Feature Extraction for Day-ahead Electricity-Load Forecasting in Residential Buildings ⋆

Aleksey V. Kychkin * **Georgios C. Chasparis** **

* *Department of Information Technologies in Business, National Research University Higher School of Economics, Studencheskaya 38, 614070, Perm, Russia (e-mail: avkychkin@hse.ru)*
** *Software Competence Center Hagenberg GmbH, Softwarepark 21, 4232 Hagenberg, Austria (e-mail: georgios.chasparis@scch.at)*

Abstract: In the context of electricity demand response, an important task is to generate accurate forecasts of energy loads for groups of households as well as individual consumers. We consider the problem of short-term (one-day-ahead) forecasting of the electricity consumption load of a residential building. In order to generate such forecasts, historical energy consumption data are used, presented in the form of a time series with a fixed time step. In this paper, we first review existing (one-day-ahead) forecasting methodologies including: a) naive persistence models, b) autoregressive-based models (e.g., AR and SARIMA), c) triple exponential smoothing (Holt-Winters) model, and d) combinations of naive persistence and auto-regressive-based models (PAR). We then introduce a novel forecasting methodology, namely seasonal persistence-based regressive model (SPR) that optimally selects between lower- and higher-frequency persistence and temporal dependencies that are specific to the residential electricity load profiles. Given that the proposed forecasting method equivalently translates into a regression optimization problem, recursive-least-squares is utilized to train the model in a computationally efficient manner. Finally, we demonstrate through simulations the forecasting accuracy of this method in comparison with the standard forecasting techniques (a)-(d).

Keywords: demand response, electricity consumption, short term load forecasting, persistence models, autoregressive models, Holt-Winters model.

1. INTRODUCTION

Recently, electricity markets' operators and policy makers have been looking on alternative ways for motivating prosumers to participate in demand-response mechanisms. Increasing the percentage of renewables' integration could significantly reduce the electricity price and costs of production. For example, marketing energy *flexibility* over the next day (e.g., in the day-ahead or intra-day spot electricity markets) is one indirect form of such demand-response services that can increase the participation of residential prosumers in exchange of a reduced tariff, Xu et al. (2016); Chasparis et al. (2019). However, the performance of such day-ahead optimization relies heavily upon the accuracy of several forecast quantities, including forecasts of the load profile, price, and Photovoltaic (PV) generation.

In this paper, we are addressing the problem of electricity load forecasting. As a general rule, the forecast of the electricity load consumption will be based upon prior available measurements recorded at regular time intervals.The type and duration of the forecasts may vary Hong and Fan (2016), ranging from one hour ahead up to one day ahead, usually referred to as *Short Term Load Forecasting* (STLF) methods Alfares and Mohammad (2002).

⋆ This work has been supported by the Austrian Research Agency FFG through the research project Flex+ (FFG # 864996).

In addition, aggregate forecasts over multiple households can usually be constructed with high accuracy. However, in case of individual level STLF, accuracy may considerably decrease, Haben et al. (2019). Such degradation may be attributed to rapid changes in residents'/users' behavior. In addition, forecasts are often requested in a fine granularity (e.g., of 15-min intervals) which creates additional challenges in maintaining high prediction accuracy.

Given the involved challenges in establishing accurate electricity-load forecasts over a short-term time horizon (of at least one-day ahead), this paper first investigates standard (black-box) forecasting methodologies as well as naive (persistence) models. Our intention is to evaluate the accuracy level that such well-known forecasting models can generate in a short-term time horizon. Furthermore, we propose a family of regression models that are specifically tailored for electricity-load forecasting in residential buildings and perform a comparative analysis with standard forecasting models. Similarly to expert-based learning Cesa-Bianchi and Lugosi (2006), our goal is to integrate persistence-based features as well auto-regressive features in a single model, and optimally determine their relative influence in generating forecasts. We demonstrate, using real-world data, that the proposed models significantly outperform standard black-box models and naive persistence models.

The remainder of this paper is organized as follows. In Section 2, we discuss related work and the main contributions of this paper. Section 3 presents standard and black-box models for electricity load forecasting, such as persistence, Holt-Winters and SARIMA models. In Section 4, we present a new class of load forecasting regression models which are specifically tailored to the problem of load forecasting, namely the *Persistence-based Auto-Regressive* (PAR) model and the *Seasonal Persistence-based Regressive* (SPR) model. In both cases, the model tries to compute the relative importance between lower- and higher-frequency persistence and temporal features. Section 5 evaluates the performance of PAR and SPR models and compares them with the standard forecasting models of Section 3 using real-world data from two residential buildings.

Notation and abbreviations

d	day index
t	time index
$y_d(t)$	electricity load at time interval t of day d
$\hat{y}_d(t)$	forecast of electricity load at time interval t of day d
AR	Auto-regressive model
CLD	"Copy-last-days" persistence model
HW	Holt-Winters model
N-day	persistence model of N previous days
PAR	Persistence-based auto-regressive model
PM	Persistence model
PV	Photovoltaic
SPR	Seasonal persistence-based regressive model
STLF	Short Term Load Forecasting
RMSE	Root-mean squared error

2. RELATED WORK AND CONTRIBUTION

2.1 Related work

There are many popular forecasting methods in the literature which can mainly be classified into three groups:

- black-box standard models;
- models specifically tailored for STLF;
- optimally combining models for STLF.

The first group can be represented by standard averaging techniques and auto-regressive models. For example, references Haben et al. (2014); Kychkin (2016) discuss simple models that calculate average day ahead values for energy consumption based on previous days. The auto-regressive-based models (AR/ARMA/ARIMA/SARIMA) discussed in Haben et al. (2019); Clements et al. (2016) and the exponential smoothing approach, like the Holt-Winters method Alfares and Mohammad (2002), has been extensively used for electricity load forecasting as well as other applications Akpinar and Yumusak (2016). Furthermore, the SARIMA and Holt-Winters method are more efficient if the data has seasonality, like day or week seasons in the electricity consumption. Such methods can also be used for long term load forecasting.

The second group includes STLF models specifically tailored for load forecasting. In particular, Cancelo et al. (2008) shows an STLF realization of the Spanish system operator and describes the influence of weather data in the load predictions. Regression analysis, as a tool for estimating the relative importance of the features, has also been used for STLF in the literature, e.g., Hong et al. (2010). A set of relevant features can be extracted from the time series of energy data that is described in references Christ et al. (2017) and Christ et al. (2018). These papers also demonstrate the time series data attribute selection problem and describe a feature extraction algorithm. Reference Kychkin and Mikriukov (2016) proposes a method for analysing the building multi-sectional lighting and climate control/conditioning energy consumption data with linear regression models. Alternatively, artificial neural networks are also quite popular in STLF applications, including the references Haben et al. (2019); Chitsaz et al. (2015); Hippert et al. (2005). Moreover, fuzzy logic and knowledge-based models can be also used as in Alfares and Mohammad (2002).

The third group of forecasting methods optimally combines several other reference/basis models. For example, Chen et al. (2004) introduces a Wavelet-ARMAX-Winters methodology that incorporates three modeling strategies: ARMAX models, trigonometric regressions sensitive to seasonality effects and Holt-Winters model. In reference Ye and Dai (2018) authors propose a hybrid time series forecasting algorithm based on transfer learning, namely Online Sequential Extreme Learning Machine that combines Kernels and ensemble learning. In case of such combinations it is important to adaptively update relative importance (or weights) of the reference/basis models. Reference Soares and Medeiros (2008) describes the two-level Seasonal Autoregressive model (TLSAR), that combines calculations for potential and irregular load, and Dummy-Adjusted SARIMA, which modifies the standard SARIMA model by using dummy variables of the day type. These methods demonstrate better forecasting performance than neural-networks-based models but they were tested only for aggregate energy data. Finally, reference Hippert et al. (2005) describes combinations of naive, smoothing and regression models in comparison with Large neural networks, which show better accuracy on the test data set.

In summary, methods from the first group (i.e., standard averaging techniques and auto-regressive models) try to smooth data and as a result cannot predict load with irregular peaks. On the other hand, the second group (i.e., models specifically tailored for load forecasting) helps to identify hidden patterns in the day consumption and can detect load peaks. Some of these models can also identify nonlinear dependencies in the energy data (as in neural-network models), but they usually require larger training times and data sets in comparison with linear regression approaches. Finally, the models of the third group (i.e., combinations of other reference/basis models) appear to have the largest potential by combining basis models and adapting their relative importance (or weights). However, such adaptation usually increases computational complexity.

2.2 Contribution

One of the greatest challenges of STLF design is the uncertainty involved even within the relatively short horizon of one day ahead, primarily due to the stochastic patterns with which residents use electrical equipment. However random the behavior of the residents might be, there are certain "persistence" factors that tend to prevail. In this work, we intend to capture such multi-faceted persistence factors and identify their relevant importance. To this end, we try to identify factors/features that are quite specific to the electricity load in residential buildings, in order to minimize the uncertainty involved. We show that such load-specific design can have significant advantages in comparison with standard black-box models, such as auto-regressive-based, SARIMA and Holt-Winters models.

In parallel, we would like to invest on methodologies that are computationally efficient, in order to avoid the assumptions of large data sets and/or large training times. Note that such assumptions are common for seasonal-based models, such as SARIMA and Holt-Winters, as well as for artificial neural-network models. Instead, we focus our design approach on regression-based architectures where training can be performed on small data sets and even recursively, e.g., using recursive least squares. This increases considerably the efficiency of the implementation, given that the forecast granularity may be rather fine (every 15min or lower).

Furthermore, the use of regression-based models serves as an optimization criterion for optimal combining forecasts from different models, in the spirit of expert-based learning of Cesa-Bianchi and Lugosi (2006). Such framework reduces considerably the uncertainty involved with respect to the inability of some forecasting models to continuously provide accurate forecasts (e.g., due to changes in the operating conditions). In addition, it provides the relative significance of the basis models or features which is also valuable for the interpretability of the derived models.

3. STANDARD FORECASTING MODELS

3.1 Framework

Let us assume that we have measured the electricity load consumption of a household over the duration of $d-1>0$ days, and measurements have been collected with a period of 15min. As a result, in each such day, we have available 96 sequential load measurements, and each measurement at time instance t represents the total electricity load consumption during the last 15min interval. Using the available measurements over all previous $d-1$ days, we would like to predict the load consumption over the next day d (*day-ahead forecast*) and with the fine granularity of 15min intervals.

In the following subsections, we present several classes of standard or black-box forecasting models that have been used for the formulation of such day-ahead forecasts.

3.2 Persistence models

Persistence forecasting models are usually utilized to establish reference (baseline) models which can then be used for comparison tests. It is beneficial in many cases to know whether a developed forecasting model can provide better predictions than a baseline model. Persistence models are among the most trivial ones and they are based on the principle that "*things stay the same*", i.e., the forecast is always equal to the last known data point.

According to Notton and Voyant (2018), a persistence model would assume that the electricity load at time $t+1$ is equal to the load at time t. However, *how exactly time instances* $t+1$ *and* t *should be defined?* Given that we are interested in a day-ahead forecast and with 15min granularity, a persistence model that assumes that the load remains constant over the next day and equal to the current one would most likely fail. Instead, a persistence model would be more accurate if it assumes that the electricity load at time t of day d (briefly (t,d)) would be the same with the corresponding load at the same time t on the previous day $d-1$ or on the previous same day $d-7$. An additional variation of such model would also consider more than one previous days (e.g., the average consumption at the same time on N previous days).

More formally, let $y_d(t)$ denote the electricity load of a household at time instance t on day d. Then, the 1-day persistence model assumes that

$$\hat{y}_d^{\mathrm{PM}}(t) = y_{d-1}(t).$$

Analogously, we can define the ***N-day persistence model*** (or briefly N-day) as follows

$$\hat{y}_d^{\mathrm{PM}}(t) = \frac{1}{N}\sum_{i=d-N}^{d-1} y_i(t). \tag{1}$$

In other words, the N-day persistence model takes an average of the load of N previous days and at exactly the same time.

Since electricity load is highly correlated with the residents' presence in a household (i.e., with the residents' schedule), we can further improve the above N-day persistence model by only considering the N previous same days. Informally, if d corresponds to a "Monday", then to establish our forecast for time t, we need to create the average load at the same time on the most recent N previous Mondays. We will refer to this model as ***copy-last-days persistence model*** (or briefly CLD), according to which the forecasts are computed as follows:

$$\hat{y}_d^{\mathrm{PM}}(t) = \frac{1}{N}\sum_{i=d-7N}^{d-7} y_i(t). \tag{2}$$

3.3 Auto-regressive model

As we have seen in the previous subsection, the persistence models discussed try to capture temporal dependencies whose frequency extends over multiple days or weeks (due to, e.g., similarities in the schedule of the residents on similar days). We will refer to such dependencies as low-frequency temporal dependencies.

However, there might also be temporal dependencies of the electricity load within the same day. Informally, it is highly likely that the load measured at time interval t depends on the load measured at the previous time interval $t-1$ of the same day. We will refer to such temporal dependencies of the non-flexible load as high-frequency temporal dependencies.

Auto-regressive forecasting models can be used to capture such (high-frequency) temporal dependencies of the load within the same calendar day. Possibly the simplest such model is the Auto-Regressive model (briefly AR model), according to which the prediction of the load at time t is given by a linear combination of the load at previous time instances. In particular, we have

$$\hat{y}_d^{\mathrm{AR}}(t) = a_1 \cdot y_d(t-1) + \cdots + a_n \cdot y_d(t-n), \qquad (3)$$

which results from a maximum-a-posteriori predictor of an original white-noise perturbed process, cf., (Ljung, 1999, Chapter 4).

In case we would like to create predictions over several time instances ahead (e.g., one day ahead), then we can implement a variation of the above model, usually referred to as pseudo-regression model, which takes on the following form:

$$\hat{y}_d^{\mathrm{AR}}(t) = a_1 \cdot \hat{y}_d^{\mathrm{AR}}(t-1) + \cdots + a_n \cdot \hat{y}_d^{\mathrm{AR}}(t-n) \qquad (4)$$

In other words, if the non-flexible load at time $t-j$, $j = 1,,n$, is not known, it is replaced by the available prediction at that time instance.

Note that alternative auto-regressive-based models can be defined. For example, models that also incorporate moving-average (MA) noise terms are commonly used, which try to capture the effects of low-frequency perturbation terms in the profile.

One of the main drawbacks of such methodologies is the fact that are more appropriate for predictions in the range of only a few hours ahead. In fact, it is straightforward to see that even small prediction errors in one-step ahead predictions can propagate in an unpredictable way when formulating future predictions that extend over one day ahead.

3.4 Triple exponential smoothing (Holt-Winters) model

Contrary to the previous models, the Holt-Winters model tries to capture seasonal phenomena as well as temporal trends. The seasonal component in the model will explain the repeated fluctuations, and it will be characterized by the length of the season, which is the period after which the repetition of the oscillations begins, Szmit et al. (2012). For each observation in the season, its own component is formed. In our data the length of the season is set to 96×7 (where 96 corresponds to the daily seasonality; and 7 corresponds to the weekly seasonality), which results in 672 seasonal components, one for each 15 minute interval in a day of the week. In particular, the future estimate of a quantity y at time t, according to the Holt-Winters model, is given by:

$$\hat{y}_d^{\mathrm{HW}}(t) = L(t-k) + kP(t-k) + S(t-T) \qquad (5)$$

where $L(t)$ is the *level component*, given by

$$L(t) = \alpha(y(t) - S(t-T)) + (1-\alpha)(L(t-1) + P(t-1)), \qquad (6)$$

$P(t)$ is the *trend component*, given by

$$P(t) = \beta(L(t) - L(t-1)) + (1-\beta)P(t-1), \qquad (7)$$

and $S(t)$ is *season component*, given by

$$S(t) = \gamma(y(t) - L(t)) + (1-\gamma)S(t-T). \qquad (8)$$

We have used the following notation: k is the forecasting range k=96, $y(t)$ is the real (measured) value of the electricity load at time t, T is the time series period, α is the data smoothing factor, β is the trend smoothing factor, and γ is the seasonal change smoothing factor. Furthermore, α, β, $\gamma \in (0,1)$. In the case of creating sequential forecasts, then $y(t)$ is replaced by the corresponding estimate at the same time.

In the above model, the level component tries to capture a baseline load level (or reference), while the trend component approximates (through a low-pass filter) how the level component varies with time (within a few days). Finally, the season component tries to capture lower frequency dependencies (over longer periods of time). Thus, overall the Holt-Winters model can be thought of as a combination of both high- and low-frequency temporal dependencies.

To train this model, we use the root-mean squared error (RMSE) as the loss function, which measures the quality of the fitting to the train data set. Then, we evaluate the cross-validation value of the loss function using the α, β, γ parameters of the model, and then we change the parameters in accordance to its gradient. Since there is a constraint on the values of the smoothing parameters, which should remain within $(0,1)$, we implement the truncated Newton conjugate gradient to update the parameters.

3.5 Seasonal auto-regressive integrated moving average (SARIMA) model

The ARIMA model (Auto-Regressive Integrated Moving Average) is one of the most common methods for analyzing and forecasting time series. It is an extension of ARMA models for non-stationary time series, which can be made stationary by taking differences (of some order) from the original time series. ARIMA uses three main parameters (p,d,q), which are expressed as integers. These three parameters together take into account seasonality, tendency, and noise in the data sets. In particular,

- p is the auto-regressive order, which allows to incorporate previous values of the time series;
- d is the order of integration, which allows to incorporate previous differences of the time series; and
- q is the order of the moving average, which allows for setting the model error as a linear combination of previously observed error values.

The main disadvantage of this model is that it does not support seasonal time series, which makes it impossible to use it to predict time series of energy consumption, characterized by strong seasonality, as in case of one day or one week seasons.

A variation of the ARIMA model, namely SARIMA can be used instead to also track seasonality. In this model, the parameters (p,d,q) are considered as the non-seasonal parameters, that remain the same as above. Additionally to these parameters, we also introduce parameters (P,D,Q) which are defined similarly to (p,d,q), but apply instead to the seasonal component of the time series. Finally, parameter S describes the period of the season in the time series (96 if the season corresponds to one day, 7×96 if the season corresponds to one week, etc., where 96 refers to the granularity of sensor data within one day). Similarly to the Holt-Winters model, this is also a black-box model that captures the seasonal effects and it will be used in order to better evaluate the performance of the derived auto-regressive-based forecasting models.

The selection of the parameters $(p, d, q)(P, D, Q)S$ was based upon the recommendations presented in Akpinar and Yumusak (2016). Therein, it is recommended that the conditions $d + D \leq 2$, $P + Q \leq 2$ should be satisfied. In order to compute the most appropriate set of parameters, we first created an enumeration of the model parameters, which were then compared by using the Akaike information criterion (AIC). This process led to the following SARIMA parameters: (1,1,1)(1,1,1)96.

4. PERSISTENT-BASED REGRESSIVE MODELS

In this section, we introduce two classes of persistence-based models which try to integrate either multiple models, as in *persistence-based auto-regressive model* (PAR), or a large number of persistence parameters that better try to capture seasonal phenomena, as in the *seasonal persistence-based regressive model* (SPR). In the following subsections, we present the details of these models.

All models in this section have been trained by setting up a linear regression optimization problem, which was iteratively solved using the Recursive Least Squares algorithms, cf., Sayed (2003).

4.1 Persistence-based auto-regressive (PAR) model

As we have discussed above, the persistence models can capture low-frequency temporal dependencies in the load profile (extending over multiple days or weeks), while auto-regressive models can capture high-frequency temporal dependencies (within the same calendar day). Furthermore, auto-regressive type of models can work well only within a short-term future horizon of a few hours. For this reason, in this subsection, we would like to also consider the possibility of an optimal combination between the two types of models. In principle, this idea fits well to the expert-based forecasting methods discussed in Cesa-Bianchi and Lugosi (2006) and transfer learning methodologies, such as Grubinger et al. (2017).

Briefly, such an optimal combination of the two forecasting methods assumes a combined prediction of the form:

$$\hat{y}_d^{\mathrm{PAR}}(t|a_1, ..., a_n, b_0) = a_1 \cdot \hat{y}_d^{\mathrm{AR}}(t-1) + \cdots + a_n \cdot \hat{y}_d^{\mathrm{AR}}(t-n) + b_0 \cdot \hat{y}_d^{\mathrm{PM}}(t) \quad (9)$$

In this case, we would like to compute the new set of weights a_1, a_2, ..., a_n, b_0 that corresponds to the optimal combination of the high-frequency temporal dependencies (captured by the auto-regressive terms) and the low-frequency temporal or seasonal dependencies (captured by the last persistence term).

4.2 Seasonal persistence-based regressive (SPR) model

Standard persistence models formulate forecasts through averaging of the load consumption during the same time intervals on previous days. The main assumption lies on the fact that residents/users behave almost the same during the same time of each day. However, when forecasts are requested with the fine granularity of 15min intervals, even small modifications in the schedule of the users may have a significant impact on the forecast accuracy. For example, a half-hour difference in the time a morning schedule is executed will result in large prediction errors on the following day.

For this reason, we would like to create models that are more robust to such small variations in the users' schedule. For example, instead of using the load consumption on the same 15min interval on the previous day, we may use the average load consumption over a larger time window on the previous day (e.g., one-hour window). In this way, we may better capture persistence in the schedule of the users even with small variations in its execution times. Furthermore, the total energy consumption could be another persistence factor that can reduce the uncertainty of the forecasts. For example, note that, although the execution of a day's schedule may slightly vary, usually people consume roughly about the same energy every day.

To this end, we introduced a set of persistence factors that try to reduce the level of uncertainty involved due to small variations in the schedule of the users. In particular, the set of factors considered are the following ones:

- L, electricity load;
- rs, rolling sum of the electricity load within one hour time-window (four 15-min intervals);
- d, flag (integer variable) of the type of day (e.g., business day versus weekend);
- Lh, the total energy consumption within one hour time-window (four 15-min intervals);
- Ld - percentage of the current 15min-interval's energy consumption over the average total energy consumption in one day;
- DLh - difference in hourly energy consumption within the last two hours;
- LC - low energy consumption flag (boolean variable), when the total load is less than 20 percent of the mean load consumption value during one day;
- PC - high energy consumption flag (boolean variable), when the total load is more than 150 percent of the mean load consumption value during one day.

Using these persistence factors, we defined a linear regression model using the features depicted in Table 1.

Table 1. Energy data features for different days

Short name	Target day	Previous day	Day one week ago
L	-	w_0	w_1
rs	-	w_2	w_3
d	w_4	-	-
Lh	-	w_5	w_6
Ld	-	w_7	w_8
DLh	-	w_9	w_{10}
LC	-	w_{11}	w_{12}
PC	-	w_{13}	w_{14}

5. EVALUATION

The RMSE metric was chosen for evaluating the performance of the forecasting models. In Tables 2–3, we can see the overall RMSE performances of the previously discussed models, where we have used real load consumption data from two buildings (*Building A* and *Building B*).

We have observed that the predictions of the two persistence models of Section 3.2, namely the CLD model (which uses three prior same days) and the N-day model (which uses the average value over $N = 10$ previous consecutive days), were often overestimating the electrical load. Also, these models were not able to identify the main load peaks during one day.

Table 2. RMSE of forecasting models for *Building A*'s electricity consumption

Model	One day 03-02-2017 -	One week 01-02-2017 07-02-2017	One month 17-11-2016 16-12-2016	Half a year 16-11-2016 15-05-2017
CLD	124	122	116	126
N-day	123	118	107	115
HW	164	131	130	135
SARIMA	127	118	112	119
PAR	126	118	107	114
SPR	115	107	102	100

Table 3. RMSE of forecasting models for *Building B*'s electricity consumption

Model	One day 22-10-2016 -	One week 17-10-2016 23-10-2016	One month 01-10-2016 31-10-2016	Half a year 16-10-2016 15-04-2017
CLD	153	163	189	128
N-day	153	188	205	120
HW	380	289	284	138
SARIMA	211	141	163	117
PAR	174	121	159	110
SPR	103	65	82	69

The Holt-Winters model turned out to be very sensitive to seasonality, while the performance was better when the season corresponds to 1 week, as opposed to 1 day. However, almost all the predictions turned out to overestimate the electricity load, which was probably caused by the attempt of the model to track the seasonal trend. Despite the changes in average daily energy consumption, there is no seasonal trend in this series of data.

The SARIMA model tried to smooth out the profile and practically did not predict single peaks or rapid decreases of energy consumption during one day. It should be noted that SARIMA very well described the night mode of energy consumption, as well as the falling edge of the load profile from 18:00 evening peak to 00:00 hours approximately.

The combined PAR model significantly improves the prognostic qualities of the separate persistence models. The model was less likely to produce erroneous local peaks and was very good at repeating some patterns in the load profile, in particular, a slight increase in the electricity load during the morning, specifically from 6:00 to 8:00 am.

Figure 1 depicts the forecast accuracy of PAR and SPR in *Building A* and in comparison with the standard models previously presented. Furthermore, in Tables 2–3, we see that SPR performs significantly better than PAR, as well as traditional HW and SARIMA methods. This difference should be attributed to the fact that SPR utilizes a larger family of persistence factors that significantly reduces the impact of the users' schedule uncertainty.

6. CONCLUSIONS AND FUTURE WORK

The proposed SPR forecasting model, which is specifically tailored for load forecasting in residential buildings, and integrates a large family of persistence factors, significantly improved the forecasting accuracy, as compared to standard black-box models. In particular, we presented an extensive comparative analysis with Holt-Winters, SARIMA, and Auto-Regressive-based models. The proposed SPR model is also computationally efficiently, since it is based on a linear regression and can be implemented recursively. A potential drawback though is that it assumes a linear dependence between predictors and features which might be restrictive in energy data.

As a future work, we would like to also investigate the further improvement that we could attain by also allowing nonlinear dependencies in the SPR model. Also, the results of this study can be used in the future to assess the quality of load forecasts of more complex models obtained by neural network integration using the ensemble technique.

The energy consumption in a building is mainly influenced by human behavior. In the future, data related to user management teams along with meteorological parameters can be taken into account in forecasts. The energy prediction system can be implemented for forecasting in the day-ahead mode, as well as for forecasting in modes close to soft real time.

REFERENCES

Akpinar, M. and Yumusak, N. (2016). Year Ahead Demand Forecast of City Natural Gas Using Seasonal Time Series Methods. *Energies*, 9(9), 727.

Alfares, H. and Mohammad, N. (2002). Electric load forecasting: Literature survey and classification of methods. *International Journal of Systems Science - IJSySc*, 33, 23–34.

Cancelo, J.R., Espasa, A., and Grafe, R. (2008). Forecasting the electricity load from one day to one week ahead for the Spanish system operator. *International Journal of Forecasting*, 24(4), 588–602.

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning, and Games.* Cambridge University Press, New York, NY, USA.

Chasparis, G.C., Pichler, M., Spreitzhofer, J., and Esterl, T. (2019). A cooperative demand-response framework for day-ahead optimization in battery pools. *Energy Informatics*, 2, 1–17.

Chen, H., Vidakovic, B., and Mavris, D. (2004). Multiscale forecasting method using armax models. *Current Development in Theory and Applications of Wavelets*, 4.

Chitsaz, H., Shaker, H., Zareipour, H., Wood, D., and Amjady, N. (2015). Short-term electricity load forecasting of buildings in microgrids. *Energy and Buildings*, 99, 50–60.

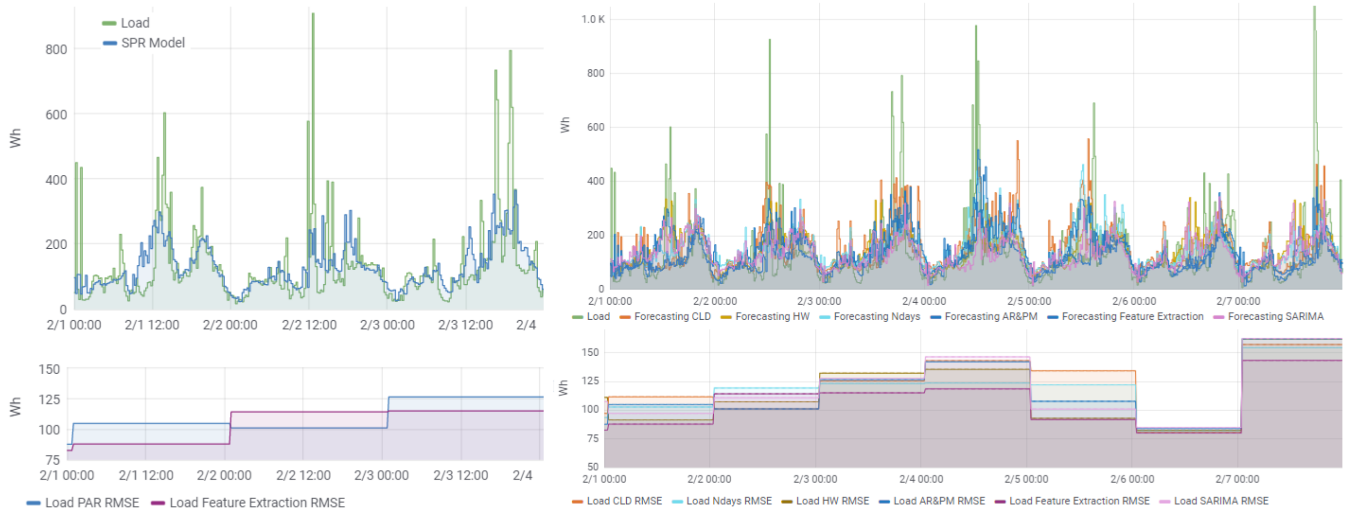

Fig. 1. Comparison of seasonal persistence-based regressive (SPR) and persistence-based auto-regressive (PAR) model with other models and *Building A* measured electricity consumption data. (The SPR and PAR models are also referred to as “Feature Extraction” and “AR&PM” models, respectively.)

Christ, M., Braun, N., Neuffer, J., and Kempa-Liehr, A.W. (2018). Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh A Python package). *Neurocomputing*, 307, 72–77.

Christ, M., Kempa-Liehr, A.W., and Feindt, M. (2017). Distributed and parallel time series feature extraction for industrial big data applications. *arXiv:1610.07717 [cs]*.

Clements, A., Hurn, A., and Li, Z. (2016). Forecasting day-ahead electricity load using a multiple equation time series approach. *European Journal of Operational Research*, 251(2), 522–530.

Grubinger, T., Chasparis, G.C., and Natschlger, T. (2017). Generalized online transfer learning for climate control in residential buildings. *Energy and Buildings*, 139, 63–71.

Haben, S., Giasemidis, G., Ziel, F., and Arora, S. (2019). Short term load forecasting and the effect of temperature at the low voltage level. *International Journal of Forecasting*, 35(4), 1469–1484.

Haben, S., Ward, J., Greetham, D.V., Singleton, C., and Grindrod, P. (2014). A new error measure for forecasts of household-level, high resolution electrical energy consumption. *International Journal of Forecasting*, 30(2), 246–256.

Hippert, H., Bunn, D., and Souza, R. (2005). Large neural networks for electricity load forecasting: Are they overfitted? *International Journal of Forecasting*, 21(3), 425–434.

Hong, T., Gui, M., Baran, M.E., and Willis, H.L. (2010). Modeling and forecasting hourly electric load by multiple linear regression with interactions. In *IEEE PES General Meeting*, 1–8.

Hong, T. and Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3), 914–938.

Kychkin, A. (2016). Synthesizing a system for remote energy monitoring in manufacturing. *Metallurgist*, 59, 752–760.

Kychkin, A. and Mikriukov, G. (2016). Applied data analysis in energy monitoring system. *Problemele energeticii regionale*, 2(31), 84–92.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice Hall Ptr, Upper Saddle River, NJ, 2nd edition.

Notton, G. and Voyant, C. (2018). Forecasting of Intermittent Solar Energy Resource. In *Advances in Renewable Energies and Power Technologies*, 77–114.

Sayed, A. (2003). *Fundamentals of Adaptive Filtering*. John Wiley & Sons, Inc., New Jersey.

Soares, L.J. and Medeiros, M.C. (2008). Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. *International Journal of Forecasting*, 24(4), 630–644.

Szmit, M., Adamus, S., Bugala, S., and Szmit, A. (2012). Implementation of Brutlag’s algorithm in Anomaly Detection 3.0. In *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 685–691.

Xu, Y., Li, N., and Low, S.H. (2016). Demand Response With Capacity Constrained Supply Function Bidding. *IEEE Transactions on Power Systems*, 31(2), 1377–1394.

Ye, R. and Dai, Q. (2018). A novel transfer learning framework for time series forecasting. *Knowledge-Based Systems*, 156, 74–99.