RESEARCH ARTICLE

Optimization and validation of pumping system design and operation for water supply in high-rise buildings

Tim M. Müller[1] · Philipp Leise[1] · Imke-Sophie Lorenz[1] · Lena C. Altherr[2] · Peter F. Pelz[1]

Abstract

The application of mathematical optimization methods for water supply system design and operation provides the capacity to increase the energy efficiency and to lower the investment costs considerably. We present a system approach for the optimal design and operation of pumping systems in real-world high-rise buildings that is based on the usage of mixed-integer nonlinear and mixed-integer linear modeling approaches. In addition, we consider different booster station topologies, i.e. parallel and series-parallel central booster stations as well as decentral booster stations. To confirm the validity of the underlying optimization models with real-world system behavior, we additionally present validation results based on experiments conducted on a modularly constructed pumping test rig. Within the models we consider layout and control decisions for different load scenarios, leading to a Deterministic Equivalent of a two-stage stochastic optimization program. We use a piecewise linearization as well as a piecewise relaxation of the pumps' characteristics to derive mixed-integer linear models. Besides the solution with off-the-shelf solvers, we present a problem specific exact solving algorithm to improve the computation time. Focusing on the efficient exploration of the solution space, we divide the problem into smaller subproblems, which partly can be cut off in the solution process. Furthermore, we discuss the performance and applicability of the solution approaches for real buildings and analyze the technical aspects of the solutions from an engineer's point of view, keeping in mind the economically important trade-off between investment and operation costs.

✉ Tim M. Müller
  tim.mueller@fst.tu-darmstadt.de

  Peter F. Pelz
  peter.pelz@fst.tu-darmstadt.de

[1] Chair of Fluid Systems, Department of Mechanical Engineering, Technische Universität Darmstadt, Darmstadt, Germany

[2] Faculty of Energy, Building Services and Environmental Engineering, Münster University of Applied Sciences, Stegerwaldstr. 39, Steinfurt, Germany

Ⓢ Springer

1 Introduction

More than half of the world's population (55 %) lived in urban areas in 2018, as shown in UN (2018). Based on this data published by the United Nations, it is estimated that 2.5 billion people will be added to the urban population by 2050 leading to an increase by more than half of the number of people living in urban areas today. This continuous trend towards urbanization will lead to an increasing number of mega-cities, cities with more than 10 million inhabitants.

As the space for living in cities is limited, one effective way to cope with the steady urbanization and to create living space are high-rise buildings. In those buildings distributed pressure boosting pumping stations, also called booster stations, provide water for every floor, especially in higher pressure zones. Consequently, their need will increase proportionally with the number of high-rise buildings. Already today there is a rapid growth of completed buildings of 200 meters or greater height, CTBUH (2018).

Water supply systems are a key component of any urban infrastructure, cf. Coelho and Andrade-Campos (2014), and urbanization will pose increasing challenges for these infrastructures in the future: Whilst a secure supply of water for the growing number of citizens is to be ensured, energy consumption for water supply must be reduced on the way to a more sustainable future. Thus, efficiency optimization of water supply systems has been in the focus of ongoing research, for a review cf. e.g. Coelho and Andrade-Campos (2014).

While existing research has so far tended to focus on the energy consumption of larger centralized water infrastructures, in this contribution, we want to investigate the energy consumption of the mentioned booster stations for buildings. Especially in tall buildings, booster stations have a considerable energy consumption, in particular if they are wrongly dimensioned and/or operated. Therefore, the potential for energy savings through the application of several pressure zones or a decentralized pump placement and an optimal selection of pump types is high, cf. e.g. Altherr et al. (2019).

In order to unlock this optimization potential, a large number of pipe network topologies and pump configurations must be considered in the planning process. However, with the increasing number of degrees of freedom, especially in tall buildings, it is a difficult challenge to investigate all possible system variants, as the number of possible design choices can be overwhelming. Moreover, one not only has to check whether a promising design is able to fulfill the different load scenarios occurring in the building, but also for an economic solution by balancing investment and operating costs against each other. To achieve these goals, we present an optimization-aided system design approach based on a problem-specific solving algorithm. In order to ensure that the underlying optimization models represent the system under consideration with sufficient accuracy, we also present an experimental validation. To model the pressure losses due to friction we use own measurements as well as literature data and compare the optimization results of both approaches.

In the following Sect. 2, we introduce relevant literature for the optimization. Afterwards, we present the technical application in detail and show different formulations of the optimization model as well as a problem specific solving algorithm. We then discuss the performance of different solving approaches as well as technical aspects of the solutions and close with an experimental validation. Finally we draw a conclusion and give an outlook for future work.

2 Related work

In this article we present a holistic approach to the optimal synthesis of technical systems, which we illustrate by means of a pumping system for water supply. In the context of this research focus, the state-of-the-art is presented in two different areas:

(i) Methods used in engineering for the design and efficiency optimization of pumping systems, in particular pressure boosting systems in buildings.
(ii) Application of mathematical optimization methods to general technical systems, in particular methods for two-stage optimization problems.

In the next two subsections we will give a short overview of relevant literature in these two areas.

2.1 Design and efficiency optimization of pumping systems

When optimizing the efficiency of a pumping system, a first step is to assemble the system from efficient components. For the transport of drinking water, booster stations consisting of centrifugal pumps are commonly used. Therefore, we focus on this pump type within this literature overview.

The development and hydraulic optimization of centrifugal pumps is already at a very high level and multiple standard works are focusing on centrifugal pump design as well as performance, cf. e.g. Japikse and Marscher (1997), Brennen (2011), or Gülich (2010). In recent decades, also through the use of computer-aided methods such as Computational Fluid Dynamics (CFD), cf. e.g. Garg (1998), the efficiency of centrifugal pumps, in particular at their design point, has been continuously improved.

In real life applications, however, it can be seen that centrifugal pumps are predominantly not operated at their design point, but in the partial load range for a large part of the time, cf. Hirschberg (2014). In this case, the excess energy is either dissipated via throttling valves, or—if variable speed pumps, so called RPM-regulated pumps, are employed—they are operated at a lower rotational speed than their nominal rotational speed. Both leads to an operation at lower efficiency, often far away from their best efficiency point. Therefore, when classifying and optimizing pump performance, a suitable approach for assessing efficiency in practical usage is crucial.

The product approach, which is part of an european standard for energy-efficiency evaluation of water pumps, c.f. EN 16480:2016 (2016), is a first step in this direction. This approach no longer considers merely the pump as a component, but the pump and electric motor as part of a system. The efficiency of this system is assessed according to the Minimum Energy Efficiency Index (MEI) which rates the pump's efficiency not

only at its design point, but also at partial load and overload. Therefore, the so-called "house of efficiency" scheme is applied, which sets efficiency criteria for 75%, 100% and 110% of the volume flow at the design point.

While this is a first step towards the right direction, when pumps are used in a real building, not only their volume flow varies, but so does the required pressure increase in relation to the demanded flow. In addition, when assessing overall energy efficiency, it is important to take the time slice of the respective volume flow and its corresponding pressure requirements into account. For this reason, the Extended Product Approach, cf. Stoffel (2015), was developed within the scope of energy efficiency guidelines of the European Union. It considers a load profile which links volume flow and pressure demands, as well as corresponding time profiles. The energy consumption of the Extended Product (pump and electric motor) considering these load profiles then yields the Energy Efficiency Index (EEI), cf. Lang et al. (2013). In this work, we present a method which considers typical load profiles within buildings and is in accordance with the EEI.

When building pumping systems, several technical aspects have to be considered. The standards DIN 1988-500 (2011) and DIN 1988-300 (2012) provide codes of practice for drinking water installations. While DIN 1988-500 (2011) focuses on the dimensioning of pressure boosting systems with RPM-regulated pumps, DIN 1988-300 (2012) describes the dimensioning of the pipe network. Particularly, the selection of suitable pipe diameters in regard to restrictions such as a peak load volume flow and a maximal water velocity for reasons of noise is outlined. In addition to a central pump configuration, the standard DIN 1988-500 (2011) also considers some decentralized pump configurations due to the specification of different pressure zones within the building. E.g., lower pressure zones of the building can be supplied directly by the public water mains and solely higher pressure zones are supplied via a pressure boosting system. Furthermore, the installation of several booster stations is described so that each pressure zone is supplied by its own pumping system. However, more sophisticated decentralized configurations, which could be automatically assessed and evaluated by our approach, are not considered.

Whereas efficient components are certainly a basic prerequisite for an efficient overall system, they are not yet a guarantee for this. With a system-wide assessment in mind, the scope of the international standard "Pump system energy assessment" is to set "the requirements for conducting and reporting the results of a pumping system energy assessment [...] that considers the entire pumping system, from energy inputs to the work performed as the result of these inputs", ISO/ASME 14414 (2019). However, the standard neither specifies how to design a pumping system nor how to validate the energy savings that would result from implementing recommendations based on this assessment. That are two aspects that we cover in this work.

In general, there are several system-wide measures to increase the energy efficiency of pumping systems:

– The optimal selection of pumps, i.e. usage of variable-speed pumps, cf. Ferreira et al. (2010), and avoiding over-dimensioning, cf. e.g. Fisher (2001), Tindall and Pendle (2015), Weber and Lorenz (2017), Weber and Lorenz (2018).

– The allocation of several pressure zones within the building and the usage of decentralized pump layouts, cf. e.g. Norgaard and Nielsen (2010); Leise and Altherr (2018).
– The optimization of the pumps' operation, cf. e.g. Pedersen and Yang (2008), Groß et al. (2017), and Nowak et al. (2018).

The approach presented in this contribution allows to cover all three of these aspects, since not only the optimal pumping system design, but also its optimal operation can be considered.

While in this work we particularly focus on the optimization of the efficiency and investment costs, other possible objectives are shown in the literature. One example is the optimization of resilient water supply networks as shown by Herrera et al. (2016), Hartisch et al. (2018), and Altherr et al. (2019).

2.2 Mathematical aspects of two-stage optimization problems in engineering

Besides literature from the engineering domain covering the technical aspects of optimizing pumping systems, also previous works regarding the mathematical aspects are relevant.

When designing an economically optimal pumping system, the overall goal often is to minimize the sum of investment costs and energy costs for the expected life cycle. Since the load within residential buildings varies, cf. Sect. 2.1, its operation at different load points has to be considered to assess the energy costs of the pumping system.

The corresponding mathematical optimization problem consists of two stages: First, finding a low-priced investment decision, i.e. a pump and pipe configuration. Second, operating a subset of the chosen pumps for each quasi-stationary demand scenario of the given load profile, such that the system satisfies this demand and at the same time the efficiency is maximized.

In this case, the considered load profile can be seen as a discrete distribution of uncertain load parameters, with the time portion of each load scenario corresponding to its probability of occurrence. The optimization program, which minimizes investment costs as well as expected value of operation costs, can then be seen as the Deterministic Equivalent of a general two-stage stochastic program with recourse, cf. e.g. Wets (1974).

As many optimization problems in the area of engineering design are subject to uncertain parameters, e.g. uncertain material characteristics or load scenarios, two-stage stochastic programs have been frequently studied in this context, and a good overview can be found in Popela (2010) and Popela et al. (2014). The life cycle of products and systems in engineering includes the two phases of planning (first stage) and operation (second stage). In two-stage stochastic optimization problems the operation is already anticipated in the planning phase, which leads to better investment decisions. The operation itself is then typically ensured by closed- or open-loop controllers.

An important aspect for the difficulty of solving these two-stage optimization problems to global optimality is whether integer variables are (1) present in the first stage solely, or (2) also present in the second stage. Regarding design and operation opti-

mization of technical systems, integer variables in the first stage, and more precisely binary variables, are often used to model purchase decisions. Moreover, in this work, binary variables in the second stage are used to model the activation or deactivation of active components, i.e. an on/off-switch, leading to problems of case (2).

Theoretical aspects of these problems, also called complete mixed-integer recourse problems, have e.g. been investigated by Schultz (1992), and sophisticated solution approaches have been proposed, cf. e.g. Carøe and Tind (1998). Given the formulation as Deterministic Equivalent Program, standard solution approaches based on Branch & Bound with LP relaxation may be applied. Yet these might suffer from high computation times, especially if multiple (load) scenarios are considered in the second stage.

3 Technical application

As mentioned in the previous sections, we generate the Deterministic Equivalent Programs of two-stage stochastic optimization models to compute energy-efficient centralized and decentralized booster stations in high-rise buildings and validate these models on a test rig. The physical properties of these systems and topology decisions are modeled within the constraints of the optimization programs. Within this contribution, the fresh cold water supply is considered solely, while waste water, hot water, and water for firefighting are not considered. Thereby, the objective is to minimize the simplified life-time cost, which consists of investment and energy costs within the lifetime. These two parts represent the economically most important goals, cf. Tolvanen (2007), when designing pumping systems. When modeling these systems, two aspects are of importance. First, reasonable computation times of model instances are required. Second, the model's level of detail has to represent the real system with an acceptable accuracy to allow for the transfer of the optimization results. The computed optimum within the model instance, should represent the true optimum of the underlying technical system. Usually, a trade-off between these two goals is necessary, as the former tends to lead to models that are as simple as possible, while the latter demands for models that are highly accurate, but computationally not tractable. To achieve a model that is as simple as possible and at the same time as accurate as needed, we conduct an experimental validation step based on a test rig that represents a scaled building with five pressure zones.

We present all relevant basics of the pump and booster station modeling, the test rig, the system modeling and the preselection of pumps in the following subsections.

3.1 Pump modeling

The most important components to fulfill the water supply in high-rise buildings are the pumps. All pumps are modeled with polynomial approximations for the pressure–volume flow characteristic, as well as the power–volume flow characteristic. This approach is based on Ulanicki et al. (2008). We model the pressure increase $\Delta p = f(q, n)$ with $f : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$ by using a quadratic approximation relating the
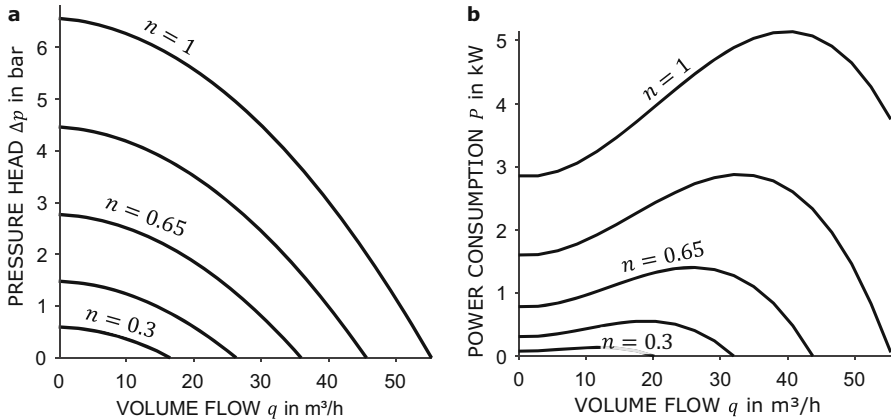
Fig. 1 Characteristics for an example pump based on the used pump model. These characteristics are usually limited in each domain according to specific pump requirements. The different curves describe several different rotational speeds

pressure head Δp, the volume flow $q \in \mathbb{R}_0^+$ and the normalized rotational speed $n \in [\underline{N}, \overline{N}]$, as given by

$$\Delta p = \sum_{m=0}^{2} \alpha_m \, q^{2-m} \, n^m, \tag{1}$$

where the parameters α_m and the lower bound of the normalized rotational speed $\underline{N}$ depend on the pump type. The rotational speed is normalized by using its maximum value. Therefore, we get an upper bound $\overline{N} = 1$. For better readability, we refer to the normalized rotational speed in the following as rotational speed or speed, only. The power $po = g(q, n)$ with $g : R_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ is modeled by using a cubic approximation,

$$po = \sum_{m=0}^{3} \beta_m \, q^{3-m} \, n^m. \tag{2}$$

The parameters α_m and β_m are derived from manufacturer provided component characteristics for each pump type. Exemplified characteristic diagrams for a pump are shown in Fig. 1.

The shown characteristics are often limited by maximum power as well as minimum and maximum volume flow requirements for each pump. These limits are added in the underlying optimization as additional constraints to model a realistic pump behavior. For simplicity they are not shown in Fig. 1.

3.2 Booster station modeling

The function of a booster station is to provide sufficient pressure to overcome the geodetic height of the floor to be provided with water as well as the friction in the pipes feeding the respective floor. We distinguish three different booster station topologies,

Fig. 2 Exemplified system
design of the underlying
optimization program
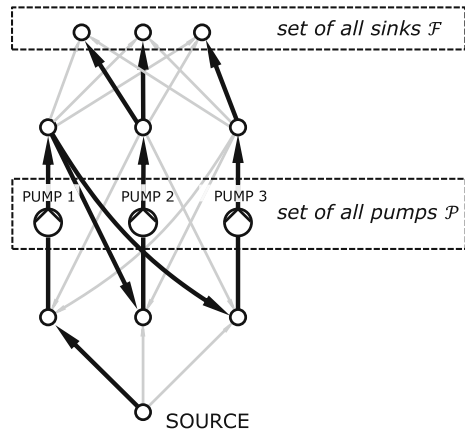
which differ in the alignment of the considered pumps. Today, pumps are commonly
aligned in parallel at the lowest pressure zone of the building, as the product range
of various manufacturers (Wilo 2020; KSB SE Co KGaA 2020; Grundfos Pumps
Ltd 2016) and the discussion in (Gülich 2010, pp. 666–669) shows. This leads to a
central, parallel booster station design providing the pressure and volume flow for the
whole building. Varying from this approach, also the serial arrangement of pumps in a
central booster station is possible, leading to higher degrees of freedom in the design
process and therefore allowing for higher energy efficiency of the booster station. For
the centralized topology the building's demand can be modeled as solely one sink
with one available water source. Additionally to these centralized system topologies,
a decentralized booster station topology is possible. In this approach, pumps can be
positioned at any pressure zone of the building and therefore a more accurate modeling
of the volume flow demand, i.e. the flow demand per pressure zone, is necessary. In the
following, we comprise floors of similar height into individual pressure zones. Each
pressure zone is modeled as an individual sink which has a different pressure demand,
depending on the height of the floors. The decentralized topology allows for an even
higher degree of freedom of pump arrangement and a reduction of the throttle losses,
facilitating the improvement of the overall efficiency even further.

The system design approach with an exemplified booster station is shown in Fig. 2.
In this exemplary system three pumps are used. The first pump is used in series with
the remaining two pumps. The second and third pump supply different pressure zones
in the building individually.

3.3 Test rig for validation

As indicated before, an experimental validation of the optimization results is crucial to
guarantee the technical functionality. A validation using real buildings is not possible
for reasons of high effort. In addition, novel concepts, i.e. the decentralized booster
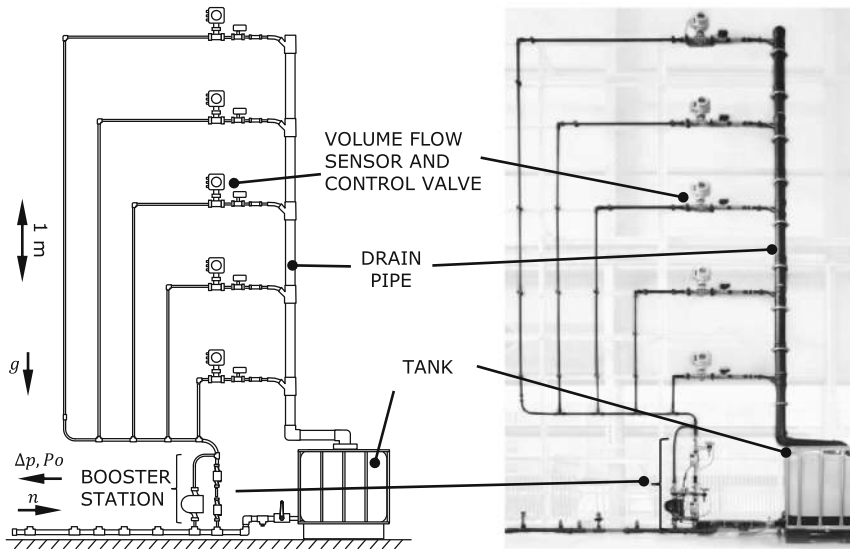station design, shall be investigated, which cannot be realized in already existing
buildings.

Fig. 3 Sketch (left) and photo (right) of the modular test rig with a central booster station design

For this reason, a down-scaled, modular test rig has been set up, cf. Figure 3 and Müller et al. (2019b). The down-scaled system's booster station has the same characteristics as those of high-rise buildings. The maximum geodetic height difference is 5 m divided into five different pressure zones, which can be supplied individually. The water is pumped from a tank with constant pressure to the five pressure zones by the booster station. As in real buildings, the water flows in at a constant supply pressure and flows out in an open drain pipe at ambient pressure (open system). The water returns to the tank to keep the water level and therefore the supply pressure constant. The test rig is equipped with sensors for measuring the volume flow in each pressure zone as well as the pressure head and power consumption of each pump. It is possible to control the speed of the pumps and the valve position on each pressure zone and thus to control the volume flow on each pressure zone. Overall, a set of thirteen pumps, consisting of six different types, is available to build and test different booster station designs.

The pumps are centrifugal circulator pumps for heating systems with an integrated frequency converter (FC) and therefore speed controllable. The sizes are chosen to fit the needs for the down-scaled booster system, even though the ratio of pressure head to volume flow is lower than in real-world booster station pumps. The pumps can be connected in any way and up to six pumps can be operated simultaneously as a booster station for the test rig. Compared to real pressure booster stations, the friction caused by installation fittings of the pumps (e.g. for pressure measurement) on the test rig is quite high. Therefore, those pressure losses are taken into account in the loss coefficient ζ_{inst}.

Table 1 Characteristics of the chosen real-world hotel buildings and the down-scaled test rig

	unit	building 1	building 2	building 3	test rig
geodetic height H	m	100	80	40	5
max. volume flow demand $\overline{Q}$	m^{3}h^{-1}	72	82	18	4.28
floors	–	28	23	10	–
pressure zones	–	5	5	5	5

3.4 System modeling and load scenarios

We consider three different real-world high-rise hotel buildings in addition to the previously described test rig as examples for the presented system optimization approach. We consider each building's geodetic height, base area and type to estimate the volume flow demand and the respective pressure increase. The estimation of volume flow demands is highly uncertain if no consumption patterns are known previously to the optimization. An overview of water demand for hotel buildings is given by Gössling et al. (2012). Besides these different measurements, Hirschberg (2014) also presented an approximation method to estimate the volume flow demand in buildings of different types. Another possibility to approximate the volume flow in buildings is given by DIN 1988-300 (2012). It shows a computational approach to estimate the volume flow demand in buildings based on a standardized estimation procedure.

We estimated the volume flow making use of the estimation method shown by Hirschberg (2014) for building 1 and 2. The volume flow demand in building 3, as shown in Table 1, is based on real consumption data in this building. In the following, the different buildings are referred to as B1, B2 and B3, respectively. We assume that all buildings in Table 1 are divided into 5 pressure zones similarly to the test rig.

The demand of the test rig system is derived by scaling the demands of real buildings according to the geodetic height and the friction losses of the test rig.

The maximum load scenario—given by the maximum volume flow demand $\overline{Q}$ and maximum pressure head p_s—allows to lay out the booster station for the peak load. This load scenario occurs only for a fraction of the operation time. To design the booster station more realistically, we also consider the events of partial load. The volume flow demand in partial load is derived according to Hirschberg (2014) and all scenarios are shown in Fig. 4.

The necessary pressure at the sink of a central booster station is calculated using the geodetic height of the building H, the minimum flow pressure $p_{\text{flow}} = 1$ bar, the supply pressure $p_{\text{sup}} = 3$ bar (i.e. provided by the connection to the regional water supply system) and the friction, which depends quadratically on the flow rate. According to Hirschberg (2014), the pressure loss due to friction in the maximum load case is approximately $\sigma = 52\%$ of both the geodetic height and minimum flow pressure. This leads to:

Fig. 4 Load scenarios based on Hirschberg (2014) and their according time shares
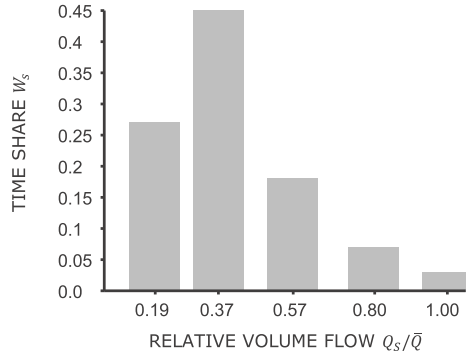
$$p_s = \varrho g H + p_{\text{flow}} + \frac{1}{2}\varrho\zeta\left(\frac{Q_s}{A}\right)^2 - p_{\text{sup}}$$

$$= (\varrho g H + p_{\text{flow}})\left(1 + \sigma\left(\frac{Q_s}{\bar{Q}}\right)^2\right) - p_{\text{sup}} \quad \forall s \in \mathcal{S}. \tag{3}$$

For the decentralized system design, we derive individual load scenarios for each pressure zone, which sum up to the load scenarios in Fig. 4. The pressure at the different sinks is derived by scaling the pressure for the central booster station according to the geodetic height and the volume flow of the individual pressure zones.

3.5 Preselection of pumps

In this work, we apply a model series of 200 different market-available pumps which were explicitly designed for the water supply of buildings. The combinatorial possibilities resulting from this wide range of variants is challenging within reasonable computation time. In addition, most pumps are not suitable for each specific building. Therefore, we aim to reduce the number of considered pumps based on domain-specific knowledge. To the best of the authors knowledge there exists no general purpose method to select multiple appropriate pumps for consideration within an optimization program. Sets of pumps for the respective buildings are currently only selected based on expert knowledge as for example shown by Larralde and Ocampo (2010a) and Larralde (2010). Therefore, a crucial step in the optimal design of booster stations for high-rise buildings is the selection of an appropriate set of available pumps, as shown by Leise and Altherr (2018). For this, we implemented a heuristic, as shown in Algorithm 1, for the pump selection, which is based on a two-step process.

In the first step of the heuristic selection of pumps we define additional partial load scenarios that should be met by individual pumps selected in the following step. These additional scenarios are derived from the original scenarios (represented by black square markers in Fig. 5) by taking advantage of the generally valid relationship of pressure and volume flow for in series and parallel connected pumps:

Algorithm 1: Heuristic Selection of Pumps

---

input : original load scenarios in Δp–q domain $\mathcal{S}$
        set of all available pumps $\mathcal{P}^{\text{total}}$
        distance parameter $d^{\max}$
        set of desired fractional positions $\mathcal{F}$
output: selected pumps $\mathcal{P}$
$\mathcal{Z} \leftarrow \emptyset$ // selected fractional loads in Δp-q domain
$\mathcal{P} \leftarrow \emptyset$ // selected pumps
// step 1: select fractional loads
for $i \leftarrow 1$ **to** $|\mathcal{S}|$ **do**
   | generate fractional loads for parallel pump usage based on $\mathcal{F}$ and add them to $\mathcal{Z}$
for $i \leftarrow 1$ **to** $|\mathcal{Z}|$ **do**
   | generate fractional loads for serial pump usage based on $\mathcal{F}$ and add them to $\mathcal{Z}$
sort $\mathcal{Z}$ by power
while *true* **do**
   compute distance matrix $d_{i,j}$ between every point z in $\mathcal{Z}$
   if *at least one $d_{i,j} < d^{max}$ $\forall z_i, z_j \in \mathcal{Z}$ without $i = j$* **then**
     | remove point with $d_{i,j} < d^{\max}$ and lowest possible power from $\mathcal{Z}$
   else
     | quit loop
// step 2: get best suitable pumps for $\mathcal{Z}$
for $i \leftarrow 1$ **to** $|\mathcal{Z}|$ **do**
   for $j \leftarrow 1$ **to** $|\mathcal{P}^{total}|$ **do**
     select best suitable pump $p_j \in \mathcal{P}^{\text{total}}$ for $z_i \in \mathcal{Z}$
   save best suitable pump in $\mathcal{P}$ if not already selected

---

On the one hand, for in series connected pump systems with N pumps the volume flow in each pump is equivalent to the total system volume flow, while the pressure in each pump is given by $\Delta p/N$ (if we assume that the pressure increase is equal in each pump).

On the other hand, a parallel usage of N pumps of the same type results in total in the same pressure increase for each pump but the volume flow for each pump reduces to q/N if all parallel pumps are operated with the same speed.

These two effects for in series and parallel system design result in a reduction of the original load scenarios to fractions of the original load scenarios in q and/or Δp direction.

In the selection heuristic we start by generating all fractional loads that can be derived for parallel combinations with up to 6 equivalent pumps based on the original load scenarios. Afterwards we then derive all loads that are based on in series connections of up to 6 equivalent pumps for the originally given loads and newly computed ones.

For each building we want to consider a predefined number of pumps (in our case 15 in total). As the number of fractional loads in the first step exceeds this predefined number and also results in a clustering of fractional load scenarios with small power, we use a set distance $d^{\max}$ in the Δp–q domain to select only loads which have a predefined distance from each other. We then iteratively delete load scenarios with a lower hydraulic power demand within the range of the set distance. With this reduction step, we assure that every deleted load scenario point remains feasible in the solution, as we could use a pump with a higher hydraulic power and a frequency converter to
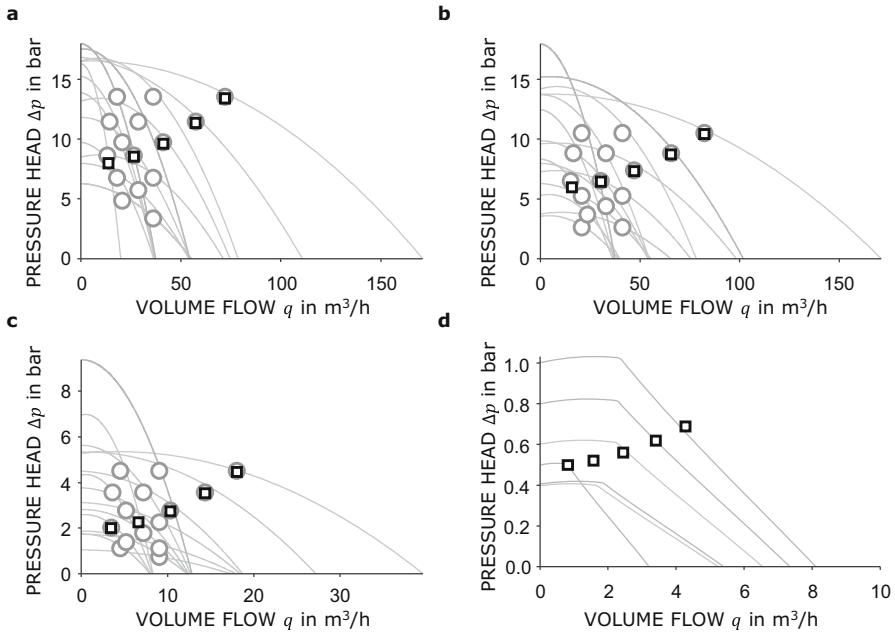
Fig. 5 Characteristic diagrams of all used pumps at their maximum rotational speed for **a** building B1 of 100 m height, **b** building B2 of 80 m height, **c** building B3 of 40 m height, and **d** the test rig. The considered scenarios are shown as black squares. The scaled points to select pumps in Algorithm 1 are shown as gray circles for the three buildings

reduce the speed according to the desired fractional demand. Additionally, it would be possible in a real building to throttle the flow to achieve the desired partial load scenario. This simple procedure ensures to derive the predefined number of pumps as well as a great coverage of the overall domain.

In the second step of the heuristic selection of pumps we select an appropriate pump for each of the remaining load scenarios in the Δp–q domain which are given by gray circles and black squares in Fig. 5. A pump is selected if it can fulfill the desired demand at one fractional load and has the lowest electric power of all considered pumps at full speed. In this manner we derive a final pump catalog that consists of a predefined number of pumps (in our case 15), which are suitable to supply each original load scenario by a single pump or a combination of in series/parallel connected pumps. Additionally, each selected pump has the lowest power demand to supply the partial load it was selected for. By implementing and applying this heuristic, we remove all pumps that are not suitable and all suitable pumps with a higher power demand for the supply of a respective building. In the following, we consider solely a set number of preselected pumps which specifically supply a certain building efficiently.

4 Optimization model

This section deals with the mathematical formulation of the previously introduced technical design problem. The technical specifications are modeled in a basic stochastic optimization model (Sect. 4.1).

We reduce the two-stage stochastic problem to its Deterministic Equivalent Program by using multiple load scenarios. In our case, the first stage variables represent investment and design decisions while second stage variables represent control decisions. Within this scenario based optimization approach binary variables are present in both stages. Even in the second stage as the model takes into account that not all pumps are always switched on to meet the demand of the different load scenarios. In contrast, some pumps may also be switched off and bypassed for efficiency reasons, especially in load cases with lower volume flow requirements.

Since pumps have nonlinear characteristics, the mixed-integer nonlinear program (MINLP) pump model (Sect. 4.2) as well as two linear pump models, allowing the formulation of a mixed-integer linear program (MILP), are presented. In a first approach, a piecewise linearized pump model (Sect. 4.3) is derived and in a second approach, a piecewise linear relaxed pump model (Sect. 4.4) is established.

4.1 Basic optimization model

All derived optimization models in this contribution have a common mixed-integer linear foundation, which we present in this section. The complete system design with all possibilities is modeled as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, in which the edges $\mathcal{E}$ model the pumps as well as the pipes. The junctions of the booster station are modeled as vertices $\mathcal{V}$. The basic model's variables, which are indicated by lower case letters, as well as sets and parameters, indicated by upper case and Greek letters, are shown in Table 2.

The shared objective of all presented optimization models is the maximization of the net present value, cf. Meck et al. (2019). This results in a minimization of lifetime cost, which are approximated considering investment costs for pumps and frequency converters as well as discounted energy costs. The objective is shown in Obj. (5a). We use the annuity-present value factor *APF* to discount all costs to the present day.

We model an upper bound of pumps, Constr. (5b), to solely select a suitable subset of all possible pumps in $\mathcal{P}$ in the final solution, i.e. maximum N^{buy} pumps may be bought. The necessary precondition to use a pump is that it was bought, Constr. (5c). Purchased pumps should be used only in their physical domain. Therefore, we limit the pump speed, power and volume flow with Constr. (5d) and (5e). Constr. (5f) models the volume flow conservation in the network, while Constr. (5g) models upper bounds for the volume flow in each pipe. We demand the volume flow in the sinks to be exactly the predefined volume flow for each scenario, as shown in Fig. 4 with Constr. (5h). It is imperative to choose at least one pump to connect the source with the sinks, Constr. (5i) and (5j). Constr. (5k)–(5m) prevent undesired topologies, as for example loops. The pressure at the inlet of a pump is set to the pressure at the source P_s^{source} in Constr. (5n) if a pump is connected to the source. The pressure increase from the inlet to the outlet of a pump is modeled by Constr. (5o). The pressure at the inlet and

Table 2 Decision variables, sets and parameters

Variable	Domain	Description
$\Delta p_{i,s}$; $q_{i,s}$; $n_{i,s}$	$\mathbb{R}_0^+$	Pressure increase, volume flow and rotational speed of pump $i \in \mathcal{P}$ in scenario $s \in S$
$po_{i,s}^{\text{pump}}$; $po_{i,s}^{\text{FC}}$; $po_{i,s}$	$\mathbb{R}_0^+$	Power consumption of pump $i \in \mathcal{P}$, the associated frequency converter and the sum of both in scenario $s \in S$
$p_{i,s}^{\text{in}}$; $p_{i,s}^{\text{out}}$	$\mathbb{R}_0^+$	Pressure at inlet and outlet of pump $i \in \mathcal{P}$, $s \in \mathcal{S}$
$q_{i,j,s}^{\text{pipe}}$	$\mathbb{R}_0^+$	Volume flow between outlet of pump $i \in \mathcal{P}$ and inlet of pump $j \in \mathcal{P}$, $s \in \mathcal{S}$
$q_{i,s}^{\text{source}}$	$\mathbb{R}_0^+$	Volume flow between source and inlet of pump $i \in \mathcal{P}$, $s \in \mathcal{S}$
$q_{i,f,s}^{\text{sink}}$	$\mathbb{R}_0^+$	Volume flow between outlet of pump $i \in \mathcal{P}$ and sink $f \in \mathcal{F}$, $s \in \mathcal{S}$
		Indicator for connection between:
$t_{i,j}$	$\{0, 1\}$	Outlet of pump i and inlet of pump j,
t_i^{source}	$\{0, 1\}$	Source and inlet of pump i,
$t_{i,f}^{\text{sink}}$	$\{0, 1\}$	Outlet of pump i and sink $f \in \mathcal{F}$
$x_{i,s}$	$\{0, 1\}$	Activation of pump i in scenario $s \in S$
y_i; y_i^{FC}	$\{0, 1\}$	Investment decision for pump $i \in \mathcal{P}$ and associated frequency converter

Set		Description
$\mathcal{P}$		Set of available pumps
$\mathcal{S}$		Set of demand scenarios with $|\mathcal{S}| = 5$
$\mathcal{F}$		Set of sinks, which corresponds to individually supplied pressure zones. For central systems applies $|\mathcal{F}| = 1$, for decentral systems $|\mathcal{F}| = 5$

Parameter	Range	Description
APF	$\mathbb{R}_0^+$	Annuity-present value factor
$C_{\text{invest},i}^{\text{pump}}$	$\mathbb{R}_0^+$	Investment costs of pump $i \in \mathcal{P}$
$C_{\text{invest},i}^{\text{FC}}$	$\mathbb{R}_0^+$	Investment costs of frequency converter for pump $i \in \mathcal{P}$
C_{elect}	$\mathbb{R}_0^+$	Electricity price (0.3 €/kWh)
W_s	$[0, 1]$	Time share of scenario $s \in \mathcal{S}$.
T_{on}	$\mathbb{R}_0^+$	Total runtime per year ($0.3 \times 365d \times 24h$)
N^{buy}	$\mathbb{N}$	Maximum number of pumps that can be purchased.
$\overline{\Delta P_i}$; $\overline{P}$; $\overline{Q}$; $\overline{Po}_i$; $\underline{Po}_i$; $\underline{N}_i$, $\overline{N}_i$	$\mathbb{R}_0^+$	Upper $^-$ or lower $_-$ bounds for the operating parameter of pump i, i.e., pressure increase, pressure, volume flow, power, speed
P_s^{source}	$\mathbb{R}_0^+$	Pressure at the source
$P_{f,s}^{\text{sink}}$; $Q_{f,s}^{\text{bound}}$	$\mathbb{R}_0^+$	Pressure and volume flow demand at sink $f \in \mathcal{F}$ in scenario $s \in \mathcal{S}$
Po_s^{hydr}	$\mathbb{R}_0^+$	Required hydraulic power $s \in \mathcal{S}$

Table 2 continued

Parameter	Range	Description
$\alpha_{i,m}$	$\mathbb{R}_0^+$	Pressure-volume flow regression coefficients for pump $i \in \mathcal{P}$ with $m \in \{0, 1, 2\}$
$\beta_{i,m}$	$\mathbb{R}_0^+$	Power-volume flow regression coefficients for pump $i \in \mathcal{P}$ with $m \in \{0, 1, 2, 3\}$
$\zeta^{\text{inst.}}$	$\mathbb{R}_0^+$	Loss coefficient for the installation fittings of the pumps
$\eta_{\text{best}}^{\text{pump}}$	$\mathbb{R}_0^+$	Best efficiency of any pump in set $\mathcal{P}$
η_i^{FC}	$\mathbb{R}_0^+$	Efficiency of frequency converter of pump $i \in \mathcal{P}$

outlet of an unused pump is set to zero, Constr. (5p). The pressure of a pump's outlet should fulfill the predefined pressure requirements at the sink if it is connected to the sink, Constr. (5q). The pressure propagation between connected pumps is modeled by Constr. (5r). If a frequency converter for a certain pump is not purchased ($y_i^{FC} = 0$), the speed of the pump has to be set to the nominal speed ($n = 1$) in case of activity, Constr. (5s). If the pump is turned off, or a frequency converter is purchased, this constraint has to be deactivated, Constr. (5s). If a frequency converter is bought, the pump is speed controllable, but a part of the transmitted electrical power is dissipated, Constr. (5t). In addition, the electrical power for the pump motor is added to the total power, Constr. (5u). Moreover, we can say that the electrical power of the system has to be higher than the hydraulic power divided by the best efficiency of the best pump, Constr. (5v). This provides a lower bound for the power consumption, which significantly speeds up the optimization process, as shown by Müller et al. (2019a).

In the subsequently introduced models, we allow for the option to restrict the usage of in series connected pumps, as pumps connected in parallel correspond to the current standard installation procedure, cf. Sect. 3.2. A solely parallel connection can be achieved by adding the following constraint to the basic model (5):

$$t_{i,j} = 0 \quad \forall i, j \in \mathcal{P}. \tag{4}$$

The objective as well as all so far presented constraints, Model (5) and Constr. (4), depend linearly on the decision variables. The according mixed-integer linear program describes the generic selection and usage of pumps to supply multiple sinks with relevant connected pumps to derive a suitable booster station. Nevertheless, this model does not contain the explicit modeling of the used pumps, as shown in Eq. (1) and Eq. (2). Therefore, we present different methods to model these nonlinear relations of type $\Delta p = f(q, n)$ and $po = g(q, n)$ in the following subsections.

$$\min \quad T_{\text{on}} APF\, C_{\text{elect}} \sum_{s \in S} W_s \sum_{i \in \mathcal{P}} po_{i,s} + \sum_{i \in \mathcal{P}} C_{\text{invest},i}^{\text{pump}}\, y_i + \sum_{i \in \mathcal{P}} C_{\text{invest},i}^{\text{FC}}\, y_i^{\text{FC}} \tag{5a}$$

subject to

$$\sum_{i \in \mathcal{P}} y_i \leq N^{\text{buy}} \tag{5b}$$

$$x_{i,s} \leq y_i \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5c}$$

$$n_{i,s} \geq \underline{N}_i x_{i,s} \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5d}$$

$$\Delta p_{i,s} \leq \overline{P} x_{i,s}, \ q_{i,s} \leq \overline{Q} x_{i,s}, \ n_{i,s} \leq \overline{N}_i x_{i,s},$$
$$po_{i,s} \leq \overline{Po}_i, \ po_{i,s}^{\text{pump}} \leq \overline{Po}_i x_{i,s} \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5e}$$

$$q_{i,s} = \sum_{j \in \mathcal{P}} q_{j,i,s}^{\text{pipe}} + q_{i,s}^{\text{source}} = \sum_{j \in \mathcal{P}} q_{i,j,s}^{\text{pipe}} + \sum_{f \in \mathcal{F}} q_{i,f,s}^{\text{sink}} \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5f}$$

$$q_{i,j,s}^{\text{pipe}} \leq \overline{Q} t_{i,j}, \ q_{i,s}^{\text{source}} \leq \overline{Q} t_i^{\text{source}}, \ q_{i,f,s}^{\text{sink}} \leq \overline{Q} t_{i,f}^{\text{sink}} \quad \forall i, j \in \mathcal{P}, f \in \mathcal{F}, s \in \mathcal{S} \tag{5g}$$

$$\sum_{i \in \mathcal{P}} q_{i,f,s}^{\text{sink}} = Q_{f,s}^{\text{bound}} \quad \forall f \in \mathcal{F}, s \in \mathcal{S} \tag{5h}$$

$$\sum_{i \in \mathcal{P}} t_{i,f}^{\text{sink}} \geq 1 \quad \forall f \in \mathcal{F} \tag{5i}$$

$$\sum_{i \in \mathcal{P}} t_i^{\text{source}} \geq 1 \tag{5j}$$

$$t_{i,i} = 0 \quad \forall i \in \mathcal{P} \tag{5k}$$

$$\sum_{j \in \mathcal{P}} t_{i,j} \leq |\mathcal{P}|(1 - t_i^{\text{source}}) \quad \forall i \in \mathcal{P} \tag{5l}$$

$$\sum_{j \in \mathcal{P}} t_{i,j} \leq |\mathcal{P}| y_i, \ \sum_{j \in \mathcal{P}} t_{j,i} \leq |\mathcal{P}| y_i, \ t_i^{\text{source}} \leq y_i, \ \sum_{t \in \mathcal{F}} t_{i,f}^{\text{sink}} \leq y_i \quad \forall i \in \mathcal{P} \tag{5m}$$

$$p_{i,s}^{\text{in}} - P_s^{\text{source}} \overset{\leq +}{\geq -} \overline{P}(1 - t_i^{\text{source}}) \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5n}$$

$$p_{i,s}^{\text{in}} + \Delta p_{i,s} - p_{i,s}^{\text{out}} \overset{\leq +}{\geq -} \overline{P}(1 - x_{i,s}) \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5o}$$

$$p_{i,s}^{\text{in}} \leq \overline{P} \left(\sum_{j \in \mathcal{P}} t_{j,i} + t_i^{\text{source}} \right), \ p_{i,s}^{\text{out}} \leq \overline{P} \left(\sum_{j \in \mathcal{P}} t_{i,j} + \sum_{f \in \mathcal{F}} t_{i,f}^{\text{sink}} \right) \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5p}$$

$$p_{i,s}^{\text{out}} - P_{f,s}^{\text{sink}} \overset{\leq +}{\geq -} \overline{P}(1 - t_{i,f}^{\text{sink}}) \quad \forall i \in \mathcal{P}, f \in \mathcal{F}, s \in \mathcal{S} \tag{5q}$$

$$p_{i,s}^{\text{out}} - p_{j,s}^{\text{in}} \overset{\leq +}{\geq -} \overline{P}(1 - t_{i,j}) \quad \forall i, j \in \mathcal{P}, s \in \mathcal{S} \tag{5r}$$

$$n_{i,s} - 1 \overset{\leq +}{\geq -} \left((1 - x_{i,s}) + y_i^{FC} \right) \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{5s}$$

$$\left(1 - \eta_i^{\text{FC}} \right) po_{i,s} - po_{i,s}^{\text{FC}} \leq \overline{Po}_i \left(1 - y_i^{\text{FC}} \right) \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5t}$$

$$po_{i,s} = po_{i,s}^{\text{pump}} + po_{i,s}^{\text{FC}} \quad \forall i \in \mathcal{P}, s \in \mathcal{S} \tag{5u}$$

$$\sum_{i \in \mathcal{P}} po_{i,s} \geq Po_s^{\text{hydr}} / \eta_{\text{best}}^{\text{pump}} \quad \forall s \in \mathcal{S} \tag{5v}$$

4.2 Mixed-integer nonlinear pump model

The most accurate modeling approach is given by a nonlinear approach in which each considered pump is modeled with nonlinear constraints according to Eq. (1) and Eq. (2). We add all pumps by adding

$$\Delta p_{i,s} = (\alpha_{i,0} - \zeta^{\text{inst.}}) \, q_{i,s}^{\ 2} + \sum_{m=1}^{2} \alpha_{i,m} \, q_{i,s}^{\ 2-m} \, n_{i,s}^{\ m} \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{6a}$$

Table 3 Additional variables, sets and parameters for PWL

Variable	Domain	Description
$\lambda_{u,v,i,s}$	[0, 1]	Additional variable to model the piecewise linear approximation based on grid point (u, v) and pump $i \in \mathcal{P}$ in scenario $s \in S$
$a_{i,s,k}$	{0, 1}	Binary variable to select simplex $k \in \mathcal{K}$ for pump $i \in \mathcal{P}$ in scenario $s \in S$

Set		Description
$\mathcal{Q}$		Set of grid points in direction of the volume flow q
$\mathcal{N}$		Set of grid points in direction of the speed n
$\mathcal{K}$		Set of simplices
$\mathcal{K}(u, v, i, s)$		Subset of simplices which contain grid point (u, v) for pump $i \in \mathcal{P}$ in scenario $s \in S$

Parameter	Range	Description
$Q_{u,v,i}$		Grid values in q direction for pump $i \in \mathcal{P}$
$N_{u,v,i}$		Grid values in n direction for pump $i \in \mathcal{P}$
$\Delta P_{u,v,i}$		Pressure value at grid point $(Q_{u,v,i}, N_{u,v,i})$ by evaluating Eq. 1 for pump $i \in \mathcal{P}$
$Po_{u,v,i}^{\text{pump}}$		Power value at grid point $(Q_{u,v,i}, N_{u,v,i})$ by evaluating Eq. 2 for pump $i \in \mathcal{P}$

$$po_{i,s}^{\text{pump}} = \sum_{m=0}^{3} \beta_{i,m} \, q_{i,s}^{\,3-m} \, n_{i,s}^{\,m} \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \qquad (6b)$$

to the basic model in Eq. (5), where $\zeta^{\text{inst.}}$ is the pressure loss factor of the fittings used for installing the pump into the system. The benefits of this approach are a fast implementation and a highly accurate model, while the drawbacks are potentially high solution times.

4.3 Piecewise linearized pump model

A common approach to solve nonlinear non-convex optimization models with integer variables, is to make use of approximations based on piecewise linearization (PWL) techniques, cf. Geißler et al. (2012). Linearization of all nonlinear constraints results in a mixed-integer linear program, which can then be solved by state-of-the-art solvers as for example GUROBI (GUROBI 2020) or CPLEX (IBM 2020). Piecewise linearization methods are therefore applied in a high variety of applications, both within classical operations research, as well as technical applications, e.g. Misener et al. (2009), Morsi et al. (2012), Rausch et al. (2016) and Mikolajková et al. (2018). The underlying modeling to reformulate a nonlinear functional relationship as linear constraints can be implemented by multiple methods, as shown by Misener and Floudas (2010), Vielma et al. (2010) and Geißler et al. (2012).

In this contribution, we use the convex combination method, as described by Vielma et al. ([2010]), to model the nonlinear relations of the pressure head $\Delta p = f(q, n)$ and of the power $po = g(q, n)$ as piecewise linearized function approximations. All additional variables, sets and parameters for the PWL are shown in Table [3]. Instead of using Eq. ([6]) to model the nonlinear pump characteristics, we model a piecewise linearized approximation by using the set of constraints ([7]).

$$q_{i,s} - \sum_{u \in \mathcal{Q}} \sum_{v \in \mathcal{N}} \lambda_{u,v,i,s} Q_{u,v,i} \underset{\geq -}{\overset{\leq +}{\gtrless}} (1 - x_{i,s})\overline{Q} \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{7a}$$

$$n_{i,s} - \sum_{u \in \mathcal{Q}} \sum_{v \in \mathcal{N}} \lambda_{u,v,i,s} N_{u,v,i} \underset{\geq -}{\overset{\leq +}{\gtrless}} (1 - x_{i,s})\overline{N}_i \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{7b}$$

$$\Delta p_{i,s} - \sum_{u \in \mathcal{Q}} \sum_{v \in \mathcal{N}} \lambda_{u,v,i,s} \Delta P_{u,v,i} \underset{\geq -}{\overset{\leq +}{\gtrless}} (1 - x_{i,s})P \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{7c}$$

$$po_{i,s}^{\text{pump}} - \sum_{u \in \mathcal{Q}} \sum_{v \in \mathcal{N}} \lambda_{u,v,i,s} Po_{u,v,i}^{\text{pump}} \underset{\geq -}{\overset{\leq +}{\gtrless}} (1 - x_{i,s})\overline{Po}_i \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{7d}$$

$$\sum_{u \in \mathcal{Q}} \sum_{v \in \mathcal{N}} \lambda_{u,v,i,s} = 1 \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{7e}$$

$$\sum_{k \in \mathcal{K}} a_{i,s,k} = 1 \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{7f}$$

$$\lambda_{u,v,i,s} - \sum_{k \in \mathcal{K}(u,v,i,s)} a_{i,s,k} \leq 0 \quad \forall u \in \mathcal{Q}, v \in \mathcal{N}, i \in \mathcal{P}, s \in \mathcal{S}. \tag{7g}$$

The index sets $\mathcal{Q}$ and $\mathcal{N}$ represent grid points in the (q, n) domain in which the pressure $\Delta P_{u,v,i,s}$ and power $Po_{u,v,i,s}^{\text{pump}}$ are evaluated for each pump $i \in \mathcal{P}$ before optimization according to Eqs. ([1]) and ([2]). The pressure losses and additional power for frequency converters are added to these values in the preprocessing of the optimization to derive an equivalent description, as shown in Constr. ([6]). Besides these grid points, we have to use additional binary variables $a_{i,s,k} \in \{0, 1\}$ and continuous variables $\lambda_{u,v,i,s} \in [0, 1]$ to ensure a piecewise linearized approximation. The set $\mathcal{K}$ in Constr. ([7f]) contains all segments of the piecewise linearized approximation. In turn, the subset $\mathcal{K}(u, v, i, s)$ in Constr. ([7g]) covers all segments that contain the grid point (u, v) for pump i and scenario s. These additional constraints are added to the optimization program in model ([5]), which results in a mixed-integer linear program that can be solved by common MILP solvers.

4.4 Piecewise linear relaxed pump model

As a third possibility we use piecewise linear relaxed pump models within the final optimization program. Relaxations are often used in the MINLP solution process, as shown by Puranik and Sahinidis ([2017]), to derive valid lower bounds. A general introduction to relaxation methods for nonlinear network design problems is given by Humpola and Fügenschuh ([2015]). Relaxation methods with piecewise linear underes-
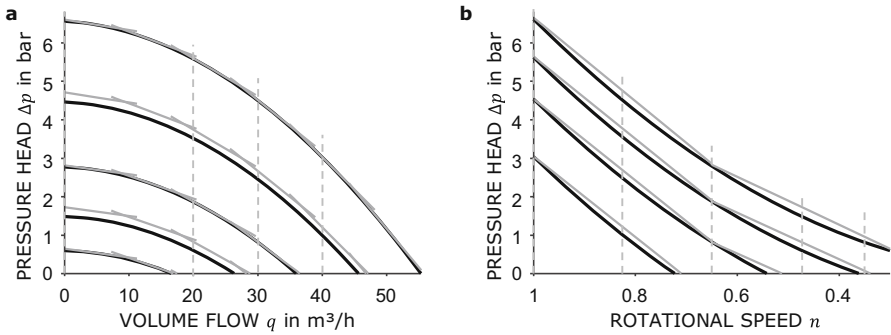
Fig. 6 Piecewise linear overestimation of the quadratic function of the pressure $\Delta p = f(q, n)$. Shown are **a** the function $\Delta p(q)$ for different speeds n and **b** the function $\Delta p(n)$ for different volume flows q. $R = 6$ elements in q direction and $K = 2$ elements in n direction are used. The dashed lines in **a** show the values of q for which the lines were plotted in **b** and vice versa

timators are also used in MINLP solvers to improve the solution speed, cf. Misener and Floudas (2012), Misener and Floudas (2013), Misener and Floudas (2014). Besides these multi-purpose implementations, these methods can also be used to improve the solution process for domain specific optimization models, like for example shown by Pecci et al. (2019) in the domain of water supply networks. The authors derived global bounds on the placement of valves by using a Branch-and-Bound algorithm with relaxed nonlinear constraints. As shown in all previous references, the usage of piecewise relaxed underestimators can improve the solution process. For this reason, we developed a relaxed pump model to improve the solution speed of the solving algorithm, cf. Sect. 5.

In order to underestimate the total cost it is necessary to overestimate the pressure increase and to underestimate the power consumption of the pumps:

$$\Delta p_{i,s} \leq \Delta p_{i,s}^{\text{relaxed}} \quad \forall i \in \mathcal{P}, s \in \mathcal{S}, \tag{8a}$$

$$po_{i,s}^{\text{pump}} \geq po_{i,s}^{\text{relaxed}} \quad \forall i \in \mathcal{P}, s \in \mathcal{S}. \tag{8b}$$

For these estimations, piecewise defined planes are determined, which serve as linear under- or overestimators. They have to be defined piecewise, since $\Delta p = f(q, n)$ is not concave in n and $po^{\text{pump}} = g(q, n)$ is not convex in q. The functions are shown in Figs. 6 and 7 along with the over- and underestimations.

We define R equidistant segments in q- and K equidistant segments in n direction. The boundaries of these segments are described by the sets $\mathcal{Q}^{\text{seg}} = \{Q_1^{\text{seg}}, ..., Q_{1+R}^{\text{seg}}\}$ and $\mathcal{N}^{\text{seg}} = \{N_1^{\text{seg}}, ..., N_{1+K}^{\text{seg}}\}$. For each segment (u, v) and each function $(\Delta p = f(q, n), po^{\text{pump}} = g(q, n))$ a plane is defined, which should approximate the function as good as possible. Each plane is defined by the parameters $\gamma_{i,k,u,v}$ ($\Delta p = f(q, n)$) and $\delta_{i,k,u,v}$ ($po^{\text{pump}} = g(q, n)$) with $i \in \mathcal{P}$, $k \in \{1, 2, 3\}$, $u \in \{1, ..., R\}$, $v \in \{1, ..., K\}$.
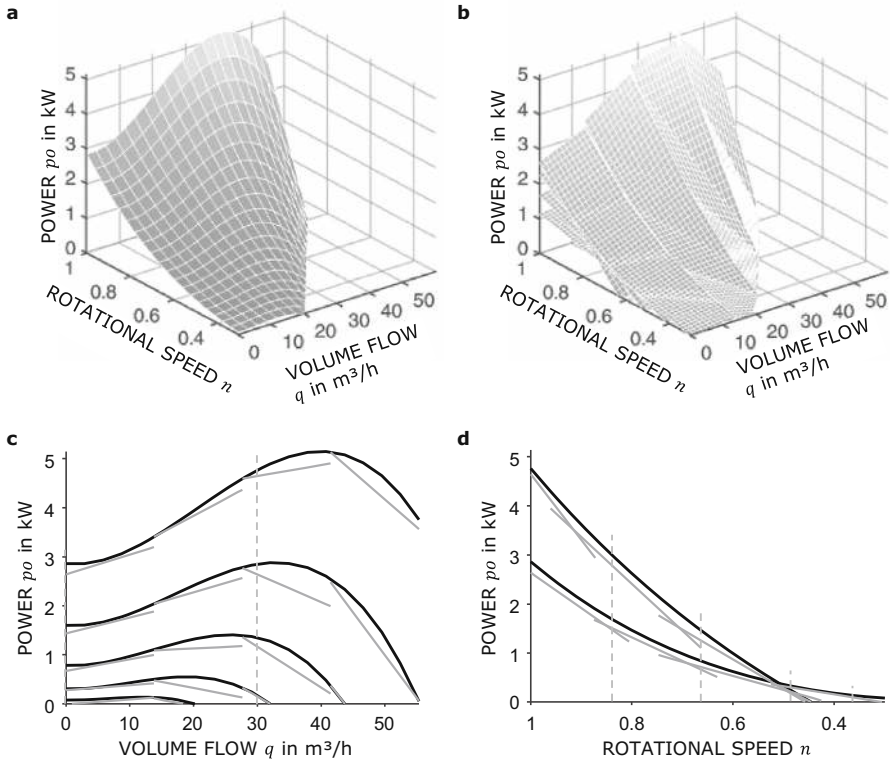
Fig. 7 Piecewise linear underestimation of the cubic function of the power consumption $po = g(q, n)$. **a** shows the nonlinear function and **b** the piecewise underestimation. **c** shows the function $po(q)$ for different speeds n and **d** $po(n)$ for different volume flows (q). $R = 4$ elements in q direction and $K = 4$ elements in n direction are used. The dashed lines in **c** show the values of q for which the lines were plotted in **d** and vice versa

To solve the overall MINLP, cf. Sect. 4, the constraints of the pump characteristics (6) are replaced by the following constraints:

$$\Delta p_{i,s} - \left(\gamma_{i,1,u,v}\, q_{i,s} + \gamma_{i,2,u,v}\, n_{i,s} + \gamma_{i,3,u,v}\right) \le (1 - x_{i,s,v}^{\gamma})\overline{M}_1$$

$$\forall u \in \{1, ..., R\},\, v \in \{1, ..., K\},\, i \in \mathcal{P},\, s \in \mathcal{S} \qquad (9a)$$

$$po_{i,s}^{\text{pump}} - \left(\delta_{i,1,u,v}\, q_{i,s} + \delta_{i,2,u,v}\, n_{i,s} + \delta_{i,3,u,v}\right) \ge -(1 - x_{i,s,u}^{\delta})\overline{M}_2$$

$$\forall u \in \{1, ..., R\},\, v \in \{1, ..., K\},\, i \in \mathcal{P},\, s \in \mathcal{S} \qquad (9b)$$

$$x_{i,s,v}^{\gamma 1} + x_{i,s,v}^{\gamma 2} \le x_{i,s,v}^{\gamma} \quad \forall v \in \{1, ..., K\},\, i \in \mathcal{P},\, s \in \mathcal{S} \qquad (9c)$$

$$n_{i,s} \ge (\mathcal{N}_v^{\text{seg}} + \varepsilon)(1 - x_{i,s,v}^{\gamma 1}) \quad \forall v \in \{1, ..., K\},\, i \in \mathcal{P},\, s \in \mathcal{S} \qquad (9d)$$

$$n_{i,s} - (\mathcal{N}_{v+1}^{\text{seg}} - \varepsilon) \le (1 - x_{i,s,v}^{\gamma 2}) \quad \forall v \in \{1, ..., K\},\, i \in \mathcal{P},\, s \in \mathcal{S} \qquad (9e)$$

$$x_{i,s,u}^{\delta 1} + x_{i,s,u}^{\delta 2} \le x_{i,s,u}^{\delta} \quad \forall u \in \{1, ..., R\},\, i \in \mathcal{P},\, s \in \mathcal{S} \qquad (9f)$$

$$q_{i,s} \geq (Q_u^{\text{seg}} + \varepsilon)(1 - x_{i,s,u}^{\delta 1}) \quad \forall u \in \{1, ..., R\}, i \in \mathcal{P}, s \in \mathcal{S} \tag{9g}$$

$$q_{i,s} - (Q_{u+1}^{\text{seg}} - \varepsilon) \leq \overline{Q}(1 - x_{i,s,u}^{\delta 2}) \quad \forall u \in \{1, ..., R\}, i \in \mathcal{P}, s \in \mathcal{S} \tag{9h}$$

Constraints (9a) and (9b) guarantee that the pressure in each segment is under-estimated and the power overestimated. Here, $x_{i,s,v}^{\gamma}$ and $x_{i,s,u}^{\delta}$ are binary activation variables which activate a plane if and only if the speed (in case of pressure overes-timation) or volume flow (in case of power underestimation) is within the segment boundaries. A small value of $\varepsilon = 10^{-3}$ is necessary to ensure that at least one plane is active at the boundaries of the piecewise defined planes. $\overline{M}_1$ and $\overline{M}_2$ are upper bounds which are necessary for the bigM-constraints. Constraints (9c) to (9h) are used to check the mentioned boundaries with binary auxiliary variables $x^{\gamma 1}$, $x^{\gamma 2}$, $x^{\delta 1}$ and $x^{\delta 2}$.

The parameters of the planes are determined using a MILP. The objective is to find planes within a segment which describe the function as good as possible and over- or underestimate it. To model this, a two-dimensional fine grid $\mathcal{Q}^{\text{grid}}$ and $\mathcal{N}^{\text{grid}}$ is defined: at each grid point it is constrained that the overestimation or underestimation is maintained and the distance to the nonlinear function is calculated. The set $\mathcal{Q}^u$ with $u \in \{1, ..., R\}$ and $\mathcal{N}^v$ with $v \in \{1, ..., K\}$ describes the grid points within a segment. The following applies:

$$\mathcal{Q}^u = \{z \in \mathcal{Q}^{\text{grid}} | Q_u^{\text{seg}} \leq z \leq Q_{u+1}^{\text{seg}}\} \quad \forall u \in \{1, ..., R\}, \tag{10a}$$

$$\mathcal{N}^v = \{z \in \mathcal{N}^{\text{grid}} | N_v^{\text{seg}} \leq z \leq N_{v+1}^{\text{seg}}\} \quad \forall v \in \{1, ..., K\}. \tag{10b}$$

With these definitions, both, MILP (11) for the pressure characteristic and MILP (12) for the power characteristics are formulated for each pump $i \in \mathcal{P}$ and for each segment (u, v) with $u \in \{1, ..., R\}, v \in \{1, ..., K\}$.

The objective (11a)/(12a) is to minimize the distance between the nonlinear function and its over-/underestimation, respectively. Constraint (11b) models the overestima-tion of the pressure characteristic and Constr. (12b) the underestimation of the power characteristic at every grid point. Since the number of grid points is limited, distance factors $\varepsilon^{\Delta p} = 0.005\,bar$ and $\varepsilon^{Po} = 0.5\,W$ are introduced. These define the minimum distance between the linear planes and the nonlinear function at the grid points. Thus, it is ensured that the planes are over/underestimating also between grid points, what has been checked in the postprocessing. The result for the pressure characteristic is shown in Fig. 6 and for the power characteristic in Fig. 7.

$$\min \sum_{x=1}^{|\mathcal{Q}^u|} \sum_{y=1}^{|\mathcal{N}^v|} \left(\gamma_{i,1,u,v} Q_x^{\text{u}} + \gamma_{i,2,u,v} N_y^{\text{v}} + \gamma_{i,3,u,v} \right) - $$
$$\left((\alpha_{i,0} - \zeta^{\text{inst.}})(Q_x^{\text{u}})^2 + \sum_{m=1}^{2} \alpha_{i,m} (Q_x^{\text{u}})^{2-m} (N_y^{\text{v}})^m \right) \tag{11a}$$

subject to

$$\gamma_{i,1,u,v} Q_x^u + \gamma_{i,2,u,v} N_y^v + \gamma_{i,3,u,v} \geq$$

$$(\alpha_{i,0} - \zeta^{\text{inst.}})(Q_x^u)^2 + \sum_{m=0}^{2} \alpha_{i,m} (Q_x^u)^{2-m} (N_y^v)^m + \varepsilon^{\Delta p} \tag{11b}$$

$$\forall x \in \{1, ..., |\mathcal{Q}^{\text{grid}}|\}, y \in \{1, ..., |\mathcal{Q}^u|\}$$

$$\min \sum_{x=1}^{|\mathcal{Q}^u|} \sum_{y=1}^{|\mathcal{N}^v|} \left(\beta_{i,1}(Q_x^u)^3 + \beta_{i,2}(Q_x^u)^2 N_y^v + \beta_{i,3} Q_x^u (N_y^v)^2 + \beta_{i,4}(N_y^v)^3 + \beta_{i,5} \right) -$$

$$\left(\delta_{i,1,u,v} Q_x^u + \delta_{i,2,u,v} N_y^v + \delta_{i,3,u,v} \right) \tag{12a}$$

subject to

$$\beta_{i,1}(Q_x^u)^3 + \beta_{i,2}(Q_x^u)^2 N_y^v + \beta_{i,3} Q_x^u (N_y^v)^2 + \beta_{i,4}(N_y^v)^3 + \beta_{i,5} \geq$$

$$\delta_{i,1,u,v} Q_x^u + \delta_{i,2,u,v} N_y^v + \delta_{i,3,u,v} + \varepsilon^{\text{Po}}$$

$$\forall x \in \{1, ..., |\mathcal{Q}^{\text{grid}}|\}, y \in \{1, ..., |\mathcal{Q}^u|\} \tag{12b}$$

5 Solving algorithm

The MINLP or piecewise linearized MILP presented in Sect. 4 can be solved with conventional solvers. This, however, leads to high computation times or large duality gaps for practically relevant problem sizes. For this reason, we developed a problem specific, exact solving algorithm, which allows significantly shorter computation times and leads to an acceptable result accuracy, even for the considered nonconvex MINLP.

The basic idea is to divide the problem into a number of smaller problems and to explore this solution space systematically. For this purpose, the set of all pumps $\mathcal{P}$ is reduced to a significantly smaller subset $\mathcal{P}^{\text{reduced}} \subseteq \mathcal{P}$ and the purchase decisions for these pumps are fixed to one. This results in a fixation of the most important first stage decision variables. The number of pumps in the reduced set $|\mathcal{P}^{\text{reduced}}|$ is less than or equal to the maximum number of pumps that can be purchased N^{buy}, which drastically reduces the number of binary variables: $|\mathcal{P}^{\text{reduced}}| \leq N^{\text{buy}} \ll |\mathcal{P}|$. If M is the number of different pump types in $\mathcal{P}$, then we can calculate the total number of tuples $I = |\mathcal{T}|$ with:

$$I = \binom{M + N^{\text{buy}}}{M + 1} = \frac{(M + N^{\text{buy}})!}{(M + 1)! \cdot (N^{\text{buy}} - 1)!}. \tag{13}$$

For the investigated instances ($M = 15$ and $N^{\text{buy}} = 4$) this results in $I = 3876$ reduced MINLP with $|\mathcal{P}^{\text{reduced}}| \leq 4$. When considering a decentral topology (e.g. test instance

---

Algorithm 2: Generic description of the solving algorithm.

---

input : set of pumps $\mathcal{P}$; load scenarios; instance-specific parameters
output: best topology and control; solution information

```
// preprocessing
```
initialize all possible feasible purchase decisions $\mathcal{T}$
$\mathcal{I} = \{1, 2, \ldots, I\}$
$\mathcal{L} = \{L_1, L_2, \ldots, L_I\}$
$\mathcal{U} = \{U_1, U_2, \ldots, U_I\}$
$\mathcal{U} \leftarrow U_j = \infty \, \forall \, j \in \mathcal{I}$
$\mathcal{L} \leftarrow L_j = \sum_{s \in S} W_s \, Po_s^{\text{hydr}} \, \forall \, j \in \mathcal{I}$
$U^{\text{global}} = \infty, L^{\text{global}} = \min(\mathcal{L})$

```
// main part
```
while $U^{\text{global}} > \mathcal{L}^{\text{global}}$ **do**

> ```
> // heuristic for upper-bound-search
> ```
> estimate primal bound for each tuple, $U_j^{\text{est}} \, \forall \, j \in \mathcal{I}$
>
> solve reduced MINLP for the most promising tuple (lowest U_j^{est})
>
> **if** *infeasible or dual bound* $> U^{\text{global}}$ **then**
> > remove tuple
>
> **else**
> > update tuple information in: $\mathcal{U}, \mathcal{L}$
> update global information: $U^{\text{global}} = \min(\mathcal{U}), L^{\text{global}} = \min(\mathcal{L})$
>
> ```
> // lower-bound-search
> ```
> solve reduced MINLP or relaxed, reduced MILP for tuple with lowest dual bound in $\mathcal{L}$
> **if** *infeasible or dual bound* $> U^{\text{global}}$ **then**
> > remove tuple
>
> **else**
> > update tuple information in: $\mathcal{U}$ (only if MINLP was solved), $\mathcal{L}$
> update global information: $U^{\text{global}} = \min(\mathcal{U}), L^{\text{global}} = \min(\mathcal{L})$

B1.D, cf. 6.1) this yields a MINLP with 425 variables (64 binary) and 1057 constraints (41 nonlinear). In contrast to this is a single MINLP with $|\mathcal{P}| = M \times N^{\text{buy}} = 60$ pumps. This yields 26581 variables (4380 binary) and 82562 constraints (601 nonlinear). The resulting reduced MINLP is therefore less challenging in terms of combinatorics and can generally be solved faster. The goal is to efficiently explore the set of possible purchase decisions $\mathcal{T}$, given the reduced set of available pumps, while also cutting off non-optimal solutions.

The solving algorithm is shown in Algorithm 2. For better clarity, we define the index set $\mathcal{I} = \{1, 2, \ldots, I\}$ based on Eq. (13). First, the set of pump-tuples $\mathcal{T}$ is created and physically inadmissible tuples are removed during preprocessing. In addition, a set of lower bounds $\mathcal{L} = \{L_1, L_2, \ldots, L_I\}$ for the energy costs of each tuple $T_j \in \mathcal{T}$ is analytically calculated based on domain-specific knowledge. Besides, a set of all primal upper bounds $\mathcal{U} = \{U_1, U_2, \ldots, U_I\}$ is defined and each part is set to ∞ before optimization. For the considered overall minimization problem, the primal bound is the minimal primal solution of a tuple $U^{\text{global}} = \min(\mathcal{U})$ (best found solution), which is infinite at the start. The dual bound of the overall minimization problem L^{global} should be raised. Therefore, the weakest dual bound in one of the subproblems/tuples is determined. Thus, the dual bound of the overall problem is

bounded by the lowest dual bound of a tuple $L^{\text{global}} \geq \min(\mathcal{L})$. It is initialized using the preprocessed bounds for the energy and the known investment costs of the pumps. Now the search for better primal and dual solutions is repeated until the overall problem has been solved to proven global optimality. The global primal and dual bounds are updated in each step and proven weak solutions or infeasible solutions are cut off. In the following, the single steps of the solving algorithm are explained in detail.

5.1 Preprocessing

First, the set of purchase decisions $T_j \in \mathcal{T} \quad \forall j \in \mathcal{I}$ is generated, each consisting of a reduced set of pumps $\mathcal{P}_j^{\text{reduced}} \subseteq \mathcal{P}$ with $j \in \mathcal{I}$. For each tuple $T_j \in \mathcal{T}$, it is determined whether the pumps of the tuple $\mathcal{P}_j^{\text{reduced}}$ can fulfill the required load demand. There are three elimination criteria:

(i) The sum of the maximum pressure head of all pumps is smaller than required. Even a series connection of all pumps could not provide the required pressure.
(ii) The sum of the maximum volume flow of all pumps is smaller than required. Even a parallel connection of all pumps could not provide the required volume flow.
(iii) The sum of the maximum hydraulic power of all pumps is smaller than required. Even if all pumps would be operated at the point of maximum hydraulic power, it would not be sufficient.

Tuples that comply with any of these criteria are deleted immediately. This can efficiently reduce the number of tuples. Nevertheless, this excludes only a subset of all infeasible solutions.

The minimum cost are underestimated for each of the remaining tuples. Since the purchase decision for a certain pump set is fixed, the investment costs of this pump set can be calculated directly. The purchase decision for frequency converters is set to zero. We underestimate the power consumption and thus the energy cost. For each $T_j \in \mathcal{T}$ the underestimated power consumption $Po_{s,j}^{\text{lb,phys}}$ is calculated by the hydraulic power Po_s^{hydr} and best efficiency of any pump in the tuple $\eta_{\text{best},j}^{\text{pump}}$ in every scenario $s \in \mathcal{S}$. For each tuple $j \in \mathcal{I}$ the following underestimation is valid:

$$\sum_{s \in S} W_s \sum_{i \in \mathcal{P}} po_{i,s} \geq \sum_{s \in S} W_s \, Po_{s,j}^{\text{lb,phys}} = \sum_{s \in S} W_s \, Po_s^{\text{hydr}} / \eta_{\text{best},j}^{\text{pump}}. \tag{14}$$

This corresponds to the hypothetical case that the pump with the highest efficiency provides the required hydraulic power in its best-efficiency point all by itself. The resulting cost according to Obj. (5a) underestimates the true cost and is thus a lower bound, which is based on physical principles. This consideration allows to cut off many tuples early. The power dissipated in the frequency converter is usually one order of magnitude smaller than the electrical power in the motor and is therefore neglected in this estimation.

Finally, the global primal and dual bounds are initialized in preprocessing. The global primal bound is the value of the best valid solution and is initially set to infinity.

The global dual bound equals the minimum value of the lower bound of any tuple. It is initialized to the minimum of the values estimated by Eq. (14).

5.2 Bound improvement

After preprocessing, the gap between the primal and dual bound is attempted to be reduced over time. On the one hand, better primal solutions are searched for. On the other hand, the minimum values of the lower bound of the tuples are improved to increase the dual bound of the overall minimization problem. If the resulting value of the lower bound of a tuple is above the global primal bound, it can be cut off.

5.2.1 Upper bound search

In order to find a new, better primal solution, the objective value of each tuple $T_j \in \mathcal{T}$ is estimated in each step. The exact MINLP is then solved for the tuple with the best estimated objective value, i.e. the most promising set of pumps.

Estimation of the Objective Value The value of the primal solution of a tuple $T_j \in \mathcal{T}$ is estimated using a heuristic. It is assumed that the difference between the lower bound of the power consumption estimated during preprocessing $Po_{s,j}^{\mathrm{lb,phys}}$, Eq. (14), and the power consumptions' exact value of the primal solution of a tuple $Po_{s,j}^{\mathrm{primal}}$ is comparable for all tuples $j \in \mathcal{I}$. We calculate this difference for all tuples in $\mathcal{T}$ for which we already know the primal solution by:

$$v_{s,j} = (Po_{s,j}^{\mathrm{primal}} - Po_{s,j}^{\mathrm{lb,phys}})/Po_{s,j}^{\mathrm{lb,phys}} \quad \forall s \in \mathcal{S}, j \in \mathcal{I}. \tag{15}$$

The expected power consumption $Po_{s,j}^{\mathrm{exp}}$ of tuples for which we have not yet computed a primal solution is estimated using the mean of the deviation $v_{s,j}$ of all tuples with known primal solution:

$$\sum_{s \in S} W_s \sum_{i \in \mathcal{P}} po_{i,s} \approx \sum_{s \in S} W_s Po_{s,j}^{\mathrm{exp}}$$
$$= \sum_{s \in S} W_s \left(1 + \frac{\sum_{m \in \mathcal{I}^{\mathrm{primal}}} v_{s,m}}{|\mathcal{I}^{\mathrm{primal}}|} \right) Po_{s,j}^{\mathrm{lb,phys}} \quad \forall j \in \mathcal{I}, \tag{16}$$

where $\mathcal{I}^{\mathrm{primal}} = \{j \in \mathcal{I} \,|\, \exists\, U_j < \infty\}$ is the index set of tuples with finite primal solution. Based on this, the total cost is estimated according to Obj. (5a). Here, the exact known investment costs of the pumps are considered while the costs for the frequency converters are neglected.

Search for Better Primal Solution When searching for a new primal solution, the tuple with the lowest expected cost is selected. The set of pumps is reduced to the purchased pumps $\mathcal{P}_j^{\mathrm{reduced}}$ in the selected tuple T_j. The MINLP from Sect. 4 is formulated with the reduced set of pumps and the purchase decision variables are set to $y_i = 1 \;\forall\, i \in \{1, ..., |\mathcal{P}_j^{\mathrm{reduced}}|\}$.

In general, the resulting reduced problems for the tuples in the set $\mathcal{T}$ can be solved considerably faster than the overall problem. Nevertheless, high computation times may occur for some tuples $T_j \in \mathcal{T}$. Depending on the count for selection of the tuple, the computation time is limited. Starting with 15 s for first time computations, the limit is increased gradually for each reselection of a tuple. By this method, we try to avoid wasting computation time on suboptimal solutions. When solving the selected tuple, for feasible tuples, the primal bound U_j and dual bound L_j are stored for the tuple. In turn, when the subproblem of the tuple $T_j \in \mathcal{T}$ is infeasible, it is removed from the tuple list $\mathcal{T}$. We expect that the converged solution can already be cut off at an early stage without the need to solve the reduced problem to global optimality. In the best case, only a single reduced MINLP needs to be solved to global optimality. If a subproblem for a tuple $T_j \in \mathcal{T}$ is solved to proven global optimality, it will be deactivated for a search for further primal solutions. Unfortunately, it is not possible to continue or warm-start the computation with the selected programming framework[1] later on. Instead, the computation has to be re-started, this time with a longer maximum computation time ($\geq$ 15 s). For this reason, it is important to carefully select the maximum computation time for the individual steps.

If there is no promising tuple left, which we define as the lowest expected cost exceeding 120% of the global primal solution, the search for better primal solutions is no longer executed. However, the lower bound search (cf. Sect. 5.2.2) also still allows to find better primal solutions, if the reduced MINLP is solved in a later step.

5.2.2 Lower bound search

Alternating with the search for a better (i.e. lower) primal bound, the dual bound is attempted to be raised. The procedure is similar to the primal bound search: a suitable tuple is selected and a reduced optimization problem is solved. The search for a better (i.e. higher) dual bound is repeated three times before again continuing the search for a lower primal solution, as we found that it is generally more difficult than finding a better primal solution. In the dual bound search the tuple with the lowest individual dual bound $min(\mathcal{L})$ is selected (called weakest tuple). This is the weakest individual dual bound and determines the dual bound of the overall problem $L^{\text{global}} \leq min(\mathcal{L})$. To increase the overall dual bound, the dual bound of the weakest tuple has to be increased by solving the corresponding subproblem. The dual bounds for tuples that have not yet been examined are the underestimated cost of the preprocessing (cf. Eq. (14)). In turn, for tuples that have already been examined, the dual bound of the tuple from the previous step is applied.

The solved optimization problem corresponds to the one from the primal bound search with a special feature: If the tuple is selected for the first time, the piecewise linear relaxed problem (PWR MILP, cf. Sect. 4.4) of the reduced MINLP is set up and solved (computation time limit of 60 s) instead of directly solving the reduced MINLP. The piecewise linear relaxed MILP is in general easier to solve and the so computed lower bound is often sufficiently high to cut off the tuple. If the solution could not

---

[1] We implemented the solving algorithm in MATLAB 2019a using amongst others YALMIP (Löfberg 2004) and OPTI Toolbox (Currie et al. 2012).

be cut off, the tuple is re-selected at a subsequent loop iteration and then the reduced MINLP is solved, whereby the initial computation time is again limited to 15 s and subsequently increased in later iterations.

5.2.3 Parallelization

To further reduce computation times, the presented solving algorithm is parallelized. In each step N optimization problems are solved simultaneously on different cores: Therefore, not only a single tuple with the best expected primal bound (in case of primal search) or lowest dual bound (in case of dual search) is selected, but N tuples with the best expected primal or lowest dual bounds are solved. If a tuple is solved before finishing the computation of any of the other first N tuples, the next best/weakest tuple is solved on the available core. The results of the computed tuples are synchronized after the computation of the first N tuples is finished.

6 Computational results

In this section the results of the optimization problems are discussed. First, the set of test instances, which are scenarios of high technological importance, are presented and the implementation and settings are clarified. Based on this, the performance of the different solving approaches is presented. Afterwards, a technical discussion of the results and the significance for practice are given.

6.1 Selection of test instances

The previously presented different systems (Sect. 3) and the wide range of variations are analyzed carefully to choose technologically meaningful test instances. Tested instances are numerated in Table 4, in which the different possible physical variants are listed. For each of the buildings B1–B3 and the test rig three different topologies are considered. (i) Parallel connection of the pumps for a central booster station ($|\mathcal{F}| = 1$, Constr. (4) active), hereinafter referred to as parallel. (ii) Unrestricted topology for a central booster station ($|\mathcal{F}| = 1$, Constr. (4) not active), hereinafter referred to as central. (iii) Unrestricted topology for a decentral booster station ($|\mathcal{F}| = 5$, Constr. (4) not active), hereinafter referred to as decentral. The complexity increases significantly in the order of the listed topologies due to their increasing degrees of freedom. The heuristic presented in Sect. 3.5 is used to select 15 different pump types for each of the buildings B1–B3. Each pump type is available $N^{\text{avail}} = 4$ times which leads to a total set of $|\mathcal{P}| = 15 \times 4 = 60$ pumps. In total, up to $N^{\text{buy}} = 4$ pumps can be purchased and installed per booster station. For the test rig the set of pumps is limited to the $|\mathcal{P}| = 13$ available pumps (6 different types), where up to $N^{\text{buy}} = 6$ can be purchased and installed. While for the buildings B1–B3 a frequency converter has to be installed additionally to allow for speed control, all considered pumps for the test rig are already equipped with frequency converters and therefore the decision to buy a frequency converter is disregarded. The characteristic curves as well as costs of the

Table 4 Instances for testing the different solution approaches

	parallel	free topology central	decentral
Building 1	B1.P	B1.C	B1.D
Building 2	B2.P	B2.C	B2.D
Building 3	B3.P	B3.C	B3.D
Test rig w/ exp. data	TRE.P	TRE.C	TRE.D

pumps and frequency converters are obtained from manufacturer data. The loss factor ζ^{inst} is determined experimentally for the test rig, whereas it is set to zero for all real buildings. The load profiles match the load profiles presented in Sect. 3.

The lifetime $T = 15$ a for pumps is derived from the depreciation table ("AfA Table") for the sector "Energy and Water Supply" of the German Federal Ministry of Finance (AfA table 1995). For the interest rate $r = 2.08\%$ the discount rate (Bundesbank 2019) is used with a remaining lifetime of 15 years (7 year average, key date Aug 31st 2019). For the energy cost ($C_{\text{elect}} = 0.3$ € $/kWh$) the mean value of the household electricity price for the second half of 2018 in Germany is consulted (Destatis 2019). This results in an annuity-present value factor of $APF = \sum_{t=1a}^{T} 1/(1+r)^t = 12.77$. For the investment costs of the pumps and frequency converters, we rely on market prices for the European Union of the year 2018. The pressure is calculated in bar and the volume flow in m^3h^{-1}. For the buildings B1–B3 the cost are calculated in k€ and the power in kW, for the test rig in € and W. Thus, the coefficients of the problem become better scaled.

The three different optimization models, namely the nonlinear (MINLP), piecewise linearized (PWL), and piecewise relaxed model (PWR), can be solved using different off-the-shelf solvers and the developed solving algorithm. In this work the Solvers GUROBI v8.1.1[2], SCIP v5.0[3] and BARON v19.7.13[4] as well as the solving algorithm (Algorithm 2 , cf. Sect. 5.1) are trialed to compare computation time and solution quality. If nothing else is specified, default settings are used. For the PWL, 32 simplices used for buildings B1-B3 and 50 simplices for the test rig, which are evenly distributed. We use a 1-4 orientation for each grid, as introduced in Misener and Floudas (2010). For the PWR a discretization of $R = 6$ elements in q direction and $K = 3$ elements in n direction for the pressure and $R = 5$, $K = 5$ for the power is used for real buildings. For the test rig a discretization of $R = 6$, $K = 2$ for the pressure and $R = 4$, $K = 4$ for the power is used. For reasons of simplicity, an equidistant distribution has been chosen. The PWR is primarily used within the solving algorithm trying to cut off solutions fast. For this reason, a comparatively small number of grid points has been chosen, which has proven to be efficient in test runs.

---

[2] GUROBI (2020)

[3] Gleixner et al. (2017) (using IPOPT v3.12.9 (Wächter and Biegler 2006) and SoPlex v3.1. (Zuse Institute Berlin 2020))

[4] Tawarmalani and Sahinidis (2005) (using CPLEX v12.9 (IBM 2020))
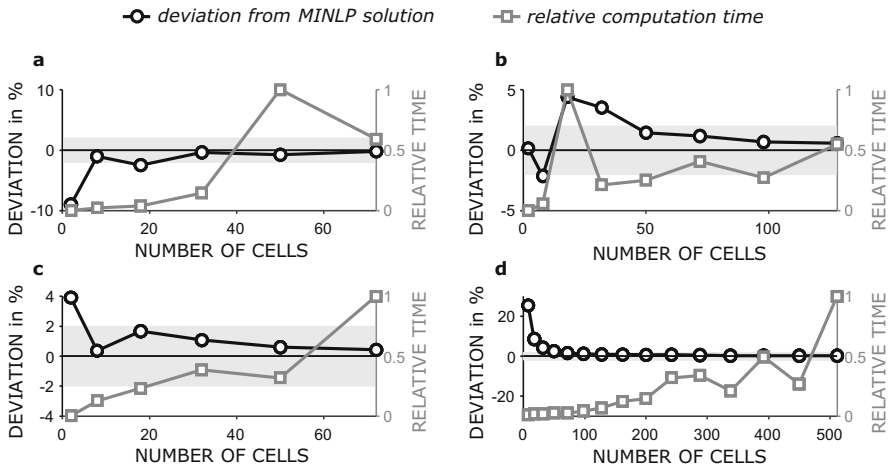
Fig. 8 Influence of grid size on deviation to MINLP solution and computation time. The 2 % limit is highlighted in gray. **a** B1.P, **b** B2.P, **c** B3.P, **d** TRE.P

The optimization model was implemented in Matlab 2019a using YALMIP (Löfberg 2004). The solver SCIP is interfaced via the OPTI toolbox (Currie et al. 2012). For all other solvers, manufacturer interfaces are used. The solving algorithm presented in Sect. 5.1 has also been developed and executed in Matlab 2019a. All computations were performed on a Windows 10 machine with an Intel i7-8700 CPU (6 cores @3.20GHz) and 16GB RAM. ε-optimality (with $\varepsilon = 0.1$) is tried to be achieved within a time limit of 3 hours.

6.2 Performance of different solution approaches

We first discuss the influence of the grid-size of the PWL-MILP on the computation time and the deviation from the MINLP solution, cf. Fig. 8. Afterwards, we compare the different solution approaches and solvers, for which we have computed the instances given in Table 4. A large part of the instances have a gap $> \varepsilon$, which is why a comparison of the gap at the end of the computation time is performed. As in Pecci et al. (2019), the share of solved instances with a gap smaller δ, $\rho(\delta) = (\text{\# instances with gap} \leq \delta)/(\text{\# test instances})$, is plotted for the remaining gap δ, cf. Fig. 9.

Comparison of MINLP and MILP The influence of the discretization is investigated using the comparatively simple instances with only parallel pump configurations (B1.P, B2.P, B3.P, TRE.P). Figure 8 shows the deviation from the best known MINLP solution and the normalized computation time over the number of cells. The deviation converges towards 0, whereby the computation time increases as a trend. Partly an overestimation, partly an underestimation of the objective value of the MINLP occurs. On the basis of the investigation, 32 cells for real buildings and 50 for the test rig are selected. With an average deviation of less than 2 % (corresponding with the order of magnitude of the measurement errors on the test rig), the level of detail is sufficient.
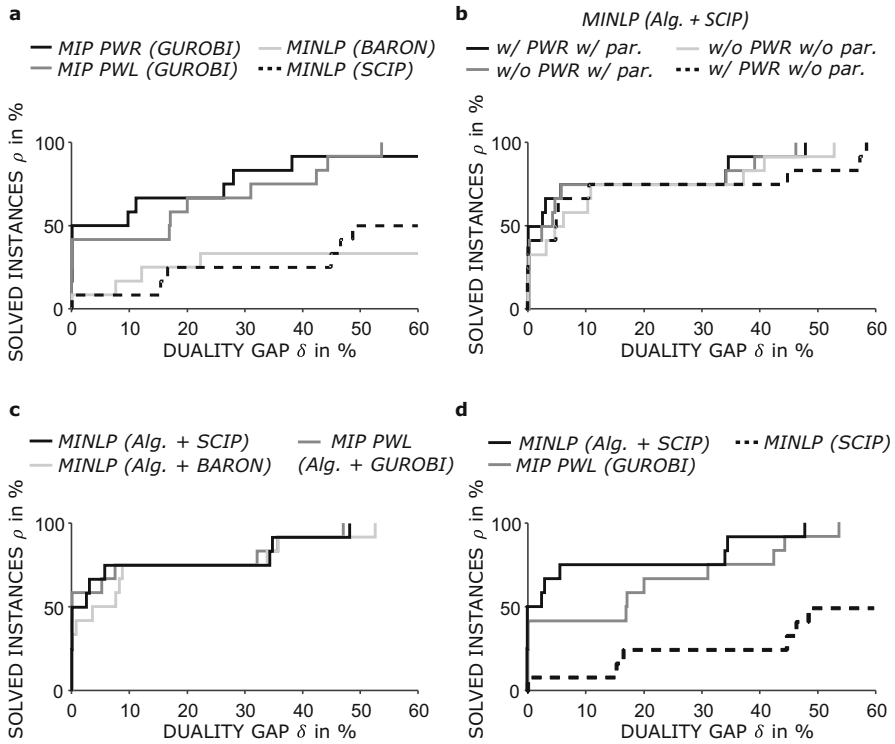
Fig. 9 Solved instances $\rho(\delta)$ with a gap smaller δ for different solution approaches. Comparison of **a** directly solved MINLP, PWL and PWR, **b** different options of solving algorithm, **c** different solvers for the solving algorithm (settings: w/ PWR, w/par.), and **d** directly solved MINLP, directly solved PWL and MINLP solved applying the solving algorithm (settings: w/ PWR, w/par.)

The comparison of the directly solved MINLP with the directly solved PWL-MILP and the PWR-MILP (cf. Fig. 9a) shows that more instances can be solved with a satisfactory gap by using one of the MILPs.

The PWL-MILP is an approximation of the MINLP. The relative deviation between the primal solution of the PWL-MILP and the best known primal MINLP solution (calculated for the MINLP with the solving algorithm) is on average 3.6 % for all 12 instances. For solutions with a gap of $\leq \varepsilon$ (4/12 instances) the deviation is 1.65 %. For 2/12 instances, the same pumps and FCs are purchased as for the best known MINLP solution. If the PWL-MILP is solved using the solving algorithm, lower gaps can be achieved (see Fig 9c). In this case, the mean deviation is 1.61 % (all instances) or 1.75 % (instances with gap $\leq \varepsilon$) and the investment decisions are the same for 5/12 instances. In comparison to the direct solution of the MINLP, more useful solutions can be achieved and the approximation quality is acceptable with around 2 % for the given discretization. However, high computation times and large duality gaps still result, especially when using an off-the-shelf solver.

The dual solution of the PWR-MILP is compared with the best found primal MINLP solution: For solutions with a gap $\leq \varepsilon$ (6/12 instances) the deviation is 7.0 % or 19.91 %

(all instances) and investment decisions are the same for 5/12 of the instances. The high deviation is due to non-converged solutions for difficult instances. In this case, the dual bound of the PWR-MILP cannot be raised sufficiently, which is why it serves only as a weak lower bound.

Figure 9 further shows, that the difference between SCIP and BARON to directly solve the MINLP is not significant. Both solvers can only solve 1/12 of the instances with a gap $\leq \varepsilon$. In addition, we occasionally encountered errors with both solvers during the solution process. For SCIP this can be reduced by a less aggressive pre-solving. A more detailed discussion can be found in Appendix A. A direct solution of the MINLP with off-the-shelf solvers is due to the large duality gap and potential wrong results an unsuitable approach. In the solving algorithm the individual problems are smaller and errors by SCIP or BARON were rare. However, for SCIP the presolving-emphasis settings for the calculation with the solving algorithm were set to "fast" to avoid any problems.

Performance of Solving Algorithm Fig. 9b shows the performance of the solving algorithm with and without the use of PWR (w/ PWR and w/o PWR) and parallelization (w/ par. and w/o par.). Without the use of PWR, the reduced MINLP is already solved within the first selection of a tuple, thus the PWR-MILP is not solved before. Without parallelization, only one tuple is solved instead of N tuples at the same time. SCIP is used as MINLP solver and GUROBI as PWR-MILP solver. All in all, the deviations of the different settings are rather small. The results show:

- The use of PWR shows only minor improvements. This could be caused by a too fine discretization, resulting in a dual bound of the PWR that can only be computed with high effort.
- The parallelization offers a slight performance increase. The reason it only yields small improvements could be that when solving N tuple at the same time, the computation of one tuple takes much longer than for the others. As long as the first N tuples are not finished, the primal and dual bounds are not updated. This holds the risk of wasting time on the computation of tuples which are cut off later anyway. In addition, parallelization causes computational overhead.

In Fig. 9c, BARON and SCIP are compared as underlying MINLP solvers for the solving algorithm (w/ PWR, w/ par.). The performance of SCIP is slightly better. In addition, the PWL-MILP is solved with the solving algorithm (GUROBI as PWL-MILP solver, w/o PWR, w/ par.). The performance is slightly below that of the MINLP. A possible cause is the computational overhead of the solving algorithm: The optimization model has to be set up for each tuple. This is more complex for the PWL model due to the increased number of equations, which can lead to a loss of performance. Nevertheless, compared to directly solving the PWL-MILP, considerably more instances can be solved (cf. Fig. 9a).

Concluding, the results for the solving algorithm (SCIP, w/ PWR, w/ par.) are compared with directly solved MINLP and PWL-MILP in Fig. 9d: the solving algorithm allows to solve 50% of the instances with a gap $\leq \varepsilon$ and all instances with a gap $\leq$ 48%. This is substantially more than directly solving the MINLP and also than directly solving the approximated PWL-MILP.
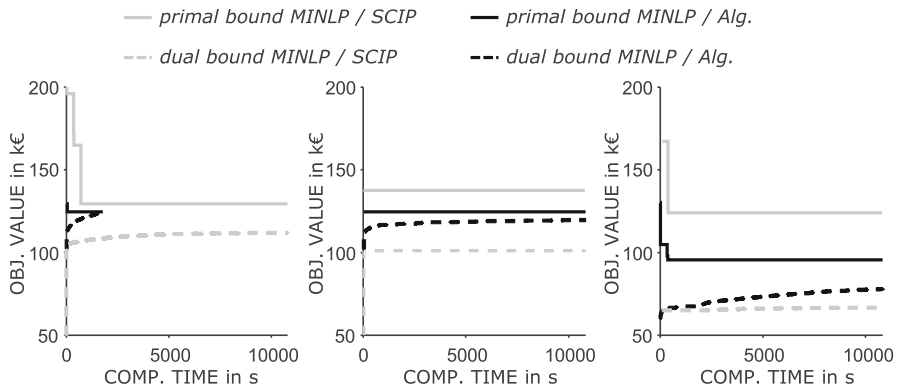
Fig. 10 Comparison of the progression of the primal and dual bound for the solving algorithm (SCIP, w/ PWR, w/ par.) and the directly solved MINLP (SCIP, presolving "fast") for the instances **a** B2.P, **b** B2.C and **c** B2.D

The progression of the dual and primal bound for the solving algorithm (SCIP, w/ PWR, w/ par.) and the directly solved MINLP (SCIP, presolving "fast") is shown in Fig. 10. The number of possible variants in the problem increases from (a) (B2.P) to (b) (B2.C) to (c) (B2.D). The solving algorithm finds the optimal primal solution very fast and at all time steps the primal bound as well as the dual bound of the solving algorithm is better than for the directly solved MINLP.

6.3 Technical discussion of the optimized booster station designs

In the following, the technical aspects of the solutions are discussed. In this section, we always refer to the best found primal solution of any solver.

Table 5 summarizes the relevant aspects of the optimization results for all instances. The following conclusions can be drawn:

- For central booster stations the restriction of the topology to solely parallel connections has no disadvantage for the tested real buildings. This may be due to the selected pump kit, since the existing pumps are designed for parallel connections. For the test rig, the central solution (unrestricted topology) is slightly superior.
- 20–25 % of the total cost can be saved with decentralized systems. The system performance increases by 50 % up to 100 %. The latter corresponds to halving the required power. This requires higher investment costs and more pumps, which is far outweighed by the energy savings.
- Although the load cases could already be covered by a single pump, always 2–4 pumps are bought due to the achieved energy savings. In some cases different pump types are used, in others (B3.P, B3.C) solely one pump type is used. Only about half of the pumps are equipped with FCs.

For building B3, the actual installed pumps are known. In the installed system four parallelly connected pumps of the same type (all with FC) are used. To be able to compare the existing system to the optimized system, the investment decision of

Table 5 Results of the best found primal solution for each instance

	building B1 parallel	central	decen.	building B2 parallel	central	decen
total cost in k€	141.7	141.7	107.4	124.4	124.4	93.9
rel. total cost in %	100	100	75.8	100	100	75.5
invest. cost in k€	32.4	32.4	34.7	30.3	30.3	32.4
# pumps	3	3	4	2	2	4
# diff. types	2	2	4	2	2	4
# FC	1	1	3	1	1	2
mean power in kW	10.9	10.9	7.2	9.3	9.3	6.1
η_{sys}^1 in %	37.8	37.8	56.9	36.6	36.7	56.1

	building B3 parallel	central	decen.	inst.	test rig TRE parallel	central	decen.
total cost in €	16377	16377	12715	24391	957	933	780
rel. total cost in %	100	100	77.6	148.9	100	97.5	81.5
invest. cost in €	6321	6321	7762	14868	276	244	226
# pumps	2	2	4	4	2	2	2
# diff. types	1	1	2	1	2	2	2
# FC	1	1	2	4	–	–	–
mean power in W	998	998	491	947	67.6	68.4	55
η_{sys}^1 in %	22.4	22.4	45.5	23.64	25.7	25.4	31.5

The system efficiency is defined as: $\eta_{sys} := \dfrac{Po_{hydr,dec.}^{mean}}{Po_{mean}} = \dfrac{\sum_{s \in \mathcal{S}} W_s \, Po_{s,hydr,dec.}}{\sum_{s \in \mathcal{S}} W_s \sum_{i \in \mathcal{P}} po_{i,s}}$

the existing system is fixed and the rotational speed optimized (see column "inst." in Table 5). In comparison, approx. 50% of the total cost can be saved by optimizing the investment costs and speed at the same time. Even if costs for an additional redundant spare pump are taken into account (2053 €, without FC), 32 % of the total cost can still be saved.

All selected pumps are highlighted in Fig. 11. Solid black lines indicate pumps, that are selected for any setting, while dashed black lines indicate pumps that are only selected for decentralized settings.

For building B1, the optimal topologies are shown in Fig. 12. The central topology (a) shows that only the largest pump is equipped with a frequency converter and operates for every load scenarios. For the load cases with higher power demand, further pumps are switched on additionally. An interpretation of this chosen topology can be that the additional pumps run for a small part of the operation time only and therefore their energy consumption and resulting costs do not weigh enough to equip either of these pumps with a frequency converter. A complex topology results for the decentralized solution (b). In scenarios with low hydraulic power demand, one pump solely supplies the pressure zones 1 and 2. Two pumps connected in series supply the upper pressure zones. After the first pump a partial volume flow is branched off to
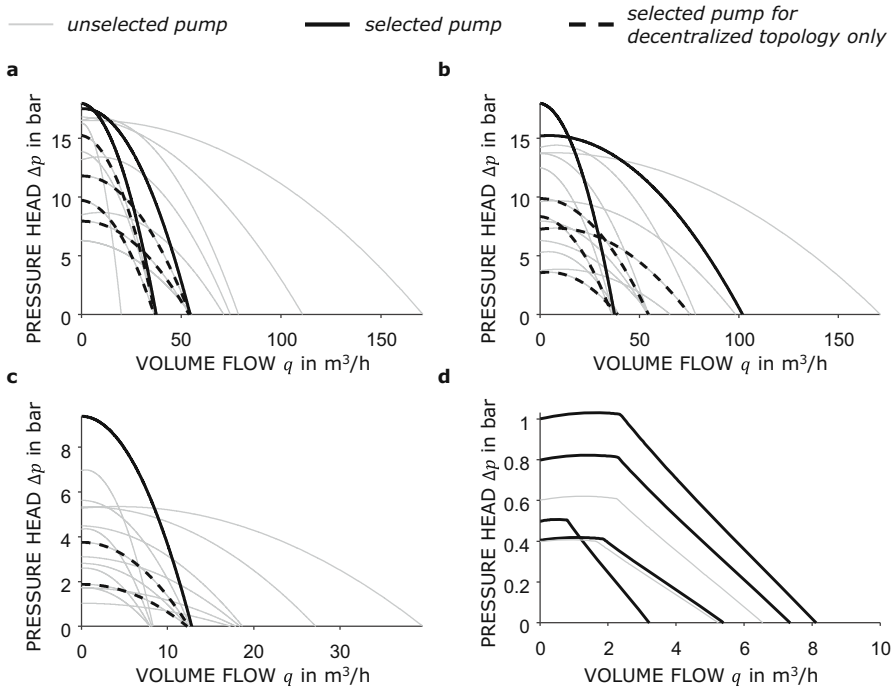
Fig. 11 Selected pumps after optimization for each building and the test rig. **a** building B1 **b** building B2 **c** building B3 **d** test rig
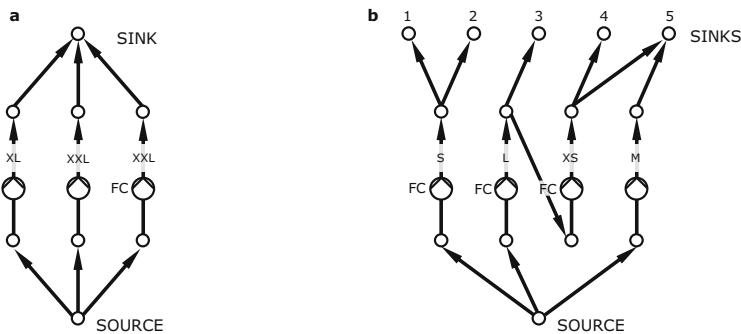
Fig. 12 Optimal topology for the instances **a** B1.P/C and **b** B1.D. The labels XS to XXL refer to the hydraulic power of the pumps at their best efficiency point

supply pressure zone 3. Only for load cases 4 and 5 a further pump is switched on, which has no frequency converter.

The trade-off between investment costs and energy costs for building B1 is shown by the Pareto front in Fig. 13. The computation was done using the ϵ-constraint method (Ehrgott 2005, pp. 98–101), in which the maximum power consumption was added as a constraint. The investment costs were weighted very high by setting the lifetime to 1 month. The optimization was performed with different upper limits on the power
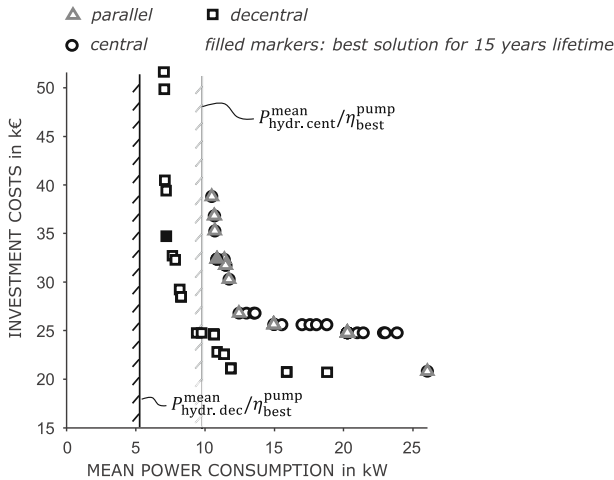
Fig. 13 Pareto Front of the objective's investment costs and power consumption for the instances B1.P, B1.C and B1.D. The filled markers represent the best solution for a lifetime of 15 years (corresponds to Table 5)

consumption. The Pareto front of the parallel and central topologies coincide exactly. As expected, the decentralized solution completely dominates either central solution. The minimum necessary power, calculated by hydraulic power and component efficiency, can almost be reached. This indicates a good choice of the pump kit and thus confirms the heuristics in Algorithm 1. The optimal solution discussed before (lifetime of 15 years), is in the range of high energy weighting. It should be noted that the computation time for high weights for the energy increases significantly, since fewer solutions can be cut off.

7 Experimental validation

In order to validate the methodology, assumptions and models, the results obtained for the test rig (Sect. 3.3) are experimentally examined. The boundary conditions applied for the optimization are obtained in two different ways. In case (i), the loss factor ζ_{loss} and the required pressure increase in the load scenarios are determined experimentally. For case (ii), a hydraulic model is developed, for which the friction coefficients of the components are estimated on the basis of data given in literature (GF Piping System 2019; Idelchik 2007). The latter case is more practical since experimental surveys cannot always be carried out in the field. In both cases, manufacturer specifications are used for the characteristic curves of the pumps and a strictly quadratic relationship between pressure losses and the volume flow is assumed. It should be noted that not all friction influences can be taken into account in the model. Depending on the topology, elbow and tee fittings are necessary to achieve the desired configuration, however, in the optimization model their friction influence is neglected. For a more detailed investigation of the friction influence, we refer to Müller et al. (2019a).
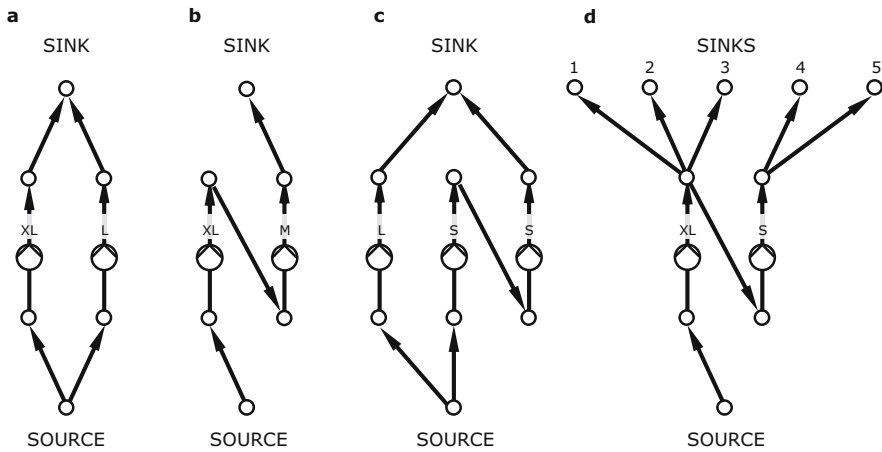
Fig. 14 Optimal topology for the test rig: **a** parallel (experimental data as well as solely literature data), **b** central (experimental data), **c** central (solely literature data) and **d** decentral (experimental data as well as solely literature data). For topology **c** the setup on the test rig is shown in Fig. 3

The optimal topologies are shown in Fig. 14. Only for the central booster station (b)/(c) there are deviations in the optimal topology between experimental and literature data. Without the use of experimental data, the pressure losses are assumed to be lower, which is why slower rotational speeds are applied.

The shown topologies are set up on the test rig, trying to keep the pressure losses due to the layout of the connections as low as possible. For every scenario the rotational speeds are applied and the valves are set so that the same volume flow is achieved on each floor. The pumps run at a constant speed, i.e. without pressure or volume flow control-loop.

In Fig. 15, the power consumption (measure for the achievement of the objective) is plotted against the volume flow (measure for the satisfaction of the load demand) for all topologies and scenarios.

The correlation between power and volume flow is well matched in computation and experiment even though there is a deviation between the required and the measured volume flows. As expected, the correlation is higher when using the experimental data.

Furthermore, the mean volume flow,

$$Q_{\text{mean}} := \sum_{s \in \mathcal{S}} W_s \sum_{f \in \mathcal{F}} q_{f,s}, \tag{17}$$

the mean power consumption,

$$Po_{\text{mean}} := \sum_{s \in \mathcal{S}} W_s \sum_{i \in \mathcal{P}} po_{i,s}, \tag{18}$$
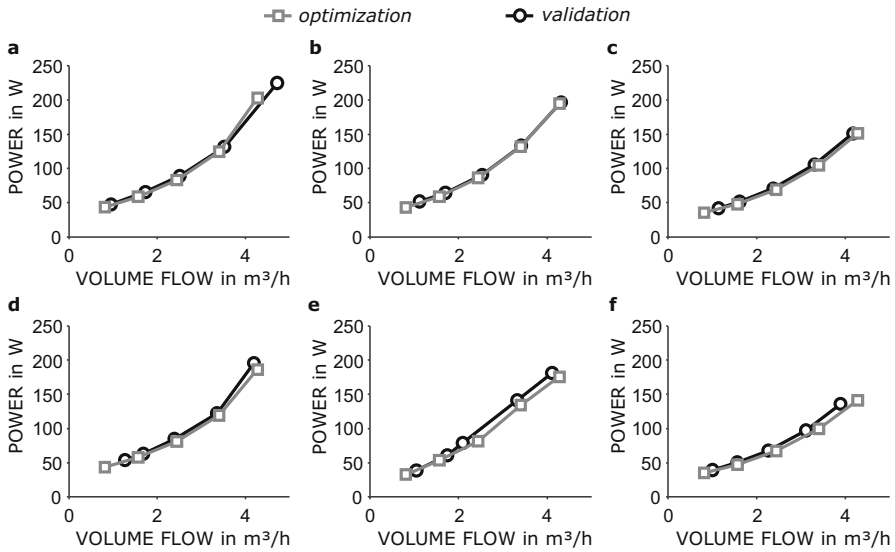
Fig. 15 Power consumption over volume flow for optimization and validation results for all five load scenarios. Results based on experimental data for **a** parallel **b** central and **c** decentral topologies, and based on literature data for **d** parallel **e** central and **f** decentral topologies are shown. The error bars are covered by the markers

and the geodetic system efficiency are investigated cf. Table 6,

$$\eta_{\text{geo}} := \frac{\sum_{s \in \mathcal{S}} W_s \varrho g \sum_{f \in \mathcal{F}} \Delta h_f q_{f,s}}{Po_{\text{mean}}}. \tag{19}$$

The values are calculated using the density $\varrho = 998 \, \text{kg m}^{-3}$, the gravitational acceleration $g = 9.81 \, \text{m s}^{-2}$, and the geodetic height difference $\Delta h_f = \{1 \, \text{m}, \dots, 5 \, \text{m}\}$ for the different pressure zones. The geodetic system efficiency η_{geo} is a measure for the deviation of the energy costs and is used to determine the expected total cost:

$$C_{\text{total}}^{\text{exp}} := C_{\text{invest}}^{\text{opt}} + C_{\text{energy}}^{\text{opt}} \left(1 + \frac{\eta_{\text{geo}}^{\text{opt}} - \eta_{\text{geo}}^{\text{exp}}}{\eta_{\text{geo}}^{\text{opt}}} \right). \tag{20}$$

The results show that the mean volume flow is higher than demanded, i.e. the pressure losses have been overestimated. As a result, the power consumption is also higher than predicted. In practice, the deviations would be compensated by a closed loop control of each pump's rotational speed. The systems are nevertheless comparable on the basis of the geodetic efficiency, which is slightly lower than assumed in the optimization. The deviation is smaller for the pump model based on the experimental data. For the instances where the pump S was installed (cf. Fig. 14b and c), the deviation is higher, which indicates inaccuracies in the used characteristics of pump S (given by the manufacturers data sheet). In spite of these uncertainties, the total costs are lower for decentralized systems. The deviations in efficiency (5-10 %) are significantly

Table 6 Comparison of optimization and validation results

		parallel	central	decentral
based on experimental data	$Q_{\text{mean}}^{\text{opt}}$ in m^{3}h^{-1}	1.732	1.732	1.732
	$Q_{\text{mean}}^{\text{exp}}$ in m^{3}h^{-1}	1.877 ± 0.027	1.895 ± 0.027	1.817 ± 0.026
	deviation in %	8.34	9.39	4.88
	$Po_{\text{mean}}^{\text{opt}}$ in W	67.578	68.448	55.043
	$Po_{\text{mean}}^{\text{exp}}$ in W	74 ± 2.36	74.16 ± 1.78	58.86 ± 1.77
	deviation in %	9.51	8.35	6.94
	$\eta_{\text{geo}}^{\text{opt}}$ in %	18.292	18.06	22.458
	$\eta_{\text{geo}}^{\text{exp}}$ in %	17.88	17.632	21.218
	deviation in %	-2.26	-2.37	-5.52
	$C_{\text{total}}^{\text{opt}}$ in €	957.3	933.7	780.6
	$C_{\text{total}}^{\text{exp}}$ in €	972.6	950	811.3
based on literature data	$Q_{\text{mean}}^{\text{opt}}$ in m^{3}h^{-1}	1.732	1.732	1.732
	$Q_{\text{mean}}^{\text{exp}}$ in m^{3}h^{-1}	1.893 ± 0.027	1.802 ± 0.026	1.715 ± 0.026
	Deviation in %	9.32	4.05	-0.98
	$Po_{\text{mean}}^{\text{opt}}$ in W	66.298	62.148	53.953
	$Po_{\text{mean}}^{\text{exp}}$ in W	72.574 ± 2.36	67.202 ± 2.36	56.318 ± 1.77
	Deviation in %	9.47	8.13	4.38
	$\eta_{\text{geo}}^{\text{opt}}$ in %	18.646	19.891	22.912
	$\eta_{\text{geo}}^{\text{exp}}$ in %	17.696	18.343	21.02
	Deviation in %	-5.1	-7.78	-8.26
	$C_{\text{total}}^{\text{opt}}$ in €	944.4	919.4	769.7
	$C_{\text{total}}^{\text{exp}}$ in €	978.4	968.1	814.5

smaller than the saving potential compared to conventional parallel design ($\approx 50\,\%$, cf. Table 5). In addition, deviations are also to be expected for the conventional design. This shows that mathematical optimization is suitable for the planning of pumping system's design and operation and moreover the level of detail realized in the models is sufficient.

8 Conclusion and outlook

We presented three different modeling approaches for the optimization of booster stations in high-rise buildings that are based on a stochastic optimization approach. The underlying optimization model is mostly linear in its decision variables, but the consideration of the pump characteristics for the booster station design leads to a nonconvex mixed-integer nonlinear program, which is hard to solve for real-world instances in its original form. To overcome this drawback, we derived two additional model variants, allowing for MILP formulations. These are a piecewise linear approximation and a

piecewise linear relaxation of the nonlinear pump characteristics. Furthermore, we implemented a problem-specific solving algorithm which is based on these different methods and additional domain-specific engineering knowledge to derive global-optimal results for real-world instances efficiently. The solving algorithm reduces the original MINLP by fixing specific pump purchase decisions and uses a selection heuristic to find feasible primal solutions in a very short time. This primal solution search is undertaken alternating with the dual solution search, which is based either on the reduced MINLP directly or on the implemented piecewise linear relaxation. The presented solving algorithm improves the solution times considerably in comparison to the original MINLP formulation. We presented the computation benefits of the solving algorithm for three different high-rise buildings as well as for a test rig. The test rig, constructed for validation purposes, represents the scaled water supply in a high-rise building. Therefore, we were able to conduct experiments for results of the optimization models considering different data sources as well as varying degrees of freedom. The comparison of optimization and experimental results on the test rig shows that the degree of detail is adequate and that the introduced method is transferable to real-world infrastructure. Allowing for different degrees of freedom for the alignment of pumps in the water supply system shows that the decentralized topology of booster stations can yield approx. 25 % of cost savings when compared to an optimized parallel pump alignment. The latter design is the current common practice. Additionally, frequency converters were installed for only approx. half of the pumps in booster stations, which is due to the trade-off of energy costs in operation and investment costs making up total cost. Compared to a conventionally designed booster station, the mathematical optimization could achieve a total cost reduction of up to approx. 50 %. These results show that we were able to derive significantly improved system designs by combining domain-specific engineering knowledge and discrete optimization methods. The developed approach assists engineers as early as in the design stage to improve the system design. This is possible, since this approach allows to consider the real load profile in the implemented two-stage stochastic optimization problem. For the future, we want to address further open research questions in this area, both in the algorithmic and engineering domain. For instance, to obtain reasonable problem sizes, we had to reduce the number of different pump types in the preprocessing with the help of an engineering knowledge applying heuristic. Therefore, improving the performance of the algorithm to consider more pumps could achieve even better system designs. This objective could be approached by an improved deployment of the underlying piecewise relaxation or by warm-starting already investigated subproblems. From an engineering perspective, we want to examine more complex, branched systems and want to look into the transfer of the methods to different fluid systems. Finally, we would like to examine the operation in more detail and ensure that the best possible performance is achieved even with load variations and deviations of the assumed parameters.

A Appendix

Within the computational evaluations we encountered problems with the two MINLP solvers SCIP (v5.0) and BARON (v.19.7.13), which we present in more detail in the following. Within modeling, we put a lot of effort in the conditioning of the underlying optimization problem to reduce errors based on an ill-conditioned optimization problem. A summary can be found in Table 7. While BARON occasionally generated no output for specific instances, SCIP either indicated larger dual bounds than the value of the true primal solutions or declared instances as infeasible while we were able to derive feasible solutions either with BARON or manually by engineering knowledge. BARON terminated without solution output for 4/12 instances for the original MINLP given by (5) and (6). SCIP output an incorrect solution for 5/12 instances in the original MINLP with default presolving settings. We were able to reduce the number of false results for SCIP by adjusting the presolving emphasis settings to "fast" or "off", as indicated in Table 7. A newer version of SCIP or BARON can potentially solve the encountered issues.

Table 7 Specification of the instances for which an error occurred in the direct solution of the MINLPs

Presolving emphasis settings	BARON Default	SCIP Default	Fast	Off
No output	B1.P, B1.D, B2.C, B2.D	–	–	B1.P
Wrongly detected infeasibility	–	B1.P, B1.D, B3.P	–	TE.D
Indicated too large dual bound	–	B3.C, B3.D	B2.P	–
# Instances with error	4/12	5/12	1/12	2/12

References

AfA table (1995) Afa table for the branch energy and water supply. (tax depreciation for wear and tear). File number IV A 8-S 1551-9/95, 45-S 1551-20

Altherr LC, Leise P, Pfetsch ME, Schmitt A (2019) Resilient layout, design and operation of energy-efficient water distribution networks for high-rise buildings using MINLP. Optimiz Eng 20(2):605–645

Brennen CE (2011) Hydrodynamics of pumps. Cambridge University Press, Cambridge

Bundesbank (2019) Discount rates in accordance with section 253 (2) HGB, 7-year average, remaining time 15 years. Time series BBK01.WX0037

Carøe CC, Tind J (1998) L-shaped decomposition of two-stage stochastic programs with integer recourse. Math Program 83(1–3):451–464

Coelho B, Andrade-Campos A (2014) Efficiency achievement in water supply systems—a review. Renew Sustain Energy Rev 30:59–84

CTBUH (2018) tall building year in review. Technical report, Council on Tall Buildings and Urban Habitat

Currie J, Wilson DI, Sahinidis N, Pinto J (2012) OPTI: Lowering the barrier between open source optimizers and the industrial MATLAB user. https://www.inverseproblem.co.nz/OPTI/. Accessed Oct 29th 2019

Destatis (2019) Daten zur Energiepreisentwicklung. Lange Reihen von Januar 2005 bis August 2019

DIN 1988-300 (2012) Codes of practice for drinking water installations—Part 300: Pipe sizing; DVGW code of practice

DIN 1988–500 (2011) Codes of practice for drinking water installations—Part 500: Pressure boosting stations with RPM-regulated pumps; DVGW code of practice

Ehrgott M (2005) Multicriteria optimization, vol 2. Springer, Berlin

EN 16480:2016 (2016) Pumps–minimum required efficiency of rotodynamic water pumps

Ferreira FJ, Fong JA, de Almeida AT (2010) Ecoanalysis of variable-speed drives for flow regulation in pumping systems. IEEE Trans Ind Electron 58(6):2117–2125

Fisher DS (2001) The impact of safety factors on pump performance and selection: Case study, a methodology for correcting pump oversizing/discussion. ASHRAE Trans 107:584

Garg VK (1998) Applied computational fluid dynamics. CRC Press, Boca Raton

Geißler B, Martin A, Morsi A, Schewe L (2012) Using piecewise linear functions for solving MINLPs. In: Lee J, Leyffer S (eds) Mixed integer nonlinear programming. Springer, New York, pp 287–314

GF Piping System (2019) JRG Sanipex MT - Technical documentation. https://www.gfps.com/appgate/ecat/common_flow/10006Q/CH/en/231565/download/document.html. Accessed Oct 29th 2019

Gleixner A, Eifler L, Gally T, Gamrath G, Gemander P, Gottwald RL, Hendel G, Hojny C, Koch T, Miltenberger M, Müller B, Pfetsch ME, Puchert C, Rehfeldt D, Schlösser F, Serrano F, Shinano Y, Viernickel JM, Vigerske S, Weninger D, Witt JT, Witzig J (2017) The SCIP Optimization Suite 5.0. Technical report, Optimization Online, http://www.optimization-online.org/DB_HTML/2017/12/6385.html. Accessed Oct 29th 2019

Gössling S, Peeters P, Hall CM, Ceron JP, Dubois G, Scott D et al (2012) Tourism and water use: supply, demand, and security. an international review. Tourism Manage 33(1):1–15

Groß T, Pöttgen P, Pelz PF (2017) Analytical approach for the optimal operation of pumps in booster systems. J Water Resour Plan Manage 143(8):04017029

Grundfos Pumps Ltd (2016) Pressure booster pumps and sets 2016 range overview. https://uk.grundfos.com/content/dam/UK/Brochure/Booster%20overview/Booster%20Pumps%20Brochure%20May%202016.pdf. Accessed March 4th 2020

Gülich JF (2010) Centrifugal pumps. Springer, Berlin

GUROBI (2020) GUROBI Optimization. http://www.gurobi.com. Accessed March 24th 2020

Hartisch M, Herbst A, Lorenz U, Weber JB (2018) Towards resilient process networks-designing booster stations via quantified programming. Appl Mech Mater Trans Tech Publ 885:199–210

Herrera M, Abraham E, Stoianov I (2016) A graph-theoretic framework for assessing the resilience of sectorised water distribution networks. Water Resour Manage 30(5):1685–1699

Hirschberg R (2014) Lastprofil und Regelkurve zur energetischen Bewertung von Druckerhöhungsanlagen (DEA). HLH Lüftung/Klima, Heizung/Sanitär, Gebäudetechnik pp 1436–5103

Humpola J, Fügenschuh A (2015) Convex reformulations for solving a nonlinear network design problem. Comput Optim Appl 62(3):717–759

IBM (2020) IBM ILOG CPLEX Optimization Studio. https://www.ibm.com/products/ilog-cplex-optimization-studio. Accessed March 24th 2020

Idelchik IE (2007) Handbook of hydraulic resistance, 4th edn. Begell House, New York

ISO/ASME 14414 (2019) Pump system energy assessment

Japikse D, Marscher WD, Furst RB (1997) Centrifugal pump design and performance. Concepts ETI Inc, Wilder

KSB SE Co KGaA (2020) Hyamat SVP pressure boosting units. https://products.ksb.com/global/products/pumps-and-pump-systems/ready-to-connect-pump-sets/pressure-boosting-units/hyamat-svp-3014. Accessed March 4th 2020

Lang S, Ludwig G, Pelz PF, Stoffel B (2013) General methodologies of determining the energy-efficiency-index of pump units in the frame of the extended product approach. In: Proceedings of the 8th international conference on energy efficiency in motor driven systems (EEMODS)

Larralde E, Ocampo R (2010a) Centrifugal pump selection process. World Pumps 2010(2):24–28

Larralde E, Ocampo R (2010) Pump selection: a real example. World Pumps 3:28–33

Leise P, Altherr LC (2018) Optimizing the design and control of decentralized water supply systems – A case-study of a hotel building. In: International conference on engineering optimization, Springer, pp 1241–1252

Löfberg J (2004) YALMIP: A toolbox for modeling and optimization in MATLAB. In: Proceedings of the CACSD conference

Meck M, Müller TM, Altherr LC, Pelz PF (2019) Designing an Efficient and Cost-Optimal Industrial Cooling System Using Mixed-Integer Nonlinear Programming. In: Operations research proceedings 2019, Springer, to appear

Mikolajková M, Saxén H, Pettersson F (2018) Linearization of an MINLP model and its application to gas distribution optimization. Energy 146:156–168

Misener R, Floudas CA (2010) Piecewise-linear approximations of multidimensional functions. J Optim Theory Appl 145(1):120–147. https://doi.org/10.1007/s10957-009-9626-0

Misener R, Floudas CA (2012) Global optimization of mixed-integer quadratically-constrained quadratic programs (MIQCQP) through piecewise-linear and edge-concave relaxations. Math Program 136(1):155–182

Misener R, Floudas CA (2013) GloMIQO: global mixed-integer quadratic optimizer. J Global Optim 57(1):3–50

Misener R, Floudas CA (2014) ANTIGONE: algorithms for continuous/integer global optimization of nonlinear equations. J Global Optim 59(2–3):503–526

Misener R, Gounaris CE, Floudas CA (2009) Global optimization of gas lifting operations: a comparative study of piecewise linear formulations. Ind Eng Chem Res 48(13):6098–6104

Morsi A, Geißler B, Martin A (2012) Mixed integer optimization of water supply networks. In: Mathematical optimization of water networks. Springer, pp 35–54

Müller TM, Altherr LC, Leise P, Pelz PF (2019a) Optimization of pumping systems for buildings: Experimental validation of different degrees of model detail on a modular test rig. In: Operations research proceedings 2019, Springer, Dresden, Germany, to appear

Müller TM, Leise P, Meck T, Altherr LC, Pelz PF (2019b) Systemic optimization of booster stations—from data collection to validation. In: International rotating equipment conference (IREC) 2019, Wiesbaden, Germany

Norgaard J, Nielsen A (2010) Water supply in tall buildings: roof tanks versus pressurised systems. Grundfos Water Boosting, Denmark

Nowak D, Krieg H, Bortz M, Geil C, Knapp A, Roclawski H, Böhle M (2018) Decision support for the design and operation of variable speed pumps in water supply systems. Water 10(6):1

Pecci F, Abraham E, Stoianov I (2019) Global optimality bounds for the placement of control valves in water supply networks. Optim Eng 20(2):457–495. https://doi.org/10.3390/w10060734

Pedersen GK, Yang Z (2008) Efficiency optimization of a multi-pump booster system. In: Proceedings of the 10th annual conference on genetic and evolutionary computation, ACM, pp 1611–1618

Popela P (2010) Stochastic programming models for engineering design problems. Eng Mech 17(5–6):351–362

Popela P, Novotnỳ J, Roupec J, Hrabec D, Olstad A (2014) Two-stage stochastic programming for engineering problems. Eng Mech 21(5):335–353

Puranik Y, Sahinidis NV (2017) Domain reduction techniques for global NLP and MINLP optimization. Constraints 22(3):338–376

Rausch L, Leise P, Ederer T, Altherr LC, Pelz PF (2016) A comparison of MILP and MINLP solver performance on the example of a drinking water supply system design problem. In: Proceedings of

the European congress on computational methods in applied sciences and engineering, Crete Island, Greece, pp 5–10

Schultz R (1992) Continuity and stability in two-stage stochastic integer programming. In: Stochastic Optimization, Springer, pp 81–92

Stoffel B (2015) Assessing the energy efficiency of pumps and pump units. Elsevier, Amsterdam

Tawarmalani M, Sahinidis NV (2005) A polyhedral branch-and-cut approach to global optimization. Math Program 103:225–249

Tindall J, Pendle J (2015) Are we significantly oversizing domestic water systems? In: CIBSE Technical Symposium 2015, 16th–17th April 2015, UCL London

Tolvanen J (2007) Life cycle energy cost savings through careful system design and pump selection. World Pumps 2007(490):34–37

Ulanicki B, Kahler J, Coulbeck B (2008) Modeling the efficiency and power characteristics of a pump group. J Water Resour Plan Manage 134(1):88–93

UN (2018) 2018 revision of world urbanization prospects. Tech. rep, United Nations

Vielma JP, Ahmed S, Nemhauser G (2010) Mixed-integer models for nonseparable piecewise-linear optimization: Unifying framework and extensions. Oper Res 58(2):303–315

Wächter A, Biegler LT (2006) On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math Program 106(1):25–57

Weber JB, Lorenz U (2017) Optimizing booster stations. In: Proceedings of the genetic and evolutionary computation conference companion, ACM, pp 1303–1310

Weber JB, Lorenz U (2018) Algorithmic system design of thermofluid systems. In: International conference on engineering optimization, Springer, pp 132–143

Wets RJB (1974) Stochastic programs with fixed recourse: the equivalent deterministic program. SIAM Rev 16(3):309–339

Wilo SE (2020) Products and expertise for pressure boosting. https://wilo.com/ie/en/Products-and-expertise/Product-Data/Pressure-boosting_32.html. Accessed March 4th 2020

Zuse Institute Berlin (2020) Sequential object-oriented simPlex. https://soplex.zib.de/. Accessed March 24th 2020