



**HAL**  
open science

## Evaluating forecasting methods in the context of local energy communities

Jonathan Coignard, Maxime Janvier, Vincent Debusschere, Gilles Moreau, Raphaël Caire, Stéphanie Chollet

► **To cite this version:**

Jonathan Coignard, Maxime Janvier, Vincent Debusschere, Gilles Moreau, Raphaël Caire, et al.. Evaluating forecasting methods in the context of local energy communities. International Journal of Electrical Power & Energy Systems, 2021, 131, pp.106956. 10.1016/j.ijepes.2021.106956 . hal-03636300

**HAL Id: hal-03636300**

**<https://hal.science/hal-03636300v1>**

Submitted on 9 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Evaluating Forecasting Methods in the Context of Local Energy Communities

Jonathan Coignard<sup>†‡</sup>, Maxime Janvier<sup>‡</sup>, Vincent Debusschere<sup>†</sup>, Gilles Moreau<sup>‡</sup>, Stéphanie Chollet<sup>§</sup>,  
Raphaël Caire<sup>†</sup>

<sup>†</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP\*, G2Elab, 38000 Grenoble, France

<sup>§</sup>Grenoble INP - LCIS, University Grenoble Alpes, Grenoble, France

<sup>‡</sup>Lancey Energy Storage, Grenoble, France

\*Institute of Engineering Univ. Grenoble Alpes

Email: [jonathan.coignard@grenoble-inp.fr](mailto:jonathan.coignard@grenoble-inp.fr)

**Abstract**—To push the boundaries of self-sufficiency, local energy communities may rely on load demand forecasts to schedule energy usage ahead of time. In the perspective of selecting a forecasting method, this work explores goodness of a forecast from two sides: its quality, and its value. Traditionally, forecasting methods are ranked based on quality metrics such as the Mean Absolute Percentage Error (MAPE). This work additionally considers the value of a forecast, quantifying practical outcomes for local energy communities such as self-sufficiency, cost of electricity, and fairness. Our contribution is two-fold: first in creating a broader framework to evaluate forecasting performance with regards to energy communities, and second, in highlighting the relationship between quality and value metrics for energy communities ranging from 2 to 95 participants. When selecting a forecasting method, our results show that simply relying on quality metrics is misleading. This paper illustrates with a study case, the clear difference in considering value metrics rather than quality metrics, and the potential impact on the design of a coordination platform for local energy communities.

## I. INTRODUCTION

By 2021, member states of the European Union are required to have a framework for renewable energy communities [1]. At the core, renewable energy communities (or "energy communities" in this paper) aim at consuming electricity when it is locally generated. From the grid's perspective, matching production and consumption of electricity is a positive improvement, which previous tariffs such as net-metering or feed-in-tariff could not achieve. However, from the community's perspective matching production and consumption is challenging [2]. One of the option to increase local energy consumption is to equip communities with controllable loads and batteries. To control such devices, numerous strategies rely on demand forecasts. This is the case for decentralized scheduling algorithms [3], as for dedicated market designs [4]. To this end, the literature and various libraries offer a wide variety of forecasting methods to predict short term electricity demand [5]. Naturally, if we are to rely on a specific forecasting method to schedule energy usage, the clarification of performance (or goodness) of a forecast appears to be a worth while objective.

To evaluate or rank load forecasting methods, performances are often determined using accuracy metrics such as Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) [6]. Those metrics are intrinsic to each forecast,

with no consideration for the context in which forecasts are used. In [7], the author defines goodness of a forecast from two aspects: *quality* and *value*. Where *quality* relates to the correspondence between forecasts and matching observations (e.g., accuracy), and *value* corresponds to the operational gain from considering forecasts inside a process of functional blocs, at the decision-making stage. In the context of local energy communities, value metrics may take multiple forms, if the community prioritizes financial aspects [8], self-sufficiency [9], fairness among members [10], or sustainability [11]. Importantly, the relation between measurements of quality and value of a forecast is far from being trivial. In particular this relation may exhibit thresholds, from which value is more or less influenced by a variation in quality. In most cases, this relation has no guarantees of being linear or monotonic [7], but could lead to the choice of a different algorithm than just relying on a quality-metric.

We propose to study both quality and value of demand forecasts in the context of energy communities. We jointly assess forecasts quality from aspects such as accuracy, bias, or reliability, and forecast value in term of self-sufficiency, cost of energy, and fairness among community members. Ultimately, looking at the incremental value of a forecast alongside with its intrinsic quality is key to evaluate the benefit and risks of relying on demand forecast to schedule energy usage. This approach provides tools to answer questions such as: Does improving load forecast accuracy improves community outcomes in a significant manner? — How much forecasting errors penalize some community members more than other? — Can larger storage or solar panels mitigate the impact of inaccurate load forecast? — What is the importance of real-time controls to limit the dependence on forecasts?

We illustrate the proposed frameworks of metrics with a list of common forecasting methods from moving averages to K-Nearest Neighbors applied to a simulated energy community. This work advances prior literature in the following important ways:

- Creating a broader framework to evaluate load forecasting performance with regards to energy communities;
- Highlighting the relationship between quality and value metrics for energy communities ranging from 2 to 95

participants.

We structure the remainder of the paper as follows: the next section is a literature review on performance metrics to assess forecasting quality, and value. In Section III, we describe a 'sandbox' energy community to illustrate value metrics. Section IV presents the results of performance metrics for common forecasting methods, then discusses the relationship between quality and value metrics. We conclude by highlighting additional research opportunities.

## II. EVALUATING LOAD FORECASTS

### A. Quantifying the quality of a load forecast

Once a sufficiently large sample of forecasts and observations has been collected, the relationship between the forecast-observation pairs can be used to describe different aspects of the quality of a forecast [12]. In [7], author described nine aspects (called "attributes") that contribute to the quality of a forecast. From those attributes accuracy, skill, association, bias, and reliability are commonly emphasized, or found in demand forecast literature.

- *Accuracy* describes the average difference between the individual forecasts and observations;
- *Skill* refers to the accuracy of a forecast relative to a benchmark;
- *Association* describes the strength of the linear relationship between the forecasts and observations;
- *Bias* describes the difference between the average forecast and the average observation;
- *Reliability* refers to a stratified bias or accuracy into different ranges or categories (e.g., bias during daytime or nighttime).

To practically measure the various attributes of forecast quality a panel of metrics are available. According to [6], of 50 papers, the most common error (i.e., accuracy) metrics for building energy forecast are MAPE (53%), RMSE (47%), the coefficient of variation of RMSE (38%), mean absolute error (36%), and the mean square error (16%). This is inline with [13], and [14] which report study outcomes in terms of MAPE, and RMSE. Although those error metrics are often used, they have been subject to criticisms [15]. Commonly referred attributes for good metrics are scale-independence — interpretability — sensitivity to small changes — protection against outliers — and defined in most numerical cases.

To cover *accuracy* aspects, we select two scale-independent metrics. First, to provide a familiar and interpretable reference for readers, we report MAPE in (1).

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \left( \frac{|e_t|}{|y_t|} \times 100 \right) \quad (1)$$

Where  $e_t = (y_t - f_t)$  represents the error between the observed value  $y_t$  and the forecasted value  $f_t$  at the timestep  $t$ . We further add the Mean Absolute Scaled Error (MASE), in (2), to increase confidence in the conclusions derived from the

MAPE. MASE is reliable, robust, and scale-independent, also used in the M4 forecasting competition [16].

$$\text{MASE} = \frac{1}{T} \sum_{t=1}^T \left( \frac{|e_t|}{\frac{1}{N-1} \sum_{n=2}^N |y_n - y_{n-1}|} \right) \quad (2)$$

Where  $N$  represents the number of in-sample values. Typical values for one-step MASE are less than one, as it is usually possible to obtain forecasts more accurate than a random walk. However, multi-step MASE values are often larger than one, as it becomes more difficult to forecast greater horizons [15]. For a further panorama of available *accuracy* metrics, we refer the readers to [17].

To evaluate the *bias* of forecasting methods towards under-prediction or over-prediction, we choose to use the Mean Error (ME), expressed in (3).

$$\text{ME} = \frac{1}{T} \sum_{t=1}^T e_t \quad (3)$$

We note that [18] proposes the Mean Root Error (MRE) to measure relative bias and error magnitude at the same time. MRE is defined in a complex plane using the real axis for positive errors, and the imaginary axis for negative errors. However, given that we already assess the magnitude of the error, we preferably select ME, the simpler metric.

To assess *reliability* or obtain quality metrics at a granular scale, we calculate bias and accuracy for different time-ranges. For instance, calculating metrics separately for each hour of the day might reveal that some algorithms have more bias during on-peak hours and less during off-peak hours. This granular look is a useful insight to understand the complex relationship between quality and value of a forecast, as the value of a forecasting method might depend more on being accurate during a few hours rather than being accurate overall. Reliability is also interestingly considered for weekdays and weekends, and across seasons.

Finally, although in this work we have implemented single-valued forecasts, load demand forecasts may also express degrees of certainty that an outcome occurs. In the latter case, comprehensive metrics to evaluate probabilistic forecasts are proposed in [19]. For instance, the Winkler score gives a penalty when true values lie outside of the constructed intervals, and rewards forecasters for narrow prediction intervals.

### B. Quantifying the value of a load forecast

A forecast have no intrinsic value in itself. The value is rather in the forecast's ability to influence the decision made by its users. This notion of value is relevant if we include the finality of forecasts in their assessment. In [7], the author presents four elements impacting the value of a forecast: (a) the courses of action available to the decision maker (i.e., the decision space) — (b) the payoff structure (e.g., benefit associated with the decision making) — (c) the quality of the information used as a basis in the absence of a forecast — and (d) the quality of the forecast. Although the notion of value is influenced by the forecast's quality, it mostly depends on the

systems deployed by the energy community in our case (i.e., solar panels, storage capacity, ...).

1) *The existence of a decision space*: Before measuring the value of a forecast in terms of financial gain or self-sufficiency, we need to establish the existence of different courses of actions (or decisions) that rely on the forecast. For instance, a community with a small battery, or a large battery but no reason to use it, does not create the decision space for a forecast to potentially create value. To create a decision space, we highlight two necessary conditions for energy communities: (a) to have some production surplus (e.g., at noon with solar panels), — and (b) to have controllable devices to absorb a surplus of energy. This decision space can be characterized using measures of self-consumption before flexible loads and batteries, as expressed in (5), in parallel with the available battery or defferable load power and energy characteristics (e.g., in kW and kWh).

2) *The definition of value metrics*: Usually, when demand forecasts are used to schedule controllable devices (or energy usage), it is to minimize operational costs. However, energy community may also prioritize other aspects at different levels: increasing self-sufficiency [9] — maintaining fairness among members [10] — maximizing system lifetime (e.g., reducing battery degradation) [11] — maximizing grid services (e.g., providing flexibility to the distribution grid) [20] — and ensuring that no-one is better off by himself [8]. Each objective calls for a specific metric to measure what is considered valuable by the energy community. In this section, we highlight the value of forecasting in terms of self-sufficiency and self-consumption, as this paper targets local energy communities. We also discuss metrics of fairness among community members, and operational cost. Those aspects are viewed from the perspective of the community, and quantified with metrics from the literature.

We use self-sufficiency, and self-consumption as defined in [2] to convey information on energy import and export. Intuitively, self-sufficiency is a metric to evaluate the ability of a community to provide to its own energy needs (i.e., less imports). Whereas, self-consumption assesses the amount of generation that remains within the community (i.e., less exports). Those two metrics are scale-independent, bounded, and easily interpretable especially when calculated over a year to account for seasonal effects. Fig. 1 illustrates a typical day where a community import energy (zone A), export energy (zone B), and consume local generation (zone C). Self-sufficiency (SSuff), and self-consumption (SCons) are formalized as follows:

$$\text{SSuff} = \frac{\sum_{t=1}^T \min(d_t, g_t)}{\sum_{t=1}^T d_t} = \frac{C}{A + C} \quad (4)$$

$$\text{SCons} = \frac{\sum_{t=1}^T \min(d_t, g_t)}{\sum_{t=1}^T g_t} = \frac{C}{B + C} \quad (5)$$

Where  $g_t > 0$  is the local energy generation, and  $d_t$  is the local energy demand from the entire community. We note that power profiles from storage systems (i.e., charging and discharging)

are merged with demand and generation profiles, such that charging is included in  $d_t$  and discharging in  $g_t$ . Achieving both 100% self-sufficiency and 100% self-consumption would indicate that the community does not import energy, and fully use local generation.

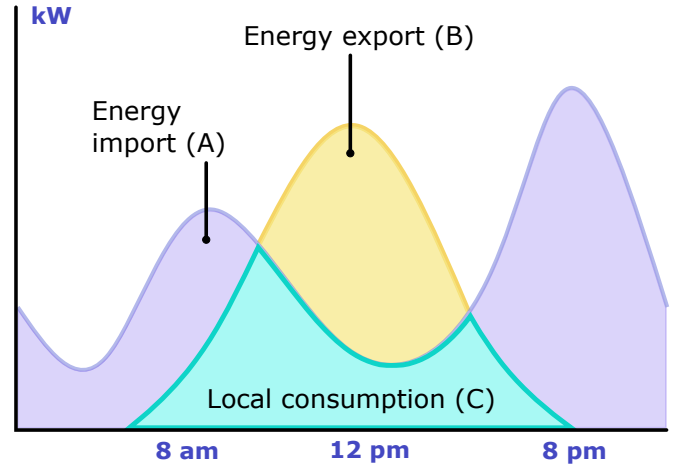


Fig. 1. Profile illustrating energy imported from the main grid (A), energy exported to the main grid (B), and local generation consumed by the community (C).

From a financial point of view, metrics are usually tailored to the specifics of a project. In [8], authors propose a cost equation for a coalition of households inspired by the “German Mieterstromgesetz” law. This equation includes fixed and variable costs to connect to the distribution grid (in € and €/kW), costs of buying electricity (in €/kWh), value from selling electricity surplus (in €/kWh), investment costs (in €), and coordination costs as a function of the coalition’s size (in €/member). In addition, this equation can be extended for communities sharing electricity on a public network (e.g., like in the French “collective self-consumption” law) by including a fee for using the public network (€/kWh).

When the objective of scheduling controllable devices is to maximize affordability, or access to energy for all members, the notion of fairness becomes a proxy for the value of the forecast. The definition of fairness is often specific to a project. For the purpose of this paper we refer to fairness as access to the locally produced energy allocated by the community. To measure fairness of access, [21] proposed a metric also used in [10], which requires a bounded range, as expressed in (6). This definition ensure an intuitive, scale-independent, and bounded metric.

$$\text{Fairness} = 1 - \frac{\sigma}{\sigma_{max}} \quad (6)$$

For energy community, we can define  $\sigma$  as the standard deviation of the percentage of local energy used by each household. The maximal standard deviation is  $\sigma_{max} = \frac{1}{2}(H - L)$ , where  $H = 100\%$  is the upper bound when one household consume all locally sourced energy, and  $L = 0\%$  the lower bound for an household which does not consume any local energy. We

obtain a relation where fairness is higher when local energy is equally shared across the community, with a maximum fairness equal to 1 whenever all members consume exactly the same amount of locally generated energy.

We should note that to measure the value of a forecast, the metrics discussed above (self-sufficiency, self-consumption, cost, and fairness) are only relevant if they are considered in the decision process (i.e., by the scheduling mechanism).

3) *The incremental payoffs:* For a forecast to acquire value (as measured by the previous metrics), it should influence users to take better decisions than the ones they would have taken without a forecast. As a result, to characterize incremental value, we consider value metrics relatively to a reference situation without forecasts. Characterizing value metrics without forecasts might be challenging. One option is to exclusively rely on a real-time controller as demonstrated in [22]. Alternatively, common naive forecasts can be used as a reference, such as moving average methods or simply using historical data from the previous day or week.

### III. MODELING ENERGY COMMUNITIES

To assess the value of forecasting methods, we model an energy community. This approach goes beyond simply looking at quality metrics (e.g., accuracy). Since we have no guarantees that a forecast with a lower MAPE necessarily leads to higher self-sufficiency, we thus developed a needed representative community model to calculate self-sufficiency.

We model the energy community as an optimization problem (described in III-B). We assume that demand forecasts are used to coordinate individual batteries in order to maximize self-sufficiency at the community level. In other words, individual homeowner batteries are charging when solar panels are generating too much power, and discharging when the community relies on the grid. This approach presumes that: (a) there is a central entity managing the resources of the community (in line with European guidelines [1]) — (b) controllable resources are scheduled for the benefit of the community as a whole - (c) the community is connected on the national grid and thus do not manage local grid constraints.

The proposed framework to model the operation of an energy community includes both a forecasting layer and an optimization layer. The overall simulation flow is sequential and repeated each day for a year. First, we ran a forecasting method using load data from the previous two weeks to predict load consumption for the upcoming two days (described in section III-A). Then, the predictions are fed to the optimization, which schedules batteries to maximize self-sufficiency over the next two days (described in section III-B). The complete Python implementation is available, as supplementary material, on GitHub<sup>1</sup>.

This model does not attempt to be generic, but rather to represent a potential energy community. We provide a simple, yet realistic modeling approach to illustrate the importance of considering value metrics when choosing a forecasting method

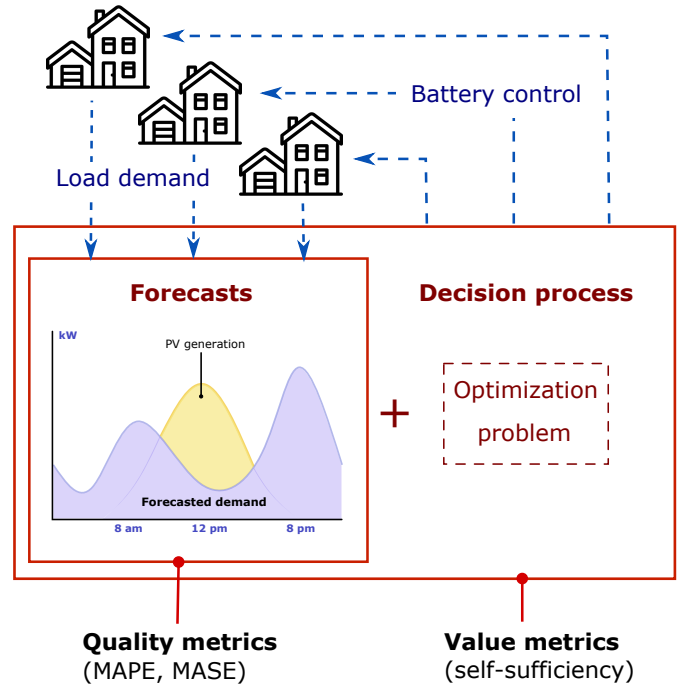


Fig. 2. Concept of quality and value of a forecast applied to an energy community.

(Fig. 2). We note that state-of-the-art energy community modeling, and control also involves numerous other concerns: grid constraints [3], consumer behavior [23], cooperative game theory [24], uncertainties and risks based approaches [25], [26]. Since our objective is to illustrate the relationship between the quality and value of a forecast, we voluntarily keep our modeling simple while being realistic (as shown in the result section).

#### A. The forecasting layer

We choose to implement common forecasting methods from the literature on load forecasts, as well as exploring various categories of techniques: traditional statistics, and classification based methods [6]. For the purpose of clarity, we cap the number of forecasting methods to six, furthermore, we do not include exogenous variables, and solely rely on historical values of the variable to predict. We note that forecasting PV production is out of our scope, and we assume that PV forecasts are available from third parties. The next paragraphs provide a rapid view into the tuning strategy for each of the forecasting methods.

1) *Moving-average:* We implement a rolling average with the prior seven days of data to account for weekly cycles. The resulting forecast is a smooth line.

2) *Seasonal naive:* The seasonal naive forecast simply reuse the prior week of data. In other words, the forecast for the upcoming Monday is the measured demand profile from the previous Monday.

<sup>1</sup>[https://github.com/Jonathan56/pylec\\_forecastpaper](https://github.com/Jonathan56/pylec_forecastpaper)

3) *Holt-Winters*: The Holt-Winters method is based on exponential smoothing [27]. For the implementation we used the statsmodels library, parametrized with the additive seasonal component.

4) *Prophet*: The Prophet method use a decomposable time series model with three main model components: trend, seasonality, and holidays [28]. For the implementation we used the facebook library, parametrized with a weekly and multiplicative seasonality.

5) *Lasso*: Least Absolute Shrinkage and Selection Operator (Lasso) is a linear model, with an added regularisation term [29]. For the implementation we used the scikit-learn library, parametrized with the constant that multiplies the L1 norm  $\alpha = 0.01$ .

6) *K-Nearest neighbors*: The K-Nearest neighbors method is based on feature similarity to predict the values of any new data points [30]. For the implementation we used the scikit-learn library, parametrized with 3 neighbors.

### B. The scheduling layer

The energy community is modeled through an optimization problem which aims at maximizing the community's self-sufficiency. Practically, this amounts for each household to charging their batteries when solar panels are generating power, and discharging their batteries in the evening. Since the optimization layer is not the focus of this publication, we keep the formulation simple. For a distributed optimization including a larger panel of controlling devices we refer the reader to [3].

The objective function minimize power consumption imported from the larger grid by the community of households  $p_{import}(t)$ . This import is the positive part of the sum over all households of uncontrolled demand ( $unctr(t)$ ) and the controlled demand ( $ctr(t)$ , i.e., batteries) as defined in (7c). The set of stationary batteries  $B$  are controlled during an horizon  $T$  to withdraw power  $p^b(t) > 0$  or inject power  $p^b(t) < 0$  within power and energy constraints ( $P_{min}^b, P_{max}^b, E_{min}^b, E_{max}^b$ ). We assume charging and discharging efficiency ( $\eta_{in}, \eta_{out}$ ), and a final state of charge defined by  $E_{end}^b$ . The optimization problem is formulated as follows:

$$\text{Minimize } \sum_{t=1}^T p_{import}(t) \quad (7a)$$

$$\text{Subject to:} \quad (7b)$$

$$\forall t = 1 \dots T$$

$$p_{import}(t) = \max(0, ctr(t) + unctr(t)) \quad (7c)$$

$$ctr(t) = \sum_{b=1}^B (p_{in}^b(t) - p_{out}^b(t)) \quad (7d)$$

$$0 \leq p_{in}^b(t) \leq P_{max}^b \quad (7e)$$

$$0 \leq p_{out}^b(t) \leq P_{min}^b \quad (7f)$$

$$E_{min}^b \leq E^b(t) \leq E_{max}^b \quad (7g)$$

$$E^b(t) = E^b(t-1) + p_{in}^b(t)\Delta t \cdot \eta_{in} - p_{out}^b(t)\Delta t / \eta_{out} \quad (7h)$$

$$E_{end}^b \leq E^b(T) \quad (7i)$$

## IV. COMPARING FORECASTS QUALITY AND VALUE

This section provide the reader with an illustrative example on the relation between quality and value of a forecast. To do so, we base our analysis on a collection of 100 real-world residential load profiles, for a year with a 15-min resolution, and solar production data from the PVGIS platform. The average yearly household consumption is 4 MWh which corresponds to residential buildings without electric heating. The PV location is set to Grenoble, France, with 1.25 kWp per household which enables between 98 % and 68 % of self-consumption for a coalition of 22 households depending if the batteries are optimized from a perfect forecast or left unused. This solar capacity installed per household ensures a sufficiently large production to justify batteries (therefore creating a decisions space for the forecast's user), while not over-sizing the solar installation. The battery capacity per household is set to 3 kWh and 2 kW output to potentially reach 98 % of self-consumption.

We illustrate our results on the concepts of quality and value for a variety of community sizes (i.e., the number of households forming the community). We create communities with 2 to 95 participants where households are randomly sampled while preserving the average yearly consumption of 4 MWh. We evaluate quality and value metrics across community sizes for several reasons:

- To increase confidence in our results for a variety of households and communities;
- To study the impact of growing communities on the quality and value of forecasting methods;
- To highlight the interest of selecting scale-independent metrics.

Each community is simulated for an entire year at a 15-min timestep based on the 'sandbox' described in the previous section. The metrics discussed in Section II are calculated at the end of the simulation. We present the results with regards to the forecast quality, value, and their interdependence.

### A. Quality of forecasts as a function of community size

We look at demand forecast quality via accuracy using scale independent metrics (i.e., MAPE, MASE) as we increase community sizes. Fig. 3 shows the evolution of MAPE, indicating as expected in [31], a better accuracy for larger community. When communities reach 95 households most forecasting methods achieve a 10 % MAPE, except the moving average forecast which only achieves 20 % MAPE. We note that for this dataset, the seasonal naive method has a similar quality as more elaborated forecasting methods such as Holt-Winters or K-Nearest Neighbors.

When we look at accuracy using MASE, Fig. 4 shows a constant accuracy for most forecasting methods except for the moving average where MASE increases linearly. The difference in trends expressed through the lens of MAPE, and MASE is interesting as it leads to different interpretations of quality: an increasing accuracy, or an accuracy rather independent on the community size.

In the light of quality metrics, It seems pretty clear that the moving average method has the worse quality both in terms

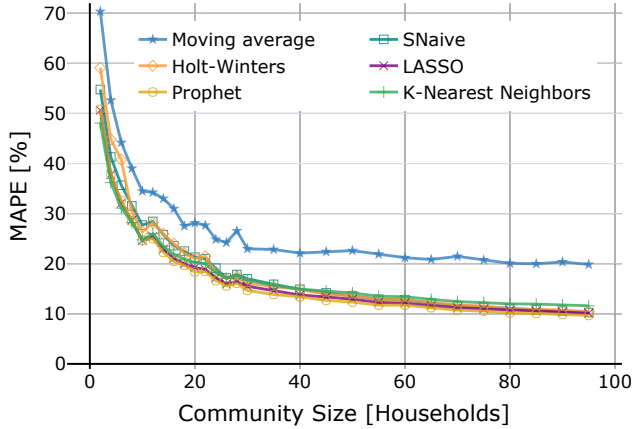


Fig. 3. MAPE as a function of community size for different forecasting methods.

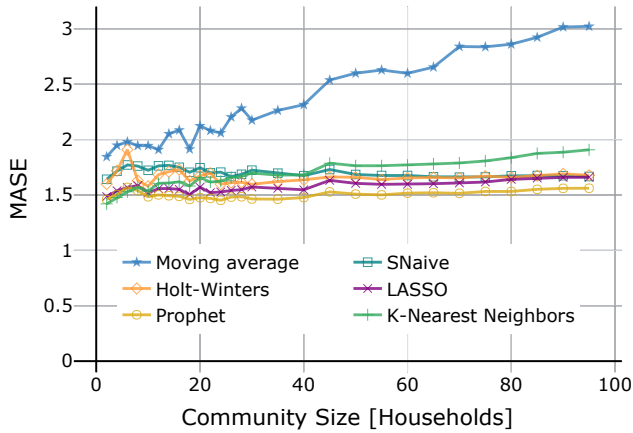


Fig. 4. MASE as a function of community size for different forecasting methods.

of MAPE and MASE. However, stratified quality provides a granular look into the results. Fig. 5 shows the average bias from the moving average forecast (to under or over estimate demand) every 15-min of the day (in green), along with when batteries deviate from the optimal schedule (in orange and blue). In this case, deviating means that batteries discharge or charge to/from energy resources outside of the community. This clearly shows that hours with the largest bias (around 4 AM and 10 PM) do not necessary lead to the largest deviation from the optimal schedule (around 11 PM, and 4 PM). In other words, a poor accuracy at night time might have a limited impact on value which is principally defined during daylight hours. This partially explains why the moving average method could lead to a value larger than expected in the light of its overall accuracy.

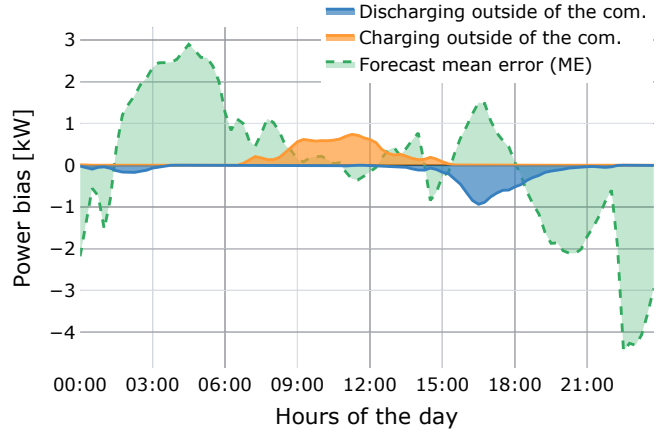


Fig. 5. Mean error from the moving average forecast (green), battery deviation from an optimal schedule leading to charging or discharging outside of the community (in orange and blue).

### B. Value of forecasts as a function of community size

To look at the value brought by forecasting methods, we calculate self-sufficiency for the community Fig. 6. It is striking to see no significant differences between forecasting methods in terms of self-sufficiency. This suggests that better quality is not synonymous with better value. We note that for community sizes above 40 households, the value gap between the perfect forecast and our proposed methods is approximately constant, around 2%. This is reasonable because self-sufficiency is already increased by 20% from a scenario without batteries.

### C. Discussion on the interest of value metrics

To select a forecasting method (our initial topic of interest), both quality metrics and value metrics agree to rank Prophet as the overall best method. In this sense, value metrics are simply confirming the perspective of quality metrics. However, we should note that quality and value metrics do not agree on the ranking of all forecasting methods. In particular, although the moving average method shows a lesser quality (both for MAPE and MASE), it performs well in terms of self-sufficiency. This behavior was illustrated in Fig. 5.

Although the forecasting value seems washed up after the optimization, it remains that MAPE is highly correlated with self-sufficiency when taking individual forecasting methods (Fig. 7). Furthermore, even though the relation between quality and value has no reasons to be linear or monotonic, it appears to be the case here. In other words, self-sufficiency increases proportionally to MAPE in the domain explored when considering a forecasting method. Interestingly, this linear relation may differ depending on the forecasting method (as illustrated by the moving average method). Thus resulting in forecasts achieving the same value for different MAPE (e.g., 50% self-sufficiency for 10% or 20% MAPE).

## V. CONCLUSIONS

Renewable energy communities should play a significant role in the energy transition toward European carbon emission goals. As demand forecasts are relied upon to control distributed energy resources, we need to verify their pertinence, not just from a quality aspect, but also from the incremental value they bring. This work explores the concept of quality and value of a forecast and proposes metrics to quantify both aspects.

Our results show that a forecast of lesser quality (e.g., in terms of MAPE) does not necessarily translate into a smaller value (e.g., self-sufficiency) for the energy community. Therefore, we highlight the importance of contextualizing forecasting methods to properly evaluate the incremental value they bring, rather than selecting a forecasting method simply from its overall accuracy. Additionally, investigating which aspect of the forecast quality is correlated with a greater value is insightful (e.g., to prioritize accuracy at certain times). Finally, estimating incremental value enables the right efforts to be invested in improving forecasting methods.

## FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## DECLARATION OF COMPETING INTEREST

The authors declared that there is no conflict of interest.

## REFERENCES

- [1] J. Jeriha, "Collective self-consumption and energy communities," tech. rep., COMPILE project, June 2019.
- [2] R. Luthander, J. Widén, D. Nilsson, and J. Palm, "Photovoltaic self-consumption in buildings: A review," *Applied Energy*, vol. 142, pp. 80–94, Mar. 2015.
- [3] E. Münsing, J. Mather, and S. Moura, "Blockchains for decentralized optimization of energy resources in microgrid networks," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 2164–2171, Aug. 2017.
- [4] B. Cornélusse, I. Savelli, S. Paoletti, A. Giannitrapani, and A. Vicino, "A Community Microgrid Architecture with an Internal Local Market," *Applied Energy*, vol. 242, pp. 547–560, May 2019.
- [5] B. Hayes, J. Gruber, and M. Prodanovic, "Short-Term Load Forecasting at the local level using smart meter data," in *2015 IEEE Eindhoven PowerTech*, pp. 1–6, June 2015.
- [6] M. Bourdeau, X. q. Zhai, E. Nefzaoui, X. Guo, and P. Chatellier, "Modeling and forecasting building energy consumption: A review of data-driven techniques," *Sustainable Cities and Society*, vol. 48, p. 101533, July 2019.
- [7] A. H. Murphy, "What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting," *Weather and Forecasting*, vol. 8, pp. 281–293, June 1993.
- [8] I. Abada, A. Ehrenmann, and X. Lambin, "On the viability of energy communities," cambridge Working Papers in Economics, Faculty of Economics, University of Cambridge, Oct. 2017.
- [9] M. Secchi and G. Barchi, "Peer to peer electricity sharing: maximising PV self-consumption through BESS control strategies," in *2019 IEEE International Conference on Environment and Electrical Engineering*, (Genova, Italy), pp. 1–6, IEEE, June 2019.
- [10] F. Moret and P. Pinson, "Energy Collectives: a Community and Fairness based Approach to Future Electricity Markets," *IEEE Transactions on Power Systems*, pp. 1–1, 2018.
- [11] H. S. V. S. K. Nunna and D. Srinivasan, "Multiagent-Based Transactive Energy Framework for Distribution Systems With Smart Microgrids," *IEEE Transactions on Industrial Informatics*, vol. 13, pp. 2241–2250, Oct. 2017.

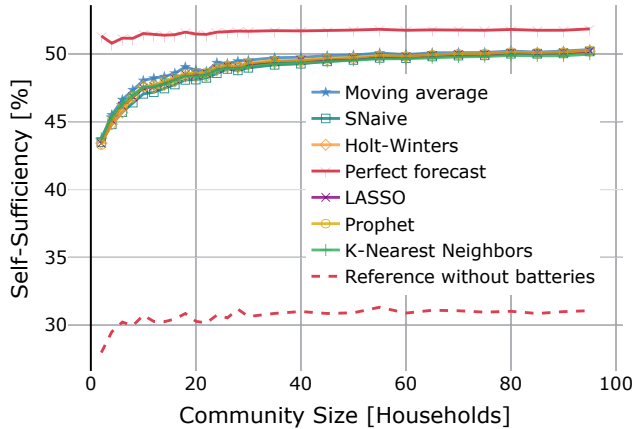


Fig. 6. Self-sufficiency as a function of community size for different forecasting methods.

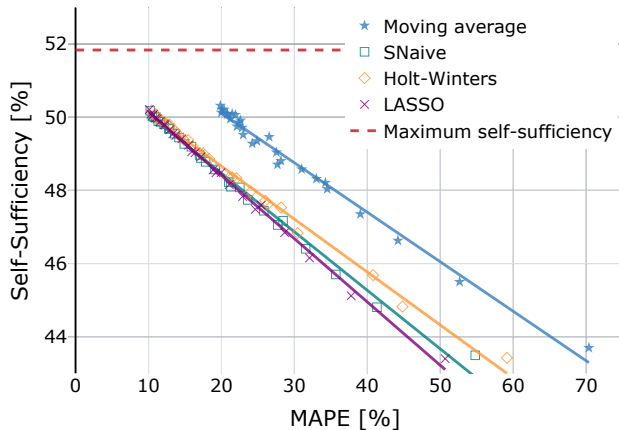


Fig. 7. Sufficiency as a function of MAPE for different forecasting methods.

Aside from gaining confidence on which forecasting method to select, value metrics are also relevant on at least two other counts: (a) selecting suitable quality metrics — (b) estimating progress margins for the community. In this study, MASE does not appear to be a suitable quality metric. Although, it is recommended in [15], here MASE is not representative of self-sufficiency, whereas, MAPE is strongly correlated with self-sufficiency (0.97 pearson coefficient). Finally, estimating the progress margin for the community is a good indicator on the amount of effort to spend improving forecasting methods. In our case, the remaining 10 % of MAPE corresponds to 2 % of self-sufficiency. From there, we can choose to work on reducing the MAPE by combining forecasting methods and adding exogenous variables, or alternatively, judge that the achieved self-sufficiency is high enough.



- [12] Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, and J. C. Schaake, eds., *Handbook of Hydrometeorological Ensemble Forecasting*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019.
- [13] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, July 2017.
- [14] F. Martínez-Álvarez, A. Troncoso, G. Asencio-Cortés, and J. C. Riquelme, "A Survey on Data Mining Techniques Applied to Electricity-Related Time Series Forecasting," *Energies*, vol. 8, pp. 13162–13193, Nov. 2015.
- [15] R. J. Hyndman, "Another Look at Forecast accuracy metrics for intermittent demand.," *Foresight*, p. 4, 2006.
- [16] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *International Journal of Forecasting*, vol. 36, pp. 54–74, Jan. 2020.
- [17] A. Botchkarev, "A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 14, pp. 045–076, 2019.
- [18] N. Kourntzes, J. R. Trapero, and I. Svetunkov, "Measuring the behaviour of experts on demand forecasting: a complex task," *Working paper*, p. 23, 2014.
- [19] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, pp. 914–938, July 2016.
- [20] F. A. Rahimi and A. Ipakchi, "Transactive Energy Techniques: Closing the Gap between Wholesale and Retail Markets," *The Electricity Journal*, vol. 25, pp. 29–35, Oct. 2012.
- [21] T. Hobfeld, L. Skorin-Kapov, P. E. Heegaard, and M. Varela, "Definition of QoE Fairness in Shared Systems," *IEEE Communications Letters*, vol. 21, pp. 184–187, Jan. 2017.
- [22] A. J. Aristizábal, D. H. Ospina, M. Castañeda, S. Zapata, and E. Banguero, "Fuzzy logic energy management for a microgrid with storage battery," *International Journal of Ambient Energy*, vol. 0, pp. 1–9, Aug. 2018.
- [23] C. D. Korkas, S. Baldi, and E. B. Kosmatopoulos, "9 - Grid-Connected Microgrids: Demand Management via Distributed Control and Human-in-the-Loop Optimization," in *Advances in Renewable Energies and Power Technologies* (I. Yahyaoui, ed.), pp. 315–344, Elsevier, Jan. 2018. 00000.
- [24] P. Chakraborty, E. Baeyens, P. P. Khargonekar, and K. Poola, "A cooperative game for the realized profit of an aggregation of renewable energy producers," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 5805–5812, Dec. 2016. 00018.
- [25] M. Tavakoli, F. Shokridehaki, M. Marzband, R. Godina, and E. Poursmaeil, "A two stage hierarchical control approach for the optimal energy management in commercial building microgrids based on local wind power and PEVs," *Sustainable Cities and Society*, vol. 41, pp. 332–340, Aug. 2018. 00070.
- [26] M. Tavakoli, F. Shokridehaki, M. Funsho Akorede, M. Marzband, I. Vechiu, and E. Poursmaeil, "CVaR-based energy management scheme for optimal resilience and operational cost in commercial building microgrids," *International Journal of Electrical Power & Energy Systems*, vol. 100, pp. 1–9, Sept. 2018. 00091.
- [27] R. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice, 2nd edition*. OTexts: Melbourne, Australia, 2018.
- [28] S. J. Taylor and B. Letham, "Forecasting at scale," preprint, PeerJ Preprints, Sept. 2017.
- [29] F. Ziel and B. Liu, "Lasso estimation for GEFCom2014 probabilistic electric load forecasting," *International Journal of Forecasting*, vol. 32, pp. 1029–1037, July 2016.
- [30] G.-F. Fan, Y.-H. Guo, J.-M. Zheng, and W.-C. Hong, "Application of the Weighted K-Nearest Neighbor Algorithm for Short-Term Load Forecasting," *Energies*, vol. 12, p. 916, Jan. 2019.
- [31] P. Goncalves Da Silva, D. Ilić, and S. Karnouskos, "The Impact of Smart Grid Prosumer Grouping on Forecasting Accuracy and Its Benefits for Local Electricity Market Trading," *IEEE Transactions on Smart Grid*, vol. 5, pp. 402–410, Jan. 2014.