Carsten Heinrich

Markov Chain Modeling of Aggregations of Electric Water Heaters for Demand Response Applications

Semester Project, Autumn 2013

Department:
EEH – Power Systems Laboratory, ETH Zürich

Supervisors:
Evangelos Vrettos, Maryam Kamgarpour, ETH Zürich

Professor:
Göran Andersson, ETH Zürich
Abstract

We develop a dynamic system model for a population of water boilers, which is simple enough for control purposes, but at the same time appreciates the complexity of the boilers internal dynamics and the volatility of such a load. For this purpose, a computationally involved model for a single water boiler is improved and then used to create a population. Based on these boilers a simple Markov Chain Model is created. Finally the ability of the Markov Model to predict the power consumption of other populations in an open loop control system is analyzed.
Contents

1 Introduction 3

2 Detailed Model of a Individual Electric Water Heater 3
   2.1 Laminar Flow 3
   2.2 Buoyancy Correction 5
   2.3 Water Draws 6
   2.4 Internal Controller 6
   2.5 Initial State 7

3 Water Heater Population 7
   3.1 State of Charge 9
      3.1.1 Calculation 1 9
      3.1.2 Calculation 2 10
      3.1.3 Calculation 3 10
   3.2 Markov Chain-based Population Model 10

4 Control Schemes 14
   4.1 Control of Set Temperature 14
      4.1.1 SOC definition 1 18
      4.1.2 SOC definition 3 18
      4.1.3 SOC definition 2 19
   4.2 Probabilistic Switching 19
      4.2.1 SOC definition 1 22
      4.2.2 SOC definition 2 22
      4.2.3 SOC definition 2 24
      4.2.4 Random Switching at every time step 25
   4.3 Prediction of Power Consumption 27

5 Conclusion 28
1 Introduction

More Renewable Energy Sources in the electricity grid require a higher participation in frequency control on the demand side. While in many countries industrial customers already participate in reserve markets ([1]), research on how to use the potential of devices in households is still going on. Special attention has been payed to thermostatically controlled load (TCL), due to their high potential and their ability to store energy in form of heat. In order to provide services for frequency control, one needs reliable models of their power consumption and its development over time. These models have to make accurate predictions for devices, that on the one hand show highly complicated dynamics and on the other hand are influenced by random costumer behavior. In many papers, TCLs are modeled with a uniform temperature distribution, e.g. [2], [3] and [4]. In [5] a model with two layers is used. Even though these methods might be computationally efficient, their abilities to predict behavior based on thermal gradient such as convection or conduction within the devices is limited. In this project we use Markov chain models, introduced in [6] and further investigated in [7] and [8], and base them on a model from [9], which calculates the temperature distribution within an electric water heater (EWH) and includes convection, conduction and friction. This way we develop a method, that both captures the complex behavior due to temperature gradients within EWHs and is still computationally efficient. Finally we test its ability to predict future power consumption in an open loop system.

2 Detailed Model of a Individual Electric Water Heater

2.1 Laminar Flow

For the detailed model of the individual EWHs, we use a model that was developed in [9], which assumes that the laminar heat flow of a EWH is described by the following equation:
\[
\frac{\partial T}{\partial t} + V \frac{\partial T}{\partial x} = a\epsilon_{eff} - \frac{\partial^2 T}{\partial x^2} - k(T - T_a) + Q(x, T)
\]  

(1)

Here \(x\) stands for the vertical axis of the EWH, \(T\) is temperature, \(t\) is time, \(V\) is the vertical water velocity during water draws, \(a\) is the thermal diffusivity, \(k\) is the heat loss coefficient, \(T_a\) is the ambient Temperature, \(\epsilon_{eff}\) accounts for turbulent mixing and \(Q(x, t)\) describes the heat generation by a heating element within the heater. Based on this PDE, the model moves on to a discrete grid in time and space. We divide the EWH into \(n\) layers of thickness \(\Delta x\), as well as time into time steps of \(\Delta t\). We apply the Crank-Nicolson scheme proposed in [10] and by use two boundary conditions for the top and the bottom layer:

\[
\frac{\partial T_2}{\partial t} = \frac{-a\epsilon_{eff}}{\Delta x^2}(T_2 - T_3) + \frac{V}{\Delta x}(T_{cw} - T_2) - k'(T_2 - T_a)
\]  

(2)

\[
\frac{\partial T_{n+1}}{\partial t} = \frac{-a\epsilon_{eff}}{\Delta x^2}(T_n - T_{n+1}) + \frac{V}{\Delta x}(T_n - T_{n+1}) - k'(T_{n+1} - T_a)
\]  

(3)

\(T_{cw}\) represents the water temperature at the inlet and \(k'\) is the heat loss coefficient of the top and the bottom layers. This way, eq. 1 can be brought into matrix form for the temperature development within the heater (see [9]):

\[
Z_1 T_{m+1} = Z_2 T_m + Z_3 u_m
\]  

(4)

or since \(Z_1\) is non-singular:

\[
T_{m+1} = A_m T_m + B_m u_m
\]  

(5)

with: \(A_m := Z_1^{-1} Z_2\) and \(B_m := Z_1^{-1} Z_3\)

Note, that two artificial layers were introduced, which represent the incoming water and the ambient temperature. Further \(T \in \mathbb{R}^{n+2}\) is the temperature vector for the \(n\)-layers of the EWH, plus two rows that contain the water inlet temperature \(T_{cw}\) as well as the ambient air temperature \(T_a\). \(u \in \{0, 1\}\) is the binary variable of heating operation, \(m\) is the time index.
and:

\[
Z_1 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\
-4d & 1/\Delta t & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & -k' \\
0 & -q_1 & q_2 & q_3 & 0 & \ldots & 0 & 0 & 0 & k \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \ldots & -q_1 & q_2 & q_3 & k \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 1/\Delta t & -k' \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]  
\[(6)\]

\[
Z_2 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\
0 & q_5 & e & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\
0 & -q_1 & q_4 & -q_3 & 0 & \ldots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \ldots & q_1 & q_4 & -q_3 & 0 \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & q_6 & q_5 & 0 \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]  
\[(7)\]

\[
Z_3 = \begin{pmatrix}
0 & \ldots & 0 & \frac{\eta P_{el}}{m_i c} & 0 & \ldots & 0 \\
\end{pmatrix}
\]  
\[(8)\]

where:

\[
q_1 = \frac{e}{2} + d, \quad q_2 = \frac{1}{\Delta t} + e + \frac{k}{2}, \quad q_3 = d - \frac{e}{2}, \quad q_4 = \frac{1}{\Delta t} - e - \frac{k}{2} \\
d = \frac{V}{\Delta x^2}, \quad e = \frac{a_{eff} \Delta x}{\Delta x^2}
\]

Here, \(P_{el}\) is the nominal electrical power of the heating element, \(\eta\) is the electrical efficiency, \(m_i\) is the water mass of the i-th layer and \(c\) is the specific heat capacity of water.

### 2.2 Buoyancy Correction

Natural convection is assumed to consist of two main forces. Buoyancy accelerates hot layers according to the formula

\[
F^b_i = \min[0, gm_i (\rho_{ref} - \rho_i)] (9)
\]

where \(\rho_{ref}\) is the minimum of \(\rho_{i+1}\) and the average density of the layers above layer \(i\). The friction force is given as the product of the friction parameter \(b\) and
and the vertical water velocity $U$. This parameter depends on the geometry of the EWH and the water viscosity.

$$F_i^{fric} = -bU_i$$ (10)

Based on these two forces, we calculate the acceleration of each water layer and the resulting distance that this water layer is going to move upwards within the next time step. The temperature of each layer $i$ within this distance is then corrected according to formula 11, using the empirical mixing ration $r$.

$$T_i = (1 - r)T_i + rT_{i+1}, \quad T_{i+1} = rT_i + (1 - r)T_{i+1}$$ (11)

2.3 Water Draws

For the water draws we create a vector of draw velocities. We model an amount of ten to fifteen water draws per day (uniformly distributed). These draws are distributed over one day according to a probability distribution from [11]. With a probability of 90% draws are short and with a probability of 10% they are long. Short water draws (e.g., hand washing) last about 60 seconds (duration normal distribution with mean value $\mu = 60s$ and standard deviation of $\sigma = 6s$), long water draws (e.g., showers) last about 600 seconds (normal distribution with mean value $\mu = 600s$ and standard deviation of $\sigma = 60s$). The speed of the water draws in liters per minute is a random variable (normal distribution with mean value $\mu = 5.4$liter/min and standard deviation of $\sigma = 1.08$liters/min). All water draws result in a vertical water velocity vector for every time step during one day.

2.4 Internal Controller

The thermostatic controller of the EWH is modeled as a measurement instrument, which is situated in the middle of the device. Whenever the measured temperature drops below or above the temperature limits, the boiler switches the heating element on or off.
2.5 Initial State

In order to start with a realistic initial condition, we run the simulation for the duration of one day and use the final temperature distribution at the end of the day as initial condition for the simulation.

Figure 1 shows the results of the simulation of one boiler. The upper graph depicts the temperature distribution within the boiler throughout one day. The middle graph shows the water draw velocities and the bottom graph the power consumption. All plots are from the same heater with set temperature of 78.98 °C, bandwidth of 8.1°C, a volume of 178.9 liters and the heating power of 3.41 kW.

3 Water Heater Population

With the model from Chap. 2 we create a population of 10 000 EWHs. In order to simulate a variety of boilers some heater properties are chosen randomly. These values are shown in table 1. Values, which were assumed to be constant throughout the population, are shown in table 2. In this project, we simulate boilers with the complex model from chapter 2. In this section we develop a Markov Chain-based Population model. We then compare the total power consumption of the “real system” form chapter 2 to the predictions made by this Markov Chain-based model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>value</th>
<th>Distrib.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_0$</td>
<td>volume of EWH [l]</td>
<td>100 - 250</td>
<td>uniform</td>
</tr>
<tr>
<td>$P$</td>
<td>power density of heating element [W/°C]</td>
<td>21.50 ± 20%</td>
<td>uniform</td>
</tr>
<tr>
<td>$v_{cw}$</td>
<td>water draw velocity [m³/s]</td>
<td>1.25, STD: 0.375</td>
<td>normal</td>
</tr>
<tr>
<td>$T_a$</td>
<td>ambient temp. [°C]</td>
<td>22, STD: 1.76</td>
<td>normal</td>
</tr>
<tr>
<td>$T_{set}$</td>
<td>set-point temp. of water heater[°C]</td>
<td>50 - 80</td>
<td>uniform</td>
</tr>
<tr>
<td>$\Delta T$</td>
<td>temp. bandwidth [°C]</td>
<td>5 - 10</td>
<td>uniform</td>
</tr>
</tbody>
</table>
Figure 1: Upper graph: temperature distribution in the layers throughout the simulated day, middle graph: water draw velocities, lower graph: power consumption
Table 2: constant population parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k)</td>
<td>tank heat loss coefficient</td>
<td>(6.3588 \times 10^{-7})</td>
</tr>
<tr>
<td>(k_{bt})</td>
<td>tank heat loss coefficient (top and bottom)</td>
<td>(1.2382 \times 10^{-6})</td>
</tr>
<tr>
<td>(r)</td>
<td>mixing parameter</td>
<td>(0.1)</td>
</tr>
<tr>
<td>(b)</td>
<td>friction parameter</td>
<td>(0.7 \frac{N_s}{m})</td>
</tr>
<tr>
<td>(a)</td>
<td>water thermal diffusivity</td>
<td>(0.1434 \times 10^{-6} \frac{m^2}{s})</td>
</tr>
<tr>
<td>(c)</td>
<td>water specific heat capacity</td>
<td>(4185.5 \frac{J}{kgK})</td>
</tr>
<tr>
<td>(g)</td>
<td>gravitational constant</td>
<td>(9.81 \frac{m}{s^2})</td>
</tr>
<tr>
<td>(\eta)</td>
<td>efficiency</td>
<td>(0.95)</td>
</tr>
<tr>
<td>(\epsilon_{eff})</td>
<td>turbulent mixing parameter</td>
<td>(1)</td>
</tr>
<tr>
<td>(T_{cw})</td>
<td>water temperature at inlet</td>
<td>(14.4 , ^\circ C)</td>
</tr>
<tr>
<td>(n_{heat})</td>
<td>layer with heating device</td>
<td>(2)</td>
</tr>
<tr>
<td>(n_{thermostat})</td>
<td>layer with thermostat</td>
<td>(6)</td>
</tr>
<tr>
<td>(n)</td>
<td>number of vertical layers</td>
<td>(10)</td>
</tr>
<tr>
<td>(\Delta t)</td>
<td>time step</td>
<td>(10 , s)</td>
</tr>
</tbody>
</table>

3.1 State of Charge

Since EWHs can store heat, it is helpful to apply the concept of state of charge (SOC). We consider a heater as fully “charged”, when the water temperature of the water within the EWH is at the upper temperature limit of the boiler bandwidth \(T_{max}\), and fully “discharged”, when the water temperature is at the lower limit \(T_{min}\), respectively. The SOC definition has a big impact on the behavior of the model. Several ways come to mind, both with advantages and disadvantages.

Since the heating element is situated in the second layer, the first layer below is not heated, neither directly, nor by convection. Conduction is the only way this layer heats up and since during a water draw, new cold water flows in at the bottom, this layer normally stays cold. Therefore this layer is not taken into account, when calculating the SOC.

3.1.1 Calculation 1

This calculation is based on the average water temperature in the layers. The SOC is set to zero in case that the temperature drops below the minimum bandwidth temperature and is set to 1, if the average temperature is above the maximal bandwidth temperature. Layers with a temperature
distribution lower than $T_{min}$, are taken into account, as if they had the temperature $T_{min}$. Eq. 12 therefore estimates the energy content of the EWH too high.

\[ SOC = \min(1, \frac{\sum_{i=2}^{N} \max(T_i - T_{min}, 0)}{N(T_{max} - T_{min})}) \] (12)

### 3.1.2 Calculation 2

Since the on-off-state of each EWH is uniquely defined through the temperature in the layer of the thermostat, this layer is of particular importance. Eq. 13 calculates the SOC using only the temperature measured at the thermostat.

\[ SOC = \min(1, \max(0, \frac{T_{thermostat} - T_{min}}{T_{max} - T_{min}})) \] (13)

### 3.1.3 Calculation 3

This calculation is similar to eq. 12. It simply rounds up to zero or down to one, after taking the average of all layers. The estimation of the energy content of each EWH is therefore closer to the actual content.

\[ SOC = \min(1, \max(0, \frac{\sum_{i=2}^{N} T_i - T_{min}}{N(T_{max} - T_{min})})) \] (14)

### 3.2 Markov Chain-based Population Model

In this simulation we use a Markov chain-based population model. In order to move on to markov chain transition matrices, we discretize the SOC into $N/2$ smaller temperature intervals. Further we differentiate states, depending on whether the boiler is currently heating or not and end up with a total of $N$ different states. For every time step, a boiler is in exactly one of these $N$ states defined by:

\[ bin_{z,t} = \left\lfloor SOC_{z,t} \frac{N}{2} \right\rfloor u_{z,t} + \left\lceil N + 1 - SOC_{z,t} \frac{N}{2} \right\rceil (1 - u_{z,t}) \] (15)
Figure 2: Schematic illustration of EWH behavior. Every bin corresponds to a temperature in the EWH. Switching on or off, corresponds to a movement to the left and to the right respectively.

$\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ are the Gaussian ceiling and floor functions. Further $z$ is the EWH index, $t$ denotes the time step and $u$ is one, when the heating element is running and zero, when the heater is off. Figure 2 illustrates the behavior of an EWH, which is not controlled externally, and from which no water is drawn. When the SOC of the EWH is close to zero and the boilers’ heating device is switched on, it is heating, while moving from bin 1 to bin $\frac{N}{2}$ until it reaches a SOC of 1, which is when it switches off and moves over to bin $\frac{N}{2} + 1$, gradually cooling down from now on, until reaching bin $N$, and turning on again. Water draws and external switching lead to more complicated dynamics of the heater, so that it can also switch on and off when the SOC is not close to zero or one.

If one assumes independent transition probabilities between the $N$-states, the behavior of the population can be described by a Markov Transition Matrix (MTM). The probability for a boiler to move from state $i$ to state $j$ is the $ji$-th entry $m_{ji}$ of the MTM $M$. Let $x(t)$ be a $\mathbb{R}^{N \times 1}$ vector, describing
the amount of EWH in the $N$ different states at time step $t$, then $x(t+1)$ is given as:

$$x(t+1) = Mx(t)$$ \hspace{1cm} (16)$$

The total power consumption (TPC) of the population can then be calculated with:

$$P_{tot}(t) = cx(t)$$ \hspace{1cm} (17)$$

where $c = (P, \ldots, P, 0, \ldots, 0)$ and $P$ is the average power of the heaters.

The entries $m_{ji}$ are calculated with the results from the detailed model. For all 10 000 boilers a bin-vector is created, that shows in which of the 20 bins the EWH was for every time step throughout the day. These transitions are then “copied” into the MTM. If a boiler at time step $t$ occupies bin $i$, and at time step $t+1$ it occupies bin $j$, $m_{ji}$ is increased by one. Note that this also includes the time steps, when a boiler stayed within the same bin. In the end all entries in the MTM are normalized by the total sum of entries in their column. This calculation is applied for every hour, leading to 24 different MTMs. If we use only one MTM for the entire day, this Matrix converges, within a couple of simulation hours, to the unique fixed point of the MTM. These fixed points exist for regular MTMs (see [12]).

$$\lim_{m \to \infty} TPC = \lim_{m \to \infty} cM^m x(0) = cx_{fix}$$ \hspace{1cm} (18)$$

The water draw probabilities throughout the simulated day vary every hour. Since the water draws have a strong influence on the stored energy within the EWHs, they also influence the transition probabilities between the bins. The transition probabilities during one day are therefore not constant, but time dependent. Hence, the longest time interval, during which these transition probabilities can be considered constant, is therefore one hour, the time during which also the water draw probabilities are constant.

Under the assumption that the transition probabilities are time invariant for constant water draw probabilities, another possibility is to use the same MTMs for every water draw probability. Since the probabilities are in real-
ity not time invariant, and the whole system is dependent on its past, this approach leads to much bigger errors and was not further investigated in this work.

Since the heating time is much shorter than the cooling time of boilers, it is crucial to use a population of EWHs, that is large enough, so that for every calculation at least one of the EWHs has been in every bin and a representative MTM can be calculated. Further, depending on the definition of the SOC, the bins, in which EWHs spend less amount of time are also different. For calculation 1 and calculation 3 times when heaters are in state $\bar{N}$ are rarer, since at many time steps, there is new cold water flowing in from below, leading to an average temperature, that is much less than $T_{max}$.

Fig. 3 shows the results of the TPC for the two simulations throughout one day for SOC definition 2. For all three SOC definitions the error stays below five percent, throughout the entire day. The error reaches its maximum, during times of high water draw probability.
4 Control Schemes

If control is applied to the population, this also has to be modeled in the Markov chain-based approach. How accurately we are able to predict changes in the TPC due to control depends on the estimation of the state-vector with the help of MTMs. These are dependent on the definition of SOC. Hence we investigate the control schemes for the three different SOC definitions.

4.1 Control of Set Temperature

In this section, we increase and decrease the TPC of the EWH population by manipulating the maximum and minimum set temperatures $T_{\text{max}}$ and $T_{\text{min}}$. Sending a signal $u(t) \in [-1, 1]$ results in a change of $T_{\text{min}}$ and $T_{\text{max}}$ as follows:

$$
\Delta T = (T_{\text{max}} - T_{\text{min}})u(t) \quad (19)
$$

$$
T_{\text{min,new}} = T_{\text{min}} + \Delta T, T_{\text{max,new}} = T_{\text{max}} + \Delta T \quad (20)
$$

A signal of 0.1 for example sent to a heater with $T_{\text{min}} = 50^\circ C$ and $T_{\text{max}} = 60^\circ C$ leads to new limits of $51^\circ C$ and $61^\circ C$ and boilers with a thermostat water temperature between $50^\circ C$ and $51^\circ C$ now have a temperature below $T_{\text{min}}$ and are switched on immediately. Devices, which were going to hit the old $T_{\text{max}}$ boundary are still continuing to heat water, until they hit the now higher limit. So by sending positive/negative signals, boilers are switched on/off, the maximum/minimum value 1/-1 switches all boilers in/off.

In the simulation this behavior is approximated as follows: when increasing the set temperature the devices from the lower two bins are put in the on-state bin with the lowest temperature and when decreasing the devices are put in the off-state bin with the highest corresponding temperature. Figure 4 illustrates this method for a set temperature increase. $X_i$ represents the number of heaters in bin $i$. After switching, in this case with an input signal that increases the limits by the width of one bin, the new bins with the highest temperature are empty.

In matrix form, a change of the set temperature can be represented as fol-
Figure 4: Change of amount of boilers in each bin when increasing the bandwidth temperature limits. Initial bin distribution on the left, new bin distribution on the right.

\[
x_{\text{new},i} = A \frac{\mu^N}{\tau} x_{\text{old},i}
\]

with

\[
A = \begin{pmatrix}
1 & 0 & \ldots & 0 & 1 \\
0 & \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad \mathbf{0} & \quad \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad \mathbf{0} & \quad \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad \mathbf{0} \\
0 & \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad \mathbf{0} & \quad \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad \mathbf{0} & \quad \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad 0 \\
0 & \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad \mathbf{0} & \quad \mathbb{I}_{\frac{N}{2} - 1 \times \frac{N}{2} - 1} & \quad 0 \\
\end{pmatrix}
\]
if $u > 0$, and

$$A = \begin{pmatrix}
0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
\textbf{1}_{\frac{n}{2}-1 \times \frac{n}{2}-1} & \vdots & \vdots & \textbf{0} & \\
0 & 0 & \vdots & \vdots & \textbf{1}_{\frac{n}{2}-1 \times \frac{n}{2}-1} \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{pmatrix} \quad (23)$$

if $u < 0$.

The way we model the reaction of the population to this switching strategy has certain drawbacks. Heaters, which after switching would lay outside the temperature bandwidth, are simply put into the closest bin within the temperature bandwidth. The water temperature within these heaters would not change, though. Hence the energy according to the new state, increases or decreases instantaneously so this approximation violates the law of energy conservation. By introducing additional states for heaters, that would have a negative SOC or a SOC greater than one, this effect could be taken care of in future work.

Further, the empty bins, which occur, when manipulating the set temperature, do not occur in the real system. Here, especially, when lowering the set temperature, there are heaters, that had a temperature outside the heater bandwidth even before the limits were manipulated, so in the real system the new bins at the bottom of the temperature bandwidth are probably not empty. According to the SOC definition, these devices would have a negative energy content. Since we do not allow for negative energy content in the Markov model, these heaters are not accounted for in the simulation. The natural behavior of unswitched heaters without water draws is a “circulation” (see fig.2). Also with switching and water draws, this still represents the main dynamics of the population. Accordingly, the fact that the markov matrices were created on the basis of an unswitched population, will lead
Figure 5: Increase of set point temperature of 0.2 for 15 min (6h until 6h15min) using the 2nd SOC definition. After lowering the set temperature back to nominal level, the oscillation starts due to the created empty bins.

Figure 6: Increase of set point temperature of 0.5 between the hours 1 and 3 using SOC calculation type one. The red graph represents the TPC of the population modelled with the MTMs. The blue graph depicts the TPC of the “real” population.

to a circulation of these empty bins within the state vector, resulting in a strong oscillation of the simulated TPC. Due to these problems, the simulation is not able to follow the real system anymore, after the bandwidth temperatures have been manipulated. This oscillation can be seen in figure 5. The simulation is able to follow the real heaters power demand even though a small error builds up, until the set temperature is set to the default value again by sending a negative signal.

Results from simulations using this switching strategy can also be seen in figures 6, 7 and 8. It becomes obvious, how the definition of the SOC in the EWH has a big influence on the simulation performance.
Figure 7: Increase in setpoint temperature of 0.5 between the hours 1 to 3 for the SOC calculation type 3. It overestimates the amount of boilers in the bins with low temperature significantly and is therefore not able to predict the change in TPC correctly.

4.1.1 SOC definition 1

The 1st SOC calculation in fig. 6 uses the average temperature, but allows no "negative" energy content, so layers with a water temperature below $T_{\text{min}}$ are treated as if they were at $T_{\text{min}}$. This leads to an overestimation of the energy content in the boiler and means, that in the simulation, there are fewer devices in the lower half of the off-states than in the real system. These are the boilers, which are switched on, when sending positive control signals like in this example, so the TPC change is lower in the simulation than for real boilers as can be seen in the graph.

4.1.2 SOC definition 3

Figure 7 shows the result from the same signal, using the 3rd definition for the SOC. This calculation allows negative energy content in the single layers, but no negative energy content in total. So the SOC can not drop below zero, whereas single layers can have negative energy content. This calculation has the disadvantage, that the cold layers on the bottom of the heater have the same influence on the SOC as the hot layers on top. It is rather unlikely for a heater to reach the hotter bins since the cold layers on
the bottom keep the average quite low, since the water inflow temperature is only 14.4 °C. In the real heaters, only the temperature at the thermostat is relevant to calculate whether the device is going to switch on or off. Often situations occur, especially after small water draws, when the lower water layers are still quite cold, but the temperature at the thermostat is still in the wanted range. In figure 1 such situations can be seen for example in the 7th hour. This fact leads to an overestimation of the boilers in the lower temperature range of the bins and hence, when the set temperature is increased the simulation overestimates the jump in TPC (see fig. 7), as there are more devices turning on.

4.1.3 SOC definition 2

Figure 8 shows the result for the same switching signal for the SOC calculation 2, which is fully based on the temperature of the layer with the thermostat. It is able estimate the change in TPC correctly and is even able to follow the real systems TPC for several minutes after the set point manipulation. Figure 9 and 10 show two more simulation results also for SOC type 2. In fig. 9 we decrease the set point temperature. Also here the simulation is able to predict the TPC for a couple of minutes, but the two created empty bins then “travel” in the on state bins, lowering the amount of heating boilers so that the power demand can not be predicted anymore. In the real system, these two bins would never be empty since, as mentioned before, there are always boilers with temperatures below $T_{m,in}$ due to water draws. In fig. 10 the set temperature is increased for 15 min. The simulation is able to predict the TPC for this time period. When the temperature is switched back again, it fails due to the effect explained above.

4.2 Probabilistic Switching

This control strategy works similar to the control strategy in [7]. For every time step $t$ a signal $u(t) \in [-1, 1]$ is sent to the population. The boilers draw a random number $rand \in [0, 1]$ and whenever $rand < |u(t)|$, the boiler
Figure 8: Increase of set point temperature of 0.5 between the hours 1 and 3 using SOC calculation type two. With this SOC definition, the model is able to predict the resulting change in TPC accurately.

Figure 9: Decrease of set point temperature of 0.1 between hours 7 and 8 using SOC calculation type two.

Figure 10: Increase of set point temperature of 0.4 for 15 min starting at hour 2 and using SOC calculation type two.
switches, on or off, depending on whether the signal is positive or negative. We only switch boilers, which are not in the bins 1 or \( \frac{N}{2} + 1 \), in order not to switch devices, with a temperature outside the set temperature bandwidth. Also in the detailed stratified model we do not switch boilers, which are in those bins, in order to switch a similar amount of heaters in both simulations. Hence also in the stratified model, the heaters are “aware” of their SOC. This approach leads to different TPC graphs for the different SOC definitions also for the detailed model. This assumption is to a certain degree unrealistic, since real heaters are only aware of the temperature at their thermostat. A more realistic approach could be to introduce additional bins for heaters with a SOC below zero and for heaters with a SOC above 1.

For the markov transition matrices, the switching is modeled, with the equation:

\[
x_{t+1} = Mx_t + Bu_t
\]

where \( t \) denotes the time step and \( B \in \mathbb{R}^{N \times \frac{N}{2}} \)

\[
B = \begin{pmatrix}
1 & 0 & \ldots & 0 & -1 \\
0 & 1 & \ldots & 0 & -1 \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
0 & \ldots & 0 & 1 & \ldots \\
-1 & 0 & \ldots & 0 & 1 \\
\end{pmatrix}
\]

Again the ability to predict the TPC is very different for the different SOC definitions. In the case of set temperature manipulation, only the set temperature was changed and the regulation in the heaters, depending on the
thermostat temperature is decisive for the switching. In this section we switch single heaters on and off, not depending on their thermostat temperature. In Figures 11, 12 and 13 one can see the results for a negative (left) and a positive (right) switching signal at 500s for the three different SOC definitions.

4.2.1 SOC definition 1

As already mentioned the first SOC definition (fig. 11) overestimates the amount of energy stored in the boilers and therefore also overestimates the amount of boilers, which are currently heating and have an average temperature within the temperature limits. So when a negative signal to turn heaters off is sent, the amount of boilers, that follow this signal is higher than in the detailed simulation. Also in the detailed simulation boilers outside the temperature bandwidth are turned off, but after the switching action, these heaters are again controlled by their internal control, and if they were laying outside the temperature bandwidth, they turn on again. When switching boilers on, this model estimates more devices to lay within the temperature boundaries as well and accordingly, more heaters are switched. In both cases the simulation overestimates the change in TPC significantly.

4.2.2 SOC definition 2

With this control strategy the SOC definition 2, which is entirely based on the temperature at the thermostat has a drawback, even though the simulation is able to predict the change in TPC with a relatively small error. The definition does not account for the energy content of a boiler but only its temperature at the thermostat and therefore can not describe dynamics, which have to do with the temperature distribution within a heater. For this SOC calculation it makes no difference, whether the layers below the
thermostat are filled with cold water due to a recent water draw, or with hot water due to heating as long as the temperature at the thermostat is the same. In the following we want to explain, how that can be very important, and how it leads to a failing calculation of the TPC development of the population. Boilers that were externally switched on show a rather homogeneous temperature distribution within the EWH (e.g. compare 1st h in fig. 1). If a boiler with a homogenous temperature distribution is switched on, the temperature around the thermostat does not instantaneously increase. It takes a while, until the hot water from the heating element rises, and reaches the thermostat. Within this time, the heater stays within the same bin. In the simulation though, the Markov matrix was based on a scenario, when heaters were not externally switched, and heaters in the on-state showed a strong temperature gradient. So when in the simulation, one heater is being switched from the off to the on state, the Markov matrix models its behavior as if the heating element of this device had already been running for a while. The switching probability to a bin with higher temperature is much higher than in the real system. This leads to the fact, that the TPC of the population in the Markov model after being switched on, drops much faster, than in reality (compare fig. 12 - right figure). The same effect appears when heaters are being switched off. The Markov matrix was again based on a situation, where all heaters in the off state were cooling down, so it can not model the behavior of a heater, that was heating before, and therefore has a
strong temperature gradient, and is now cooling down. For a real heater it takes thereafter longer to reach the lower temperature limit, than in the simulation. Hence the power consumption of the simulation increases faster after switching heaters off, than the real population. (compare fig. 12 - left figure)

4.2.3 SOC definition 2

Fig. 13 shows the result of the single switching with a neg. signal of $-0.5$ (left) and a pos. signal of 0.5 (right) at $t = 500s$ for the 3rd SOC calculation, that takes the entire energy content of the EWHs into account. With this definition, the simulation still overestimates the boilers, that are available for switching, especially for positive switching signals, but compared with the first SOC definition, this effect is much weaker. Since the states of the EWHs are calculated as an average of all water layers, this way we also account for temperature gradients within the heaters and no significant error builds up with time in addition to the initial switching error.
Figure 13: Switching signal at $t = 500$ s for SOC definition 3, left: negative signal $u_{500} = -0.5$, right: positive signal $u_{500} = 0.5$

4.2.4 Random Continuous Switching

The figures 14, 15 and 16 show how the simulation performs for a random normally distributed input signal with mean value zero and a standard deviation of 0.1 (truncated to mean value ± 3 STD). On the left, we plot the power demand (PD) of the real system in blue and the PD calculated with the markov model in red for the 3rd SOC definition. Up to 25% of the population are switched in one time step. The simulation is more or less able to follow the real power consumption. All three SOC definitions are able to predict the TPC of the heaters with an error below 15% for the first 10 min of the simulation. After that, the SOC definitions 1 and 2 fail and errors rise to up to 45%. Using the 3rd definition allows to predict the TPC of the population for nearly an hour, with an error of about 20%. After this time, also this accumulated error starts to increase significantly. As can be seen from the left graph in fig. 14, the error of the 3rd SOC definition is constantly negative, the markov chain-based model costantly overestimates the power consumption. This effect can be explained in a combinations of the errors from fig. 13. Since the overestimation of the PD is greater, when boilers are switched on, than the underestimation, when boilers are switched off, the combination of both results in an overestimation, when both positive and negative signals are sent.

We analyze the case, where only negative and positive signals are sent.
Figure 14: Left: TPC of real system and simulation with the 3rd definition of SOC, switching signal random (normal distribution: mean: 0, STD: 0.1).
Right: errors for the same simulation for the three different SOC definition.

Figure 15: Left: TPC of real system and simulation with the 3rd definition of SOC, negative random switching signal (absolute value of a normal distribution: mean: 0, STD: 0.1). Right: errors for the same simulation for the three different SOC definition.

separately. The different cases correspond to a controller trying to decrease or increase the TPC for a certain amount of time. Fig. 15 shows the simulation where only negative switching signals are sent. As can be seen on the error graph, the 1st definition of SOC fails entirely to follow the TPC of the real system, and when using the 2nd definition the above described error leads to a failure of this simulation in the long run. By using the 3rd definition of SOC we are able to predict the power consumption of the population for an entire hour with an error of about 5% only.

Fig. 16 shows the same case for positive switching signals. Again using
Figure 16: Left: TPC of real system and simulation with the 3rd definition of SOC, positive random switching signal (absolute value of a normal distribution: mean: 0, STD: 0.1). Right: errors for the same simulation for the three different SOC definition.

the 3rd SOC definition leads to the best result. But the results for positive switching signals are as expected not as good as for the negative signals. As can be seen in this graph, the positive switching also for the 3rd definition fails to follow the TPC of the population after about 30 minutes. This is also the reason, in the case of negative and positive switching (compare fig. 14) the simulation fails after about an hour.

4.3 Prediction of Power Consumption

Based on the results from this chapter, we introduce a strategy to control the TPC of the load. Using the markov model with the 3rd SOC definition, we intend to increase and decrease the TPC by 1 MW for one hour. Therefore we first simulate the power consumption without any switching with the markov chain-based model. This model then predicts, how many EWH are currently heating and how many are in the off-state. Using this prediction we derive a input signal, which increases or decreases the TPC in this model by exactly 1 MW. We then run the real system with the same input signal and check, how accurately we meet the 1 MW change.

Figure 17 and 18 show the TPC for the intended power increase and decrease of 1 MW during one 1 hour. Since, when switching boilers, the simulation overestimates the power change, the change of the real system with a switch-
Figure 17: Left: TPC of unswitched and switched system as well as the
difference between the two. Right: error of the intended power increase.

Figure 18: Left: TPC of unswitched and switched system as well as the
difference between the two. Right: error of the intended power increase.

...ing signal calculated on the basis of the simulation is lower. Therefore we do
not meet the intended 1 MW entirely. The maximal error for this switching
strategy rather big 37% but since the error is due to the static error which
results from the definition of SOC, with more investigation this error could
probably be decreased.

5 Conclusion

In this project, we developed a method to model and predict the behavior
of a population of EWHs using the Markov chain model, which is on the
one hand computationally efficient and on the other hand includes effects
resulting from the very complex dynamics of such thermal devices. In order to account for the unpredictability of customer behavior and the heterogeneity of a device population, parameters such as water draws and heater power have been simulated as random variables. It has been shown, how the performance of control strategies is very dependent on the definition of SOC used to develop the markov model. Finally we presented, how this method could be used to control the total power consumption of a population for demand response purposes in an open control loop and with very low communication requirements, while respecting the temperature bandwidth constraints and therefore the customer comfort. We found that the combination of a markov chain model based on a detailed involved model look promising but in order to improve the results obtained in this work and to move on to more realistic scenarios, one would have to investigate further, how such a model acts when the water draw probability distribution is not constant as we assumed. Since we assumed a open loop system, very little communication requirements would be necessary once a similar model has been calculated, but a big amount of information was used in this work, in order to calculate the markov matrices in the first place. A reliable basic model of EWHs is crucial for that and a real population of electric water heater might not be well enough represented by the model used here.
References


Innovative Smart Grid Technologies Europe (ISGT Europe), Berlin, Germany, October 2012, pp. 18.

