

# Short-term Power Load Forecast of an Electrically Heated House in St. John's, Newfoundland, Canada

Hla U May Marma, M. Tariq Iqbal, and Christopher Thomas Seary

**Abstract**— A highly efficient deep learning method for short-term power load forecasting has been developed recently. It is a challenge to improve forecasting accuracy, as power consumption data at the individual household level is erratic for variable weather conditions and random human behaviour. In this paper, a robust short-term power load forecasting method is developed based on a Bidirectional long short-term memory (Bi-LSTM) and long short-term memory (LSTM) neural network with stationary wavelet transform (SWT). The actual power load data is classified according to seasonal power usage behaviour. For each load classification, short-term power load forecasting is performed using the developed method. A set of lagged power load data vectors is generated from the historical power load data, and SWT decomposes the vectors into sub-components. A Bi-LSTM neural network layer extracts features from the sub-components, and an LSTM layer is used to forecast the power load from each extracted feature. A dropout layer with fixed probability is added after the Bi-LSTM and LSTM layers to bolster the forecasting accuracy. In order to evaluate the accuracy of the proposed model, it is compared against other developed short-term load forecasting models which are subjected to two seasonal load classifications.

**Index Terms**—Load forecast, Stationary wavelet transform, Long short-term memory, Neural Network.

## I. INTRODUCTION

Load forecasting with high accuracy is very important for practical power system and smart grids analysis. There are three categories of load forecasting methodologies: LTLF (long term load forecast: more than 1 year), MTLF (medium term load forecast: within 1 month to 1 year) and STL (short-term load forecast: 1 hour to 1 day or 1 week ahead) [1], [9]. Among those, short-term load forecast is more reliable and efficient. STL improves the efficiency and reliability of smart grid including home energy management, demand response implementation, electricity price market design [18]-[20], [23]. Two techniques are commonly used for STL: statistical techniques such as the linear regression model, exponential model etc.; and artificial intelligence techniques [9].

It is studied that among all STL techniques, artificial neural networks (ANNs) is most popular for short-term

electric load forecasting. ANNs has distinct advantages and more accurate prediction compared to others [3],[6]-[8]; which influence more research on neural network based STL. It is found that STL is implemented by neural fuzzy network, recurrent neural network (RNN), wavelet based neural network or hybrid neural network [5],[10]-[13].

Nowadays, hybrid neural networking has become more popular and suitable for learning non-stationary and complex time series data. A hybrid forecasting method consisting of discrete wavelet transform (DWT), autoregressive integrated moving average (ARIMA) and artificial neural network (ANN) is proposed in [14] to forecast daily peak load. Such hybrid model used load data of Fars Electrical Power Company, Iran in 2009 and predicted daily peak load of the system. Recently, it is invested that the performance of STL is improved by using LSTM model [24] but using a single LSTM model has less accuracy than the hybrid model combined with LSTM [5]. The author of [15] designed a hybrid forecasting method called a recurrent inception convolution neural network, a composition of 1-D convolution neural network and LSTM. This model is verified by using power consumption data from three large distribution complexes in South Korea. A combination of short-term wind power forecasting approach based on DWT and LSTM is proposed in [16]; and 12 months data from three wind farms in Mongolia, the Netherlands, and Yunnan, China were used to verify this model. The author of [4] developed a hybrid DWT and collaborative representation (CRT) method. In this method, DWT including CRT is used for feature extraction from the input vector composed by the lagged power load and forecasting is predicted by LSTM. The lagged load variables consist of the load values in last 3 hours of the same day, the last 3 hours and same hour of the day before, and the last 3 hours and same hour of the previous week. It is found that the load features extracted from the lagged power load variable vector provide superior forecasting performance. At the individual household level, a hybrid deep learning methodology combined with LSTM neural network and with SWT is proposed [5], in which SWT decompose the input data into signal components and each signal component is fed to LSTM separately for forecasting. This developed model accuracy is verified by using remote sensor data of five different family houses in London, United Kingdom.

They suggested that SWT alleviates the volatility and increases the data dimensions, improves the accuracy of LSTM forecasting. Although the developed neural network [4]-[5], [15]-[16] can learn the features, overfitting is a threatening problem for large neural networks and reduces the accuracy of forecasting. In order to avoid the overfitting

---

Published on May 12, 2020.

This research was supported in part by the Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2016-04170 and School of Graduate Studies at Memorial University of Newfoundland, Canada.

Hla U May Marma, Memorial University of Newfoundland, Canada.  
(e-mail: hmarma@mun.ca).

M. Tariq Iqbal, Memorial University of Newfoundland, Canada.  
(e-mail: tariq@mun.ca).

Christopher Thomas Seary, Memorial University of Newfoundland, Canada. (e-mail: cts322@mun.ca).

problem of neural network, [17] developed an hourly natural gas demand forecasting method by adding a dropout layer in the neural network which prevents units from co-adapting too much [27]. This model consists of multi-layer Bi-LSTM model and LSTM combined with DWT; DWT and multi-layer Bi-LSTM model is used to decompose the actual data into sub-components and capture the features in the sub-components, respectively; and LSTM predict the hourly natural gas demand. A dropout layer is added after each Bi-LSTM and LSTM layer in this model.

Although neural network based STLF is becoming more popular, [2] indicates that neural network was commonly used for an aggregate level load forecasting and limited research has been undertaken based on individual household [3]-[7] before. Recently, the availability of high frequency data collected by new smart metering system in individual households opens the opportunities to research on individual household load forecasting. In addition, having a large amount of data increases the prediction accuracy of load consumption at the individual household level; adding value to improve efficiency of smart grid technologies, such as home energy management and demand response implementation [18]-[20]. Individual household load forecasting will help to project future load consumption and better manage electricity use.

Although smart meter based individual household load consumption data is highly volatile [21], and univariate time series load forecasting is a challenging problem for deep learning; a hybrid model [5] resolves this problem by using SWT combined with LSTM. The developed model of [5] feeds the original energy consumption values to SWT for decomposition of original signal; but [4] mentions that the load value at any time is correlated to the loads in the previous time steps.

It is reported by Natural Resources Canada's Office of Energy Efficiency in 2013 that in the Canada residential sector, the energy was distributed 63% for space heating, 19% for water heating, 16% for appliance and lighting; and 1% for cooling [25]. In winter, the space heating load is dominant which is not included in summer. This creates complexity and difficulty of intro-class data fitting as space heater is completely unused at summer. In order to overcome the difficulty of data fitting in input and improve the accuracy of forecasting, [26] developed Solar photovoltaic (PV) power forecasting hybrid method based on DWT-CNN-LSTM models; independently established for four weather types : sunny, cloudy, rainy, and heavy rainy days.

In this paper, we propose a robust short-term electric load forecasting model for the individual household level by using the power data from smart meter, installed in house. The deep learning model is based on SWT with a Bi-LSTM and LSTM neural network. A set of lagged power load data vectors is assigned to SWT, it decomposes the vectors and creates sub-components. In order to determine the most appropriate wavelet packet function, a comparison is made among the wavelet functions. The sub-components are individually fed to Bi-LSTM to capture the features by considering the data information bidirectionally. The abstracted features from Bi-LSTM are fed to LSTM for forecast learning. The power forecast is constructed from the

predicted sub-components by using ISWT. This process is independently constructed for two seasonal load classifications. The evaluation of the developed model is verified for all seasonal classified load by using the dataset of an individual household in St. John's, Newfoundland and Labrador, Canada. The power consumption profile of individual household is a very common scenario and reflects the power consumption profile in Canada residences, so this study can represent an important load profile forecasting study.

The paper is arranged as follows: in Section II the data description and observation are analyzed, in Section III a curve fitting method for STLF is described; in Section IV the methodology included with five main parts SWT, Bi-LSTM, LSTM, Dropout layer, ISWT of the proposed short-term power load forecasting model are explained; in Section V the accuracy of the proposed load forecasting model is evaluated by case study using experimental data; and in Section VI conclusions are drawn.

## II. DATA DESCRIPTION AND OBSERVATIONS

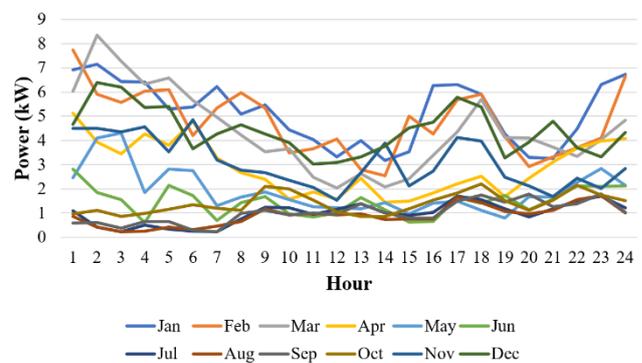


Fig. 1. 24 hourly Weekdays load profile for each month.

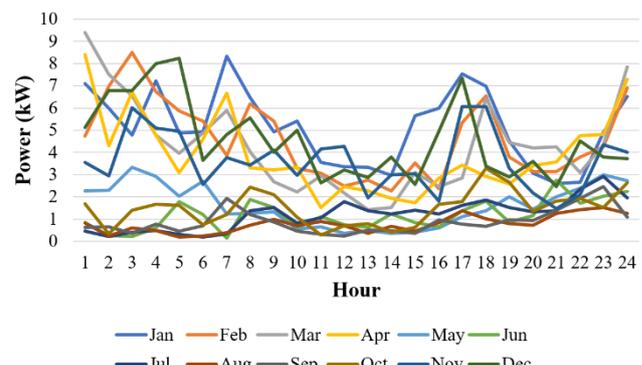


Fig. 2. 24 hourly Weekends load profile for each month.

The power load data of a house in St. Johns was collected by installing current sensors. Five current sensors are installed to measure consumption of the garage, house space heater, domestic water heater, lights and appliances. The total power consumption is calculated by combining the five current sensors dataset. The data was collected every 3 minutes for many years, however only one-year data is used for this study.

The 24-hourly weekdays and weekends load profile for all months are presented in Fig. 1 and Fig. 2 respectively. It is shown that 24 hourly load consumption profile varies from each month. A box and whisker plot of the seasonal

profile is shown in Fig. 3. For each month, the top and bottom line corresponds to that month's overall maximum and minimum, respectively. The middle line of the blue box is the overall average for the whole month. The top and bottom of the blue box present the average of the daily maxima and minima of all of the days in the month, respectively. It shows that the overall average is higher from November to May than from June to October.

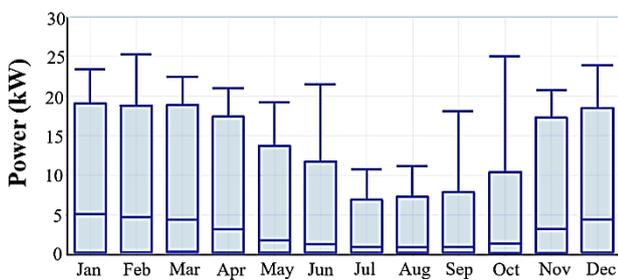


Fig. 3. One-year seasonal load profile of the house

### III. CURVE FITTING METHOD

Firstly, a curve fitting method is used to develop a prediction model for load forecasting. Two variable regression models have been considered here. Here  $x$  stands for hours and  $y$  stands for power consumption. MATLAB Surface fitting toolbox has been used to develop the models and to analyze the developed models' accuracy. Five regression models, known as Sum of Sine (1), Polynomial (2), Power (3), Rational (4), and Weibull (5) models are tested to fit the load forecasting model shown in Fig. 4. It is shown that the actual power load data at the individual household is non-stationary time series data and the developed five regression models are significantly far from the fitting with the actual power load data

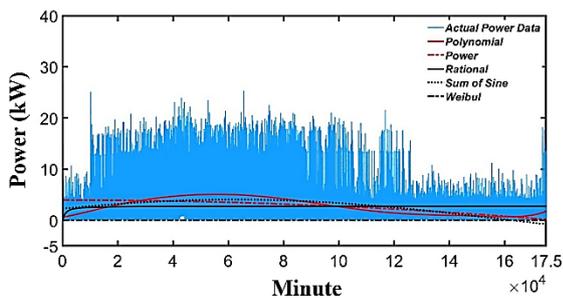


Fig. 4. Curve fitting method for short-term power load forecasting

$$y = a_1 \sin(b_1 x + c_1) \quad (1)$$

$$y = p_1 x^8 + p_2 x^7 + p_3 x^6 + p_4 x^5 + p_5 x^4 + p_6 x^3 + p_7 x^2 + p_8 x + p_9 \quad (2)$$

$$y = ax^b + c \quad (3)$$

$$y = \frac{p_1 x^5 + p_2 x^4 + p_3 x^3 + p_4 x^2 + p_5 x + p_6}{x^5 + q_1 x^4 + q_2 x^3 + q_3 x^2 + q_4 x + q_5} \quad (4)$$

$$y = abx^{b-1} e^{-ax^b} \quad (5)$$

The accuracy of the developed regression models is evaluated by sum square error (SSE), R-square values and root mean square error (RMSE) between experimental and calculated data using these equations. The R-square value represents how closely the fitted model can follow the

variance of the actual data set. It ranges from 0 to 1 where a value closer to 1 and RMSE value closer to 0 represents a better fit. Regression models along with sum square error, R-square values and RMSE are shown in Table I.

From the Table I, it can be seen that for all regression models the R-square value is close to 0; and the SSE and RMSE values are very high, indicating that the curve fitting models are unable to predict load consumption for an individual household load.

TABLE I: REGRESSION MODELS ALONG WITH SUM SQUARE ERROR, R-SQUARE VALUES AND RMSE

Model	SSE	R-square	RMSE
Sum of Sine	$2.648 \times 10^6$	0.1052	3.882
Polynomial	$2.517 \times 10^6$	0.1495	3.785
Power	$2.733 \times 10^6$	0.0766	3.944
Rational	$2.951 \times 10^6$	0.0028	4.099
Weibull	$4.252 \times 10^6$	0.4369	4.92

### IV. METHODOLOGY

This research presents short-term power load consumption forecasting for a house. The residential power is consumed by space heating, water heating, appliances and lighting. Water heater, lighting and other appliance are used daily, on the other hand space heater usage is influenced by variable weather conditions and human behavior. Space heater is used according for certain weather condition specifically November to May in a year as shown Fig. 5; and remaining months of the year, it is excluded from the load. Based on Fig. 1 and Fig. 2, although 24 hourly load profile for June, July, August, September and October vary slowly for both Weekdays and Weekends, the load profile changes abruptly for November, December, January, February, March, April and June. The yearly seasonal load profile Fig. 3 shows that the load with space heater consumed higher average power than the load without space heater. Therefore, the actual power load data is classified as two different types: 1) power loads without space heater, and 2) power loads with space heater. Classified data are processed individually by the proposed short-term power load forecast model as shown in the overall flowchart of proposed model in Fig. 6.

The structure of the proposed forecasting model is illustrated in Fig. 7. There are five main steps for the proposed model: (1) data preprocessing; (2) a lagged power load variable vector is decomposed by SWT; (3) Bi-LSTM feature extraction from each sub-components; (4) LSTM based prediction with dropout layer are applied for each sub components; and (5) finally, ISWT based re-constructor is used to generates actual power load forecast signal by combining predicted sub-components.

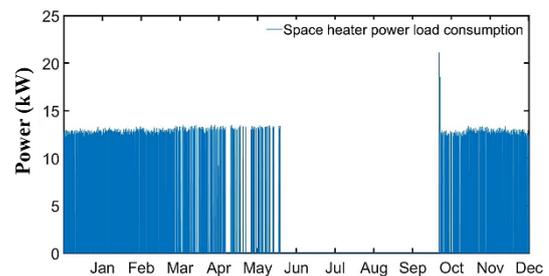


Fig. 5. Yearly space heater load profile for the house

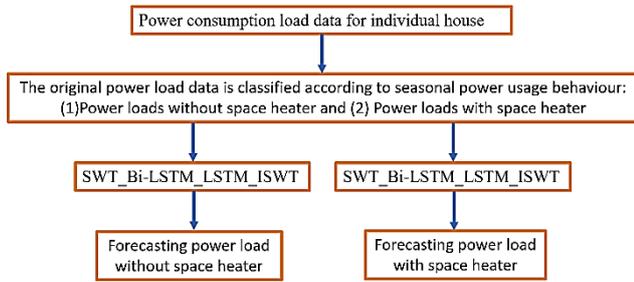


Fig. 6. Overall Flowchart of the short-term power load forecasting model

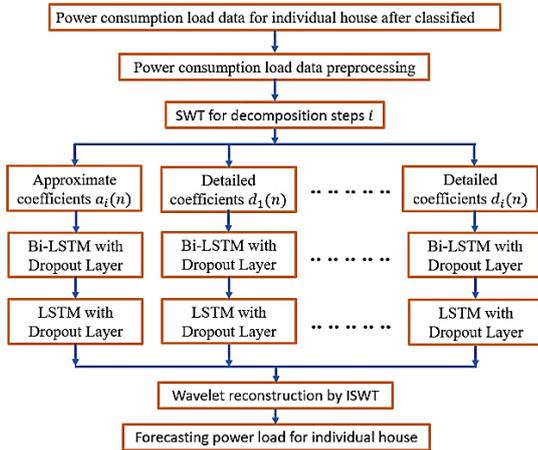


Fig. 7. The structure of the short-term power load forecasting model

### A. Data Preprocessing

There are 175200 data samples collected per year by 3 minutes interval steps; 116160 and 47040 samples are used for power loads with space heater and without space heater deep learning framework.

The power load value for a certain time is correlated to the previous time steps load value. Hence, in order to increase the data dimension, a set of lagged power load variable vector is created from actual data similar to [4]. The vector is generated by considering only the load values in the last three hours of the same day, the last three hours and same hour of the day before, and the last three hours and same hour of the previous week, in 30 minute time steps as shown in (6). The developed variable vector is assigned to each point of the load curve.

$$X = [x(t - 3431), \dots, x(t - 3361), x(t - 551), \dots, x(t - 481), \dots, x(t - 61), \dots, x(t - 1)]^T \quad (6)$$

$X$  is a set of lagged variable vectors for a certain point with dimension  $l \times 1$ ,  $l=20$ . For the  $m$  power load samples data, the training data  $n = m - 3232$ , is considered for the RNN.

### B. SWT Decomposition and Reconstruction

The SWT algorithm is adopted to decompose a signal into wavelets as shown in Fig. 8. SWT is known as non-sampling wavelet transform and a time invariance extension of DWT. At the 1st level of SWT, the original signal  $x(n)$  splits into approximation coefficients  $a_1(n)$  and the detail coefficients  $d_1(n)$ . Then the next level,  $a_1(n)$  splits into two:  $a_2(n)$  and  $d_2(n)$ ; and this step is continued for until the number of decomposition steps  $i$ . The original signal  $x(n)$  can be reconstructed through ISWT whereas  $x(n)$  is

the last level approximation coefficients  $a_i(n)$  and summation of all levels detail coefficients (i.e.  $d_1(n), d_2(n), \dots, d_i(n)$ ).

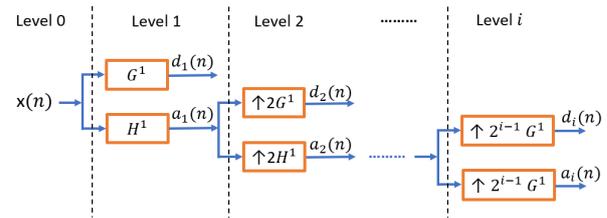


Fig. 8. SWT decomposition for  $i$  level

### C. LSTM and Bi-LSTM

LSTM is a special type of RNN and it is effective for processing time series data. LSTM consists of 'self-connected' memory cells, multiplicative gate units in the hidden state as shown in Fig. 9. Where,  $i, f, g$  and  $o$  represent input gate, forget gate, cell candidate and output gate, respectively. LSTM neural network can be explained by following (7) to (12):

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \quad (7)$$

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \quad (8)$$

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \quad (9)$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \quad (10)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (11)$$

$$h_t = o_t \odot \sigma_c(c_t) \quad (12)$$

$i_t, f_t, g_t, o_t, c_t$  and  $h_t$  denote input gate, forget gate, cell candidate, output gate, cell state and hidden state at time step  $t$ , respectively.  $W, R$  and  $b$  represent the input weights, the recurrent weights, and the bias of  $i, f, g$  and  $o$ ; respectively. State and gate activation function are denoted by  $\sigma_c$  and  $\sigma_g$ ; respectively.

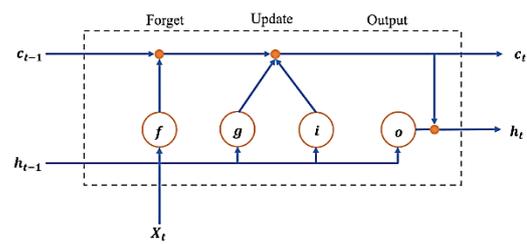


Fig. 9. The architecture of LSTM

In Bi-LSTM model, the same output layer connects the two separate hidden layers known as forward LSTM layer and backward LSTM layer [31]-[33]; as shown in Fig. 10. The previous and future data information are used in this model. The forward output layer sequence,  $\vec{h}_t$  and the backward output layer sequence,  $\overleftarrow{h}_t$  are computed by using input in a positive time sequence and reversed time sequence, respectively. The output layer  $y_t$  can be expressed as follows:

$$y_t = G(\vec{h}_t, \overleftarrow{h}_t) \quad (13)$$

Where,  $G$  is a function, used to generate output based on

$\vec{h}_t$  and  $\overleftarrow{h}_t$ .

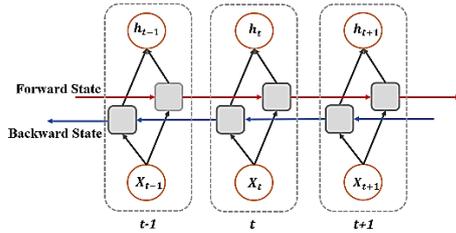


Fig. 10. The architecture of Bi-LSTM

#### D. Dropout Layer

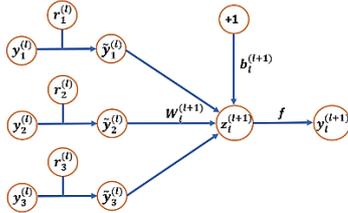


Fig. 11. The dropout network

The dropout layer improves the accuracy of forecasting by preventing the neural network from overfitting. The dropout network is shown in Fig. 11 and can represent as follows:

$$r_j^{(l)} \sim \text{Bernoulli}(p) \quad (14)$$

$$\tilde{y}^{(l)} = r^{(l)} * y^{(l)} \quad (15)$$

$$z_i^{(l+1)} = W_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)} \quad (16)$$

$$y_i^{(l+1)} = \sigma(z_i^{(l+1)}) \quad (17)$$

Where \* presents element-wise product.

A vector of independent Bernoulli random variables for the specific layer  $l$ ,  $r^{(l)}$  each of which has probability being 1.  $y^{(l)}$  and  $\tilde{y}^{(l)}$  denote outputs and thinned outputs for the specific layer  $l$ .  $\tilde{y}^{(l)}$  is obtained by randomly sampled  $r^{(l)}$  multiplied element-wise with  $y^{(l)}$ . Then generated  $\tilde{y}^{(l)}$  are fed to the layer as input. The same process is repeated at each layer.

#### V. CASE STUDY

The developed STLF model is verified by comparing with DWT Bi-LSTM LSTM, SWT LSTM, SWT Bi-LSTM model through MATLAB/Simulink simulation using a case study. In this case study, two different cases are considered: 1) Case 1 – load without space heater and 2) Case 1 – load with space heater. For both cases, in order to select the best basis wavelet functions, the mean absolute percentage error (MAPE) of reconstructed signal and actual signal are compared among Haar (haar), Fejer-Korovkin filters (fk), Coiflets (coif), Symlets (sym), Daubechies (db), and Discrete approximation of Meyer (dmey) wavelets as shown in Table II. It is found that Haar wavelet has comparatively less percentage of reconstruction error among other wavelets for both cases. Hence Haar wavelet is used for SWT and ISWT. The Pearson correlation coefficient (PCC) was used in [5] to determine the number of wavelets and recommended three decomposition steps of SWT.

$$MAPE = \frac{1}{n} \times \sum_{i=1}^n \left| \frac{(L_a - L_f)}{L_f} \right| \times 100 \quad (18)$$

TABLE II: MAPE OF RECONSTRUCTED SIGNAL AND ACTUAL SIGNAL FOR DIFFERENT BASIS WAVELET FUNCTION

Case	Wavelet	MAPE(%)
Case 1	Haar	4.5113e-14
	Fejer-Korovkin filters	1.3556e-14
	Coiflets	2.6704e-10
	Symlets	1.5876e-10
	Daubechies	9.4170e-11
	Discrete approximation of Meyer	0.0012
Case 2	Haar	5.2138e-14
	Fejer-Korovkin filters	2.0130e-14
	Coiflets	4.9429e-10
	Symlets	3.3480e-10
	Daubechies	1.8231e-10
	Discrete approximation of Meyer	0.0021

For both cases, during the deep learning training process, one Bi-LSTM layer with 120 hidden units and one LSTM layer with 120 hidden units are designed for this model. Each layer has dropout layer with 0.3 probability. Adam optimizer is used for this case study. Table III shows the used training option for both cases study to train data in the deep learning toolbox.

TABLE III: TRAINING OPTION TO TRAIN DATA IN THE DEEP LEARNING TOOLBOX

Option	Parameter
Max Epochs	150
Gradient Threshold	1
Initial Learn Rate	0.010
Mini Batch Size	30
Learn Rate Drop Period	75
Learn Rate Drop Factor	0.316

Three statistical featured including average, median and L1 norm are determined as tabulated in Table IV for evaluating the forecasting accuracy.

TABLE IV: STATISTICAL FEATURES

Features	Equation
Average	$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i$
Median	$med = \frac{1}{2} (x_{\lfloor \frac{n+1}{2} \rfloor} + x_{\lfloor \frac{n}{2} \rfloor + 1})$
L1 Norm	$\ L\ _1 = \sum_{i=1}^n  x_i $

Where,  $x_i$  is the  $i^{\text{th}}$  sampled measurement point,  $i = 1, 2, \dots, n$  for  $n$  observations.

Three different comparison: 1) Comparison 1 – among SWT\_Bi\_LSTM\_LSTM with dropout layer for a set of lagged power load data vectors (i.e. proposed model) and other developed STLF model Comparison 2 – between proposed model and SWT\_Bi\_LSTM\_LSTM with dropout layer for instantaneous power load data; and 3) Comparison 3 – between proposed model and SWT\_Bi\_LSTM\_LSTM without dropout layer for a set of lagged power load data vectors are studied.

$$\text{Percentage of accuracy} = \left(1 - \left| \frac{\text{Test-Forecast}}{\text{Forecast}} \right| \right) \times 100 \quad (19)$$

A. Comparison 1

TABLE V: STRUCTURE OF OTHER STLF MODEL

Model	Structure
DWT_Bi-LSTM_LSTM	<ul style="list-style-type: none"> <li>• DWT (Haar wavelet, three decomposition steps) decomposition for original data</li> <li>• Bi-LSTM (120 hidden units)</li> <li>• Dropout (0.3 probability)</li> <li>• LSTM (120 hidden units)</li> <li>• Dropout (0.3 probability)</li> <li>• Fully connected layer</li> <li>• Regression layer</li> <li>• IDWT</li> </ul>
SWT_LSTM	<ul style="list-style-type: none"> <li>• SWT (Haar wavelet, three decomposition steps) decomposition for original data</li> <li>• LSTM (120 hidden units)</li> <li>• Dropout (0.3 probability)</li> <li>• Fully connected layer</li> <li>• Regression layer</li> <li>• ISWT</li> </ul>
SWT_Bi-LSTM	<ul style="list-style-type: none"> <li>• SWT (Haar wavelet, three decomposition steps) decomposition for original data</li> <li>• Bi-LSTM (120 hidden units)</li> <li>• Dropout (0.3 probability)</li> <li>• Fully connected layer</li> <li>• Regression layer</li> <li>• ISWT</li> </ul>

In Comparison 1, the proposed model is compared with other developed STLF models DWT\_Bi-LSTM\_LSTM, SWT\_LSTM, SWT\_Bi-LSTM to evaluate the accuracy of proposed model. The structure of the three developed STLM models are listed in Table V. The same neural network

properties are used to train the data for all models. The comparison of the proposed model vs. other STLF models are shown in Table VI and VII for Case 1 and Case 2, respectively. Table VI show that the proposed forecast model is comparatively much more accurate than the other developed models in Case 1. For example, the predicted daily average power consumption accuracy for the proposed model, DWT\_Bi-LSTM\_LSTM, SWT\_LSTM and SWT\_Bi-LSTM are 97.5338%, 89.4967%, 44.4417% and 84.7646%; respectively. For Case 2 as shown in Table VII, the proposed model forecast is also more accurate compared to the other STLF models but with a much higher error than Case 1. Case 2 has less prediction accuracy because Case 2 – 24 hourly daily load profile varies rapidly than Case 1 as shown in Fig. 1 and 2.

B. Comparison 2

In Comparison 1, SWT\_Bi\_LSTM\_LSTM with dropout layer for a set of lagged power load data vectors (i.e. proposed model) and SWT\_Bi\_LSTM\_LSTM with dropout layer for instantaneous power load data are considered. Table VII shows the forecasting result of Comparison 1. The SWT\_Bi\_LSTM\_LSTM with dropout layer for instantaneous power load data model creates significantly higher error compared to proposed model for both cases. For example, for Case 1: the accuracy of daily average power, median and L1 norm with respect to proposed model are 97.5338%, 85.5504% and 97.5411% respectively; and for the other model, the values are 80.6084%, 52.093% and 80.6089% respectively. Therefore, it is concluded from Comparison 1 that considering a set of lagged power load data vectors improves forecasting accuracy.

TABLE VI: COMPARISON AMONG PROPOSED MODEL AND OTHER DEVELOPED STLF MODEL FOR CASE 1

Model	Average power (KW)	Percentage of accuracy for average power (%)	Median	Percentage of accuracy for median (%)	L1 norm	Percentage of accuracy for L1 norm (%)
Test	0.7988		0.6450		19.1712	
Proposed model	0.7791	97.5338	0.7150	89.1473	18.6998	97.5411
DWT_Bi-LSTM_LSTM	0.7149	89.4967	0.7382	85.5504	17.1567	89.4921
SWT_LSTM	1.2426	44.4417	1.2607	95.4574	29.8233	44.4370
SWT_Bi-LSTM	0.9205	84.7646	0.9211	42.8062	22.0923	84.7631

TABLE VII: COMPARISON AMONG PROPOSED MODEL AND OTHER DEVELOPED STLF MODEL FOR CASE 2

Model	Average power (KW)	Percentage of accuracy for average power (%)	Median	Percentage of accuracy for median (%)	L1 norm	Percentage of accuracy for L1 norm (%)
Test	3.5190		2.4825		84.4554	
Proposed model	3.6878	95.2032	2.7958	87.3797	88.5070	95.2027
DWT_Bi-LSTM_LSTM	3.7352	93.8562	3.7355	49.5267	89.6450	93.8552
SWT_LSTM	3.7165	94.3876	3.712	50.4733	89.1951	94.3879
SWT_Bi-LSTM	4.3726	75.7431	4.3741	23.8026	104.9413	75.7435

TABLE VIII: COMPARISON BETWEEN PROPOSED MODEL AND SWT\_Bi\_LSTM\_LSTM WITH DROPOUT LAYER FOR INSTANTANEOUS POWER LOAD DATA FOR BOTH CASES

	Model	Average power (KW)	Percentage of accuracy for average power (%)	Median	Percentage of accuracy for median (%)	L1 norm	Percentage of accuracy for L1 norm (%)
Case 1	Test	0.7988		0.6450		19.1712	
	Proposed model	0.7791	97.5338	0.7382	85.5504	18.6998	97.5411
	SWT_Bi-LSTM_LSTM for instantaneous power load data	0.9537	80.6084	0.9540	52.093	22.8887	80.6089
Case 2	Test	3.5190		2.4825		84.4554	
	Proposed model	3.6878	95.2032	2.7958	87.3797	88.5070	95.2027
	SWT_Bi-LSTM_LSTM for instantaneous power load data	4.4892	72.4297	4.4909	80.9023	107.7406	72.4290

TABLE IX: COMPARISON BETWEEN PROPOSED MODEL AND SWT\_Bi\_LSTM\_LSTM WITHOUT DROPOUT LAYER FOR A SET OF LAGGED POWER LOAD DATA VECTORS FOR BOTH CASES

	Model	Average power (KW)	Percentage of accuracy for average power (%)	Median	Percentage of accuracy for median (%)	L1 norm	Percentage of accuracy for L1 norm (%)
Case 1	Test	0.7988		0.6450		19.1712	
	Proposed model	0.7791	97.5338	0.7150	89.1473	18.6998	97.5411
	SWT_Bi-LSTM_LSTM without dropout layer	0.6692	83.7757	0.6423	99.5796	16.0613	83.7783
Case 2	Test	3.5190		2.4825		84.4554	
	Proposed model	3.6878	95.2032	2.7958	87.3797	88.5070	95.2027
	SWT_Bi-LSTM_LSTM without dropout layer	4.1392	17.6243	4.1162	65.8087	99.3401	82.3757

### A. Comparison 3

In Comparison 3, a comparison between the proposed model and SWT\_Bi\_LSTM\_LSTM without dropout layer for a set of lagged power load data vectors are studied. Comparison 3 is conducted in order to evaluate whether adding a dropout layer after each Bi-LSTM and LSTM layer leads to better prediction accuracy. A comparison between proposed model and SWT\_Bi-LSTM\_LSTM without dropout layer is shown in Table IX. It is found that the proposed model has higher forecasting accuracy for both cases because adding a dropout layer after the Bi-LSTM and LSTM neural network boost the prediction accuracy.

## VI. CONCLUSION

The nature of variable weather conditions and random human behaviour cause randomness in the power consumption profile at the individual household level and create a serious difficulty to improve the short-term power load accuracy. A robust short-term power load forecasting has been developed by using wavelet transform and deep learning method in this paper. The developed model consists of SWT and a Bi-LSTM and LSTM neural network with dropout layers.

In order to evaluate the accuracy of the proposed model, three comparisons for each of two cases are studied by using the real power load dataset of a house in St. John's, Newfoundland and Labrador, Canada. From Comparison 1, it is found that the proposed model has higher forecasting accuracy compared to the other developed STLF model. From Comparison 2, it is found that considering a set of lagged power load data vectors introduced significantly

better forecasting accuracy than the original data input, as this increases the dimensions of the training data. Comparison 3 has been conducted to evaluate the effectiveness of adding dropout layers after the Bi-LSTM and LSTM neural network and it shows that the dropout layer improves the accuracy of the proposed model.

## REFERENCES

- [1] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, 2015.
- [2] F. McLoughlin, A. Duffy, and M. Conlon, "Evaluation of time series techniques to characterise domestic electricity demand," *Energy*, vol. 50, pp. 120–130, 2013.
- [3] M. Aydinalp-Koksal and V. I. Ugursal, "Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector," *Applied Energy*, vol. 85, no. 4, pp. 271–296, 2008.
- [4] M. Imani and H. Ghassemian, "Residential load forecasting using wavelet and collaborative representation transforms," *Applied Energy*, vol. 253, p. 113505, 2019.
- [5] K. Yan, W. Li, Z. Ji, M. Qi, and Y. Du, "A Hybrid LSTM Neural Network for Energy Consumption Forecasting of Individual Households," *IEEE Access*, vol. 7, pp. 157633–157642, 2019.
- [6] M. Aydinalp, V. I. Ugursal, and A. S. Fung, "Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks," *Applied Energy*, vol. 79, no. 2, pp. 159–178, 2004.
- [7] Ringwood JV, Bofell D, Murray FT, "Forecasting electricity demand on short, medium and long time scales using neural networks," *Journal of Intelligent and Robotic Syst* 2001;31:129–47.
- [8] Z. W. Geem and W. E. Roper, "Energy demand estimation of South Korea using artificial neural network," *Energy Policy*, vol. 37, no. 10, pp. 4049–4054, 2009.
- [9] A. Baliyan, K. Gaurav, and S. K. Mishra, "A Review of Short Term Load Forecasting using Artificial Neural Network Models," *Procedia Computer Science*, vol. 48, pp. 121–125, 2015.

- [10] A. Ghanbari, S. Abbasian-Naghneh, and E. Hadavandi, "An intelligent load forecasting expert system by integration of ant colony optimization, genetic algorithms and fuzzy logic," *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, 2011.
- [11] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, "Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches †," *Energies*, vol. 11, no. 7, p. 1636, 2018.
- [12] N. Amjady and F. Keynia, "Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm," *Energy*, vol. 34, no. 1, pp. 46–57, 2009.
- [13] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network," *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, 2017.
- [14] A. K. Fard and M.-R. Akbari-Zadeh, "A hybrid method based on wavelet, ANN and ARIMA model for short-term load forecasting," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 26, no. 2, pp. 167–182, 2013.
- [15] J. Kim, J. Moon, E. Hwang, and P. Kang, "Recurrent inception convolution neural network for multi short-term load forecasting," *Energy and Buildings*, vol. 194, pp. 328–341, 2019.
- [16] Y. Liu, L. Guan, C. Hou, H. Han, Z. Liu, Y. Sun, and M. Zheng, "Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform," *Applied Sciences*, vol. 9, no. 6, p. 1108, 2019.
- [17] H. Su, E. Zio, J. Zhang, M. Xu, X. Li, and Z. Zhang, "A hybrid hourly natural gas demand forecasting method based on the integration of wavelet transform and enhanced Deep-RNN model," *Energy*, vol. 178, pp. 585–597, 2019.
- [18] C. Keerthisinghe, G. Verbic, and A. C. Chapman, "A Fast Technique for Smart Home Management: ADP With Temporal Difference Learning," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3291–3303, 2018.
- [19] Y. Wang, D. Gan, M. Sun, N. Zhang, Z. Lu, and C. Kang, "Probabilistic individual load forecasting using pinball loss guided LSTM," *Applied Energy*, vol. 235, pp. 10–20, 2019.
- [20] B. Yildiz, J. I. Bilbao, J. Dore, and A. B. Sproul, "Short-term forecasting of individual household electricity loads with investigating impact of data resolution and forecast horizon," *Renewable Energy and Environmental Sustainability*, vol. 3, p. 3, 2018.
- [21] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, "Short-Term Residential Load Forecasting Based on Resident Behaviour Learning," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1087–1088, 2018.
- [22] M. Imani and H. Ghassemian, "Electrical Load Forecasting Using Customers Clustering and Smart Meters in Internet of Things," *2018 9th International Symposium on Telecommunications (IST)*, 2018.
- [23] P. Guo, J. C. Lam, and V. O. Li, "Drivers of domestic electricity users' price responsiveness: A novel machine learning approach," *Applied Energy*, vol. 235, pp. 900–913, 2019.
- [24] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019.
- [25] "Energy Efficiency Trends in Canada 1990 to 2013," *Natural Resources Canada*, 31-Jan-2019. [Online]. Available: <https://www.nrcan.gc.ca/energy/publications/19030>. [Accessed: 15-Apr-2020].
- [26] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, "Wavelet Decomposition and Convolutional LSTM Networks Based Improved Deep Learning Model for Solar Irradiance Forecasting," *Applied Sciences*, vol. 8, no. 8, p. 1286, Jan. 2018.
- [27] Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*. Vol. 15, pp. 1929-1958, 2014.
- [28] Yarin, Ghahramani, and Zoubin, "A Theoretically Grounded Application of Dropout in Recurrent Neural Networks," *arXiv.org*, 05-Oct-2016. [Online]. Available: <https://arxiv.org/abs/1512.05287>. [Accessed: 15-Apr-2020].
- [29] A. Graves and J. Schmidhuber, "Frame-wise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, 19-Aug-2005. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0893608005001206>. [Accessed: 15-Apr-2020].
- [30] A. Graves, N. Jaitly, and A.-R. Mohamed, "Hybrid speech recognition with Deep Bidirectional LSTM," *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, 2013.
- [31] S. Wang, X. Wang, S. Wang, and D. Wang, "Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 109, pp. 470–479, 2019.
- [32] J. Kim and N. Moon, "BiLSTM model based on multivariate time series data in multiple field for forecasting trading area," *Journal of Ambient Intelligence and Humanized Computing*, 2019.
- [33] K. B. Sahay, S. Sahu, and P. Singh, "Short-term load forecasting of Toronto Canada by using different ANN algorithms," *2016 IEEE 6th International Conference on Power Systems (ICPS)*, 2016.



#### Hla U May Marma

She was born in Bangladesh, 1992, 3<sup>rd</sup> January.

She received her B.Sc degree from Bangladesh University of Engineering and Technology (BUET), Bangladesh in 2015 and presently enrolled in the Master of Electrical Engineering program at Memorial University of Newfoundland. Her research is based on load modeling of power systems, and load forecasting.

She has worked for two years in Stichting Bangladesh Accord Foundation as Electrical Safety Engineer.



#### M. Tariq Iqbal

He completed his B.Sc.(EE) at University of Engineering and Technology, Lahore in 1986, his M.Sc. Nuclear Engineering at CNS, Quaid-e-Azam University, Islamabad in 1988 and his Ph.D. in Electrical Engineering at Imperial College London in 1994. From 1988 to 1991 and again from 1995 to 1999, he worked at Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad. From 1999 to 2000, he worked as an Associate Professor at Riphah International University. Since 2001, he is working at Faculty of Engineering and Applied Science, Memorial University of Newfoundland. Currently, his research focuses on modeling and control of hybrid renewable energy systems and he has strong interest in consumer electronics and instrumentation



#### Christopher Thomas Seary

He was born in St. John's, Canada, 1991, 26<sup>th</sup> July.

He received the B. Eng. degree in electrical engineering from Memorial University of Newfoundland, Canada, in 2014. He is currently pursuing an M. Eng. degree from same. His masters research is focused on multi-megahertz resonant converters for wireless power

applications using capacitive links. His hobbies include woodworking and cycling