



A long-term unit commitment problem with hydrothermal coordination for economic and emission control in large-scale electricity systems

Alexander Franz¹ · Julia Rieck² · Jürgen Zimmermann¹

Received: 21 September 2018 / Accepted: 29 January 2020 / Published online: 19 February 2020
© The Author(s) 2020

Abstract

The paper describes a long-term scheduling problem for thermal power plants and energy storages. In addition, renewable energy sources are integrated by considering the residual demand. Besides the classical minimization of the production costs, emission-related costs are taken into account. Thereby, emission costs are determined by market prices for CO₂ emission certificates (i.e., using the EU emissions trading system). For the proposed unit commitment problem with hydrothermal coordination for economic and emission control, an enhanced mixed-integer linear programming model is presented. Moreover, a new heuristic approach is developed, which consists of two solution stages. The heuristic first performs an isolated dispatching of thermal plants. Then, a re-optimization stage is included in order to embed activities of energy storages into the final solution schedule. The considered approach is able to find outstanding schedules for benchmark instances with a planning horizon of up to one year. Furthermore, promising results are also obtained for large-scale real-world electricity systems. For the German electricity market, the relationship of CO₂ certificate prices and the optimal thermal dispatch is illustrated by a comprehensive sensitivity analysis.

Keywords Emission control · Hydrothermal coordination · Power generation scheduling · Unit commitment problem

✉ Julia Rieck
rieck@bwl.uni-hildesheim.de

Alexander Franz
alexander.franz@tu-clausthal.de

Jürgen Zimmermann
juergen.zimmermann@tu-clausthal.de

¹ Operations Research Group, Institute of Management and Economics, Clausthal University of Technology, 38678 Clausthal-Zellerfeld, Germany

² Operations Research Group, Institute of Business Economics and Information Systems, University of Hildesheim, 31141 Hildesheim, Germany

1 Introduction

In the electricity industry, a challenging problem deals with the economic scheduling of generating plants, where several plant-specific aspects and technical requirements have to be taken into account. System-wide constraints such as a steady equilibrium between electricity provision and consumption are of great importance. Hence, a sufficient amount of electricity to cover the customer demands at any point in time must be ensured for a prescribed planning horizon. The corresponding optimization problem is known as *unit commitment problem* (UCP); see e.g., Soliman and Mantawy (2012).

Definition 1 The UCP determines the production levels of generating plants with different characteristics over a prescribed planning horizon in order to meet the electricity demands of customers. These demands may vary over time. Possible objective functions are the minimization of operating costs or the maximization of profit (cf. Wood et al. 2013).

The problem is not only of considerable interest from a theoretical perspective, but also from a practical point of view. A solution can help to find a cost-efficient energy schedule which avoids deficits in energy supply.

Most research on the UCP focuses on short-term planning horizons (e.g., 1 day up to 1 week). In short-term, production control activities are performed, where shifts of demands or unexpected outages are included in real time (cf., e.g., Bartsch et al. 2008). In contrast, long-term problems aim at production planning, where plant performances concerning full-load hours, fuel demand, or operating costs are analyzed. In what follows, we consider long-term planning horizons (e.g., 1 year) in order to provide tactical and strategic information for supervisory institutions like, e.g., system operators and generating companies. Those institutions are then in a position to estimate fuel demands and maintenance activities as well as to determine the economic and environmental performance of the whole generation portfolio.

The increasing share of renewable energies (e.g., in Germany about 35 % of the gross electricity consumption is supplied by renewables) leads to new challenges for generating companies and system operators. Due to the volatile characteristic of renewable feed-ins, highly flexible but cost-efficient power plant operations are required. The power output for these plants must be determined in the scheduling process (cf. Definition 1). *Energy storages*, mainly hydro storages, are promising means to support the flexible requirements and to level the remaining (i.e., the residual) demand for thermal power plants. Both, storage activities and the sharp rise of the renewable feed-in lead to clean energy solutions ensuring great progress toward the European energy and climate objectives. Hence, the scheduling of generating plants should be enhanced by energy storages (and their coordination with generating plants) as well as renewable sources. Moreover, it should not only address economic goals. In addition, a minimization of the *emission* of harmful substances such as CO₂ is necessary. In this way, the trade-off between production costs and emissions, e.g., for the German electricity market, can be analyzed. The resulting

problem may be described as a UCP with *hydrothermal (HT) coordination* (i.e., the UCP-HT, cf. Wood et al. 2013) and *environmental aspects* (e.g., Raglend and Padhy 2006; Saravanan et al. 2014).

Definition 2 The UCP-HT is an UCP in which not only power plants, but also energy storages are taken into account in order to meet customer demands. In addition, environmental aspects may be considered in the objective function through a combined minimization of operating and emission costs.

Generally, an exact solution of the proposed problem can be obtained (particularly for small-scale instances) by using mixed-integer linear programming (MILP). However, practitioners require fast algorithms for long-term, large-scale, volatile, and non-convex electricity systems in order to solve various scenarios dealing with, e.g., different renewable feed-ins (cf. Viana 2004). Hence, heuristics are necessary to provide near-optimal solutions in reasonable time. Simple procedures are based on priority lists or truncated dynamic programming. In recent years, Lagrange relaxation or fix-and-optimize methods based directly on the mathematical model of the problem have become standard (e.g., Ongsakul and Petcharak 2004; Gollmer et al. 1999; Franz et al. 2019). Metaheuristics have also been developed for UCPs, but typically without pump storages (e.g., Kazarlis et al. 1996; Viana et al. 2008). A survey of solution methods for the UCP can be found, e.g., in Padhy (2004).

Within this article, we consider a heuristic approach for large-scale UCP-HT applications with environmental aspects. Our heuristic is characterized, in particular, by its fast solution finding. A two-stage hierarchical methodology is used in contrast to the often monolithic solution methods for pure UCPs (e.g., Kazarlis et al. 1996; Carrión and Arroyo 2006; Delarue et al. 2013; Morales-España et al. 2013, or the survey in Saravanan et al. 2013). The problem is divided into the optimization of thermal plants as well as the planning and coordination of hydro storages (see Franz and Zimmermann 2018b).

The remainder of this paper is organized as follows. In Sect. 2, we propose a mixed-integer linear programming formulation for the UCP-HT in order to describe the problem structure. Moreover, an approach for the extension toward emission minimization is treated. The novel two-stage heuristic approach for the UCP-HT with and without emission costs is explained in detail in Sect. 3. Section 4 is devoted to the results of an experimental performance analysis. In addition, a sophisticated case study for the German electricity market is presented. Finally, conclusions are given in Sect. 5.

2 Specification of the problem

The aim of planning and scheduling thermal, renewable, and storage units is to determine the operation status of each unit in order to meet a given demand over a prescribed planning horizon at minimum total production costs (e.g., Wood et al. 2013). Due to substantially low marginal costs (compared to thermal electricity

provision), renewable feed-ins are typically prioritized (e.g., Delarue et al. 2013).¹ Hence, without any loss of generality, we consider the residual demand instead of the real customer demand while solving the problem at hand.

Definition 3 The residual demand is the energy demand minus the volatile feed-in of renewables.

As a consequence, only thermal plants and energy storages have to be tackled. In Sect. 2.1, a mixed-integer linear model for the UCP–HT is presented, where the economic goal of minimizing the operating costs is considered. In Sect. 2.2, the objective function is extended by environmental aspects.

2.1 Unit commitment problem with hydrothermal coordination

The basic idea of our model formulation comes from Carrión and Arroyo (2006), Ostrowski et al. (2012) and Morales-España et al. (2013), who formulated the problem without energy storages. Particularly, the latter model is characterized as tight, due to a relatively small integrality gap, i.e., the LP relaxation is close to the convex hull of the feasible integer solutions. Therefore, we extended this tight model in order to manage the coordination between thermal and storage units (cf. Franz et al. 2019, where also explanations for the tightness of the described model are given). Furthermore, a reformulation and an adjustment of constraints (i.e., constraints for start-up processes and reserve capacity) are performed so that the resulting model provides near-optimal solutions significantly faster than the model formulation presented by Morales-España et al. (2013).

Let \mathcal{I} be the set of thermal plants and \mathcal{J} be the set of energy storages. For each thermal plant $i \in \mathcal{I}$, we define the continuous power generation level p_{it} as well as the binary on/off status u_{it} at each point in time $t \in \mathcal{T}$, where \mathcal{T} represents the set of hourly time periods ($\mathcal{T} := \{1, 2, \dots, T\}$). The power generation level p_{it} denotes the power level above the minimum power output \underline{P}_i . For each energy storage $j \in \mathcal{J}$, we specify the continuous power generation level p_{jt}^g as well as the continuous pumping level p_{jt}^p . We assume that $p \geq 0$ is satisfied for all power-related decision variables.

The objective function of the problem is firstly chosen as the minimization of total operating costs, which covers thermal production costs, thermal start-up costs, and non-served energy costs. The thermal production cost for each unit $i \in \mathcal{I}$ and point in time $t \in \mathcal{T}$ is typically modeled using the convex heat consumption function $Q(\cdot)$ for the overall power output $p'_{it} = p_{it} + \underline{P}_i$, e.g., as

$$c_i^{\text{pc}} Q_{it}(p'_{it}, u_{it}) \quad \text{with} \quad Q_{it}(p'_{it}, u_{it}) = \begin{cases} q_i^a + q_i^b p'_{it} + q_i^c p'_{it}{}^2, & \text{if } u_{it} = 1 \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

¹ Furthermore, the renewable energy law (EEG) regulates the priority feed-in of electricity from renewable energy sources in Germany.

c_i^{pc} represents a unit-specific fuel cost coefficient and q_i^a, q_i^b as well as q_i^c are nonnegative heat coefficients relating to the respective technical efficiency of unit i . Please note that no further cost additives (e.g., for maintenance, lubricants etc.) are taken into account.

Function (1) can be linearized, where fixed operation costs c^{fix} and variable production costs c^{var} in accordance to the power output are used. A sophisticated linearization approach is, e.g., presented by Carrión and Arroyo (2006). In our study, we set $q_i^c := 0$ for all plants $i \in \mathcal{I}$, similar to Gollmer et al. (2000) and Martinez Diaz (2008), since the curvature of function (1) is typically quite small. As costs for generating the minimum power level can be integrated in the fixed costs (they always occur if a generator is running), function (1) simplify to

$$c_i^{fix} u_{it} + c_i^{var} p_{it}, \tag{2}$$

where variable costs c_i^{var} have to be considered for each megawatt p_{it} that is delivered above the minimum power level \underline{P}_i [cf. constraints (6)]. Therefore, the variable and fixed cost components can be defined as $c_i^{var} := c_i^{pc} q_i^b$ and $c_i^{fix} := c_i^{pc} (q_i^a + q_i^b \underline{P}_i)$, respectively.

The thermal start-up costs usually vary between warm and cold starts (cold starts of $i \in \mathcal{I}$ are necessary after a downtime of at least β_i time periods). We introduce binary decision variables indicating whether or not plant i is started ($u_{it}^s = 1$) and performs a cold start ($u_{it}^{cs} = 1$) in $t \in \mathcal{T}$. Then, start-up costs may be formulated as $c_i^{ws} u_{it}^s + c_i^A u_{it}^{cs}$, where c_i^{ws} indicates warm start costs and c_i^A represents additional costs for a cold start. Both cost components result from reheating the plant's boiler. An extension to continuously modeled start-up costs using an exponential heat consumption approach instead of the two-piece linearized one can be found in Wood et al. (2013).

A feasible solution must ensure that the net energy provision delivered by thermal plants and storages is equal to the (residual) demand $D_t, t \in \mathcal{T}$ [cf. constrains (4)]. In order to relax the demand balance conditions, a demand overfulfillment is permitted but penalized by higher production costs. Furthermore, non-served energy (e.g., in case of undercapacity of all units) is allowed by introducing nonnegative slack variables n_t . Particularly for long-term planning horizons, $n_t > 0$ can be allowed for some t , since the resulting schedule is used for production planning purposes and not for real-time production control activities. The non-served energy is penalized in the objective function by $c^{non} n_t$, with a penalty cost factor c^{non} (its value must be higher than the most expensive plant). Due to this trick, a first feasible solution (with $n_t = D_t + R_t$) can be found immediately by a standard solver like CPLEX or Gurobi that will have a positive impact on the computation time.

The entire UCP with hydrothermal coordination may now be formulated as follows

$$\text{Min.} \quad \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} (c_i^{fix} u_{it} + c_i^{var} p_{it} + c_i^{ws} u_{it}^s + c_i^A u_{it}^{cs}) + \sum_{t \in \mathcal{T}} c^{non} n_t \tag{3}$$

$$\text{s.t. } \sum_{i \in \mathcal{I}} (\underline{P}_i u_{it} + p_{it}) + \sum_{j \in \mathcal{J}} (p_{jt}^g - p_{jt}^p) + n_t \geq D_t \quad (t \in \mathcal{T}) \tag{4}$$

$$\sum_{i \in \mathcal{I}} \bar{P}_i u_{it} + \sum_{j \in \mathcal{J}} \bar{P}_j^g + n_t \geq R_t + D_t \quad (t \in \mathcal{T}) \tag{5}$$

$$p_{it} \leq (\bar{P}_i - \underline{P}_i) u_{it} \quad (t \in \mathcal{T}, i \in \mathcal{I}) \tag{6}$$

$$\sum_{\tau=t-\pi_i+1}^t u_{i\tau}^s \leq u_{it} \quad (i \in \mathcal{I}, \pi_i \leq t \leq T) \tag{7}$$

$$\sum_{\tau=t-\delta_i+1}^t u_{i\tau}^d \leq 1 - u_{it} \quad (i \in \mathcal{I}, \delta_i \leq t \leq T) \tag{8}$$

$$u_{it} - u_{i,t-1} = u_{it}^s - u_{it}^d \quad (i \in \mathcal{I}, 2 \leq t \leq T) \tag{9}$$

$$u_{it}^{cs} + \sum_{\tau=1}^{\beta_i-1} u_{i,t-\tau}^d \geq u_{it}^s \quad (i \in \mathcal{I}, \beta_i \leq t \leq T) \tag{10}$$

$$e_{jt} = e_{j,t-1} + p_{jt}^p n_j^p - p_{jt}^g / \eta_j^g \quad (j \in \mathcal{J}, 2 \leq t \leq T) \tag{11}$$

$$\underline{E}_j \leq e_{jt} \leq \bar{E}_j \quad (j \in \mathcal{J}, t \in \mathcal{T}) \tag{12}$$

$$p_{jt}^g \leq \bar{P}_j^g; p_{jt}^p \leq \bar{P}_j^p \quad (j \in \mathcal{J}, t \in \mathcal{T}) \tag{13}$$

$$e_{jt}, n_t, p_{it}, p_{jt}^g, p_{jt}^p \geq 0 \quad (i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}) \tag{14}$$

$$u_{it}, u_{it}^{cs}, u_{it}^d, u_{it}^s \in \{0, 1\} \quad (i \in \mathcal{I}, t \in \mathcal{T}). \tag{15}$$

In objective function (3), total operating costs are minimized. The fulfillment of energy demands either by thermal or by hydro storages is ensured with system constraints (4). Let \bar{P}_i be the maximum power level of thermal plant $i \in \mathcal{I}$ and let \bar{P}_j^g be the maximum hydropower output by turbining at storage j . Then, system constraints (5) provide a sufficient spinning reserve (in the event of an unpredictable failure) for each point in time. Constraints (6) ensure that power outputs can only be determined between a minimum and a maximum level. Once a thermal plant $i \in \mathcal{I}$ is started (decommitted), it has to be online (offline) for at least its minimum uptime π_i (minimum downtime δ_i), which is guaranteed by inequalities (7) and (8). Binary decision

variables u_{it}^d indicate whether a plant i is decommitted in period t or not. Then, logical constraints (9) and (10) guarantee that all binary variables $u \in \{0, 1\}$ take appropriate values. For hydro storages, energy flow conservation constraints (11) are to be considered to determine the amount of energy e_{jt} maintained in each storage j in period t . The maintained energy e_{jt} corresponds to the maintained energy in period $t - 1$ plus the energy obtained by pumping and minus the power output for turbinning, where respective efficiencies η_j^p and η_j^g are taken into account. Inequalities (12) refer to a lower \underline{E}_j and an upper level \overline{E}_j in energy storage $j \in \mathcal{J}$. Moreover, inequalities (13) restrict the technical limits of generation and pumping levels. Finally, constraints (14) and (15) define feasible domains for all continuous and binary decision variables. For the sake of simplicity, constraints affecting the initial operating status are omitted. However, they can easily be added (cf., e.g., Morales-España et al. 2013) and are considered within our performance analysis (cf. Sect. 4).

2.2 Unit commitment problem with production and emission costs

A UCP-HT which addresses a minimization of costs and emissions can be formulated using the principle of the *European Union Emissions Trading System* (EU ETS). The EU ETS was established in 2005 to manage a systematic reduction in greenhouse gases. Particularly, it aims at limiting the overall CO₂-production (cap principle).² Therefore, each CO₂-emitting plant (i.e., thermal power plants for the electricity industry) has to buy allowances for the emission of each produced ton of CO₂ by bidding on an auction infrastructure (trade principle). The resulting “cap-and-trade” mechanism provides time-dependent market clearing prices $c_t^{\text{CO}_2} [\frac{\text{€}}{\text{t CO}_2}]$, $t \in \mathcal{T}$, for emission certificates (cf., e.g., Ellermann et al. 2010 or Kruger and Pizer 2010).

In order to enhance objective function (3) of the UCP-HT by environmental aspects, CO₂-related costs have to be added. These costs comprise emission costs for steady-state operations

$$c_t^{\text{CO}_2} \epsilon_i Q_{it}(p_{it}, u_{it}) \quad (t \in \mathcal{T}, i \in \mathcal{I}) \tag{16}$$

and emission costs caused by start-up procedures

$$c_t^{\text{CO}_2} \epsilon_i (q_i^{\text{ws}} u_{it}^s + q_i^A u_{it}^{\text{cs}}) \quad (t \in \mathcal{T}, i \in \mathcal{I}), \tag{17}$$

where ϵ_i represents the carbon dioxide content $[\frac{\text{t CO}_2}{\text{MWh}_{\text{th}}}]$ of each produced mega watt by power plant i . Please note that the emission factor ϵ_i is mainly influenced by the fuel (e.g., lignite, hard coal, gas), its quality and plant parameters. As in Sect. 2.1, we again choose $q_i^c := 0$ for all plants $i \in \mathcal{I}$ within heat consumption function $Q(\cdot)$. Consequently, expression (16) simplifies to $c_{it}^{\text{CO}_2, \text{fix}} \epsilon_i u_{it} + c_{it}^{\text{CO}_2, \text{var}} \epsilon_i p_{it}$, similar to expression (2).

² For prior and current U.S. efforts in emission trading see, e.g., Tietenberg (2010).

The resulting objective function

$$(3) + (17) + \sum_{i \in \mathcal{I}} \sum_{i \in \mathcal{I}} c_{it}^{\text{CO}_2, \text{fix}} \epsilon_i u_{it} + c_{it}^{\text{CO}_2, \text{var}} \epsilon_i p_{it} \quad (18)$$

of the UCP–HT with economic and emission aspects contains two conflicting targets. Typically, gas-fired (e.g., combined cycle) power plants have high production costs and low CO_2 emissions (since the emission factor ϵ_i is small), whereas lignite fired power plants offer low production costs, but high CO_2 emissions. Therefore, a highly cost-efficient dispatch increases the usage of emission-intensive power plant activities. For a more sustainable and environmentally friendly electricity provision, gas-fired plants should be preferred. The market price $c_t^{\text{CO}_2}$ for emission certificates in expressions (16) and (17) weights the two preferences, where $c_t^{\text{CO}_2}$ is exogenously given by the EU ETS and hence driven by market processes.

3 Two-stage heuristic approach

The problem of finding an optimal production schedule for all involved thermal, renewable, and storage units represents an \mathcal{NP} -hard optimization problem, which means that no polynomial time algorithm exists that will generally solve instances of the problem in reasonable time (cf., e.g., Garey and Johnson 1979 and for general UCPs, see, e.g., Tseng 1996 and Bendotti et al. 2017). However, heuristic algorithms provide valuable alternatives. In order to find a feasible as well as near-optimal solution for our considered UCP–HT quickly, we present a two-stage heuristic approach. Please note that a two-stage approach for the problem variant without explicit emission costs has already been outlined in previous conference proceedings (Franz et al. 2017; Franz and Zimmermann 2018a, b). These basics were fine-tuned and provided with further performance increasing features (e.g., smart choice of solution representation, Fig. 3, as well as the concept of shift demand step, Fig. 5). In the paper under consideration, a detailed description of the resulting algorithm is presented for the first time. Accordingly to our explanations in Sect. 2.2, we consider the objective function (18), where economic and emission aspects are involved.

The starting point of our approach is the fact that an isolated scheduling of all thermal plants $i \in \mathcal{I}$ generally provides a possibility to fulfill the energy demands described in constraints (4) and the spinning reserve in constraints (5). To be more precise, a feasible solution for the problem under consideration can be found without using any energy storages if the overall thermal capacity is sufficient, i.e.,

$$\sum_{i \in \mathcal{I}} \bar{P}_i \geq \max_{i \in \mathcal{I}} (D_t + R_t) \quad (19)$$

is satisfied. Once a feasible solution is found by dispatching thermal plants, the solution can be enhanced using energy storages in order to reduce the total costs in objective function (3). These thoughts lead to the basis of our *two-stage heuristic approach*, which is illustrated in Fig. 1.

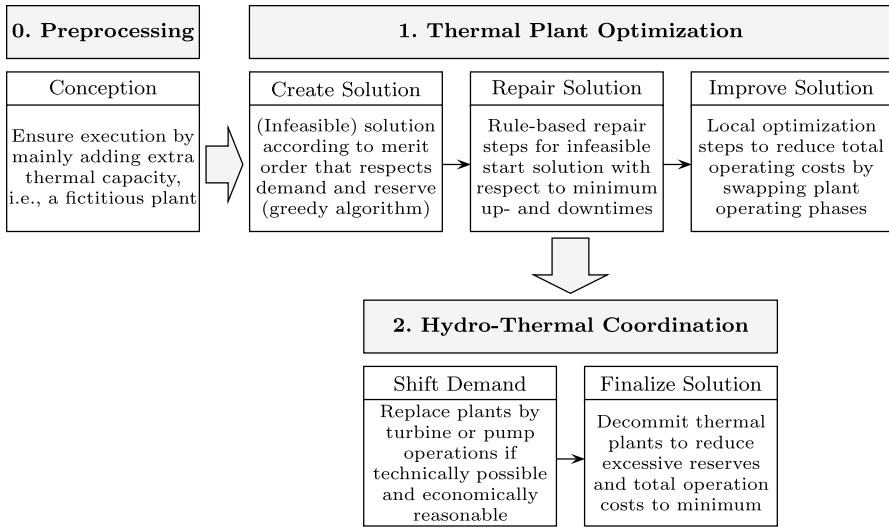


Fig. 1 Concept of the two-stage heuristic approach for the UCP-HT

In the **first stage** (cf. Sect. 3.1), a *thermal plant optimization* is performed and in the **second stage** (cf. Sect. 3.2) energy storages are considered, i.e., the *hydro-thermal coordination* is implemented. In addition, a *preprocessing step* is included (prior to the first stage) to ensure the following conditions required for a successful execution of the algorithm. Particularly, in case of missing total thermal capacity, of high gradients of residual demands, or of negative residual demands (i.e., in situations of high renewable feed-in), the preprocessing is necessary (cf. Fig. 1). In order to avoid thermal undercapacity (if inequality (19) is not fulfilled), a fictitious plant with high operating costs (significantly higher than those of all other units) is added to set \mathcal{I} . For treating, e.g., high demand gradients, a highly flexible fictitious plant must be introduced. Furthermore, in case of time intervals with negative residual demands, wind and solar infeed have to be throttled and consequently limited (e.g., to zero or to the overall capacity of the pumped storages). It must be ensured that at the end of the algorithm the inserted fictitious plant has to be eliminated (cf. “finalize solution” step, p. 16).

3.1 Thermal plant optimization: first stage of the heuristic approach

The *thermal plant optimization* preselects certain plants to fulfill the fluctuating residual demand as well as the spinning reserve requirements without using the energy storage operations of turbinning or pumping. In detail, the following three steps are executed within the **first stage** (cf. Fig. 1):

Create start solution An initial solution is obtained by applying a greedy algorithm, assuming $\delta_i = \pi_i = 1, i \in \mathcal{I}$. Thereby, all plants are sorted and ranked according to non-decreasing marginal costs (i.e., to their merit order). Then, for each point in time $t \in \mathcal{T}$, a subset of plants is committed, i.e., plants are stepwise activated

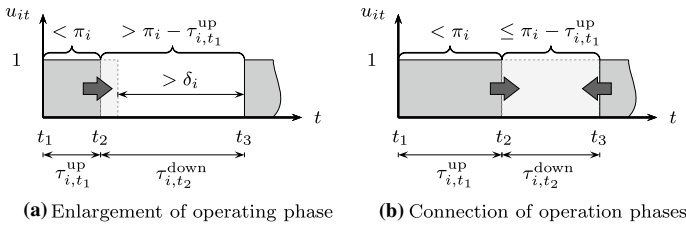


Fig. 2 Repairing the solution schedule in terms of minimum uptimes π_i

in the sequence of their merit order until the demand and reserve requirements $\sum_{i \in \mathcal{I}} \bar{P}_i u_{it} \geq D_t + R_t$ are satisfied. Some plants might be skipped in the considered order to fulfill the following inequalities $\sum_{i \in \mathcal{I}} P_i u_{it} \leq D_t$ which are necessary due to the power output specifications of all plants i . Within this first step, the power outputs p_{it} of all committed plants $i \in \mathcal{I}$ (with $u_{it} = 1, t \in \mathcal{T}$) are always set to the maximum value $\bar{P}_i - P_i$, except for the last committed plant i' . This plant is typically characterized by high marginal costs and has to be operated at part-load to balance the demand. In the event that the part-load operation of plant i' is not sufficient for demand balancing, the power of the second last committed plant must also be reduced, and so on.

Repair infeasible solution Typically, the initial solution of the first step is infeasible and needs to be repaired with respect to the minimum up- and downtime requirements. In general, a fulfillment of minimum up- and downtime requirements of a unit $i \in \mathcal{I}$ can be achieved by enlarging the corresponding *operating phases* or *idle phases* until constraints (7) and (8) are satisfied. This procedure is similar to that one in Delarue et al. (2013). An operating phase indicates that plant i is active ($u_{it} = 1, t \in \mathcal{T}$), whereas *idle phases* denote that plant i is inactive ($u_{it} = 0, t \in \mathcal{T}$).

Let τ_{i,t_1}^{up} be the duration of an operating phase of some unit $i \in \mathcal{I}$ that starts at uptime t_1 . Moreover, let τ_{i,t_2}^{down} be the duration of the following idle phase that starts at time t_2 (cf. Fig. 2). In the case of a violation of the minimum uptime requirements [cf. constraints (7)], the duration τ_{i,t_1}^{up} may be extended to the right or left as long as the time boundaries of \mathcal{T} are observed. We first consider the extension of τ_{i,t_1}^{up} to the right (cf. Fig. 2a). This extension may only be performed if

- (i) the additional uptime $\pi_i - \tau_{i,t_1}^{up} > 0$ needed to fulfill constraints (7) is smaller than τ_{i,t_2}^{down} and
- (ii) the remaining downtime duration $\tau_{i,t_2}^{down} - (\pi_i - \tau_{i,t_1}^{up})$ satisfies the minimum downtime δ_i .

In the event that the extension to the right is not sufficient to fulfill constraints (7), we extend the operating phase according to items (i) and (ii) as much as possible. Additionally, the repair procedure is carried out to the left. Thereby, the downtime duration of the previous idle phase must be taken into account [similar to (i) and (ii)]. Furthermore, if both attempts fail (and thus are not realized), the operating

phase between t_1 and t_2 is connected to the successive operating phase, that starts at time t_3 (i.e., the idle phase is eliminated, cf. Fig. 2b). In case the minimum uptime is still not fulfilled, the already extended operating phase is merged with the previous operating phase. This procedure will finally lead to an operating phase that is at least equal to the required minimum uptime.

Similar repair steps have to be executed for violations against the minimum downtime δ_i . In contrast to the repair procedure described above, an enlargement of an idle phase is only possible if constraints (4) and (5) are still satisfied. Otherwise, the idle phase must be eliminated.

Once a solution schedule is repaired, the power output of all active plants has to be adjusted. For this purpose, the plants' power outputs are set according to the merit order. This means that the most expensive, active plants are operated at part-load again (cf. the first step of our heuristic). However, it might happen that the demand is overfulfilled [possible due to constraints (4)] and redundant plants exist which are decommitted within the next steps.

To economize plant operations in terms of start-up and fixed operation costs, quite short operating and idle phases are canceled, before the repair step is considered. In order to identify such a short operating or idle phase, it is reasonable to use the minimum uptime or downtime as a reference value, i.e., $\tau_{it}^{\text{up}} \leq \xi \cdot \pi_i$ and $\tau_{it}^{\text{down}} \leq \xi \cdot \delta_i$. Please note that a cancelation applies for the whole algorithm. A canceled operating phase can result in an energy shortage and a canceled idle phase in an energy overfulfillment. In the case of an energy shortage, additional plants (e.g., flexible peak plants) must be committed according to their minimum uptime and downtime requirements. In general, this action is significantly cheaper than a comprehensive enlargement of the current operating phase (which would be necessary while executing the repair step). In preliminary studies, $\xi := 0.2$ produces the best results.

Improve feasible solution Further improvements with respect to the objective function value of the constructed feasible solution can be achieved by local search-based algorithms. The basic idea here is to swap operating phases of one plant by a set of operating phases of other plants (i.e., to search in a neighborhood of the current solution). Thereby, a possible overfulfillment of demand caused by our repair step may be reduced which results in lower production and emission costs. Within the improvement step, the feasibility of the solution is always maintained.

Let S^{thermal} be the current solution that considers thermal plants. The schedule is encoded in a binary $|\mathcal{I}| \times |\mathcal{T}|$ -matrix, where for each plant $i \in \mathcal{I}$ and point in time $t \in \mathcal{T}$ the status variables u_{it} are included (cf. Fig. 3). All other relevant information like power outputs p_{it} and the amount of time periods operated in the current operating status (particularly τ_{it}^{up} , cf. Fig. 2) can be calculated on the basis of the matrix entries (cf. "start solution" step).

For applying a neighborhood movement, a unit $i \in \mathcal{I}$ (iteratively chosen in accordance to the merit order) and the first suitable operating phase of this unit are selected. Please note that it is generally not advantageous (in terms of solution time and quality) to swap plants with long operating phases. Therefore, we denote a plant with an operating phase that is smaller than a multiple ζ of the minimum uptime as suitable for swapping. For our purposes, $\zeta := 3$ performs best. If

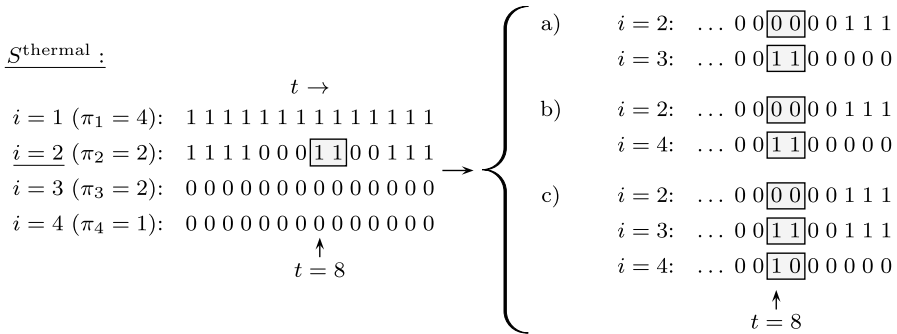


Fig. 3 Solution representation and possible neighborhood movements

the operating phase of plant i is redundant (i.e., the demand can be fulfilled even without i), then, the operating phase is decommitted by updating the status variables within schedule S^{thermal} . Otherwise, unit i is temporarily decommitted, where the decommitment of i has to be compensated by operating phases of one or several plants $i' \in \mathcal{I} \subset \mathcal{I} \setminus \{i\}$. Generally, several subsets \mathcal{I}' can be chosen that result in different objective function values determined by objective function (3). In order to limit the number of subsets, we generate all subsets \mathcal{I}' with a prescribed maximal cardinality k (e.g., $k \leq 3$) consisting of units i' , where

- i' is positioned after i in the merit order and
- i' is decommitted so far.

When all subsets are created, a subset \mathcal{I}' is further investigated in our neighborhood search if constraints (4) are satisfied without slack as well as constraints (5), (7), and (8) are fulfilled. Due to a best improve hill climbing approach, the (feasible) neighbor solution with the best objective function value replaces the current solution (in case it is better). Once the analysis of the current operating phase of plant i is completed, S^{thermal} is updated and the next suitable operating phase is selected.

Figure 3 shows the solution representation of a schedule S^{thermal} on the left-hand side. Moreover, the neighborhood movement for unit $i = 2$ is demonstrated, where the second operating phase starting at $t = 8$ is considered. We assume that the chosen operating phase is not redundant. Therefore, unit i is decommitted and subsets $\mathcal{I}' = \{3\}$, $\mathcal{I}' = \{4\}$, $\mathcal{I}' = \{3, 4\}$ are investigated. Each of them leads to a fulfillment of the demand. The relevant excerpts of the three neighborhoods (a)–(c) are given on the right-hand side of Fig. 3. Please note that the current schedule S^{thermal} is feasible in our simple example and contains no overfulfillment. Therefore, the demand requires a commitment of at least two units in $t \in \{8, 9\}$. Hence, the duration of the new operating phase of unit $i = 4$ in neighbor (b) is two periods (although $\pi_4 = 1$). In contrast, the operating phase of unit $i = 4$ in neighbor (c) equals the minimum uptime, as unit $i = 3$ fully covers the demand in $t \in \{8, 9\}$. Since neighbor (c) results in an overfulfillment, the objective function values of (a) and (b) are always

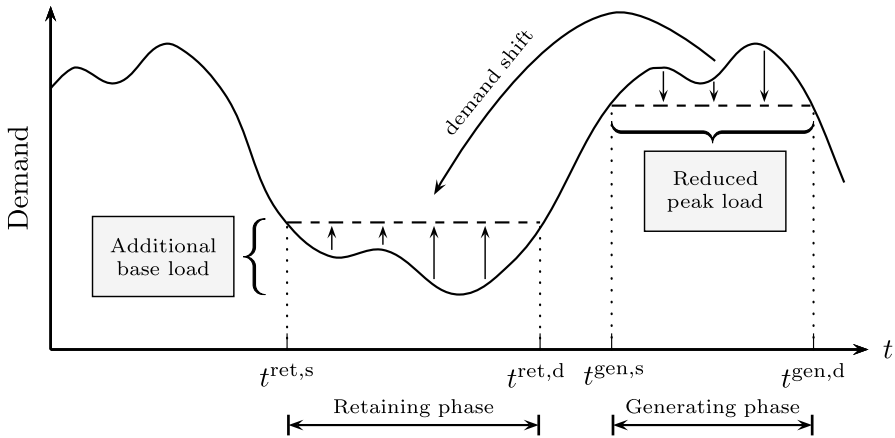


Fig. 4 Demand shift due to generating and retaining

better. The solution with the best objective function values is considered as a new starting point for further neighborhood search steps.

3.2 Hydrothermal coordination: second stage of the heuristic approach

The second stage in Fig. 1 improves the solution S^{thermal} determined in the first stage by implementing the “hydrothermal coordination” problem for pumped storages. Besides S^{thermal} , a second schedule S^{hydro} for the hydro storage activities is introduced (i.e., generating electricity or retaining energy back to the storages). Analogous to the thermal schedule, S^{hydro} is encoded in a $|\mathcal{J}| \times |\mathcal{T}|$ -matrix and contains information about energy levels e_{jt} for all $j \in \mathcal{J}$ and $t \in \mathcal{T}$. Please note that the corresponding power levels p_{jt}^g and p_{jt}^p can always be derived using constraints (11). In detail, the following two steps are executed in the **second stage**:

Shift demand using energy storages Energy storages represent a way to decouple the interdependency of demand and supply by shifting demands of one or several periods to other points in time. They can be operated in the state of generating energy due to turbining (*generating phase*), retaining energy due to pumping (*retaining phase*), and *maintaining energy* in the storage (cf. Fig. 4).

In our demand shift step, an allocation of energy storages is found by stepwise replacing operating phases of thermal plants by storage operations, i.e., by generating and retaining phases in order to achieve cost reductions. Typically, generating phases (starting at $t^{\text{gen,s}}$ and ending at $t^{\text{gen,d}}$) can be identified in high energy demand intervals and retaining phases (starting at $t^{\text{ret,s}}$ and ending at $t^{\text{ret,d}}$) in low energy demand intervals (cf. Fig. 4). Therefore, our basic idea is to search for demand maxima and minima. In a demand maximum, possible candidate plants for a decommitment are positioned (i.e., candidates for generating phases). Moreover, the already committed plants in a demand minimum may run at higher-power levels, i.e., these are possible candidates for retaining energy (an additional

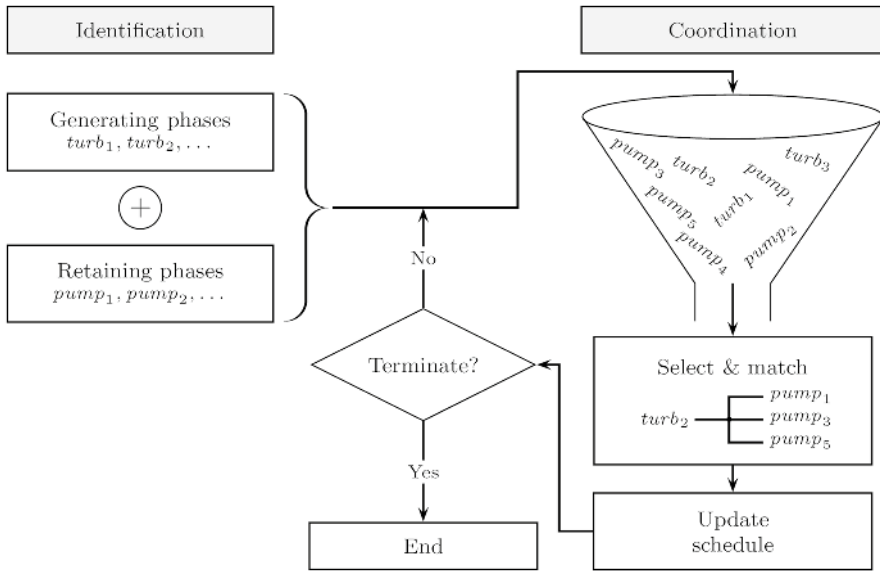


Fig. 5 Concept of the shift demand step

commitment of plants is usually connected with high fixed costs and therefore not suitable).

The principle of our algorithm is demonstrated in Fig. 5, where we select candidates of generating and retaining phases (referred to as *turb* and *pump*) in the *identification* part and suitably match them in the *coordination* part.

In the identification part, we move along the time-axis and search for situations in which $D_{t-1} < D_t \geq D_{t+1}$ (generating phase) or $D_{t-1} > D_t \leq D_{t+1}$ (retaining phase) hold. Once such a point in time $t \in \mathcal{T}$ is identified, a differentiated procedure is required for demand maxima and minima.

For a demand maximum (e.g., at time t^{\max}), all at $t^{\max} \in \mathcal{T}$ committed plants have to be taken into account. An operating phase of a committed plant $i \in \mathcal{I}$ (i.e., a phase between $t^{\text{gen,s}}$ and $t^{\text{gen,d}}$, see Fig. 4) is identified as a candidate for a generating phase if

- cost savings may occur (i.e., total costs of operating phase of i are larger than additional costs for operating the cheapest thermal plant in order to refill the hydro storage with the highest efficiency, assuming that this storage is used to compensate the operating phase of i),
- the power level of i during its whole operating phase is lower or equal to the total generating capacity of all hydro storages,
- the duration $t^{\text{gen,d}} - t^{\text{gen,s}} + 1$ of the generating phase is limited to a threshold τ^{\max} (e.g., $\tau^{\max} := 12$ h), since hydro storages are typically operated only hours not days.

For a demand minimum (e.g., at time t^{\min}), all at $t^{\min} \in \mathcal{T}$ committed plants have to be taken into account. A committed plant $i \in \mathcal{I}$ is identified as a candidate for a retaining phase if the power level $p_{i,t^{\min}}$ in t^{\min} is lower than \bar{P}_i . Then, $\bar{P}_i - (p_{i,t^{\min}} + \underline{P}_i)$ can be utilized to retain energy in hydro storages. Furthermore, all time periods on the left- and on the right-hand side of t^{\min} within the operating phase of i are also considered for retaining energy.

Within the identification part, the generating and retaining phases found previously are stored in separated lists (namely, L^{gen} and L^{ret}). Prior executing the coordination part, candidates in list L^{gen} are sorted according to non-increasing cost reductions and candidates in L^{ret} are sorted according to non-decreasing variable operating costs. Both costs relate to the changed use of thermal power plants through the incorporation of hydro storages. In the case of a generating phase, there are cost reductions in the amount of fixed and variable costs incurred if the thermal plant is possibly shutdown. In retaining phases, the higher-power output results in additional costs in the amount of variable operating costs. Algorithm 1 depicts the subsequent coordination part, where iteratively generating and retaining phases are selected and matched.

Algorithm 1: Coordination of generating and retaining phases

```

Data: current  $S^{\text{thermal}}, S^{\text{hydro}}$ , lists  $L^{\text{gen}}$  and  $L^{\text{ret}}$ 
1 foreach ( $turb \in L^{\text{gen}}$ ) do
2   check if  $turb$  can be realized completely, can be shortened and realized, or
   cannot be realized;
3   if ( $turb$  can be realized) then
4     identify suitable  $pump$  phases in  $L^{\text{ret}}$ ;
5     if ( $the\ decommitment\ of\ i\ can\ fully\ be\ compensated$ ) and
      ( $cost\ savings\ can\ be\ achieved$ ) then
6       update  $S^{\text{thermal}}$  and  $S^{\text{hydro}}$ ;
7     end
8   end
9 end
10 return  $S^{\text{thermal}}, S^{\text{hydro}}$ 

```

The first line of Algorithm 1 shows that we iterate over all elements in list L^{gen} . Assuming that generating phase $turb$ in L^{gen} is considered, which belongs to the operating phase of plant i with a start in time period $t^{\text{gen},s}$ and an end in $t^{\text{gen},d}$ (see also Fig. 4). In accordance to line 2, the realization of $turb$, i.e., the decommitment of i and the compensation with hydro storage energy, has to be verified by ensuring the following prerequisites for each $t \in \{t^{\text{gen},s}, t^{\text{gen},d}\}$:

- (i) The energy level in each hydro storage $j \in \mathcal{J}$ in t must not fall below \underline{E}_j if additional generating power is provided. Please note that we first suppose fully filled hydro storages in $t^{\text{gen},s}$ (a sufficient storage level will be guaranteed in later steps).
- (ii) The additional generating power of all hydro storages (under consideration of already existing power levels) must compensate the power level of i in t without exceeding the maximum power of the respective turbines.

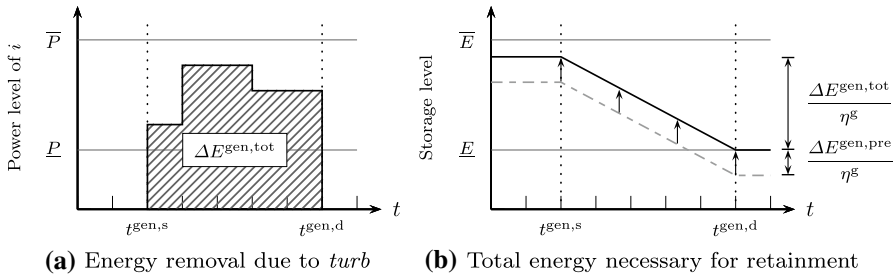


Fig. 6 Concept of the shift demand step

(iii) Constraints (4) and (5) are satisfied in t .

If the prerequisites are fulfilled only until a time $t' < t^{\text{gen,d}}$ with $t' \in \mathcal{T}$, *turb* is shortened and realized in $t^{\text{gen,s}}, \dots, t'$.

In order to compensate an energy removal $\Delta E^{\text{gen,tot}}$ according to *turb* (cf. Fig. 6a), suitable retaining phases *pump* have to be selected in compliance with the sequence in list L^{ret} (lines 3 and 4 of Algorithm 1). Please note that according to item (i) (cf. previous enumeration) fully filled hydro storages were assumed in period $t^{\text{gen,s}}$, which is not guaranteed in any case. Therefore, a specific amount $\Delta E^{\text{gen,pre}}$ ($0 \leq E^{\text{gen,pre}} \leq \Delta E^{\text{gen,tot}}$, cf. Fig. 6b) has to be retained before period $t^{\text{gen,s}}$ in order to guarantee that the level of no hydro storage falls below \underline{E} in $t \in \{t^{\text{gen,s}}, t^{\text{gen,d}}\}$ or any other period. Using a specific phase *pump* in list L^{ret} , a hydro storage $j \in \mathcal{J}$ is filled as long as

- (i) the surplus $\bar{P}_i - (p_{i,t'} + P_i)$ of committed plant $i \in \mathcal{I}$ connected to the current *pump* is depleted ($t' \in \{t^{\text{ret,s}}, \dots, t^{\text{ret,d}}\}$),
- (ii) the maximal pumping capacity of j is observed,
- (iii) constraints (12) are satisfied or
- (iv) $E^{\text{gen,pre}}$ is reached.

We start the procedure with the most efficient hydro storage $j \in \mathcal{J}$. Once the second or third condition is satisfied for storage j , the next efficient storage is selected. If the first condition is fulfilled, the next period $t' + 1$ (in case $t' < \min\{t^{\text{ret,d}}, t^{\text{gen,s}}\}$) or the next *pump* in list L^{ret} (in case $t' = \min\{t^{\text{ret,d}}, t^{\text{gen,s}}\}$) is chosen. In the event that the fourth condition occurs, the procedure terminates. The remaining energy $\Delta E^{\text{gen,tot}} - \Delta E^{\text{gen,pre}}$ (under consideration of efficiencies) can then be retained before time $t^{\text{gen,s}}$ or after time $t^{\text{gen,d}}$ in an analogous way to (i)–(iv).

After the selecting and matching step, all schedules must be updated if the energy removal $\Delta E^{\text{gen,tot}}$ can be compensated and cost savings can be achieved (cf. line 5 in Algorithm 1). The update (line 6) is performed as follows: unit i that corresponds to the selected phase *turb* has to be decommitted, i.e., we set $u_{i,t} := 0$ for $t \in \{t^{\text{gen,s}}, \dots, t^{\text{gen,d}}\}$ in S^{thermal} . Furthermore, according to the designated additional pumping power and the designated generating power (due to the selected and used *pump* and *turb* phases), the energy levels $e_{j,t}$, $j \in \mathcal{J}$ (if used) and $t \in \mathcal{T}$ are adjusted.

We use the updated (and feasible) schedules S^{thermal} and S^{hydro} in order to start a new coordination iteration. Hence, the next phase *turb* in list L^{gen} is selected and matched with other phases *pump* (cf. Algorithm 1), where list L^{ret} has to be updated according to the already selected pump phases.

Finalize solution A last finalizing step is performed to replace or decommit (redundant) operating phases of thermal plants if cost savings are thereby achievable. The procedure is similar to the local optimization steps in the **first stage** of the algorithm visualized in Fig. 1. However, in addition to thermal plants, hydro storages for the provision of spinning reserve have to be taken into account. If excessive spinning reserve is provided (due to the updated schedules), further decommitments of thermal plants can be performed, always maintaining solution feasibility. Additionally, the provided energy of the fictitious plant that was introduced in the preprocessing (cf. Fig. 1) is replaced by variable n_t in each period t . Please note that the usage of the fictitious plant is reduced significantly due to the shift demand and improve/finalize steps.

4 Experimental performance analysis

Within the performance analysis (cf. Sect. 4.1), we compared the results obtained by our two-stage heuristic approach to CPLEX results using our MILP formulation (cf. Sect. 2.1). We initially assume that emission penalty costs are zero ($c_t^{\text{CO}_2} := 0, t \in \mathcal{T}$); therefore, only thermal production costs are to be minimized. Further results analyzing production and emission costs in the UCP-HT can be found in Sect. 4.2.

4.1 Performance results of the two-stage heuristic approach

The following study makes use of two test sets T_1 and T_2 , where each set consists of 50 instances. Test set T_1 was first introduced by Kazarlis et al. (1996), and it is commonly used for analyzing short-term UCP models (e.g., Ongsakul and Petcharaks 2004; Viana et al. 2008; Delarue et al. 2013, and Morales-España et al. 2013). All instances in T_1 are based on the same plant system with 10 thermal units. Then, this 10-plant system is exactly replicated in order to obtain a system with 20,30,...,100 plants. Moreover, demand values for a planning horizon of 24 h are adopted by Kazarlis et al. (1996) and proportionally adjusted relatively to the total capacity of the replicated plant system. Since we need planning horizons for investigating our long-term problem, we copied the 24-h horizon in order to get instances of 1, 4, 12, and 20 weeks as well as 1 year (i.e., 24 to 8760 hourly time periods). Please note that on weekends only 80 % of the demand is chosen. The spinning reserve requirements are set to 10 % of the corresponding demand in each period. Test set T_2 is derived from T_1 by replacing the demand by a highly volatile residual demand that appeared in Germany in 2015 (relatively adjusted to the total thermal capacity). Furthermore, hydro storages are added to each instance of T_1 and T_2 . Thereby, hydro

Table 1 Computational results for T_1 - and T_2 -instances

Horizon	MILP (T_1)		Heuristic (T_1)		MILP (T_2)		Heuristic (T_2)	
	Gap (%)	t^{CPU} (s)	Gap (%)	t^{CPU} (s)	Gap (%)	t^{CPU} (s)	Gap (%)	t^{CPU} (s)
1 week	0.2 (0.8)	11	0.8 (0.9)	0.1	0.3 (0.6)	31	0.8 (1.0)	0.1
4 weeks	0.5 (0.9)	177	0.8 (0.9)	0.5	0.6 (1.0)	198	0.6 (0.8)	0.3
12 weeks	0.7 (0.9)	558	0.9 (1.0)	4.2	0.5 (0.9)	3406	0.7 (0.9)	2.4
20 weeks	0.6 (0.9)	1103	0.9 (1.0)	10.6	7.4 (68.9)	3586	0.8 (1.0)	5.9
1 year	6.3 (48.4)	14,245	0.9 (1.0)	72.1	62.8 (99.8)	14,167	0.9 (1.0)	58.7

storages had a share of about 7 % of the total thermal capacity (as it was the case in Germany in 2015).

All optimization runs are performed on an Intel i7-2760QM with 2.7 GHz and 16 GB RAM. The mixed-integer linear models were solved with CPLEX 12.6 (using GAMS) applying an optimality gap of 1 % and a time limit of 5 h. Under preliminary tests, we found out that flow-cover, Gomory fractional, and mixed-integer rounding cuts provided by CPLEX perform best and should be used (all other cuts are excluded). Our two-stage heuristic approach proposed in Sect. 3 has been implemented in C# (.NET 4.0).

Computational results for test sets T_1 and T_2 with 10–100 plants and planning horizons ranging from one week to one year are presented in Table 1.³ Columns 2–5 refer to test set T_1 and columns 6–9 to test set T_2 . “Gap” displays the average gap [%] between the best integer solution found and the best lower bound obtained by CPLEX. The worst-case gap values are written in parentheses. “ t^{CPU} ” shows the average computation times in seconds [s].

Applying the MILP model, all instances in test set T_1 with a planning horizon of up to 20 weeks are solved to (near-)optimality (i.e., a gap ≤ 1 % is reached) within the specified time limit of 5 h. For instances with a one-year planning horizon, the time limit often prevents good outcomes in terms of gap values. Instances in T_2 -instances seem to be more difficult to solve, since the gap values are larger compared to the gaps of T_1 -instances. Here, only instances with a planning horizon of up to 12 weeks can be tackled with gap values lower than 1 %. Moreover, average gaps for instances with a one-year planning horizon are quite high (approx. 63 %). The solution quality of the two-stage heuristic approach is comparable to MILP for all medium-term instances in T_1 and T_2 , as gap values are lower or equal to 1 %. For long-term instances, that are mainly focused in this article, great differences between heuristic approach and MILP can be observed. The heuristic approach is always able to find significantly better solutions (even worst-case gap values are ≤ 1 % in contrast to the respective MILP gaps that are < 100 %) within considerable less computation time ($t^{\text{CPU}} < 2$ minutes in contrast to MILP $t^{\text{CPU}} \leq 5$ h). Due to the

³ Test sets and detailed results can be downloaded from <http://www.wiwi.tu-clausthal.de/abteilungen/unternehmensforschung/forschung/benchmark-instances/>.

Table 2 Relative improvements Δ and solution times t^{cpu} for each step of the heuristic approach (for one-year instances)

#Plants	Thermal plant optimization					Hydrothermal coordination			
	Create Sol.	Repair Sol.		Improve Sol.		Shift Demand		Finalize Sol.	
		t^{cpu} (s)	Δ (%)	t^{cpu} (ms)	Δ (%)	t^{cpu} (ms)	Δ (%)	t^{cpu} (ms)	Δ (%)
10	7	-2.50	9	2.27	47	3.04	695	1.23	49
20	11	-2.63	11	2.18	72	3.39	2481	1.02	120
30	17	-4.15	13	3.53	107	3.98	4290	1.09	121
40	21	-3.37	15	2.75	147	3.97	6780	1.01	155
50	25	-4.15	17	3.15	185	3.72	8591	1.36	212
60	31	-4.11	18	2.98	233	3.36	12,062	1.44	303
70	34	-4.08	21	2.95	282	3.95	15,205	1.45	366
80	40	-4.44	22	3.12	339	4.09	18,831	1.64	619
90	44	-4.49	24	3.08	406	4.09	20,764	1.73	831
100	49	-4.53	27	3.09	482	3.19	26,444	1.75	1102
avg	28	-3.84	18	2.91	230	3.68	11,614	1.37	388

fast run times (within which outstanding solutions are generated), the two-stage heuristic seems to be a good choice for practitioners.

A detailed analysis of the two-stage heuristic approach is illustrated in Table 2. Each optimization step (cf. Fig. 1) is evaluated individually using a third test set T_3 with long-term instances. T_3 consists of the ten one-year problem instances of T_1 , where the recurrent demand is again replaced by volatile residual demands. We here use the six yearly residual demands occurred in Germany from 2010 to 2015 (cf. ENTSO-E and German TSOs) and adjusted them proportionally to the total capacity of the respective plant system. Thus, 60 long-term instances are obtained with a 10 to 100 plant system. In Table 2, average solution times t^{cpu} and average relative improvements Δ , compared to the respective previous optimization step, are presented. Hence, for the first step, i.e., the creation of a start solution, improvements do not exist. Negative improvements indicate a deterioration of the objective function value. In the transition from a created to repaired solution negative improvements occur. From there to the finalized solution, positive improvements are realized. Additionally, solution times given in milliseconds [ms] are negligibly small in all steps of the thermal plant optimization stage (cf. Sect. 3.1). Most of the overall solution time is spent within the integration of energy storages, in particular for larger systems. Performance gains, that are achieved within this hydrothermal coordination stage (cf. Sect. 3.2), strongly depend on the characteristics of the problem instances. More precisely, for instances with high demand differences, improvements are increasing. For instances with primarily base load situations (e.g., due to a low volatility or a low feed-in of the renewable sources), the smoothing effect of energy storages are negligible and only small improvements could be reached. Furthermore, the demand shifting offers the overall best improvements. Further improvements within the last step are generated by avoiding an excessive spinning reserve.

Table 3 Objective function values and gaps compared to the best-known solutions for $|T| = 24$ h

#Plants	MIP	Heuristic	MIP-MOR	DPLR	GA
10	563,938	565,278	566,218	564,049	565,825
20	1,123,871	1,128,200	1,128,558	1,128,098	1,126,243
40	2,243,270	2,249,814	2,245,746	2,256,195	2,251,911
60	3,361,998	3,372,620	3,370,333	3,384,293	3,376,625
80	4,482,358	4,495,092	4,485,972	4,512,391	4,504,933
100	5,600,181	5,617,841	5,611,847	5,640,488	5,627,437
avg	0.00 %	0.33 %	0.24 %	0.50 %	0.39 %

To further assess the solution quality of the heuristic approach, we isolated its first stage (cf. Sect. 3.1) and compared the objective function values to results of procedures devised by Kazarlis et al. (1996), Morales-España et al. (2013), and Ongsakul and Petcharakas (2004). Hence, a feasible production schedule for thermal plants is determined. The computational results of the widely used Kazarlis et al. (1996) test set (for a 24 h planning horizon without hydro storages) are given in Table 3. For the solver solution process, a gap of less than or equal to 1% is chosen as a stopp criterion. Columns 2 and 3 state the objective function values generated by our procedures, whereas “MIP-MOR” illustrates the results using the mixed-integer linear programming approach by Morales-España et al. (2013). Column “DPLR” displays the results for a Lagrange relaxation combined with a dynamic programming approach for the economic dispatch problem given by Ongsakul and Petcharakas (2004). Finally, column “GA” represents objective function values using the genetic algorithm presented by Kazarlis et al. (1996). Since “DPLR” and “GA” use squared production costs [cf. Eq. (1)], we also considered those costs in our procedures (“MIP” and “Heuristic”) as well as in our implementation of “MIP-MOR” (same settings as in “MIP”) to facilitate comparability. Please note that due to the quadratic objective, our mixed-integer programming model is named “MIP” (in contrast to the “MILP” model in Table 1). Line “avg. gap” represents the average gap between the objective function values obtained by the respective procedure and the best-known objective function values (within this paper). Overall, the models “MIP” and “MIP-MOR” give the best results in terms of average gap, followed by the two-stage heuristic, which also provides a significantly lower computation time. Please note that only “MIP”, “MIP-MOR”, and “Heuristic” are comparable with respect to CPU-times, since “DPLR” and “GA” are solved on different computers (Ongsakul and Petcharakas 2004 present solution times of 2–3 minutes for “DPLR” and “GA” vs. less than a second for our “Heuristic”). To sum up, the first three steps within the first stage (cf. Sect. 3) seem to provide a quite good basis for our two-stage heuristic approach solving the UCP–HT.

Table 4 Results for different ψ for the German electricity market for one year (#instances: 27 for $\psi \in \{0.25, 0.5, \dots, 5.00\}$)

ψ (-)	#Plants (-)	$\sum_i P_i^{\max}$ (GW)	$\sum_t D_t$ (TWh)	$\sum_t \text{wind}_t$ (TWh)	$\sum_t pv_t$ (TWh)	t^{cpu} (s)	Gap (%)
0.25	80	22.8	138.1	22.0	9.6	23.5	0.98
0.50	148	45.6	276.2	44.0	19.2	52.7	1.22
0.75	221	68.4	414.3	66.0	28.8	28.8	1.04
1.00	289	91.3	552.4	88.0	38.4	150.6	1.46
1.50	436	136.9	828.6	132.0	57.6	186.3	(6.13)
2.00	595	182.5	1104.8	176.0	76.8	369.0	(5.59)
3.00	870	273.8	1657.2	264.0	115.2	805.8	(6.08)
5.00	1467	458.4	2762.1	440.0	192.0	1711.5	n.a.

Bold values indicate the benchmark value

4.2 Case study for the German electricity market

The numerical results obtained in Sect. 4 make the heuristic procedure highly suitable for solving real-world energy problems. The following case study is based on the characteristics of the German electricity market that comprises of nearly 300 large-scale power plants with about 91 GW of thermal power in 2015. Moreover, 126.4 GW of volatile renewable units, 6.5 GW of pumped storages, and a cumulated net demand of 550 TWh are used. Furthermore, in 2015 the average price for CO₂ emissions was $c^{\text{CO}_2} := 7.94$ € per ton CO₂ (since c^{CO_2} is an average price, period t in $c_t^{\text{CO}_2}$ is eliminated). We choose a long-term planning horizon of one year. In order to extensively assess the capability of the proposed heuristic for real-world instances, a scaling factor ψ is introduced. Thereby, the total thermal capacity $\sum P_i^{\max}$ [GW], the yearly demand $\sum D_t$ [TWh], as well as the renewable feed-in $\sum \text{wind}_t$ and $\sum pv_t$ [TWh] are multiplied by $\psi \in \{0.25, 0.50, 0.75, 1.00, 1.50, 2.00, 3.00, 5.00\}$. Based on the scaled values, the thermal plant portfolio is derived such that it represents a ψ -adjusted electricity system of Germany. The portfolio of hydro storages comprises only storages that participate on the wholesale market (here, 6.5 GW of generating power) and is scaled accordingly. Realistic hourly values for the demand and for renewable feed-ins are derived by the methodology in Wagner (2014) applying the calibration data from 2010 to 2015 (cf. ENTSO-E and German TSOs). The stochastic Ornstein–Uhlenbeck process of Wagner (2014) is used to generate three scenarios for the demand as well as three scenarios for wind and solar infeeds, respectively. Consequently, 27 possible combinations are considered, with the residual demand specified in each combination (see Definition 3). This might have some small impacts on the results compared to a detailed modeling of renewable sources, where, e.g., each wind farm is considered individually. However, these small differences seem to be negligible against the background of the underlying long-term planning horizon and the fact that according to Wagner (2014), the behavior of the modeled renewable feed-ins is quite realistic. For run-of-river and biomass provided electricity, a constant monthly profile given by the ENTSO-E is utilized.

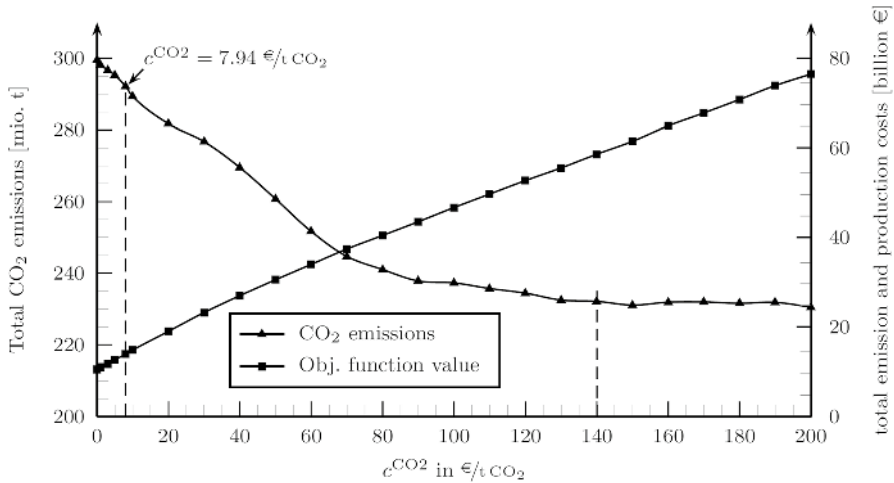


Fig. 7 Impact of CO₂ market prices c^{CO_2}

Results of our two-stage heuristic are depicted in Table 4. As expected, instances with a low scaling factor ψ are solved quite fast, but also for the entire German electricity system ($\psi = 1$), solutions can be generated in about 2.5 min. Moreover, a feasible hourly production schedule for a system with a scaling of $\psi = 5$ (with almost 1500 plants and in terms of the demand nearly comparable to the system of the EU-28) is received in less than half an hour. Within the test set with one scaling factor, only small solution time variations are observed. The related solution quality is represented by gap values that are determined using the best MILP solution found within a time limit of 48 h. Here, gap values are always below 1.5 %, which results again in quite high solution qualities. For higher scaling factors ($\psi \geq 1.50$) no feasible MILP-solutions can be achieved in 48 h due to the rapidly increasing model size. Therefore, the LP-relaxations of our MILP-model are used to assess the solution quality. Note that the resulting gap values are not as tight as the previous values and are hence written in parentheses. However, they indicate the maximal deviation to the optimum (all deviations are lower than 7 %). For $\psi = 5$, “n.a.” means that CPLEX cannot provide solutions within the prescribed time limit. Considering the results, it can be concluded that the two-stage heuristic procedure provides not only outstanding results for theoretical test instances, but is also excellently for comprehensive practical purposes.

Furthermore, the proposed heuristic approach can be applied in order to receive trade-off results between the minimization of emission costs and the minimization of production costs. In what follows, the weighted-sum method is used (cf. Sect. 2.2), where weighting factors for CO₂ emissions are set according to the market price c^{CO_2} . Trade-offs are approximated by varying the weight c^{CO_2} from zero emission costs to high costs (e.g., 200 €/t CO₂). Figure 7 illustrates the resulting trade-off curve for the German electricity market with raising market prices c^{CO_2} .

Please note that the corresponding solution times and solution gaps are comparable to the results in Table 4 for $\psi = 1.00$.

It might be deduced from Fig. 7 that a shift from coal- to gas-fired plants (with less CO₂ emissions) appears with increasing CO₂ prices and thereby CO₂ emissions are reduced. Therefore, the principle of the EU ETS can generally be used to control the management of CO₂ emissions. Reductions in CO₂ emissions are only achieved by accepting much higher production costs. In our case study, a reduction of nearly 15 % due to increase in c^{CO_2} from 7.94 to 140 € per ton CO₂ results in a rise in total production cost of 350 % (from ca. 17 to 60 billion €). All in all, the presented technique offers a possibility to evaluate the effectiveness of the EU ETS.

5 Conclusion

In this paper, we considered the UCP with hydrothermal coordination, i.e., the UCP–HT. In order to meet recent challenges of reducing emissions, particularly CO₂ emissions, we enhanced the UCP–HT to a UCP–HT with production and emission costs. The emission costs are included on the basis of the cap-and-trade principle of the EU ETS. Nowadays, the trend to low-carbon electricity supply prefers units with less emissions (e.g., gas-fired plants); consequently, cost-efficient power plants (e.g., coal-fired plants) are no longer prioritized at all. In the presence of (high) market prices for CO₂ allowances, units with low emissions may achieve superior positions in the merit order and thereby higher utilizations. In contrast, carbon-intensive plants move significantly down in the merit order.

For the proposed UCP–HT with and without emission costs, a MILP formulation as well as a new and fast two-stage heuristic approach are presented. The proposed approach significantly outperforms the MILP as well as other tested techniques for long-term planning horizons. Particularly, real-world instances on the basis of the German electricity market can be solved in minutes, where an outstanding solution quality is realized. Furthermore, the two-stage heuristic allows sophisticated analyses of trade-off situations between thermal production costs and emission costs, e.g., to determine suitable CO₂ prices for an effective reduction in the overall emissions. Our case study emphasizes that current low market prices for CO₂ allowances are quite effectless, which goes in line with most other studies.

It can be concluded that with regard to the need of fast scheduling procedures for the operational planning of thermal power plants, renewable sources, and storages, the two-stage heuristic approach is highly suitable for real-world applications and their analysis. In order to further improve the presented approach, it can be embedded into a multi-start scheme, where different solutions are generated and the best solution found is returned. Future research might also observe plant availabilities, combined heat-and-power techniques for thermal plants or electricity trading with foreign countries. These aspects extend either the constraint set or the objective function and could affect the targets of different decision makers (e.g., system operators, generating and industrial companies). Generally, the two-stage approach would be a good basis for those extensions due to its modular architecture.

Acknowledgements Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bartsch M, Röhling A, Salje P, Scholz U (2008) Stromwirtschaft: ein Praxishandbuch. Carl Heymanns, Köln
- Bendotti P, Fouilhoux P, Rottner C (2017) On the complexity of the unit commitment problem. *Optimization Online*, http://www.optimization-online.org/DB_HTML/2017/06/6061.html, pp 1–9. Accessed 26 Sept 2018
- Carrión M, Arroyo JM (2006) A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Trans Power Syst* 21:1371–1378
- Delarue E, Cattrysse D, D'haeseleer W (2013) Enhanced priority list unit commitment method for power systems with a high share of renewables. *Electr Power Syst Res* 105:115–123
- Ellermann AD, Convery F, Perthuis C, Alberola E (2010) Pricing carbon: the European union emissions trading scheme. Cambridge Univ. Press, Cambridge
- Franz A, Zimmermann J (2018a) A two-stage heuristic approach for solving the long-term unit commitment problem with hydro-thermal coordination in large-scale electricity systems. In: *IEEE international conference on industrial engineering and engineering management*, Nusa Dua, Indonesia, pp 70–74
- Franz A, Zimmermann J (2018b) A two-stage heuristic procedure for solving the long-term unit commitment problem with pumped storages and its application to the german electricity market. In: Fink A, Fügenschuh A, Geiger G (eds) *Oper Res Proc 2016*. Springer, Berlin, pp 149–155
- Franz A, Rieck J, Zimmermann J (2017) Two-stage heuristic approach for solving the long-term unit commitment problem with hydro-thermal coordination. In: Doerner K, Ljubic I, Pflug G, Tragler G (eds) *Oper Res Proc 2015*. Springer, Berlin, pp 667–673
- Franz A, Rieck J, Zimmermann J (2019) Fix-and-optimize procedures for solving the long-term unit commitment problem with pumped storages. *Ann Oper Res* 274:241–265
- Garey MR, Johnson DS (1979) *Computers and intractability: a guide to the theory on NP-completeness*. Freeman, New York
- Gollmer R, Möller A, Nowak MP, Römisch W, Schultz R (1999) Primal and dual methods for unit commitment in a hydro-thermal powersystem. In: 13th Power system computation conference, Trondheim, Norway, pp 724–730
- Gollmer R, Nowak MP, Römisch W, Schultz R (2000) Unit commitment in power generation—a basic model and some extensions. *Ann Oper Res* 96:167–189
- Kazarlis S, Bakirtzis JM, Petridis V (1996) A genetic algorithm solution to the unit commitment problem. *IEEE Trans Power Syst* 11:83–92
- Kruger JA, Pizer WA (2010) Greenhouse gas trading in Europe: the new grad policy experiment. *Environment* 46:8–23
- Martínez Díaz DJ (2008) Production cost models with regard to liberalised electricity markets. Ph.D. thesis, Karlsruhe Institute of Technology, Karlsruhe
- Morales-España G, Latorre JM, Ramos A (2013) Tight and compact MILP formulation for the thermal unit commitment problem. *IEEE Trans Power Syst* 28:4897–4908
- Ongsakul W, Petcharak N (2004) Unit commitment by enhanced adaptive Lagrangian relaxation. *IEEE Trans Power Syst* 19:620–628
- Ostrowski J, Anjos MF, Vannelli A (2012) Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Trans Power Syst* 27:39–46
- Padhy NP (2004) Unit commitment—a bibliographical survey. *IEEE Trans Power Syst* 19:1196–1205

- Raglend IJ, Padhy NP (2006) Solutions to practical unit commitment problems with operational, power flow and environmental constraints. In: IEEE power engineering society general meeting, Montreal, Canada, pp 1–8
- Saravanan B, Das S, Sikri S, Kothari DP (2013) A solution to the unit commitment problem—a review. *Front Energy* 7:223–236
- Saravanan B, Vasudevan ER, Dwarkadas K (2014) Emission constrained unit commitment problem solution using invasive weed optimization algorithm. *Int J Electr Power Energy Syst* 55:21–28
- Soliman SA, Mantawy AH (eds) (2012) *Modern optimization techniques with applications in electric power systems*. Springer, New York
- Tietenberg TH (2010) *Emissions trading: principles and practice*. Routledge, Washington
- Tseng CL (1996) On power system generation unit commitment problems. Ph.D. thesis, University of California, Berkeley
- Viana AM (2004) *Metaheuristics for the unit commitment problem: the constraint oriented neighbourhoods search strategy*. Ph.D. thesis, University of Porto, Porto
- Viana AM, Pinho de Sousa J, Matos MA (2008) Fast solutions for UC problems by a new metaheuristic approach. *Electr Power Syst Res* 78:1385–1395
- Wagner A (2014) Residual demand modeling and application to electricity pricing. *Energy J* 35:45–72
- Wood AJ, Wollenberg BF, Sheble GB (2013) *Power generation, operation and control*. Wiley, New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.