Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

Sustainable operation of district heating systems using dynamic hierarchical optimisation

Markus Hofmeister[1,4], Sebastian Mosbach[1,2,4], Jörg Hammacher[5], Martin Blum[5], Gerd Röhrig[5], Christoph Dörr[5], Volker Flegel[6], Amit Bhave[4], Markus Kraft[1,2,3,4]

released: July 6, 2021

[1] Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

[2] CARES
Cambridge Centre for Advanced Research and Education in Singapore
1 Create Way
CREATE Tower, #05-05
Singapore, 138602

[3] School of Chemical and Biomedical Engineering
Nanyang Technological University
62 Nanyang Drive
Singapore, 637459

[4] CMCL Innovations
Sheraton House
Cambridge
CB3 0AX
United Kingdom

[5] Stadtwerke Pirmasens Versorgungs GmbH
An der Streckbrücke 4
66954 Pirmasens
Germany

[6] Celron GmbH
Ludwigstraße 8
80539 Munich
Germany

Preprint No. 275

UNIVERSITY OF CAMBRIDGE

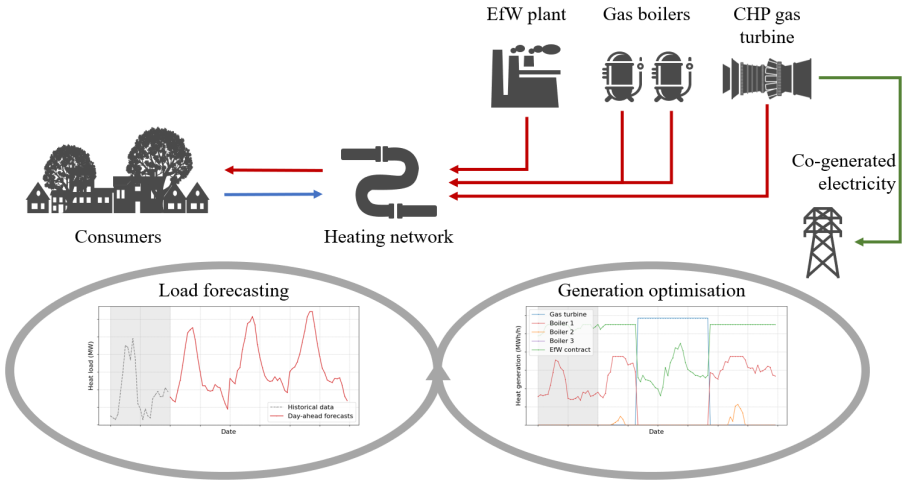
Keywords: district heating, generation optimisation, model predictive control, load forecasting

Edited by
Computational Modelling Group
Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

E-Mail: mk306@cam.ac.uk
World Wide Web: https://como.ceb.cam.ac.uk/

Abstract

Resource-optimised management of district heating networks needs to consider a wide range of factors, including demand forecasting, flexibility of the heat provision mix, and volatile market conditions. While traditional approaches often rely on static models and rather simple heuristics, dynamic cross-domain interoperability that allows the consideration of all these factors is essential to holistically optimise thermal grid operations. This paper demonstrates a proof-of-concept for a knowledge graph-based optimisation problem to minimise total heat generation cost for a district heating provider. The optimisation follows a hierarchical approach based on a merit-order principle and is embedded in a model predictive control (MPC) framework to allow the system to incorporate most recent information and react to disturbances promptly. A detailed sensitivity study is conducted to identify key design criteria and input parameters. Simulation-based optimisation is used to determine the short-term heat generation mix based on data-driven gas consumption models and day-ahead forecasts for the network's energy demand and grid temperatures. The proposed forecasting models deliver reliable and accurate predictions without any noteworthy difference in generation optimisation results when evaluated for forecasted versus actual historical values. The effectiveness of the approach is demonstrated for an existing heating network of a midsize city in Germany, where a reduction of approximately 20 % in operating cost and approximately 40 % in CO_2 emissions is obtained compared to baseline operational data.



Highlights

- Developed an MPC framework to optimise heat generation for a district heating network.
- Identified key cost drivers using a detailed sensitivity analysis.
- Developed high accuracy time series forecasting models for key input variables.
- Assessed improvement potential based on actual data from a municipal utility.
- Identified approximately 20 % and 40 % reduction potential for generation cost and CO_2 emissions, respectively.

Contents

1 Introduction

Global warming and climate change impose severe threats to human lives and livelihoods, environmental stability, and biodiversity. Hence, the reduction of greenhouse gas emissions is one of the predominant themes across scientific research, with a particular focus on the energy sector and the process industry. As the reduction of carbon emissions is a complex multi-dimensional problem, one of the greatest challenges to overcome is low interoperability between different domains and stakeholders.

Despite significant digitalisation advances in individual fields, the majority of data is still managed in isolated silos with limited information exchange capabilities due to different types, requirements, definitions, protocols and standards for data and software. Semantic modelling overcomes this limitation and offers the potential to incorporate an ever increasing variety of data, support automatic control operation, provide computer-aided decision making, and enable data sharing among different tools [26]. Such interoperability between individual components as well as heterogeneous information systems is essential for the next generation of smart grid applications [20] and requires various components and systems involved in the same application scenario to "automatically interpret each other's roles and 'understand' each other" [13]. Particular benefits of semantic technologies encompass separating the information semantics from the underlying data, formalising knowledge otherwise implicit in code or database schemas, and reducing modelling efforts by allowing the reuse of established ontologies [47].

District heating (DH) is expected to play an essential role in the cost-effective decarbonisation strategy of many countries [19, 27]. However, this requires increased utilisation of intermittent renewable energy sources and waste heat and shall also include district cooling systems to allow for holistically optimised future energy systems [29]. To stimulate the inclusion of (industrial) waste heat, a heat merit order is proposed to increase transparency on marginal generation cost [34]. A growing share of waste heat increases primary energy efficiency and may reduce cost; however, the value of a particular amount of heat supplied is usually not publicly available, which prevents further efficiency gains and liberalisation of the market. Moreover, consumer-side energy efficiency measures (*e.g.* building refurbishment) lead to decreasing overall heat demand and lower heat density, which increases cost and efficiency pressure on DH providers [19]. Although national policies (*e.g.* subsidies, co-generation incentives) influence the development and profitability of DH systems, the role of DH largely depends on local conditions and locally specific features, like municipal level energy strategies, available infrastructure, and local waste management systems. Hence, case studies on particular DH systems need to be performed rather than abstract country-level analyses [48].

Driven by their size and complexity, operational optimisation of energy networks, *i.e.* electricity, water, gas and heating networks can lead to significant energy and cost savings and are an active field of research. However, most previous work on optimising large-scale DH network operations only considers one aspect of the overall system and focus on either generation optimisation, *i.e.* load dispatch, or heat distribution, *i.e.* heat supply scheduling and temperature optimisation. However, an integrated optimisation perspective is a key prerequisite to minimise cost and resource intensity within a growing smart grid environment with increasing prosumer interaction. While current operations are of-

ten still based on heuristic methods, simple logic, or expert judgement [16], the intelligent control of DH systems is considered one of the key challenges by [29]. Given the systems' complexity and high parameters combinatorics, the determination of optimal production and distribution plans (*e.g.* load distribution across generator units, flow temperature) is practically impossible without any automated decision support tools.

A few approaches to conceptualise DH networks by means of ontologies can be found in the literature. However, these works are either quite limited in scope or focus on providing a semantic middleware between high-level energy management software and local components instead of representing actual digital twins [20, 28]. Knowledge graph-based models provide the capability to create sophisticated cross-domain digital twins that can be used to achieve substantial improvements in terms of energy and resource efficiency. Knowledge graphs leverage Semantic technologies [6] and the concept of Linked Data [5] to expressed data as a directed graph, where nodes represent concepts or instances and edges denote links between related concepts or instances. An essential prerequisite for implementing dynamic knowledge graph-based models is to virtualise and semantically represent the operations covered by the digital twin. This includes creating software models of the operations, creating ontologies to define the concepts and relationships covered by the software models, and deploying the ontologies and software models in a software ecosystem.

The **purpose of this paper** is to explore optimised heat generation strategies for a municipal utility of a midsize city. A hierarchical model to minimise generation cost is developed based on the merit-order principle and coupled to statistical forecasting models for key system variables in a model predictive control (MPC) framework. This model represents early progress towards the development of a semantic software ecosystem for municipal utilities.

The paper is structured as follows: Section 2 provides an overview of recent developments in district heating and suitable load forecasting methods. Section 3 describes the industrial case study setup and objectives. Model development is detailed in Section 4, with the results being discussed in Section 5 and conclusions being presented in Section 6.

2 Background

2.1 District Heating

Compared to decentralised heating options, DH systems often provide a more resource efficient alternative due to electricity co-generation and economy of scale [19]. DH networks have the potential to play an important role in sector coupling by compensating for fluctuations in renewable power generation and usage of waste heat from industrial processes instead of fossil fuels [32]. However, new trends in DH go beyond the shift towards greener generation and need to address the transformation of the network itself, including prosumer involvement and smart integration of mixed energy systems. Smart energy systems denote integrated smart electricity, gas and thermal grids, which are coordinated collectively to achieve an optimal solution for each individual sector as well as

for the overall energy system [29].

The role of DH in sector coupling and a low-carbon energy mix depends on more flexible operating modes, deeper insight into the system's dynamic behaviour, and more sophisticated simulation and optimisation tools. Building on previous work on gas and electric networks, the optimal control of a district heating networks has been studies using multiple techniques, including reduced order models [18, 39], the method of characteristics [44], and network abstraction introducing virtual single pipelines from source to consumption points [3]. While some approaches focus purely on minimising pumping cost as a result of pressure and mass flow variations, excluding thermal transport [18], other efforts study the efficient planning of input energy density using incompressible Euler equations and dynamic transport of thermal energy to describe a hot water DH network [39]. These works focus on precisely modelling the thermal transport dynamics between source(s) and consumption, address central difficulties in operating DH networks, and can help to optimise the distribution of injected energy density to better match the time-dependent demand. Deeper understanding of the time-dependent advection of injected heat can help to minimise the required peak load [39], allow for better integration of external over-capacities from renewable energies, and identify the best compromise between pumping costs and heat losses [16].

Besides optimising the injection and distribution, effective measures to plan and optimise resources for efficient energy production are essential [25]. Several studies focus on the sole production optimisation, *i.e.* addressing the unit commitment and load dispatch problem, but only a few works study both the load dispatch problem together with the supply temperature determination. A short-term production optimisation has been investigated by [42] by separating the unit commitment and the economic dispatch problem; however, relevant heat load forecasts have been created manually. The optimal generation allocation between individual plants has been studies in [14] using real-time smart metering customer data and a genetic algorithm with a static DH system model. Simple heat demand forecasts are constructed based on ambient temperature; however, this study considers simple heat boilers and simplified production cost only. Production optimisation and distribution including network storage has been investigated in [16]; however, this work relies on actual historical heat demand and does not include heat load forecasting. Several load forecasting methods have been proposed in the literature, including various filtering techniques, adaptive linear time-series models, grey-box approaches, and neural networks [2, 25]. While some forecasting approaches are solely focused on the production side, some argue that reliable load prediction methods need to account for the network storage effect as a key characteristic of large-scale DH networks [3]. The network storage effect describes the phenomenon that a significant amount of energy can be stored within the network due to varying supply volumes and temperatures. This introduces a temporal dependency between individual injected heat amounts and is particularly relevant as real-time consumption data is mostly not available for DH.

Semantic modelling is a suitable approach for integrated system optimisation and could help to address the difficulties in embracing the increasing complexity of current and future energy systems, minimise repetitive work and improve knowledge extraction from assessing numerous individual case studies. Ontology-based software ecosystems support the seamless integration and consolidation of dynamic data from different devices

(*e.g.* sensors, smart meters, actuators) and systems (*e.g.* building management system, energy managements systems) across heterogeneous data models and/or communication protocols. The feasibility of ontology-driven automation of the optimal coordination of district energy resources has been studied in [20]. A socio-technical ontology to conceptualise a DH network and its constituent buildings, including enhanced sensing and actuation infrastructure, as well as stakeholders has been proposed and web-based services for real-time decision making have been drafted. The role of the ontology is envisaged as intermediate layer between high-level energy management software applications and local devices. This work has been extended in [28] demonstrating its applicability for a real-time optimal control heat generation use case. A three-layer approach consisting of a sensoring and actuation layer, an interoperability layer using data models in the form of semantics, and an intelligence layer is proposed. The proposed ontology covers dedicated aspects of any district energy system; however, it is not publicly available.

2.2 Model Predictive Control

Model Predictive Control represents an important advanced control technique for difficult multivariate control problems [45]. MPC is applicable to any multiple-input, multiple-output process, which can be described by a dynamic process model with reasonable accuracy and has become the de facto standard algorithm for advanced control in process industries [36].

MPC algorithms consist of a dynamic process model, a cost function over a finite prediction horizon, and an (online) optimisation algorithm to derive the cost minimising sequence of control inputs: At each time step the current system state is sampled and the cost minimising control sequence is computed for a pre-defined prediction horizon; however, only the first step of the control strategy actually gets implemented. Afterwards, the system state is sampled again and the process model is used to predict future system states and derive an updated sequence of optimal control actions. The main advantage of this approach is that it allows the current time step to be optimised, while already considering future time steps. This forward shifting finite prediction horizons is a distinguishing feature of MPC [45], which is also referred to as receding horizon control. MPC can handle different kinds of system constraints efficiently, accounts for outside disturbances, and reacts to forecasting errors by re-optimising every time step.

MPC strategies have been extensively applied to various energy management and scheduling problems [40] and have been shown to improve operational decision making compared to traditional day-ahead control [49]. MPC has been applied to a variety of control scenarios related to DH networks, including the optimisation of supply temperatures considering network storage effects [3]. In a DH context, an MPC is usually composed of a load prediction module and an optimisation procedure to determine the best possible path for control variables while meeting different technical and operational constraints [16]. Depending on the objective function, operating costs and/or CO_2 emissions may be significantly improved.

2.3 Time Series Forecasting

Time series forecasting is essential for decision-making in a variety of disciplines, *e.g.* scheduling and demand planning, and identifying the best forecasting technique for each data set remains one of the key challenges [37]. Exploratory time series analysis is a powerful tool to identify relevant patterns and potential outliers in a time series to tailor one's forecasting strategy [22]. A short introduction of key features can be found in appendix A. The availability of historical data, the strength of relationships between the forecast variable and any explanatory variable, and the forecast length need to be considered when selecting a suitable family of methods and the subsequent selection of a single method with the best goodness-of-fit to complexity trade-off. An overview of selected error metrics used in this paper can be found in appendix B. While *pure* time series forecasting models predict the future value $\hat{Y}_{t+1}$ of the variable Y at time step $t+1$ solely based on historical observations at previous time steps and some random variation represented by an error term ε

$$\hat{Y}_{t+1} = f(Y_t, Y_{t-1}, Y_{t-2}, ..., \varepsilon), \tag{1}$$

dynamic regression models also include the influence of external predictors p [22]

$$\hat{Y}_{t+1} = f(Y_t, Y_{t-1}, Y_{t-2}, , ..., p_1, p_2, ..., \varepsilon). \tag{2}$$

Depending on the information used to compute a forecast, *ex-ante* and *ex-post* forecasts are distinguished. While *ex-ante* forecasts only use information available in advance, *ex-post* forecasts also consider forecasted information for external predictors.

Time Series Regression Time series regression predicts future values of the forecast variable based on linear relationships with external variables or time series, referred to as predictor variables. Categorical predictor variables can be modelled by *indicator* or *dummy* variables taking values of 0 or 1, respectively [22]. *Dummy* variables can also be used to account for outliers or to capture permanent level shifts in the data.

Naïve Forecast The naïve forecast, sometimes also referred to as persistence forecast, represents the simplest form of any forecasting method and is often used as benchmark for more elaborate forecasts. It assumes that the underlying time series can be described by a random walk. Hence, the best prediction for any future value $\hat{Y}_{t+h|t}$ within h future time steps corresponds simply to the last observed value Y_t

$$\hat{Y}_{t+h|t} = Y_t. \tag{3}$$

The seasonal naïve forecast considers the best prediction for any future value in h time step to correspond to the equivalent value of the last seasonal cycle of length m

$$\hat{Y}_{t+h|t} = Y_{t+h-m}. \tag{4}$$

Autoregressive Integrated Moving Average (ARIMA) Models

ARIMA models refer to a well-established and widely used class of statistical forecasting models and are still considered the dominant benchmark in empirical forecasting [37].

ARIMA models derive forecasts based on historical observations (*i.e.* autoregression (AR) model) and past forecast errors (*i.e.* moving average (MA) model). To fulfil the method's prerequisite of a stationary time series, one or multiple (seasonal) differencing steps might be required (*i.e.* integration). Differencing a time series is done by computing the differences between consecutive observations. Seasonal difference is the difference between an observation and the equivalent observation from the previous season.

An ARMA model of AR order p and MA order q, *i.e.* ARMA(p,q), for a sequence Y is given by

$$Y_t = c + \varepsilon_t + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{5}$$

where the drift parameter c denotes the average change between consecutive observations and ε_t represents white noise. The first summation term denotes the autoregressive part of the model as linear combination of the past p lagged values of the variable Y_{t-i} and the respective AR parameters ϕ_i, and the second summation term represents the moving average part as a weighted sum of the past q forecast errors ε_{t-j} and the respective MA parameters θ_j. Using the lag operator L, which shifts the index back one time step and can be treated using ordinary algebraic rules, *i.e.*

$$LY_t = Y_{t-1} \tag{6}$$

this expression can be significantly shortened using lag polynomials [22]

$$\begin{gathered} \phi(L) Y_t = c + \theta(L) \varepsilon_t \\ \text{with} \\ \phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i \\ \theta(L) = 1 + \sum_{j=1}^{q} \theta_j L^j. \end{gathered} \tag{7}$$

By including d differencing steps, a non-seasonal ARIMA(p,d,q) model is derived as

$$\phi(L)(1-L)^d Y_t = c + \theta(L) \varepsilon_t. \tag{8}$$

To incorporate seasonality, the Seasonal ARIMA (SARIMA) method has been developed by including further linear combinations of seasonal historical values and forecast errors, which are simply multiplied with the non-seasonal terms. An ARIMA model with a seasonal period m is abbreviated as ARIMA (p,d,q)(P,D,Q)$_m$ and defined by

$$\Phi_P(L^m) \phi_p(L) (1-L)^d (1-L^m)^D Y_t = \Theta_Q(L^m) \theta_q(L) \varepsilon_t, \tag{9}$$

with the lowercase indices p, d, q and uppercase indices P, D, Q denoting the non-seasonal and seasonal model orders, respectively. The AR orders of the model are given by p and P, the MA orders are specified by q and Q, and the number of differencing steps are specified by d and D. The non-seasonal AR and MA model parameters are given by ϕ_p and θ_q, while Φ_P and Θ_Q denote the seasonal AR and MA model parameters. Defining

z_t as the differenced time series, the full model can be expressed using the seasonal lag operator L^m as

$$\Phi_P(L^m)\phi_p(L)z_t = \Theta_Q(L^m)\theta_q(L)\varepsilon_t. \tag{10}$$

ARIMA models are remarkably flexible at handling a wide range of different time series patterns; however, appropriate hyperparameter selection might be difficult. Box and Jenkins proposed an iterative three-step approach to determine best-fitting SARIMA models [8], and even automatic model detection algorithms have been developed [23]. Having specified the hyperparameters of the model, the parameters ϕ_p, Φ_P, θ_q, and Θ_Q are estimated using maximum likelihood estimation. This technique estimates the values of the parameters which maximise the probability of obtaining the observed data from the model.

Although principally capable of modelling non-stationary data, significant forecasting accuracy improvements can be achieved with proper time series pre-processing to ensure stationarity, *i.e.* transformation and decomposition. Seasonal ARIMA models are not designed for long seasonal periods, and face both runtime and memory constraints due to the increase in parameters to be estimated. However, external predictor variables can easily be incorporated as exogenous regressors [12, 23]. Such SARIMAX models allow multiple and long term seasonalities to be included via external Fourier covariates.

Dynamic Regression Models

Dynamic regression models extend classical time series forecasting methods by including further external predictors, *e.g.* weather or holidays [22]. Classical ARIMA models can be rephrased to fit a regression model with ARIMA errors to a series of exogenous regressors.

While regression models generally require uncorrelated error terms, ARIMA models allow the errors from a regression to contain autocorrelation. If Y_t is a linear function of k predictor variables x_t, an regression with ARIMA errors η_t can be expressed as

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 x_{1,t} + \ldots + \beta_k x_{k,t} + \eta_t \\ \phi(L)(1-L)^d \eta_t &= \theta(L)\varepsilon_t, \end{aligned} \tag{11}$$

where β denotes the parameters from the regression model and ϕ and θ refer to the AR and MA parameters of the ARIMA model, respectively. As one can see, there are two error terms, η_t from the regression and ε_t from the ARIMA model; however, only the errors from the ARIMA model ε_t are white noise and shall be minimised when estimating the model parameters.

Stationarity of the time series as well as all predictors is an essential prerequisite to fit an regression model with ARMA errors. In case any of the variables needs to be differenced, it is best practice to difference all variables to maintain the form of the relationship between them. While forecasting with a regression model with ARIMA errors, both the regression and the ARIMA part are forecasted, and the results combined.

3 Case Study Description

The municipal utility provides district heating to several hundred private as well as multiple industrial customers and relies on two major heat sources: 1) internal heat generation using three conventional gas boilers and one combined heat and power (CHP) gas turbine (GT), and 2) external heat sourcing from an energy-from-waste (EfW) plant. **Figure** 1 depicts the general setup of the municipal utility operated DH system. Based on this setup, the utility provider has several degrees of freedom to satisfy the heat demand of its customers by scheduling (*i.e.* time and load) individual aggregates and external heat sourcing.

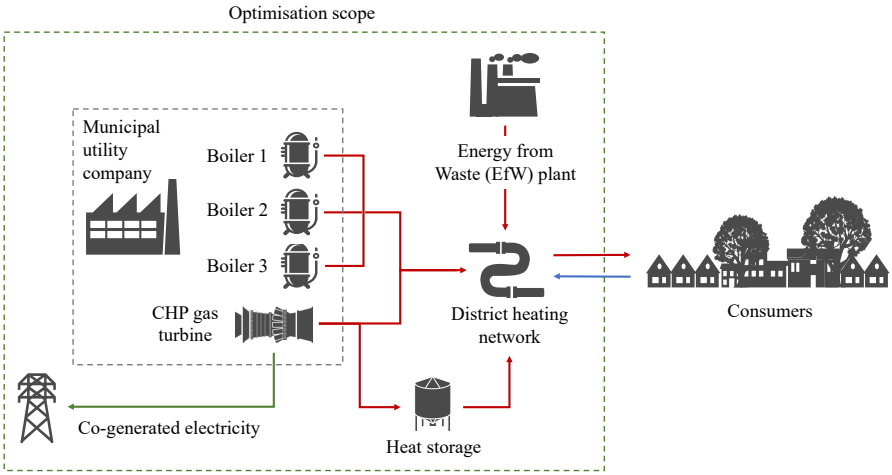


Figure 1: *District heating setup of municipal utility company and optimisation scope for total generation cost minimisation. Red arrows indicate heat provision by hot water, the blue arrow depicts return flow, and the green arrow denotes co-generated electricity provided to the grid.*

The operational cost of individual aggregates depends on a variety of constant parameters (*e.g.* heat capacity, efficiency, wear cost, *etc.*) as well as time-dependent parameters (*e.g.* gas price, CO_2 emission certificate (EUA) price, labour cost and shift surcharges, *etc.*). Furthermore, the availability of individual aggregates can change over time due to revisions, absent personnel, outages, *etc.* To also include potential revenues from co-generated electricity of the CHP GT, the electricity spot price is another key time-dependent variable. Additionally, several dependencies between individual variables and between subsequent time steps need to be considered. Functional dependencies between variables are primarily determined by the technical design of the system, *e.g.* the minimum amount of heat to be sourced from the EfW plant is a function on the spread of the flow and return temperature and the minimum circulation volume at the respective grid entry point. Temporal dependencies describe the influence of certain conditions in the past on any future decisions, *e.g.* the GT requires a minimum rest duration between any two subsequent operations, or the marginal heat unit price from the EfW plant decreases step-wise with an increasing annual volume sourced (*i.e.* volume rebate). Hence, the minimisation of total cost of internal heat generation and external sourcing represents a

complex time-series optimisation problem with dozens of variables as well as dynamic and interdependent constraints.

Currently, the optimal heat generation mix is determined several times a year based on high-level heuristics and long-term forecasts. Such a *traditional* approach heavily relies on personal experience and judgement, and is consequently rather limited to stable markets with known or easily predictable conditions. On the contrary, optimised decision making in dynamic markets with various interdependencies needs continuous adjustments to incorporate all information and data available. Therefore, an automatable *data-driven* optimisation framework is required based on the continuous integration of data from classical engineering, market data, and procurement and scheduling.

Therefore, this work aims to develop 1) a model to minimise total generation cost for the municipal utility, 2) understand key influencing factors by conducting a detailed sensitivity analysis, 3) develop reliable forecast models for relevant *internal* time series (*i.e.* heat demand, grid flow temperature, grid return temperature), and 4) ensure the continuous incorporation of updated *external* data (*i.e.* weather forecast, electricity spot price forward curve, gas spot price, *etc.*). Lastly, the optimisation framework will be assessed on the basis of real historical data from the utility provider.

4 Model Development

This section provides an overview of the methodology employed to develop the proposed optimisation framework for minimising total heat generation cost for a district heating network. This framework includes an operation optimisation (section 4.1), an MPC integration (section 4.3), and a prediction engine to forecast heat demand and grid temperatures for the upcoming optimisation horizon (section 4.4). An object oriented programming approach in Python has been used to ease future conceptualisation of used classes, attributes, and relationships into an intelligent semantic system.

4.1 Optimisation Problem

The proposed optimisation aims to minimise total generation cost by controlling the operational schedules of individual generation units, while safeguarding that required heat demand is met. Operational expenses consist of fix cost per hour c^{fix} irrespective of the generated amount of heat, *e.g.* personnel and wear cost, and purely demand driven variable cost c^{var}, such as fuel and emission certificate cost as well as potential revenue for co-generated electricity sold to the grid (including potential subsidies and saved grid charges). The total unit cost for generating heat amount q on generation unit g at time step t can be calculated by

$$\text{opex}(q_{g,t}) = \frac{1}{q_{g,t}} \cdot \sum_{c^{\text{fix}} \in C^{\text{fix}}} c_t^{\text{fix}} + \sum_{c^{\text{var}} \in C^{\text{var}}} c_t^{\text{var}} - P_{\text{CHP, el}}(q_{g,t}) \cdot p_t^{\text{el}} \tag{12}$$

where C^{fix} and C^{var} denote all applicable hourly fix and marginal variable cost, respectively, $P_{\text{CHP, el}}$ is the electrical co-generation power and p_t^{el} summarises all revenues per

generated unit of electricity. Modelling any external heat sourcing as generation unit with only demand dependent variable cost, the total cost minimisation can be expressed as

$$\begin{aligned} \min_{q} \quad & \sum_{t=1}^{T} \sum_{g \in G} \left(\text{opex}(q_{g,t}) \cdot q_{g,t} + c_{g,t}^{\text{switch}} \right) \\ \text{s.t.} \quad & \sum_{g \in G} q_{g,t} = Q_t \end{aligned} \tag{13}$$

with T being the number of optimisation time steps, G denoting the collection of available heat generation units and Q_t the actual heat demand to be satisfied at any given time step. The switching cost $c_{g,t}^{\text{switch}}$ for generator g at time step t include both potential start-up (in case generator g is used at time step t, but has not been active at time step $t-1$) and shut-down cost (in case generator g is not used at time step t, but has been active at time step $t-1$).

Several further constraints need to be considered when optimising overall heat generation. Key constraints include: 1) A minimum amount of heat is supplied to the grid at all times by maintaining a minimum circulation to ensure hydraulic stability. This amount of heat must be provided by municipal utility owned heat generators. 2) The availability of individual generator units can change over time due to (un-)planned downtime or personnel constraints. 3) The GT can only be restarted after a certain minimum resting period after each operation. 4) Several generator units can only operate above a certain minimum threshold (*i.e.* minimum GT load due to emission restrictions, minimum external sourcing due to hydraulic conditions). 5) External sourcing from the EfW plant is based on a take-or-pay contract, which incentivises external sourcing.

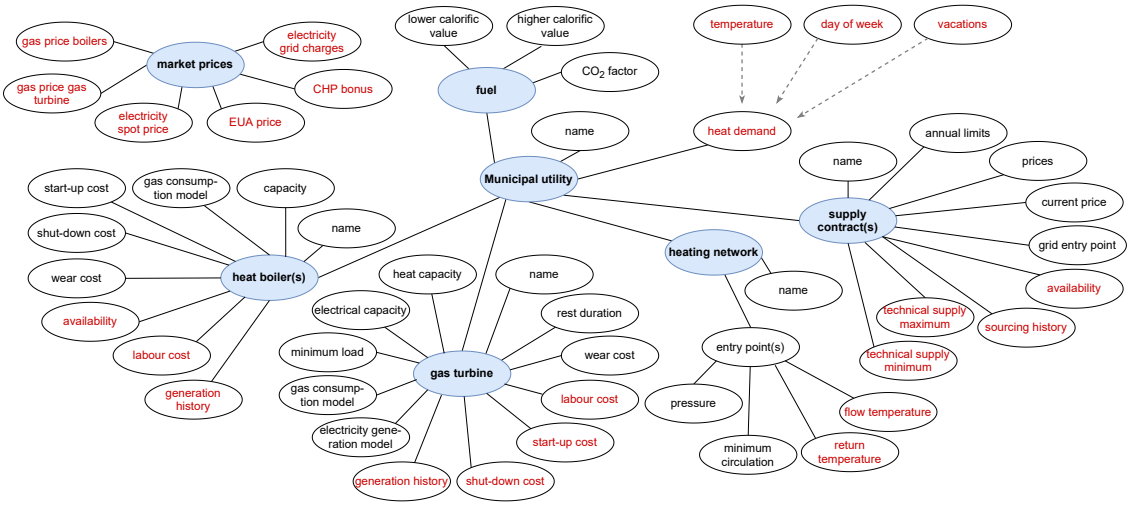


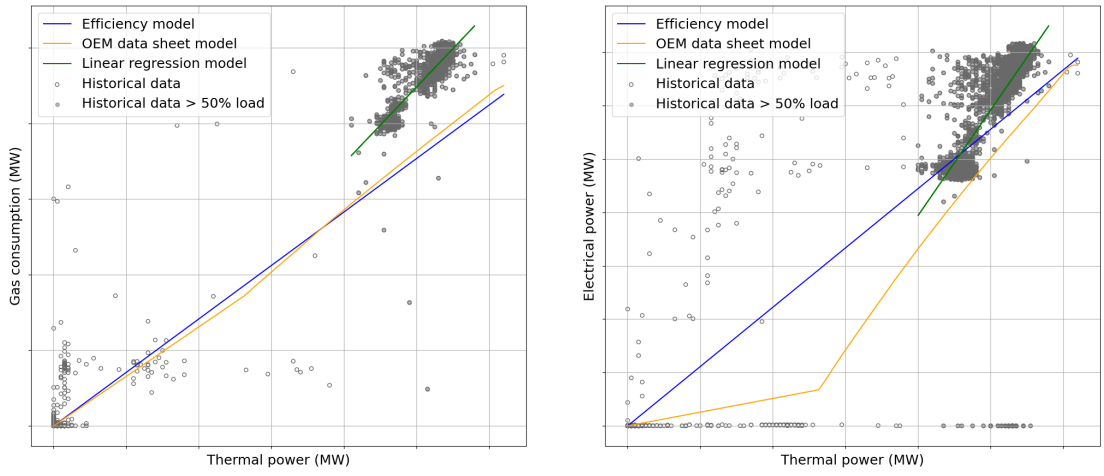
Figure 2: *Class structure of virtual municipal utility representation: blue shaded elements represent classes, black elements denote constant attributes, and red elements represent time series attributes. Dashed arrows indicate the influence of external variables on total heat demand.*

Figure 2 depicts the proposed class structure to represent all required information, including relevant concepts, attributes and relationships. As the available heat storage is

associated with significant heat losses and was barely used in recent years, it has not been considered in the overall optimisation setup. While time-independent parameters are implemented as constant values or sets of constant values (*e.g.* annual volume limits and associated prices for EfW purchases), time-dependent variables are implemented as time series with hourly frequency. Gas consumption, electricity co-generation, and minimum heat supply are implemented as separate functions to account for internal dependencies. Gas demand and electricity co-generation models have been derived based on historical generation and consumption data as depicted in **Figure** 3. Linear regression models are proposed, as both original equipment manufacturer (OEM) data sheet parameters and provided efficiency values could not explain the relationship between actual observed values with sufficient accuracy.

The proposed optimisation framework follows a hierarchical approach consisting of two key steps. First, two principal heat generation modes are defined and associated operational cost are minimised individually. A *mode_excl_GT* represents heat generation without GT (*i.e.* the entire heat demand needs to be provided by conventional gas boilers and external sourcing), and *mode_incl_GT* denotes heat generation with prioritised GT contribution. Subsequently, the cost optimal sequence of both modes is derived as the final optimisation result. This two-step approach allows to initially optimise both modes purely with regards to demand-driven cost (equation 12) and incorporates event-driven switching cost between modes only later, after the successful evaluation that incurred switching benefit justifies associated cost. This logic also allows for temporary negative incremental benefit, if it is overcompensated later.

Algorithm 1 summarises the key steps in optimising the individual generation modes: For each time step, the minimal generation cost are evaluated by ranking available gen-



(a) Total gas consumption.

(b) Electricity co-generation.

Figure 3: *Total gas consumption and electricity co-generation models for CHP gas turbine (axes anonymised for confidentiality reasons). As historical values cannot be explained with constant efficiency or OEM data sheet parameters, linear regression models have been implemented.*

Algorithm 1: minimise_interval_cost: Minimise heat generation cost for a given set of heat generators and a certain number of time steps.

Input: A set *heat generators* of all heat generators to be considered, and a number *optimisation period* of optimisation time steps to perform
Output: Cost optimised heat generation distribution, *i.e.* heat amount per time step per generator in *heat generators*

for $t \leftarrow 0$ **to** *optimisation period* **do**
    /* Assess minimum amount of heat to be supplied to the grid to ensure hydraulic stability */
    min supply $\leftarrow$ function(T_{flow}, T_{return}, Q_{min}, p,)
    demand $\leftarrow$ max(*min supply*, actual demand[t])
    /* Derive *ranked capacities* of *heat generators*, ranked by priority and unit price (*i.e.* merit-order) [algorithm 3] */
    ranked capacities $\leftarrow$ get_ranked_capacities(*heat generators*, *min supply*)
    /* Derive optimal heat generation mix to satisfy heat demand with minimum total cost [algorithm 4] */
    opt generation $\leftarrow$ minimise_current_cost(*ranked capacities*, *demand*)
    /* Include potential start-up and shut-down cost incurred by changes in active heat generators [algorithm 5] */
    if $t > 0$ **then**
        switching cost $\leftarrow$ assess_switching_cost(*opt generation*$[t-1]$, *opt generation*$[t]$)
        Add potential *switching cost* to current operational expenses
end
return *opt generation*

eration capacities based on ascending unit price and subsequently selecting the cheapest combination of capacities required to satisfy the demand for that time interval (*i.e.* merit-order-principle [34]). While ranking capacities, temporarily unavailable aggregates are excluded for the respective time interval and further constraints (*e.g.* provision of grid stabilisation heat only by own generators) are considered. The *mode_incl_GT* prioritises heat generation with the CHP gas turbine over all other capacities, once higher-order constraints are met. Finally, unavoidable switching cost are included, *e.g.* GT shut-down cost in case GT has been active and heat demand drops below GT's minimum load. Details on the subroutines mentioned in Algorithm 1 can be found in the appendix.

Assuming rather constant pressure conditions in the hot water district heating network, the minimum heat supply Q_{min} can be derived by

$$Q_{\text{min}} = \dot{V}_{\text{min}} \cdot \rho \cdot c_{\text{p}} \cdot \Delta T \tag{14}$$

with

$$c_{\text{p}} = \frac{c_{\text{p}}(T_{\text{flow}}, p) + c_{\text{p}}(T_{\text{return}}, p)}{2} \qquad \rho = \frac{\rho(T_{\text{flow}}, p) + \rho(T_{\text{return}}, p)}{2}$$

with $\dot{V}_{\text{min}}$ as minimum volumetric flow rate (given by minimum pump speed), ΔT denoting the difference between flow and return temperature, and c_{p} and ρ being the average water heat capacity and density, respectively.

Having derived the optimal generation mix and cost of both modes for each time interval, a holistic time-integrated optimisation across both modes is conducted as described by **Algorithm** 2. Iteratively stepping through the individually optimised generation time

Algorithm 2: optimise_generation: Optimise heat sourcing from internally optimised generation modes excl. and incl. gas turbine.

Input: Two multivariate time series (*mode_excl_GT* and *mode_incl_GT*) of internally optimised heat generation (*i.e.* cost-optimised generation mix per time step) and flag *gt_active* whether GT has been active in previous time step

Output: Multivariate time series of overall cost-optimised generation mix

```
   /* Assign current and alternative heat generation mode                    */
1  if gt_active then current ← mode_incl_GT else current ← mode_excl_GT
2  if gt_active then alternative ← mode_excl_GT else alternative ← mode_incl_GT
   /* Initialise time stepping, optimal generation, and maximum and latest
      cumulative benefit from current GT operation                           */
3  t ← 0
4  optimum ← []
5  gt_benefit ← [0,0]
6  while t < length(current) do
7      if current generation cost ≤ alternative generation cost then
8          optimum[t] ← current[t]
9          if gt_active then
10             Update maximum and cumulative gt_benefit
11         t ← t+1
12     else
           /* Assess benefit of potential switch between modes if alternative
              becomes cheaper than current [algorithm 6]                     */
13         switch, period, cost ← assess_switching_period(current, alternative, gt_benefit)
14         for dt in period do
15             if switch = True then
16                 optimum[t] ← alternative[t]
17                 Include switching cost cost in optimum at time step dt := 0
18                 if gt_active then
19                     Reset gt_benefit ← [0,0] at time step dt := 0
20             else
21                 optimum[t] ← current[t]
22                 if gt_active then
23                     Update maximum and cumulative gt_benefit
24             t ← t+1
25         end
26 end
27 return optimum
```

series, the algorithm selects the most cost effective operation of both modes based on the previous setpoint and identifies all points in time where potentially switching between modes could result in lower total cost. Subsequently, it evaluates whether the cumulative benefit from a prospective switch exceeds the associated switching cost (*i.e.* start-up and shut-down cost) and, if at all, for what duration the generation mode shall be switched. Details of this evaluation logic can again be found in the appendix and are schematically depicted in **Figure** 4. Each switching between pre-optimised modes needs to be justified by additionally generated profit. In case the GT is currently inactive, all intervals with a cumulative benefit exceeding associated switching cost between modes are considered profitable. In case the GT is currently active, GT operation will be maintained until cumulative benefit drops by at least switching cost from previous maximum. The duration

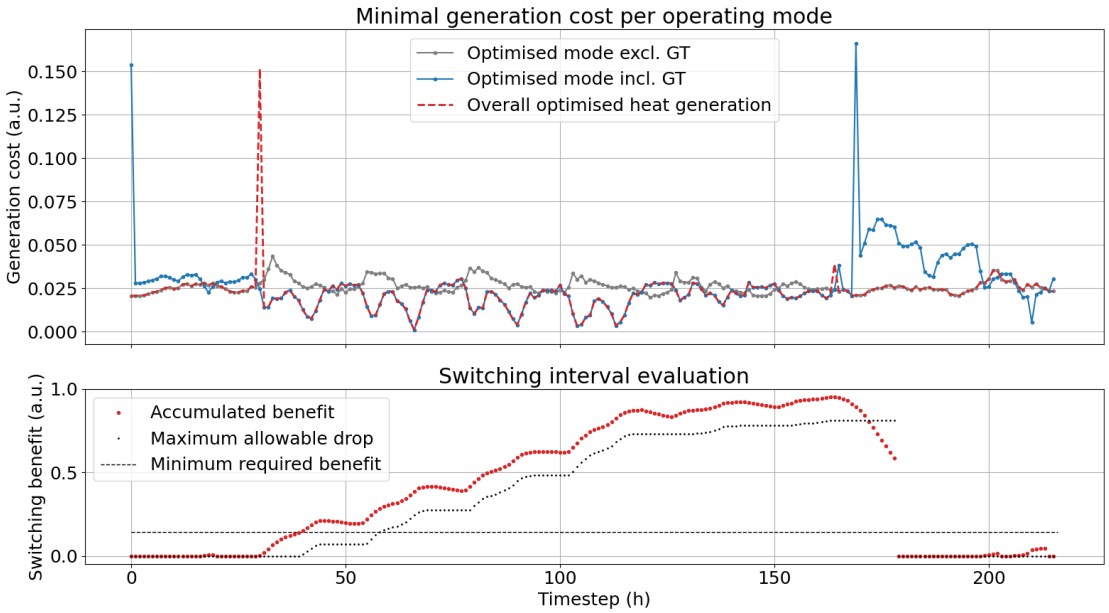


Figure 4: *Schematic depiction of switching period evaluation with monetary values displayed in arbitrary units for confidentiality reasons: Switching operating modes shall be conducted if resulting cumulative benefit exceeds associated switching cost and shall be maintained until cumulative benefit drops by at least switching cost from maximum.*

for which operating modes shall be switched is determined by the time step of maximum accumulated benefit. Initial switching cost from current mode to alternative mode are included at the beginning of the switching interval, while final switching cost back from alternative mode to current mode are considered at the end of the interval.

4.2 Sensitivity Analysis

Sensitivity analysis is a tool to study the effect of changes in input parameter values on the output value of a simulation model and can act as a screening process to identify the key model parameters [17]. Local sensitivity analyses how a small perturbation in the vicinity of an input space value influences the final output value. A classical approach to assess local sensitivities is to compute the normalised sensitivity coefficients using the *one factor at time* method, for which one input parameter is varied while holding all other parameters fixed [33].

Given a function $\eta(\xi^{(n)}, \theta)$ with the input space vector $\xi^{(n)}$, the normalised sensitivity coefficient A_j with respect to the j^{th} model parameter θ_j is defined as [33]:

$$A_j = \frac{\theta_j}{\eta(\xi^{(n)}, \theta)} \cdot \frac{\partial \eta(\xi^{(n)}, \theta)}{\partial \theta_j}. \tag{15}$$

A small positive relative perturbation r of model parameter j can be denoted as:

$$\tilde{\theta}^j := (\theta_1, \ldots, \theta_{j-1}, (1+r)\cdot\theta_j, \theta_{j+1}, \ldots, \theta_n). \tag{16}$$

Applying a finite difference approximation, the normalised sensitivity coefficient can be calculated by:

$$A_j = \frac{\theta_j}{\eta(\xi^{(n)}, \theta)} \cdot \frac{\eta(\xi^{(n)}, \tilde{\theta}^j) - \eta(\xi^{(n)}, \theta)}{(\tilde{\theta}^j - \theta)_j} = \frac{\eta(\xi^{(n)}, \tilde{\theta}^j) - \eta(\xi^{(n)}, \theta)}{r\eta(\xi^{(n)}, \theta)}. \tag{17}$$

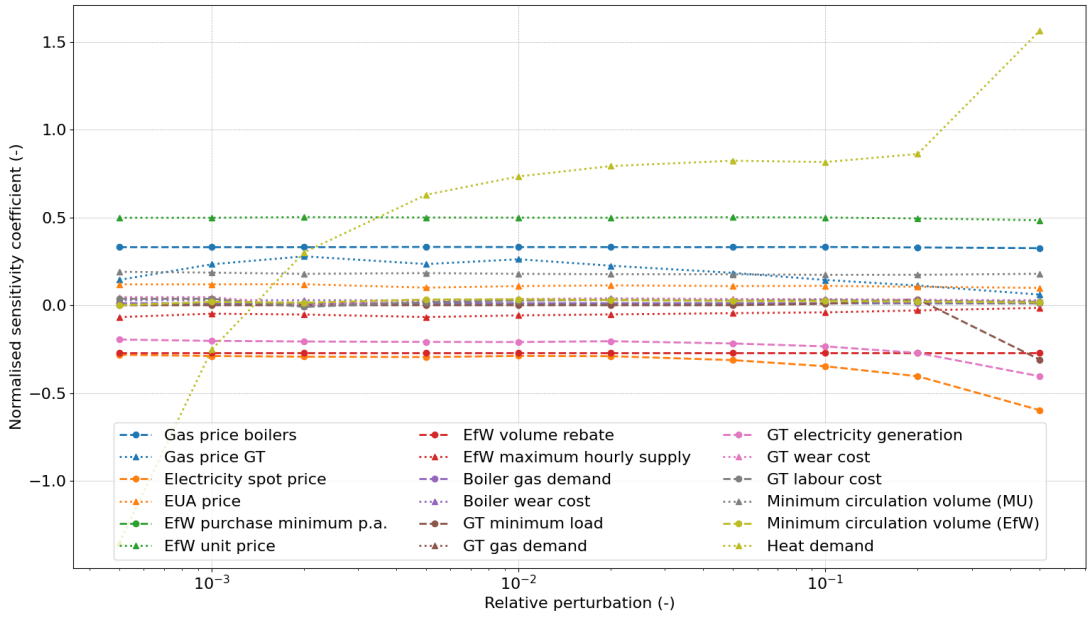


Figure 5: *Normalised sensitivity coefficients of key model parameters for a range of relative perturbations: heat demand, EfW heat unit price, and fuel prices are the most influential input parameters.*

Figure 5 depicts the normalised sensitivity coefficients for key input parameters of the proposed optimisation model. This analysis is based on the total heat generation cost for one full heating period (*i.e.* one entire year) and has been conducted with actual historical values for any time series input to exclude potential forecasting distortion. Furthermore, a relative perturbation of any input parameter represented by a time series means that all values of this series are adjusted accordingly.

It can be obtained from Figure 5 that heat demand, the EfW heat unit price, and fuel cost (gas prices for both conventional boilers and GT) are the most influential input parameters. Furthermore, revenues from co-generated electricity are decisive, with the electricity spot price and GT co-generation power as further key inputs. However, a potential volume rebate on the EfW unit price has been found more relevant than the GT fuel cost or co-generation power alone. The minimum heat supply requirement due to the municipal utility's constraint on minimum pump speed, and the EUA price are further sensitive inputs.

4.3 Model Predictive Control Implementation

The optimisation problem is integrated into a MPC framework to allow the system to continuously react to updated information. **Figure** 6 depicts the proposed MPC loop: In each time step (1h in this study) actual measurements from the system (*e.g.* observed heat load) along with external forecasts are provided to the MPC controller. Within the controller, internal forecasts (*i.e.* heat load and grid temperatures) are updated and the model minimises total generation cost over a pre-determined *optimisation horizon* t_{mpc} (24h in this study). The model determines the optimal generation mix over that entire period in single time step intervals. Although a sequence of t_{mpc} control moves is calculated, only the first setpoint is actually implemented. Subsequently, the *optimisation horizon* shifts forward and the MPC loop starts again, either indefinite or until a maximum number of time steps to evaluate t_{end} has been reached.

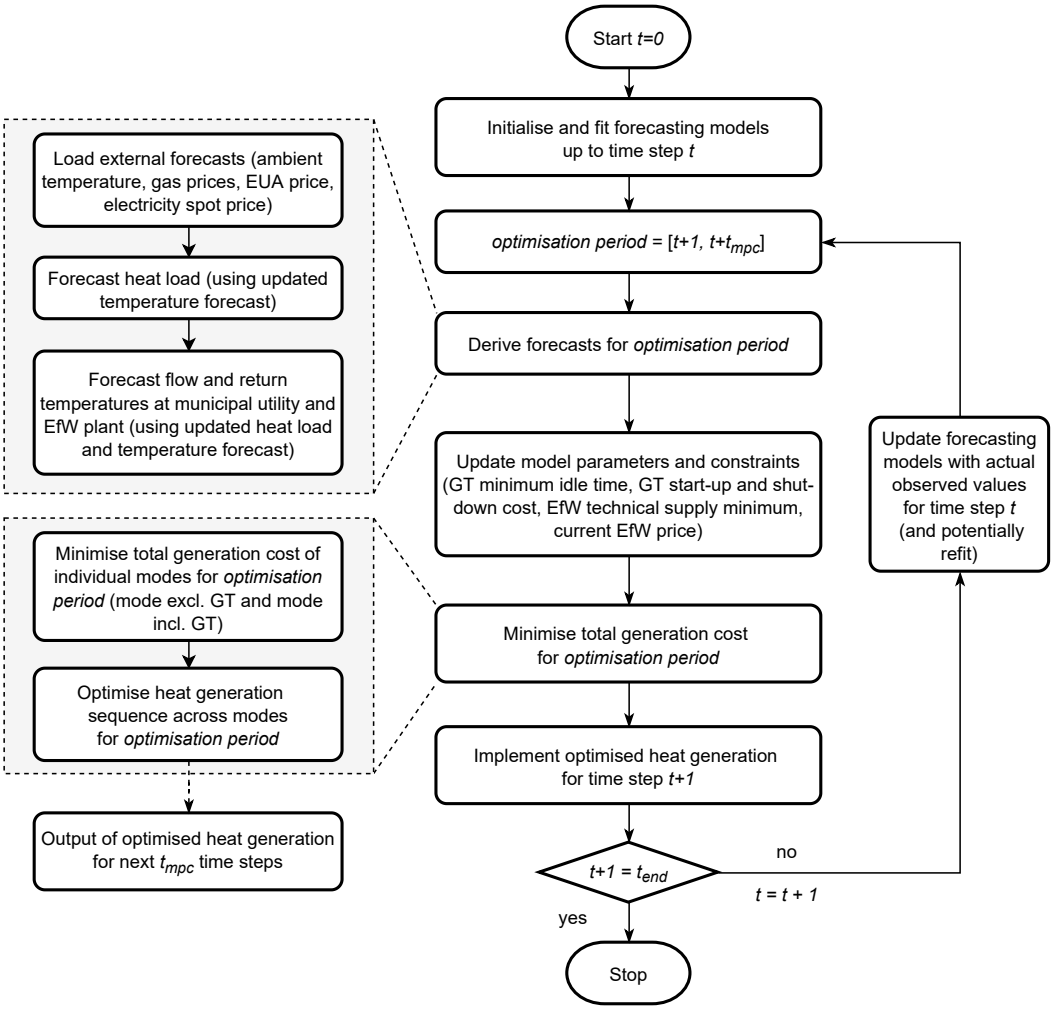


Figure 6: *Flow chart of MPC implementation of heat generation optimisation for optimisation horizon* t_{mpc} *and total number of optimisation time steps* t_{end}.

4.4 Forecasting Models

Continuous operation optimisation requires reliable forecasts for heat demand and grid temperatures as input parameters. Considering the nature of these variables, a suitable forecasting method needs to satisfy several requirements: 1) capability to handle high-resolution data (*i.e.* hourly data points on a time scale of years), 2) flexibility to model seasonalities on different time scales (*i.e.* daily and annual consumption patterns), 3) multi-step forecasting capabilities to allow for extended forecasting periods (*e.g.* annual forecasts for long-term planning), and 4) incorporation of continuous (*e.g.* temperature) and discrete (*e.g.* day of week, holidays) external data. Although having issues with long seasonalities, seasonal ARIMA models with exogeneous regressor variables can satisfy all these requirements.

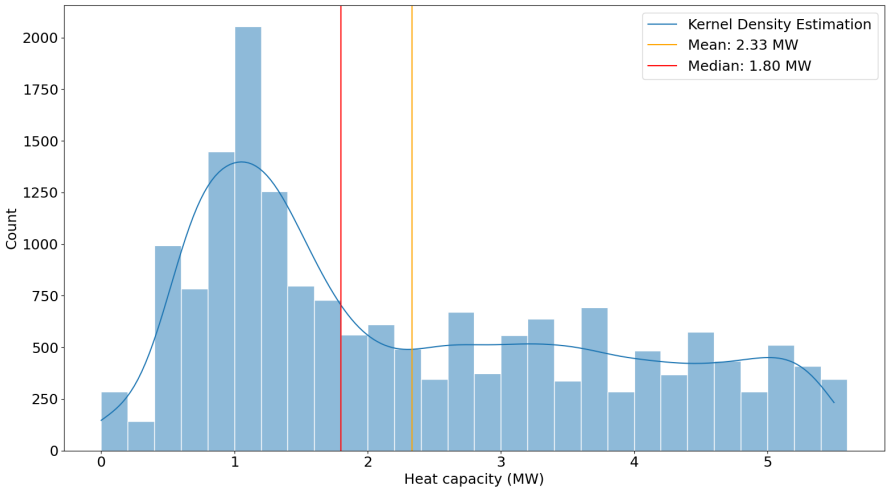


Figure 7: *Swing capacity on most expensive heat generator for optimised generation without forecasting. Swing capacity is the minimum of used and idle capacity and can be used to assess the minimum required accuracy of a suitable heat demand forecast.*

To estimate the minimum required accuracy of a suitable heat load forecast, an analysis of the swing capacity on the most expensive heat generator unit in an optimised generation scenario has been conducted. Swing capacity refers to the minimum of the used and idle capacity on the respective generator. The underlying logic is that any heat demand variation below that swing capacity would not have triggered a different optimised scheduling decision. Hence, any forecast method with an error below the average swing capacity is well suited. **Figure** 7 depicts the swing capacity distribution for the optimised generation scenarios of two full heating periods. It can be seen that a suitable forecast method needs to yield an average absolute forecast error below $\sim$2.0 MW.

As shown in section 4.2, heat demand is the single most important input parameter. Hence, external predictor variables shall be incorporated to increase prediction accuracy. Previous load forecasting works for both electricity and heat demand revealed previous load history (*i.e.* optimal number of lags), calendar effects (*i.e.* weekday vs. weekend, hour of day), and ambient temperature as important input features [2, 7, 9, 10]. Further weather vari-

ables, such as humidity, wind speed, and solar radiation have been considered by [7, 25]. A distinction between holiday and non-holiday has been proposed by [9] and [10] included categorial variables to capture the effect of greatly changing loads on Saturday and Monday, *i.e.* the change caused by the transition from working days to rest days. In this work, ambient temperature is used as a key external predictor variable. Furthermore, categorial variables for weekday, weekend, Monday, Saturday, holiday, and school vacation are implemented as one-hot encoded variables. To capture annual seasonal effects, additional Fourier terms are considered. No feature scaling or normalisation has been conducted due to comparable scales of all input parameters.

Although ACF and PACF plots are established tools to derive the orders of moving average and auto-regressive models, respectively, these techniques are only valid for *pure* AR and AM models [38]. It is possible for AR and MA terms to cancel each other's effects in a mixed ARMA model, which is why visual identification of appropriate model orders becomes problematic and other techniques for hyperparameter estimation shall be used. In this study, Sobol points are used to identify best fitting hyperparameters. Sobol sequences describe quasi-random low-discrepancy sequences, which have the ability to cover any given domain approximately evenly. Besides this advantage over purely random numbers, Sobol sequences do not require a predefined number of samples and their accuracy improves continually as more data points are added, which offers benefits compared to deterministic sequences [46].

A seasonal ARIMA model of the form SARIMA(X)$(p,d,q)(P,D,Q)$ requires fitting of 6 hyperparameters. To constrain the hyperparameter space, two established guiding principles are used: 1) the order of seasonal differencing shall not exceed one and the number of total differences shall not exceed two, and 2) seasonal coefficients larger than two are barely required [38]. The remaining 6-dimensional hyperparameter space encompasses d and D in the range [0 to 1] and p,q,P, and Q in the range [0 to 2].

For each variable to be forecasted, several models (*i.e.* 200 models for heat load and 100 models for grid temperatures) have been sampled from the hyperparameter space using Sobol points (and subsequently converted into integer model parameters). Each model has been fitted to 12 individual data histories and its prediction accuracy for day-ahead in sample predictions has been evaluated. A representative selection of these 12 forecast evaluations is ensured using equidistant spacing throughout a full heating period with a comparable share of each weekday. Based on the average RMSE and maximum error of these in-sample predictions, a subset of 30 models has been shortlisted for each variable and evaluated on another set of 50 day-ahead in-sample predictions.

It has been observed, that several shortlisted SARIMAX model configurations perform comparably well and yield average absolute errors well below 2.0 MW. **Figure** 8 compares the forecast errors for day-ahead heat load predictions of four different methods. It can be seen that the SARIMAX model clearly outperformes a simple naïve, seasonal naïve and pure temperature regression model. The distribution of forecast errors is both narrower compared to the other methods and unbiased (*i.e.* mean error of zero).

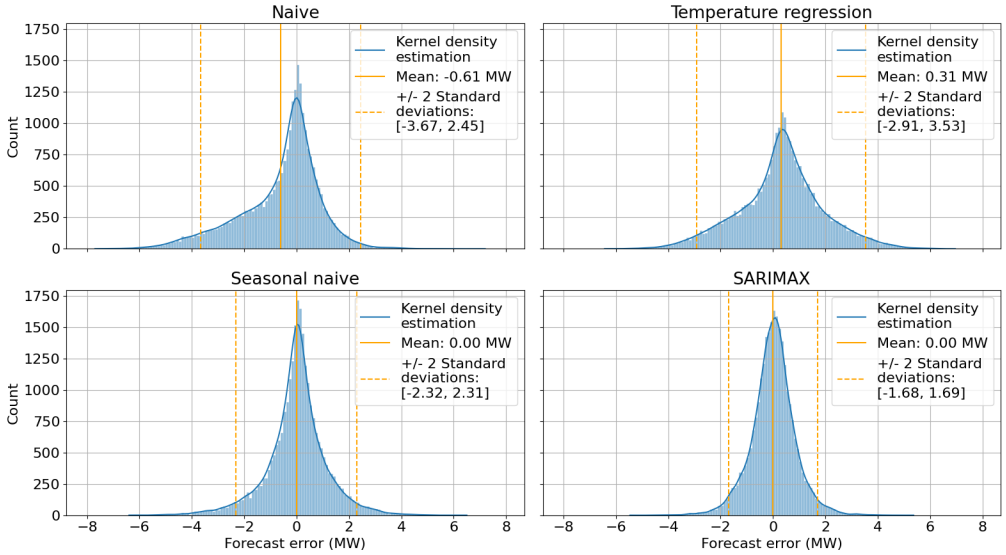


Figure 8: *Comparison of heat load prediction errors for different forecasting methods. Depicted forecast errors refer to the performance for hourly values of one-day ahead forecasts for a period of three consecutive years.*

5 Results and Discussion

The development of the MPC embedded generation optimisation revealed several key insights. The sensitivity study re-emphasised the importance of a reliable heat demand forecast as key input parameter and a minimum required accuracy has been quantified. The importance of the contract details with the EfW plant have been shown, with the base unit price and a potential volume rebate having a major influence on annual heat provision cost. As historical data shows that the current contract has never been fully exploited, this information shall be used to foster future contract negotiations with the EfW plant. Furthermore, the minimum circulation volume in the hot water district heating system and the associated minimum heat provision has been identified as another important cost driver, which has been neglected so far.

Although initially intended as sole benchmark for more advanced machine-learning forecasting methods (a short description can be found in appendix C), statistical SARIMA(X) models have been found sufficiently accurate and precise for this case study. SARIMA(X) models are well-established and robust forecasting models and the prediction errors of best-fitting models are shown to be insignificant with regards to the results of the optimisation model (see section 5.3).

5.1 Data Quality Issues

Scrutinising real industry data during this work revealed several severe data quality issues and re-emphasised the importance of a semantic data representation. While some drawbacks are associated with data handling (*e.g.* high manual work load and long provision

times, inconsistent formats, *etc.*), the most severe issues were associated with the data quality itself: The current data management system does not support any automated sense checking of newly integrated data, leading to large amounts of meaningless and nonphysical values, *e.g.* the majority of historical GT gas consumption contained negative values. Furthermore, large variances in historical values (*e.g.* specific gas consumption of conventional boilers) have been detected, which potentially indicate calibration issues at selected meters.

While single outliers can be detected and removed, *e.g.* by using median absolute deviation filters, systematic measurement inconsistencies cannot easily be accounted for by any mathematical data transformations. One fundamental issue in the current data management is the blending of hourly average and instantaneous values. The amount of heat provided to the grid is measured and stored as hourly average, while the generation of individual aggregates is recorded as instantaneous values, which makes a meaningful historical evaluation of generation and provision basically impossible. Consistent data representation and reporting frequencies are crucial to ensure precise comparisons and build models, which can effectively learn from the past.

The observed shortcomings in the current data management emphasise the necessity for (near) real-time data plausibility checking and data consolidation. Furthermore, semantically represented data can help to improve anomaly detection by automatically understanding the influence of a certain value of one variable on a related variable.

5.2 Forecasting Results

Having screened multiple SARIMA(X) model configurations for each relevant variable, the final forecasting model has been defined as the configuration with the lowest average RMSE. It could be observed that forecast quality is largely influenced by the length of the used fitting history and the interval between potential model refitting. **Table** 1 provides an overview of the best-fitting SARIMA(X) configurations for heat load and grid temperatures at both relevant grid entry points. While the optimal length of data history used to fit the model varies between the variables, a refitting interval of 2 to 3 months has been found as a good balance between additional computational cost and updating model parameters to more recent data.

Figure 9 depicts the forecasting performance of the derived heat load model for a representative winter month. Although peak loads tend to be underestimated, the model is able to predict the actual data reasonably well. However, it can be seen that forecast accuracy for the first half of a day is tentatively higher than for the second half. As heat demand in this study actually refers to the amount of heat provided to the network (*i.e.* due to the absence of reliable heat consumption data on the consumer side), slight over- and underpredictions of heat demand within a day can easily be compensated for by the inertia of the district heating grid and under- or overproduction in a subsequent time step, respectively (*i.e.* network storage effect). However, it needs to be noted that these forecasts are generated with actual observed temperature data and no forecast values as external predictor variable. The influence of temperature forecast errors, especially with regards to longer forecasting horizons, has not explicitly been investigated.

Table 1: *Summary of best-fitting SARIMA(X) forecasting models.*

Forecast variable	Non-seasonal			Seasonal part			m	Fitting history	Exogeneous predictors
	p	d	q	P	D	Q			
Heat load	1	0	2	1	1	1	24	365 days	ambient temperature, annual Fourier seasonality, weekday, weekend, Saturday, Monday, holiday, vacation
Municipal utility flow temperature	1	0	0	2	0	1	24	21 days	heat load, ambient temperature
Municipal utility return temperature	2	1	1	1	0	1	24	21 days	-
EfW plant flow temperature	1	0	1	1	1	1	24	14 days	heat load, ambient temperature
EfW plant return temperature	2	1	1	1	1	1	24	14 days	-

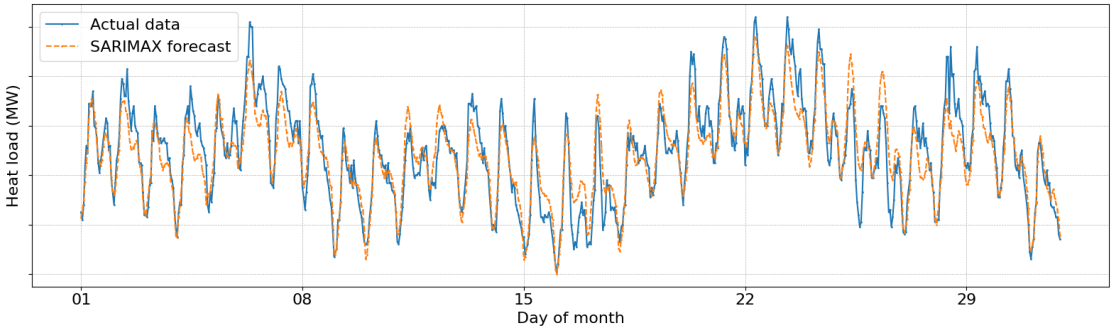


Figure 9: *Comparison of actual heat demand and one-day ahead forecast values for a representative winter month (axes anonymised for confidentiality reasons).*

Figure 10 illustrates the performance for the two worst days within a forecasting period of three consecutive years, revealing an average hourly deviation between actual and predicted values of 2.6 MW and 2.1 MW, respectively. However, it needs to be mentioned that these forecasts refer to multi-step day-ahead forecasts without hourly updates of actual observed heat load values.

Table 2 provides an overview of key accuracy metrics for the best-fitting SARIMA(X) models. It can be obtained that all SARIMA(X) models yield better performance than their seasonal naïve benchmarks. The average heat load error is well below 1.0 MW and grid temperatures can be predicted with an average accuracy of ± 2.0 °C. The high MAPE value for the heat load is primarily an effect of low heat demand during summer periods, with average heat load being well below 2.0 MW.

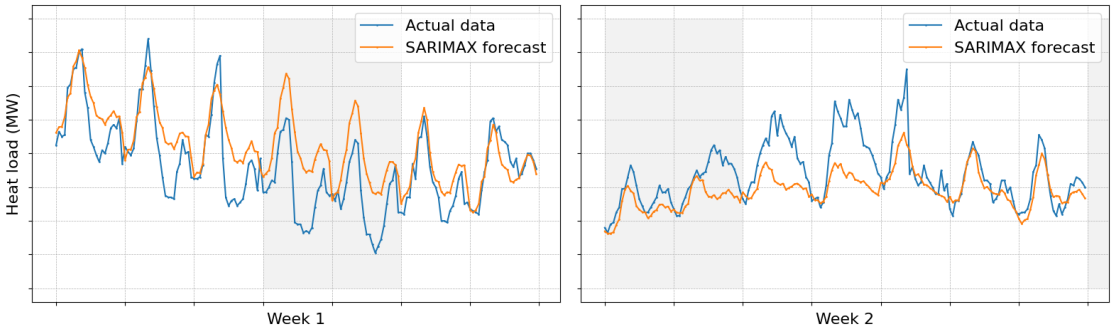


Figure 10: *Forecast performance for days with highest daily deviation from actual observed values in a forecasting period of three consecutive years (axes anonymised for confidentiality reasons). Gray shaded areas denote weekends.*

Table 2: *Overview of several error metrics for best-fitting SARIMA(X) models. MAE denotes the Mean Absolute Error, RMSE refers to the Root Mean Square Error, and MAPE stands for Mean Absolute Percentage Error. Percentage Better refers to the share of SARIMA(X) predictions with lower absolute error than seasonal naïve benchmark.*

Forecast variable	**MAE**	**RMSE**	**Maximum Error**	**MAPE**	**Percentage Better**
Heat load	0.62 MW	0.84 MW	5.46 MW	16.3 %	59.0 %
Municipal utility flow temperature	1.36 °C	2.00 °C	20.22 °C	1.5 %	61.9 %
Municipal utility return temperature	1.21 °C	1.64 °C	11.51 °C	1.6 %	57.7 %
EfW plant flow temperature	1.46 °C	2.08 °C	15.37 °C	1.6 %	61.8 %
EfW plant return temperature	1.1 °C	1.42 °C	7.65 °C	1.6 %	54.4 %

5.3 Optimisation Results

This section discusses the results obtained from the MPC embedded optimisation problem including internal forecasting of heat load and grid temperatures and external data updated within each time step, *i.e.* every hour. Based on the nature of the evaluation, it needs to be noted that these results have been generated using the historical electricity spot price instead of the hourly price forward curve (HPFC), which would be applied for forward-looking applications. The effect of using historical spot prices instead of HPFC prices, especially for longer forecasting periods, has not been investigated. Two separate full years with different market situations have been studied: *Year 1* refers to a year with relevant CHP co-generation incentive, while *Year 2* lacks any comparable subsidies. Furthermore, *Year 2* is associated with higher labour cost.

Table 3 summarises the key results of both optimisation studies and **Figures** 11, 12, and

13 illustrate the optimised generation for *Year 1*. All figures are to scale, but anonymised due to confidentiality reasons. Optimised scheduling of heat generator units and EfW sourcing resulted in annual total cost savings of ∼26 % and ∼20 %, respectively. Increasing the current share of EfW sourcing can, furthermore, significantly reduce the amount of gas consumed. Associated CO_2 reduction potential has been assessed solely based on the amount of saved gas. Two current limitations to be noted: First, the model assumes that all heat generator units can be activated within an one hour period. While this holds true from a technical perspective, this might need to be adjusted to account for less flexible personnel planning. Second, the CO_2 contribution from the EfW plant is not explicitly considered and the reduction potential is derived by the single-objective cost minimisation. An overarching and multi-objective optimisation, including cost, emissions, *etc.*, is likely to yield even higher reduction potential.

Table 3: *Summary of optimisation results for two full heating periods.*

	Identified improvement potential p.a.			
Year	**Heat generation cost**	**Gas demand**	**CO_2 emissions**	**EfW purchase**
Year 1	−26.2 %	−50.7 %	−4913 t	+22.2 p.p.
Year 2	−20.3 %	−38.3 %	−2181 t	+9.0 p.p.

Figure 11 compares the actual historical GT operation with the optimised generation schedule. It can be seen that the optimisation utilises the GT more hesitantly. Besides a more thorough economic evaluation prior to any potential GT start-up, this observation is likely to be attributed to the omission of the available heat storage. As the GT requires a minimum heat load of ∼9.0 MW, already slight drops in the anticipated heat load will suggest that the generated heat cannot immediately be accommodated in the grid and consequently suppress GT operation.

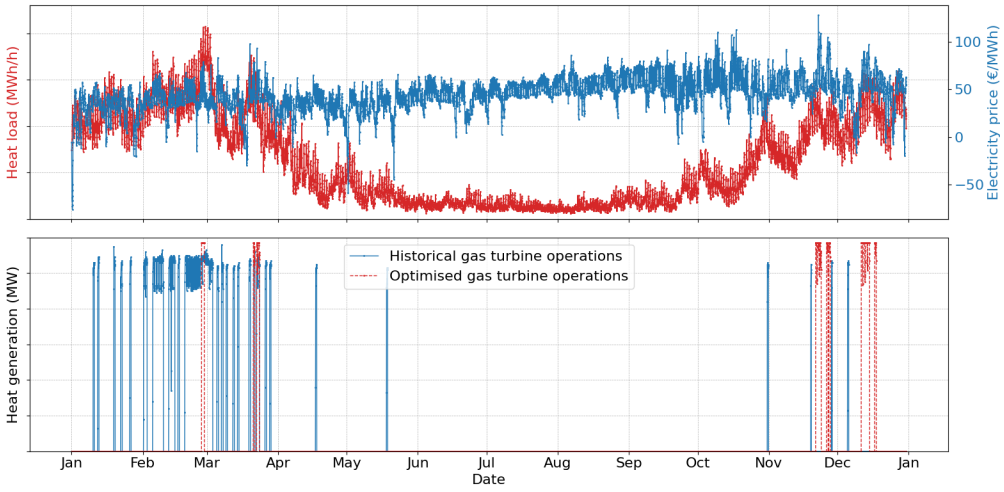


Figure 11: *Comparison of optimised and historical CHP gas turbine operations (bottom) and associated heat load (axes anonymised for confidentiality reasons) and electricity spot price (top).*

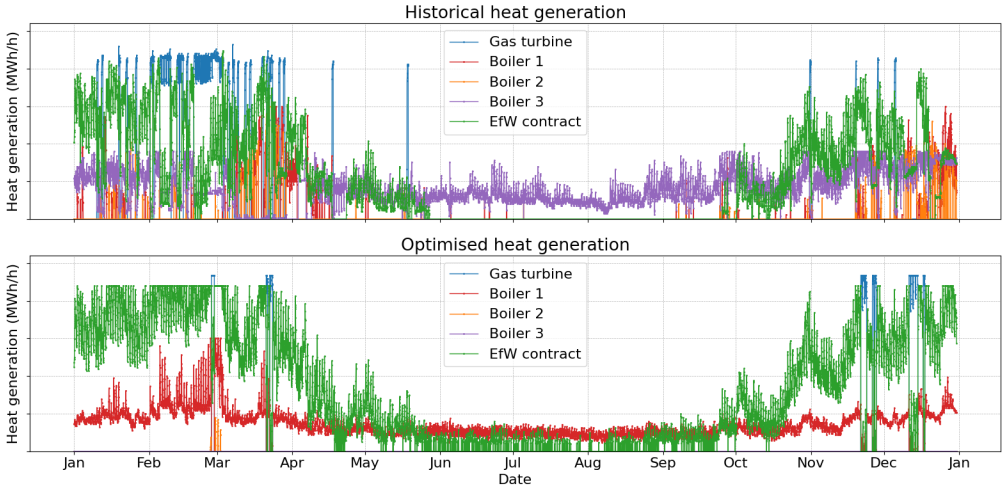


Figure 12: *Comparison of optimised and historical heat generation mix (axes anonymised for confidentiality reasons).*

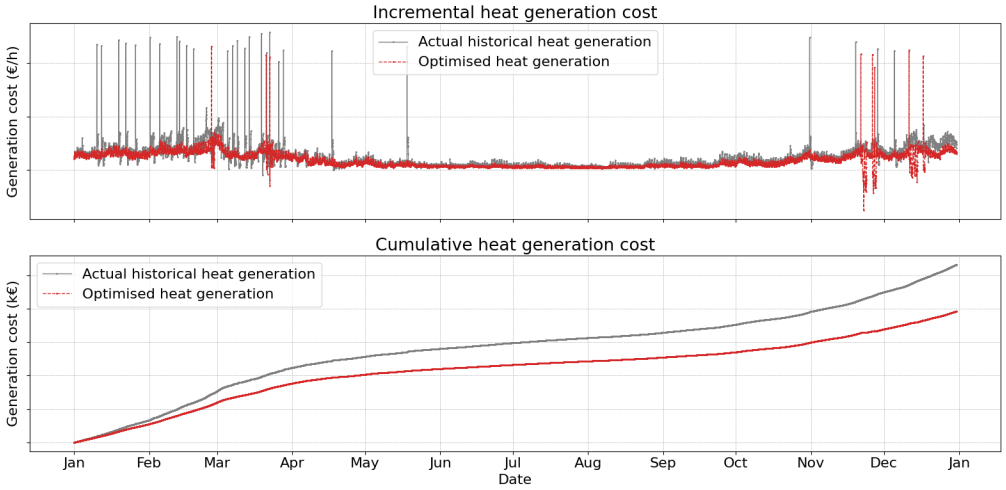


Figure 13: *Comparison of incremental and cumulative cost for optimised and historical heat generation (axes anonymised for confidentiality reasons).*

Figure 12 illustrates the historical and optimised generation schedule for *Year 1* and **Figure** 13 depicts the associated incremental and cumulative generation cost. It can be observed that the optimised generation mix follows a more rigorous selection logic and, hence, is characterised by a less variable generator composition. Furthermore, it becomes obvious that sourcing heat from the EfW plant is of superior importance to minimise generation cost. Historically, the negotiated heat provision volumes from the EfW plant have only been used to ∼48 % and ∼36 % in *Year 1* and *Year 2*, respectively. After the optimisation this percentage increases to ∼55 % for both years. Although this value suggest even further room for external sourcing, this share is limited by 1) a maximum hourly provision limit due to maximum pump speed at the EfW plant and 2) the constraint that the minimum circulation of the network needs to be maintained by the municipality itself.

It needs to be noted that the actual network hydraulics are not explicitly modelled. Hence, assuming constant pressure and forecasting flow and return temperatures at the relevant grid entry points solely based on heat load and ambient temperature is a significant simplification and does not account for any hydraulic induced supply limitations.

A comparative analysis to assess the influence of forecast errors has been conducted for both years. It has been revealed that an optimisation with actual historical values instead of forecasts from the proposed models results in no considerable deviation for the optimised generation schedule and total cost. This supports the initial analysis of minimum required heat load forecast accuracy and the suitability of the developed SARIMA(X) models.

To assess the impact of increasing regulatory pressure, *i.e.* increasing CO_2 certificate and fuel prices, normalised local sensitivity coefficients for representative perturbations have been compared for the actual historical market environment and an anticipated market situation with respective price increases (see **Figure** 14). While gas prices have been adopted based on the latest contract data from the municipal utility, an EUA price of 50 €/t is assumed [43]. It can be obtained that anticipated price changes will further increase the relevance of contractual terms with the EfW plant (*i.e.* unit price and potential volume rebate) and significantly increase the importance of an efficient CHP GT operation. Increasing emission cost will foster the share of cost-effective CHP generation and further increases the importance of a dynamic and automatable control due to highly volatile electricity spot prices.

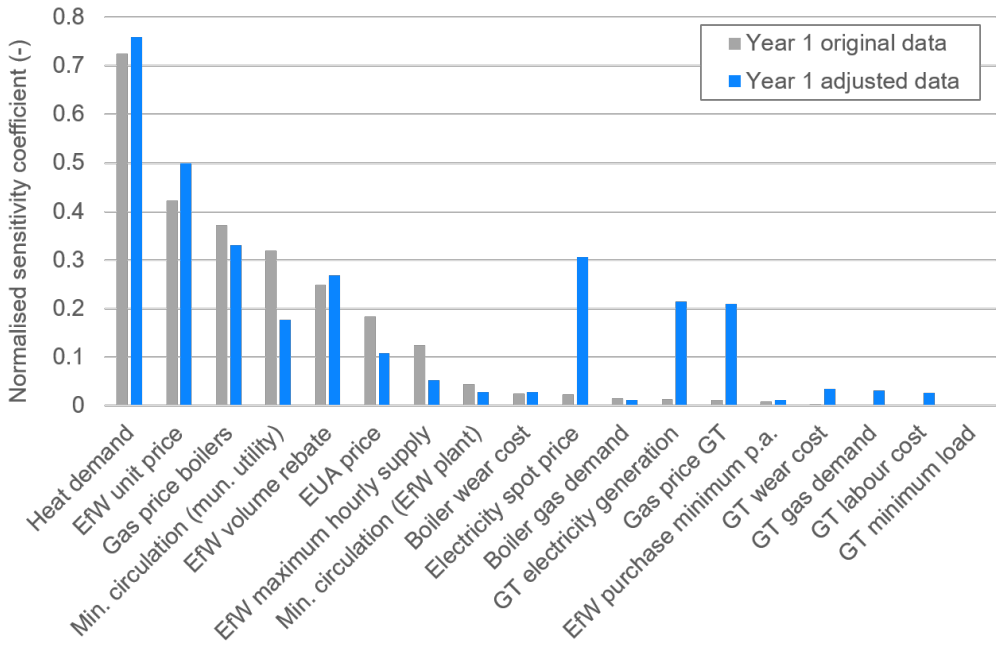


Figure 14: *Comparison of normalised local sensitivity coefficients of key model parameters for actual historical market environment (i.e. Year 1 original data) and anticipated changes in EUA and gas prices (i.e. Year 1 adjusted data).*

A fundamental limitation of the proposed approach is the omission of temperature and load-dependent effects on gas consumption and electricity co-generation. Although the

current models describes the historical data sufficiently close, these limitations need to be considered when comparing optimised and actual historical generation data.

6 Conclusions

An MPC framework to optimise heat generation and sourcing for a district heating provider, including the operation of its CHP gas turbine has been accomplished. A detailed analysis of the data and the current operations has revealed both technical and economic effect mechanisms and dependencies to foster fact-based decision making. A hierarchical optimisation approach has been derived as a set of transparent algorithms to improve current operations, which are still primarily driven by heuristics and personal judgement. A systematic understanding of long-term cost-drivers and short-term optimisation measures has been identified.

The optimisation problem has been integrated into a model predictive control framework to continuously update short-term forecasts and incorporate latest observations. One day-ahead heat demand and grid temperature forecasting models have been developed, considering also external predictor variables (*e.g.* weather and calendar effects). The forecasting accuracy of classical statistical seasonal ARIMA models has been found sufficiently accurate for all variables, and especially precise for anticipated heat load. The optimisation framework has identified annual cost saving potential above 20 % for two individually assessed years with different market dynamics and revealed a prospective CO_2 abatement potential exceeding 40 kg/MWh. The increasing importance of a dynamic and holistic generation optimisation is illustrated using local sensitivity analyses, including anticipated shifts in general market conditions.

Several severe data quality issues have been identified and clear areas of mitigation measures have been proposed, outlining the superior importance of a systematic semantic representation of data.

Acknowledgements

This research was supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. M. Hofmeister acknowledges financial support provided by the Cambridge Trust and CMCL. M. Kraft gratefully acknowledges the support of the Alexander von Humboldt Foundation.

A Time Series Analysis

A time series is a collection of observations ordered by time and is usually modelled as a sequence of a random variable $Y(t)$. Time series forecasting is focused on estimating future realisations of this stochastic process $Y(t+h)$, using only information available at time t. While forecasting could generally be understood as a subset of supervised regression problems, special temporal dependencies need to be accounted for in time series data.

Relevant time series components include trends, seasons, and cycles. A *trend* refers to a long-term increase or decrease in the data (linear or non-linear). A *season* denotes a recurring behaviour with fixed and known period and *cycles* reflect fluctuations without fixed frequency. The average length of cycles tends to exceed seasonal patterns, combined with a more variable magnitude than seasonal patterns [22]. As time series data often exhibits superimposed patterns, several decomposition approaches have been proposed in the literature [4, 11]. Furthermore, transformations (*e.g.* logarithmic and power transformations) and adjustments (*e.g.* calendar and inflation adjustments) are often applied on complex time series [15, 22]. Various statistical methods have been developed to forecast seasonality, cycle, trend, and randomness in an integrated manner [37].

To describe the strength of temporal dependency within a time series, autocorrelation measures the linear relationship between values of a sequence at two different points in time [8]. The autocorrelation coefficient r of a sequence x at lag k is defined by

$$r_k = \frac{\sum_{t=k+1}^{T}(x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{T}(x_t - \bar{x})^2}, \tag{A.1}$$

where $\bar{x}$ denotes the mean of the sequence and T is the length of the time series. As correlation between two variables can result from both direct and indirect dependence, the partial autocorrelation defines the autocorrelation between values of a sequence at two different times after the removal of any linear dependence on observations at intermediate lags, *i.e.* confounding [8]. The (partial) autocorrelation function (PACF/ACF) plots the (partial) autocorrelation coefficient as a function of consecutive time lags and can be used to detect periodicity and patterns in time series data [1]. Trended data shows strong positive autocorrelation for small lags, with decreasing values for increasing lags. Seasonal data shows strong autocorrelation at multiples of the seasonal frequency.

B Goodness of Fit Measures

Multiple measures have been proposed to evaluate the performance of different forecasting methods, each associated with characteristic advantages and disadvantages depending on how individual forecast error are summarised [12]. A forecast error e_t denotes the difference between an observed value Y_t and its prediction $\hat{Y}_t$:

$$e_t = Y_t - \hat{Y}_t. \tag{B.1}$$

A good forecasting method leads to uncorrelated residuals with zero mean. While correlation between residuals indicates information left in the residuals, which should rather be

used to improve the forecast, a mean other than zero results in a biased forecast. Furthermore, normally distributed residuals with constant variance ease the calculation of prediction intervals [22]. To compare the overall performance of various forecasting methods, a selection of measures should be considered. A brief overview of selected methods is provided in the following.

Scale-dependent errors Scale-dependent errors have the same scale as the data. Hence, they cannot be used for comparisons between different time series. The most commonly used accuracy metrics are the mean absolute error (MAE) and root mean square error (RMSE) (equation (B.2)). While minimising MAE leads to forecasts of the median, minimising the RMSE results in forecasts of the mean of the time series. RMSE penalises extreme errors during forecasting.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \tag{B.2}$$

Percentage errors Percentage errors are unit free and frequently used for comparisons between data sets, with the mean absolute percentage error (MAPE) being the most commonly used metric (equation (B.3)). One of the key disadvantages of these measures are their extreme values for Y_t values close to zero [30].

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{e_t}{Y_t}\right| \cdot 100\% \tag{B.3}$$

Prediction comparisons To provide an alternative to percentage errors when comparing forecasts across series with different units, *scaled errors* have been proposed [24]. This approach ensures comparable error scales by scaling the individual errors based on the training MAE from a simple forecast method. Similarly, the *percentage better* methodology benchmarks the accuracy of one method against the accuracy of a reference (and often simpler) forecast method [12].

C Artificial Neural Network Forecasting Models

Neural Network Autoregression (NNAR) A NNAR model is a feedforward artificial neural network with lagged values as inputs for forecasting univariate time series. The number of lagged observations to be considered defines the number of input neurons. For seasonal time series, those neurons may also include the seasonally lagged sequence [22]. Although the model has initially been developed for univariate time series, exogenous variables can be incorporated as additional regressors. Such NNARX models have been shown to significantly improve forecasting performance compared to both SARIMAX and NNAR models without external data [31, 41]. For forecasting purposes, the network is applied iteratively. While one-step forecasts simply use the available historical inputs, multi-step forecasts use the one-step forecasts as inputs, along with the historical data.

Long Short-Term Memory (LSTM) LSTM denotes a recurrent artificial neural network architecture, mostly used with unstructured data (*e.g.* audio, video) [21]. Unlike feedforward neural networks, a LSTM model has feedback connections, which allow to keep track of dependencies of new observations with historic ones. LSTM networks are well-suited for time series forecasting due to their ability to detect and consider lags of unknown duration between important events in a sequence. They are capable of automatically learning features from sequence data, support multivariate data, and can infer and forecast non-linear dependencies among multiple time series. Furthermore, LSTMs can output sequences of variable length and therefore be used for multi-step forecasting. Compared to classical methods, no assumptions regarding the underlying nature of the time series are required. Hence, physical time series such as metering or monitoring sensor outputs tend to be very good use cases for LSTM models. It has been show that LSTMs can clearly outperform classical SARIMAX forecasts [35] and yield more accurate predictions than feedforward neural networks [7] for electric load forecasts, especially in the short-term.

D Optimisation Algorithm (continuation)

Algorithm 3: get_ranked_capacities: Derive list of available capacities sorted by increasing priority and unit price

Input: A set *heat generators* of all heat generators to be considered, and a minimum amount of heat *min supply* to be supplied to the grid to keep it stable

Output: A set *capacities* as subset of *heat generators* with available generation capacities sorted by increasing priority (1st), increasing unit price (2nd), and decreasing capacity(3rd)

```
1  Initialise capacities as empty list
2  Retrieve min heat suppliers as subset of heat generators capable of providing stabilisation heat
   /* Perform 2 consecutive runs through heat generators to rank available
      generation capacities by increasing unit price and prioritise
      min supply requirement                                              */
3  for run ← 0 to 1 do
4      foreach generator in heat generators do
5          if generator is gas turbine then priority ← 2 else priority ← 3
6          if generator is available then
7              if generator in capacities then
8                  Add new generator entry with remaining capacity, only marginal operational
                     cost, and priority to capacities
9              else
10                 Add new generator entry with total generator capacity, full operational cost
                     (incl. potential hourly fix cost), and priority to capacities
11     end
12     if run = 0 then
13         Delete all capacities entries for generators not in min heat suppliers
14         Sort capacities by increasing priority (1st), increasing unit price (2nd), and decreasing
             capacity(3rd)
15         capacities sorted ← first n entries of capacities required to satisfy min supply
16         foreach entry in capacities sorted do
17             if entry's capacity is only required partially then
18                 Adjust entry's capacity to reflect only required quantity
19                 Re-evaluate generation cost for adjusted capacity
20             capacities ← Updated capacities sorted with priority := 1 for all entries
21         end
22 end
23 Sort capacities by increasing priority (1st), increasing unit price (2nd), and decreasing
     capacity(3rd)
24 return capacities
```

Algorithm 4: minimise_current_cost: Assess minimal cost required to satisfy current heat demand

Input: A ranked (i.e., merit-order) list *capacities* of available generation capacities, and required heat *demand* to be satisfied

Output: Cost-optimised heat generation distribution across available sources *opt generation*, and total generation/sourcing expenses *cost*

```
1  if gas turbine in capacities' heat generators then
2      if demand ≥ gas turbine minimum load then
3          Summarise all individual gas turbine entries in capacities into single entry
4      else
5          Remove all gas turbine entries from capacities
6          capacities ← get_ranked_capacities(remaining heat generators in capacities,
             min supply)
7  cost ← 0
8  active ← []
9  foreach entry in capacities do
       /* Add capacities until demand is satisfied                          */
10     if demand > 0 then
11         demand ← demand - entry's capacity
           /* If entry's capacity is needed in full                         */
12         if demand ≥ 0 then
13             cost ← cost + entry's capacity × entry's unit price
14             if entry's generator in active then
15                 Increment required capacity for entry's heat generator by entry's capacity
16             else
17                 Add entry to active
           /* If entry's capacity is only needed partially                  */
18         else
19             if entry's generator is gas turbine or sourcing contract then
20                 if demand < generator's minimum requirement then
21                     demand ← demand + entry's capacity
22                     continue
23             if entry's generator in active then
24                 opex ← Re-evaluate marginal variable cost for demand
25                 cost ← cost + opex
26                 Increment required capacity for entry's generator by demand
27             else
28                 opex ← Re-evaluate full generation cost for demand (incl. potential hourly fix
                     cost and demand dependent variable cost)
29                 cost ← cost + opex
30                 Update entry's capacity to demand and add to active
31 end
32 return active, cost
```

Algorithm 5: assess_switching_cost: Assess switching cost between two heat generator setups (from $setup1$ to $setup2$)

Input: Two sets of heat generators $setup1$ and $setup2$
Output: Total shut-down and start-up cost to switch from $setup1$ to $setup2$

```
1  cost ← 0
2  Remove all external heat sourcing contracts from setup1 and setup2
3  foreach generator in setup1 do
4  |   if generator not in setup2 then
5  |   |   Add generator shut-down cost to cost
6  end
7  foreach generator in setup2 do
8  |   if generator not in setup1 then
9  |   |   Add generator start-up cost to cost
10 end
11 return cost
```

Algorithm 6: assess_switching_period: Assess whether switching between two internally optimised heat generation modes is beneficial

Input: Two multivariate time series (*current* and $alternative$) of internally optimised heat generation modes and tuple of maximum and latest cumulative $gt\_benefit(max, cum)$ from current gas turbine operation

Output: Decision whether heat generation mode shall be switched, associated cost of switching, and the period for which the modes shall/shall not be switched

```
1  switch ← False
2  first neg ← True
3  acc_benefit[max, cum] ← gt_benefit(max, cum)
4  for t ← 0 to length(current) do
       /* Assess incremental benefit/loss inc and update cumulative benefit
          from potentially switching modes acc_benefit[cum]                  */
5      inc ← current generation cost - alternative generation cost
6      if gt_benefit > 0 then inc = −inc
7      acc_benefit[cum] ← acc_benefit[cum] + inc
8      if inc < 0 then
           /* Find maximum accumulated benefit if it starts to drop         */
9          if first neg = True then
10             t_max ← Derive time step t of max. accumulated benefit in current interval
11             cost ← Assess switching cost from current to alternative setup and back for period
                [0, t_max]
12             first neg ← False
13         if acc_benefit[cum] drops by at least cost from previous maximum then
               /* If GT inactive, switch if max. generated benefit exceeds
                  associated switching cost                                  */
14             if gt_benefit = 0 and acc_benefit[max] > cost then
15                 switch ← True
               /* If GT active, switch if benefit continuously decreased from
                  previously generated maximum                               */
16             else if gt_benefit > 0 and t_max = 0 then
17                 switch ← True
18                 t_max ← t
19             return switch, t_max, cost
20     else
21         first neg ← True
22     end
23 end
   /* In case end of time series gets reached, switch modes if accumulated
      benefit exceeds initial switching cost                                 */
24 cost ← Assess switching cost from current to alternative at t = 0
25 if gt_benefit = 0 and acc_benefit[cum] > cost then
26     switch ← True
27 return switch, t, cost
```

References

[1] R. Adhikari and R. K. Agrawal. An Introductory Study on Time Series Modeling and Forecasting, 2013. Available at https://arxiv.org/abs/1302.6613 (accessed Feb 2021).

[2] P. Bacher, H. Madsen, H. A. Nielsen, and B. Perers. Short-term heat load forecasting for single family houses. *Energy and Buildings*, 65:101–112, 2013. doi:10.1016/j.enbuild.2013.04.022.

[3] R. Bavière and M. Vallée. Optimal Temperature Control of Large Scale District Heating Networks. *Energy Procedia*, 149:69–78, 2018. doi:10.1016/j.egypro.2018.08.170.

[4] E. Bee Dagum and S. Bianconcini. *Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation*. Statistics for Social and Behavioral Sciences. Springer International Publishing, Cham, 2016. doi:10.1007/978-3-319-31822-6.

[5] T. Berners-Lee. Linked data - design issues, 2006. Available at http://www.w3.org/DesignIssues/LinkedData.html (accessed Dec 2020).

[6] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):28–37, 2001.

[7] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies*, 11(7), 2018. doi:10.3390/en11071636.

[8] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time series analysis : forecasting and control*. Wiley series in probability and statistics. Wiley, Hoboken, N.J., 4th edition, 2008. doi:10.1002/9781118619193.ch5.

[9] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, and J. He. Short-Term Load Forecasting With Deep Residual Networks. *IEEE Transactions on Smart Grid*, 10(4):3943–3952, 2019. doi:10.1109/TSG.2018.2844307.

[10] Y. Chen and D. Zhang. Theory-guided deep-learning for electrical load forecasting (TgDLF) via ensemble long short-term memory. *Advances in Applied Energy*, 1: 100004, 2021. doi:10.1016/j.adapen.2020.100004.

[11] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1): 3–73, 1990.

[12] J. G. De Gooijer and R. J. Hyndman. 25 years of time series forecasting. *International Journal of Forecasting*, 22(3):443–473, 2006. doi:10.1016/j.ijforecast.2006.01.001.

[13] A. Eibeck, M. Q. Lim, and M. Kraft. J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. *Computers & Chemical Engineering*, 131:106586, 2019. doi:10.1016/j.compchemeng.2019.106586.

[14] T. Fang and R. Lahdelma. Genetic optimization of multi-plant heat production in district heating networks. *Applied Energy*, 159:610–619, 2015. doi:10.1016/j.apenergy.2015.09.027.

[15] J. Faraway and C. Chatfield. Time series forecasting with neural networks: a comparative study using the air line data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(2):231–250, 2008. doi:10.1111/1467-9876.00109.

[16] L. Giraud, M. Merabet, R. Baviere, and M. Vallée. Optimal Control of District Heating Systems using Dynamic Simulation and Mixed Integer Linear Programming. In *Proceedings of the 12th International Modelica Conference*, pages 141–150, 2017. doi:10.3384/ecp17132141.

[17] E. G. Goh and K. Noborio. Sensitivity Analysis and Validation for Numerical Simulation of Water Infiltration into Unsaturated Soil. *International Scholarly Research Notices*, 2015(824721):7, 2015. doi:10.1155/2015/824721.

[18] E. Guelpa, C. Toro, A. Sciacovelli, R. Melli, E. Sciubba, and V. Verda. Optimal operation of large district heating networks through fast fluid-dynamic simulation. *Energy*, 102:586–595, 2016. doi:10.1016/j.energy.2016.02.058.

[19] A. Hast, S. Syri, V. Lekavičius, and A. Galinis. District heating in cities as a part of low-carbon energy system. *Energy*, 152:627–639, 2018. doi:10.1016/j.energy.2018.03.156.

[20] J.-L. Hippolyte, Y. Rezgui, H. Li, B. Jayan, and S. Howell. Ontology-driven development of web services to support district energy applications. *Automation in Construction*, 86:210–225, 2018. doi:10.1016/j.autcon.2017.10.004.

[21] S. Hochreiter and J. Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, 1997. doi:10.1162/neco.1997.9.8.1735.

[22] R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, Melbourne, Australia, 2nd edition, 2018. Available at `https://www.otexts.org/fpp` (accessed Mar 2021).

[23] R. J. Hyndman and Y. Khandakar. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, 27(3), 2008. doi:10.18637/jss.v027.i03.

[24] R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006. doi:10.1016/j.ijforecast.2006.03.001.

[25] S. Idowu, S. Saguna, C. Åhlund, and O. Schelén. Forecasting heat load for smart district heating systems: A machine learning approach. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 554–559, 2014. doi:10.1109/SmartGridComm.2014.7007705.

[26] R. L. King. Information services for smart grids. In *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–5, 2008. doi:10.1109/PES.2008.4596956.

[27] A. Lake, B. Rezaie, and S. Beyerlein. Review of district heating and cooling systems for a sustainable future. *Renewable and Sustainable Energy Reviews*, 67:417–425, 2017. doi:10.1016/j.rser.2016.09.061.

[28] Y. Li, Y. Rezgui, and S. Kubicki. An intelligent semantic system for real-time demand response management of a thermal grid. *Sustainable Cities and Society*, 52: 101857, 2020. doi:10.1016/j.scs.2019.101857.

[29] H. Lund, S. Werner, R. Wiltshire, S. Svendsen, J. E. Thorsen, F. Hvelplund, and B. V. Mathiesen. 4th Generation District Heating (4GDH): Integrating smart thermal grids into future sustainable energy systems. *Energy*, 68:1–11, 2014. doi:10.1016/j.energy.2014.02.089.

[30] S. G. Makridakis, S. C. Wheelwright, and R. J. Hyndman. *Forecasting : methods and applications*. John Wiley and Sons, New York, 3rd edition, 1998. ISBN 0471532339.

[31] A. Maleki, S. Nasseri, M. S. Aminabad, and M. Hadi. Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics. *KSCE Journal of Civil Engineering*, 22(9):3233–3245, 2018. doi:10.1007/s12205-018-1195-z.

[32] J. Mohring, D. Linn, M. Eimer, M. Rein, and N. Siedow. District heating networks – dynamic simulation and optimal operation. In S. Göttlich, M. Herty, and A. Milde, editors, *Mathematical Modeling, Simulation and Optimization for Power Engineering and Management*, volume 34. Springer, Cham., 2021. doi:10.1007/978-3-030-62732-4.

[33] J. Morio. Global and local sensitivity analysis methods for a physical system. *European Journal of Physics*, 32(6):1577–1583, 2011. doi:10.1088/0143-0807/32/6/011.

[34] S. Moser, S. Puschnigg, and V. Rodin. Designing the Heat Merit Order to determine the value of industrial waste heat for district heating systems. *Energy*, 200(117579): 117579, 2020. doi:10.1016/j.energy.2020.117579.

[35] S. Muzaffar and A. Afshari. Short-Term Load Forecasts Using LSTM Networks. *Energy Procedia*, 158:2922–2927, 2019. doi:10.1016/j.egypro.2019.01.952.

[36] M. Nikolaou. Model predictive controllers: A critical synthesis of theory and industrial needs. *Advances in Chemical Engineering*, 26:131–204, 2001. doi:10.1016/S0065-2377(01)26003-7.

[37] F. Petropoulos, S. Makridakis, V. Assimakopoulos, and K. Nikolopoulos. 'Horses for Courses' in demand forecasting. *European Journal of Operational Research*, 237(1):152–163, 2014. doi:10.1016/j.ejor.2014.02.036.

[38] R. Nau, Duke University. ARIMA models for time series forecasting, 2020. Available at https://people.duke.edu/~rnau/seasarim.htm (accessed Mar 2021).

[39] M. Rein, J. Mohring, T. Damm, and A. Klar. Optimal control of district heating networks using a reduced order model. *Optimal Control Applications and Methods*, 41(4):1352–1370, 2020. doi:10.1002/oca.2610.

[40] J. Reynolds, M. W. Ahmad, Y. Rezgui, and J.-L. Hippolyte. Operational supply and demand optimisation of a multi-vector district energy system using artificial neural networks and a genetic algorithm. *Applied Energy*, 235:699–713, 2019. doi:10.1016/j.apenergy.2018.11.001.

[41] L. Ruiz, M. Cuéllar, M. Calvo-Flores, and M. d. C. Pegalajar Jiménez. An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. *Energies*, 9:684, 2016. doi:10.3390/en9090684.

[42] H. Runvik, P.-O. Larsson, S. Velut, J. Funkquist, M. Bohlin, A. Nilsson, and S. M. Razavi. Production Planning for Distributed District Heating Networks with JModelica.org. In *Proceedings of the 11th International Modelica Conference*, pages 217–223, 2015. doi:10.3384/ecp15118217.

[43] S. Twidale. Analysts raise EU carbon price forecasts as tougher climate targets loom. Reuters, 2021. Available at https://www.reuters.com/business/sustainable-business/analysts-raise-eu-carbon-price-forecasts-tougher-climate-targets-loom-2021-04-20/ (accessed June 2021).

[44] G. Sandou, S. Font, S. Tebbani, A. Hiret, C. Mondon, S. Tebbani, A. Hiret, and C. Mondon. Predictive Control of a Complex District Heating Network. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 7372–7377, 2005. doi:10.1109/CDC.2005.1583351.

[45] D. E. Seborg. *Process dynamics and control*. Wiley, Hoboken, N.J., 3rd edition, 2011. ISBN 978-0-470-64610-6.

[46] I. M. Sobol. Distribution of points in a cube and approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7:86–112, 1967.

[47] T. Baker and N. Noy and R. Swick and I. Herman. W3C's Semantic Web Case Studies and Use Cases, 2021. Available at https://www.w3.org/2001/sw/sweo/public/UseCases/ (accessed May 2021).

[48] S. Werner. International review of district heating and cooling. *Energy*, 137:617–631, 2017. doi:10.1016/j.energy.2017.04.045.

[49] Y. Zhang, T. Zhang, R. Wang, Y. Liu, and B. Guo. Optimal operation of a smart residential microgrid based on model predictive control by considering uncertainties and storage impacts. *Solar Energy*, 122:1052–1065, 2015. doi:10.1016/j.solener.2015.10.027.