Sustainable Energy, Grids and Networks 34 (2023) 101033

Contents lists available at ScienceDirect

Sustainable Energy, Grids and Networks

journal homepage: www.elsevier.com/locate/segan

On the trade-off between profitability, complexity and security of forecasting-based optimization in residential energy management systems

Nils Müller, Mattia Marinelli, Kai Heussen, Charalampos Ziras *

Wind and Energy Systems Department, Technical University of Denmark, Building 330, Risø campus, 4000 Roskilde, Denmark

ARTICLE INFO

Article history:
Received 12 January 2023
Received in revised form 14 March 2023
Accepted 17 March 2023
Available online 21 March 2023

Keywords:
Energy management system
Prosumer
Flexibility
Forecasting
Machine learning
Security

ABSTRACT

With the emergence of affordable access to data sources, machine learning models and computational resources, sophisticated control concepts for residential energy management systems (EMSs) are on the rise. At the heart of those are production and consumption forecasts. Given the wide spectrum of implementation opportunities, selection of appropriate forecasting strategies is challenging. This work systematically evaluates forecasting-based optimization for residential EMSs in terms of trade-offs between economic profitability, computational complexity and security. The foundation of the study is two real prosumer cases equipped with a photovoltaic-battery system. Results demonstrate that, within the considered scenarios, best trade-offs are achieved based on forecasts of a default gradient-boosted decision trees model, using a short initial training set, weather forecast inputs and regular retraining. Over 90% of the theoretical maximum economic benefit is achieved in this scenario, at significantly lower computational complexity than others with similar savings, while being applicable to new systems without large data history. In terms of security, this scenario exhibits tolerance against weather input manipulation. However, sensitivity to price tampering may require data integrity checking in residential EMSs.

1. Introduction

Electricity grids are facing increasing shares of volatile renewable generation and variable consumption due to the electrification of the mobility and heating sectors [1,2]. The resulting larger temporal changes in supply and demand are entailing a need for electricity flexibility, with a great potential found within the residential sector [3]. At the same time, the increase and fluctuations in electricity prices motivate consumers to optimize their consumption. In that context, residential energy management systems (EMSs) constitute a promising solution, as controlling flexible energy resources allows to simultaneously provide (1) flexibility to the power system, and (2) financial benefits to consumers.

Given their steady cost decrease, photovoltaic (PV)-battery systems have become prominent examples of residential flexibility assets [4]. Existing EMSs typically apply simple myopic heuristics or rule-based controls for battery scheduling, without consideration of future electricity prices, production and consumption [5]. More advanced approaches combine optimization techniques with PV production and load forecasts. Given the affordable or even free access to weather and electricity prices data, machine learning (ML) models and computational resources, such more advanced concepts slowly find their way into application.

The use of forecasts is at the center of these optimization-based approaches. However, the wide range of implementation options and limitations makes selection of an appropriate forecasting strategy a compelling task. Overly complex models may provide minimal improvements at the cost of computational overhead. On the contrary, the lack of historical data for newly installed PV-battery systems may render the implementation of advanced models infeasible. Finally, strategies relying on data integration via the Internet (e.g., weather data or cloud-based forecasts) may open new opportunities for adversaries aiming at financial damage, for example, through data manipulation. These observations illustrate the need for a systematic and holistic assessment of different forecasting strategies for optimization in residential EMSs, considering trade-offs of profitability, complexity and security (see Fig. 1). In a nutshell, this can be expressed by the following research question: *"Under which conditions of data availability, computing resources and model complexity can forecasting-based battery scheduling in residential EMSs provide best trade-offs regarding economic profitability, computational complexity and security?"* To address this question, this work evaluates

* Corresponding author.
E-mail address: chazi@dtu.dk (C. Ziras).

https://doi.org/10.1016/j.segan.2023.101033

Nomenclature

Abbreviations

ANN	Artificial neural network
BMS	Battery management system
EMS	Energy management system
GBDT	Gradient-boosted decision trees
GHI	Global horizontal irradiance
HPC	High-performance computing
IoT	Internet of things
ML	Machine learning
PV	Photovoltaic
PVMS	Photovoltaic management system
RMSE	Root mean squared error
SM	Smart meter
SOC	State of charge
TPE	Tree Parzen Estimator

Parameters

α	Aggressiveness of the attack [–]
ΔT	Normalized duration of a time step [–]
η	Battery efficiency [–]
$\overline{p}_{\text{inv}}$	Inverter power capacity limit [kW]
$\overline{s}$	Battery upper SOC limit [kWh]
$\underline{s}$	Battery lower SOC limit [kWh]
d	Feature dimension [–]
k	Depth of decision trees [–]
L	Dataset length [–]
m	Number of decision trees [–]
N	Number [–]
v	Number of nodes in decision trees [–]
w	Time series window length [–]

Sets

$\boldsymbol{\Omega}$	Hyperparameter space
$\boldsymbol{\omega}$	Set of hyperparameters
$\mathcal{T}$	Set of steps in the optimization horizon
$\mathcal{W}$	Set of optimization variables

Indices

σ	Time step in moving average window
τ	5-min time step
C	Forecasting case
C'	EMS scenario
h	1-h time step
i	Dataset observation
j	1-h time steps ahead index
q	Time-series cross-validation fold

Variables

$\boldsymbol{X}$	Vector of covariate values [–]
δ	Battery charging/discharging status [–]
$\hat{\Lambda}$	Forecast of spot price [€/kWh]
$\hat{P}^{\text{L}}$	1-h avg. load forecast [kW]
$\hat{P}^{\text{PV}}$	1-h avg. PV production forecast [kW]
Λ	Spot price of a 1-h time step [€/kWh]
λ	Spot price of a 5-min time step [€/kWh]
$nRMSE$	Normalized RMSE [–]
$rRMSE$	Relative RMSE [–]
$\tilde{\Lambda}$	Manipulated spot price [€/kWh]
$\widehat{GHI}$	1-h average GHI forecast [W/m^2]
$\widehat{O}$	1-h avg. cloud opacity forecast [–]
A	Indicator of prosumer absence [–]
B	Economic benefit [€]
D	Day of the week [–]
F	Fees and taxes [€/kWh]
f	Fees and taxes of a 5-min time step [€/kWh]
H	Hour of the day [–]
K	Energy cost [€]
M	Memory need [–]
O	Approximated number of operations [–]
p	5-min avg. net demand ($p^{\text{L}} - p^{\text{PV}}$) [kW]
p^{b}	5-min avg. power from the grid [kW]
p^{c}	5-min avg. battery charging power [kW]
p^{d}	5-min avg. battery discharging power [kW]
P^{L}	1-h avg. load consumption [kW]
p^{L}	5-min avg. load consumption [kW]
P^{PV}	1-h average PV production [kW]
p^{PV}	5-min avg. PV production [kW]
p^{s}	5-min avg. power sold to the grid [kW]
R	Random number drawn from uniform distribution [–]
rB	Relative economic benefit [–]
rM	Relative memory need [–]
rO	Relative approximated number of operations [–]
s	SOC at end of a 5-min time step [kWh]

optimization-based control in residential EMS under several forecasting cases defined by a variety of model types, data availability scenarios and modeling strategies on two real prosumer cases.

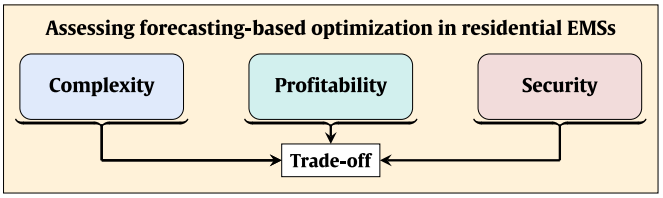

Fig. 1. Aspects for evaluating forecasting-based optimization in EMSs.

1.1. Related work

Most works on optimization in residential EMSs assess forecasting only by means of profitability. Typical approaches include the comparison of state-of-the-art rule-based control with forecasting-based optimization techniques [6,7], and the evaluation of different levels of forecast accuracy [8–12]. The authors of [7] compare a rule- and optimization-based strategy over the period of one year. Results demonstrate an up to 25% cost reduction by applying the latter. However, perfect weather and load forecasts (i.e., actual measured values) are assumed. The authors of [9] evaluate the impact of forecast uncertainty. Instead of evaluating real forecasts, random errors are artificially added to measurements to model forecasting uncertainty. Moreover, the evaluation is only based on simulation data for one week on a 30-min resolution. As shown in [11,13], time resolution, and thus the frequency at which the system is re-optimized (optimization frequency), has significant impact on economic performance assessment. In [8,10], the authors compare the impact

of perfect and realistic forecasts on economic performance. Both consider data of hourly resolution. Moreover, [8] is based on an artificially created dataset. Separate consumption and production data sources are combined, which breaks any existing correlation among those. The authors of [11] compare various forecasts for PV generation and load consumption. Several types of persistence models as well as perfect forecasts are considered for both PV and load forecasts. Additionally, artificial neural network (ANN)-based forecasts are used for load consumption, and an irradiation forecast-based PV model for PV forecasts. Compared to the previously described works, optimization frequency is two minutes. Moreover, several sensitivity analyses are conducted, including varying forecast errors, optimization frequencies and battery capacities. Results demonstrate that advanced forecast models allow for further price savings compared to persistence. However, prices are assumed to be always known for the next 24 h, which is not the case for spot prices. Moreover, PV generation and load consumption profiles of different buildings from different regions are combined, and PV production is artificially scaled. Finally, the impact of data availability (e.g., different amounts of training data) and modeling strategies (e.g., retraining) is not explored.

In contrast to the above-mentioned works, some extend economic evaluation with considerations of computational complexity. In [5,6], the authors compare multiple EMS strategies, covering both rule-based heuristics and forecasting-based optimization. The authors claim that the former achieve near-optimal solutions with lower computing resources compared to the latter. However, the optimization only runs on a 30-min frequency. Moreover, only persistence forecasts for PV generation and load consumption are considered as realistic forecasting approach. The authors of [14] propose a multi-objective predictive energy management strategy. The proposed prediction model is compared to several ML-based PV and load forecast models regarding profitability and computational complexity. However, only hourly data and one-step ahead predictions are considered.

In summary, the review of related literature demonstrates that most works only evaluate forecasting-based optimization in terms of profitability. Further, a large fraction exhibits methodologically shortcomings as they use short evaluation sets with low data resolution (30–60 min), assume to know prices for the entire optimization period or rely on artificially constructed prosumer datasets. To the best of the authors knowledge, no work systematically assesses several forecasting strategies for optimization in residential EMSs regarding economic profitability, computational complexity and security.

1.2. Contribution and paper structure

The main contributions of this work are as follows:

- Systematic and holistic evaluation of multiple scenarios of forecasting-based battery schedule optimization in residential EMSs.
- Consideration of various forecast cases defined by different model types, data availability and modeling strategies.
- Investigation of two real prosumer cases on an evaluation period of more than one year, considering an optimization frequency of 5 min and realistic price availability.
- Recommendations on optimal strategies for forecasting-based optimization in residential EMSs regarding trade-offs between economic profitability, computational complexity and security.

The remainder of this paper is structured as follows: In Section 2, the investigated prosumer scenarios are described. Section 3 introduces the applied methodology with regards to control strategies and underlying forecasting cases. In Section 4, details on the experimental setup and metrics are provided. Results are presented and evaluated in Section 5. Finally, a discussion of result implications is provided in Section 6, followed by a conclusion and view on future work in Section 7.

2. Prosumer concept and case description

In this study, two different residential prosumers are considered, which are each equipped with rooftop PV, a stationary storage system and an EMS (see Fig. 2). A common PV-battery inverter is assumed. The EMS comes with a dedicated smart meter smart meter (SM) that measures power at the grid connection point. PV and battery measurements are provided by the battery management system (battery management system (BMS)) and PV management system PV management system (PVMS), respectively, while load consumption is deducted from these measurements. The BMS further provides the current state of charge state of charge (SOC). While the measurements are sampled at high rates, they are usually available to users in extracted reports at, for example, 5-min resolution. The present work is based on such 5-min average values. Load and PV variations within the averaging period are not taken into account. Note that another meter is installed by the utility company for billing, but typically these meters provide only accumulated energy import and export values at 15- to 60-min rate. The prosumers are subject to instantaneous summation netting, that is imports and exports are summed up separately on the net result of all three phases [13].

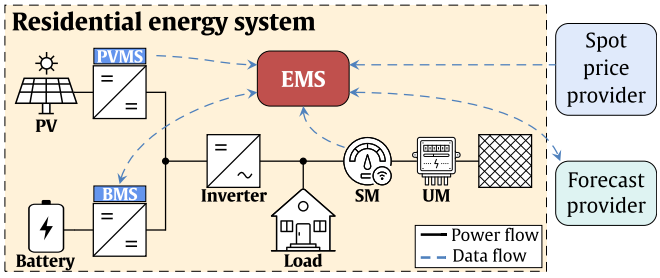

Fig. 2. Schematic representation of the residential energy system of both considered prosumers.

Apart from residential energy system-specific data, such as power measurements or the battery's SOC, the EMS has access to spot prices, which are published every day at 13:00 for the following day. Additionally, prosumers may acquire forecasts from third parties. One scenario is receiving weather forecasts, which can be used to generate PV and load forecasts locally. Another is direct procurement of the latter, for example, from providers of cloud-based forecasts. In this case, prosumers may need to provide historical and/or real-time measurements.

The consideration of two prosumers is justified by different production and consumption levels and patterns, allowing a broader evaluation of forecasting-based optimization. Production levels differ due to a higher nominal power of the PV plant of prosumer 1. Differences in load consumption mainly result from electric vehicle (EV) charging in case of prosumer 1, which entails a higher load level and less predictable patterns compared to the second one. Moreover, prosumer 1 purchased another EV in 2022, resulting in a change of load level and patterns during the recordings. Finally, prosumer 1 exhibits a higher self-consumption due to alignment of EV charging to PV production. Based on these characteristics, prosumer 2 can be considered a more traditional passive consumer, while prosumer 1 represents an already active future consumer. More details on the prosumers' setup follow in Section 4.1.

3. Methodology

This section introduces the underlying methodology of the different scenarios of forecasting-based optimization and a rule-based control benchmark, which together are referred to as *EMS scenarios* in the following. Section 3.1 describes the considered battery control strategies. In Section 3.2, different cases of PV and load forecasts are presented, which constitute the central foundation of the EMS scenarios. Finally, the realization of spot price forecasts is addressed in Section 3.3, followed by the introduction of two data manipulation scenarios in Section 3.4.

3.1. Battery control strategies

This subsection describes the two considered battery control strategies, namely an offline rule-based control benchmark (Section 3.1.1) and forecasting-based rolling-horizon optimization (Section 3.1.2).

3.1.1. Offline rule-based control benchmark

Rule-based controllers not depending on any prosumer-external information are common in real-life applications because they are robust, easy to implement and have minimal computational requirements. Thus, this offline approach will serve as a benchmark in the present study. The control mode that is used in this work, and many commercial PV-battery systems, minimizes the exchange of energy between the prosumer and the grid [5]. Let p_τ^{L} and p_τ^{PV} denote the 5-min average consumption and PV generation at time step τ, respectively. The difference of the two is the net demand $p_\tau = p_\tau^{\mathrm{L}} - p_\tau^{\mathrm{PV}}$. Each time step τ has a normalized duration of ΔT, where $\Delta T = 1/12$ for a 5-min step duration. When there is power surplus from the prosumer side, energy is stored in the battery. Once the battery is fully charged, excess energy is fed to the grid. If there is power deficit, energy from the battery supplies the load. If this is not possible because there is no sufficient energy stored, power is drawn from the grid. In all cases, battery inverter constraints are taken into account. A common inverter for the PV and the battery is assumed (see Fig. 2), so that the total power produced by the PV and flowing out of the battery cannot exceed the inverter's power capacity. The described control logic is summarized in the form of an algorithmic description in Algorithm 1, where η, p_τ^{c} and p_τ^{d} denote the battery's efficiency, charging and discharging power, respectively. s_τ is the SOC at the end of a 5-min period τ. The average power bought from/sold to the grid at τ is represented by p_τ^{b} and p_τ^{s}, while $\overline{s}$, $\underline{s}$ and $\overline{p}_{\mathrm{inv}}$ represent the upper/lower SOC limit and inverter power capacity, respectively.

Algorithm 1 Rule-based strategy for minimizing energy exchanges with the network.

$p_\tau \leftarrow p_\tau^{\mathrm{L}} - p_\tau^{\mathrm{PV}}$
if $p_\tau \geq 0$ **then**
  $p_\tau^{\mathrm{d}} \leftarrow \min(\overline{p}_{\mathrm{inv}} - p_\tau^{\mathrm{PV}}, p_\tau, (s_{\tau-1} - \underline{s})\eta/\Delta T)$
  $p_\tau^{\mathrm{b}} \leftarrow p_\tau - p_\tau^{\mathrm{d}}$
  $p_\tau^{\mathrm{s}} \leftarrow 0$
  $p_\tau^{\mathrm{c}} \leftarrow 0$
else
  $p_\tau^{\mathrm{c}} \leftarrow \min(\overline{p}_{\mathrm{inv}}, -p_\tau, (\overline{s} - s_{\tau-1})/(\eta\Delta T))$
  $p_\tau^{\mathrm{s}} \leftarrow -p_\tau - p_\tau^{\mathrm{c}}$
  $p_\tau^{\mathrm{b}} \leftarrow 0$
  $p_\tau^{\mathrm{d}} \leftarrow 0$

3.1.2. Forecasting-based rolling-horizon optimization

More advanced battery control schemes consider schedule optimization, which typically relies on prosumer-external information such as spot prices (see Fig. 2). In this work, an optimization problem according to

$$\min_{\mathcal{W}} \quad \sum_{\tau\in\mathcal{T}} \left[(\lambda_\tau|\hat{\lambda}_\tau + f_\tau)p_\tau^{\mathrm{b}} - \lambda_\tau|\hat{\lambda}_\tau p_\tau^{\mathrm{s}}\right]\Delta T, \tag{1a}$$

$$\text{s.t.} \quad 0 \leq p_\tau^{\mathrm{b}}, \quad 0 \leq p_\tau^{\mathrm{s}} \tag{1b}$$

$$0 \leq p_\tau^{\mathrm{c}} \leq \delta_\tau \overline{p}_{\mathrm{inv}} \tag{1c}$$

$$0 \leq p_\tau^{\mathrm{d}} \leq (1-\delta_\tau)\overline{p}_{\mathrm{inv}} \tag{1d}$$

$$\hat{p}_\tau^{\mathrm{PV}} + p_\tau^{\mathrm{d}} \leq \overline{p}_{\mathrm{inv}} \tag{1e}$$

$$p_\tau^{\mathrm{b}} - p_\tau^{\mathrm{s}} = \hat{p}_\tau^{\mathrm{L}} - \hat{p}_\tau^{\mathrm{PV}} + p_\tau^{\mathrm{c}} - p_\tau^{\mathrm{d}} \tag{1f}$$

$$s_{\tau+1} = s_\tau + \left[p_{\tau+1}^{\mathrm{c}}\,\eta + p_{\tau+1}^{\mathrm{d}}\,/\eta\right]\Delta T \tag{1g}$$

$$s_0 = s^{\mathrm{s}}, \quad s_{w_{\mathrm{opt}}} = s^{\mathrm{e}} \tag{1h}$$

$$\underline{s} \leq s_\tau \leq \overline{s} \tag{1i}$$

is considered. The set of optimization variables is denoted by $\mathcal{W}$. At every 5-min time step τ, the problem is solved over a look-ahead horizon of size w_{opt}, corresponding to a set of steps $\mathcal{T} = \{1, 2, \ldots, w_{\mathrm{opt}}\}$. Spot prices Λ and imposed fees and taxes F are hourly. Thus, constant values are used for each 5-min step τ within the respective hour, represented by λ_τ and f_τ. Unknown future values of p_τ^{PV}, p_τ^{L} and λ_τ within the look-ahead horizon are supplemented with forecasts. These are provided as hourly values and denoted by $\hat{P}_h^{\mathrm{PV}}$, $\hat{P}_h^{\mathrm{L}}$ and $\hat{\Lambda}_h$ for a 1-h time step h. For all 5-min time steps τ within the corresponding hour h, constant forecasts are considered, which are referred to as $\hat{p}_\tau^{\mathrm{PV}}$, $\hat{p}_\tau^{\mathrm{L}}$ and $\hat{\lambda}_\tau$. Whether true or forecasted prices are used in (1a) depends on the step τ within $\mathcal{T}$, which is indicated through a $\lambda_\tau|\hat{\lambda}_\tau$ notation. PV and load forecasts for a prediction horizon of size w_{pr} are performed every hour in a rolling fashion. Consequently, new forecasts are available every hour. As a default, this work assumes $w_{\mathrm{pr}} = 36$, which corresponds to an optimization horizon of $w_{\mathrm{opt}} = 432$. Initial studies suggest that $w_{\mathrm{pr}} = 36$ is sufficiently large to approximate the performance for $w_{\mathrm{pr}} \to \infty$. The impact of varying prediction horizons is evaluated in Section 5. Prices are published every day at 13:00 for the upcoming day. Price forecasts are performed at the same time and extend the published prices by another day. Further details on the PV, load and price forecasts follow in Section 3.2.

The battery charging/discharging status is represented by δ_τ and constitutes a binary decision variable. Starting and ending SOC values are denoted by s^{s} and s^{e}. The related constraint in (1h) requires the battery to be half charged at the end of the optimization period. All constraints (1b)–(1i) are imposed $\forall\ \tau \in \mathcal{T}$ except for (1g) and (1h), which hold $\forall\ \tau \in \mathcal{T}\backslash w_{\mathrm{opt}}$. After implementing the resulting optimal battery schedule at τ, a new optimization problem is solved at the following step based on the latest measurements and forecasts. *GLPK* [15] and *CVXPY* [16] are used as open-source optimization solver and modeling language, respectively.

3.2. Forecasting cases

This subsection describes the considered cases of PV and load forecasts required for battery schedule optimization (see Section 3.1.2). The set of forecasting cases is created by varying model type, data availability and modeling strategies. While Sections 3.2.1 and 3.2.2 introduce the considered scenarios of data availability and modeling strategies, respectively, applied forecast model types are introduced in Section 3.2.3–3.2.5. An overview of the 18 resulting cases (C1–C18) is provided in Table 1.

3.2.1. Data availability

In this work, different availability and usage scenarios are considered with respect to historical training data and external weather forecasts (see Table 1). As for historical data, a small and large set are considered. The former represents a scenario of minimal available data history, which might result from a newly installed PV plant, metering device or EMS. For the latter, a long operation history of the PV-battery system is assumed. Varying the training set size aims to assess whether (1) data-driven forecasting can be applied to new systems "out-of-the-box", and (2) extensive data histories allow for significant forecast improvements and additional economic benefits.

As for external weather data, GHI and cloud opacity forecasts are considered. These can be obtained against payment or partly free of charge [17,18]. The comparison of cases with and without use of weather forecasts enables to assess if additional effort and potential costs of incorporating external weather data are justified by savings through higher forecast accuracy. Moreover, it allows contrasting profitability gains with security concerns arising from the dependency on potentially manipulated external data.

Both larger training sets and additional features increase the amount of data to be processed, which ultimately impacts computational complexity. If the computational burden exceeds the capabilities of typical EMS hardware, cloud-computing may be necessary. In this case, sensitive consumption data may need to be provided to third-parties, which should be taken into account in the assessment of different forecasting strategies.

3.2.2. Modeling strategies

The considered modeling strategies comprise model (hyperparameter) selection and regular retraining (see Table 1). Both are typical procedures which usually improve accuracy, however, at the cost of increased computational complexity. Contrasting cases with and without applying these strategies provides insight regarding optimal trade-offs between profitability and computational burden. Similar to the processing of extensive data (see Section 3.2.1), complex modeling strategies may necessitate cloud computing. Resulting privacy concerns should be considered in the evaluation of forecasting strategies.

3.2.3. Naïve forecast

The first model type is a naïve persistence forecast, which is a popular benchmark and frequently applied by studies on forecasting-based optimization of PV-battery systems [6,11]. The term *persistence* stems from the fact that values within the prediction horizon of size w_{pr} are assumed to be the same as in a previous period. Daily persistence is considered for PV forecasting according to

$$\hat{P}^{\mathrm{PV}}_{h+1}, \ldots, \hat{P}^{\mathrm{PV}}_{h+w_{\mathrm{pr}}} = P^{\mathrm{PV}}_{h+1-24}, \ldots, P^{\mathrm{PV}}_{h+w_{\mathrm{pr}}-24}. \tag{2}$$

To account for different consumption patterns between weekdays and weekends, weekly persistence is applied for load forecasting as given by

$$\hat{P}^{\mathrm{L}}_{h+1}, \ldots, \hat{P}^{\mathrm{L}}_{h+w_{\mathrm{pr}}} = P^{\mathrm{L}}_{h+1-7\cdot24}, \ldots, P^{\mathrm{L}}_{h+w_{\mathrm{pr}}-7\cdot24}. \tag{3}$$

Eqs. (2) and (3) are implemented as rolling forecasts, generating predictions at each time step h for the following w_{pr} hours. The model is implemented in Python using the open-source forecasting library *Darts* [19]. Persistence forecasts are independent of training data, external weather data as well as model selection or training processes (see Table 1). This simplicity renders it also an attractive strategy for residential EMSs, as shown by the frequent consideration in many related works.

Table 1
Overview of forecasting cases with regards to model type, data availability and modeling strategies.

Case	Model	Data availability		Modeling strategies	
		Train set size	Weather data	Selection	Retraining
C1	Naïve	None	No	No	No
C2	GBDT	Small	No	No	No
C3	GBDT	Small	No	No	Yes
C4	GBDT	Small	No	Yes	No
C5	GBDT	Small	No	Yes	Yes
C6	GBDT	Small	Yes	No	No
C7	GBDT	Small	Yes	No	Yes
C8	GBDT	Small	Yes	Yes	No
C9	GBDT	Small	Yes	Yes	Yes
C10	GBDT	Large	No	No	No
C11	GBDT	Large	No	No	Yes
C12	GBDT	Large	No	Yes	No
C13	GBDT	Large	No	Yes	Yes
C14	GBDT	Large	Yes	No	No
C15	GBDT	Large	Yes	No	Yes
C16	GBDT	Large	Yes	Yes	No
C17	GBDT	Large	Yes	Yes	Yes
C18	Oracle	–	–	–	–

3.2.4. GBDT forecasts

To enable a fair comparison of different data availability and modeling strategy scenarios for PV and load forecasting, the same model type (gradient-boosted decision trees (GBDT)) is considered for the cases C2-C17 (see Table 1). Although comparing different ML models would provide additional insights, it is out of the scope of this work. GBDT [20] is a widely applied ML technique. Its popularity arises from its efficiency, interpretability and state-of-the-art accuracy, as, for example, demonstrated by regularly winning data mining and time series forecasting competitions [21]. Moreover, they are actively researched and improved as many recent versions, such as XGBoost [22], LightGBM [23] and CatBoost [24], demonstrate. GBDT combines the predictions of many individual decision trees, which constitute a set of weak learners. The trees are connected in series, so that each learner tries to minimize the residual between ground truth and prediction of the previous tree. The simultaneous high accuracy and efficiency renders GBDT a promising candidate for residential EMS applications.

GBDT is applied to predict the expected values for a prediction horizon of w_{pr} steps at time step h based on lag values $P^{\mathrm{PV|L}}_{h}, \ldots, P^{\mathrm{PV|L}}_{h-w_{\mathrm{hist}}}$ and covariates $\boldsymbol{X}^{\mathrm{PV|L}}_{h+1}, \ldots, \boldsymbol{X}^{\mathrm{PV|L}}_{h+w_{\mathrm{pr}}}$ according to

$$\hat{p}^{\mathrm{PV}}_{h+1}, \ldots, \hat{p}^{\mathrm{PV}}_{h+w_{\mathrm{pr}}} = \Phi\left(P^{\mathrm{PV}}_{h}, \ldots, P^{\mathrm{PV}}_{h-w_{\mathrm{hist}}}, \boldsymbol{X}^{\mathrm{PV}}_{h+1}, \ldots, \boldsymbol{X}^{\mathrm{PV}}_{h+w_{\mathrm{pr}}}\right) \tag{4}$$

and

$$\hat{P}^{\mathrm{L}}_{h+1}, \ldots, \hat{P}^{\mathrm{L}}_{h+w_{\mathrm{pr}}} = \Phi\left(P^{\mathrm{L}}_{h}, \ldots, P^{\mathrm{L}}_{h-w_{\mathrm{hist}}}, \boldsymbol{X}^{\mathrm{L}}_{h+1}, \ldots, \boldsymbol{X}^{\mathrm{L}}_{h+w_{\mathrm{pr}}}\right), \tag{5}$$

where a history window of w_{hist} steps is considered. The forecasts are generated every hour in a rolling fashion to provide up-to-date predictions. Covariates comprise calendric features, prosumer absence and external weather forecasts (see Table 2). Absence feature A assumes that prosumers can enter holidays in their EMS to allow load forecast models for better predictions in these periods. Since no correlation between PV generation and day of the week D and prosumer absence A exists, they are excluded for PV forecasting. As a result, $\boldsymbol{X}^{\mathrm{PV}} = \{H, \widehat{GHI}, \widehat{O}\}$ and $\boldsymbol{X}^{\mathrm{L}} = \{H, D, A, \widehat{GHI}, \widehat{O}\}$ in cases considering use of weather forecasts, and $\boldsymbol{X}^{\mathrm{PV}} = \{H\}$ and $\boldsymbol{X}^{\mathrm{L}} = \{H, D, A\}$ in cases without.

For scenarios applying model selection, the automatic hyperparameter optimization software *optuna* [25] is used in combination with three-fold time-series cross-validation [26]. The tuned hyperparameters and respective search spaces are listed in

Table 2
Covariates considered for GBDT-based forecasting.

Covariate	Sign	Value range
Hour of the day	H	$\{H \in \mathbb{N} \mid H = [0, \ldots, 23]\}$
Day of the week	D	$\{D \in \mathbb{N} \mid D = [0, \ldots, 6]\}$
Prosumer absence	A	$\{A \in \mathbb{N} \mid A = [0, 1]\}$
GHI forecasts	$\widehat{GHI}$	$\{\widehat{GHI} \in \mathbb{R} \mid \widehat{GHI} \geq 0\}$
Cloud opacity forecasts	$\widehat{O}$	$\{\widehat{O} \in \mathbb{R} \mid \widehat{O} = [0, \ldots, 1]\}$

Table 3. For all other hyperparameters, default values according to [27] are used. To find optimal sets of hyperparameters for PV ($\boldsymbol{\omega}^{\mathrm{PV}}_{\mathrm{opt}}$) and load forecasting ($\boldsymbol{\omega}^{\mathrm{L}}_{\mathrm{opt}}$) within the hyperparameter space $\boldsymbol{\Omega}$, the average root mean squared error (RMSE) over all N_{folds} folds and w_{pr} prediction steps is minimized by

$$\boldsymbol{\omega}^{\mathrm{PV|L}}_{\mathrm{opt}} = \arg\min_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \frac{\sum_{q=1}^{N_{\mathrm{folds}}} \sum_{j=1}^{w_{\mathrm{pr}}} \sqrt{\sum_{i=1}^{L^{(q)}_{\mathrm{val}}} \frac{\left(\hat{P}^{\mathrm{PV|L}}_{j,i}(\boldsymbol{\omega}) - P^{\mathrm{PV|L}}_{i}\right)^2}{L^{(q)}_{\mathrm{val}}}}}{N_{\mathrm{folds}} \cdot w_{\mathrm{pr}}}, \tag{6}$$

where $L^{(q)}_{\mathrm{val}}$ is the length of the validation set of the qth fold, $\hat{P}^{\mathrm{PV|L}}_{j,i}(\boldsymbol{\omega})$ the j-steps ahead forecast for the ith observation in the validation set of the qth fold based on a hyperparameter set $\boldsymbol{\omega}$ and $P^{\mathrm{PV|L}}_{i}$ the corresponding ground truth. The Bayesian optimization algorithm Tree Parzen Estimator (TPE) [28] is applied to find $\boldsymbol{\omega}_{\mathrm{opt}}$ according to (6) within a predefined number of hyperparameter set samples of $N_{\mathrm{trails}} = 2000$. While a large data history is considered sufficient for identification of optimal hyperparameters, small training datasets are likely to require regular model reselection, for example, due to lack of samples of all seasons of a year. Therefore, in cases considering a large data history (see Table 1), model selection is only conducted once based on the initial training set. For the ones with little training data, model selection is conducted repeatedly every three months. Scenarios without model selection use default parameters [27] and $w_{\mathrm{hist}} = 48$.

Table 3
Tuned hyperparameters and associated search spaces for GBDT-based forecasting.

No.	Hyperparameter	Search space
1	w_{hist}	[4, …, 192]
2	L1 regularization	[0, …, 100]
3	Bagging fraction	[0.1, …, 1]
4	Max. number of leaves in one tree	[20, …, 3000]
5	Feature fraction	[0.1, …, 1]
6	Max. depth of a tree	[3, …, 21]
7	Number of decision trees	[100, …, 10000]
8	Learning rate	[0.001, …, 0.3]

In cases which consider retraining (see Table 1), the GBDT model is repeatedly trained every week based on the entire previous data history. If no retraining is considered, only initial training is conducted. All GBDT-based cases (C2-C17) are implemented in Python using the open-source forecasting library *Darts* [19].

3.2.5. Oracle forecast

In addition to the naïve lower-end forecast benchmark, an oracle forecast is considered to quantify the theoretical optimum. The oracle forecast is characterized by perfect knowledge of the future, which includes that time resolution and w_{pr} are converging to infinity. This behavior is approximated with assuming perfect forecasts for every 5-min time step τ within a prediction horizon of $w_{\mathrm{pr}} = 2016$ steps (seven days) according to

$$\hat{p}^{\mathrm{PV}}_{\tau+1}, \ldots, \hat{p}^{\mathrm{PV}}_{\tau+w_{\mathrm{pr}}} = p^{\mathrm{PV}}_{\tau+1}, \ldots, p^{\mathrm{PV}}_{\tau+w_{\mathrm{pr}}} \tag{7}$$

and

$$\hat{p}^{\mathrm{L}}_{\tau+1}, \ldots, \hat{p}^{\mathrm{L}}_{\tau+w_{\mathrm{pr}}} = p^{\mathrm{L}}_{\tau+1}, \ldots, p^{\mathrm{L}}_{\tau+w_{\mathrm{pr}}}. \tag{8}$$

As the oracle forecast only constitutes a theoretical benchmark, no reasonable definition of data availability scenarios and modeling strategies can be made (see Table 1).

3.3. Spot price forecast

Hourly spot prices are typically published every day at 13:00 for the following day [29]. This leads to varying spot price knowledge horizons between 12 and 35 h, depending on the time of the day. Thus, if optimization horizons of more than 12 h are considered, price forecasts are required. Spot price forecasting is a complex task which depends on inputs such as wind production, consumption, calendric features and many more [30]. Recently, first providers offer access to advanced forecasts [31]. However, for the evaluation period considered in this work, historical forecasts could not be acquired. To avoid the assumption of knowing true prices for the entire optimization period, price forecasts are generated based on a GBDT model. Together with the publishing of spot prices for the next day, prices for the day after tomorrow are predicted at 13:00 according to

$$\hat{\Lambda}_{h+36}, \ldots, \hat{\Lambda}_{h+60} = \Phi\left(\Lambda_{h+35}, \ldots, \Lambda_{h-w_{\mathrm{hist}}}, \boldsymbol{X}^{\Lambda}_{h+36}, \ldots, \boldsymbol{X}^{\Lambda}_{h+60}\right), \tag{9}$$

with $\boldsymbol{X}^{\Lambda} = \{H, D\}$. Since advanced spot price forecasting is not the focus of this work, only lag values and calendric covariates are considered. Optimal hyperparameters are selected on a two-year history following a similar approach to (6). During the prediction of the evaluation set, the model is retrained on a daily basis. Note that varying spot price forecasts is not explicitly part of the case study. However, to validate this comparatively simple approach and assess if more complex price forecasts can be justified by significant economic benefits, a comparison to assuming true prices is included in Section 5.2.2.

3.4. Data manipulation

Cost-optimal scheduling of batteries requires external data such as spot prices and weather forecasts (see Fig. 2). While the required connection to the internet is the foundation for such smart applications, it also introduces new cyber vulnerabilities. Events such as the Mirai botnet in 2016 have shown that attacks on distributed internet of things (IoT)-devices are a reality [32]. Thus, also potential damage should be taken into consideration whenever assessing the advancements of IoT-based applications. Among the most famous and critical attacks in power systems are false data injections [33]. Based on an impact quantification of such attacks, the different EMS scenarios can be better assessed in terms of trade-offs between profitability and security. For that purpose, this subsection introduces two data manipulation scenarios. While Section 3.4.1 describes manipulation of spot price data, Section 3.4.2 is concerned with tampering of external weather forecasts.

3.4.1. Spot price manipulation

The objective of the considered price manipulation is to approximate an opposite behavior of cost-optimal operation. For that purpose, the attack model mirrors prices on their moving average according to

$$\underset{\sim}{\Lambda}_h = \Lambda_h - 2 \cdot \left(\Lambda_h - \sum_{\sigma=0}^{w_{\mathrm{avg}}} \frac{\Lambda_{h-\sigma}}{w_{\mathrm{avg}}}\right), \tag{10}$$

where $w_{\mathrm{avg}} = 23$. One attacker's motivation could be financial damage of prosumers. However, more critical is the potential switch from peak shaving to peak reinforcing behavior of flexible residential loads. If able to manipulate price input of multiple EMSs, an attacker could target overloading situations entailing disconnection of customers. The attack model in (10) is exemplarily depicted in Fig. 3.

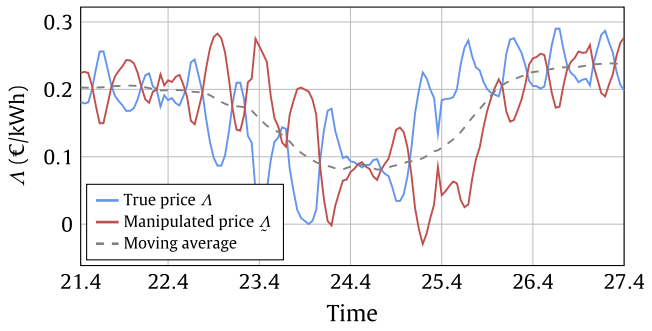

Fig. 3. Exemplary depiction of the spot price manipulation according to (10) on an excerpt from April 2022.

3.4.2. GHI and cloud opacity forecast manipulation

Reducing the accuracy of inputs for PV and load forecasts is likely to translate to lower economic benefits due to sub-optimal battery scheduling. Therefore, an attacker's motivation for manipulating weather forecast inputs could be financial damage. As damage increases over time, attackers may try to keep modifications discreet to avoid detection. This behavior is imitated by adding noise of different intensity levels to GHI and cloud opacity forecasts according to

$$\underline{\widehat{GHI}}_{h+j} = \widehat{GHI}_{h+j} \cdot (1 + \alpha R)\,, \forall j \in [1, 2, ..., w_{\text{pr}}] \tag{11}$$

and

$$\underline{\widehat{O}}_{h+j} = \widehat{O}_{h+j} \cdot (1 + \alpha R)\,, \forall j \in [1, 2, ..., w_{\text{pr}}], \tag{12}$$

with R being a random number drawn from the uniform distribution $R \sim U(-1, 1)$ and α the aggressiveness of the attack with $\alpha \in [0.2, 1, 10]$. To further hide the attack, physically implausible values are avoided by containing manipulated GHI values between zero and the maximum value in the respective region, and cloud opacity between zero and one. The attack models in (11) and (12) are exemplarily depicted in Fig. 4.

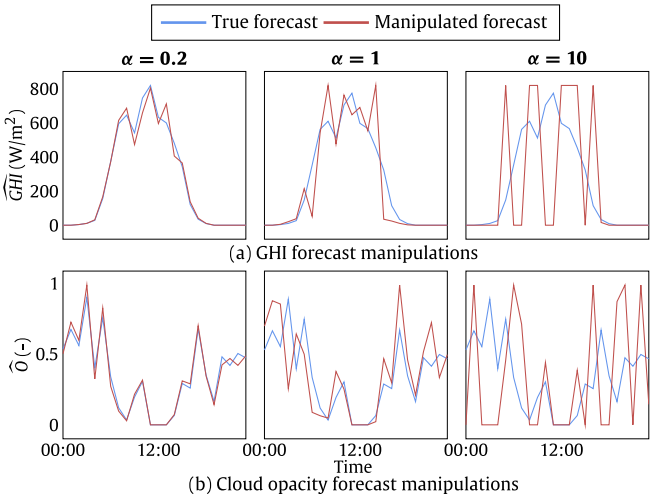

Fig. 4. Exemplary depiction of (a) GHI and (b) cloud opacity forecast manipulations on an excerpt from June 2022.

4. Experimental setup and metrics

This section is concerned with the experimental setup and applied metrics of the present study. Section 4.1 provides prosumer and data specifications. Thereafter, the applied forecasting performance metrics (Section 4.2) and economic performance indicators (Section 4.3) are introduced.

4.1. Prosumer and data specification

The two prosumers are located in Roskilde, Denmark, and are subject to the DK2 day-ahead price zone. Both follow the schematic representation in Fig. 2. They are equipped with a PV system of 6 kW_p (prosumer 1) and 5 kW_p (prosumer 2), respectively. For the battery system, $\bar{s} = 8$ kWh, $\underline{s} = 0.8$ kWh and $\eta = 0.95$ is considered. Moreover, the inverter comes with $\overline{p}_{\text{inv}} = 6$ kW for prosumer 1 and $\overline{p}_{\text{inv}} = 5$ kW for prosumer 2. All components of the residential energy systems follow a dimensioning typical for the Danish case. The self-consumption of the two prosumers without battery is 67% and 44%, respectively. In both cases the EMS receives PV and load measurements as 5-min averages. External weather forecasts (GHI and cloud opacity) are provided with an hourly resolution. For the two prosumers, historical data of different length is available. The dataset of prosumer 1 comprises approximately three years, beginning on the 1st of September 2019 and ending on the 30th of October 2022. For prosumer 2, 14.5 months between the 15th of August 2021 and the 30th of October 2022 are available. In both cases, the last 14 months (1st of September 2021 to 30th of October 2022) are reserved for evaluation in Section 5. As detailed in Section 3.2.1, historical training data of different size are considered within the forecasting cases. While the small set comprises the last two weeks of August 2021, the large one spans over two years from 1st of September 2019 to 31st of August 2021. Since historical data of prosumer 2 are not available before the 15th of August 2021, forecasting cases considering two years of training data are only evaluated on prosumer 1.

4.2. Forecasting performance metrics

This subsection introduces the performance metrics applied to evaluate accuracy (Section 4.2.1) and computational complexity (Section 4.2.2) of the forecasting cases.

4.2.1. Accuracy

The accuracy for a j-steps ahead PV or load prediction under forecasting case C is quantified based on the normalized RMSE according to

$$\begin{aligned} nRMSE_j^{\text{PV|L},C} &= \frac{\sqrt{\frac{\sum_{i=1}^{L_{\text{eval}}} \left(\hat{P}_{j,i}^{\text{PV|L},C} - P_i^{\text{PV|L}}\right)^2}{L_{\text{eval}}}}}{\frac{\sum_{i=1}^{L_{\text{eval}}} P_i^{\text{PV|L}}}{L_{\text{eval}}}} \\ &= \frac{\sqrt{L_{\text{eval}} \sum_{i=1}^{L_{\text{eval}}} \left(\hat{P}_{j,i}^{\text{PV|L},C} - P_i^{\text{PV|L}}\right)^2}}{\sum_{i=1}^{L_{\text{eval}}} P_i^{\text{PV|L}}}, \end{aligned} \tag{13}$$

where L_{eval} is the length of the 14 month evaluation set. Normalization removes the impact of the scale of PV production and load consumption and thus facilitates comparison among *different* prosumers. The overall performance of the considered multi-step forecasts is quantified as the average over the entire forecasting horizon of size w_{pr} according to

$$nRMSE_{\text{avg}}^{\text{PV|L},C} = \frac{\sum_{j=1}^{w_{\text{pr}}} nRMSE_j^{\text{PV|L},C}}{w_{\text{pr}}}. \tag{14}$$

For comparison of different forecasting cases on the *same* prosumer, the relative averaged normalized RMSE $rRMSE_{\text{avg}}^{\text{PV|L},C}$ is

considered, which follows from dividing $nRMSE_{\text{avg}}^{\text{PV},C}$ and $nRMSE_{\text{avg}}^{\text{L},C}$ by the respective highest value among all cases[1] as given by

$$rRMSE_{\text{avg}}^{\text{PV|L},C} = \frac{nRMSE_{\text{avg}}^{\text{PV|L},C}}{\max\left\{nRMSE_{\text{avg}}^{\text{PV|L,C1}}, \ldots, nRMSE_{\text{avg}}^{\text{PV|L,C17}}\right\}}. \tag{15}$$

4.2.2. Computational complexity

The computational complexity of an ML model in terms of training time, prediction time and space can be expressed with the respective big O notation [34]. For the considered GBDT model, these are given as $\mathcal{O}_{\text{train}}\,(L_{\text{train}}\log(L_{\text{train}})dm)$ (train time complexity), $\mathcal{O}_{\text{pred}}\,(km)$ (prediction time complexity) and $\mathcal{O}_{\text{space}}\,(vm)$ (space complexity), where L_{train} is the length of the training set, d the feature dimension, m the number of trees, k the depth of the trees and v the number of nodes in the trees [35,36]. Based on these notations, the number of operations for training and prediction as well as memory needs of a PV and load forecasting model of case C are approximated by

$$O_{\text{train}}^{\text{PV|L},C} \approx N_{\text{trails}}^{C} L_{\text{train}}^{C} \log(L_{\text{train}}^{C}) d^{\text{PV|L},C} m^{\text{PV|L},C}, \tag{16}$$

$$O_{\text{pred}}^{\text{PV|L},C} \approx k^{\text{PV|L},C} m^{\text{PV|L},C} \tag{17}$$

and

$$M^{\text{PV|L},C} \approx v_{\text{max}}^{C} m_{\text{max}}^{C}, \tag{18}$$

where m_{max}^{C} and v_{max}^{C} constitute the maximum implemented number of trees and nodes in trees, respectively.[2] For cases considering retraining, $L_{\text{train}}^{\text{PV|L},C}$ is defined by the size of the largest retraining set during evaluation. To facilitate comparison among the different forecasting cases, $O_{\text{train}}^{\text{PV|L},C}$, $O_{\text{pred}}^{\text{PV|L},C}$ and $M^{\text{PV|L},C}$ are divided by the respective maximum value across all cases according to

$$rO_{\text{train}}^{\text{PV|L},C} = \frac{O_{\text{train}}^{\text{PV|L},C}}{\max\left\{O_{\text{train}}^{\text{PV|L,C1}}, \ldots, O_{\text{train}}^{\text{PV|L,C17}}\right\}}, \tag{19}$$

$$rO_{\text{pred}}^{\text{PV|L},C} = \frac{O_{\text{pred}}^{\text{PV|L},C}}{\max\left\{O_{\text{pred}}^{\text{PV|L,C1}}, \ldots, O_{\text{pred}}^{\text{PV|L,C17}}\right\}} \tag{20}$$

and

$$rM^{\text{PV|L},C} = \frac{M^{\text{PV|L},C}}{\max\left\{M^{\text{PV|L,C1}}, \ldots, M^{\text{PV|L,C17}}\right\}}. \tag{21}$$

4.3. Economic performance indicators

The set of evaluated EMS scenarios comprises rolling-horizon optimization based on the forecasting cases C1–C18 and the offline rule-based benchmark. For simplicity, optimization-based EMS scenarios are referred to as their underlying forecasting case. The energy cost under an EMS scenario C' is calculated by

$$K^{C'} = \sum_{i=1}^{L_{\text{eval},\tau}} \left[p_i^{\text{b},C'}(\lambda_i + f_i) - p_i^{\text{s},C'}\lambda_i\right] \Delta T, \tag{22}$$

where $L_{\text{eval},\tau}$ is the length of the evaluation period in 5-min resolution. Based on the costs, the economic benefit under a scenario C' is expressed as the difference to the baseline cost without a battery (K^{base}) according to

$$B^{C'} = K^{\text{base}} - K^{C'}. \tag{23}$$

To simplify comparison among the scenarios, their benefit is assessed by comparing with the theoretical maximum. As the maximum benefit is achieved in case of assuming oracle forecasts (C18), the resulting relative benefit under an EMS scenario C' can be expressed as

$$rB^{C'} = \frac{B^{C'}}{B^{\text{C18}}}. \tag{24}$$

Note that in case of a monthly cost analysis of the evaluation period, the relative benefit $rB^{C'}$ will be denoted as $rB_m^{C'}$.

5. Results

This section evaluates the considered EMS scenarios regarding profitability, complexity and security. In preparation of that, Section 5.1 examines the underlying forecasting cases in terms of accuracy and computational complexity. Thereafter, the evaluation of EMS scenarios follows in Section 5.2. Unless otherwise stated, results are based on the default forecasting horizon $w_{\text{pr}} = 36$ (see Section 3.1.2).

5.1. Performance evaluation of forecasting cases

In this subsection, the forecasting cases are first evaluated regarding accuracy (Section 5.1.1) and computational complexity (Section 5.1.2). In Section 5.1.3, a conclusion on the trade-off between these factors is provided. Finally, Section 5.1.4 analyzes the error behavior of PV and load forecasts over the forecasting horizon.

5.1.1. Accuracy

A performance overview for all cases is provided in Table 4. Note that $rRMSE_{\text{avg}}$ is depicted, which shows the relative performance of each case against the worst forecast for a given variable and prosumer. Highest accuracy is written in bold, second best is underlined and third best dotted underlined. The *absolute* averaged RMSEs of the best performing cases are $RMSE_{\text{avg}}^{\text{PV,C17}} = 0.3196\,\text{kW}$ and $RMSE_{\text{avg}}^{\text{L,C9}} = 0.5065\,\text{kW}$ (prosumer 1) as well as $RMSE_{\text{avg}}^{\text{PV,C9}} = 0.2569\,\text{kW}$ and $RMSE_{\text{avg}}^{\text{L,C9}} = 0.2555\,\text{kW}$ (prosumer 2).

Model type. The persistence model (C1) forms the lower performance end for load forecasting. For PV forecasting, the GBDT model without large training data, external weather forecasts, model selection and retraining (C2) performs worst. Apart from this exception, GBDT-based forecasts outperform the persistence benchmark by at least 12.3 percentage points, which proves the existence of learnable patterns in P^{PV} and P^{L}, justifying the application of ML models.

Historical data size. The availability of comprehensive training data improves accuracy significantly in cases without use of external weather input or retraining (e.g., C2 vs. C10 and C6 vs. C14). Other cases only exhibit minor improvements, which is caused by two factors. On the one hand, using highly correlated weather forecasts as input simplifies the problem and makes comprehensive training data obsolete. On the other hand, retraining exploits newly incoming data and thus minimizes the need for large data histories. These findings suggest that data-driven forecasting can also be applied in scenarios of small data histories, such as newly installed EMSs. For load forecasts, large training data can even worsen results (e.g., C9 vs. C17). The reason is a changing consumption behavior of prosumer 1 due to purchase of a second EV in 2022, which renders older load data less useful and hinders model selection and training.

[1] C18 (oracle) is excluded as it constitutes no realistic forecasting case.

[2] In cases considering model selection, combinations of v and m may occur which entail higher memory needs than the finally selected model.

Table 4
Accuracy and computational complexity for PV and load forecasts. Highest accuracy in bold, second best underlined and third best dotted underlined.

Case	Model	Train set size	Weather data	Selection	Retraining	Accuracy				Computational complexity									
						$rRMSE_{avg}$ (–)				rO_{train} (–)				rO_{pred} (–)				rM (–)	
						Prosumer 1		Prosumer 2		Prosumer 1		Prosumer 2		Prosumer 1		Prosumer 2		Prosumer 1		Prosumer 2	
						PV	Load	PV	Load	PV	Load	PV	Load	PV	Load	PV	Load	PV	Load	PV	Load
C1	Naïve	–	No	No	No	0.916	1	0.954	1	0	0	0	0	0	0	0	0	0	0	0	0
C2	GBDT	Small	No	No	No	1	0.999	1	0.78	0	0	0	0	0	0	0	0	0	0	0	0
C3	GBDT	Small	No	No	Yes	0.754	0.782	0.778	0.715	0	0	0	0	0	0	0	0	0	0	0	0
C4	GBDT	Small	No	Yes	No	0.795	0.838	0.835	0.722	0.9	0.3	0.9	0.5	0.6	0.9	0.3	0.8	1	1	1	1
C5	GBDT	Small	No	Yes	Yes	0.719	0.759	0.753	0.702	1	0.3	1	0.5	0.6	0.9	0.3	0.8	1	1	1	1
C6	GBDT	Small	Yes	No	No	0.541	0.877	0.567	0.769	0	0	0	0	0	0	0	0	0	0	0	0
C7	GBDT	Small	Yes	No	Yes	0.401	0.678	0.429	0.709	0	0	0	0	0	0	0	0	0	0	0	0
C8	GBDT	Small	Yes	Yes	No	0.412	0.714	0.446	0.725	0.9	0.3	0.9	0.9	1	1	1	1	1	1	1	1
C9	GBDT	Small	Yes	Yes	Yes	0.381	**0.664**	**0.411**	**0.701**	1	0.4	1	1	1	1	1	1	1	1	1	1
C10	GBDT	Large	No	No	No	0.705	0.828	–	–	0	0	–	–	0	0	–	–	0	0	–	–
C11	GBDT	Large	No	No	Yes	0.701	0.761	–	–	0	0	–	–	0	0	–	–	0	0	–	–
C12	GBDT	Large	No	Yes	No	0.694	0.791	–	–	0.2	0.2	–	–	0	0.1	–	–	1	1	–	–
C13	GBDT	Large	No	Yes	Yes	0.693	0.752	–	–	0.4	0.3	–	–	0	0.1	–	–	1	1	–	–
C14	GBDT	Large	Yes	No	No	0.377	0.77	–	–	0	0	–	–	0	0	–	–	0	0	–	–
C15	GBDT	Large	Yes	No	Yes	0.373	0.693	–	–	0	0	–	–	0	0	–	–	0	0	–	–
C16	GBDT	Large	Yes	Yes	No	0.363	0.753	–	–	0.5	0.6	–	–	0.1	0.7	–	–	1	1	–	–
C17	GBDT	Large	Yes	Yes	Yes	**0.36**	0.681	–	–	0.8	1	–	–	0.1	0.7	–	–	1	1	–	–
C18	Oracle	–	–	–	–	0	0	0	0	–	–	–	–	–	–	–	–	–	–	–	–

Weather forecasts. Incorporating GHI and cloud opacity forecasts improves both PV and load forecasting in all cases (e.g., C3 vs. C7). While the impact is significant for PV forecasts as anticipated, the potential for load forecast improvement depends on the prosumer. As prosumer 1 is aligning consumption (EV charging) with PV production, the impact is stronger compared to prosumer 2.

Model selection. Model selection improves accuracy in all cases. However, the improvement compared to using default hyperparameters usually is only in the range of 1 to 3 percentage points (e.g., C15 vs. C17).

Retraining. Cases considering retraining exhibit lower *rRMSE* values compared to their non-retrained counterparts (e.g., C6 vs. C7). As expected, the impact is larger for models with small initial training sets. Nevertheless, even for cases assuming two years of training data, retraining can significantly improve accuracy, given that new changes in data distributions appear. As mentioned above, this is the case for load consumption of prosumer 1, which explains the comparatively large improvement for P^{L} forecasts, for example, between C16 and C17. In contrast, PV models trained on two years of historical data have learned most patterns, rendering improvement through frequent retraining marginal (C16 vs. C17). Compared to model selection, the improvement potential of retraining is at least in the same range and in many cases larger.

5.1.2. Computational complexity

An overview of the computational complexities for all cases is provided in Table 4. As for $rRMSE_{avg}$, relative complexities of each case against the highest value for a given variable and prosumer are shown.

Time complexity. From Table 4 it can be seen that model selection is the dominating factor for both training and prediction time complexity (rO_{train} and rO_{pred}). Cases without selection appear as 0 because their complexities are orders of magnitude smaller. The time required for the considered model selection process (see Section 3.2.4) and training of one GBDT model is in the range of 2 to 24 h on a high-performance computing (HPC) cluster [37]. This process could not be realized locally using the hardware of a residential EMS and would require cloud computing. In contrast, training a default model is in the order of seconds to minutes on a standard laptop, allowing local realization. The strong impact of model selection results from the entailed increase of model size, which is a key driver of rO_{train} and rO_{pred}. For example, in most cases the number of decision trees raises from 100 (default value) to over 7000. In addition, selection requires training of N_{trails} models, which further increases rO_{train}.

The use of weather forecasts as model input also impacts rO_{train} and rO_{pred}, as can be seen from comparison of C13 and C17. This can be mainly explained by a higher feature dimension. Retraining does not affect rO_{pred} as it does not increase model size (e.g., C12 vs. C13). In contrast, since retraining is conducted on larger data than initial training, rO_{train} increases as the comparison of C16 and C17 shows. Nevertheless, the impact of additional weather input and retraining is negligible in cases without model selection.

Space complexity. Space complexity is also dominated by model selection (see Table 4). In these cases, rM is orders of magnitude larger since the defined hyperparameter space (see Table 3) requires evaluation of models of significantly larger size compared to the default model.

5.1.3. Trade-off between accuracy and complexity

The highest accuracy (apart from oracle case C18) is achieved by GBDT-based forecasting using external weather forecasts, model selection and retraining (C9 and C17). As can be seen from Table 4, this comes at the cost of high time and space complexity. However, second or third best results are achieved by models with significantly smaller complexity (C7 and C15). In these cases, a default model is combined with the use of external weather data and retraining. The loss in accuracy in terms of $rRMSE_{avg}$ is in the range of 1 to 2 percentage points. It can be concluded that C7 and C15 provide best trade-offs with regards to accuracy and complexity. An example of the forecasts under C7 is depicted in Fig. 5.

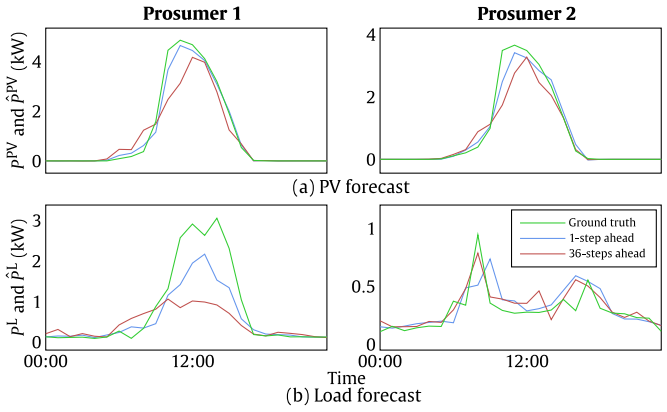

Fig. 5. Representative excerpts of 1- and 36-steps (hours) ahead forecasts of (a) PV production and (b) load consumption.

5.1.4. Error development over the prediction horizon

Understanding the error behavior over the forecasting horizon adds further insight to the comparison of average performance values in Section 5.1.1. The $nRMSE^{PV,C9}$ and $nRMSE^{L,C9}$ are depicted in Fig. 6 over the prediction horizon.

In all cases the error increases with the horizon. This constitutes a typical behavior as predictions further into the future are usually more difficult. PV forecast errors of both prosumers follow a similar trend, which can be explained by their local proximity. The error increases rapidly up to a four-steps ahead horizon. Thereafter, the gradient drops. The accuracy of load forecasts is lower than for PV forecasts, which can be attributed to the randomness of consumer behavior. Moreover, the $nRMSE^{L,C9}$ evolves differently among the prosumers. While errors begin in a similar range at one-step ahead, they follow different gradients over the consecutive horizon. The error of prosumer 2 almost stays constant. This stems from comparatively similar load patterns among different days. The resulting correlation with time of the day renders load forecasting rather a regression problem, explaining the stable $nRMSE^{L,C9}$ over the horizon. For prosumer 1 the error increases rapidly until five-steps ahead, followed by a saturation phase. In this case, load patterns exhibit more variation between different days due to less predictable EV charging. The error can be kept small within the first hours, due to similarity with the most recent lag values of P^{L}. Once this effect cancels out, the stronger variation results in higher $nRMSE^{L,C9}$ values compared to prosumer 2.

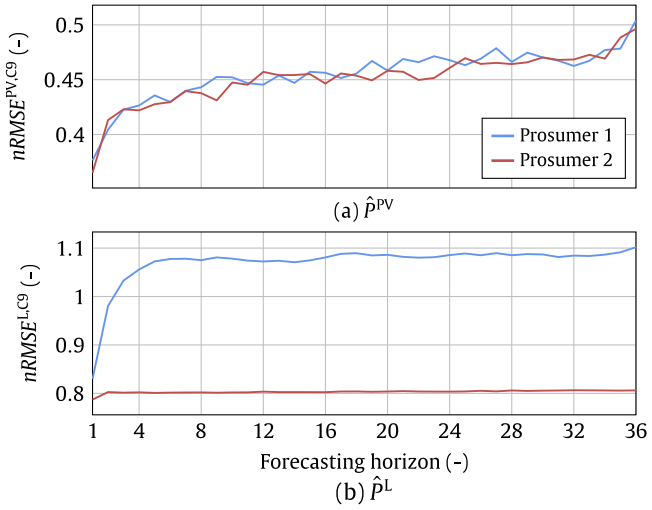

Fig. 6. *nRMSE* of (a) $\hat{P}^{PV}$ and (b) $\hat{P}^{L}$ for both prosumers over the forecasting horizon exemplary on C9.

5.2. Performance evaluation of EMS scenarios

This subsection evaluates the economic benefit of the EMS scenarios, considering trade-offs with computational complexity and security. Section 5.2.1 provides a summary of the prosumers energy and cost quantities. In Section 5.2.2, scenarios are compared and recommendations on best trade-offs provided. Based on the suggested scenario, the impact of time of the year (Section 5.2.3), optimization horizon (Section 5.2.4) and data manipulation (Section 5.2.5) on the economic value is assessed.

5.2.1. Prosumers overview

Table 5 provides an overview of the prosumers energy and cost characteristics. Prosumer 1 exhibits higher production and consumption values over the 14-month evaluation period. Both exhibit negative energy costs[3] without using a battery, equal to $K^{base} = -282$€ and -43€, respectively. The addition of a storage system enables a maximum additional benefit of $B^{C18} = 466$€ and 555€ for the theoretical EMS scenario considering optimization based on perfect forecasts (C18). Consequently, the maximum total revenue over the evaluation period amounts to 748€ and 598€, respectively. The offline rule-based benchmark, which minimizes energy exchanges, achieves 45% and 56% of B^{C18} for prosumer 1 and 2, respectively. The following evaluation examines what fraction of the maximum theoretical benefit B^{C18} the various EMS scenarios achieve based on their underlying forecasting case.

Table 5
Overview of energy and cost quantities of prosumer 1 and 2 based on the full 14-month evaluation period.

Energy and cost quantities	Prosumer 1	Prosumer 2
PV production (kWh)	7505	5748
Consumption (kWh)	4766	3210
Base cost K^{base} (€)	−282	−43
Max benefit B^{C18} (€)	466	555
Rule-based benefit B^{rb} (€)	209	310

5.2.2. Scenario comparison and recommendation

In Fig. 7, the relative benefit *rB* over the entire 14-month evaluation period is depicted for all EMS scenarios.

Impact of price forecast. The impact of spot price forecast accuracy is assessed on C17 (prosumer 1) and C9 (prosumer 2). Perfect price forecasts increase *rB* by 0.0004 (prosumer 1) and 0.0019 percentage points (prosumer 2). This translates to an additional benefit of 0.2051€ and 1.072€, respectively, over a period of 14 months. It can be concluded that sophisticated price forecasts as extension of available prices are not required for cost-optimal control of residential PV-battery systems.

Impact of PV and load forecasts. From C1–C9, it can be seen that differences of *rB* between scenarios exhibit similar trends for the two prosumers. In both cases, even naïve persistence-based optimization (C1) achieves significant improvements compared to offline rule-based control (78% and 86% of the theoretical optimum), without need for model training and selection or external weather forecasts. Consequently, even in the simplest case, rolling-horizon optimization enables additional gains of 152.11€ and 168.2€ to the prosumers compared to the offline rule-based scheme. Nevertheless, with the exception of C2, all GBDT-based scenarios outperform persistence-based optimization, motivating the use of ML. This is in line with the $rRMSE_{avg}$-based findings presented in Section 5.1.1.

As can be seen from Fig. 7, highest relative benefits are achieved by C7, C9, C15 and C17 ($rB = 0.9$ for prosumer 1 and $rB = 0.93$ for prosumer 2). This translates to additional income of 56.01€ and 36.53€, respectively, through use of ML models compared to simple persistence forecasting. Among these scenarios, C7 exhibits advantages from a computational complexity perspective, since it avoids model selection (opposed to C9 and C17) and dependency on large historical data (in contrast to C15 and C17). On the one hand, this indicates that ML-based forecasts are also economically beneficial for new systems without extensive data history. On the other hand, it suggests that the computational burden of extensive model selection (see Section 5.1.2) is economically not justified.

[3] Negative energy costs result from the fact that revenues for PV production exceed electricity costs in the evaluation period.

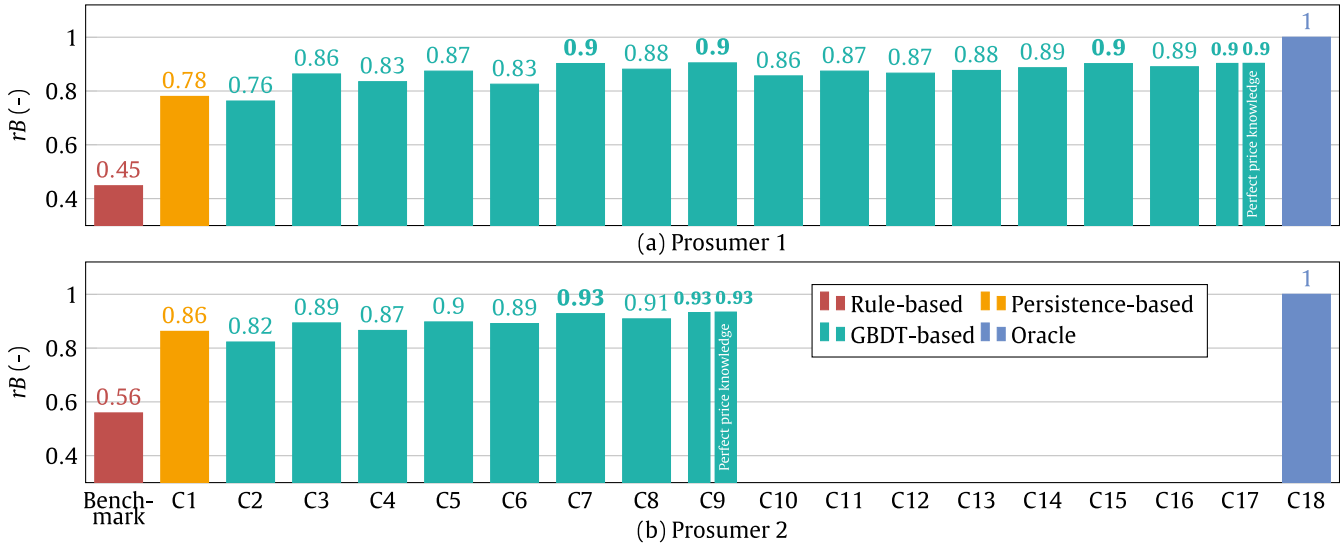

Fig. 7. Relative benefit rB of all EMS scenarios over the full evaluation period for (a) prosumer 1 and (b) prosumer 2. Best values are written in bold (excluding oracle). Impact of assuming perfect future spot price knowledge is exemplary shown on C17 and C9.

The computational simplicity of C7 also translates to benefits in terms of security. The use of default parameters facilitates local implementation, avoiding any need for externalization of PV and load forecasts to cloud-based solutions (see Section 5.1.2). Thus, no (potentially sensitive) data need to be provided to third-parties. Nevertheless, C7 is dependent on external weather forecasts, which might introduces opportunities for adversaries, as further evaluated in Section 5.2.5. In contrast, C3 avoids using external weather data, while exhibiting the same advantages in terms of computational complexity. However, the potential security advantage comes at the cost of a benefit reduction of 3-4 percentage points, which translates to 19.09 € for prosumer 1 and 17.76 € for prosumer 2. Note that all optimization-based scenarios (C1–C18) require external price data. Thus, to run the residential energy system (Fig. 2) isolated from public networks, offline rule-based control is the only opportunity among the considered EMS scenarios. The impact of price manipulations is evaluated in Section 5.2.5.

To conclude, the EMS scenario providing best trade-offs in terms of economic profitability and computational complexity is seen in rolling-horizon optimization based on forecasts of a default GBDT model, using a short two weeks initial training set, external weather forecast inputs and weekly retraining (C7). It achieves the same financial gain as models resulting from extensive selection processes at significantly lower computational costs. Moreover, it can be applied to new systems with short data history. The small computational burden also eases local implementation, providing data security advantages. Nevertheless, for a holistic assessment of C7, the sensitivity to attacks on required external data streams (price and weather forecasts) must be quantified, which follows in Section 5.2.5. In the subsequent sections, C7 is considered as representative case for GBDT-based forecasts.

5.2.3. Impact of time of the year

Fig. 8 depicts the monthly relative benefit rB_m for offline rule-based control as well as persistence-, GBDT- and oracle-based optimization. Relative benefits are volatile under the rule-based scheme in both prosumer cases, ranging from $rB_m = 0.102$ to $rB_m = 0.887$. The pronounced under-performance of rule-based control around December 2021 and July 2022 is driven by two factors. On the one hand, the margin for battery utilization is low, either due to small PV production (December) or low consumption because of holidays (July). While optimization exploits the remaining benefits, the battery is barely used under rule-based control. On the other hand, the respective months exhibit particularly high and volatile prices. As the considered rule-based scheme only discharges to cover load demand, high prices cannot be actively exploited by grid exports.

Optimization-based battery scheduling provides more stable values over the year, even in case of persistence forecasts. Therefore, the difference between persistence- and GBDT-based optimization is largely stable. While persistence-based optimization is outperformed by rule-based control in some months, GBDT-based optimization achieves the best results across the entire evaluation period.

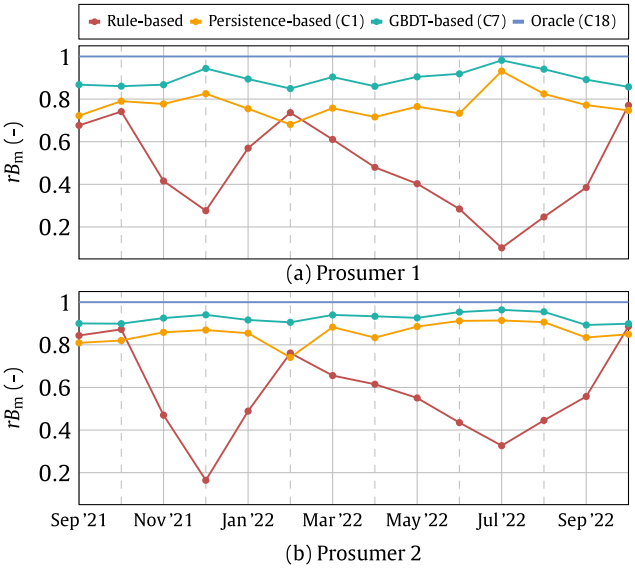

Fig. 8. Monthly relative benefit rB_m for (a) prosumer 1 and (b) prosumer 2.

5.2.4. Impact of prediction horizon

In Fig. 9, the relative benefit rB is depicted for varying forecasting horizons.[4] The oracle-based scenario (C18) and offline rule-based benchmark are included for the sake of easier comparison, although they do not depend on w_{pr} and therefore remain constant. On horizons $w_{pr} < 4$, rule-based control outperforms persistence- and GBDT-based optimization. This can be explained by under-utilization of the battery due to the 50% SOC condition at the end of the optimization horizon (see Section 3.1.2). To satisfy this constraint for short horizons, the battery remains at a SOC close to 50% even in periods of abundant PV production, instead of charging. Removing this constraint further deteriorates performance, as the myopic optimal decision is to fully discharge the battery. For $w_{pr} \geq 4$, optimization outperforms rule-based control. In case of GBDT-based optimization, the relative benefit saturates around $w_{pr} = 16$ at $rB = 0.9$ (prosumer 1) and $rB = 0.93$ (prosumer 2), respectively. Thus, horizons between $w_{pr} = 16$ and $w_{pr} = 20$ should be favored because the additional economic benefit of a longer horizon is negligible and unnecessarily increases complexity.

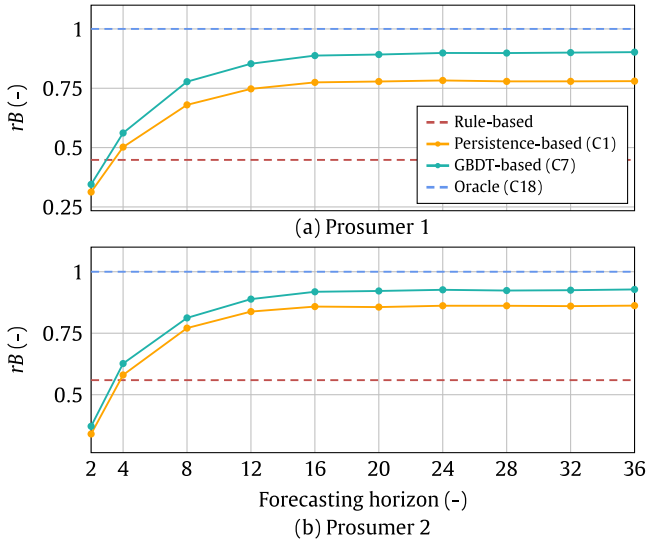

Fig. 9. Relative benefit rB over the forecasting horizon for (a) prosumer 1 and (b) prosumer 2.

5.2.5. Impact of data manipulation

The impact of weather forecast and price manipulation on the economic benefits is evaluated on C7 and shown in Fig. 10.

Weather forecast manipulation. The manipulation of weather forecast model input exhibits minor economic impact on both prosumers. Even for $\alpha = 10$, the relative benefit over the entire evaluation period only drops from $rB^{C7} = 0.902$ to $rB^{C7} = 0.877$ (prosumer 1) and $rB^{C7} = 0.928$ to $rB^{C7} = 0.904$ (prosumer 2), which translates to loss of 11.62 € and 13.53 €, respectively. This behavior can be explained by two factors. (1) Regular retraining allows the model to recognize and react on reduced information content of weather forecasts by putting less weight on these inputs. If the weather forecasts would contain no information, the model would approximate C3, which neglects external weather data. Therefore, the maximum reduction which can result from weather forecast manipulation can be quantified for C7 as $rB^{C7} = rB^{C3} = 0.863$ (prosumer 1) and $rB^{C7} = rB^{C3} = 0.894$ (prosumer 2), respectively. (2) As the introduced errors are randomly distributed around the true values, simultaneously processing a sequence of $w_{pr} = 36$ steps allows the model to derive a rolling average of the weather inputs. Therefore, they provide information even under high noise levels as the case for $\alpha = 10$ (see Fig. 4). It can be concluded that manipulation of external weather forecasts constitutes only a small economic risk for prosumers applying forecasting-based optimization on their PV-battery system. Therefore, avoiding use of external weather forecasts for security reasons is not well justified. This supports the highlighting of C7 in Section 5.2.2 as best trade-off in terms of profitability, complexity and security within the considered scenarios.

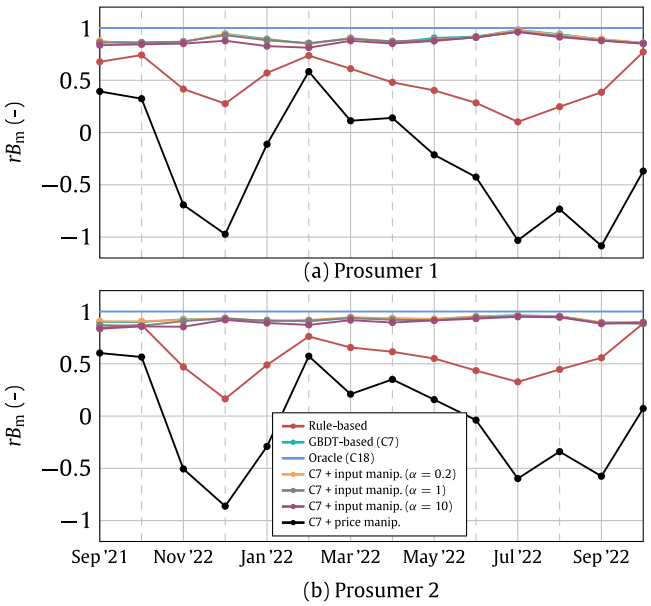

Fig. 10. Impact of data manipulation on the monthly relative benefit rB_m for (a) prosumer 1 and (b) prosumer 2.

Spot price manipulation. As can be seen from Fig. 10, price manipulation severely decreases the relative benefits in both prosumer cases. Benefits are lower than under offline rule-based control for each month of the evaluation period. Over the entire evaluation period, values drop from $rB^{C7} = 0.902$ to $rB^{C7} = -0.362$ (prosumer 1) and $rB^{C7} = 0.928$ to $rB^{C7} = -0.099$ (prosumer 2). This translates to a loss of 579.32 € and 570.34 € respectively, compared to the non-manipulated cases. It can be concluded that a manipulation of spot prices would reduce the economic benefit of optimization-based battery scheduling drastically and even generate additional cost compared to a scenario without a battery. Since all optimization-based scenarios (C1–C18) depend on price data, only offline rule-based control mitigates such risk within the considered EMS scenarios.

Although the potential impact of price manipulation is high, it would require an attack to last for months. To avoid price manipulations remaining undetected over long periods, residential EMSs should be equipped with concepts for spot price data integrity checking. In this case, forecasting-based rolling-horizon optimization according to C7 still provides the best trade-off in terms of economic profitability, computational complexity and security within the evaluated scenarios.

6. Discussion

In this section, implications of the results from Section 5 are discussed in a broader context. Considered aspects include the

[4] Note that $w_{pr} = 1$ is not included, since it may often lead to infeasible optimization problems due to the ending SOC constraint.

use of ML in residential EMSs (Section 6.1), result transferability (Section 6.2) and market readiness (Section 6.3).

6.1. ML for residential EMSs

Scenario C7 demonstrates that highest economic benefits from PV-battery systems can even be obtained by using default ML forecasting models with almost no initial training data. Instead of model selection and extensive training data, the incorporation of weather forecasts and frequent retraining is of key relevance. This importance results from characteristics of residential production and consumption patterns. On the one hand, the strong correlation of PV production and weather features enables accurate forecasts even with simple models. On the other hand, residential consumption exhibits regular changes (e.g., through new electronic devices), which reduces the value of extensive data histories and explains the importance of frequent retraining. These findings suggest that ML-based forecasting is beneficial for residential EMSs. However, data- and computation-intensive approaches, including deep learning, are not suitable and justifiable.

An alternative to combining ML-based forecasting with optimization may be seen in more advanced rule-based concepts, which include price and weather forecast information. An argument often used in favor of the latter is low computational burden. However, the simplicity of C7 can hardly be undercut, and it avoids manual development and tweaking of control rules. Both concepts can be locally implemented and thus avoid provision of sensitive consumption data to third parties. However, ML-based forecasting additionally exhibits robustness against weather input manipulations, since it automatically puts less weight on affected features through retraining. Last but not least, rule-based approaches can only approximate the economic benefits achieved by forecasting-based optimization.

6.2. Transferability

Section 5 shows similar impact of forecasting strategies on both prosumers. For example, relative benefits exhibit the same trends over varying forecasting cases (see Fig. 7) and horizons (see Fig. 9). Since prosumers with substantially different production and consumption levels, patterns and uncertainties are considered, this similarity points towards transferability of recommendations on forecasting strategies to other prosumers. Nevertheless, case studies with large and versatile prosumer portfolios covering different locations, weather conditions and component dimensions are required to substantiate the findings.

The main difference between the two is a lower benefit level across all scenarios for prosumer 1. This is explained by the prosumers' contrast with respect to predictability and optimization potential. While prosumer 2 exhibits strong repetitiveness in load patterns, EV charging introduces more randomness in the other case. Moreover, the active alignment of EV charging to PV production reduces the margin for further load optimization in contrast to the passive behavior of prosumer 2. Given that the two considered users represent rather extreme cases, it can be expected that the relative benefits of other prosumers in many cases will lie between those two.

Although results are based on prosumers located in Denmark, findings transfer to other regions with similar instantaneous netting schemes, for example, Belgium and parts of the United States (Nevada, Arizona and New York) [13]. Further states and confederations move towards instantaneous metering (e.g., Netherlands) or are promoting the roll-out of SMs (European Union), which provide the technical means for employing time-varying prices and short netting intervals.

6.3. Market readiness

The data [17], models [19] and optimization algorithms [15] used in C7 can already be acquired and used free of charge for private use. Moreover, simple hardware in the range of existing EMSs or small single-board computers is capable of hosting such applications. Therefore, forecasting-based optimization as in C7 can be considered market ready.

7. Conclusion and future work

In this work, trade-offs between economic profitability, computational complexity and security of forecasting-based optimization in residential EMSs are evaluated. Two PV-battery systems of real prosumers exhibiting different production and consumption characteristics serve as the foundation of the study. Several forecasting cases are considered, which result from variations of model type, data availability and modeling strategies. The resulting EMS scenarios and underlying forecasts are systematically quantified and assessed regarding forecasting accuracy, computational complexity and economic benefits, including sensitivity analyses on time of the year and length of the forecasting horizon. Moreover, two data manipulation scenarios are included to quantify possible attack impact and assess the EMS scenarios in terms of security. Results show that the theoretical maximum benefit over a 14-month period in the two prosumer cases is 466 € and 555 €, respectively, compared to a scenario without battery. Optimization based on naïve persistence forecasts achieves 78% (prosumer 1) and 86% (prosumer 2) of this upper limit. The relative benefits further raise to 90% and 93%, respectively, in scenarios considering GBDT-based forecasts. This performance increase already is achieved in a scenario which (1) can be applied to new systems with short data history and (2) can be implemented locally without need for extensive computing resources (e.g., cloud computing). The highlighted scenario does not depend on sophisticated price forecasts and is tolerant against manipulations of weather model inputs. However, due to sensitivity to price manipulations, incorporation of concepts for price data integrity checking into residential EMSs should be considered.

Future studies should evaluate the profitability, complexity and security trade-offs for other residential energy systems, such as electric vehicle- or heat-pump-based setups. Another aspect of interest is to understand if intra-hourly forecasts can provide further benefits despite the high randomness of load consumption. Finally, new error metrics for forecast model selection should be developed, which improve translation of forecasting accuracy to financial gains and thus might increase efficiency and benefits of hyperparameter selection.

CRediT authorship contribution statement

Nils Müller: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Mattia Marinelli:** Resources, Writing – review & editing, Supervision. **Kai Heussen:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Charalampos Ziras:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft.

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.

Data availability

The data that has been used is confidential.

Acknowledgment

This work is funded by the Innovation Fund Denmark (IFD) under File No. 91363, and the EU Horizon project EV4EU under grant agreement 101056765.

References

[1] H. Holttinen, J. Kiviluoma, D. Flynn, J.C. Smith, A. Orths, P.B. Eriksen, N. Cutululis, L. Söder, M. Korpås, A. Estanqueiro, J. MacDowell, A. Tuohy, T.K. Vrana, M. O'Malley, System impact studies for near 100% renewable energy systems dominated by inverter based variable generation, IEEE Trans. Power Syst. 37 (4) (2022) 3249–3258, http://dx.doi.org/10.1109/TPWRS.2020.3034924.

[2] J. Ostergaard, C. Ziras, H.W. Bindner, J. Kazempour, M. Marinelli, P. Markussen, S.H. Rosted, J.S. Christensen, Energy security through demand-side flexibility: The case of Denmark, IEEE Power Energy Mag. 19 (2) (2021) 46–55, http://dx.doi.org/10.1109/MPE.2020.3043615.

[3] L. Söder, P.D. Lund, H. Koduvere, T.F. Bolkesjø, G.H. Rossebø, E. Rosenlund-Soysal, K. Skytte, J. Katz, D. Blumberga, A review of demand side flexibility potential in northern europe, Renew. Sustain. Energy Rev. 91 (2018) 654–664, http://dx.doi.org/10.1016/j.rser.2018.03.104.

[4] X. Han, J. Garrison, G. Hug, Techno-economic analysis of PV-battery systems in Switzerland, Renew. Sustain. Energy Rev. 158 (2022) http://dx.doi.org/10.1016/j.rser.2021.112028.

[5] D. Azuatalam, K. Paridari, Y. Ma, M. Förstl, A.C. Chapman, G. Verbič, Energy management of small-scale PV-battery systems: A systematic review considering practical implementation, computational requirements, quality of input data and battery degradation, Renew. Sustain. Energy Rev. 112 (2019) 555–570, http://dx.doi.org/10.1016/j.rser.2019.06.007.

[6] D. Azuatalam, M. Förstl, K. Paridari, Y. Ma, A.C. Chapman, G. Verbič, Techno-economic analysis of residential PV-battery self-consumption, in: 2018 Asia-Pacific Solar Research Conference (APSRC), Vol. 186, 2018, pp. 171–178, https://scholar.google.com/scholar?hl=de&as_sdt=0%2C5&q=Techno-economic+Analysis+of+Residential+PV-battery+Self-consumption&btnG=.

[7] J. Salpakari, P. Lund, Optimal and rule-based control strategies for energy flexibility in buildings with PV, Appl. Energy 161 (2016) 425–436, http://dx.doi.org/10.1016/j.apenergy.2015.10.036.

[8] C. Sun, F. Sun, S.J. Moura, Nonlinear predictive energy management of residential buildings with photovoltaics & batteries, J. Power Sources 325 (2016) 723–731, http://dx.doi.org/10.1016/j.jpowsour.2016.06.076.

[9] Y. Zhang, R. Wang, T. Zhang, Y. Liu, B. Guo, Model predictive control-based operation management for a residential microgrid with considering forecast uncertainties and demand response strategies, IET Gener., Transm. Distribution 10 (10) (2016) 2367–2378, http://dx.doi.org/10.1049/iet-gtd.2015.1127.

[10] Y. Iwafune, T. Ikegami, J.G. da Silva Fonseca, T. Oozeki, K. Ogimoto, Cooperative home energy management using batteries for a photovoltaic system considering the diversity of households, Energy Convers. Manage. 96 (2015) 322–329, http://dx.doi.org/10.1016/j.enconman.2015.02.083.

[11] M. Elkazaz, M. Sumner, S. Pholboon, R. Davies, D. Thomas, Performance assessment of an energy management system for a home microgrid with PV generation, Energies 13 (13) (2020) http://dx.doi.org/10.3390/en13133436.

[12] D. van der Meer, G.C. Wang, J. Munkhammar, An alternative optimal strategy for stochastic model predictive control of a residential battery energy management system with solar photovoltaic, Appl. Energy 283 (2021) http://dx.doi.org/10.1016/j.apenergy.2020.116289.

[13] C. Ziras, L. Calearo, M. Marinelli, The effect of net metering methods on prosumer energy settlements, Sustain. Energy Grids Netw. 27 (2021) http://dx.doi.org/10.1016/j.segan.2021.100519.

[14] K. Shivam, J.-C. Tzou, S.-C. Wu, A multi-objective predictive energy management strategy for residential grid-connected PV-battery hybrid systems based on machine learning technique, Energy Convers. Manage. 237 (2021) http://dx.doi.org/10.1016/j.enconman.2021.114103.

[15] GNU project, 2023, https://www.gnu.org/software/glpk/, (Accessed: 2023-03-07).

[16] S. Diamond, S. Boyd, CVXPY: A python-embedded modeling language for convex optimization, 2022, https://www.cvxpy.org, (Accessed: 2022-12-15).

[17] Free rooftop solar forecasting, 2022, https://solcast.com/free-rooftop-solar-forecasting, (Accessed: 2022-11-24).

[18] Current weather and forecast - OpenWeatherMap, 2022, https://openweathermap.org/, (Accessed: 2022-11-24).

[19] J. Herzen, F. Lässig, S.G. Piazzetta, T. Neuer, L. Tafti, G. Raille, T. Van Pottelbergh, M. Pasieka, A. Skrodzki, N. Huguenin, M. Dumonal, J. Kościsz, D. Bader, F. Gusset, M. Benheddi, C. Williamson, M. Kosinski, M. Petrik, G. Grosch, Darts: User-friendly modern machine learning for time series, J. Mach. Learn. Res. 23 (124) (2022) 1–6.

[20] J.H. Friedman, Greedy function approximation: A gradient boosting machine, Ann. Statist. 29 (5) (2001) 1189–1232.

[21] C.S. Bojer, J.P. Meldgaard, Kaggle forecasting competitions: An overlooked learning opportunity, Int. J. Forecast. 37 (2) (2021) 587–603, http://dx.doi.org/10.1016/j.ijforecast.2020.07.007.

[22] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, 2016, pp. 785–794, http://dx.doi.org/10.1145/2939672.2939785.

[23] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, LightGBM: A highly efficient gradient boosting decision tree, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS '17, 2017, pp. 3149–3157, https://scholar.google.de/scholar?hl=de&as_sdt=0%2C5&q=LightGBM%3A+A+highly+efficient+gradient+boosting+decision+tree%2C+proceedings.neurips&btnG=.

[24] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, A. Gulin, CatBoost: unbiased boosting with categorical features, in: S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett (Eds.), Advances in Neural Information Processing Systems, Vol. 31, 2018.

[25] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, 2019, pp. 2623–2631, http://dx.doi.org/10.1145/3292500.3330701.

[26] R. Hyndman, G. Athanasopoulos, Forecasting: Principles and Practice, second ed., OTexts, Australia, 2018, https://scholar.google.com/scholar?hl=de&as_sdt=0%2C5&q=+Forecasting%3A+principles+and+practice+hyndman+2018&btnG=.

[27] LightGBM documentation, 2022, https://lightgbm.readthedocs.io/en/latest/Parameters.html, (Accessed: 2022-11-24).

[28] J. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyper-parameter optimization, in: Advances in Neural Information Processing Systems, Vol. 24, 2011.

[29] Day-ahead prices, 2022, https://www.nordpoolgroup.com/en/Market-data1/Dayahead/, (Accessed: 2022-11-24).

[30] A. Jedrzejewski, J. Lago, G. Marcjasz, R. Weron, Electricity price forecasting: The dawn of machine learning, IEEE Power Energy Mag. 20 (3) (2022) 24–31, http://dx.doi.org/10.1109/MPE.2022.3150809.

[31] Carnot spot price forecast, 2022, https://www.carnot.dk/, (Accessed: 2022-11-24).

[32] C. Kolias, G. Kambourakis, A. Stavrou, J. Voas, DDoS in the IoT: Mirai and other botnets, Computer 50 (7) (2017) 80–84, http://dx.doi.org/10.1109/MC.2017.201.

[33] B. Bostami, M. Ahmed, S. Choudhury, False data injection attacks in internet of things, in: Performability in Internet of Things, Springer International Publishing, 2019, pp. 47–58.

[34] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, Introduction to Algorithms, third ed., The MIT Press, 2009, https://scholar.google.com/scholar?hl=de&as_sdt=0%2C5&q=Introduction+to+algorithms+2022&btnG=.

[35] H.M. Sani, C. Lei, D. Neagu, Computational complexity analysis of decision tree algorithms, in: International Conference on Innovative Techniques and Applications of Artificial Intelligence, Springer, 2018, pp. 191–197.

[36] Time complexities of machine learning algorithms, 2022, https://7-hiddenlayers.com/time-complexities-of-ml-algorithms, (Accessed: 2022-11-24).

[37] DTU Computing Center, DTU Computing Center resources, 2022, http://dx.doi.org/10.48714/DTU.HPC.0001.