



# Model-predictive energy management system for thermal batch production processes using online load prediction

Florian Fuhrmann\*, Alexander Schirrer, Martin Kozek

*Institute of Mechanics and Mechatronics, Technische Universität Wien, Getreidemarkt 9/BA, Wien 1060, Austria*

## ARTICLE INFO

### Article history:

Received 21 October 2021

Revised 27 April 2022

Accepted 27 April 2022

Available online 3 May 2022

### Keywords:

Energy management system

Model predictive control

Online load prediction

Mixed-integer linear programming

Thermal batch processes

## ABSTRACT

Predictive energy management systems (EMS) enable industrial plants to participate in the modern power market and reduce energy cost. In this paper, a novel modular model predictive EMS specifically designed for industrial thermal batch processes is presented. The EMS consists of a two-layer mixed-integer model predictive controller and an online load predictor, and thus solves the main challenges of EMS in industry - high implementation costs and the possible reduction of production reliability. The modular formulation of the optimization problem enables system integrators to implement the EMS without time-consuming modelling tasks and elaborate parameter tuning. The online load predictor estimates the typical pulse-like heat loads of batch processes ensuring both - reliable production and maximal flexibility of the power demand. The utilization of real-time data provides additional robustness against uncertainties caused by human operators. The performance of the EMS is evaluated in a case study of an existing food plant.

© 2022 The Author(s). Published by Elsevier Ltd.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

Decarbonisation of factories is a key measure to fight climate change as the industrial sector accounts for 29% of global final energy consumption (IEA, 2019). To increase the pressure on industry to reduce energy consumption, emission trading is introduced, for example through the EU Emissions Trading System. In addition, the electricity market has been liberalized to promote flexible electricity consumption and allow for a higher share of renewable energy in the power grid. A predictive energy management system (EMS) is necessary for industrial plants to participate and profit from this modern power market. The goal of an EMS is to operate energy supply systems (ESS) optimally in terms of energy costs, energy efficiency, machine wear and CO<sub>2</sub>-emissions, while complying with operating limits ensuring production safety. Today, only few factories employ an EMS, but EMSs are attracting increased interest (Sirola and Edgar, 2012) and are the subject of intense research in various areas like microgrids (Zia et al., 2018), fuel cells (Teng et al., 2020), urban energy systems (Moser et al., 2020; Powell et al., 2016), heating, ventilation and air conditioning (HVAC) (Rawlings et al., 2018; Dullinger et al., 2018; Touretzky and Baldea, 2014; Risbeck et al., 2020), homes

(Shareef et al., 2018; Touretzky and Baldea, 2016), and powertrains (Biswas and Emadi, 2019).

Due to the modern power market EMSs are becoming more widely used in large industrial plants (Petek et al., 2018). Nevertheless, there are still obstacles to the widespread application of EMSs in industry in general and in batch production processes in particular. May et al. (2017) provided an overview of the state of the art in EMS for the manufacturing sector, stating the potential negative impact on production performance as the main barrier to their adoption. The production performance is reduced in case the EMS violates process constraints in an effort to increase energy efficiency. Model predictive control (MPC) is a suitable method to overcome this barrier. It can incorporate economic and operational objectives - in the form of a flexible cost function addressing multiple goals - while respecting various technical, regulatory, and process constraints (Dengiz et al., 2021). Prediction accuracy is critical to the performance and reliability of MPC. Srinivasan et al. (2003) detect uncertainty as the main bottleneck in using optimization based methods at the industrial level. Incorrect predictions can lead to the violation of critical constraints which can affect the quality of end-products (Thombre et al., 2020). Due to this fact, the EMSs are often operated overly conservatively (Shareef et al., 2018; Touretzky and Baldea, 2016), reducing the flexibilities and thereby economic benefits of EMSs. Different approaches to tackle the challenges of conservative EMSs for systems with uncertainty have been presented: Thombre et al. (2020) employ multivariate data

\* Corresponding author.

E-mail address: [florian.fuhrmann@tuwien.ac.at](mailto:florian.fuhrmann@tuwien.ac.at) (F. Fuhrmann).

## Nomenclature

### Abbreviations

BC	batch consumer
COP	coefficient of performance
EMM	energy management in manufacturing
EMS	energy management system
ESS	energy supply system
HLC	higher-level control
HP	heat pump
HS	heat source
HT	heat treatment
HVAC	heating ventilation and air conditioning
LLC	low-level control
MILP	mixed-integer linear programming
MPC	model predictive control
OLP	online load predictor
RU	ramp-up
RD	ramp-down
SOC	state of charge
SU	start-up
SD	shut-down
TES	thermal energy storage

### Symbols

$\alpha, \beta$	linearization coefficients
$\epsilon_{\text{fix}}$	fix loss
$\epsilon_{\text{SOC}}$	loss proportional to SOC
$\eta$	compressor efficiency
$\rho$	density
$\Delta T$	temperature difference
$\Delta N$	time-step shift
$\dot{m}$	mass flow
$\mathbf{C}$	cost coefficient vector in $\epsilon$
$c$	cost factor in e.g. $\epsilon/\text{MWh}$
$c_p$	mass-specific heat coefficient
$n$	number of time-steps
$N_p$	prediction horizon
$p$	pressure
$P$	power consumption
$Q$	heat
$RU$	ramp up
$RD$	ramp down
$S$	slack variable
$t$	time
$t_s$	sampling time
$T$	temperature
$U$	plant input
$u$	operation condition
$v$	start-up integer
$V$	volume
$w$	shut-down integer

### Indexes

avg	average
C	charge
D	discharge
full	full load
in	incoming mass flow
$j$	running index
$k$	current time step
lhs	left hand side
lim	limit
min	minimum value

max	maximum value
out	outgoing mass flow
part	partial load
rhs	right hand side
sink	heat sink of the heat pump
source	heat source of the heat pump

### Further nomenclature

$X$	scalar variable
$\mathbf{X}$	vector variable
$\dot{X}$	time derivative of $X$
$\hat{X}$	estimate of $X$

analysis on historical industrial data to implement a multistage nonlinear MPC scheme based on a scenario-tree formulation. [Moretti et al. \(2020\)](#) use an affine adjustable robust formulation for the optimization problem.

These methods require intense data processing and are of significant complexity, hindering their implementation in factories. To reduce complexity and increase the robustness against prediction errors, a combination of optimization-based and (meta-)heuristic approaches of EMS can be used ([Dengiz et al., 2019](#)). These (meta-)heuristic EMSs are naturally only valid for their specific area of application. Implementing an EMS is especially challenging for batch processes. Therein, the energy demand is peak-shaped and exceeds the maximum energy supply rate during certain processes. Only predictive measures can prevent bottlenecks in the energy supply for batch processes and thereby ensure production reliability. Further, batch processes are typically semi-automated which causes uncertainties in the load predictions. No methods addressing these specific problem characteristics in a modular fashion could be found in available literature. This represents a significant research gap for an EMS structure which ensures production safety for batch processes.

The second big barrier of implementing EMSs in producing industry are the high implementation costs. Model-based EMSs rely on a system model which considers all significant system dynamics and constraints of the plant. The modelling process is highly challenging and typically requires expert knowledge. ESSs of production sites are often grown structures which hinders a standardized integration of EMSs ([Fluch et al., 2017](#)). Additionally, ESSs are continuously changed for example due to the expansion of production or the incorporation of renewable energy sources. To avoid the repetition of the modeling process after each modification, the system has to be structured in a modular, easily adaptable way. According to [Isaksson et al. \(2018\)](#), the modeling effort is the most important issue of MPC from an industrial perspective, and cost-efficient formulation and maintenance of the models is crucial. Modular mixed-integer linear programming (MILP) formulations of the system model have proven to be easily adaptable and simple to implement ([Moser et al., 2020](#)). The MILP-formulation has a further benefit: ESS consist of multiple components with switching behavior (heat pumps, gas boilers, etc.). Integer variables are the method of choice to model switching behavior – therefore, MILP is an efficient choice for this optimization problem. On the other hand, MILP requires non-linear effects to be linearized. The most evident nonlinearity in thermal energy systems is the mixture of fluids with different temperatures. Thermal batch processes typically have different required temperature levels and thereby varying demand temperatures. [Moser et al. \(2020\)](#) presented a modular MILP-based EMS for urban multi-energy systems. The pulse-like heat demand of batch processes and the high uncertainty of load predictions are major differences between urban ESSs and industrial ESSs. Therefore a research gap concerning a modular MILP-formulation of ESSs for thermal batch pro-

cesses and the incorporation of the occurring nonlinearities is evident.

Recently scheduling for energy management is in focus of research (Touretzky and Baldea, 2014; Risbeck et al., 2020; Touretzky and Baldea, 2016; Santander and Baldea, 2021; Beykal et al., 2022; Schäfer et al., 2020). Santander and Baldea presented a EMS for batch processes including scheduling of the production (Santander and Baldea, 2021). Beykal et al. (2022) suggest a two-layer architecture for integrating planning and scheduling problems under demand uncertainty. Touretzky and Baldea (2014, 2016) introduce a hierarchical economic MPC combining scheduling and control in the context of buildings with thermal energy storage. Risbeck et al. (2020) present a real-time capable MILP formulation for EMS with scheduling for HVAC. Scheduling increases the flexibility of ESS as heat demand can be synchronized with times with small electricity costs. Further, scheduling helps to prevent infeasibilities as unenforceable production schedules are avoided. However the implementation of scheduling in manufacturing plants demands cost-intense adaptations of the logistics and production system. On-demand production and the lack of acceptance by human operators can impede the implementation of scheduling. For instance, the food factory used as a case study in this paper cannot execute fully automated scheduling due to logistical reasons. Therefore, and as it is the goal of this paper to present lightweight EMS with little implementation effort, scheduling was not included in the EMS. This paper aims to close the aforementioned research gaps concerning EMSs for thermal batch processes. Therefore a modular EMS, consisting of a two-layer model predictive controller (MPC) and an online load predictor (OLP), is introduced in this paper. The two-layer MPC allows an efficient separation of the optimization goals production reliability and operation optimization. The higher-level controller (HLC) considers the predicted energy prices, CO<sub>2</sub>-certificates, and machine wear to calculate optimal trajectories, while the lower-layer controller (LLC) reacts to disturbances and ensures production reliability

In the current paper the optimization problem is formulated as a modular - component-based - MILP to enable fast implementation and easy adaptability. The given formulation of the mixed-integer linear program allows a straightforward parameterization from datasheets, enabling system integrators to implement and configure the energy management system without time-consuming modelling tasks and elaborate parameter tuning. The OLP utilizes the production schedule to estimate the heat load of future heat treatments (HT), reacts to measured deviations of the actual heat load and defines the time-dependent minimal state of charge  $SOC_{min}$  online in every time step. These three functionalities ensure process reliability while simultaneously maximizing the flexibility of the EMS. Furthermore, the formulation of the  $SOC_{min}$  avoids the typical nonlinearities in the MILP-formulation of thermal batch processes. The performance of the EMS is demonstrated for the use case of a food production plant. The EMS is compared against the installed baseline controller in simulations. The system models utilized for the simulation are detailed nonlinear models and validated by industrial measurement data.

The manuscript presents a model predictive energy management system for industrial batch processes that ensures production reliability and optimal operation.

The main contributions presented in this paper are:

- A novel energy management system structure consisting of a two layer model predictive controller and an online load predictor for batch processes is presented.

- A modular formulation of the arising mixed-integer linear program is suggested which enables a fast implementation and simple parameterization of the energy management system.
- An online load predictor for thermal batch processes ensures robustness of the energy management system against uncertainties caused by the discontinuous nature of batch processes and uncertainties of the production schedule.
- The performance of the novel energy management system is demonstrated in a case study based on the validated model of an existing food factory utilizing real industrial measurement data.

The remainder of the paper is structured as follows. In Section 2, the problem statement is given. In Section 3, the novel EMS is presented, and the design of the simulation study is described. In Section 4 the industrial use case is provided.

## 2. Problem statement

In this section, first, the considered energy supply systems' (ESS) structure and its properties in factories with batch production are described. After that, the premises of the load prediction are defined. Then, the optimization problem is outlined, and the main challenges in implementing an EMS in factories are discussed.

### 2.1. Structure of the energy supply systems

The factory processes considered in this publication are batch processes, where in contrast to continuous processes a certain amount of products is produced in a timeframe. This induces - often energy intense - start-up and shut-down processes for each batch. A typical example for energy intensive batch processes are heat treatments. During a heat treatment products undergo specific temperature trajectories to alter their physical or chemical properties (e.g., annealing, tempering or pasteurization). Heat treatments start with a heating phase, where the treated material is brought from the initial temperature  $T_0$  to the desired temperature level  $T_{end}$ . Therefore, the heating phases induce short pulse-like heat loads, as displayed in Fig. 1. Delayed or incomplete heating phases caused by insufficient heat supply may affect product quality and cause economic losses. As the demand peaks typically exceed the maximum energy supply rate during these process steps, ESS for batch factories include at least one buffer storage. Fig. 2 displays the resulting ESS structure for heat supply systems, which is considered in this paper. Essential components are a heat source (HS), a thermal energy storage (TES), and  $N$  batch-type heat consumers (BC) with temperatures  $T_{BC,n}$  ( $n = 1, 2, \dots, N$ ).

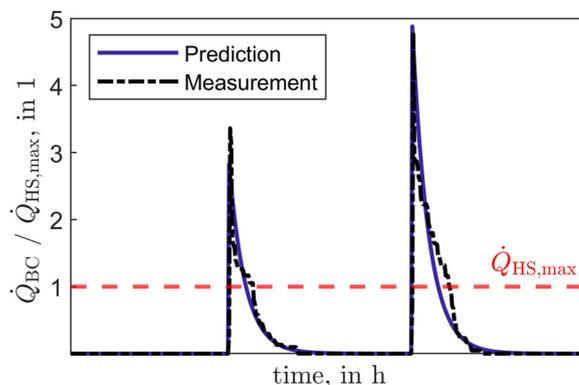


Fig. 1. Typical pulse-like heat loads of a batch consumer (BC) and the prediction compared to the maximum heat production of the heat source (HS). Adapted from (Matthew et al., 2020).

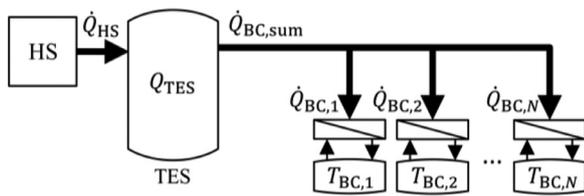


Fig. 2. Structure of energy supply systems for thermal batch processes consisting of a heat source (HS), a thermal energy storage (TES) and N batch consumers (BC). Adapted from (Matthew et al., 2020).

The sum of the heat loads  $\dot{Q}_{BC,sum}$  is the disturbance affecting the heat stored in the TES,  $Q_{TES}$ , which is controlled utilizing the heat supply unit's heat flow  $\dot{Q}_{HS}$  as the manipulated variable.

### 2.2. Load prediction

The prediction accuracy is critical for the performance and reliability of an MPC because wrong predictions can lead to the violation of critical constraints affecting the quality of end-products (Thombre et al., 2020). For example an under-estimated heat demand can result in a temperature constraint being violated during heat treatment. The too low temperature can lead to a reduced shelf life of food or undesirable material properties of hardened steel. To predict thermal loads caused by batch processes, the heat demand  $\dot{Q}_{BC,sum}$  and the associated demand temperature  $T_{Demand}$  have to be predicted based on the production schedule. The production schedule includes the starting temperature  $T_{0,HT,m}$ , end temperature  $T_{end,HT,m}$  and starting times  $t_{0,m}$  ( $m = 1, 2, \dots, M$ ) of all  $M$  heat treatments. In factories, the actual production process usually diverges from the production schedule because processes are not fully automated, and the duration of process steps depends on hardly predictable factors like educt quality or operator availability. Therefore, the starting time of the heat treatments,  $t_{0,HT,m}$ , deviates with a maximum process dependent deviation  $\Delta t_{0,HT,max}$ . The prediction error of the integral heat amount needed for a HT  $\Delta Q_{HT,max}$  is usually negligible as it is dependent on the well-known parameters starting temperature  $T_{HT,0,m}$ , end temperature  $T_{HT,end,m}$ , and heat capacity of the product  $C_{HT,m}$ . In the case study, the robustness of the suggested EMS against these deviations will be tested and the energy cost caused by different  $\Delta t_{start,max}$  values will be quantified for different control strategies.

### 2.3. Optimization problem

The goal of EMS is to operate energy supply systems (ESS) optimally in terms of energy costs, energy efficiency and machine wear while maintaining operating limits in the interest of production reliability. The critical constraint ensuring production reliability is that for each heat treatment, the desired heat flow  $\dot{Q}_{BC,m}$  must be available with the minimum temperature of  $T_{m,end}$ . The heat supply unit introduces additional constraints of minimal partial load, maximum load, and minimal stop time. To reduce wear and avoid an ineffective transient operation, the heat source should not be operated less than a desired minimal continuous operation time  $t_{up,desired}$ .

The power, gas, and fuel consumption of the ESS determine the energy costs. Due to decarbonization, fluctuating power prices become the main drivers of energy costs. To fully exploit the flexibility of the power market, a minimal prediction horizon of one day is needed to utilize the day-ahead market. Furthermore, the billing of power consumption is usually based on records at 15 min intervals. Therefore, a sampling time below 15 min is needed to effectively react to disturbances on the power consumption. Batch pro-

cesses induce short, intense loads and thereby demand a sampling time of few minutes.

Bottlenecks in the energy supply cause a product quality reduction, resulting in economic losses that are magnitudes higher than the energy costs of the product. Therefore, production safety is not seen as a cost factor but instead modelled as a constraint. Since thermal batch processes are usually semi-automated, the EMS must be robust to heat load uncertainties, as described in Section 2.2.

Bemporad and Morari (1999) describe a control system as robust if stability is maintained and performance indicators are met for a specified range of model deviations and noise signals. In this publication, the designed control structure is not obtained by classical approaches of robust control design. Instead robustness with respect to specific application aspects is realized by appropriate design decisions. In particular, the robustness against the uncertain noise signal heat demand  $\dot{Q}_{BC,sum}$  is discussed. The range of noise signals is specified by the maximum process dependent deviation  $\Delta t_{0,HT,max}$  and a prediction error of the integral heat amount needed for a HT  $\Delta Q_{HT,max}$ . The considered ESS are stable as the controlled system state  $Q_{TES}$  is physically limited by the minimum operation temperature  $T_{min}$  and maximum operation temperature  $T_{max}$  as described in Section 3.1.3. The relevant performance indicator for the system is that the heat flow  $\dot{Q}_{BC,m}$  must be available with the minimum temperature of  $T_{m,end}$ . The presented EMS aims to ensure the quality indicator and avoiding overly conservative reactions causing high energy cost or machine wear.

To enable a fast and efficient implementation, it is desirable that all model and control parameters are standard component parameters known to the plant operator. This also enhances the acceptance of the method by the plant operators. Further changes in the desired behavior shall be easily executable. Summing up, the implementation and maintenance of the control system shall be as intuitive as possible so that no control engineering expert is needed. Implementation cost and running costs need to be compensated by energy cost reduction.

## 3. Methods

The EMS consists of three components: the higher-level controller (HLC), the lower-level controller (LLC) - both utilizing MILP optimization - and the online load predictor (OLP), as displayed in Fig. 3. This section will first introduce the simple parameterizable formulation of the arising two-layer mixed-integer linear program. Then the OLP ensuring reliable production, robustness against uncertainties of the production schedule and maximal flexibility of the power demand is presented.

### 3.1. Higher-level and lower-level controller (HLC, LLC)

In this subsection, the architecture of the structure of the optimization problem is introduced, then the optimization models are introduced component-wise and finally the assembling of the plant-wide optimization problem is discussed.

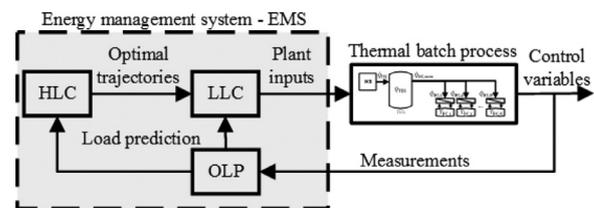


Fig. 3. Architecture of the energy management system (EMS) including the higher level controller (HLC), the lower level controller (LLC) and the online load predictor (OLP).

### 3.1.1. Optimization structure

The basic idea of separating the optimization problem into HLC and LLC is to execute the economic optimization with the HLC while ensuring production safety with the LLC. Furthermore, the separation allows to increase the robustness to uncertainties in the load prediction by setting the boundary conditions accordingly.

The robustness of the control system against the uncertainties in starting time  $\Delta t_{0,HT,max}$  and integral heat amount  $\Delta Q_{HT,max}$  is increased by:

- Adapting the heat load prediction as described in Section 3.2.2 according to possible uncertainties in starting time  $\Delta t_{0,HT,max}$ .
- Considering safety buffer  $\Delta Q_{robust}$  when calculating  $SOC_{min}$  to ensure robustness to  $\Delta Q_{HT,max}$  and mitigate the effect of uncertainties on EMS performance as described in Section 3.2.3.

The HLC and the LLC are model predictive controllers utilizing a mixed-integer linear programming (MILP) formulation. The MILP-optimization problem is defined by given inputs, a set of constraints and an objective function. The presented EMS is based on a modular – component-wise – formulation of the optimization problem similar to Moser et al. (2020). The modular framework enables efficient implementation and adaption of the EMS.

The two-layer structure of the EMS is necessary due to the high calculation effort caused by the requirements for the prediction horizon. The prediction horizon  $N_p$  should be defined long enough to fully exploit the power market but no longer to avoid unnecessary calculation effort. Typically, the prediction horizon is specified with 24 h because of the day-ahead power market. Furthermore, the optimization interval cannot be longer than few minutes to enable a control of the 15 min power consumption, which is the time basis of the billing of power consumption and to ensure production safety. In the two-layer concept, the HLC considers the long-term effects with a prediction horizon  $N_{p,HLC}$  of 24 h with a large sampling time  $t_{s,HLC}$  of 15 min. The LLC considers the short-term effects, with a prediction horizon  $N_{p,LLC}$  of 1.5 h with a short sampling time  $t_{s,LLC}$  of 1 min. Without the two-layer concept, the number of binary variables would impede the real-time applicability of the optimization. There are other solutions to reduce the computational burden of long forecast horizons caused by time-varying electricity prices, e.g., adaptive grid algorithms, but they have so far only been used for scheduling and the performance is highly influenced by operational constraints (e.g., ramp constraints) (Schäfer et al., 2020).

The two optimization layers both consider the same constraints and inputs. For both controllers, the initial condition of the heat source  $U_0$ , the initial state of charge  $SOC_0$ , the heat demand  $\dot{Q}_{BC,demand}$  and the power costs  $c_{power}$  are inputs. The HLC differs from the LLC in terms of cost weights, optimization parameters and the considered minimal state of charge  $SOC_{min}$ . The differently chosen  $SOC_{min}$  constraints ensure, robustness to uncertainties in the load prediction and mitigate the effect of uncertainties on EMS performance. Aggressive reactions of the EMS to deviations of the SOC in its prediction are only executed when the production safety is endangered. The detailed definition of  $SOC_{min}$  is given in Section 3.2.3.

To enable a fast and convenient implementation, a component-wise formulation of the optimization problem is used. Each component model consists of constraints defining the operation limits of a component and objective function terms for the optimization. The constraints and objectives are adjustable through a set of parameters and weights, respectively. The parameters and the weight factors are defined so that they can be easily read out of datasheets or defined by the operator without intense data analysis or control knowledge. Constraints that can, but should not be violated, are defined as soft constraints to ensure optimization robust-

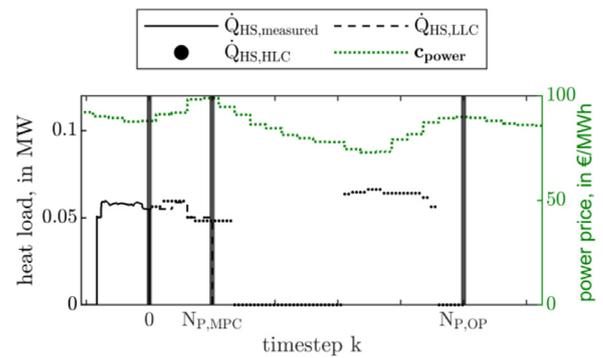


Fig. 4. Measured and predicted heat flow of the heat source  $\dot{Q}_{HS}$  and the power price  $c_{power}$ .

ness. The individual components are connected by so-called nodes that represent the required mass and energy balances. The modular component-wise definition of the optimization problem - from now on referred to as optimization models - enables a fast implementation to arbitrary energy supply systems. Optimization models can easily be extended with additional effects, and adaptations of the energy supply systems can be incorporated with minor modelling effort.

Fig. 4 shows an example of optimal trajectories calculated by the controller. The optimal trajectory of the plant input predicted by the HLC  $\dot{Q}_{HS,HLC}$  is shown by as dots, the predicted trajectory of the LLC  $\dot{Q}_{HS,LLC}$  as dashed line, the measured heat flow of the past time-steps as solid line and the flexible power price used in the optimization as green dotted line. Furthermore, the prediction horizons  $N_{p,HLC}$  and  $N_{p,LLC}$  are indicated. The switching behavior of the heat pump and the desired usage in times of low energy prices are evident.

In the following, the constraints and objectives are defined component-wise for heat pumps (as heat source), thermal energy storages and batch consumers. Finally, the global optimizer configuration is given. The optimization framework could easily be adapted and extended with further components like latent heat storages or gas vessels. Further, each component model could easily be extended with additional effects like heat loss. The constraints and objectives are formulated with the toolbox YALMIP (Lofberg, 2004).

### 3.1.2. Heat pump model

The heat pump (HP) transfers thermal energy from a cooler heat source to a warmer heat sink using the refrigeration cycle. The heat flows are calculated as follows:

$$\dot{Q}_{HP,sink} = (h_{HP,sink,out} - h_{HP,sink,in}) \cdot \dot{m}_{sink} \quad (1)$$

$$\dot{Q}_{HP,source} = (h_{HP,source,out} - h_{HP,source,in}) \cdot \dot{m}_{source} \quad (2)$$

The mass-specific enthalpy of the sink inflow  $h_{sink,in}$ , sink outflow  $h_{sink,out}$ , source inflow  $h_{source,in}$ , and source outflow  $h_{source,out}$  are calculated using the "CoolProps" physical property database (Bell et al., 2014).

The coefficient of performance  $COP$  is calculated as Carnot efficiency for full-load operation as given in (3) and calculated identically for minimal partial load:

$$COP_{full} = \frac{T_{HP,sink,out} + \Delta T_{HP,sink,full}}{T_{HP,sink,out} + \Delta T_{HP,sink,full} - T_{HP,source,out} + \Delta T_{HP,source,full}} \cdot \eta_{HP,full} \quad (3)$$

where  $\eta_{HP,full}$  is the compressor efficiency at full load,  $\Delta T_{HP,sink,full}$  is the temperature difference between HP working fluid and sink

fluid at full load and  $T_{HP, \text{sink}, \text{out}}$  the outlet temperature of the sink fluid. The same definitions are valid for the HP source. For all further operation points, a linearly interpolated COP is used to calculate the power demand:

$$COP = \alpha_{COP} + \beta_{COP} \cdot \frac{\dot{Q}_{HP, \text{sink}}}{Q_{HP, \text{sink}, k, \text{max}}}, \quad (4)$$

$$P_{HP} = \frac{u_{HP, \text{sink}}}{\alpha_{COP}} + \frac{\dot{Q}_{HP, \text{sink}}}{\beta_{COP}}, \quad (5)$$

where  $\alpha_{COP}$  and  $\beta_{COP}$  are linearization coefficients,  $\dot{Q}_{HP, \text{sink}, \text{max}}$  is the maximum sink heat flow,  $P_{HP}$  is the power consumption of the HP, and  $u_{HP}$  is an integer variable indicating operating ( $u_{HP} = 1$ ) or standstill ( $u_{HP} = 0$ ) condition. Furthermore, the following inequalities restrict the operating range to reasonable boundaries.

$$\begin{aligned} 0 &\leq \dot{Q}_{HP, \text{source}} \\ 0 &\leq \dot{Q}_{HP, \text{sink}} \\ 0 &\leq \dot{m}_{HP, \text{source}} \\ 0 &\leq \dot{m}_{HP, \text{sink}} \\ 0 &\leq P_{HP, \text{sink}} \end{aligned} \quad (6)$$

The energy balance is given as:

$$\dot{Q}_{HP, \text{sink}} = \dot{Q}_{HP, \text{source}} + P_{HP} \quad (7)$$

Eq. (8) connects the plant input  $U_{HP}$  with the sink heat flow  $\dot{Q}_{HP, \text{sink}}$  utilizing the scalar maximum sink heat flow  $\dot{Q}_{HP, \text{sink}, \text{max}}$ .

$$U_{HP} = \frac{\dot{Q}_{HP, \text{sink}}}{\dot{Q}_{HP, \text{sink}, \text{max}}} \quad (8)$$

The following introduction of the binary variables  $u_{HP}$ ,  $v_{HP}$  and  $w_{HP}$  is needed for operation constraints given in (10)-(16), e.g., minimum downtime. The binary decision variable  $u_{HP}$  indicates the operation condition,  $v_{HP}$  the start-up and  $w_{HP}$  the shutdown-events.

$$\begin{aligned} u_{HP} + w_{HP} &\leq 1 \\ u_{HP, k} - w_{HP, k} &= u_{HP, k} - u_{HP, k-1} \end{aligned} \quad (9)$$

Constraint (10) ensures the minimum uptime  $n_{SU}$  time-steps. Many components of energy supply systems demand minimal durations of operation conditions due to safety or wear reduction. Heat-pumps usually demand a minimal standstill duration of several minutes.

$$\sum_{j=1+n_{\text{max}}-n_{SU}}^{N_p+n_{\text{max}}-n_{SU}} u_{HP, t=j, \dots, j+n_{SU}} \leq u_{HP, t=j+n_{SU}} \quad (10)$$

In words, the condition requires that the sum of all startup events in the period of  $n_{SU}$  steps must be either 0 if the component is shut down at the end of the period, or a maximum of 1 if the component is in operation at the end of the period. This must be valid for all possible periods in the prediction horizon  $N_p$ . In the same way constraint (11) ensures the minimum standstill time-steps  $n_{SD}$ .

$$\sum_{j=1+n_{\text{max}}-n_{SD}}^{N_p+n_{\text{max}}-n_{SD}} w_{HP, t=j, \dots, j+n_{SD}-1} \leq 1 - u_{HP, t=j+n_{SD}} \quad (11)$$

The maximum heat flow at the heat sink  $\dot{Q}_{HP, \text{sink}, \text{max}}$ , maximum heat flow after start-up  $\dot{Q}_{HP, \text{sink}, \text{SUlim}}$  and before shutdown  $\dot{Q}_{HP, \text{sink}, \text{SDlim}}$  are ensured by constraint (12). Large energy supply units like gas turbines usually cannot operate with full power directly after start-up.

$$\begin{aligned} \dot{Q}_{HP, \text{sink}} &\leq \dot{Q}_{HP, \text{sink}, \text{max}} \cdot u_{HP, t=k, \dots, k+N_p} \\ &\quad - (\dot{Q}_{HP, \text{sink}, \text{max}} - \dot{Q}_{HP, \text{sink}, \text{SUlim}} \cdot v_{HP, t=k, \dots, k+N_p}) \\ &\quad - (\dot{Q}_{HP, \text{sink}, \text{max}} - \dot{Q}_{HP, \text{sink}, \text{SDlim}} \cdot [w_{HP, t=k+1, \dots, k+N_p}, 0]) \end{aligned} \quad (12)$$

The vector indicating shutdown-events  $w_{HP}$  is extended with a 0 to obtain dimension equality.

The initial condition is defined in (13):

$$\begin{aligned} u_{HP, t=k} &= u_{HP, 0} \\ U_{HP, t=k} &= U_{HP, 0} \end{aligned} \quad (13)$$

The minimal partial load is ensured in (14) utilizing the binary variable  $u_{HP}$  to enable shutdown events. Minimal part loads are typical among others for turbines and combined heat and power plants.

$$\dot{Q}_{HP, \text{sink}} \geq \dot{Q}_{HP, \text{sink}, \text{min}} \cdot u_{HP, t=k, \dots, k+N_p} \quad (14)$$

The ramp constraint is defined in (15). Especially large energy supply units have restricted ramp limits due to wear and safety reasons.

$$\begin{aligned} U_{HP, t=k, \dots, k+N_p} - U_{HP, t=k-1, \dots, k+N_p-1} \\ \leq RU_{HP} \cdot u_{HP, t=k, \dots, k+N_p} \\ + \frac{\dot{Q}_{HP, \text{sink}, \text{SUlim}} \cdot v_{HP, t=k, \dots, k+N_p}}{\dot{Q}_{HP, \text{sink}, \text{max}}} \end{aligned} \quad (15)$$

$$\begin{aligned} U_{HP, t=k+1, \dots, k+N_p+1} - U_{HP, t=k, \dots, k+N_p} \\ \leq RD_{HP} \cdot u_{HP, t=k, \dots, k+N_p} \\ + \frac{\dot{Q}_{HP, \text{sink}, \text{SDlim}} \cdot w_{HP, t=k, \dots, k+N_p}}{\dot{Q}_{HP, \text{sink}, \text{max}}} \end{aligned} \quad (16)$$

Where  $RU_{HP}$  and  $RD_{HP}$  are parameters defining the slope of the ramp. The second term of the right hand side is needed to enable higher changes in the utilization at start-up or shut-down events.

The objective function of the heat pump is given in (17) where  $C$  are cost vectors with length  $N_p$  and  $t_s$  is the sampling time. The cost function considers costs of the electric power  $C_{\text{power}}$ , costs of changes in the utilization to ensure a smooth operation of the heat pump  $C_{\Delta U}$ , costs to penalize starting maneuvers  $C_{SU}$  and shut down events  $C_{SD}$  and, costs for deviation of the utilization from the calculated trajectory  $C_{\text{traj}}$ .

$$\begin{aligned} C_{HP} &= C_{\text{power}} + C_{\Delta U} + C_{SU} + C_{SD} + C_{\text{traj}} \\ C_{\text{power}} &= \sum_{j=k}^{k+N_p} c_{\text{power}, j} \cdot P_j \cdot \frac{t_s}{60} \cdot c_{\text{power}, \text{factor}} \\ C_{\Delta U} &= \sum_{j=k}^{k+N_p-1} \text{abs}(U_{HP, j+1} - U_{HP, j}) + \text{abs}(U_{HP, t=k} - U_{HP, 0}) \cdot C_{\Delta U, \text{factor}} \\ C_{SU} &= \sum_{j=k}^{k+N_p} v_j \cdot C_{SU, \text{factor}} \\ C_{SD} &= \sum_{j=k}^{k+N_p} w_j \cdot C_{SD, \text{factor}} \\ C_{\text{traj}} &= \sum_{j=k}^{k+N_p} \text{abs}(U_{HP, j} - U_{HP, \text{traj}}) \cdot c_{\text{traj}, \text{factor}} \end{aligned} \quad (17)$$

Note that absolute values in the objective function have to be reformulated to retain linearity. One linearization method is displayed in (18) (Matthew et al., 2020).

$$\text{abs}(X) = X' \Leftrightarrow \begin{cases} X \leq X' \\ -X \leq X' \end{cases} \quad (18)$$

The parameters which have to be defined for the heat pump model are listed in [Appendix A \(Table 6\)](#) and are basic component parameters usually known by plant operators. The cost factors  $c$  are further discussed in [Section 3.1.5](#).

### 3.1.3. Thermal energy storage model

The following constraints ([19–23](#)) define the operation limits of sensible thermal energy storage (TES) with sampling time  $t_s$ . First the maximum and minimum stored sensible heat is defined utilizing the maximum operation temperature  $T_{\max}$  and minimum operation temperature  $T_{\min}$ .

$$\begin{aligned} Q_{\text{TES,max}} &= T_{\max} \cdot c_p \cdot V \cdot \rho \\ Q_{\text{TES,min}} &= T_{\min} \cdot c_p \cdot V \cdot \rho \end{aligned} \quad (19)$$

Where  $c_p$  is the mass specific heat coefficient  $V$  the Volume of the tank and  $\rho$  the density of the storage fluid. The specific-heat coefficient  $c_p$  is calculated using the "CoolProps" physical property database ([Bell et al., 2014](#)).

The [Eqs. \(20\)](#) and [\(21\)](#) define the course of the heat stored in the TES considering the SOC dependent losses with parameter  $\epsilon_{\text{SOC}}$ , the fix losses with parameter  $\epsilon_{\text{fix}}$ , and the charging and discharging heat flows, respectively.

$$\begin{aligned} Q_{\text{TES},j=k} &= Q_{\text{TES},0} \cdot \left(1 - \epsilon_{\text{SOC}} \cdot \frac{t_s}{60}\right) - \epsilon_{\text{fix}} \cdot \frac{t_s}{60} \\ &+ \dot{Q}_{\text{TES},C,t=k} \cdot \frac{t_s}{60} - \dot{Q}_{\text{TES},D,t=k} \cdot \frac{t_s}{60} \end{aligned} \quad (20)$$

$$\begin{aligned} Q_{\text{TES},j=k+1,\dots,k+N_p} &= Q_{\text{TES},j=k,\dots,k+N_p-1} \cdot \left(1 - \epsilon_{\text{SOC}} \cdot \frac{t_s}{60}\right) - \epsilon_{\text{fix}} \cdot \frac{t_s}{60} \\ &+ \dot{Q}_{\text{TES},C,j=k+1,\dots,k+N_p} \cdot \frac{t_s}{60} - \dot{Q}_{\text{TES},D,j=k+1,\dots,k+N_p} \cdot \frac{t_s}{60} \end{aligned} \quad (21)$$

In [\(22\)](#) the state of charge (SOC) of the TES is defined. The SOC is defined between 0 - where the storage temperature  $T_{\text{TES}} = T_{\min}$  and therefore no heat can be discharged from the TES - and 1, where  $T_{\text{TES}} = T_{\max}$  and therefore the TES cannot be charged further.

$$\text{SOC} = \frac{Q_{\text{TES}} - Q_{\text{TES,min}}}{Q_{\text{TES,max}} - Q_{\text{TES,min}}} \quad (22)$$

To ensure production safety, the online load predictor (OLP) defines a maximal SOC  $\text{SOC}_{\max}$  and a minimal SOC  $\text{SOC}_{\min}$ . Although violations of these constraints should be avoided at all costs, they can occur, for example, in the case of prediction errors. These violations could lead to the infeasibility of the optimization problem. To avoid infeasibilities during optimization the constraint is implemented as a soft constraint. Violations of soft constraints trigger costs multiple magnitudes higher than all other cost terms. As a result, all measures are taken to avoid or minimize violations.

$$\begin{aligned} \text{SOC} - \mathbf{S}_{\text{TES}} &\leq \text{SOC}_{\max} \\ \text{SOC} + \mathbf{S}_{\text{TES}} &\leq \text{SOC}_{\min} \\ \mathbf{S} &\geq 0 \end{aligned} \quad (23)$$

$$C_{\text{TES}} = \mathbf{S}_{\text{TES}} \cdot c_{\text{S, factor}} \quad (24)$$

The slack constraint is the only contribution of the TES to the objective function [\(24\)](#) where  $c_{\text{S, factor}}$  is the slack cost factor and  $\mathbf{S}_{\text{TES}}$  the slack variable.

The parameters which have to be defined for the TES model are listed in [Appendix A \(Table 7\)](#) and are basic component parameters usually known by plant operators.

### 3.1.4. Batch consumer model

The batch consumer (BC) is defined by the enthalpy balance and the imposed heat demand [\(24\)](#).

$$\begin{aligned} \dot{Q}_{\text{BC}} &= (h_{\text{BC,out}} - h_{\text{BC,in}}) \cdot \dot{m}_{\text{BC}} \\ \dot{Q}_{\text{BC}} &= \dot{Q}_{\text{BC,demand}} \\ \dot{m}_{\text{BC}} &\geq 0 \end{aligned} \quad (25)$$

The mass-specific enthalpy of the inflow  $h_{\text{BC,in}}$  and outflow  $h_{\text{BC,out}}$  are calculated using the "CoolProps" physical property database ([Bell et al., 2014](#)). The BC does not affect the objective function. The parameters which have to be defined for the BC model are listed in [Appendix A \(Table 8\)](#) and are basic component parameters usually known by plant operators.

The BC optimization model considers no heat loss or conversion rate as it maps all significant effects of the use case presented in [Section 4](#). Additional effects like heat losses or conversion rates could easily be implemented in the model similar to the TES or HP model.

### 3.1.5. Nodes

Nodes represent the required mass and energy balances to connect the single components. To increase numerical stability, these nodes are implemented as soft constraints triggering a warning in case of balance residuals higher than numerical deviations caused by the optimization. For the examined use case, the overall enthalpy balance [\(26\)](#) is sufficient to connect all components:

$$\begin{aligned} \dot{Q}_{\text{HP,sink}} + \dot{Q}_{\text{TES,D}} + \mathbf{S}_{\text{lhs}} &= \sum_{m=1}^4 \dot{Q}_{\text{BC},m} + \dot{Q}_{\text{TES,C}} + \mathbf{S}_{\text{rhs}} \\ \mathbf{S}_{\text{lhs}} &\geq 0 \\ \mathbf{S}_{\text{rhs}} &\geq 0, \end{aligned} \quad (26)$$

where  $\mathbf{S}_{\text{lhs}}$  and  $\mathbf{S}_{\text{rhs}}$  are slack variables. The slack constraint [\(27\)](#) is the only contribution to the cost function where  $c_{\text{S, factor}}$  is the slack cost factor.

$$C_{\text{Nodes}} = (\mathbf{S}_{\text{lhs}} + \mathbf{S}_{\text{rhs}}) \cdot c_{\text{S, factor}} \quad (27)$$

### 3.1.6. Assembling of the constraints and objectives

For holistic optimal control of the ESS, the constraints and objectives of the components and nodes have to be combined to one single MILP. Due to the component-wise structure and the connection nodes, the overall optimization problem is given by the summation of all component cost functions and stacking of all component constraints. This minimizes the implementation effort. Changes in the ESS, such as the installation of a photovoltaic system or an expansion of production can be incorporated into the EMS with little effort. For each component the respective component constraints and objectives have to be added to the optimization and the nodes adapted accordingly.

A remaining implementation effort is the correct choice of the weight factors  $c$  and optimization parameters listed in [Table 1](#). To define parameters efficiently, the following rules have been found useful:

- To obtain an optimization problem formulation which is simple to interpret and plausible to adjust in the individual objective weightings, an a-priori normalization of the different objective terms in mandatory.
- Monetary factors are intuitive for plant operators, and the power price is defined externally. Therefore, the power price  $c_{\text{power}}$  can be used for normalization.
- Start-up and shut-down costs should include potential wear and tear for these events, as well as labor costs and additional fuel costs.

**Table 1**  
Weight factors and optimization parameters which need to be defined for each controller.

Weights	Description
$C_{\text{power, factor}}$	Electric power cost weight to reduce power costs
$C_{\Delta U, \text{ factor}}$	Cost weight to smoothen the utilization of the HP
$C_{\text{SU, factor}}$	Cost weight to consider costs triggered by start-up events
$C_{\text{SD, factor}}$	Cost weight to consider costs triggered by shut-down events
$C_{\text{traj, factor}}$	Cost weight to penalize deviations from a given trajectory
$C_{\text{slack}}$	Cost weight to avoid critical operation conditions
Opt. Params.	Description
$N_P$	Prediction horizon of the optimizer
$t_s$	Sampling time in min

- For many components, the definition of a minimal uptime or downtime as constraints (see (10) and (11)) is sufficient.
- The slack cost should be orders of magnitude higher than all other cost terms.
- To fully exploit the flexibility of the power market, a minimal prediction horizon of one day is needed to utilize the day-ahead market.
- The billing of power consumption is usually based on records at 15 min intervals. Therefore, a sampling time below 15 min is needed to effectively react to disturbances on the power consumption.

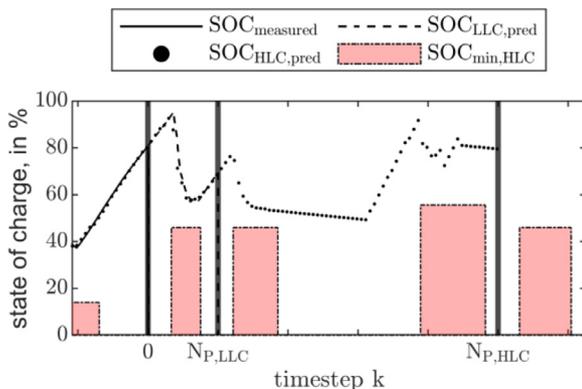
Detailed rules for the optimal weighting of the cost function are currently investigated by the authors.

### 3.2. Online load predictor

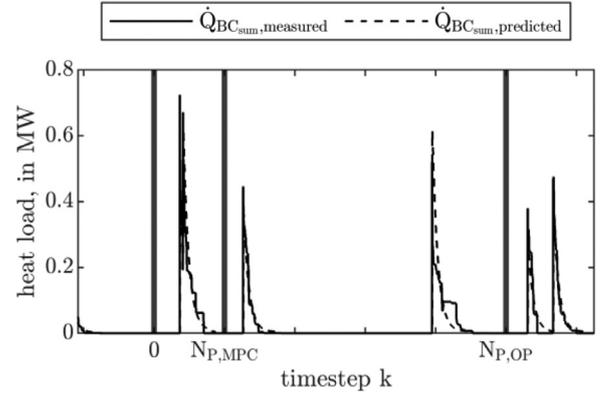
The online load predictor utilizes the production schedule to estimate the heat load of future heat treatments (HT), reacts to measured deviations of the actual heat load and defines the time-dependent minimal state of charge  $\text{SOC}_{\text{min}}$  online in every single time step. These three functionalities are described in the following Sections. They ensure process reliability while simultaneously maximizing the flexibility of the EMS. Further, the formulation of the  $\text{SOC}_{\text{min}}$  avoids the typical nonlinearities in the MILP-formulation of thermal batch processes system. In Fig. 5 the calculated  $\text{SOC}_{\text{min}}$  is visualized and Fig. 6 shows estimated and measured heat loads of five heat treatments.

#### 3.2.1. Heat load estimation

The utilized estimation method for pulse-like heat loads is presented in Fuhrmann et al. (2020). The method can be implemented in a straightforward way because historical data from few measurement points are sufficient for parameterization and it is robust against measurement noise. The method estimates the total



**Fig. 5.** Measured and predicted state of charge (SOC) of the thermal energy storage and the critical state of charge  $\text{SOC}_{\text{crit}}$ .



**Fig. 6.** Measured and predicted heat load of all batch consumers  $\dot{Q}_{\text{BC, sum}}$ . Adapted from (Matthew et al., 2020).

heat load of the batch consumers  $\dot{Q}_{\text{BC, sum}}$  by estimating the heat load of each single HT  $\dot{Q}_{\text{HT, m}}$ . The resulting heat load prediction is shown in Fig. 6. During operation, the method needs the estimated starting temperature  $\hat{T}_{\text{HT, 0, m}}$  and end temperature  $\hat{T}_{\text{HT, end, m}}$  of the HT, which are typically part of the production schedule as they are crucial for a successful HT. Furthermore, the planned starting time  $\hat{t}_{0, \text{HT, m}}$  is necessary for the prediction of the total heat load of the batch consumers  $\dot{Q}_{\text{BC, sum}}$ .

#### 3.2.2. Prediction error compensation

Heat loads caused by batch processes typically show two kinds of deviations from the predicted heat load: deviations of the starting time of the HT  $t_{0, \text{HT, m}}$  caused by the manual starting procedure and deviations of the time course of the heat flow  $\dot{Q}_{\text{BC, sum}}$  caused by deviating heat conductivity of the products. In contrast, the prediction error of the integral heat amount needed for a HT  $Q_{\text{HT, m}}$  is usually negligible as it is dependent on the well-known parameters starting temperature  $T_{\text{HT, 0, m}}$ , end temperature  $T_{\text{HT, end, m}}$ , and heat capacity of the product  $C_{\text{HT, m}}$ . These circumstances are utilized by the online heat load predictor, which executes three correction measures.

The first measure is actually a prevention measure. The estimated starting time of all HT  $\hat{t}_{0, \text{HT, m}}$  is shifted forward by a maximum starting time deviation  $\Delta t_{\text{start, max}}$  to the earliest possible start of the HT. Together with the later introduced  $\text{SOC}_{\text{min}}$ , this ensures that the heat supply is sufficient at all possible starting times. Thereby, the control system is robust against time shifts in the production schedule smaller than  $\Delta t_{\text{start, max}}$ . The parameter  $\Delta t_{\text{start, max}}$  can either be defined by the plant operator or calculated from historical measurement data.

While the first measure avoids bottlenecks caused by HT starting sooner than expected, the second measure counteracts delays of HT. The heat load predictor shifts the heat load prediction of a HT  $\dot{Q}_{\text{HT, m}}$  backward by one time-step in case the HT does not start at its predicted starting time  $\hat{t}_{0, \text{HT, m}}$ .

Thirdly, the prediction of the heat flow trajectory is corrected at each time step. The integrated difference between measured and predicted heat flow is distributed to the remaining load prediction utilizing:

$$\dot{Q}_{\text{HT, m, t=k, \dots, t}_{\text{HT, end}}} = \dot{Q}_{\text{HT, m, t=k, \dots, t}_{\text{HT, end}}} + \frac{\sum_{j=\hat{t}_{\text{HT, 0, m}}}^k \dot{Q}_{\text{HT, m, j}} - \dot{Q}_{\text{HT, m, j}}}{t_{\text{HT, end}} - k}, \quad (28)$$

where  $k$  is the current time step.

### 3.2.3. Calculation of $SOC_{min}$

The online load predictor (OLP) calculates a time variant minimal state of charge  $SOC_{min}$  for energy storages which is incorporated as constraint to the MPC. The  $SOC_{min}$  is recalculated each time-step allowing a fast and optimal reaction to disturbances. This ensures process reliability while simultaneously maximizing the flexibility of the EMS. Further the formulation of the  $SOC_{min}$  as enthalpy level instead of temperature allows a consideration of different demand temperatures and avoids the typical nonlinearities of the system. The purpose of the  $SOC_{min}$  is to avoid bottlenecks in the heat supply. Bottlenecks in the heat supply occur when the temperature difference at the heat exchanger of the batch consumers  $\Delta T_{BC,n}$  is too small to provide the necessary heat flow within the heat-exchanger. To avoid bottlenecks, the  $SOC_{min}$  is defined as follows:

$$SOC_{min} = \frac{(T_{HT,max,end} + \Delta T_{BC,desired}) \cdot c_p \cdot V \cdot \rho + \Delta Q_{robust} - Q_{TES,min}}{Q_{TES,max} - Q_{TES,min}} \quad (29)$$

where  $T_{HT,max,end}$  is the vector with the length  $N_p$  consisting of the maximum end temperature of a HT  $T_{HT,end,m}$  occurring at each time-step up to the control horizon,  $\Delta T_{BC,desired}$  is the desired minimal temperature difference at the heat exchangers and  $\Delta Q_{robust}$  is a safety margin to increase robustness against prediction errors. The robustness of the EMS is limited to realizable production schedules. The EMS cannot prevent a violation of  $SOC_{min}$  when the heat demand is higher than the ESS can provide. This could only be avoided by a scheduling algorithm. Due to the high implementation cost and low acceptance of scheduling in some manufacturing plants the EMS presented in this manuscript does not include a scheduling algorithm but assumes the production schedule to be given.

Given a realizable production schedule the EMS is robust for all prediction errors smaller than  $\Delta Q_{robust}$ . The  $SOC_{min}$  is calculated separately for each optimizer with different values for  $\Delta Q_{robust}$ , where  $\Delta Q_{robust,HLC}$  is always larger than  $\Delta Q_{robust,LLC}$ . This difference mitigates the effect of uncertainties on EMS performance. Reactions of the EMS uncertainties are only executed when the production safety is endangered. In both HLC and LLC, the  $SOC_{min}$  is implemented as a slack constraint to avoid infeasibility in case of constraint violations.

The  $SOC_{min}$  ensures maximum flexibility of the optimization as no minimal state of charge is demanded when no HT is active while ensuring robustness against uncertainties in the integral heat demand smaller than  $\Delta Q_{robust}$ .

There are two challenges which can cause an undershooting of  $SOC_{min}$ . First, a production schedule can be unenforceable when it demands a heat load which is too high to be provided by the given ESS. Including scheduling to the EMS would systematically exclude the possibility of infeasible production schedules. The EMS pro-

posed in this paper can detect unenforceable production schedules by checking for violations of  $SOC_{min}$  in the  $SOC$  prediction. When a significant violation is detected the EMS can display a warning to the operator.

The second reason for an undershooting of the  $SOC_{min}$  are large prediction errors of the heat loads. By including a safety margin  $\Delta Q_{robust}$  and by using the heat load prediction method presented in Fuhrmann et al. (2020) the likelihood of infeasibilities is decreased.

To avoid mathematical infeasibilities in the optimizer the  $SOC_{min}$  constraints are implemented as soft constraints. Thereby the MPC minimizes unpreventable violations of  $SOC_{min}$ .

## 4. Case study of an industrial food plant

The presented methods can be used for the optimal predictive control of the energy supply for all kinds of thermal batch processes (e.g. annealing, tempering, pasteurization). Also, every type of heat source can be integrated into the optimization with the according optimization model. In the case study presented in this section, an industrial food plant is considered. The plant manufactures meat products that undergo specific temperature trajectories to alter the flavor and structure of the meat and extend the expiration date. These heat treatments are typical batch processes. Therefore, the use case is suitable to study the performance of EMS, and the results can be easily transferred to other batch processes. The performance of the novel EMS is compared to the installed baseline controller in a simulation study. Therefore, simulation models of the industrial plant were developed and validated by industrial measurement data.

The simulation study was executed utilizing Matlab-Simulink® as co-simulation platform. The EMS was implemented as Matlab-Function-block, the optimization problems were defined utilizing YALMIP, and Gurobi® was used as a solver (Lofberg, 2004; Gurobi, 2018).

### 4.1. Simulation model

The structure of the plant and the installed sensors are displayed in Fig. 7. A list of all sensors can be found in Table 9 in the Appendix. The plant is comprised of a constant heat source, a heat pump (HP) as heat source, a thermal energy storage (TES) and four batch consumers (BC). The constant heat source - a continuous heat recovery system - has no effect on the energy management and is thus neglected in the optimization problem.

The simulation model considers additional effects and nonlinearities compared to the optimization models and consists of multiple component models. The TES model considers nonlinear mixing effects and free convection and was developed utilizing Matlab®. The simulation model of the BC takes nonlinear temperature-gradient-dependent effects into account. The heat-pump model

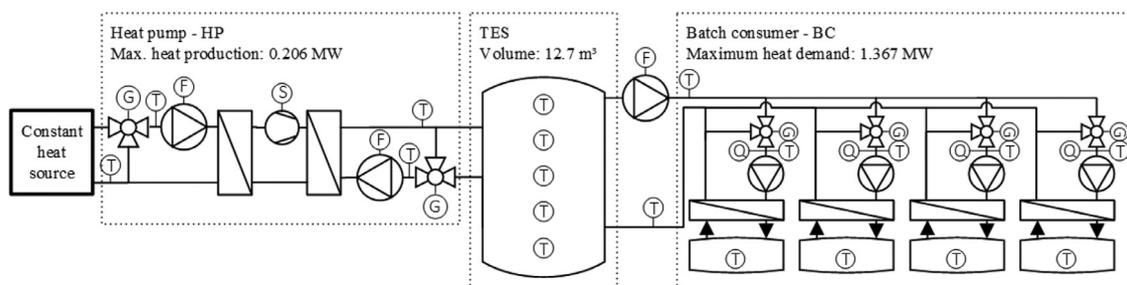
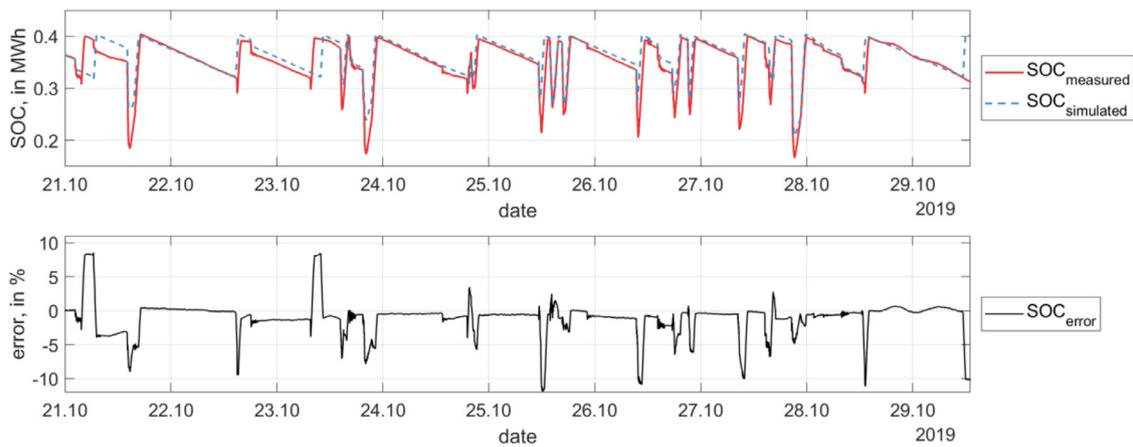


Fig. 7. Structure and sensors of the food production plant considered in the use case consisting of a constant heat source, a heat pump (HP), a thermal energy storage (TES) and four batch consumer (BC). The sensors measure temperature (T), volume flow (F), frequency (S), or valve position (G).



**Fig. 8.** The top graph displays the measured state of charge (SOC) from real measurement data and the simulated SOC calculated utilizing the simulation model. The lower graph displays the relative error of the simulated SOC. Adapted from Gurobi (2018).

was built using Dymola, considering underlying controllers and a model of the vapor-compression cycle. Furthermore, the hysteresis controller utilized in the real industrial plant was replicated for the simulations and serves as a baseline. Simulink was chosen as a co-simulation platform.

The simulation model of the energy system was validated with measurement data from the industrial use case as part of a diploma thesis (Sack, 2021). To illustrate the model accuracy, Fig. 8 shows the measured and modelled state of charge (SOC) and the relative model error for ten production days. The SOC of the thermal storage is the most informative measurement for model validation as it integrates all possible errors and displays possible trends. The measured heat demand of the BC was used as single input in this simulation study. The HP utilization was simulated using the hysteresis controller model. It is evident that short-term deviations occur during transient events, but the model overall shows sufficient accuracy to quantify the EMS performance.

#### 4.2. Design of the case study

To achieve qualitative and quantitative valid results of the EMS performance, a simulation duration of one month, including 102 heat treatments, was chosen. The production schedule, heat load and weather data were taken from actual industrial measurement data. The EMS is compared with two other controllers: First, the original hysteresis controller, which is currently used in the industrial plant. Second, an optimized hysteresis controller with optimized parameters to avoid bottlenecks in the heat supply. The weights of the cost function were determined according to the specifications of the industrial plant operator, taking into account the rules presented in Section 3.1.5. The goal was to maximize cost reduction  $C_{Power,red}$  while keeping the number of short starts  $n_{SU,short}$  within the admissible bound of one per day. To demonstrate the potential effect of different weights, the simulation was repeated with a strongly increased weight on start-up operations  $C_{SU,factor}$ . All optimization parameters of the controllers can be found in Appendix A (Table 5).

The control performance assessment is done utilizing the energy costs reduction  $C_{Power,red}$  and the number of heat treatments affected by bottlenecks in the heat supply  $n_{HT,affected}$  for every control method.

To investigate the robustness of the control method and the performance of the online load predictor (OLP), the simulation is repeated with different load predictions. First, in the case of perfect information (PI) the applied heat load  $\dot{Q}_{BC,sum}$  is used as per-

**Table 2**

Performance indicators used in the simulation study.

Parameter	Description
$C_{Power,red}$	Power cost reduction compared to the hysteresis control
$n_{SU,short}$	Number of start-ups shorter than $t_{SU,desired}$
$n_{SD,short}$	Number of shut-downs shorter than $t_{SD,desired}$
$\Delta U_{HP,mean}$	Average change of plant input $U_{HP}$

fect load prediction. In the further experiments, the prediction model presented in Section 3.2 is utilized for load prediction. To further investigate the robustness against prediction errors of the starting time of the heat treatments,  $t_{0,HT,m}$  was altered using random time shifts:

$$t_{0,m} = t_{0,m} + t \cdot \text{ceil}(\Delta N_{Shift,max} \cdot U(-1, 1)) \quad (30)$$

where  $U(-1, 1)$  is a uniform distribution in the interval  $[-1, 1]$  and  $\Delta N_{Shift,max}$  is the maximal deviation of the starting time. As a maximum time deviation of 30 min is usual in the considered industrial use case,  $\Delta N_{Shift,max}$  was altered between 0 and 30 in the simulations.

In Appendix A, the optimization parameters used in the simulation study are listed. Table 2 gives descriptions of the performance indicators used to quantify the performance of the controller in the simulation study.

#### 4.3. Results of the case study

The simulation study allows a qualitative and quantitative assessment of the performance of the proposed EMS for batch factories. The results are shown in Table 3. The hysteresis controller with safe settings avoids potentially unacceptable affected heat treatments in trade-off to 0.3% higher energy cost. The suggested two-layer EMS can successfully prevent bottlenecks in the heat supply, while simultaneously reducing power costs by 9% com-

**Table 3**

Comparison of different control strategies.

	$C_{Power,red}$	$n_{HT,affected}$	$n_{SU}$	$n_{SU,short}$
Hysteresis Control				
Installed	-	7	73	0
Safe settings	-0.3%	0	74	0
EMS	9.0%	0	85	17
EMS ( $C_{SU,factor} \cdot 100$ )	3.52%	0	44	7

**Table 4**

Performance of the EMS with activated and deactivated OLP for different load predictions (PI: perfect information).

OLP	Pred.	$C_{\text{Power,red}}$	$n_{\text{HT,affected}}$	$n_{\text{SU}}$	$n_{\text{SU,short}}$
on	PI	9,0%	0	85	2
	$\Delta t = 0$	9,0%	0	85	2
	$\Delta t = 10$	9,0%	0	91	6
	$\Delta t = 30$	8,5%	0	106	24
off	PI	5,3%	0	234	155
	$\Delta t = 0$	5,3%	0	234	155
	$\Delta t = 10$	4,65%	0	236	162
	$\Delta t = 30$	4,22%	0	240	165

pared to the installed hysteresis controller. The number of start-ups of the HP  $n_{\text{HP,SU}}$  increased by only 16%. In case fewer start-ups are desired, the weights of the cost function can easily be adapted accordingly. This was demonstrated with a strongly increased weight on start-up operations  $c_{\text{SU, factor}}$ , where the cost reduction  $C_{\text{Power,red}}$  decreased to 3.52% but also the number of start-up operations was reduced by 66% compared to the hysteresis controller.

Next, the performance of the online load predictor is analyzed. The results of simulations with activated and deactivated OLP and for different load predictions are given in Table 4. The results show that the OLP increases the performance of the EMS significantly in all cases. Only 0.5% of the energy consumption are lost for the maximum deviation of starting times  $\Delta t = 30$  min, the number of start-ups  $n_{\text{SU}}$  is increased by 25.7%. Deviations in the heat load  $\dot{Q}_{\text{BC,sum}}$  from the prediction are neglected by the EMS because of the knowledge of the certain integral heat amount of a heat treatment  $Q_{\text{HT}}$ . Thereby the operation of the HP remains more steady. The number of short start-ups of the HP  $n_{\text{SU,short}}$  increases strongly but remains a magnitude smaller than without OLP. These start-ups are especially disadvantageous as a HP is inefficient during start-up maneuvers.

Without OLP, the benefit of the EMS is decreased strongly. Even in the case of perfect information (PI), the model errors of the optimization models leads to an unsteady and inefficient operation of the HP. Furthermore, it is striking, that in both cases – with and without the OLP – the performance of the EMS is nearly equal for the case PI where the measurement data is used for heat load prediction and the case  $\Delta t = 0$  where the model is used for heat load prediction. This shows that the heat load model developed in Fuhrmann et al. (2020) is very accurate and suitable for disturbance prediction.

## 5. Conclusions

In this paper, a novel energy management system (EMS) for thermal batch production processes using an online load predictor (OLP) is proposed and characterized. The formulation of the arising mixed-integer linear program presented in this paper enables system integrators to implement an EMS for batch processes with little effort spent on modelling tasks and parameter tuning. Furthermore, the EMS avoids bottlenecks in the energy supply and thereby tackles the two major obstacles of EMS in industry – implementation cost and reduction of production reliability. By implementing the EMS, existing control structures for energy supply systems can be retrofitted to meet modern requirements like demand side control. The straightforward implementation is especially striking for small factories without access to control experts. The method utilizes the expert knowledge of the plant operators to increase performance and acceptance and focuses on the balance between effort and performance.

The presented case study based on validated simulation models of the plant clearly shows the benefits of the novel EMS. The results can be easily transferred to other thermal batch processes (e.g. annealing, tempering, pasteurization), since the typical characteristics of batch processes are strongly pronounced in the considered plant.

Still there are drawbacks to the presented EMS. The implementation costs are higher compared to heuristic solutions. On the other hand the EMS has decisive advantages. First, for complex energy supply systems, heuristics are not straightforward to implement, while the modular structure of the EMS enables a fast and convenient extension or adaption to arbitrary energy supply systems. Furthermore, the formulation allows the consideration of multiple predictions like electricity price, weather and production schedule as well as multiple objectives.

Further research is currently carried out with the presented EMS. The EMS is applied to larger and more complex energy supply systems, including steam systems and cooling circuits. To further facilitate the implementation of the EMS, rules for the weighting of the cost function and setting of optimization parameters are elaborated. Also the robustness of the optimization against failures of sensors of components is currently being increased to allow an implementation of the EMS at the industrial plant.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRedit authorship contribution statement

**Florian Fuhrmann:** Conceptualization, Methodology, Software, Validation, Writing – original draft. **Alexander Schirrer:** Conceptualization, Writing – review & editing, Supervision. **Martin Kozek:** Writing – review & editing, Supervision, Funding acquisition.

## Acknowledgments

The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Program. This work was supported by the project 'EDCSproof', which is part of the energy model region NEFI – New Energy for Industry and is funded by the Austrian Climate and Energy Fund [FFG, No.868837]

## Appendix A. Parameters of the EMS

Parameters of the EMS.

**Table 5**

Weights and optimization parameters (opt.par.) of the higher-level control (HLC) and lower-level control (LLC).

Weights	Description	HLC	LLC
$c_{\text{power, factor}}$	Power cost weight	1	0
$c_{\text{dU, factor}}$	$\Delta U$ cost weight	0,62	0
$c_{\text{SU, factor}}$	Start-up cost weight	20,7	0
$c_{\text{SD, factor}}$	Shutdown cost weight	20,7	0
$c_{\text{traj, factor}}$	Trajectory cost weight	0	69
$c_{\text{slack}}$	Slack cost weight	69000	69000
Opt. Par.	Description	HLC	LLC
$N_p$	Prediction horizon	96	60
$t_s$	sampling time in min	15	1

**Table 6**  
Parameters of the heat pump model.

Parameter	Value	Description
$\dot{Q}_{HP,sink,min}$	0.2	Minimum part load heat flow at heat sink in MW
$\dot{Q}_{HP,sink,max}$	0.5	Maximum heat flow at heat sink in MW
$RU_{HP}$	10	Maximum ramp-up rate in %
$RD_{HP}$	10	Maximum ramp down rate in %
$n_{HP,SU}$	15	Minimum uptime in steps
$n_{HP,SD}$	15	Minimum downtime in steps
$n_{HP,max}$	15	Maximum length of logic constraints to facilitate notation of constraints
$\dot{Q}_{HP,sink,SUlim}$	0.2	Maximum heat generation after start-up/before shutdown in MW
$\dot{Q}_{HP,sink,SDlim}$	0.5	Maximum heat generation before shutdown in MW
$T_{HP,sink,out}$	393.15	Outlet temperature at sink in K
$T_{HP,sink,in}$	433.15	Inlet temperature at sink in K
$T_{HP,source,out}$	353.15	Outlet temperature at source in K
$T_{HP,source,in}$	363.15	inlet temperature at source in K
$\Delta T_{HP,sink,full}$	10	Temperature difference between HP working fluid and sink fluid at full load in K
$\Delta T_{HP,sink,part}$	5	Temperature difference between HP working fluid and sink fluid at full load in K
$\Delta T_{HP,source,full}$	10	Temperature difference between HP working fluid and source fluid at full load in K
$\Delta T_{HP,source,part}$	5	Temperature difference between HP working fluid and source fluid at part load in K
$\eta_{HP,part}$	0.7	Compressor efficiency at part load
$\eta_{HP,full}$	0.8	Compressor efficiency at full load
$p_{HP,sink}$	$3 \cdot 1e5$	Operating pressure at heat sink in Pascal
$p_{HP,source}$	$1 \cdot 1e5$	Operating pressure at heat sink in Pascal

**Table 7**  
Parameters of the storage model.

Parameter	Value	Description
$T_{TES,max}$	373.15	Maximum temperature in K
$T_{TES,min}$	273.15	Minimum temperature in K
$\eta_{TES,C}$	1	Charging efficiency
$\eta_{TES,D}$	1	Discharging efficiency
$\epsilon_{TES,SOC}$	5	Loss proportional to SOC per hour in $h^{-1}$
$\epsilon_{TES,fix}$	15	Fix loss per hour in $MWh/h$
$SOC_{max}$	100	Maximum state of charge in %
$c_p$	$4196 \cdot 1e^{-6}$	Specific heat capacity in $MJ/kgK$
$V_{TES}$	12.7	Volume in $m^3$
$\rho_{TES}$	971.8	Density in $kg/m^3$

**Table 8**  
Parameters of the batch consumer model.

Parameter	Value	Description
$T_{BC,out}$	353.15	Inlet temperature in K
$T_{BC,in}$	363.15	Outlet temperature in K
$p_{BC}$	$3 \cdot 1e5$	Operating pressure in Pascal

**Table 9**  
List of all measurement points of the use case.

Component	Type	Description
HP	S	Frequency of the compressor
HP	T	Source inflow temperature
HP	G	Source mixture valve position
HP	T	Source outflow temperature
HP	F	Mass-flow in the source
HP	T	Sink inflow temperature
HP	G	Sink mixture valve position
HP	T	Sink outflow temperature
HP	F	Mass-flow in the sink
TES	T	Temperature measurements in five different heights
-	F	Mass-flow from the TES to the BC
-	T	T of the water from the TES to the BC
-	T	T of the water from the BC to the TES
BC	T	Temperatures of the four BC
BC	Q	Heat flows transferred to the four BC
BC	G	Valve positions of the four BC
BC	T	Temperatures of the inflow to the BC.

**References**

Bell, I.H., Wronski, J., Quoilin, S., Lemort, V., 2014. Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library coolprop. *Ind. Eng. Chem. Res.* 53 (6), 2498–2508.

Bemporad, A., Morari, M., 1999. Robust model predictive control: a survey. *Robustness Identif. Control* 245, 207–226.

Beykal, B., Avraamidou, S., Pistikopoulos, E.N., 2022. Data-driven optimization of mixed-integer bi-level multi-follower integrated planning and scheduling problems under demand uncertainty. *Comput. Chem. Eng.* 156, 107551.

Biswas, A., Emadi, A., 2019. Energy management systems for electrified powertrains: state-of-the-art review and future trends. *IEEE Trans. Veh. Technol.* 68 (7), 6453–6467.

C. Matthew, Y. Yin, B. Amada, P. Williams, and D. Xiao, (2020) "Optimization with absolute values," *Cornell University Computational Optimization Open Textbook*. [Online]. Available: [https://optimization.cbe.cornell.edu/index.php?title=Optimization\\_with\\_absolute\\_values](https://optimization.cbe.cornell.edu/index.php?title=Optimization_with_absolute_values).

Dengiz, T., Jochem, P., Fichtner, W., 2019. Uncertainty handling control algorithms for demand response with modulating electric heating devices. In: *Proc. 2019 IEEE PES Innov. Smart Grid Technol. Eur. ISGT-Europe 2019*, pp. 1–5.

Dengiz, T., Jochem, P., Fichtner, W., 2021. Demand response through decentralized optimization in residential areas with wind and photovoltaics. *Energy* 223 (42).

Dullinger, C., Struckl, W., Kozek, M., 2018. A general approach for mixed-integer predictive control of HVAC systems using MILP. *Appl. Therm. Eng.* 128, 1646–1659.

Fluch, J., et al., 2017. Evaluation of innovative integration concepts of combined solar thermal and heat pump systems for efficient thermal supply of industrial processes. In: *EuroSun 2016*, pp. 1–13.

Fuhrmann, F., Schirrer, A., Kozek, M., Jakubek, S., 2020. Prediction of pulsed heat loads in manufacturing plants. *IFAC-PapersOnLine* 53 (2), 10449–10454.

O. Gurobi, "Gurobi Optimizer Reference Manual, Version 5.0," [www.gurobi.com](http://www.gurobi.com), 2018.

IEA, "World Energy Outlook 2019," Paris, 2019,p 42.

Isaksson, A.J., Harjunkoski, I., Sand, G., 2018. The impact of digitalization on the future of control and operations. *Comput. Chem. Eng.* 114, 122–129.

Lofberg, J., 2004. YALMIP : a toolbox for modeling and optimization in MATLAB. In: *2004 IEEE Int. Symp. Comput. Aided Control Syst. Des.*, pp. 284–289.

May, G., Stahl, B., Taisch, M., Kiritsis, D., 2017. Energy management in manufacturing: from literature review to a conceptual framework. *J. Clean. Prod.* 167, 1464–1489.

Moretti, L., Martelli, E., Manzolini, G., 2020. An efficient robust optimization model for the unit commitment and dispatch of multi-energy systems and microgrids. *Appl. Energy* 261 (May 2019), 113859.

Moser, A., et al., 2020. A MILP-based modular energy management system for urban multi-energy systems: performance and sensitivity analysis. *Appl. Energy* 261.

Petek, J., Artz, M., Pözl, M., Pusch, M., Moser, T., 2018. On-line optimization of a complex industrial power plant using a novel fast and accurate modeling approach. In: *Proceedings of Electrify Europe Conference*, pp. 1–15.

Powell, K.M., et al., 2016. Thermal energy storage to minimize cost and improve efficiency of a polygeneration district energy system in a real-time electricity market. *Energy* 113, 52–63.

Rawlings, J.B., Patel, N.R., Risbeck, M.J., Maravelias, C.T., Wenzel, M.J., Turney, R.D., 2018. Economic MPC and real-time decision making with application to large-scale HVAC energy systems. *Comput. Chem. Eng.* 114, 89–98.

Risbeck, M.J., Maravelias, C.T., Rawlings, J.B., Turney, R.D., 2020. Mixed-integer optimization methods for online scheduling in large-scale HVAC systems. *Optim. Lett.* 14 (4), 889–924.

Sack, M., 2021. Implementation and validation of a multi layer model predictive controller for energy supply systems. TU Wien.

Santander, O., Baldea, M., 2021. Control-aware batch process scheduling. *Comput. Chem. Eng.* 152, 107360.

Schäfer, P., Schweidtmann, A.M., Lenz, P.H.A., Markgraf, H.M.C., Mitsos, A., 2020. Wavelet-based grid-adaptation for nonlinear scheduling subject to time-variable electricity prices. *Comput. Chem. Eng.* 132.

Shareef, H., Ahmed, M.S., Mohamed, A., Al Hassan, E., 2018. Review on home energy management system considering demand responses, smart technologies, and intelligent controllers. *IEEE Access* 6, 24498–24509.

Siirola, J.J., Edgar, T.F., 2012. Process energy systems: control, economic, and sustainability objectives. *Comput. Chem. Eng.* 47, 134–144.

Srinivasan, B., Bonvin, D., Visser, E., Palanki, S., 2003. Dynamic optimization of batch processes II. Role of measurements in handling uncertainty. *Comput. Chem. Eng.* 27 (1), 27–44.

Teng, T., Zhang, X., Dong, H., Xue, Q., 2020. A comprehensive review of energy management optimization strategies for fuel cell passenger vehicle. *Int. J. Hydrogen Energy* 45 (39), 20293–20303.

Thombre, M., Mdoe, Z., Jäschke, J., 2020. Data-driven robust optimal operation of thermal energy storage in industrial clusters. *Processes* 8 (2).

Touretzky, C.R., Baldea, M., 2014. Integrating scheduling and control for economic MPC of buildings with energy storage. *J. Process Control* 24 (8), 1292–1300.

Touretzky, C.R., Baldea, M., 2016. A hierarchical scheduling and control strategy for thermal energy storage systems. *Energy Build* 110, 94–107.

Zia, M.F., Elbouchikhi, E., Benbouzid, M., 2018. Microgrids energy management systems: a critical review on methods, solutions, and prospects. *Appl. Energy* 222, 1033–1055.