Solution space of optimal heat pump schedules

David Kraljić[1], Miha Troha[1]

[1]*COMCOM d.o.o Idrija, Lapajnetova 29, Idrija, Slovenia*
E-pošta: david.kraljic@comcom.si

Abstract

We study the space of optimal schedules of a heat pump with thermal energy storage used in heating a residential building. We model the heating system as a Mixed Integer Linear Program with the objective to minimise the cost of heating. We generate a large number of realistic daily heat demands and calculate the optimal schedule for the heat pump. In addition to cost savings stemming from optimal running, we find that the space of optimal schedules is large in practice, even for the simplest model of the heating system we use, and that the optimal schedules are difficult to reproduce with statistical models. These findings strengthen the case for the use of mathematical optimisation in real-life applications.

1 Introduction

Mathematical optimisation is a generic framework for formulating and solving problems where some goal (or objective) needs to be optimised (maximised or minimised) given some constraints. It has been applied successfully across many domains such transport and logistics[1], power plant dispatch and electric grid balancing [2], economics [3], machine learning [4], biology [5] and many others.

With the advent of smart buildings and smart devices[1] which enable users to actively control them, the problem arises of how to do that optimally. Mathematical optimisation framework is well suited to tackle this problem. The goal of a smart device owner is to minimise utility costs (or equivalently maximise profits) subject to the physical characteristics[2] of a smart device.

Usually, in practice, smart devices are not actively controlled by mathematical optimisation algorithms, but instead they follow simple rules or some predefined daily schedules. The question appears whether modelling the devices and surrounding processes to produce their optimal schedules adds any value in comparison to schedules or logic fixed in advance. The aim of this contribution is to study this question in a simplified but realistic heating system for an actual residential building. We focus on the properties of the solution space for optimal running of the heating system given the realistic demand for heat for the building.

---

[1]Examples include air conditioning units, heat pumps, charging electric cars.

[2]Examples include size, losses, speed of charging or discharging etc.

In Section 2 we describe the devices and other data, in Section 3 we discuss the mathematical and statistical modelling used, in Section 4 we show the results and discuss them.

2 Set-up, Data and Benchmark

The optimisation problem that we address in this contribution aims to minimise costs of operating a heat pump by considering physical constraints of the heat pump and the requirement that enough heat must be generated to satisfy heat demand of the building. Data for costs, physical constraints of the heap pump and demand of the building were determined based on realistic historical data of a residential building[3].

2.1 Costs

The main cost of heating comes from the cost of electric energy which is split into two tariffs: lower cost tariff (MT) applicable between 10pm and 6am and a higher cost tariff (VT) applicable between 6am and 10pm. We assume that VT stands at 1.5 of MT, which is a realistic split given the current electricity prices.

2.2 Physical constraints

The heat pump is capable of providing roughly $160\,\text{kW}$ of heat at a coefficient of performance[4] ~ 1.6. The heat storage is capable of storing up to $200\,\text{kWh}$ of energy.

2.3 Demand

We use real heat demand data which is measured by a calorimeter at the exit of the heating system. A typical daily demand profile based on 480 days during three years of heating season (mid October to April) is depicted in Fig. 1.

2.4 Benchmark

The heat pump operates in the 'on/off' fashion, that is, either it runs at maximum power or is idle at low power. Schedules produced by the mathematical optimisation algorithms are to be benchmarked against actual historical

---

[3]Retirement home 'Dom Sv. Lenart', Lenart, Slovenia

[4]The coefficient of performance determines the heat generated by one unit of electrical power.
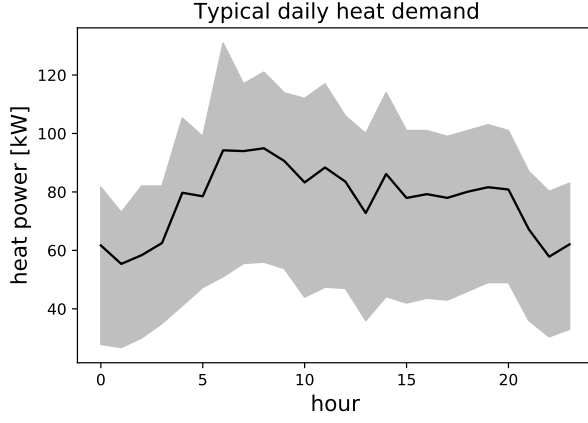
Figure 1: Typical daily heat demand. The shaded region around the mean line corresponds to area between the first and third quantile.

schedules of the heat pump. For an example of a historical schedule, as captured by a power meter located at the heat pump, see Fig. 2.
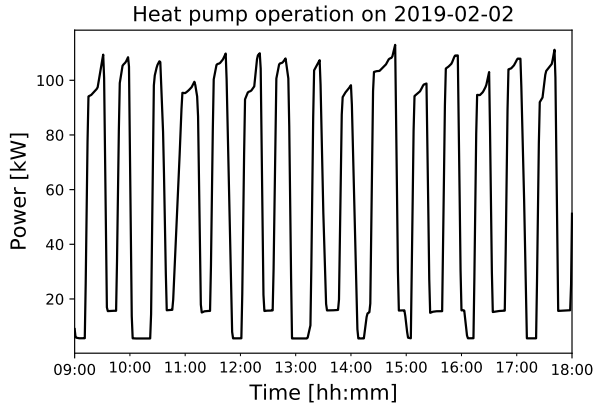

Figure 2: Typical operation of the heat pump.

3 Methods

3.1 Optimal schedule of the heat pump

We obtain the optimal daily schedule for the heat pump by building an optimisation model for the heating set-up with the objective to minimise the cost of operation. We follow the modelling of [6], [7].

Our aim is to evaluate the variability of optimal solutions given the heat demand. Therefore, we simplify a realistic description as far possible without losing the crucial elements of the set-up - that is the variation in schedules is only determined by the variations in heat demand and not in the variation of the modelling parameters. Thus, we neglect the variable, fixed, and cycling costs of the heat pump as well as heat losses in the storage. We also assume the coefficient of performance is constant throughout the day (even though it varies with outside temperature). We also model the heat pump as

only being on at maximum power or off. Finally, we chose hourly granulation for all the variables and parameters in our model. This means that the solution space for the heat pump can be encoded as a sequence of zeros and ones of length 24, corresponding to heat pump either running or not in a specific hour. The size of the solution space is 2^{24}.

The optimisation problem we solve reads as follows:

$$\min_{P_t, \overline{P}_t, S_t} \quad \sum_{t \in T} P_t^{\text{HP}} \cdot \pi_t^{\text{elec}} \tag{1}$$

$$\text{s. t.} \quad S_{t+1} = S_t + \text{cop}_t \cdot P_t^{\text{HP}} - \text{hd}_t \ t \in \overline{T} \tag{2}$$

$$P_t^{\text{HP}} = p^{\text{max}} \cdot \overline{P}_t^{\text{HP}} \ t \in T \tag{3}$$

$$s^{\text{min}} \leq S_t \leq s^{\text{max}} \ t \in T \tag{4}$$

$$P_t, S_t \in \mathbb{R} \ t \in T \tag{5}$$

$$\overline{P}_t \in \mathbb{B} \ t \in T \tag{6}$$

where P_t^{HP} is the heat pump power, π_t^{elec} is the electricity price, S_t is the heat energy stored, cop_t is the coefficient of performance, hd_t is the building heat demand, p^{max}, s^{max}, s^{min} are the power and energy limits, and $\overline{P}_t^{\text{HP}}$ is the binary decision variable denoting whether the heat pump is on or off. Time sets T and $\overline{T}$ are defined as $\{0, 1, ..., 23\}$ and $\{0, 1, ..., 22\}$.

The objective in Eq. 1 is the cash paid for the electricity consumed to heat the building, which we minimise[5]. Eq. 2 describes the energy balance for the heat storage. The heat pump is constrained to be either off or on at max power whcih is enforced by Eq. 3. The storage limits are set by Eq. 4.

We use the following parameters: $p^{\text{max}} = 100\,\text{kW}$, $s^{\text{max}} = 200\,\text{kWh}$, $s^{\text{min}} = 0\,\text{kWh}$, $\text{cop}_t = 1.6$. For all runs we set the initially stored heat, $S_0 = 100\,\text{kWh}$.

3.2 Daily heat demand profiles

We are interested in running the optimisation problem from Sec. 3.1 a large number of times to explore as much of solution space as computationally possible. For a 24 hour granulation of the daily optimal schedule of an 'on/off' heat pump the total number of potential solutions is $2^{24} \approx 1.7 \cdot 10^7$. Since we only have the actual data for 480 days during the heating season, we need to generate synthetic but realistic heat demand profiles.

Heat demand for each hour of the day is a random variable with a mean and a variance in the simplest terms. Because the demand is bounded from below with 0, a simple normal distribution cannot be used to model it. We chose a lognormal distribution[6] which is bounded by 0 and is also only described by a mean and variance. We illustrate our modelling for a specific hour in Fig. 3.

Because the heat demand from hour to hour within a single day is not independent, we need to also model these hourly cross-correlations. The size of these is illustrated by the covariance matrix for the hourly demand in Fig. 4.

---

[5]We add to electricity prices a very small random increase with each hour of the day, thus ensuring that the optimal objective has a unique solution that prefers taking decisions earlier in the day.

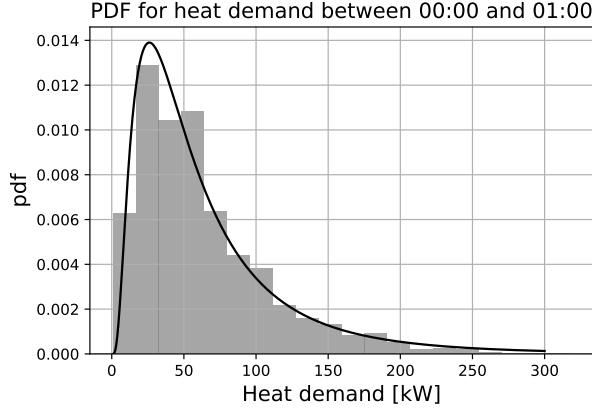[6]logarithm of a random variable is distributed normally

Figure 3: Shaded is the empirical pdf for the heat demand between hour 00:00 and 01:00. The line is our best fit lognormal model.
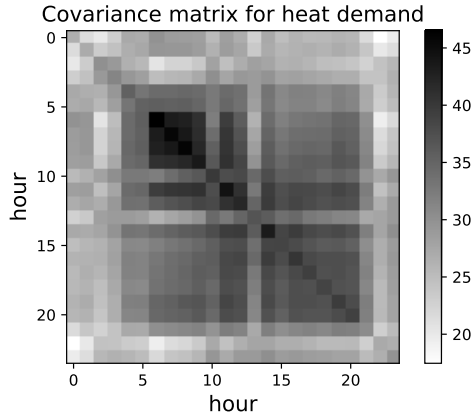

Figure 4: Empirical covariance matrix for the hourly heat demand.

Our final model for the daily demand is the lognormal distribution derived from empirical mean (Fig. 1) and the covariance matrix (Fig. 4) of the actual data.

4 Results & Discussion

We have generated $100\,000$ daily heat demand profile samples and solved the optimisation problem (Sec. 3.1) the same number of times, obtaining $100\,000$ optimal solutions.

4.1 Size of solution set

To determine if our solution set is localised in the space of all possible solutions, we compare it to the baseline case – the set where all 2^{24} solutions for the heat pump schedule are possible and equally probable. For such a uniformly distributed set a single solution appears on average $\lambda = n/2^{24}$ times, where $n = 100\,000$ is the number of generated solutions. Then it follows that the *number* of times a solution appears is Poisson distributed with mean λ.

In Fig. 5 we plot the fraction of solutions we found against the multiplicity of the solutions. That is, we plot the fraction of solutions that appear at most once, twice, three times etc., in our solution set. We also plot the expectations stemming from the uniform distribution of solutions, as explained above.

We see (Fig. 5) that the set of optimal schedules we obtain has solutions appearing more often and at higher multiplicities than expected if the baseline case would be true. Note, that the fraction of solutions that appear only once in the baseline case is practically all of them (0.9999823). This suggests that the solution space to our optimisation problem is localised and not completely random. This is expected, as the variability in the heat demand is not completely random, but constrained.
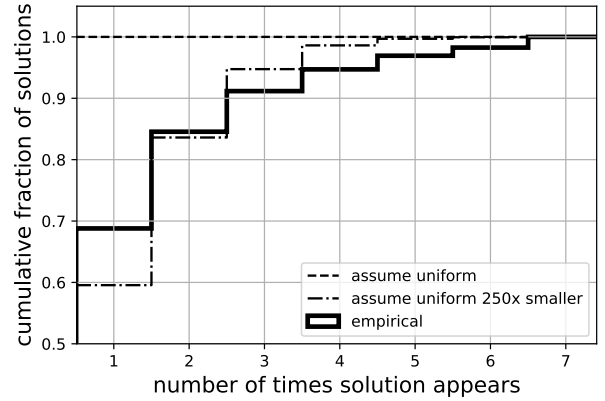

Figure 5: Empirical multiplicities of our solution set compared to expectations if solutions were uniformly distributed in the space of all possible solutions.

However, the fraction of solutions that appear once, twice etc., is still small. For example, there are no solutions that would appear more than 7 times in our set of $100\,000$, suggesting that different demand profiles lead to very different optimal solutions.

If we assume the solution set to our optimisation problem is uniformly distributed (but smaller) then by fitting the Poisson distribution to our empirical data in Fig. 5 we deduce that the solution space is roughly 250 times smaller than 2^{24} ($\approx 7 \cdot 10^4$). This means a single solution to our optimisation problem is expected to repeat roughly every $7 \cdot 10^4$ daily runs or 180 years.

4.2 Complexity of solution set

We measure how easy it is to describe the solutions to our heat pump optimisation problem by training machine learning models to predict optimal heat pump schedules given the heat demand and evaluating their performance.

The generated dataset of $100\,000$ heat demand profiles forms the input set with the corresponding optimal solutions being the output set. We split the dataset into train, validation, and test subsets. We selected tree-based (Decision tree, Random Forest) models and linear models (Logistic regression)[7]. All model parameters were fine-tuned on the validation set, with the final performance

---

[7]These models are selected as they can be expressed easily in code and thus can be deployed to control the heat pump

evaluated on the test set. We also evaluate a trivial model which produces random schedules.

We measure how well machine learning models predict the optimal solution by counting the average number of hours that differ between true and predicted schedules. In other words, we count the minimum number of errors that transform the true optimal solution to the one predicted by a machine learning model (known as the Hamming distance[8] in coding theory). We summarise the results in Table 1.

Table 1: Performance of machine learning models predicting optimal solutions.

Model	Best param.	Hamming error [hour]
Trivial model	-	10.1
Decision Tree	max_depth = 10	5.94
Random Forest	max_depth = 10 num_trees = 400	5.72
Log. Regression	-	6.36

Solving the optimisation problem in Sec. 3.1 performs a mapping from the daily heat demand profile to the optimal schedule. Mimicking this mapping with either linear or tree-based models does not work well since on average the predicted optimal solutions are still wrong on about 6 hours of the day.

4.3 Cost savings

Currently (i.e. in our benchmark data), the heat pump operates in short bursts of 15-30minutes (e.g. Fig. 2), and follows the real-time demand for heat almost exactly without utilising the large heat storage available. We evaluate the cost of this current strategy on the 480 days of heating seasons using the actual data. We run the simplified optimisation problem from Sec. 3.1 for the same days and plot the differences in Fig. 6
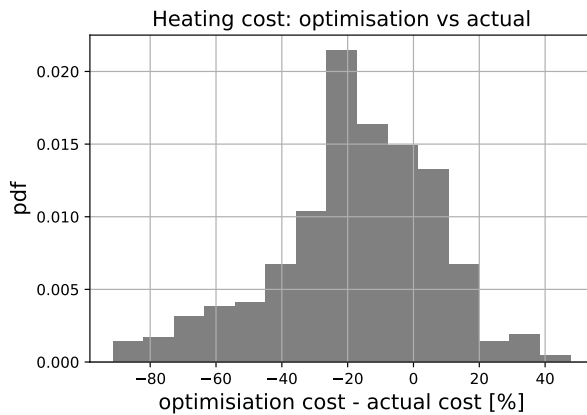

Figure 6: Percentage difference between the costs of schedules produced by our simple optimisation model and actual costs.

We note that even the simple modelling of the heating system we use manages to save on average 15% in heating costs. We expect that formulating the optimisation problem with higher granularity (sub-hourly), as well as using time-varying COP, would make the gains even bigger, as the problem would be less constrained.

5 Conclusion

We studied the set of optimal solutions for a simplified heating system consisting of a heat pump&storage satisfying realistic heat demand profiles. The set-up and data are based on an actual residential building. We find that the solution space for optimal heat pump schedules is smaller than the maximum solution space, but still large enough to justify using the optimisation framework in practice. We also show that the optimal schedules, once they are known, are not easy to predict using statistical models. Finally, we show that even simplistic modelling of the optimisation problem can lead to significant cost savings.

Acknowledgement

References

[1] Pursula, Matti & Niittymäki, Jarkko. (2001). Mathematical Methods on Optimization in Transportation Systems. 10.1007/978-1-4757-3357-0.

[2] Lin, J., & Magnago, F. H. (2017). Electricity markets: Theories and applications. Somerset: John Wiley & Sons, Incorporated.

[3] Mikulás Luptácik, 2010. Mathematical Optimization and Economic Analysis, Springer Optimization and Its Applications, Springer, number 978-0-387-89552-9.

[4] Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright. 2011. Optimization for Machine Learning. The MIT Press.

[5] Reali Federico, Priami Corrado, Marchetti Luca. 2017. Optimization Algorithms for Computational Systems Biology, Frontiers in Applied Mathematics and Statistics. 10.3389/fams.2017.00006

[6] Renaldi, Renaldi & Kiprakis, Aristides Friedrich, Daniel. (2016). An optimisation framework for thermal energy storage integration in a residential heat pump heating system. Applied Energy. 10.1016/j.apenergy.2016.02.067.

[7] Schütz, Thomas & Harb, Hassan & Streblow, Rita Mueller, Dirk. (2015). Comparison of models for thermal energy storage units and heat pumps in mixed integer linear programming.

[8] R. W. Hamming, Error detecting and error correcting codes, The Bell System Technical Journal, 10.1002/j.1538-7305.1950.tb00463.x.