs ................................................................. 76
energy tax ................................................................. 76
error term ................................................................. 135
event ................................................................. 85
event detection ............................................................. 85
exogenous factor ............................................................ 57

financial effects .............................................................. 35

global warming potential .................................................. 78
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak hours</td>
<td>14</td>
</tr>
<tr>
<td>peak load</td>
<td>94</td>
</tr>
<tr>
<td>peak time</td>
<td>94</td>
</tr>
<tr>
<td>periodic device</td>
<td>95</td>
</tr>
<tr>
<td>permanent device</td>
<td>83</td>
</tr>
<tr>
<td>prebound-effect</td>
<td>39</td>
</tr>
<tr>
<td>primary factor</td>
<td>61</td>
</tr>
<tr>
<td>pro environmental spillover</td>
<td>39</td>
</tr>
<tr>
<td>process-efficiency effects</td>
<td>35</td>
</tr>
<tr>
<td>profile fraction</td>
<td>63</td>
</tr>
<tr>
<td>prosumer</td>
<td>48</td>
</tr>
<tr>
<td>psychological effects</td>
<td>35</td>
</tr>
<tr>
<td>PV performance model</td>
<td>51</td>
</tr>
<tr>
<td>reactive power</td>
<td>23</td>
</tr>
<tr>
<td>rebound effect</td>
<td>33</td>
</tr>
<tr>
<td>recurrent neural network</td>
<td>88</td>
</tr>
<tr>
<td>regional supplement</td>
<td>76</td>
</tr>
<tr>
<td>regression analysis</td>
<td>133</td>
</tr>
<tr>
<td>regression function</td>
<td>133</td>
</tr>
<tr>
<td>reinvestment effect</td>
<td>36</td>
</tr>
<tr>
<td>rising edge</td>
<td>111</td>
</tr>
<tr>
<td>saturation shift</td>
<td>115</td>
</tr>
<tr>
<td>secondary factor</td>
<td>61</td>
</tr>
<tr>
<td>segmented regression</td>
<td>138</td>
</tr>
<tr>
<td>semi-supervised method</td>
<td>85</td>
</tr>
<tr>
<td>sequence</td>
<td>15</td>
</tr>
<tr>
<td>simple regression</td>
<td>133</td>
</tr>
<tr>
<td>sine-regression</td>
<td>72</td>
</tr>
<tr>
<td>single-state appliance</td>
<td>82</td>
</tr>
<tr>
<td>smart grid</td>
<td>73</td>
</tr>
<tr>
<td>smart meter feedback</td>
<td>26</td>
</tr>
<tr>
<td>split billing</td>
<td>92</td>
</tr>
<tr>
<td>standard annual consumption figure</td>
<td>62</td>
</tr>
<tr>
<td>standby power</td>
<td>94</td>
</tr>
<tr>
<td>state</td>
<td>82</td>
</tr>
<tr>
<td>steady-state feature</td>
<td>84</td>
</tr>
<tr>
<td>stray voltage</td>
<td>109</td>
</tr>
<tr>
<td>supervised method</td>
<td>85</td>
</tr>
<tr>
<td>supplier fees</td>
<td>76</td>
</tr>
<tr>
<td>system services fee</td>
<td>76</td>
</tr>
<tr>
<td>take-back effect</td>
<td>33</td>
</tr>
<tr>
<td>transformational rebound effect</td>
<td>35</td>
</tr>
<tr>
<td>transient current spike</td>
<td>46</td>
</tr>
<tr>
<td>transient feature</td>
<td>84</td>
</tr>
<tr>
<td>Term</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Transtheoretical Model of Change</td>
<td>26</td>
</tr>
<tr>
<td>two-sigma interval</td>
<td>111</td>
</tr>
<tr>
<td>underfitting</td>
<td>137</td>
</tr>
<tr>
<td>unsupervised method</td>
<td>85</td>
</tr>
<tr>
<td>weighted average</td>
<td>54</td>
</tr>
</tbody>
</table>