

# Control strategy for battery-supported photovoltaic systems aimed at peak load reduction

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø

Norwegian University of Science and Technology, Department of Energy and Process Engineering  
Kolbjørn Hejes v 1B, NO-7491 Trondheim, Norway

**Abstract.** The use of photovoltaic (PV) technologies is one of the key means for achieving the balance between operational power demand and generation in net Zero Energy Buildings (nZEBs). However, direct use of PV power on-site is limited due to wide variability and uncertainty of PV output, the temporal mismatch between PV generation and load and other factors. Consequently, in addition to low self-consumption rates, the problem of peak grid load and peak PV feed into the grid persists. Batteries that are coupled to PV units may partially offer the solution to these problems, if operated under an intelligent control strategy. In this paper we proposed a forecast-based control strategy for battery-to-grid interaction aimed at enhancing self-consumption and at reducing peak load. Python programming environment was used for data processing and algorithm development. Exemplification was made based on the reported hourly energy demand in one office building of 3000 m<sup>2</sup> heated floor area located in Trondheim, Norway. Forecasting of electricity load profiles was based on the seasonal autoregressive integral moving average (SARIMA) model. For PV power forecasting, the algorithm communicated with external service – Solcast API. The search method for optimal scheduling of operational time and the extent of charging/discharging was proposed. The results showed that as opposed to conventional battery use, this control strategy allowed to achieve significantly more consistent grid interaction. It offered highly accurate battery scheduling on a day-ahead basis while utilising minimum historical data and computational resources. The algorithm may be beneficial for end-users and grid operators, and thus, it has a high potential for future integration into building energy supply systems.

## 1 Introduction

Meeting strategic energy and environmental targets for nations and communities requires significant performance improvements in building sector. Net Zero Energy Buildings (nZEB) would play a crucial role in achieving these targets through low operational energy use and production of renewable energy at a quantity that meets building's delivered energy over the service lifetime [1]. Photovoltaic (PV) energy is considered as the key source that enables nZEB under various climate conditions and related political circumstances. Although grid-connected nZEBs do not require a match between instantaneous PV generation and instantaneous load, it is often preferred to increase the share of PV energy directly consumed on-site (self-consumption rate). One of the reasons for that is profitability as a result of the difference between self-produced and retail electricity prices, grid feed-in tariffs and numerous incentives [2]. Another reason is the need for more consistent grid interaction, i.e. the solutions for frequency regulations and future power grid reinforcement to handle high PV feed-in and high loads are required otherwise [3]. Maximising the building self-consumption while avoiding peak grid loads, therefore, is one of the key challenges that is focused by the research initiatives in the discipline. Two distinct but not mutually

exclusive approaches can be used to tackle these challenges: demand side management (DSM) and the use of battery energy storage systems (BESS). The former is concerned with shifting the deferrable loads to the time when PV system has high output. It often includes rescheduling of heating, ventilation, air conditioning and some other types of household equipment. The other approach enables energy flexibility by using PV energy stored in batteries at the time when load occurs. BESS is recognized as having a larger potential to increase self-consumption rates [4]. Currently, best practices in the field of energy flexibility involve a combination DSM and BESS [2, 5].

For the effective use of both these approaches, a reasonably accurate forecast of energy demand and PV generation in the (nearest) future is required [2]. Instantaneous load and on-site PV power, however, are the sources of variabilities and uncertainties. In addition to thermal properties of the building envelope, climate, energy supply systems, building purpose, occupant behaviour and maintenance practices determine the unique shape of load profile that changes on day-to-day basis. Solar power profile is a function of daily and seasonal variations of both direct and diffuse radiation, outdoor temperature and other factors. For a short-term load forecasting in real-time, building energy

\* Corresponding author: [ruslan.zhuravchak@ntnu.no](mailto:ruslan.zhuravchak@ntnu.no)

performance simulation techniques are rarely used. Detailed simulation is constrained by computational capabilities and the need for exhaustive building-related information. Data-driven approaches, on the contrary, may resolve some of these limitations [6]. Their use for electric energy demand prediction has been mainly done by regression analysis, machine learning techniques, and time series forecasting. With regression analysis, one aims at identifying a set of influential variables and the underlying (often non-linear) model structure. Machine learning techniques offer the benefits of automated pattern discovery instead. Amongst the limitations of both these approaches is the need for several explanatory variables. With time series modelling, however, the prediction of target variables can be made based only on the past observations of itself, i.e. applicable for univariate problems. Therefore, such adaptive, parsimonious time series forecasting is expected to play an important role in facilitating energy flexibility and more consistent grid interaction for nZEBs.

Likewise, real-time PV power forecasting under uncertain weather conditions is one of the fundamental challenges to enable energy flexibility. Practical solutions are offered by persistence models, physical and statistical approaches [7] that utilise sky models or data streams from either satellite imagery or reference PV systems. Recent advancements in the field succeeded at combining these into hybrid approaches which allows to achieve high forecast accuracy [8].

Given that two components, the forecasted PV output and load profiles are available, battery charge/discharge operations can be scheduled with one or more objectives, most often these are maximising self-consumption or battery service lifetime, minimising energy cost,  $CO_2$

intensity or peak grid loads. A comprehensive overview of operational strategies for PV-coupled battery systems, their objectives and methods are available in studies [4, 9, 10].

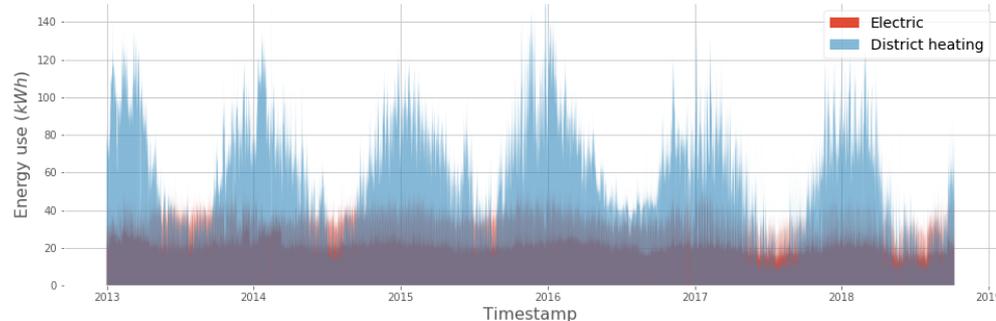
In this study, some of the best practices related to forecast-based control were implemented and elaborated further. Analysis and modelling tasks were carried out with Python programming language using NumPy, SciPy and Sk-learn libraries for numerical computing, Statsmodels for model training/testing, Pandas for data wrangling, Matplotlib and Seaborn for data visualisation.

## 2 Methodology

### 2.1 Case study

A case study was one non-residential building located in Trondheim, Norway. It is used for health-related and recreational purposes and has the total heated floor area of more than 3000  $m^2$ . Both electric and district heating energy use for the building were reported on hourly basis and accessible as datetime-indexed arrays through an energy monitoring platform.

Historical data over the entire building's monitoring period of more than five years is illustrated in Fig. 1. Because district heating was utilised in cold periods, electric energy use exhibits weak seasonality through the year. Also, no monthly/weekly seasonality was observed. Since such seasonality in electric energy use is negligible here, further data analysis and algorithm exemplifications are made using the most recent observation (Fig. 2) on the span of 90 days (Jul. 10<sup>th</sup>, 2018 through Oct. 8<sup>th</sup>, 2018, accounting for 2160 observations total).



**Fig. 1.** Historical data (hourly measurements) on electric and district heating energy use for entire observation period.



**Fig. 2.** Electric energy use (hourly measurements) in most recent observations.

It can also be observed from Fig. 2 that the series has a strong daily pattern. The lowest electric energy use for

the building occurs at around 6:00. It increases in linear or exponential manner through the day, achieves its peak

between 15:00 and 23:00 followed by decrease through the night.

The building is assumed to have a PV system of 150 kW total installed capacity in place with south-oriented, 30° tilted modules.

## 2.2 Energy use forecasting

For univariate time series forecasting, the Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model is one of the most commonly used [11, 12]. For applications with seasonal effects, ARIMA( $p, d, q$ ) has been modified to the multiplicative seasonal SARIMA( $p, d, q$ )  $\times$  ( $P, D, Q$ )<sub>s</sub>. Studies [13, 14] exemplify its application for energy forecasting purposes. Model's parameters are following:

$p, P$  – non-seasonal and seasonal autoregressive (AR) components;

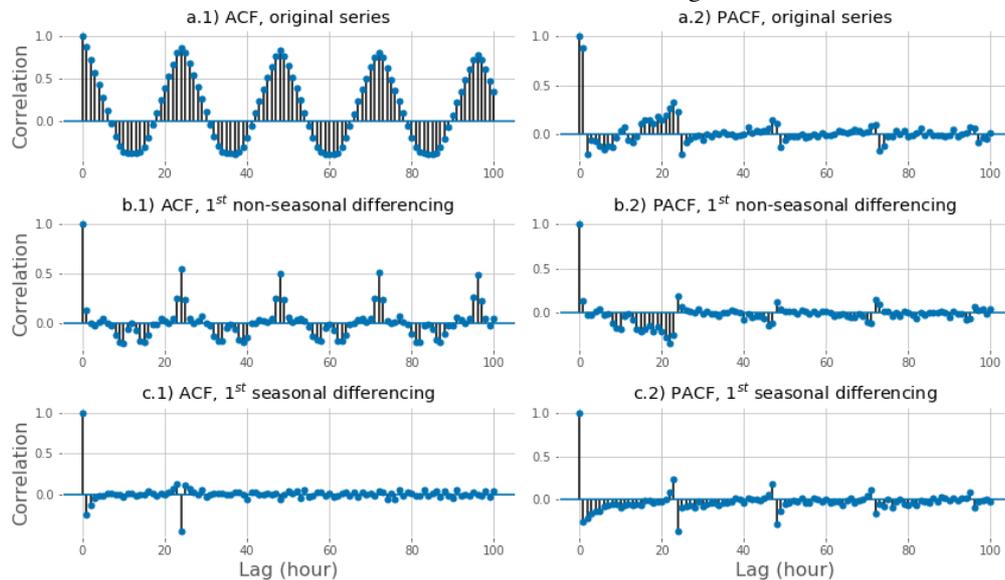
$d, D$  – non-seasonal and seasonal order of differencing;

$q, Q$  – non-seasonal and seasonal moving average (MA) components;

$s$  – periodicity of the season.

Detailed formulation for both ARIMA and SARIMA is provided in literature [15]. In this study, an open-source statistical computing package Statsmodels [16] has been used for SARIMA training/testing. Package documentation contains the details on SARIMA implementation in Python.

Model's application requires the series to have constant mean, variance and autocorrelation, i.e. stationary process. This is not the case for original series, as can be seen in Fig. 2. A sequence of steps for process stationarization here involves seasonal and non-seasonal differencing. The corresponding autocorrelation and partial autocorrelation functions (ACF and PACF) are illustrated in Fig. 3.



**Fig. 3.** ACF and PACF correlograms. a) Original data; b) Non-seasonal differencing; c) Seasonal differencing.

To induce stationarity, the first non-seasonal differencing was taken on the original series:

$$y'_t = y_t - y_{t-1} = (1 - B)y_t \quad (1)$$

where:

$y$  – original series;

$B$  – backshift operator, such that:

$$B^k y_t = y_{t-k} \quad (2)$$

The ACF of the resulting series shown in Fig. 3b.1 still indicates the presence of statistically significant lags. Therefore, an additional seasonal differencing step was applied:

$$y''_t = (1 - B^{24})y'_t \quad (3)$$

Referring to the original series, the additional seasonal differencing was obtained as:

$$y''_t = (1 - B - B^{24} + B^{25})y_t =$$

$$y_t - y_{t-1} - y_{t-24} + y_{t-25} \quad (4)$$

The newly acquired time series had the mean value close to zero (0.0028) and a reasonably small number of significant autocorrelations as shown in Fig. 3c.1, at lag 1, 2, near- or at the first seasonal lag that did not repeat afterwards. This series was considered stationary, which

is also confirmed by the Dickey-Fuller [17] test. The partial autocorrelation function (PACF) plot shown in Fig. 3c.2 exhibits slow decrease at the non-seasonal level and a spike at the seasonal lag. Although further differencing may result in even more consistent series, this step could induce the negative effect of overcomplication and overfitting the model. Moreover, as reported by others [18], these seasonal lags could merely reflect large correlation of non-seasonal ones.

The PACF correlogram in Fig 3c.2 indicates that the first 4 lags are significantly different from zero, thus suggesting the feasible non-seasonal component to be within the range  $AR \in [0,4]$ . Similarly, from the ACF plot in (Fig. 3c.1) non-seasonal  $MA \in [0,4]$ . Because PACF and ACF in final series indicate significant near-seasonal lag, the seasonal AR and MA chosen here are 1 and 1.

Since non-seasonal components AR ( $p$ ) and MA ( $q$ ) are essential for the model performance, their identification in this study is facilitated by grid search method. Grid search over the parameters' space implies training and evaluating the models with all feasible

combinations of parameters  $p, q$ . The root mean squared error (RMSE) can serve as a reference metrics for the model performance evaluation used here, defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

where:

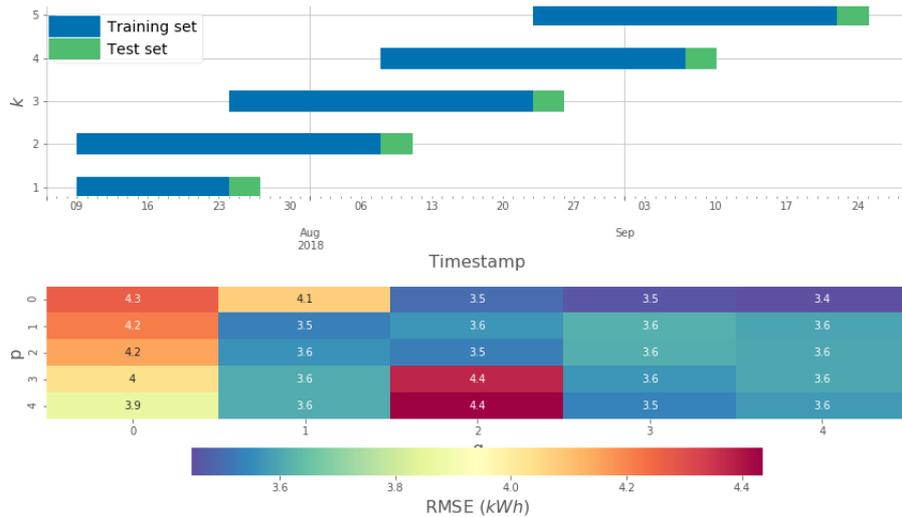
$n$  – number of samples;

$y_i$  – actual observation;

$\hat{y}_i$  – model prediction.

The preferred model parameters  $(p, q)$ , therefore, correspond to the lowest RMSE obtained on the test set.

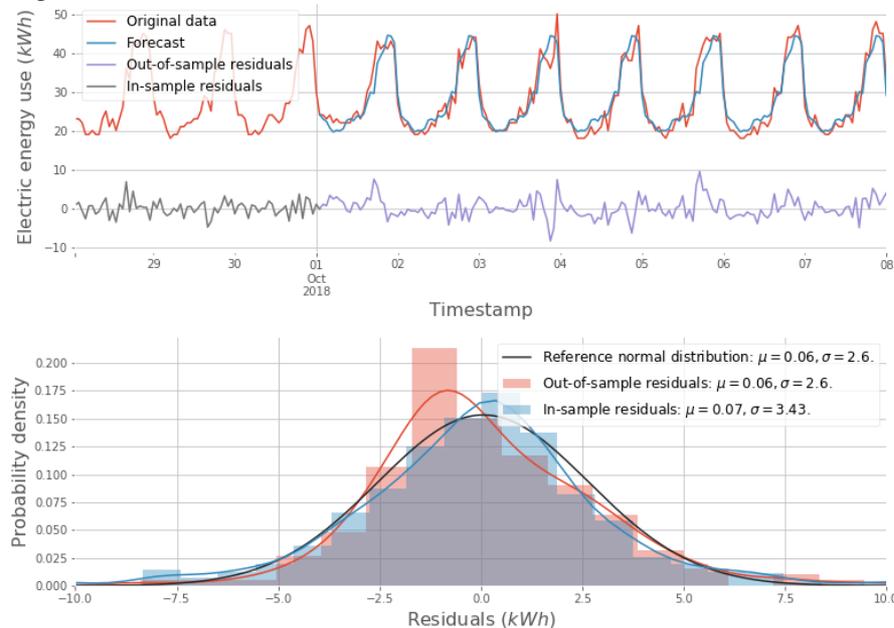
As an additional measure used to ensure that the selection of  $p$  and  $q$  is not affected by coincidental patterns in the series, a  $k$ -fold cross-validation (CV) procedure is applied. CV for time series models implies splitting the dataset into  $k$  subsets where a training set is immediately followed by a test set. Thus, 90 most recent days of monitored data are split into  $k = 5$  subsets as illustrated in Fig. 4. Grid search is conducted for each CV step. Further, the parameters  $(p, q)$  are selected based on average RMSE for all CV steps, as illustrated in Fig. 4 with color-coded heatmap.



**Fig. 4.** Parameters tuning and validation. CV intervals split (top) and grid search results (bottom).

For a given case, the most favourable model parameters  $p$  and  $q$  are 0 and 4 accordingly, since the average over all CV steps RMSE is the least for this set of

parameters. The performance of  $SARIMA(0,1,4) \times (1,1,1)_{24}$  is illustrated in Fig. 5.



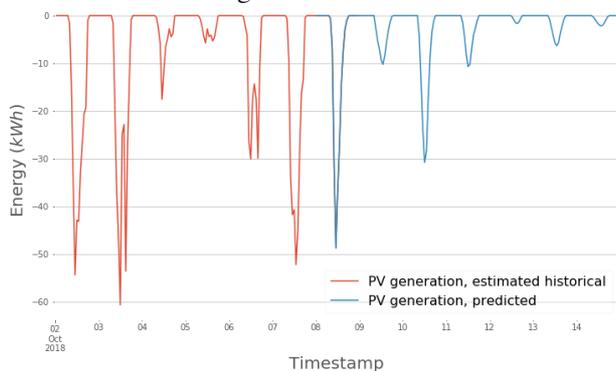
**Fig. 5.** Model performance. Top: original, forecasted and residuals series. Bottom: distribution of residuals.

Forecasted series overlays closely with the test series. Both, in- and out-of-sample residuals do not exceed 10 kWh. In-sample residuals follow normal distribution. Out-of-sample residuals density, however, is skewed to the left compared to the shape of equivalent normal distribution which indicates the loss of accuracy in long-

term forecasting. Although such systematic error can be neglected in some applications, we restrict this control strategy to day-ahead horizon, where the accuracy is the highest.

### 2.3 PV power forecasting

As was mentioned before, the proposed control strategy involved communication with external service Solcast API [19] to obtain high-quality PV output forecast. Data request included geospatial (latitude and longitude) and technical information about the PV system (installed capacity, azimuth and tilt angle). The API's response is a datetime-indexed array of the forecasted PV power with 30 minutes temporal resolution over a span of 7 days. Further data wrangling involves two steps. Since the response dataset is UCT-indexed, a conversion to local time is needed. Also, original data is converted from power to energy units and from 30 minutes to hourly intervals. An example of postprocessed API's response based on PV system's parameters mentioned in Section 2.1 is illustrated in Fig. 6.



**Fig. 6.** PV generation (reversed), historical and forecasted.

### 2.4 Load matching

Given that the energy use and PV generation profiles were known, the task of load matching was to allocate the aggregated daily PV energy to the time when energy use is the highest on that day. This implies finding threshold value ( $p$ ) that separates grid from battery energy use. The task could be formulated as follows:

$$\sum_1^{24} P_{PV} = \sum_{n=1}^n (\sum_{T_1}^{T_2} (P - p)) \pm \varepsilon \quad (6)$$

for all  $(P - p) > 0$  and if  $\sum_1^{24} P_{PV} > 0$ .

Here:

$P_{PV}$  – hourly PV energy,  $kWh$ ;

$P$  – hourly energy use in the building,  $kWh$

$p \in [0, P_{max}]$  – threshold value for the day,  $kWh$ ;

$T_1, T_2 \in [0, 24]$  – the beginning and the end of interval  $n$  when battery should be used;

$n$  – number of time intervals, used to account for the possibility of battery use during more than one period through the same day;

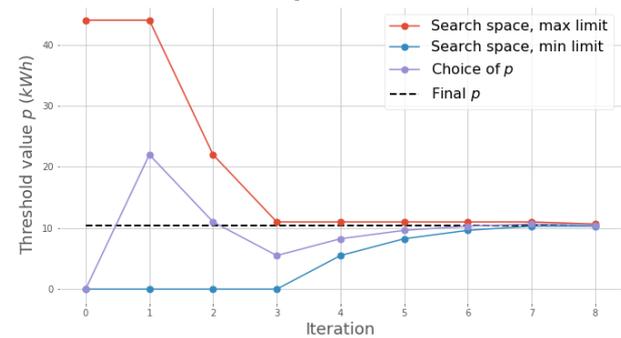
$\varepsilon$  – absolute error tolerance,  $kWh$ .

Solving this problem analytically would require approximating the load and PV curves and their subsequent integration. To avoid the associated loss of

accuracy, a numerical search technique has been implemented instead. Namely a dichotomous search algorithm was used to find the threshold value  $p$ . Fig. 7 illustrates the search for  $p$  on Oct. 9<sup>th</sup>, 2018. By gradually eliminating half of the continuous search space that was known at previous iteration, the condition (eq. 6) is achieved in a recursive manner after 8 iterations with  $p = 11 kWh$ , which yields:

$$\sum_1^{24} P_{PV} = \sum_{T_1}^{T_2} (P - p) = 52.8 \pm 0.1 kWh.$$

Finding the time intervals  $T_1, T_2$  further, is based on the condition where  $P > P - p$ .

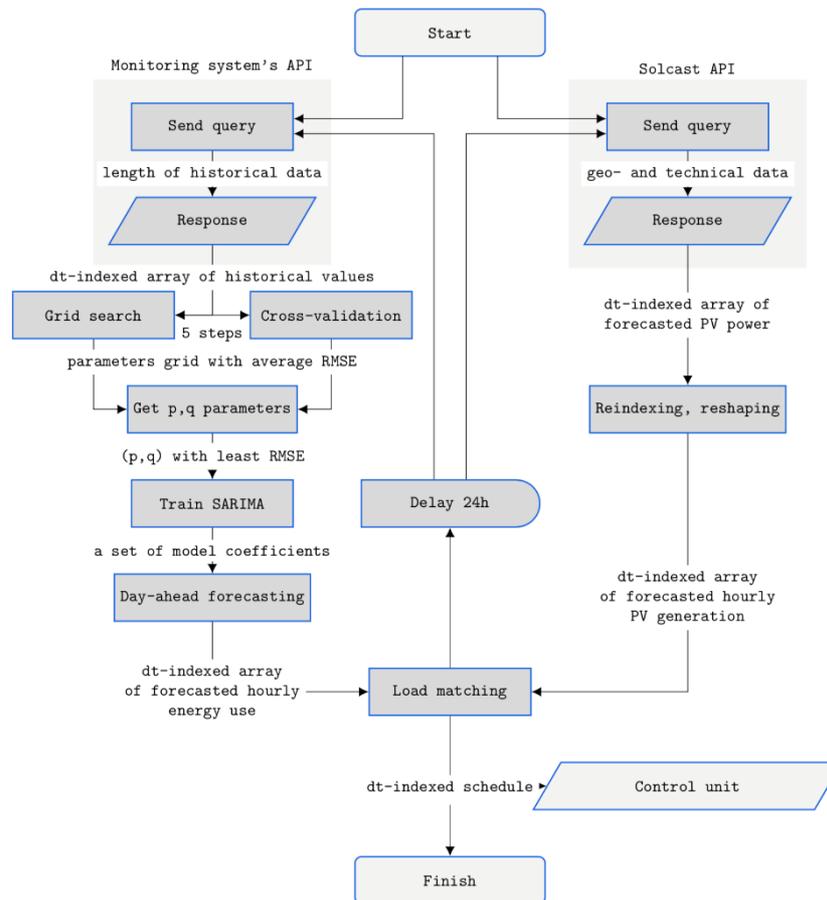


**Fig. 7.** Search progress.

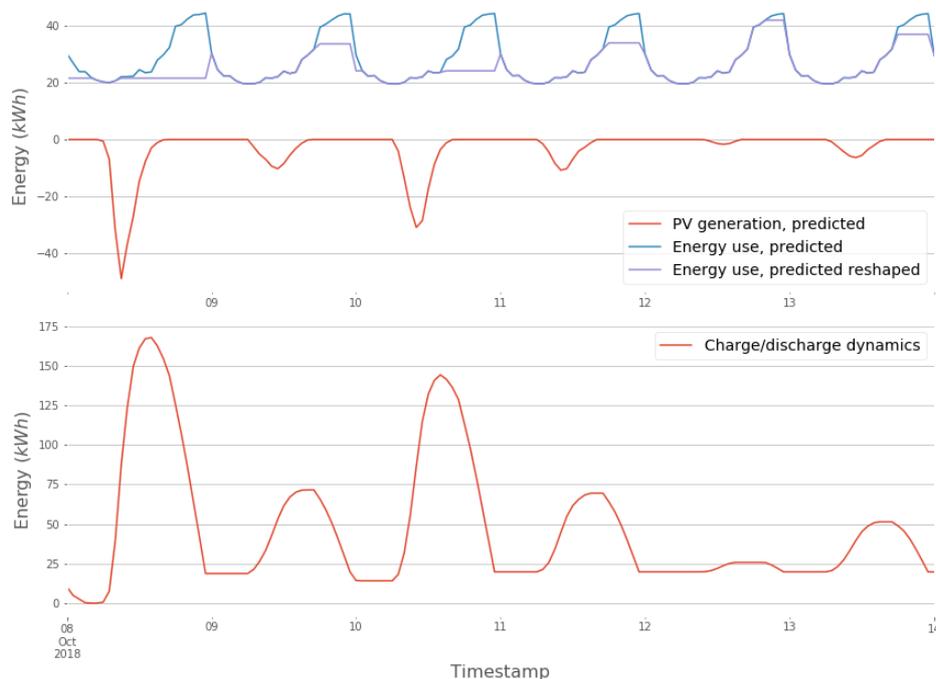
## 3 Results and discussion

Given the techniques discussed above, this study proposed the algorithm for day-ahead battery operation scheduling that follows a block diagram in Fig. 8.

The first step was to send two data requests: one for historical energy use data (of a desired length, e.g.  $N = 90$  days) and one for PV output forecast (installed capacity, latitude, longitude, the azimuth and the tilt of the PV modules). The response from Solcast API was then converted to local time, from power to estimated energy units and from 30 minutes to hourly time resolution. Since the day-ahead forecast was used, the series is cut at 24<sup>th</sup> element. The response from the energy monitoring system's API is a datetime-indexed array of length  $L = 24N$ . 5 grid search and cross-validation steps were taken after which the SARIMA model parameters ( $p, q$ ) were selected such that the corresponding average RMSE is the smallest. With these parameters, SARIMA( $p, 1, q$ )  $\times$  (1,1,1)<sub>24</sub> was trained using the entire  $L$ -length training set. The trained model was further used for predicting energy use for the following day. Both arrays, the forecasted energy use and PV output enter the load matching unit and follow the procedure described in Section 2.4. After the load matching task was completed, the delay of 24 hours is initiated. The output of load matching unit was the reshaped grid load profile and the recommended battery charge/discharge dynamics (as exemplified in Fig. 9) to achieve such peak shaving effect.



**Fig. 8.** Process flow diagram.



**Fig. 9.** Algorithm's output. Reshaped load profile (top); Charge-discharge dynamics (bottom).

Technical and climate conditions under the test case did not entail the need for controlling grid feed-in, because daily aggregated PV energy is always smaller than aggregated energy use. In case if daily PV generation was higher, a search algorithm would separate PV profile

into two parts, one for self-consumption and one for grid export:

$$\sum_1^{24} P = \sum_{n=1}^n (\sum_{T_1}^{T_2} (P_{PV} - p_{PV})) \pm \varepsilon \quad (7)$$

for all  $(P_{PV} - p_{PV}) > 0$  and if  $\sum_1^{24} P > 0$ . Here  $p_{PV} \in [0, P_{PV_{max}}]$  – threshold value specific for daily PV profile, kWh. Under such circumstances, PV energy in range

$[0, p_{PV}]$  was excessive and could be exported instantaneously.

Thus, an algorithm suggests a reshaped profile, where the peak energy use and/or PV feed-in are avoided. Significantly more consistent energy use profile can be achieved for the days with high PV output, e.g. Oct. 8<sup>th</sup> and Oct. 10<sup>th</sup>. Over the entire period between Oct. 8<sup>th</sup> and Oct. 14<sup>th</sup>, 461 kWh of available PV energy offer 5% to 51% reduction of daily peak energy use from the grid.

This approach for day-ahead scheduling of battery operational time and intensity remains relatively simple and requires little historical data. It may also be used to reshape load profiles in response to dynamic energy price or CO<sub>2</sub> intensity.

The selection of some parameters for SARIMA was justified and automated. However, a different case study might reveal the need to conduct grid search over a larger set of parameters, e.g. seasonal components ( $P, Q$ ) or to make more steps to achieve stationary series ( $d, D$ ). Additional strong seasonality in building operation, e.g. weekly/holiday-based, monthly or annually, may require added complexity of SARIMA, resampling the underlying data or considering another time series model. The needed length of historical data must be satisfied accordingly. Further automation of these procedures, therefore, is needed to achieve more robust battery control that would suit the buildings of various types and more complex load profiles. For on-site algorithm deployment, as opposed to forecasting carried out in data centres, optimization of computational time and resources usage may be necessary.

Forecasting and control principles used in this study exclude the adaptation to sudden, unforeseen changes that may occur on the short term. Also, no attention is given to the potential technical limitations associated with charge/discharge intensity, critical depletion and system's losses. The authors anticipate the best application of the algorithm in combination with model predictive control or similar.

Expanding the forecasting horizon and considering a more long-term control strategy may offer even more consistent load profile. This option, however, is limited by accuracy (as can be seen in Fig. 5) and data needs in forecasting techniques currently used. Potential improvements of energy use forecasting can be achieved through including exogenous variables in time series modelling, e.g. thermal properties of building envelope, HVAC equipment specifications/setpoints and detailed data of occupant behaviour. The improvement of forecasting accuracy is also expected through supplementing time series models with other techniques, as discussed elsewhere [6].

Longer-term control strategies are likely to require more battery storage capacity. Techno-economic feasibility studies of such measures would have to address the trade-off between self-consumption rate, smoothness of grid interaction, battery capacity and utilization rate. This should be in line with the measures to prolong battery service lifetime as discussed in study [20]. As outlined in another study [4], the acceptance of such control

strategies in buildings and communities highly depends on economic, technological and policy-related factors.

## 4 Conclusions

In order to meet energy efficiency and environmental targets without compromising grid stability, an increased self-consumption while smoothing grid interaction has to be promoted in nZEBs. This challenge requires not only capacity gains for PV systems and energy storage units, but more intelligent approaches for their utilization. The latter relies strongly on accurate energy and PV output forecasting.

With this article, an autonomous, parsimonious forecast-based control strategy was proposed. High performance of building energy use forecasting was achieved with time series model SARIMA supplemented by grid search for parameter identification and 5-fold cross validation. The proposed algorithm benefits from high quality satellite-derived PV power forecasting services. It is shown that for peak shaving purposes, the battery operation time and intensity can be scheduled on a day-ahead basis.

Such control strategies are likely to contribute to achieving the energy and environmental targets under the growing demand for efficient built environment. An increasing accessibility of PV systems and energy storage solutions are expected to positively influence their adoption.

The authors gratefully acknowledge the support from the Research Council of Norway through the research project Methods for Transparent Energy Planning of Urban Building Stocks - ExPOSE (project number 268248) under EnergiX program.

## References

1. Sartori, A. Napolitano, and K. Voss, Energy and Buildings, *Net zero energy buildings: A consistent definition framework*, **48**, 220-232, (2012).
2. R. Luthander, et al., Applied Energy, *Photovoltaic self-consumption in buildings: A review*, **142**, 80-94, (2015).
3. A. Zeh, and R. Witzmann, Energy Procedia, *Operational Strategies for Battery Storage Systems in Low-voltage Distribution Grids to Limit the Feed-in Power of Roof-mounted Solar Power Systems*, **46**, 114-123, (2014).
4. G. Angenendt, et al., Applied Energy, *Comparison of different operation strategies for PV battery home storage systems including forecast-based operation strategies*, **229**, 884-899, (2018).
5. M. Castillo-Cagigal, et al., Solar Energy, *PV self-consumption optimization with storage and Active DSM for the residential sector*, **85(9)**, 2338-2348, (2011).
6. C. Deb, et al., Renewable and Sustainable Energy Reviews, *A review on time series forecasting techniques for building energy consumption*, **74**, 902-924, (2017).

7. M.Q. Raza, M. Nadarajah, and C. Ekanayake, *Solar Energy, On recent advances in PV output power forecast*, **136**, 125-144, (2016).
8. J.M. Bright, et al., *Solar Energy, Improved satellite-derived PV power nowcasting using real-time power data from reference PV systems*, **168**, 118-139, (2018).
9. L. Schibuola, M. Scarpa, and C. Tambani, *Renewable Energy, Influence of charge control strategies on electricity import/export in battery-supported photovoltaic systems*, **113**, 312-328, (2017).
10. J. Moshövel, et al., *Applied Energy, Analysis of the maximal possible grid relief from PV-peak-power impacts by using storage systems for increased self-consumption*, **137**, 567-575, (2015).
11. D.W. van der Meer, J. Widén, and J. Munkhammar, *Renewable and Sustainable Energy Reviews, Review on probabilistic forecasting of photovoltaic power production and electricity consumption*, **81**, 1484-1512, (2018).
12. H. Verdejo, et al., *Renewable and Sustainable Energy Reviews, Statistic linear parametric techniques for residential electric energy demand forecasting. A review and an implementation to Chile* **74**, 512-521, (2017).
13. N. Elamin, and M. Fukushige, *Energy, Modeling and forecasting hourly electricity demand by SARIMAX with interactions*, **165**, 257-268, (2018).
14. A. Tarsitano, and I.L. Amerise, *Energy, Short-term load forecasting using a two-stage sarimax model*, **133**, 108-114, (2017).
15. G.E.P. Box, et al., *Time Series Analysis: Forecasting and Control*. (Wiley, 2015).
16. S. Seabold, and J. Perktold. *Statsmodels: Econometric and Statistical Modeling with Python*, (9th Python in Science Conference, USA, 2010).
17. R.H. Shumway, and D.S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, (Springer International Publishing, 2017).
18. R.E. Abdel-Aal, and A.Z. Al-Garni, *Energy, Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis*, **22(11)**, 1059-1069, (1997).
19. Solcast, *Solar irradiance data*, <http://solcast.com.au/>, (2018).
20. J. Li, and M.A. Danzer, *Journal of Power Sources, Optimal charge control strategies for stationary photovoltaic battery systems*, **258**, 365-373, (2014).