A Scalable Long-Horizon Forecasting of Building Electricity Consumption

Naveen Kumar Thokala[✉], S. Spoorthy[✉], and M. Girish Chandra[✉]

TCS Research and Innovation, Bangalore 66, Karnataka, India
{naveen.thokala,s.spoorthy,m.gchandra}@tcs.com

Abstract. Load Forecasting plays a key role in the efficient operation of the building energy management systems. In this work, a framework is proposed for effective scalable implementation of long-term (month and quarter ahead) building load forecasting. It comprises of techniques to deal with outliers and missing values, dynamic input feature selection as well as a hybrid algorithm combining direct and recursive strategies for forecasting. The solution is successfully validated using the real consumption data of six office buildings and further the average accuracies of 92–95% and 88–92% for month and quarter ahead respectively, corroborates its usefulness.

1 Introduction

Continuous power supply is vital for the effective functioning of commercial buildings. Electrical load forecasting solutions help commercial building managers assess their energy demands, and at the same time, help electrical utilities in planning their supply operations. These aspects help to avoid revenue losses due to power supply disruptions and align supply with the demand and vice-versa. Needless to say, this is vital for present-day demand-supply conditions. To help in such energy management, a range of load forecasting solutions have been developed of late for short term, medium term and long term, depending on the horizon of the forecast. These horizons may cater to the requirements of hours ahead, day-ahead, quarter-ahead and month-ahead forecasting of power consumption. Usually, *Short-term Load Forecasts*, such as hours-ahead and day-ahead forecasts, help the building manager to streamline the power consumption by adopting peak-load shaving, time-of-use pricing/demand response and energy bidding approach [1,3,4]. The *Medium-term* or *Long-term* building energy forecasts, i.e month and quarter-ahead forecasts, respectively, are useful in assessing fuel resources required for the continuous operation of the building, budgeting etc. Medium-term and Long-term forecasts also can be used at the distribution level so that electrical utilities can plan their operations of the electrical power system efficiently.

In the literature, various solutions were proposed for *Short-term load forecasting* of building-level power consumption, based on statistical and machine

learning approaches [2,3,5,6,10]. However limited attention was given to *Long-term forecasting*. Recently, Naveen et al. [4] have proposed Non-linear AutoRegressive with eXogenous (NARX) Neural Network and SVR based month ahead forecasting. One of the major challenges in long-term building load forecasting is the time horizon for which independent (explanatory) variables also need to be forecasted. Therefore, in the current work, we focus mainly on the strategy to be used for long-term forecasting rather than resolve the function to be used to model the time-series. There are two well-known strategies in the literature for forecasting (i) Direct Forecast (DF) and (ii) Recursive Forecast (RF). Direct Forecast uses the data until the current instant and maps all the future values as a function of the past values of the time series. On the other hand, RF forecasts one value in future and the forecasted value is augmented to the input while forecasting the second value and so on until the last forecast value. Both of these strategies have their own strengths and weaknesses [11]. RF performs better when the model is correctly specified [9], else the forecasts will be largely biased. With DF, it is robust to model misspecification but the approach may lead to too much variance in the forecast with respect to the time-series data input. The selection of any one of these methods is a compromise between bias and variance. In order to take advantage of strengths in both the methods, a Hybridized Direct-Recursive multi-step ahead strategy is proposed in this paper. The whole proposed solution consists of many stages to implement long-term building load forecasting.
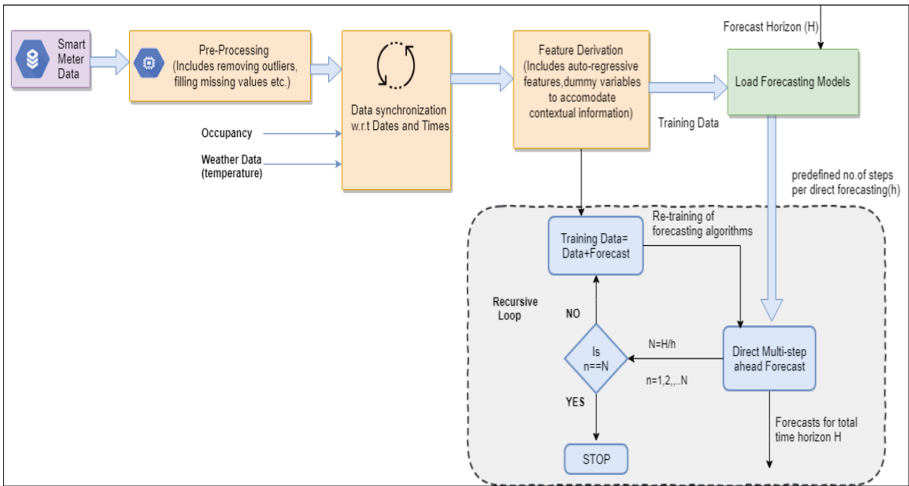

Fig. 1. Block diagram for building load forecasting

The block diagram in Fig. 1 depicts the framework for the proposed hybrid recursive-direct forecasting strategy. The solution includes a pre-processing stage to deal with outliers/missing values, followed by synchronization of smart meter

data with other sensory data. In feature derivation stage additional features necessary for the forecast are derived. The algorithm or the method to be employed for modelling the building load consumption depends on the time horizon of forecasting and the granularity of the data. Figure 1 captures this aspect as well.

The major contributions of this paper are

– An effective algorithm for detecting the outliers and treatment of missing values.
– An useful strategy for long-horizon forecasting using hybridized direct and recursive methods.

The organization of the paper is as follows. Data pre-processing and Synchronization steps are detailed in Sect. 2. The proposed hybridized algorithm is explained in Sect. 3.3. The month ahead load forecasting algorithms and the associated results are explained in Sect. 4. Corresponding details for quarter-ahead are provided in Sect. 5. There is a discussion on the results in Sect. 6 and conclusions are presented in Sect. 7.

2   Data Pre-processing

While developing the models for long-horizon forecasting, we have considered the past data of buildings' power consumption for a period of 1.5 years. The other sensory information useful in modelling the building's power consumption like occupancy and weather information are also considered for the same duration. Often, such data needs to be cleaned first before being used for using the data for analysis, due to the discrepancies that enter into the databases during the data acquisition phase.

2.1   Missing Values and Outliers in Power Data

It is not uncommon for outliers or abnormal deviations from the general or historical data values at a given time to show up in power consumption values. Such values are mostly due to cases such as errors arising from sensor placements, logging failures and other data-acquisition based problems. Further, different influencing factors such as temperature, weekday/weekend, time of the day etc. on the pattern of building energy consumption may some times cause sudden variations in energy consumption. The average power consumption also changes with seasons and the building location (psychrometric influences) [14]. Therefore, the possibility to mistake these variations to be outliers also exists. Such being the case, to detect and replace outliers, we can make use of historical data.

We propose the following process to detect and replace outliers. Let us consider detecting and replacing a value at an instant i ($Y(i)$) for a particular building. We prepare a block of data values from the historical data by picking the values existing during the same season, same day of the week and at the same time of the day. This block can be further refined by considering whether the instant i belongs to either a working day or holiday, and correspondingly

choosing the values from the historical data; this block can be represented by the vector $\mathbf{d}(i)$. The entries in this constructed block are averaged and the resulting average is compared with the value in question $(Y(i))$. If the latter deviates from the average by a large amount, we can declare the value as an outlier and is replaced with the average. The deviation can be ascertained by considering a suitable threshold; denoting the average by μ_i and the threshold by τ, $Y(i)$ is an outlier if $Y(i) > \tau * \mu_i$. Similarly, we can handle the missing values; for this purpose, the missing value at the instant i is estimated as μ_i, that is, we would be using $Y(i) = \mu_i$. It is useful to note that this simple technique is effective in handling long duration of missing values or a burst of no data, which is often encountered in the realistic scenarios.

With the good amount of historical data at our disposal, the aforementioned strategy is adopted in the present work. In fact, when experimented with the other existing sophisticated techniques, the proposal outscored them; comparisons are not captured here. When the prior data is limited, other techniques (for example, using interpolation/filtering in the graph signal domain) can be used to negotiate the missing values and outliers to some extent. These are again not covered in this paper.

2.2   Data Synchronization

The next step, *data synchronization*, is carried out to have an appropriate mapping of the power consumption data with other sensory information like occupancy and temperature. The occupancy and temperature data available with us are at one-hour granularity. To synchronize, the temperature is interpolated every 15 min using the hourly values (as temperature varies slowly), occupancy is maintained constant and replicated every 15 min during the one-hour time-period.

3   Hybridized Recursive-Direct (HRD) Multi-step Ahead Forecast

Time-series forecasting is defined as an extrapolation of the time series for the future dates or times and it requires modelling the time-series in terms of its components like the trend, seasonality, cyclic patterns, and exogenous variables if any. Forecasting involves developing the models using the historical data and forecasting the future values of the time-series [12].

Let Y be a stationarized time series and $Y(t+1)$ the value of the time series Y at $(t+1)$. Then $Y(t+1)$ can be modeled as follows:

$$\hat{Y}(t+1)|_t = f(Y, X) \tag{1}$$

where f is a function and its properties are decided by the learning algorithm considered for modelling the time series. In the linear case, $Y(t+1)$ would be the linear function of lag values of Y and other independent variables x_i, then the function f takes the following form

$$\hat{Y}(t+1)|_t = \sum_{i=1}^{n} \phi_i * Z_i + \epsilon_t \tag{2}$$

where, ϕ represents the parameters of the function learnt and Z represents the vector consisting of both lagged values of Y and independent variables X that impact Y. In general, all the past data might not be used for modelling, only appropriate lags that are useful are used. There are many ways suggested in the literature to decide upon the lag-length, such as Akaike information criterion (AIC), Bayesian Information Criteria (BIC) and auto-correlation function or correlogram.

In multi-step ahead forecasting, it requires to forecast values at multiple steps ahead, for example forecasting $Y_{(t+1)-(t+h)}$ at t. There are two methods that are used in the state-of-art to do multi-step ahead forecasting namely (i) Direct forecast and (ii) Recursive forecast.

3.1   Direct Forecast

Direct forecast uses a static mapping to forecast future values. Direct forecasts are made using a horizon specific estimated model. For example, multi-step ahead forecast for h steps at t for the series y is done in the following way using direct forecast methodology.

$$\hat{Y}(t+1)|_t = f_1(Y, X)$$
$$\hat{Y}(t+2)|_t = f_2(Y, X)$$
$$.$$
$$.$$
$$\hat{Y}(t+h)|_t = f_h(Y, X)$$

$$\tag{3}$$

where $Y(t), Y(t-1), ..Y(1)$ is the time series Y at time t. X represents the eXogenous variables. $\hat{Y}(t+1), \hat{Y}(t+2), ..., \hat{Y}(t+h)$ are the forecasted values given the time series Y at t. $f_1, f_2, ..., f_h$ are the different functions trained to forecast the values of the series at different instants of time. For every step in the forecast horizon, separate function is trained using the past data of Y.

3.2   Recursive Forecast

On the other hand, recursive forecast involves forecasting the multiple values of the series each at a time in a recursive fashion. In recursive forecasting, a single function is trained and parameters of the function are re-estimated at every time step with the new sample adding to the time series. The recursive forecasting for multi-step ahead forecast is of the form

$$\hat{Y}(t+1)|_t = f(Y, X)$$
$$\hat{Y}(t+2)|_t = f(\hat{Y}(t+1)|_t, Y, X)$$
$$.$$
$$.$$
$$\hat{Y}(t+h)|_t = f(\hat{Y}(t+h-1)|_t, \hat{Y}(t+h-2)|_t, ..Y, X)$$

$$\tag{4}$$

where $\hat{Y}(t+1)$ is forecasted using the data from the beginning to the current instant t, while forecasting $\hat{Y}(t+2)$, the forecasted value at instant $t+1$ i.e $\hat{Y}(t+1)$ is added to the input series to re-train the function f, similarly $\hat{Y}(t+h-1)$, $\hat{Y}(t+h-2), ..., \hat{Y}(t+1)$ are used while forecasting $\hat{Y}(t+h)$ in a recursive fashion.

3.3   Hybridized Direct-Recursive (HDR) Forecast

In theory *recursive forecasts* are more accurate than *direct forecasts*, if models are specified correctly [9]. As the direct forecast uses the separate models for each and every step in the forecast horizon, information between the consecutive points of the time-series is not considered resulting in high variance in the forecast. Recursive forecasts suffer from biases as the forecasted values are iteratively used as inputs to forecast future values in the time series. The error in forecasted values propagates as these are used as inputs for further forecasts. Choosing between these two forecasts is a tradeoff between the bias and estimation variance. Therefore a **Hybrid Strategy** using both direct and recursive methods is devised to address the weaknesses of both the methods. The proposed hybridized direct-recursive (HDR) strategy is to have a total forecast horizon H divided into n slots each of length h. Direct multi-step forecast strategy is used to forecast the first h values and then these forecasted values of the initial slot (h steps) are used as an input for forecasting the second slot. This continues until the last slot.

$$
\begin{aligned}
&\hat{Y}(t+1, ., t+h)|_t = f(Y, X) \\
&\hat{Y}(t+h+1, ., t+2h)|_t = f(\hat{Y}(t+h), ., \hat{Y}(t+1), Y, X) \\
&. \\
&. \\
&\hat{Y}(t+(n-1)h+1, ., t+nh)|_t = f(\hat{Y}(t+(n-1)h), ., \hat{Y}(t+1), Y, X)
\end{aligned}
\tag{5}
$$

Estimation Variance and Bias: In Direct-Forecasting method, multiple functions are used to obtain forecasts at multiple time-instants due to which the total variance in the forecast is the addition of all the individual variances at each time-step. In Recursive-Forecasting, a certain bias, b_i, is induced at every time-instant i due to the recursive nature of forecasting. This bias is additive in nature, in that it adds up to the previous bias at every recursive step.

In HDR, we observe that the sum total variance is restricted to $\Sigma_1^h \sigma_i$, as against $\Sigma_1^n \sigma_i$ in Direct-Forecasting ($h << n$). Similarly, the total bias in HDR is $\Sigma_h^n b_i$ as against $\Sigma_i^n b_i$ for Recursive-Forecasting. This observation re-affirms the robustness of HDR against Direct-Forecasting and Recursive-Forecasting methods.

The proposed HDR forecast strategy is used to forecast month and quarter ahead buildings' total power consumption. The actual office buildings' data is used for demonstrating the performance of the proposed approach.

4   Buildings' Month Ahead Load Forecasting

Smart meter measures total buildings' power consumption once in every 15 min. The buildings' power consumption is influenced by many factors like temperature (HVAC loads), occupancy and other factors like the working day or holiday etc. making it difficult for the linear algorithms like linear regression (LR), ARIMA etc., to forecast accurately. Artificial Neural Networks (ANN) and the Support Vector Regression (SVR) are the two well-known techniques for modelling non-linear and complex time series as mentioned in [4,7,8]. The modelling of the time series for the month ahead forecasting includes (i) Data Pre-processing as explained in Sect. 2. (ii) Feature Derivation and Selection (iii) Modelling of the time series using ANN/SVR.

4.1   Feature Derivation and Selection

New features derivation and selection are carried out for time-series modelling of buildings' power consumption. Dummy variables are created to capture the contextual information like Day of the week, Time of the Day, Holidays etc. The lags of buildings' power consumption are selected using the partial autocorrelation function (pacf). The power consumption lags (order selected using the pacf function) together with dummy variables used to capture contextual information forms the *Predictor Matrix*. Predictor Matrix is the input to the learning model and the buildings' power consumption is the output.

4.2   HDR Based Month Ahead Forecasting Using ANN and SVR

The modelling of the month ahead forecasting is carried out as mentioned in Sect. 3.3. The total forecast horizon for the month ahead forecasting at 15 min granularity is $H = 2880$ (30 days with 96 samples every day). The total forecasting horizon is divided into 5 slots, i.e. each slot will have 576-time steps. The direct forecast is used to forecast 576-time steps (6 days) and these forecasted values are added to input to forecast for the next 6 days of buildings power consumption. This is repeated five times to get a month ahead forecasting. The implementation is as explained in the below set of equations (Eq. 6).

$$
\begin{aligned}
&\hat{Y}(t+1,.,t+576)|_t = f(Y,X) \\
&\hat{Y}(t+577,.,t+1152)|_t = f(\hat{Y}(t+576),.,\hat{Y}(t+1),Y,X) \\
&\quad . \\
&\quad . \\
&\quad . \\
&\hat{Y}(t+2305,.,t+2880)|_t = f(\hat{Y}(t+2304),.,\hat{Y}(t+1),Y,X)
\end{aligned}
\tag{6}
$$

ANN is used as the function f in the above equations for performance comparison. The performance of the proposed approach is demonstrated using the actual buildings' power consumption data.
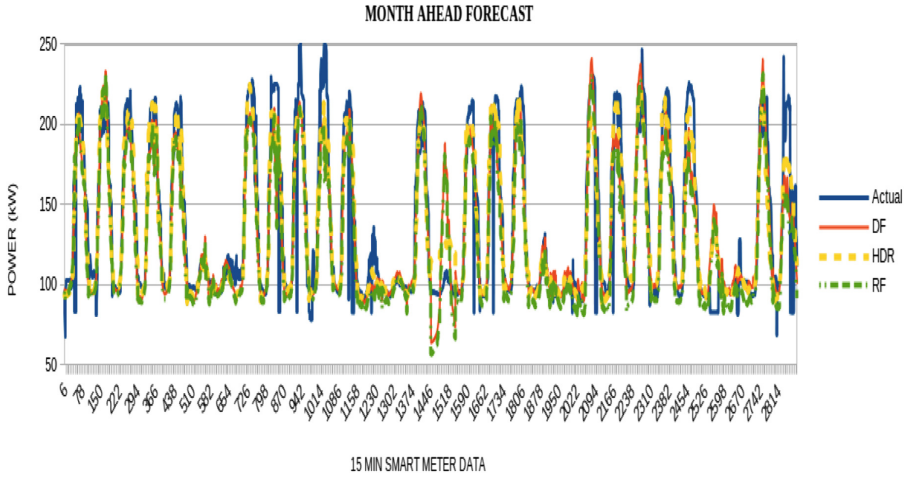
Fig. 2. Month ahead forecast comparison

4.3   Results

The proposed strategy is tested on six buildings for the month ahead forecasting. From Fig. 2, It is clear that the performance of the proposed strategy HDR is either better than the other two approaches DF and RF or matched with the best of the two approaches. For performance evaluation, Symmetric Mean Absolute Percentage Error (sMAPE) and Normalized Root Mean Squared Error (NRMSE) are considered as the error terms as these are scale independent making them applicable for comparing algorithms' performance across buildings' of different capacities. Symmetric MAPE (sMAPE) is considered over MAPE to avoid over penalty to the negative errors. Figures 3 and 4 clearly indicate that the proposed approach has improved forecasts for most of the buildings.
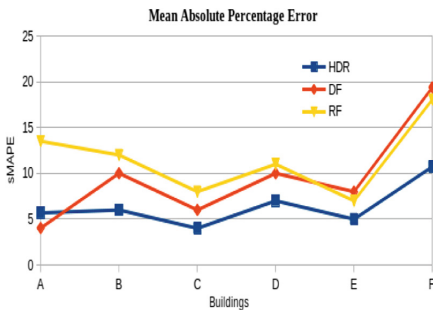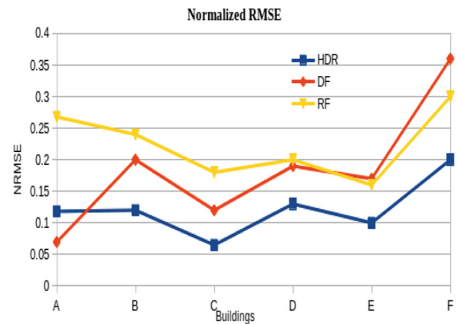

Fig. 3. Symmetric MAPE comparison

Fig. 4. Normalized RMSE comparison

5   Quarter Ahead Load Forecasting

Quarter ahead load forecasting is carried out on buildings' day-wise total energy consumption (kiloWatt hours). Day-wise aggregated building's power consumption is calculated by aggregating 15 min smart meter for the whole day. Due to aggregation, the time series has become less dynamic compared to 15 min granular smart meter data. But, the major challenge in the long forecast horizon is to capture the trend, i.e, average increase or decrease in power consumption with respect to change in season, average temperature, occupancy etc. Linear regression is considered for modelling the energy time-series because (i) aggregated data does not have too many variations to be captured like 15 min granular data and (ii) the linear model can help analyze the impact of various factors on the building energy consumption.

5.1   Feature Extraction

Similar to day-wise energy consumption, day-wise maximum and minimum temperatures, as well as occupancy, are considered as time-series. For the forecast horizon, maximum and minimum temperature forecasts are taken from the weather websites like [13], and the occupancy is extrapolated using ARIMA. Contextual information such as working day, month of the year is captured in the form of dummy variables.

5.2   HDR Based Quarter Ahead Forecasting Using Linear Regression

Multivariate Linear Regression (LR) is used to model energy time-series. The input feature vector for the LR function consists of contextual information (day of the week, working day), minimum and maximum temperatures for a day, day-wise average occupancy and auto-regressive terms (lagged values). The output of the function is buildings day-wise energy consumption. As explained in the Sect. 4.2, the total forecast horizon 90 days is divided into 9 slots, each having 10 days; first slot i.e first 10 days is forecasted using the data available until the current day. These forecasted values are added to the input for forecasting the second slot of days (i.e 11th day to 20th day), the parameters of the learning function are estimated again to forecast. This continues until the last slot. The linear regression function parameters are re-estimated for every slot as shown in the following way.

$$\hat{Y}(t+1,.,t+10)|_t = f(Y, X)$$
$$\hat{Y}(t+11,.,t+20)|_t = f(\hat{Y}(t+10),.,\hat{Y}(t+1), Y, X)$$
$$.$$
$$.$$ \hfill (7)
$$\hat{Y}(t+81,.,t+90)|_t = f(\hat{Y}(t+80), \hat{Y}(t+79)., \hat{Y}(t+1), Y, X)$$

In the above equation, the learning function f is of the form,

$$\hat{Y}(t+1,.,t+10)|_t = \phi_1 * X_1 + \phi_2 * X_2.. + \phi_n * X_n + \epsilon_t \tag{8}$$

where, $X_1, X_2, ... X_n$ are the inputs for the function, ϕ_i represents the weights/parameters trained which could signify the impact of the input features on the output i.e buildings' overall consumption. This model could be of help in understanding and taking control measures for building energy management.
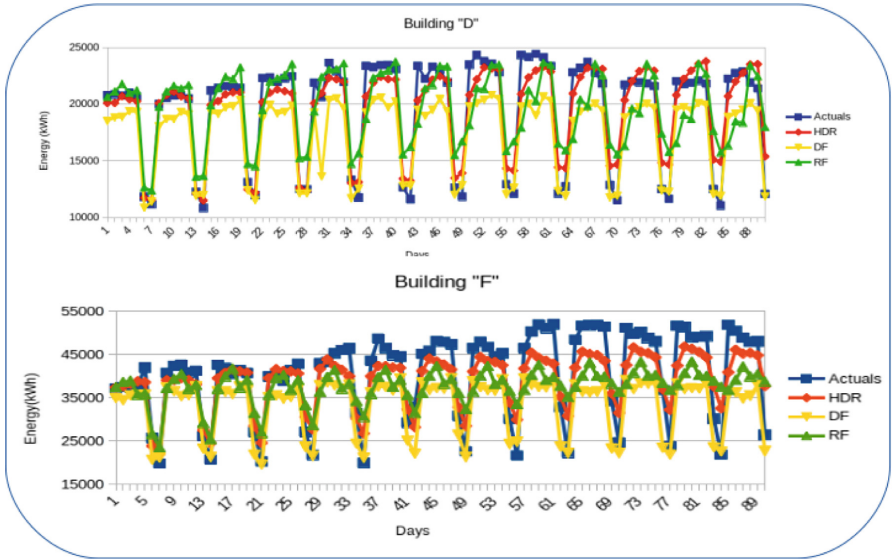

Fig. 5. Quarter ahead forecast comparison

5.3   Results

The proposed approach is used to forecast the six office buildings future energy consumption, the forecast is compared over actual consumption data in real-time. The performance of the proposed algorithms is given in Fig. 5. It could be clearly noticed that HDR forecast strategy has improved performance compared to both the techniques (Direct Forecast and Recursive Forecast). Figures 6 and 7 show that HDR forecast strategy has out-performed the other two strategies
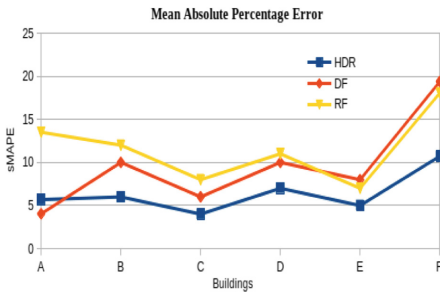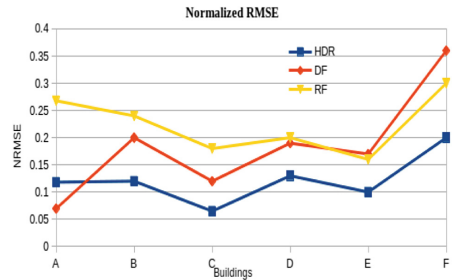

Fig. 6. sMAPE comparison

Fig. 7. NRMSE comparison

for all the buildings (*A-F*). Building *A* is exceptional, where direct forecast (DF) has superior performance compared to the HDR approach.

6   Discussion

HDR strategy is a clear winner compared to Direct and Recursive forecast strategies as shown in the Figs. 3, 6 and 4, 7. From Fig. 5, it could be noticed that (i) the forecasted values using direct multi-step ahead strategy are much lower than the actual consumption (the line with *triangles*); it is because of the static mapping of the future values with the auto-regressive terms and the relation among the consecutive data points is not considered in the forecast horizon and (ii) the forecasted values of the recursive forecast started predicting well, but the error in the initial forecasts accumulated as the forecasting horizon increases making the forecasted values highly biased. The proposed hybrid strategy (HDR) as explained in Sect. 3.3 is able to forecast accurately as the forecast horizon increases as well due to the way it is modelled.

7   Conclusion

An efficient forecasting solution, Hybridized Direct-Recursive (HDR) algorithm is proposed for the long-horizon buildings' load forecasting. The proposed framework is efficient due to its (i) effective logic for handling outliers and missing values (ii) additional contextual features derived to capture the dynamics of the buildings' energy consumption and (iii) the algorithms capability in re-estimating the functional parameters for every new slot in forecast horizon making it more efficient in forecasting accurately even as the horizon increases. We have been able to test the performance of the proposed solution on actual buildings' energy consumption in real-time and the efficacy of the proposed algorithm is demonstrated in the results Sect. 5.3. The proposed framework covers all the steps required for the real-time implementation of the algorithm. Additionally, it scales well in terms of using it across a large number of buildings. Further, the framework can be adapted for different applications.

References

1. Swaroop, R., Abdulqader Hussein, A.: Load forecasting for power system planning and operation using artificial neural network At Al Batinah Region Oman. J. Eng. Sci. Technol. **7**(4), 498–504 (2012)
2. Almeshaiei, E., Soltan, H.: A methodology for Electric Power Load Forecasting. Alexandria Eng. J. **50**(2), 137–144 (2011)
3. Friedrich, L., Afshari, A.: Short-term forecasting of the Abu Dhabi electricity load using multiple weather variables. Energy Procedia **75**, 3014–3026 (2015)
4. Thokala, N.K., Bapna, A., Girish Chandra, M.: A deployable electrical load forecasting solution for commercial buildings. In: 2018 IEEE International Conference on Industrial Technology (ICIT). IEEE (2018)

5. Fernández, I., Borges, C.E., Penya, Y.K.: Efficient building load forecasting. In: 2011 IEEE 16th Conference on Emerging Technologies and Factory Automation (ETFA). IEEE (2011)

6. Mai, W., Chung, C.Y., Wu, T., Huang, H.: Electric load forecasting for large office building based on radial basis function neural network. In: 2014 IEEE PES General Meeting — Conference and Exposition, National Harbor, MD, pp. 1–5 (2014)

7. Ojemakinde, B.T.: Support vector regression for non-stationary time series (2006)

8. Tealab, A., Hefny, H., Badr, A.: Forecasting of nonlinear time series using ANN. Future Comput. Inf. J. **2**(1), 39–47 (2017)

9. Marcellino, M., Stock, J.H., Watson, M.W.: A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. J. Econometrics **135**(1–2), 499–526 (2006)

10. Kaur, N., Kaur, A.: Electricity demand prediction using artificial neural networks in data mining. Int. J. Technol. Res. Eng. **2**(8) (2015)

11. Hamzaçebi, C., Akay, D., Kutay, F.: Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting. Expert Syst. Appl. **36**(2), 3839–3844 (2009)

12. Brockwell, P.J., Davis, R.A., Calder, M.V.: Introduction to Time Series and Forecasting, vol. 2. Springer, New York (2002). https://doi.org/10.1007/978-3-319-29854-2

13. http://www.accuweather.com/en/in/india-weather

14. Paresh, S., Thokala, N., Girish Chandra, M.: Facility energy management based on adaptive thermal comfort protocols: a case-study. In: 5th IEEE International Conference on Signal Processing and Integrated Networks (SPIN) (2018)