A novel approach to address spatial uncertainties and simultaneously optimize concept, scope, and equipment design in district heating system

Tong Guo 1, Yusheng Chen 2, Max Bachmann 1, Martin Kriegel 1
1 Technical University of Berlin, Hermann-Rietschel-Institut, Berlin, Germany
2 Technical University of Munich, Campus Straubing for Biotechnology and Sustainability, Munich, Germany

Abstract

The design of district heating systems often involves the consideration of the appropriate system scope and the trade-off between decentralized and centralized systems, referred to as spatial uncertainties. This study presents a novel optimization method aimed at sustainable dimensioning of the district heating system by simultaneously optimizing the system concept, scope, and design of equipment. The method quantifies the spatial uncertainty and assists in the decision between decentralized and centralized systems by using clustering algorithms and the genetic algorithm. The proposed approach was tested in a city-scale case study, resulting in improved solutions with lower levelized energy cost and carbon dioxide emissions, and providing insights into the energy system and district heating network scope relationship.

Highlights
- A framework to simultaneously optimize the system concept, scope, and equipment is developed.
- The method combines various ML algorithms to quantify spatial uncertainty.
- The relationship between the scope of the district heating network, the type of equipment and system, and the key performance indicators (KPI) is proposed.

Introduction

Governments worldwide are actively addressing the current environmental challenges faced by urban and suburban areas, focusing on setting long-term targets such as the 20-20-20 package (Brigitte et al. 2015). The heat transition is a major social project as part of the energy transition. More than half of the current final energy consumption in Germany is accounted for by heating and hot water preparation (Umweltbundesamt 2020).

Heating supply system is complex and consists of networks of pipelines that connect buildings in urban neighborhoods, districts, or entire cities, which can be supplied by central or decentralized producers. In order to adapt to changes in energy supply and demand and to increase the efficiency of sustainable energy use, the heating supply system has evolved from the original central heating network of the first generation to the present-day fourth-generation heating network (Lund et al. 2016). Regarding the 4th district heating (DH) system, supply concepts based on renewable energies play an important role in implementing future sustainable heat supply systems. Therefore, designing district heating systems often requires considering the appropriate DH network scope. Due to spatial uncertainties, district heating systems in the early design phase often rely on experience. As a result, the options obtained are often the locally optimal solution, and the designed system may not be suitable for future scenarios (Sameti und Haghighat 2017).

Few studies have investigated the relationship between spatial uncertainty and energy systems, including studies on the potential for district heating networks according to regional scales and energy systems (Chambers et al. 2019; Pampuri et al. 2019). In this context, the main difficulties encountered in research are divided into DH network scope issues and system selection issues. The main method used by previous researchers to address system scope issues is to establish a threshold for preliminary assessment of the scope. Typical threshold values are spatial demand density of 150 MWh/ha/a (Persson et al. 2014) or linear demand density of 1.4 MWh/Tm/a (Gudmundsson et al. 2013), above which a site is considered potentially suitable for a district heating system. It is pointed out that these indicators are empirical values based on conventional heating networks and vary from region to region (Sameti und Haghighat 2017).

Another aspect of research focuses on finding the optimal solution within a fixed scope. (Marquant et al. 2018) studied the possibility of interlinked energy networks within a specific scope. (Su et al. 2022) studied the issue of selecting the optimal layout of pipelines within a region and the best location of energy supply equipment. However, the optimization of pipeline topology and equipment location or type requires the prior determination of the spatial uncertainties. Therefore, determining the scope of the network system is the most critical aspect in the early stages of the construction process and has yet to be effectively resolved. In terms of the system selection issue or the trade-off between centralized and decentralized systems, previous qualitative studies have provided comprehensive analyses of the pros and cons of centralized and decentralized heating systems (Burger et al. 2019). However, research on quantitative research, such as specific trade-off and boundaries of these systems in particular areas, is still in its infancy.

This study aims to introduce a novel optimization approach for the simultaneous optimization of the concept,
the network’s scope, and the heat supply systems' equipment to enable sustainable dimensioning. The proposed method leverages a quantitative approach to address spatial uncertainty during the early planning stages of heat supply systems, enabling informed decision-making between centralized and decentralized systems. Furthermore, this approach provides a robust framework for optimizing heating systems, integrating multiple objectives and considering the complex trade-offs between factors such as energy efficiency, investment cost, and environmental impact. By utilizing this approach, it is possible to achieve optimal system design, taking into account the unique requirements and limitations of a designated region, thereby advancing the sustainable development of energy infrastructure.

**Methodology**

**Optimization framework**

To simultaneously optimize the system scope and type and quantify the impact of spatial uncertainty on system design, this paper constructs a framework, as outlined in Figure 1. The framework is divided into three major steps (labeled in different colors).

![Figure 1](https://example.com/f1.png)

**Figure 1: framework for multi-level optimization of DH system.**

The initial step (The yellow-labeled module) involves configuring the clustering algorithm, which utilizes building-specific information, such as the geographical coordinates and load profiles of buildings, as input parameters. The problem is divided into multiple sub-problems in the clustering stage using the chosen clustering algorithm. This generates all possible clustering scenarios based on the set number of clusters and produces a clustering map representing all sub-problems. For each cluster scenario, the pipeline network’s topology is optimized, and the generated results are used as input for the next step of system modeling. In this way, pipeline network design at different scales is confirmed. A more detailed description is provided in the section: Cluster algorithm selection and settings.

In the second step (The blue-labeled module), the simulation model of the DH system, which includes critical components and variables, is constructed using Python and Modelica. Through this process, we determine vital parameters and input variables and their corresponding ranges, which reflect the thermal performance of the district heating system. After setting the decision variables, the framework generates scenarios based on the range of these variables to ensure process feasibility. By combining the EnergyHub concept with the established models, the framework effectively simulates the multi-energy input and output characteristics of various sources and load nodes in the district heating system. Each sub-problem is treated as an energy hub, and for each energy hub, the optimal solution of the objective function is found by Mixed-Integer Linear Programs (MILPs). Once the optimal solution or selected combination of solutions for each sub-problem is determined, it represents optimization at the equipment system level. These individual solutions are subsequently aggregated across the entire cluster scenario and used as input for the next step of the Genetic Algorithm (GA) optimization. The establishment of the district heating system is described in detail in section: District Heating system model.

The third step (The orange-labeled module) involves the integration of a Multi-Objective Genetic Algorithm (MOGA). This step mainly addresses the optimization at the system scope level, with the number of clusters being one of the optimized input variables, effectively incorporating the scope of the system into the optimization process. This step primarily describes selecting the design and operational strategies for multiple interconnected energy systems within a city and using them as appropriate objective functions. Constraints are established to formulate a multi-objective optimization problem to obtain the optimal Pareto frontier. Once the Pareto frontier is generated based on the objective functions, an overview of the optimization result demonstrates the trade-off between centralized and decentralized district heating systems at a city scale. A more detailed description is provided in section: Optimization problem.

**Cluster algorithm selection and settings**

In this paper, the clustering algorithm is set, selected, and executed mainly based on the Python framework. The input and output are combined with QGIS software (QGIS Development Team 2009). Figure 2 shows buildings’ geographical and foundational information, such as locations from the QGIS system and the energy consumption information from the Modelica model. Based on this, information is given as input to the clustering algorithm and serves as the basis for
The input data dimensions are partitioned and compared in the process. The Hopkins Statistic is used as the selection criterion to evaluate whether a given dataset contains significant, non-random structural values that can be clustered (Banerjee and Dave 2004). This indicator ranges between 0 and 1, with values between 0.7 and 0.99, indicating a robust clustering trend. Ultimately, the clustering criteria selected were building location, required temperature levels for the building system, and energy consumption levels, with a Hopkins Statistic value of 0.87. The setting of the number of clusters differs from the traditional approach of selecting the optimal number of clusters. Here, the clustering is performed based on the number of buildings by finding different combinations of building clusters from 1 to the total number of buildings in the district. A cluster number of 1 means all buildings in the region belong to the district heating network, indicating a centralized system, while a cluster number equal to the total number of buildings results in a completely decentralized system. Through this unique setting, DH networks of different scopes and hybrid system configurations are established.

Therefore, the number of clusters needs to be defined in advance instead of letting the system decide. To achieve this, clustering algorithms such as k-means, hierarchical algorithm Birch, and density-based algorithm DBSCAN were compared, and both k-means and Birch were found to be suitable for the task with little difference in performance, unlike a density-based algorithm which does not require entering the number of clusters. The chosen clustering algorithm is Birch (Balanced Iterative Reducing and Clustering using Hierarchies) (Boris et al. 2018), which offers faster efficiency. According to the clustering results, the pipeline topology structure was optimized using the Thermos tool (Aunedi et al. 2020), and the results were then input into QGIS. After processing with Python and removing some non-science Outliers manually, clustering sets containing the pipeline optimization results were generated.

**District Heating system model**

This work formulated the Energy Hub concept as a Mixed-Integer Linear Programming (MILP) problem. The Energy Hub is treated as a black box that takes in multiple forms of energy, such as electricity, gas, and heat, and outputs the final form of energy supplied to the user side (Geidl et al. 2007). This work is based on the premise of District Heating, so the final output energy form is heating.

The primary purpose of this paper is to illustrate the novel framework, and therefore the selection of technologies is not comprehensive. Photovoltaic (PV), as the most widely applicable form of renewable energy, is chosen as the preferred new energy technology. Other renewable technologies, such as biomass and wind energy, which have higher environmental requirements and lower applicability to urban areas, are not included. Additionally, battery storage and hydrogen systems are also excluded due to the high costs associated with these technologies.

**Figure 2: workflow of combined clustering algorithm.**

**Figure 3: A schematic drawing for the DH system.**

As indicated by the colors in Figure 3, the system mainly consists of three major modules covering producers, consumers, and transformers used to convert forms of energy, as well as important components of the district heating system, such as storage and pipeline networks. The gray module is the building energy simulation, which is a model constructed using Modelica and Python with the Aixlib library (Maier et al. 2016), as detailed in the author's previous research (Guo et al. 2023). This model effectively outputs the building heating demand data and hourly load profile based on the input building information.

The green module is the pipeline network model built on the Modelica framework, as described in DHC research.
(Bachmann et al. 2021). The model outputs pipe heat losses and pump energy requirements based on input variables such as network type, topology, and other relevant physical parameters of the pipeline.

The blue part represents the system equipment component model using the energy hub concept, which is built in a Python environment using the Open Energy Modelling Framework (oemof) (S. Hilpert et al. 2018) as introduced earlier in (Kersten et al. 2021). The proposed system is divided into three primary circuits: heating, electricity, and gas circuits, each with its corresponding equipment. The available energy systems considered in this study include natural gas boiler (NG), combined heat and power engine (CHP), photovoltaic panel (PV), thermal storage (TS) and district heating network (DHN), with the assumption that gas and electrical networks are already connected to all buildings.

Omeof employs the MILP algorithm to optimize the corresponding devices’ selection and operation. Each cluster scenario generated in the first step is composed of multiple sub-clusters. Each sub-cluster, considered as an energy hub, is optimized by MILP within its group in the second step. The optimal solutions within each sub-cluster are then summed and fed into the next step of the GA algorithm to obtain a global solution, thus achieving a bi-level optimization.

Optimization problem

This stage primarily aims to address the challenge of multi-objective optimization. An elitist genetic algorithm NSGA-II (Deb et al. 2002) is introduced in the current framework. In this paper, the NSGA II algorithm is implemented using the pymoo library in python environment (Blank und Deb 2020), which is capable of parallel computing and effectively solving multi-objective problems, along with the ability to handle both continuous and discrete variables and equality and inequality constraints.

![Figure 4: KPI of multi-objective optimization from different stakeholders' perspective.](https://example.com)

Environmental KPIs:

KPI CO2 emission
KPI Percentage of renewable energy

Technical KPIs:

KPI System efficiency
KPI Heat loss
KPI Supply temperature

Economic KPIs:

KPI Operational Expenditures (OPEX)
KPI Capital Expenditure (CAPEX)
KPI Levelized cost of heating (LCOH)
KPI Payback period (PB)

The NSGA-II algorithm requires the specification of optimization objectives and constraints. In this paper, various stakeholders have different focus points, and the optimization objectives are defined based on three main Key Performance Indicators (KPIs), which include technical, economic, and environmental criteria, as shown in Figure 4. In the subsequent equations, we present the decisive ultimate indicators within each category as representative metrics.

Environmental KPIs:

The energy-specific CO2 emission \( e_{CO2} \) is selected as the environmental impact indicator to assess the proposed concepts.

\[
e_{CO2} = \frac{\int (W_{gas} \times H_{gas} \times f_{CO2gas} + P_{el from grid} \times f_{CO2el}) \, dt}{\int Q_{demand} \, dt}
\]  

where the \( f_{CO2gas} \) denotes the CO2 emission factors for the natural gas and \( f_{CO2el} \) represents the CO2 emission factor for electricity from the public power grid.

Another important environmental indicator is the proportion of renewable energy, which reflects the proportion of renewable energy supply in the overall system.

Technical KPIs:

In this study, the Primary Energy Factor (PEF) is calculated using the methodology described in the (Latôšov et al. 2017), which takes into account the heat delivered at the primary side of the district heating system. The PEF is a metric that represents the ratio of delivered energy to primary energy, which is the amount of primary energy required to produce one unit of useful energy.

The heat loss in the network is considered an important indicator of the network’s efficiency.

\[
\eta_{loss} = \frac{\sum \int Q_{pipe} \, dt}{\int Q_{demand} \, dt}
\]

The heat loss in the DH network is represented by the sum of the heat loss of each pipe \( Q_{pipe} \) in the network.

Economic KPIs:

The economic feasibility of the concepts is evaluated using the levelized cost of heat (LCOH) for the entire system. LCOH is a criterion that represents the specific cost of covering the total heating demands of the network and is expressed as:

\[
LCOH = \frac{TAC_{year} \times Q_{year}}{Q_{demand}} = \frac{CAPEX_{year} + OPEX_{year}}{\int Q_{demand} \, dt}
\]  

In the formula, \( TAC_{year} \) represents the total annualized cost of the network, which consist of the annualized capital cost \( CAPEX_{year} \) and the annual operating cost \( OPEX_{year} \). These costs are determined using the annuity method and evaluated based on the net present value of the German guideline VDI 2067 (VDI 2067 2012).

In economic analysis, the payback period (PB) is a widely used methodology that calculates the time period required to recover the initial investment by considering the economic returns of the project over time (Başoğul und Keçebaş 2011). Therefore, the payback period of DH system is considered in this work.

The framework is utilized to address the inherent conflicts among multiple objectives in optimizing energy performance in conjunction with economic and environmental indicators. Therefore, the economic
criterion is represented by the levelized cost of heat (LCOH), while the environmental criterion is represented by CO₂ emission. In addition, other criteria, such as the payback period and the proportion of buildings with satisfactory temperatures, are considered as constraints in the optimization algorithm. Specifically, the calculated payback time must not exceed 40 years, while the proportion of buildings with satisfactory temperatures should not be lower than 95%.

Case

The developed framework's effectiveness is demonstrated through an illustrative case study in which the methodology is applied to a small German suburban area. This suburban area, located in Leeste, Wehye, Germany, near Bremen, covers approximately 16 hectares and consists of 255 buildings of different construction years and diverse usages, including residential and mixed residential-commercial buildings. This case area is chosen for its large amount of yearly collected energy consumption data and relatively comprehensive input data.

![Figure 5: case study map and basic info.](image)

Economic and environmental input parameters

The key input parameters for economic and environmental evaluations include equipment efficiency, equipment and pipeline costs, performance parameters, and energy prices are outlined in Table 1. In this study, the overall efficiency of the CHP is determined in the model by the device size, the ratio of power and heat. The efficiency of the condensing boiler is set to 0.94. The maximum allowable size of PV panels is limited by the total available roof area per building, and the efficiency is set to 0.14 based on (Omu et al. 2013). The efficiency of the heat pump is determined in the model by the supply water temperature, the outside temperature, and the device size. The efficiency of the network is calculated using the Modelica model, and the heat loss and pump energy depend on the length of the pipes and the topology of the network. The cost of the network is determined by the piping cost and the trench cost, based on (Nussbaumer und Thalmann 2016). In this case, the equipment in the district heating system has a lifespan of 20 years, and the network has a lifespan of 40 years (Wagner et al. 2015).

In the current study, the carbon factor for natural gas is assumed to be 0.24 kg CO₂/kWh, while that for electricity is 0.34 kg CO₂/kWh based on (German Association of Energy and Water Industries, BDEW) (Kiesel 2021). The retail price of natural gas and electricity is set to be 0.053 and 0.18 Euro/kWh. These prices are assumed to be based on the data from 2020 (Kiesel 2021).

<table>
<thead>
<tr>
<th>Components</th>
<th>Size range [kW]</th>
<th>Efficiency η</th>
<th>Levelized cost [Euro/kWh]</th>
<th>Lifespan [a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHP</td>
<td>10-2000</td>
<td>[-]</td>
<td>475.9–81.5</td>
<td>20</td>
</tr>
<tr>
<td>Condensing Boiler (NG)</td>
<td>50-2000</td>
<td>0.94</td>
<td>18.7–6.7</td>
<td>20</td>
</tr>
<tr>
<td>PV panels</td>
<td>[-]</td>
<td>0.14</td>
<td>152.8–73.3</td>
<td>25</td>
</tr>
<tr>
<td>Heat pump (Air)</td>
<td>10-2000</td>
<td>[-]</td>
<td>90.4–40.7</td>
<td>20</td>
</tr>
<tr>
<td>Storage (Thermal)</td>
<td>10-20000</td>
<td>0.99</td>
<td>1.183</td>
<td>20</td>
</tr>
<tr>
<td>DH-network</td>
<td>[-]</td>
<td>[-]</td>
<td>308-1569 [Euro/m]</td>
<td>40</td>
</tr>
</tbody>
</table>

Other inputs data

The Test Reference Year (TRY) weather file provided by the German meteorological service (Wetter und Klima April 07) was used for building and piping simulation, while the data for PV simulation, including incident solar radiation, ambient temperature, relative humidity, and other relevant information, were obtained from the EnergyPlus database (U.S. Department of Energy 2015). The GA optimization in this study utilized the NSGA-II algorithm with a population size of 100 and 300 generations.

Results and discussion

The results demonstrate the effectiveness of the framework from two perspectives. The first perspective showcases the practical results of the framework as a tool in its application. The second perspective delves into the theoretical significance behind the results on a theoretical level.

Results on the application level

Figure 6 depicts the optimized system network graph achieved by combining the clustering algorithm with the topological structure. Systems with a network connection coverage (the ratio of buildings connected to the network within the research area) greater than 80% are classified as centralized systems, corresponding to cluster groups 1 to 30. Systems with a network connection coverage between 30% and 80% are classified as hybrid systems, corresponding to cluster groups 31 to 70. Hybrid systems consist of multiple small networks connecting most of the buildings, while the remaining buildings are not connected to the network and use decentralized heating.

Table 1: Size band, cost and efficiency of components in DH system.
Decentralized systems only have less than 30% of the buildings connected to the network, corresponding to cluster groups 70 to 255, with the remaining buildings using decentralized heating.

Figure 6: Cluster algorithm results correspond to QGIS pipeline diagrams.

Figure 7: Pareto sets for optimal DH solutions.

Figure 8: Energy source ternary plot for the different Pareto solutions.

As the proposed district heating (DH) system incorporates multiple energy sources to optimize both environmental and economic parameters, the distribution of energy sources is a crucial factor in determining the feasibility of the system as a future fourth-generation DH system. The ternary representation of energy shares in Figure 8 illustrates the energy source composition corresponding to the Pareto solutions.
to the Pareto optimal solutions. The three sides represent the percentages of PV renewable energy, gas, and electricity obtained from the grid in relation to the total energy supply. The sum of these percentages amounts to 100%. Under the energy price scenario set in the study, natural gas remains the primary energy source, with centralized systems having a higher dependence on it compared to hybrid and decentralized systems. Among the three systems, the hybrid system has a slightly higher utilization rate of solar energy, with a maximum of 28.2% of the energy supplied by photovoltaics.

**Results on the theoretical level**

At a theoretical level, to enhance the understanding of the relationship between the system scope, concepts, equipment, and district heating evaluation indicators, a parallel coordinate graph is used in Figure 9 to conduct a breakdown for each Pareto optimal solution. Figure 9 shows that larger network scopes lead to more centralized systems and increased CHP technology usage, while decentralized systems primarily rely on boilers. Hybrid systems have a higher heat pump utilization rate than decentralized systems. (The centralized system utilizes CHP with a range of 600 to 800. The decentralized system employs boilers with a range of 700 to 900. The hybrid system incorporates HP with a range of 200 to 600.) Figure 9 also demonstrates the correlation between regional scope and type with key performance indicators, such as LCOH, environmental impact, and system robustness. Centralized systems have a higher LCOH due to network infrastructure and CHP equipment costs, while hybrid and decentralized systems perform better regarding environmental indicators. Decentralized systems have the best performance in meeting temperature requirements, fully satisfying the supply temperature needs of all buildings.

The focus of this case study is currently on the relationship between heating systems and regional scales within Germany. Different target regions in diverse climates and locations may present distinct research conclusions. For instance, in densely populated urban areas in China, the relationship between heating systems and regional scales would differ from that observed in German regions. However, this paper primarily presents a novel framework that can be equally applied to regions with varying climates and conditions, enabling the derivation of corresponding conclusions that align with local climate conditions.

**Conclusion**

This study presents a new optimization framework for district heating systems. The framework considers the concept, scope, and equipment of the DH system to enable sustainable dimensioning. The case study, conducted in an urban community in Bremen, Germany, demonstrates the effectiveness of the framework. The contribution of this study lies in the integration of multiple machine learning algorithms to quantify spatial uncertainty in district heating systems and consider the trade-offs between centralized and decentralized systems. The focus of this study is on early-stage design, and in the future, a graphical user interface (GUI) will be developed to make the framework more accessible to users such as engineers and policymakers.

**Reference**


Computation 6 (2), S. 182–197. DOI: 10.1109/4235.996017.


Maier, Laura; David Jansen; Fabian Wülfhorst; Martin Kremer; Alexander Kümpel; Tobias Blacha, Dirk Mü ller; RWTH Aachen University, E.ON Energy Research Center, Institute for Energy Efficient Buildings and Indoor Climate (2016): AixLib: An open-source Modelica library for compound building energy systems from component to district level with automated quality management.

Marquant, Julien F.; Bollinger, L. Andrew; Evins, Ralph; Carmeliet, Jan (2018): A new combined clustering method to Analyse the potential of district heating networks at large-scale. In: Energy 156, S. 73–83. DOI: 10.1016/j.energy.2018.05.027.


Omu, Akomeno; Choudhary, Ruchi; Boies, Adam (2013): Distributed energy resource system optimisation using mixed integer linear programming. In: Energy Policy 61, S. 249–266. DOI: 10.1016/j.enpol.2013.05.009.

Pampuri, Luca; Belliardi, Marco; Bettini, Albedo; Cereghetti, Nerio; Curto, Ivan; Caputo, Paola (2019): A method for mapping areas potentially suitable for district heating systems. An application to Canton Ticino (Switzerland). In: Energy 189. DOI: 10.1016/j.energy.2019.116297.


