

# Utilizing Historical Operating Data to increase Accuracy for Optimal Seasonal Storage Integration and Planning

**Maximilian Sporleder<sup>a,b</sup>, Michael Rath<sup>a,c</sup>, Yuwei Xu<sup>a</sup>, Mathias van Beek<sup>d</sup>,  
and Mario Ragwitz<sup>a,b</sup>**

<sup>a</sup> Fraunhofer Research Institution for Energy Infrastructures and Geothermal Systems (IEG),  
Cottbus, Germany, maximilian.sporleder@ieg.fraunhofer.de

<sup>b</sup> Brandenburg University of Technology Cottbus-Senftenberg (BTU CS), Cottbus, Germany

<sup>c</sup> Bochum University of Applied Sciences, Bochum, Germany, michael.Rath@hs-bochum.de

<sup>d</sup> Fraunhofer Institute for Environmental, Safety and Energy Technology (UMSICHT),  
Oberhausen, Germany, mathias.van.beek@umsicht.fraunhofer.de

## Abstract:

Policies reasoned by global climate change and increasing commodity prices due to the international energy crisis force district heating providers to transform their assets. Pit thermal energy storage combined with solar energy can improve this transformation process. Optimal energy planning of district heating systems is often achieved by applying a linear programming model due to its fast computing. Unfortunately, depicting those systems in linear programming requires complexity reduction. We introduce a method capable of designing and operating the system with the complexity increase of considering the top and bottom temperatures of the pit thermal energy storage in linear programming.

Firstly, we extract and clean data from existing sites and simulations of seasonal storages. Secondly, we develop a polynomial regression model based on the extracted data to predict the top and bottom temperatures. Lastly, we develop a mixed-integer linear programming model using the predictions and compare it to existing sites. The model uses solar thermal energy, a pit thermal energy storage, and other units to meet the demand of a district heating system.

The polynomial regression results show an accuracy of up to 92 % with only a few features to base the prediction. The optimization model can design the storage and depict the correlation between decreasing specific costs and thermal losses due to an increasing volume. The control strategy of the heat pump requires further improvement.

## Keywords:

MILP; Seasonal Storage; Optimization; District Heating; Design.

## 1. Introduction

The energy transition to a carbon-free system is one of the key challenges in the 21<sup>st</sup> century. Therefore, district heating systems (DHS) need renewable options and flexibility to reduce the impact of global warming. [1] predicts an expansion of DHS and emphasizes the importance of seasonal storages. Seasonal storages could enhance the integration of renewables like solar thermal energy and increase the system's flexibility. [2] identifies pit thermal energy storages (PTES) as economically feasible compared to other technologies. Combining different renewable technologies and seasonal storages supplying a DHS and the connected consumers is a tested concept in Denmark [3]. However, the planning process of a DHS with different technologies and different locations can be complex and overwhelming. In order to support decision-making processes for planners, we analyze historical and simulation data from seasonal storages. We develop a method to design a DHS with a PTES and other technologies inside a discretized thermal grid.

Figure 1 illustrates the concept of storing solar thermal energy in a PTES and unloading it through a heat pump into a district heating network. The goal is to design and operate energy converters and storages connected to the grid by modeling pipes, consumers, and production units. This approach tries to model a district heating network with all components and allows the integration of PTES in this system. The main research question is how a PTES integration into a DHS with different components can be optimized.

## 1.1. State-of-the-art

In recent years, several different strategies to model PTES and DHS, design and operate them, have been introduced (see Table 1). Most studies focus on a detailed depiction of the storage to calculate the operation and thermal losses due to charging, discharging, and environmental temperatures. A more in-depth review of those studies is given in the following paragraphs.

Appropriate simulation tools are helpful for the project's economic viability. TRNSYS is found to be the most widely used simulation tool for PTES simulations due to its component sets. [2] used TRNSYS to optimize the efficiency by using the data from a pilot solar heating system combining PTES. Results show that control strategies significantly improve heat collection performance and exergy efficiency, and stratification of seasonal storage impacts collection efficiency, especially at the end of the non-heating season [4]. EnergyPlus has some advantages in modeling heating systems on the demand side to satisfy various loads and is focused on building simulations [5]. FLUENT and COSMOL specialize in the PTES device temperature and velocity field analysis rather than system integration. SDH calculation tools is a bundle of tools developed by EU project SDHp2m [6]. It specializes in calculating solar district heating systems [7].

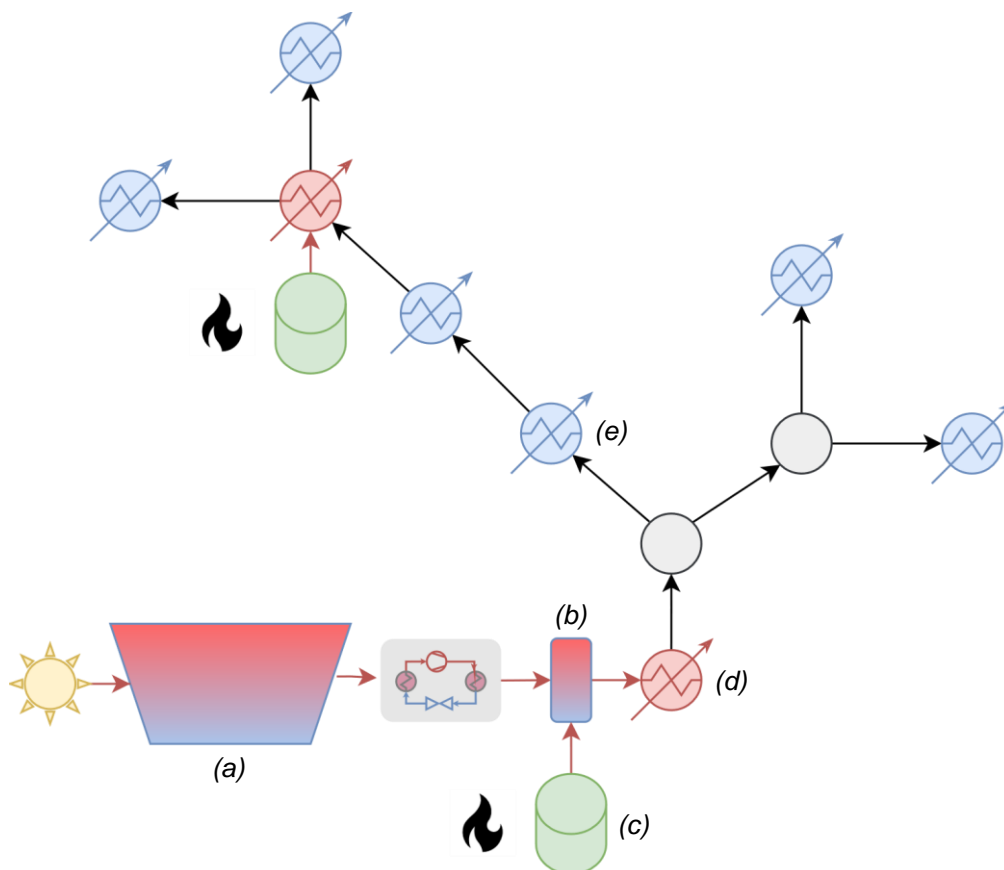


Figure 1: Fictional grid for this research with: a) PTES, solar thermal field, and heat pump as central units, b) buffer storage, and c) biomass boiler as a decentral unit; d) is referred to as a production unit; e) is referred to as consumer.

[8] considers using a biomass trigeneration system and an absorption heat pump to supplement the temperature gap and stabilize the water temperature in PTES at 85 to 90 °C. The model considers the overall efficiency of a power plant and maximizes the net present value to meet the district heating demand [8]. [9] uses TRNSYS to model the combined heating supply system to determine the performance of each plant and the ideal size. [10] studies the Dronninglund water pit thermal energy storage focusing on the balance model of energy and mass flow and providing the charging and discharging data for further research. In winter, the district heating supply temperature in Dronninglund is 75 °C; when the top temperature of the PTES is below the supply temperature, the heat pump works as a compensator. In Marstal, when the PTES top temperature drops below 70 °C, the PTES serves as the heat source for the heat pump to provide the gap between the district heating water temperature and the storage unit [11]. We apply the concept of Marstal using the PTES as a heat source for the heat pump.

The space for the PTES and solar panels affects the economics of the system. [12] compares the parameters of PV panels with thermal solar collectors. Using TRNSYS, three locations across Poland are examined, and the results show solar thermal collectors' advantages in area occupation and economics [12]. [12] uses an

electrode boiler to feed into the PTES. [13] calculates an energy system with additional units and suggests a ratio of the seasonal storage volume to the solar field area of 2 to 3 m. [14] compares distributed and centralized thermal solar collectors in Finland with 1231 MWh yearly heating demand and concludes that the centralized units have a significant advantage in the heat generation costs. The lower costs for central units are caused by decreasing specific investments with increasing dimensions [2].

Table 1. Typical energy system model comparison [15]

Energy system model	Min. Time step	Space range	Related function	Ref.
TRNSYS	Seconds	PTES/System	<ul style="list-style-type: none"> <li>• Economic analysis [9]</li> <li>• Control strategy [16]</li> <li>• Validation [17]</li> <li>• Efficiency [18]</li> <li>• PTES structure optimization [19]</li> <li>• Model development [17]</li> </ul>	[20]
COMSOL	Not available	PTES	<ul style="list-style-type: none"> <li>• Validation [21]</li> <li>• PTES structure optimization [22]</li> <li>• Stratification model [21]</li> </ul>	[23]
FLUENT	Not available	PTES	<ul style="list-style-type: none"> <li>• PTES structure optimization [24]</li> <li>• Stratification model [25]</li> </ul>	[26]
Mathematical model		PTES/System	<ul style="list-style-type: none"> <li>• PTES structure optimization [27]</li> <li>• Validation [28]</li> <li>• Scenario analysis [28]</li> <li>• Stratification model [27]</li> </ul>	
SDH calculation tools	Hourly	System	<ul style="list-style-type: none"> <li>• Combined five mixed system</li> <li>• Economic analysis</li> <li>• Scenario analysis</li> <li>• Efficiency (sizing)</li> </ul>	[7]
PyLESA	Hourly	System	<ul style="list-style-type: none"> <li>• Economic analysis [29]</li> <li>• Control strategy [29]</li> <li>• Temperature dependence for heat pump models [29]</li> <li>• Stratification model [29]</li> </ul>	[29]

The mentioned research gives and models the storage geometry and volume inside the simulation. We set the geometry as a variable and contribute to this extensive research with an approach in mixed-integer linear programming (MILP), including

- dynamic network behavior by depicting temperatures inside the grid,
- designing and controlling PTES in combination with other energy converters and storages located at different positions in the grid,
- introducing realistic PTES temperatures from simulations and measured data to obtain more realistic, dynamic coefficients of performance (COPs) (see Marstal concept),
- modeling and calculating the geometry of the PTES inside the optimization to account for thermal losses, space requirements, and specific investments.

The major problem of designing seasonal storage systems in energy system models based on MILP is the temperature modeling inside the storage – which would introduce nonlinearity – and the lack of information concerning the geometry. Our approach is based on the fact that already-designed systems have been operated for several years. [30] gathers data concerning the ratio between volume and surface, while the storages in Marstal and Dronninglund provide operational data. This information provides a generic approach supporting planners in the design phase of a carbon-neutral DHS.

The paper is structured in the method section, result section, discussion, and conclusion. In the method, we explain how we utilize the data from existing sites to model temperatures of PTES supplied by solar thermal energy. Additionally, we explain how we model the geometry of the PTES leading to specific investments, thermal losses, and space requirements. In the result section, we validate our approach by simulating the temperatures in the PTES and comparing the geometry and other parameters to existing sites. We continue by discussing our results and conclude with the most important findings.

## 2. Method

To find the optimal design for our energy system, we utilize MILP. MILP is a mathematical optimization technique for solving problems where some variables are constrained to be integers. It is used to model real-world problems where decisions are made based on discrete choices [31]. Our approach uses historical data from existing, and simulated PTES projects supplied by solar thermal fields and predicts temperature profiles for the location where the energy system model is applied. Figure 2 gives an overview of the methodology. We train a regression model based on the extracted data to predict the top and bottom temperature. With that temperature information, the MILP model optimizes the heat pump operation, and we also obtain the temperature losses at the surfaces. To create a dynamic PTES surface model, we used existing sites to build a piecewise function in the MILP model that takes the variable volume as an input and delivers the areas of the bottom, top, and sides as an output. The specific investments are modeled analogously. The MILP model includes energy balances around pipes, consumers, energy converters, and storages. The mass flow inside the pipes is estimated a priori, preventing nonlinearity.

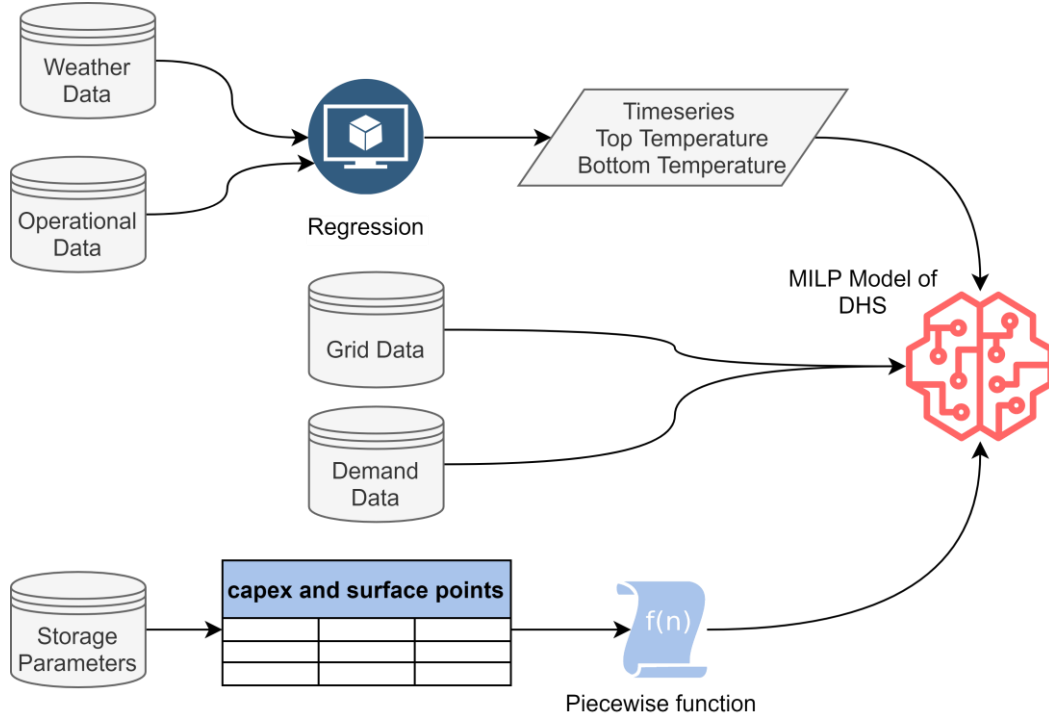


Figure 2: Illustrating the method dividing into the prediction of the temperatures, the piecewise functions for the capex and the surface areas of the storage depending on the volume, and the MILP model for the DHS.

### 2.1. Regression Model predicting PTES Top and Bottom Temperature

For our approach, we assume that the temperatures in the PTES behave similarly independent of the system due to the seasonal storing of solar thermal energy. To examine this assumption, four datasets were used in this research to predict the internal temperature of the PTES. We use the existing sites in Marstal and Dronninglund [3] and the simulation data from Wadelheim [32] and Florence [33]. The temperatures are extracted using the tool WebPlotDigitizer [34].

Many features can affect the temperature in the PTES. First of all, the environmental factors should be considered. The database of photovoltaic geographical information [35] provides the ambient temperature – two meters above the surface – and irradiance. [30] analyzes the slope and azimuth of solar thermal collectors' installation. Based on this work, the slope is set at 35 degrees, and the azimuth is chosen at 180 degrees. The soil temperature is relatively constant according to the monitoring results in [3]. So, the effect of the deep soil temperature on the tank's bottom temperature is neglected in our model. Additionally, the geometry – volume and surface area – of the PTES could affect the top temperature. However, this effect cannot be considered because the volume is a decision variable in the MILP model; therefore, the storage's geometry is unknown when we apply the regression model (see Figure 2). In conclusion, we use the features time, solar irradiance, and ambient temperature at the given location.

Before applying a polynomial regression [36], we restructure the data input to improve the accuracy. The feature time is separated into hour  $t^{\text{hour}} \in [1, \dots, 24]$  and day  $t^{\text{day}} \in [1, \dots, 365]$ . The time is then split into sine and cosine such as  $\sin/\cos\left(2\pi \frac{t^{\text{hour}}}{24}\right)$  and  $\sin/\cos\left(2\pi \frac{t^{\text{day}}}{365}\right)$ . The time is now categorized into four features.

Afterward, we apply the StandardScaler from scikit-learn [37] to all features. Then we use three data sets to train the model and one to test it. To evaluate the performance of the regression, mean squared error, root mean squared error (RMSE), mean average error (MAE), R-squared, explained variation, and accuracy are set as measurements.

## 2.2. The Energy System Model

A description of all used nomenclature can be found in the Nomenclature section after the conclusion. The energy system model is a MILP model minimizing the costs of the system represented by an objective function

$$\min(f^{\text{opex,var}}, g^{\text{opex,fix}}, h^{\text{capex}}). \quad (1)$$

$f^{\text{opex,var}}$  is the variable operational cost for energy converters combining costs for energy carriers, maintenance, and repair.  $g^{\text{opex,fix}}$  are the fixed costs for storages or energy converters for the operation, maintenance, and repair, and  $h^{\text{capex}}$  are the investments. The economic calculation is based on [38]. The calculation is performed for one year in 3-hour timesteps. The model was solved with Gurobi [39].

The mass flow in the network is calculated a priori per day based on a fixed temperature delta of 30 K, meeting the demand of the highest peak in the grid of that day. Additionally, the velocity in the pipes is limited by a minimum pressure loss of 80 Pa/m and a maximum pressure loss of 300 Pa/m [40]. The energy balance for the pipes is formulated as

$$c_p m_a \frac{T_{a,t-1}^{\text{out}} - T_{a,t}^{\text{out}}}{\Delta t} + \dot{m}_{a,t} c_p (T_{a,t}^{\text{in}} - T_{a,t}^{\text{out}}) - U_a A_a^m (T_{a,t}^{\text{out}} - T_t^{\text{soil}}) = 0 \text{ for } a \in Z^{\text{ff}}, Z^{\text{bf}}, t \in \tau, \quad (2)$$

where the first term represents the storage capacity of the pipe, the second term is the enthalpy rate entering and exiting the pipe, and the last term is the loss of the pipe. The energy balance of the consumers is given by

$$\frac{\dot{Q}_{a,t}^{\text{con}}}{\mu^{\text{con}}} = \dot{m}_{a,t} c_p (T_{a,t}^{\text{in}} - T_{a,t}^{\text{out}}) \text{ for } t \in \tau, a \in Z^{\text{con}}, \quad (3)$$

and the energy balance of the production units is formulated analogously with

$$\dot{Q}_{a,t}^{\text{pro}} \mu_a^{\text{pro}} = \dot{m}_{a,t} c_p (T_{a,t}^{\text{out}} - T_{a,t}^{\text{in}}) \text{ for } t \in \tau, a \in Z^{\text{pro}}. \quad (4)$$

The temperatures in the grid are limited by the consumers, with  $T_{a \in Z^{\text{con}}, t}^{\text{in}} \geq T_{a \in Z^{\text{con}}, t}^{\text{min}}$  and a technical limitation between 0 °C and 130 °C.

The heat flow at a production unit is the summation of all heat flows by storages and energy converters at that location  $a$ . Therefore,  $\dot{Q}_{a,t}^{\text{pro}}$  is given by

$$\dot{Q}_{a,t}^{\text{pro}} = \sum_{k \in Z^{\text{conv}}(a)} \dot{Q}_{k,t}^{\text{conv}} + \sum_{k \in Z^{\text{stor}}(a)} (\dot{Q}_{k,t}^{\text{out,stor}} - \dot{Q}_{k,t}^{\text{in,stor}}) \text{ for } t \in \tau, a \in Z^{\text{pro}}, \quad (5)$$

$Z^{\text{conv/stor}}(a)$  is the set of energy converters and storages at that location. The solar thermal field is modeled with

$$\dot{Q}_{k=\text{solar}, t}^{\text{conv}} = \mu_{k=\text{solar}} \gamma_t^{\text{rad}} A^{\text{collector}} \text{ for } t \in \tau, \quad (6)$$

where  $A^{\text{collector}}$  is the variable dimensioned by the optimizer to calculate the area of the collectors.  $\mu_{k=\text{solar}}$  is the efficiency of the solar field here assumed to be 0.5 [41]. The solar field charges the PTES, and a heat pump lifts the temperature of the PTES, if necessary, before injecting the heat into the grid. The energy balance of the PTES can be described as

$$E_{k=\text{PTES}, t} = E_{k=\text{PTES}, t} + \Delta t \left( \mu_{k=\text{PTES}} \dot{Q}_{k=\text{PTES}, t}^{\text{in}} - \frac{\dot{Q}_{k=\text{PTES}, t}^{\text{out}}}{\mu_{k=\text{PTES}}} - \dot{Q}_{k=\text{PTES}, t}^{\text{loss}} \right) \text{ for } t \in \tau. \quad (7)$$

The energy balances for the buffer storages are modeled analogously. The PTES has a cyclic condition where  $E_{k=\text{PTES}, t=\text{start}} = E_{k=\text{PTES}, t=\text{finish}}$ . The PTES energy is limited by

$$E_{k=\text{PTES}, t} \leq V_{k=\text{PTES}} \rho c_p (T_{k=\text{PTES}}^{\text{top, max}} - T_{k=\text{PTES}}^{\text{top, min}}) \text{ for } t \in \tau. \quad (8)$$

The heat flow when discharging the storage can only be directly injected into the grid if the supply temperature is lower than the temperature of the PTES at the top. This is given by

$$\dot{Q}_{k=\text{PTES}, t}^{\text{out,stor}} = \begin{cases} \dot{Q}_{k=\text{PTES}, t'}^{\text{out,stor}}, & T_{k=\text{PTES}, t}^{\text{top}} \geq T_t^{\text{supply}} \\ \frac{\text{COP}_t}{\text{COP}_t - 1} \dot{Q}_{k=\text{PTES}, t'}^{\text{out,stor}}, & T_{k=\text{PTES}, t}^{\text{top}} < T_t^{\text{supply}} \end{cases} \text{ for } t \in \tau. \quad (9)$$

The COP is calculated a priori based on the predicted temperature profile of  $T_{k=\text{PTES}, t}^{\text{top}}$ . The heat losses  $\dot{Q}_{k=\text{PTES}, t}^{\text{loss}}$  correlate with the geometry of the storage and decrease with increasing volume. The heat losses  $\dot{Q}_{k=\text{PTES}, t}^{\text{loss}}$  are calculated identically to [42]. However, in this study, we do not know the size of the surface yet due to the unknown size of the storage. Therefore, we applied a piecewise function [43] calculating the sides, top, and bottom area of the PTES depending on the volume. We use existing sites to get grid points and summarize them in Table 2. We calculate the areas for each storage based on the geometry of an obelisk and assume that the bottom and top areas are quadratic. Furthermore, we assume a standard correlation between the top area side length and the bottom area side length of 78 to 48 due to the detailed information of Dronninglund from [42].

Table 2: PTES information about the geometry and specific investments used for the piecewise functions. [2, 30, 44]

Site	Volume [m <sup>3</sup> ]	Total surface [m <sup>2</sup> ]	Angle [°]	Height [m]	Calculated bottom area [m <sup>2</sup> ]	Calculated top area [m <sup>2</sup> ]	Calculated side area [m <sup>2</sup> ]	Specific investments [€/m <sup>3</sup> ]
Stuttgart	1050	835	45	5	118.4	312.65	403.95	
Ottruggård	1500							150
Eggstein	4500	1924.9	30	9	182.9	482.97	1259.04	113.02
Sunstore 2 Marstal	10000							67
Dronninglund	60000	17076	26	16	2247.49	5934.78	8893.73	38
Toftlund	70000	19204	27	14.5	2826.33	7463.28	8914.4	
Sunstore 4 Marstal	75000	20298	32	16	3174.36	8382.31	8741.33	36
Gram	122000	28893	20	15	3957.6	10450.53	14484.87	34
Vojens	210000							24

### 3. Results

The result section divides into examining the results of the regression model to predict the top and bottom temperature of a PTES charged by a solar thermal field. Afterward, the MILP model uses this profile to design the PTES and other energy converters in a district heating network. The design is evaluated by comparing it to the existing sites of Marstal and Dronninglund. Additionally, we simulate the storage heat losses in Marstal with our predicted temperatures and compare the error. To evaluate the heat pump control, we perform two optimization runs: the first is with the predicted temperature profile, and the second is with averaged temperatures – not using the regression model in the pre-processing. We then re-simulate the actual temperatures in the storage and calculate the electrical demand for the heat pump in both optimization runs. For the re-simulation, we take the energy level  $E_{k=PTES,t}$  and divide by the volume, the specific heat capacity, and the density resulting in the current temperature delta. This temperature delta is added to  $T_{k=PTES}^{\text{top,min}}$  and the COP of the heat pump is recalculated.

#### 3.1. Polynomial Regression Results

Table 3: Polynomial regression results for three sites as training and one site as testing.

Dataset		Polynomial Regression		
train	test	RMSE	MAE	Accuracy
Marstal, Dronninglund, Florence	Wadelheim	0.6336	0.5687	0.5986
Marstal, Florence, Wadelheim	Dronninglund	0.8395	0.6668	0.2953
Dronninglund, Florence, Wadelheim	Marstal	0.5964	0.4906	0.6443
Marstal, Dronninglund, Wadelheim	Florence	0.2896	0.2337	0.9161

Table 3 shows the results for using three data sets as training and one dataset as testing. The accuracies for Wadelheim and Marstal are around 60 % due to the daily fluctuations in the demand. The result for Florence is relatively high with an accuracy of 91.6 %. The prediction accuracy for Dronninglund as testing is relatively low, at 29.5 %. The deterministic solution of the polynomial regression is given by

$$T_t^{\text{PTES,top}} = 4.0255T_t^{\text{ambient}} + 2.2715\gamma_t^{\text{rad}} + 0.2065 \sin\left(2\pi\frac{t^{\text{hour}}}{24}\right) + 2.7253 * \cos\left(2\pi\frac{t^{\text{hour}}}{24}\right) - 23.6073 * \sin\left(2\pi\frac{t^{\text{day}}}{365}\right) - 9.9604 * \cos\left(2\pi\frac{t^{\text{day}}}{365}\right) + 56.6761, \quad (10)$$

where the temperatures have the unit of °C and the irradiance W/m<sup>2</sup>. Before Eq. 10 can be used, the StandardScaler (see method section) has to be applied to every feature.

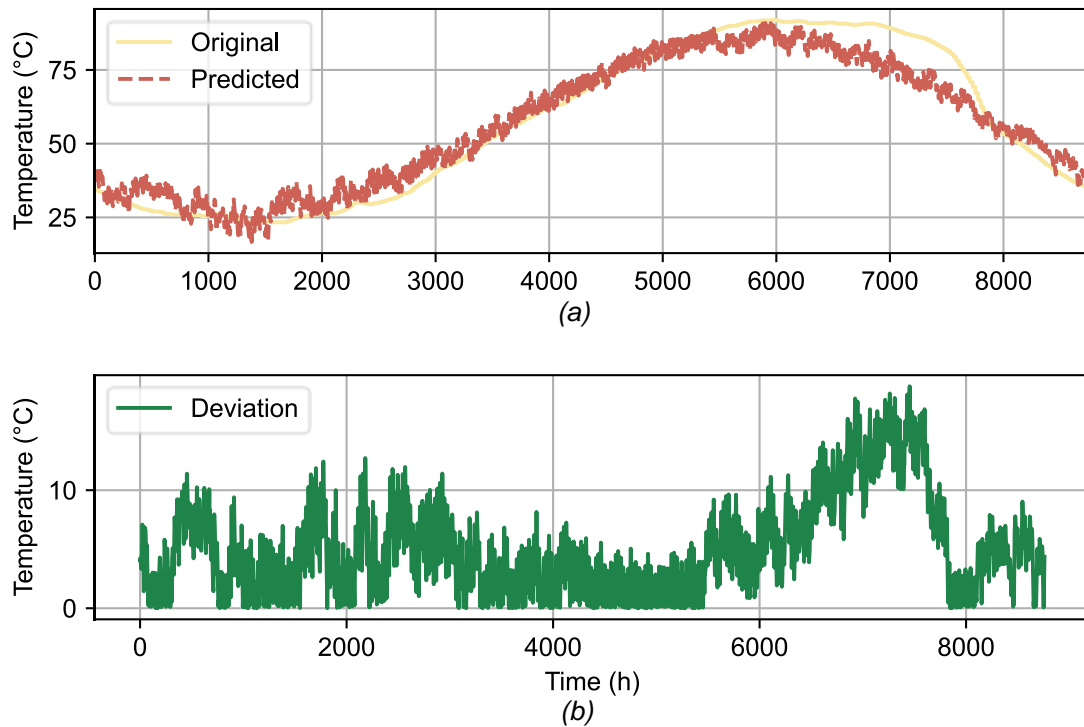


Figure 3: Polynomial regression results for Florence showing a) the simulated top temperature and the predicted top temperature and b) the deviation between the two temperature profiles. [33]

Figure 3 shows the simulated top temperature for the PTES in Florence [33] compared to the predicted temperature. A high correlation can be examined in Figure 3 a) between the two temperature profiles. However, the deviation has a maximum of 18 °C and an average of 5 °C. A deviation of 18 °C in the calculation for the COP of a 1 MW heat pump supplying 90 °C would result in an error of around 100 kW for the electrical input. This error indicates potential weaknesses in the modeling approach. To evaluate the effects on the geometry, we simulated the heat losses of the Marstal storage in 2014 based on the predicted top and bottom temperatures. The simulation results in a heat loss of 2391.85 MWh, causing an error of 17.7 % compared to the measured data [3].

### 3.2. Results for Designing the Energy System

The MILP model is applied to a fictional grid (see Figure 1). The design optimization of a central production unit results in a 250 kW central biomass boiler, a 12555 m<sup>2</sup> solar thermal field, a 26725 m<sup>3</sup> PTES, a 62 m<sup>3</sup> buffer storage, and a 958 kW heat pump. Furthermore, a decentral production unit consisting of a 50 kW biomass boiler is installed. The heat generation costs are 14.5 ct/kWh. The operation and investments of the system are displayed in Figure 4. The costs are based on the year 2020 using the Day-Ahead prices. The solar thermal field mainly loads the PTES in the summer. The heat pump also has the opportunity to charge the PTES. Due to the heat pump's partial load limit of 50 %, the biomass boilers cover peak loads. The buffer storage serves as a day-to-day flexibility supporting the heat pump operation. However, the heat flows of the buffer storage are not included in Figure 4, for clarity and due to the low impact on the energy system. The storage is mainly loaded in spring and summer, and the discharging starts mainly in October. In October, the temperatures in the storage are high enough to inject into the grid directly; therefore, the other units do not operate.

The temperature in the storage affects the electrical input for the heat pump, influencing the operational expenditures (opex). This effect is measured by resimulating the temperatures inside the storage and recalculating the COP. Based on the COP, the electrical power for the heat pump is recalculated. We also perform an optimization run without the predicted temperature profiles assuming a constant temperature inside the PTES. Using the constant temperature inside the PTES leads to a deviation of the electrical input for the heat pump of 45 %, while the predicted temperature profiles of the PTES cause an error of 29 %.

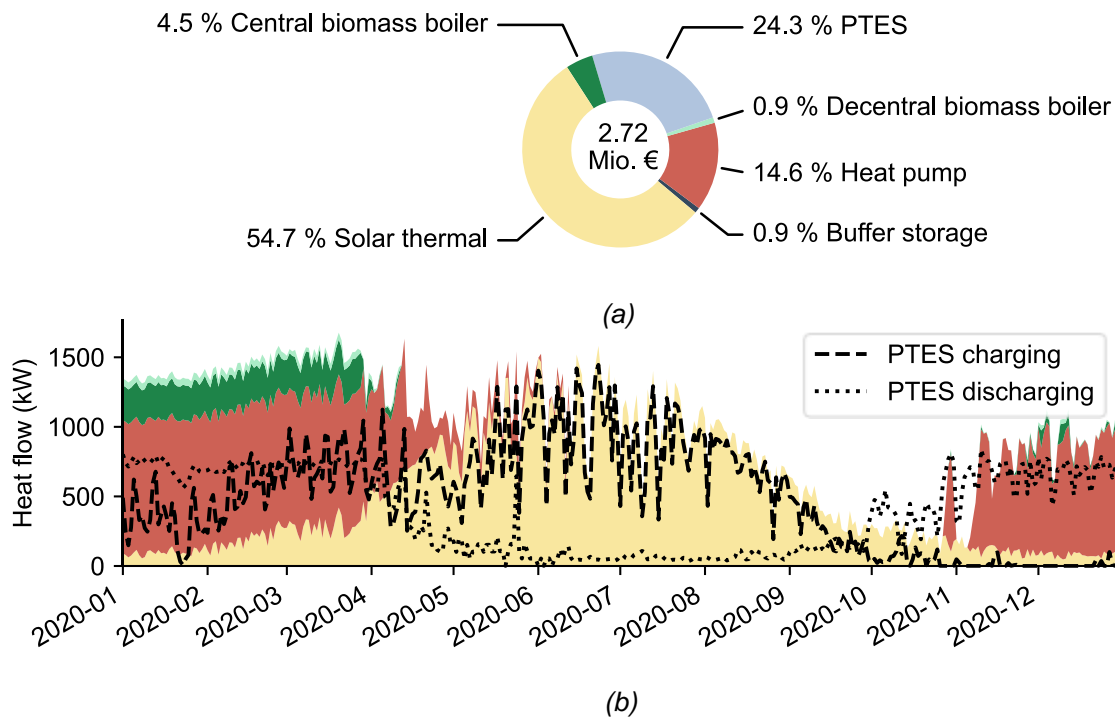


Figure 4: b) optimized operation of the energy system model and a) the investments in the energy system with the different shares for the energy converters and storages.

## 4. Discussion

The discussion divides into highlighting the benefits and drawbacks of the regression model and comparing the results of the energy system model with the existing sites in Marstal and Dronninglund.

### 4.1. Discussion of the Polynomial Regression Results

The difficulty of the regression model is the prediction of temperature profiles without knowing the demand or the volume of the PTES. Therefore, the model only has the time, the solar irradiance, and the ambient temperature as features. Using three storages as training data and Dronninglund as testing leads to a relatively low accuracy due to the considerable fluctuations in the temperature profile of Dronninglund. This clearly shows the disadvantages of the model because it cannot depict hourly or daily fluctuations. However, the prediction results improve up to 91 % for Florence. This improvement is due to the data being a simulation; therefore, no rapid changes are in the gradient. The regression model cannot replace a detailed storage simulation but can estimate the seasonal behavior at a given location for a pre-analysis.

### 4.1. Discussion of the Energy System Results

Utilizing the predicted temperature profiles in the energy system model leads to reasonable results, see Table 4. It should be noted that the energy system model is not applied to the demand structure in Marstal or Dronninglund due to the lack of data. Therefore, the results can only be compared relatively. The storage is about half the size compared to Marstal and Dronninglund; the same applies to the solar thermal field. The storage charging is 4362 MWh, while the storages in Marstal and Dronninglund are charged with 7813 MWh and 12760 MWh. The ratio volume to solar field is 2.13 m, while for Marstal and Dronninglund, it is 2.25 and 1.68 m. [13] suggests a ratio between 2 and 3 m, with 3 being relatively compared to our calculations and the existing sites. This comparison shows a correlation between the optimization results and the existing sites. Additionally, the results verify the approach of the piecewise functions to calculate the areas of the storage based on the volume due to the moderate deviation of 17.7 % compared to the measured data of Marstal in 2014.

A significant drawback is the deviation of the electrical input for the heat pump – 29 %. Utilizing the predicted temperature profiles increases the accuracy, but a deviation of 29 % is still not accurate. We also tested a binary-based model leaving the temperature calculation inside the optimization. We applied this model to the same use case and obtained a solution for a timestep of 24 h after 8 h computational time on a windows machine with 488 GB and an AMD EPYC 7542 32-Core Processor. Reducing the timestep length to 8 h led to computational times over 24 h. Therefore, we conclude that the model is not applicable. Feeding the optimization model with fixed temperature profiles already suggests a control strategy for the heat pump and

influences the results. The last option would be to assume a constant temperature inside the storage and calculate the COP based on a constant value. Based on this study, we suggest using a constant temperature inside the storage for the COP calculation and performing a more detailed simulation of the PTES with TRYNSYS or other simulation software mentioned in the state-of-the-art. This has the advantage that the energy system model already calculates the design of the PTES, and the simulation can focus on the control strategy. Assuming a constant temperature inside the PTES for the energy system model would allow loading the storage with other technologies. The regression model can only work if a seasonal behavior is present in the charging process; however, other technologies like air heat pumps would load the storage based on the electricity market.

*Table 4: Comparison of results with the existing sites in Marstal and Dronninglund. [3]*

Parameter	Marstal (2015)	Dronninglund (2015)	Energy System Results (2020)
Charging, MWh	7813	12760	4362
Discharging, MWh	5435	11983	3433
Thermal losses, MWh	2946	1275	853
Heat capacity, MWh	5430	5500	2086
T-max, °C	84	89	92
T-min, °C	20	10	23
Volume, m <sup>3</sup>	75000	63000	26725
Solar gain, kWh/m <sup>2</sup> /a	395	447	355
Solar field, m <sup>2</sup>	33300	37573	12555
Ratio volume / solar field, m	2.25	1.68	2.13

## 5. Limitation of Results

The method is applied to a fictional grid with six consumers. The grid size does not represent the usual DHS, and larger grids would lead to higher computational times. However, the solver needed ca. 1 h, and the model is applied in the planning phase; therefore, the computational time can be slightly higher than in a control optimizations. Nevertheless, the method must be applied to larger grids to evaluate its performance.

## 6. Conclusion

In this study, a method is developed to design PTES supplied by solar thermal energy inside a DHS. Our approach utilizes historical data and identifies strong correlations between the modeling results and the existing sites. The results suggest a ratio of ca. 2 m for volume vs. solar field. For 1 GWh capacity of a PTES, an area of around 1500 m<sup>2</sup> combined with 37500 m<sup>2</sup> solar thermal collectors would be needed. Achieving a more accurate result, our model can be applied in the planning phase for a grid-based DHS supplying heat from and to different locations. The model can depict the decreasing specific investments and thermal losses due to an increasing volume. This behavior is achieved by a piecewise function taking the volume as input and the investments and geometry of the PTES as output.

The difficulty within the optimization is the control strategy of the heat pump, depending on the temperature inside the storage. We examined three possibilities:

1. Predicting the temperatures of the storage before the optimization and calculating the COP a priori
2. Assuming one constant temperature of the storage and calculating the COP a priori
3. Calculating the temperature and the COP during the optimization

The first option already induces a control strategy inside the optimization and causes an error of 29 % for the electrical input. The second option did not depict any temperature changes inside the storage and caused an error of 45 %. The third option was not applicable due to a high computational time. In conclusion, we suggest the second option, followed by a detailed simulation. This study showed the computational limits for a mathematical optimization in the design stage due to the fact that the dimension – the volume of the PTES – and operation – the temperature of the PTES – is a variable. In the future, new methods could be developed depicting volume and temperature as a variable and computing results in a practical manner.

Furthermore, it would be interesting to investigate different technologies supplying a PTES and compare it to solar thermal energy due to its high investments. For Power-to-Heat technologies, the investigation should optimize at the Day-Ahead market to react to price fluctuations. In addition, large grids should be examined due to the large space requirements for PTES charged by solar thermal energy.

## Contributions and Acknowledgements

MS advised the research of the polynomial regression model, developed the energy system model, and wrote the manuscript. YX developed the polynomial regression model. MiR and MvB advised the development of the energy system model. MR acquired the finance for this research via the project ODH@Jülich.

# Nomenclature

## Letter symbols

$A$  area, m<sup>2</sup>

$c_p$  specific heat capacity, kJ/(kg K)

$COP$  coefficient of performance for heat pumps

DHS district heating system

$E$  energy, kJ

$f^{opex,var}$  variable opex of the components in the energy system, €

$g^{opex,fix}$  fixed opex of the components in the energy system, €

$h^{capex}$  capex of the components in the energy system, €

$m$  mass, kg

$\dot{m}$  mass flow, kg/s

MAE mean average error

MILP mixed-integer linear programming

PTES pit thermal energy storage

$\dot{Q}$  heat flow, kW

RMSE root mean squared error

$t$  time

$\Delta t$  timestep length of the optimization, s

$T$  temperature, K

$U$  heat transfer coefficient, kW/(m<sup>2</sup> K)

$Z$  set of arcs – pipes, consumers, producers

## Greek symbols

$\tau$  set of timesteps

$\gamma$  solar irradiance, kW/m<sup>2</sup>

$\mu$  efficiency

## Subscripts and superscripts

$a$  pipe

capex capital expenditures

collector solar collectors

con consumer

conv energy converter

ff forward-flow

bf backward-flow

in entering a component

$k$  energy converter or storage

m shell

min minimum

loss losses

opex, var variable operational expenditures

opex, fix fixed operational expenditures

out leaving a component

pro producer

rad irradiance

soil soil/ground

stor storage

supply for the supply / to the consumer

$t$  timestep

top at the top level

## References

- [1] Paardekooper S, Lund RS, Mathiesen BV, *et al.* Heat Roadmap Europe 4: Quantifying the Impact of Low-Carbon Heating and Cooling Roadmaps: Aalborg University; 2018.
- [2] Yang T, Liu W, Kramer GJ, Sun Q. Seasonal thermal energy storage: A techno-economic literature review. *Renewable and Sustainable Energy Reviews* 2021; 139: 110732.
- [3] Schmidt T, Sorensen PA. Monitoring Results from Large Scale Heat storages for District Heating in Denmark. 14th International Conference on Energy Storage 2018.
- [4] Li X, Wang Z, Li J, *et al.* Comparison of control strategies for a solar heating system with underground pit seasonal storage in the non-heating season. *Journal of Energy Storage* 2019; 26: 100963.
- [5] Crawley DB, Lawrie LK, Winkelmann FC, *et al.* EnergyPlus: creating a new-generation building energy simulation program. *Energy and Buildings* 2001; 33(4): 319–31.
- [6] Pauschinger. Solites Bericht.
- [7] Lyden A. Modelling and design of local energy systems incorporating heat pumps, thermal storage, future tariffs, and model predictive control 2020.
- [8] Dominković DF, Čosić B, Bačelić Medić Z, Duić N. A hybrid optimization model of biomass trigeneration system combined with pit thermal energy storage. *Energy Conversion and Management* 2015; 104: 90–9.
- [9] Reiter P, Poier H, Holter C. BIG Solar Graz: Solar District Heating in Graz – 500,000 m<sup>2</sup> for 20% Solar Fraction. *Energy Procedia* 2016; 91: 578–84.
- [10] Sifnaios I, Gauthier G, Trier D, Fan J, Jensen AR. Dronninglund water pit thermal energy storage dataset. *Solar Energy* 2023; 251: 68–76.
- [11] Sifnaios I, Jensen AR, Furbo S, Fan J. Performance comparison of two water pit thermal energy storage (PTES) systems using energy, exergy, and stratification indicators. *Journal of Energy Storage* 2022; 52: 104947.
- [12] Słomczyńska K, Mirek P, Panowski M. Solar Heating for Pit Thermal Energy Storage – Comparison of Solar Thermal and Photovoltaic Systems in TRNSYS 18. *Adv. Sci. Technol. Res. J.* 2022; 16(5): 40–51.
- [13] Lindenberger D, Bruckner T, Groscurth H-M, Kümmel R. Optimization of solar district heating systems: seasonal storage, heat pumps, and cogeneration. *Energy* 2000; 25(7): 591–608.
- [14] Rämä M, Mohammadi S. Comparison of distributed and centralised integration of solar heat in a district heating system. *Energy* 2017; 137: 649–60.
- [15] Martins F, Patrão C, Moura P, Almeida AT de. A Review of Energy Modeling Tools for Energy Efficiency in Smart Cities. *Smart Cities* 2021; 4(4): 1420–36.
- [16] Bai Y, Wang Z, Fan J, *et al.* Numerical and experimental study of an underground water pit for seasonal heat storage. *Renewable Energy* 2020; 150: 487–508.
- [17] Pan X, Xiang Y, Gao M, *et al.* Long-term thermal performance analysis of a large-scale water pit thermal energy storage. *Journal of Energy Storage* 2022; 52: 105001.
- [18] Xie Z, Xiang Y, Wang D, *et al.* Numerical investigations of long-term thermal performance of a large water pit heat storage. *Solar Energy* 2021; 224: 808–22.
- [19] Gauthier G. Benchmarking, and improving models of subsurface heat storage dynamics: Comparison of Danish PTES and BTES installation measurements with their corresponding TRNSYS models 2020.
- [20] Preuss J. 250520\_Kurzbeschreibung\_Trnsys-1.
- [21] Dahash A, Michele Bianchi Janetti M, Ochs F. Numerical Analysis and Evaluation of Large-Scale Hot Water Tanks and Pits in District Heating Systems. In: Corrado V, Fabrizio E, Gasparella A, Patuzzi F, editors. *Numerical Analysis and Evaluation of Large-Scale Hot Water Tanks and Pits in District Heating Systems*; 2020. IBPSA; 1692–9.
- [22] Detailed 3-D models of a large-scale underground thermal energy storage with consideration of groundwater conditions 2018.
- [23] COMSOL. COMSOL Release Notes.
- [24] Chang C, Nie B, Leng G, *et al.* Influences of the key characteristic parameters on the thermal performance of a water pit seasonal thermal storage. *Energy Procedia* 2017; 142: 495–500.
- [25] Chang C, Leng G, Li C, *et al.* Investigation on transient cooling process in a water heat storage tank with inclined sidewalls. *Energy Procedia* 2017; 142: 142–7.
- [26] Ansys. MASTER SOFTWARE LICENSE AGREEMENT.
- [27] Bai Y, Yang M, Fan J, *et al.* Influence of geometry on the thermal performance of water pit seasonal heat storages for solar district heating. *Build. Simul.* 2021; 14(3): 579–99.

- [28]Narula K, Oliveira Filho F de, Villasmil W, Patel MK. Simulation method for assessing hourly energy flows in district heating system with seasonal thermal energy storage. *Renewable Energy* 2020; 151: 1250–68.
- [29]Lyden A, Flett G, Tuohy PG. PyLESA: A Python modelling tool for planning-level Local, integrated, and smart Energy Systems Analysis. *SoftwareX* 2021; 14: 100699.
- [30]Xiang Y, Xie Z, Furbo S, Wang D, Gao M, Fan J. A comprehensive review on pit thermal energy storage: Technical elements, numerical approaches and recent applications. *Journal of Energy Storage* 2022; 55: 105716.
- [31]Achterberg T, Bixby RE, Gu Z, Rothberg E, Weninger D. Presolve Reductions in Mixed Integer Programming. *INFORMS Journal on Computing* 2020; 32(2): 473–506.
- [32]Anthrakidis A, Merten F, Buddeke M, *et al.* Die kommunale Effizienzrevolution für den Klimaschutz in den deutschen Städten – „KomRev“. Jülich: Solar-Institut Jülich der FH Aachen; 2017.
- [33]Salvestroni M, Pierucci G, Fagioli F, *et al.* Design of a seasonal storage for a solar district heating in Florence. *IOP Conference Series Materials Science and Engineering* 2019; 556(1): 12026.
- [34]Rohatgi A. Webplotdigitizer: Version 4.6; 2022. Available at: <<https://automeris.io/WebPlotDigitizer>>.
- [35]Energy, DG Joint Research Centre /. JRC Photovoltaic Geographical Information System (PVGIS) - European Commission. Available at: <[https://re.jrc.ec.europa.eu/pvg\\_tools/en/](https://re.jrc.ec.europa.eu/pvg_tools/en/)> [accessed 01.02.2023]
- [36]Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 2011; 12(85): 2825–30.
- [37]scikit-learn developers. StandardScaler. Available at: <<https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>> [accessed 13.03.2023]
- [38]VDI The Association of German Engineers. Economic efficiency of building installations - Fundamentals and economic calculation. Berlin: Beuth Publisher; 2012 2012 Sep 1. Available at: <<https://www.vdi.de/richtlinien/details/vdi-2067-blatt-1-wirtschaftlichkeit-gebaeudetechnischer-anlagen-grundlagen-und-kostenberechnung-1#:~:text=Die%20Richtlinienreihe%20VDI%202067%20behandelt,Richtlinienreihe%20in%20mehrere%200Bl%C3%A4tter%20gegliedert.>> [accessed 15.02.2023]
- [39]Gurobi Optimization LL. Gurobi Optimizer Reference Manual. Available at: <<https://www.gurobi.com>> [accessed 07.03.2023]
- [40]Verenum. Planungshandbuch Fernwärme. Version 1.1 vom 21. September 2017. Ittigen, Bern: EnergieSchweiz Bundesamt für Energie 2017 Sep 21.
- [41]Sara Ortner, Martin Pehnt, Sebastian Blömer, Andreas Auberger, Jan Steinbach, Jana Deurer, Eftim Popovski, Oliver Lösch, Nora Langreder, Nils Thamling, Malek Sahnoun, Dominik Rau. Analyse des wirtschaftlichen Potenzials für eine effiziente Wärme- und Kälteversorgung: Beitrag zur Berichtspflicht EnEff-RL, Artikel 14 - Anhang VIII: Institut für Energie- und Umweltforschung Heidelberg gGmbH, Heidelberg; IREES GmbH, Karlsruhe; Prognos AG, Berlin.
- [42]Sorknæs P. Simulation method for a pit seasonal thermal energy storage system with a heat pump in a district heating system. *Energy* 2018; 152: 533–8.
- [43]Hart WE. Pyomo - optimization modeling in Python. 2cd ed. Cham, Switzerland: Springer Optimization and its Applications 2017.
- [44]Danish Energy Agency. Technology Data for Energy storage. Copenhagen; 2018 [accessed 15.02.2023] Available at: <[https://solarthermalworld.org/wp-content/uploads/2019/05/technology\\_data\\_catalogue\\_for\\_energy\\_storage.pdf?x32997](https://solarthermalworld.org/wp-content/uploads/2019/05/technology_data_catalogue_for_energy_storage.pdf?x32997)>.