International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 11 Nov 2023, pp: 76-84 www.ijaem.net ISSN: 2395-5252

A hybrid model of ILSTM, SVM and GPC with signal decomposition by SD for short-term electricity load forecasting in the electricity market

Mesbaholdin Salami

Date of Submission: 10-11-2023 Date of Acceptance: 20-11-2023

ABSTRACT
The demand for accurate short-term load forecasting (STLF) models has increased with the rapid growth of power systems and the integration of renewable energy sources. STLF plays a crucial role in optimizing grid operations, ensuring grid stability, and promoting energy efficiency. Various techniques have been proposed to improve the accuracy of STLF models, including the application of advanced optimization algorithms and machine learning algorithms. In this article, we provide a hybrid model of the latest advancements in STLF, focusing on the utilization of Spectral density (SD), Improved Long Short-Term Memory (ILSTM), Giza Pyramids Construction (GPC) and Support Vector Machine (SVM) algorithms.
Keywords : STLF,SD, ILSTM,GPC,SVM

I. INTRODUCTION

Energy planning is one of the important issues in the field of energy supply. The limited energy resources and its high price have caused the supply chain managers to be more accurate in estimating the parameters of this supply chain. This chain starts from the power plants and finally reaches the final household or industrial electricity consumer by transferring and distributing it. One of the most important parameters of supply chain planning is the correct forecasting of the amount of electrical load consumption in different time frames. Consumption in the daily time frame is one of the most important time frames for the manager of the electricity supply chain, according to which manages this chain. Therefore, short-term electricity consumption is one of the attractive fields for many articles to achieve a model that has the best performance in forecasting electricity consumption in the short term. Among them are the ASTLF model by Bellahsen et al. [1], SMN by Wu et al. [2], FTS-BP by Li et al. [3], TgDLF by Chen et al. [4], DDM by Wang and colleagues [5], AGRU by Wang et al. [6], XGB-LGBM-MLP by Masoudi et al. [7], LSTM-NN by Memarzadeh et al. [8]. hybrid models have also been presented, which have shown higher accuracy in forecasting electricity consumption in past research. Among them are SARIMA-LSTM [9], NN-LSTM [10], SAM-LSTM [11], WT-NNPSO-SO [12], Complete Linkage [13], SUBHSO [14], GALSTM [15], EMD-WOA [16] pointed out. In this article, the trend of hybrid models is followed, and from the combination of methods, the data is first analyzed by theSpectral density(SD) method, and then the high-frequency data is entered into the improved long short-term memory(ILSTM) model, whose weights are determined by the Giza Pyramids Construction (GPC) metaheuristic method, and the signals are analyzed. It will enter the support vector machines (SVM) model with low frequency. Then the output results of these two models are summed together to obtain the forecasting values. Finally, using performance measurement criteria, the model will be compared with other existing models for short-term load forecasting. This article includes 2. methods 3. Measurement criteria 3. Results and 4. Conclusion summary.

II. METHODS :

1.1. the Spectral density (SD)

Spectral density describes the distribution of power to the frequency components of the signal. According to Fourier analysis, any physical signal can be decomposed into a number of discrete frequencies, or a range of frequencies in a continuous range. The statistical average of a given signal or a type of signal (including noise) when analyzed from the point of view of its frequency content, is called the spectrum of that signal.

When the signal energy is concentrated around a limited time interval, especially if its total energy is limited, its energy spectral density can be calculated.[17]

The energy spectrum density describes the way in which the energy of a signal or a time series is distributed along the frequency. Here, the term energy is used in its general sense in signal processingthat is, the energy E of a signal x(t) is:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (1)$$

The energy spectral density is well suited for transients—that is, pulse-like signals, which then have finite total energy. Whether bounded or not, Parsvall's theorem (or Plancherl's theorem) gives us an alternative expression for the signal energy:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{x}(f)|^2 df \quad (2)$$

where in:

$$\hat{x}(f) = \int_{-\infty}^{\infty} e^{i2\pi ft} x(t) dt \quad (3)$$

The value of the Fourier transform x(t) (at frequency f) is in Hertz. This theorem is also true in discrete-time modes. Because the integral on the right is the energy of the signal, the sub integral term$|\hat{x}(f)|^2$ can be interpreted the form of a density function, which describes the energy contained in the signal at frequency f. Hence, the spectral energy density for x(t) is defined as:

$$\acute{s}_{xx}(f) = |\hat{x}(f)|^2 \quad (4)$$

The function $\acute{s}_{xx}(f)$and the autocorrelation for $x(t)$ form a Fourier transform pair. This definition is directly generalized to a discrete signal with a countably infinite number of countersx_n , such as a signal that at discrete times $x_n = x_0 + (n\, \Delta t)$has been sampled:

$$\acute{s}_{xx}(f) = \lim_{N\to\infty} (\Delta T)^2 \left| \sum_{n=-N}^{N} x_n e^{i2\pi ftn\Delta T} \right|^2 \quad (5)$$

Here $\left|\sum_{n=-N}^{N} x_n e^{i2\pi ftn\Delta T}\right|^2$ is the discrete-time Fourier transform for x_n. A sampling interval of ΔT is required to maintain the correct physical units, and also to ensure that we recover the continuous state in the limit ΔT→0. But in mathematical sciences this interval is usually set to 1, which simplifies the results at the cost of generalization.

2.2. ILSTM

LSTM is a special type of RNNs capable of learning long-term dependencies. They were introduced by Hochreiter & Schmid Huber (1997), and modified by many in subsequent studies. LSTMs are generally designed to avoid the long-term dependence problem. LSTM network is an almost old network. But it is used in a large number of problems and is still very popular among researchers. According to the diagram below; You can see that over time, the use of LSTM has grown tremendously. [18]

Memorizing information for long periods of time is practically their default behavior, not something they are trying to learn! All recurrent neural networks are in the form of chains of recurrent neural network modules. In standard RNNs, this recurrent module has a very simple structure, like a single layer.

LSTMs also have a chain-like structure, and the iterative module has a different structure. Instead of having a single neural network layer, there are four layers that interact with each other in a very special way.

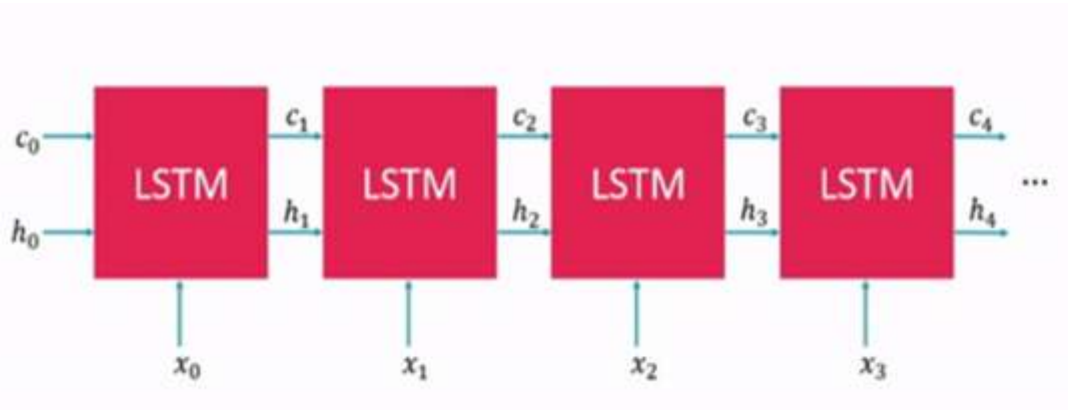

Figure 1. Schematic model of LSTM method

The first step in our LSTM is to decide what information to extract from the cell state. This decision is made by a sigmoid layer called "forget gate layer.

In and it goes to the sigmoid function and outputs a number between 0 and 1 for each number in the cell state. 1 indicates "keep this" while 0 indicates "delete this.

$$Z_t = g\,(W_z x_t + R_z y_{t-1} + b_z) \qquad (6)$$

The next step is to decide what new information to store in the cell state. It has two parts. First, a sigmoid layer called the "input gate layer" decides which values to update. Next, a tanh layer creates a vector of newly selected values, C ~ t, that can be added to the state. Next, the two are combined to create an update for the status.

$$i_t = \sigma\,(W_i x_t + R_i y_{t-1} + p_i \odot c_{t-1} + b_z)(7)$$
$$i_t = \sigma\,(W_i x_t + R_i y_{t-1} + p_i \odot c_{t-1} + b_z)(8)$$

Now it's time for the network to update the old cell mode, , to the new cell mode. It breaks the legacy mode and things that the network previously decided to remove will now be removed. Then we add it C ~ t. These new values are chosen, which scales with the decision to update each state value.

$$c_t = z_t \odot i_t + c_{t-1} \odot f_t(9)$$
$$f_t = \sigma\,(W_f x_t + R_f y_{t-1} + p_f \odot c_{t-1} + b_f)(10)$$

Ultimately, the network must decide what to produce. This output is based on the state of the cell, but will be a filtered version. First, a sigmoid layer is implemented that decides what parts of the cell state to generate. Then, it sets the cell state to tanh (so that the values are between −1 and 1 and multiplies it to the output of the sigmoid gate, so that only the parts it decides to be output are considered as output. Takes.

$$y_t + h(c_t) \odot o_t \qquad (11)$$
$$o_t = \sigma\,(W_o x_t + R_o y_{t-1} + p_o \odot c_t + b_o)(12)$$

What I have described so far is a typical LSTM. But not all LSTMs are like the above. In fact, almost every paper that includes LSTMs seems to use a slightly different version of the original. The differences are minor, but some of them are mentioned. A popular variant of LSTM, introduced by Gers &Schmid Huber (2000), is the addition of peephole connections. This means that we let the gate layers look at the state of the cell.

Another model of input forgetting in LSTM is that instead of deciding separately what to forget and what to add, we make these decisions together. The network forgets to put something in its place. The network only inserts new values into the state cell when it forgets something older.

The figure below is a summary of the LSTM network. You don't have to worry too much about complicated

LSTM formulas because that's all the logic behind forgetting gates and information retention.

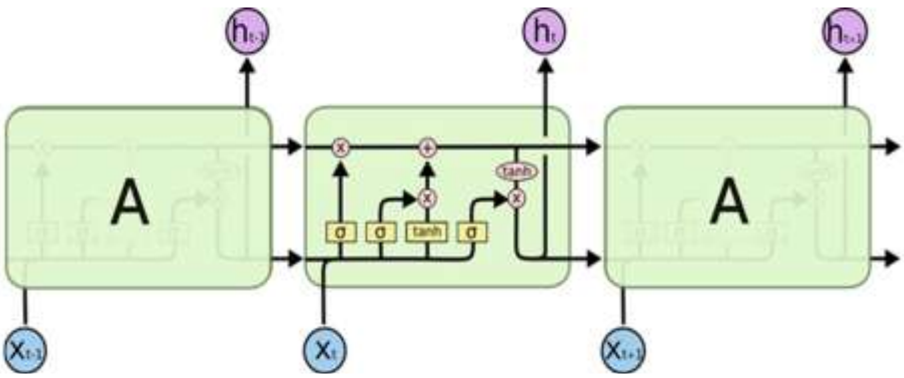

Figure 2. How LSTM network layers work

2.3.SVM

Support vector machine is one of the supervised learning methods used for classification and regression. The working basis of the SVM classifier is the linear classification of the data, and in the linear division of the data, we try to choose a hyperplane that has a higher confidence margin. Solving the equation of finding the optimal line for the data is done by linear programming methods, which are well-known methods in solving constrained problems. Before the linear division, in order for the machine to be able to classify the data with high complexity, we move the data to a much higher dimensional space using the phi function

[4]. In order to be able to solve the problem of very high dimensions using these methods, from Lagrange's dual theorem [5] to transform the desired minimization problem into its dual form, where instead of the complex function phi that takes us to a high-dimensional space, we use a simpler function called the kernel function, which is the vector multiplication of the phi function. Different kernel functions can be used, including exponential, polynomial and sigmoid kernels.[19]

We have the test data set D consisting of n members (points) defined as follows:

$$D = \{(x_i, y_i) \mid x_i \in R^p, y_i \in \{-1,1\}_{i=1}^{n}\} \quad (13)$$

where the value of y is equal to 1 or -1 and each x_i is a p-dimensional real vector. The goal is to find the separating hyperplane with the largest distance from the edge points that separates the points with y_i=1 from the points with y_i=-1. Any hyperplane can be written as a set of points satisfying the following condition:

$$w.x - b = 0 \quad (14)$$

where is the multiplication sign. w is the normal vector, which is perpendicular to the hyperplane. We want to choose w and b to maximize the distance between the parallel super planes that separate the data. These hyperplanes are described using the following relation.

Any data above the separating hyperplane is marked with label 1:

$$w.x - b = 1 \quad (15)$$

and any data at the bottom of the separating hyperplane with the label -1:

$$w.x - b = -1 \quad (16)$$

If the training data is linearly separable, we can consider two hyperplanes on the edge of points such that they have no points in common, and then try to maximize their distance. Using geometry, the distance between these two planes is $2/\|w\|$; So, we need to minimize ‖w‖. To prevent points from entering the margin, we add the following condition: for each can be written as:

$$w.x - b = 1, \quad if\ y_i = 1 \quad (17)$$

$$w.x - b = -1, if\ y_i = -1 \quad (18)$$
$$y_i(w.x - b) \geq 1, \quad \forall 1 \leq i \leq n \quad (19)$$

Putting these two together yields an optimization problem:

$$min\ 0.5\ \|w\|^2 \quad (20)$$
$$\forall 1 \leq i \leq n$$
$$y_i(w.x - b) \geq 1$$

2.4.Giza pyramids construction algorithm

Suppose that stone blocks are scattered around the construction site and the workers have to push the stone blocks to the installation site. The initial position of each block and its cost must be calculated first. In the second step, a ramp is used to transport the stone blocks to the installation site. The slope of the ramp and its friction affect the movement of stone blocks. So the next measurable factors were determined. Pay attention to the figure below : [21]

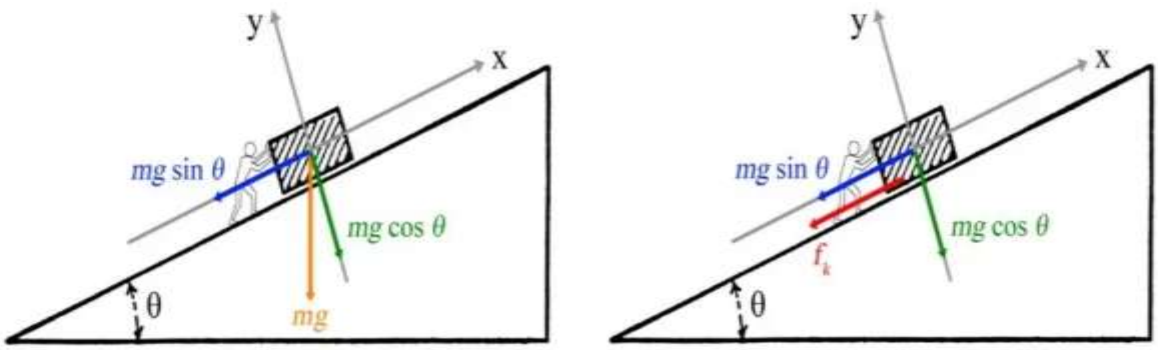

Figure 3. Ramp slope and its friction on displacement in GPC model

In the third step of the algorithm, we check that the workers are constantly moving to find the best position to dominate the stone block. According to the different characteristics of each worker, it is possible to replace the workers to balance the strength of the workers to carry the stone block. Therefore, the position of some workers will be replaced by others. This replacement changes the stone block movement system and power balance. The following algorithm describes the pseudocode of the GPC algorithm.

Due to the nature of the algorithm, each stone block moves up the ramp with an initial velocity v0. Therefore, the block stops after traveling a distance d on the ramp. f_k is the kinetic friction force, and since the rock block is on the verge of displacement, f_k can be obtained from the following equation.

$$f_k = \mu_k mg \cos\theta \quad (21)$$

where m is the mass of the stone block, g is the earth's gravity, and is the coefficient of the angle that makes the inclined surface with the horizon. Because according to Newton's second law, we are on the x-axis:

$$\sum \vec{F} = m\vec{a} \quad (22)$$
$$-mg\sin\theta - f_k = ma$$

Next, by inserting equation 1 into equation 2, the upward acceleration of the stone block on the inclined surface is obtained. So, here we need a time-independent equation of motion under constant acceleration, which can be used to calculate the displacement of a stone block on an inclined surface using the following equation.

$$a = -g(\sin\theta + \mu_k \cos\theta\,) \quad (23)$$
$$d = \frac{v_0^2}{2g(\sin\theta + \mu_k\cos\theta\,)} \quad (24)$$

v0 is the initial velocity of the stone block and is determined in the algorithm by a uniformly distributed random number in each iteration. In this way, if the worker applies force to a stone block; The block of stone starts moving with an initial speed and according to physical science, the force of friction causes the block of stone to stop after some time. So, the worker again applies another force on the block of stone so that the block of stone starts moving again at the initial speed. In each iteration of the algorithm, the initial speed is considered a random number. Because every time the worker tries to move the stone block, the applied force changes according to the worker's consumption power. Therefore, for v0 we have:

$$v_0 = rand\,(0{,}1) \quad (25)$$
$$0 < v_0 = rand\,(0{,}1) < 1$$

In this step of the tutorial, we will need a kinetic coefficient of friction between the stone block and the inclined surface, which is determined in the algorithm by a uniformly distributed random number.

$$\mu_k = rand\,[\mu_{k\_min}, \mu_{k\_max}] \quad (26)$$
$$\mu_{k\ min} \le \mu_k \le \mu_{k\ max}$$

The main idea of the algorithm is that the workers pushing the stone block are constantly moving or shaking to gain the best grip and control over the stone block. These blows make the worker perform non-repetitive movements to better push the stone block. Friction is not considered for the worker. Therefore, the new position of the worker who pushes the stone block is obtained from the following equation:

$$x = \frac{v_0^2}{2g\sin\theta} \quad (27)$$

Other algorithm formulas are used in the same way.

$$\vec{p} = (\vec{p}_i + d)x\vec{\epsilon}_i \quad (28)$$
$$\Phi = (\rho_1, \rho_2, \dots, \rho_n) \quad (29)$$
$$\Psi = (\Psi_1, \Psi_2, \dots, \Psi_n) \quad (30)$$
$$Z = (\zeta_1, \zeta_2, \dots, \zeta_n) \quad (31)$$
$$\zeta_k = \begin{cases} \Psi_k \ if\ rand\,[0{,}1] \le 0.5 \\ \rho_k \ otherwise \end{cases} \quad (32)$$

2.5.Hybrid model:

The proposed model of this article is a combination of models based on artificial intelligence, which consists of the following components and steps:

1- Historical data is analyzed by the SD method. (x_n)

2- Then high frequency data is analyzed by ILSM method.

3- The low frequency data is analyzed by the SVM method whose weights are optimized by the GPC artificial intelligence-based method.

4- Then the results of steps 1 and 2 are added together to get the predicted amount of electric charge.

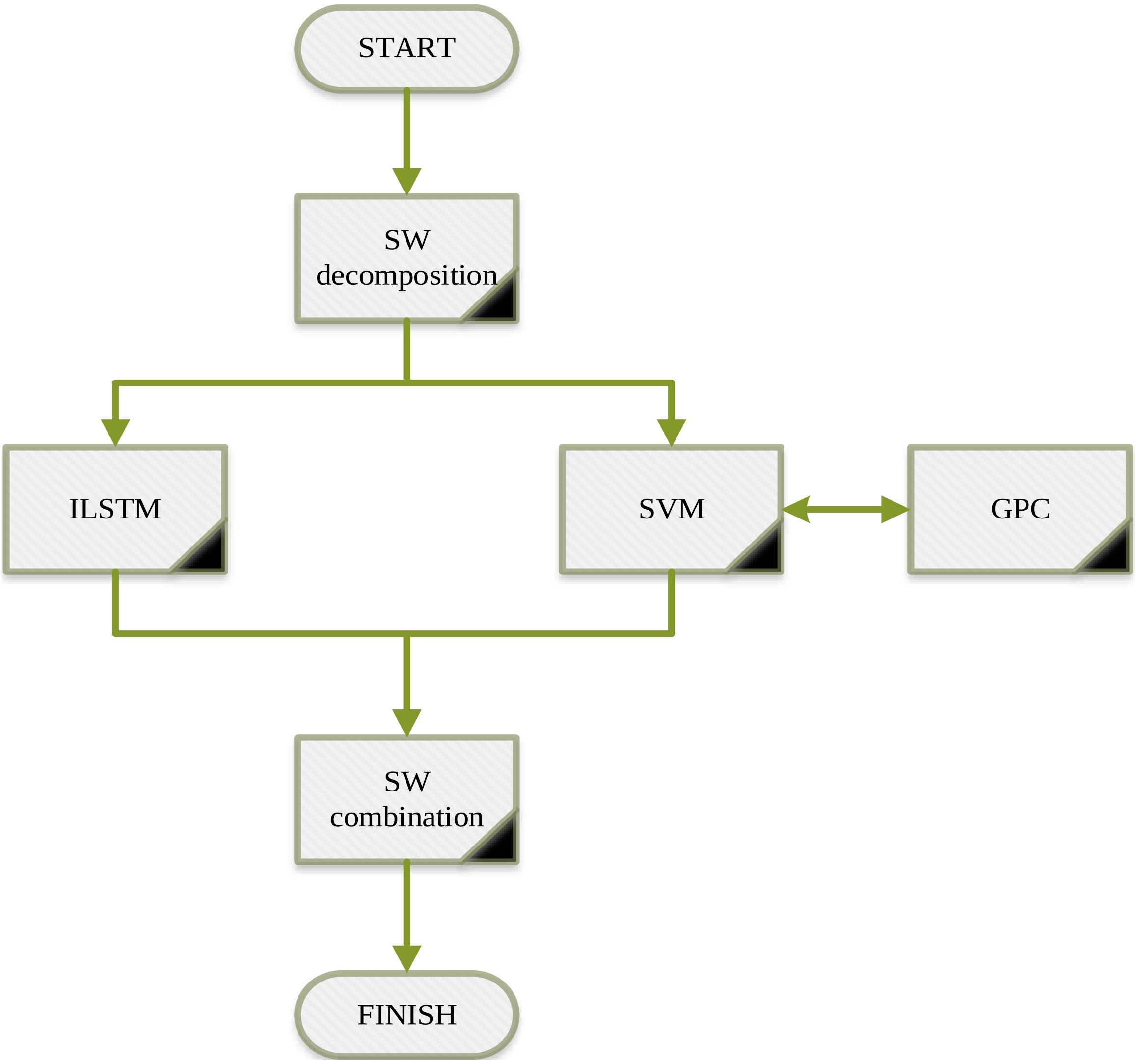

Figure 4. Proposedmodel components

III. MEASUREMENT CRITERIA

To compare the performance of the target model with other models and to evaluate the improvement in predictions among the methods proposed in reference [21] of the Mean Absolute Percent Error (MAPE) method in equation 45, Mean Absolute Error (MAE) in Equation 46 and Root Mean Square Error (RMSE) are used in Equation 47. These relationships are used to compare the performance of the target model and the existing models in Section 2.3.

$$\mathrm{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\vartheta_t - \hat{\vartheta}_t}{\vartheta_t}\right| \quad (33)$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|\vartheta_t - \hat{\vartheta}_t\right| \quad (34)$$

$$\mathrm{RMSE} = \frac{1}{N}\sqrt{\sum_{t=1}^{N}\left(\vartheta_t - \hat{\vartheta}_t\right)^2} \quad (35)$$

IV. RESULTS

To check the prediction performance results of the target model and the existing models, the data of Iran's electricity market, which is available at the address of www.igmc.ir has been used. The data is selected for the year 2020. Actual and predicted values by the models under investigation are displayed in Figure 5.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 11 Nov 2023, pp: 76-84 www.ijaem.net ISSN: 2395-5252

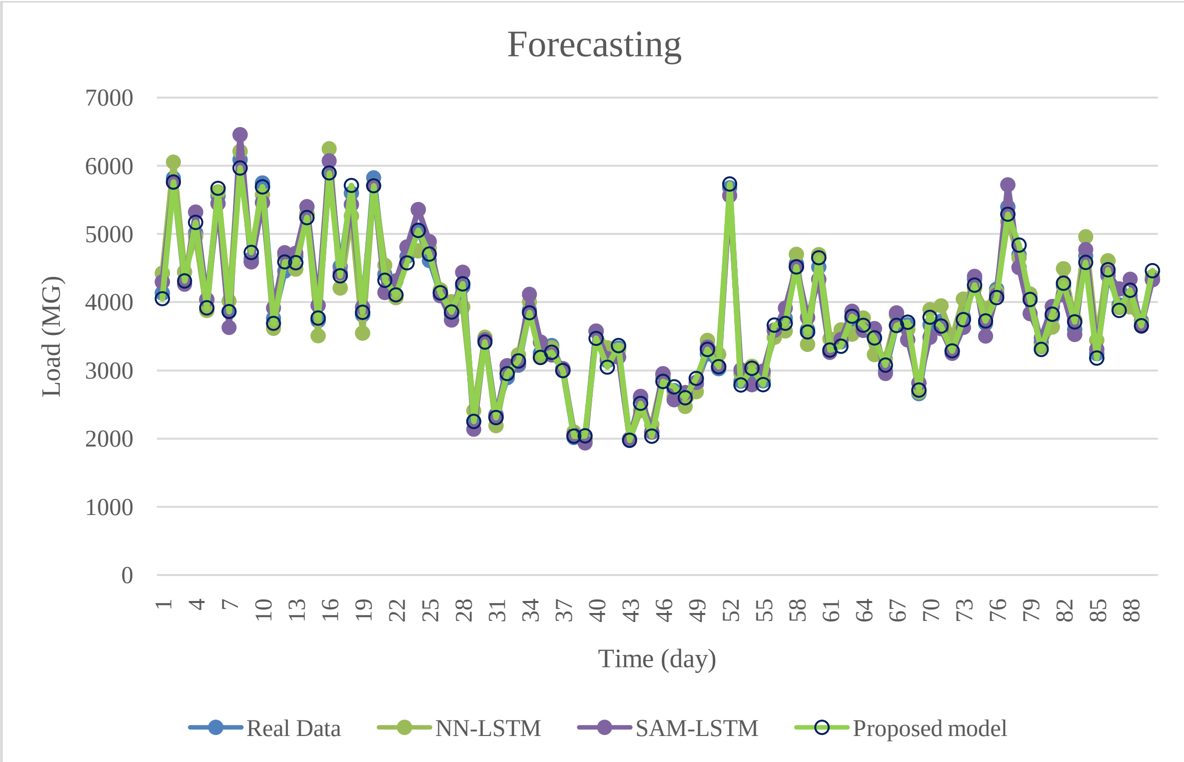

Figure 5. The performance of the target model and existing models in short-term load forecasting

The results of the measurement error resulting from the daily forecast of electricity consumption based on the measurement criteria defined in section 3 are displayed in table number 1.

Table 1. Performance of target model and existing models based on measurement criteria

Model	Criterion	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Average
SARIMA-LSTM	MAPE (%)	0.099	0.096	0.096	0.107	0.100	0.098	0.092	0.102	0.096	0.101	0.096	0.099	0.098
	MAE(MW)	326.71	346.57	329.45	335.88	368.81	347.83	326.88	307.63	354.21	327.16	353.98	325.72	337.57
	RMSE	40.12	43.61	41.31	41.98	44.49	44.74	40.05	38.23	43.38	40.39	43.88	40.51	41.89
NN-LSTM	MAPE (%)	0.074	0.079	0.075	0.082	0.087	0.086	0.072	0.070	0.078	0.080	0.078	0.074	0.078
	MAE(MW)	284.46	284.43	303.93	286.30	316.48	339.61	333.27	270.79	266.57	295.62	300.93	295.07	298.12
	RMSE	36.24	35.95	37.47	35.68	38.80	40.73	40.83	33.09	33.87	35.94	35.90	35.79	36.69
SAM-LSTM	MAPE (%)	0.042	0.045	0.043	0.042	0.044	0.041	0.044	0.045	0.045	0.037	0.042	0.042	0.043
	MAE(MW)	118.69	133.17	124.61	121.54	129.83	117.82	126.26	132.61	133.83	103.10	125.95	122.63	124.17
	RMSE	14.68	16.12	15.42	15.11	15.51	14.79	15.45	16.24	16.32	13.13	15.62	15.23	15.30
Proposed model	MAPE (%)	0.0298	0.0304	0.0289	0.0284	0.0286	0.0298	0.0284	0.0288	0.0271	0.0288	0.0288	0.0301	0.029
	MAE(MW)	48.85	51.12	44.65	44.29	44.20	48.85	43.65	45.41	37.73	45.63	45.64	49.50	45.79
	RMSE	6.00	6.34	5.74	5.76	5.78	6.01	5.60	5.81	5.20	5.76	5.86	6.13	5.83

The results show that the target model has improved prediction error by 0.069, NN-LSTM model by 0.049, and SAM-LSTM model by 0.014, .based on MAPE measurement criteria

V. CONCLUSION

Accurately predicting the amount of load consumed in the electricity market helps buyers and sellers of electricity along the supply chain to make their decisions based on more accurate values and, as a result, to avoid excess load in the network or lack of load in the network. be prevented. In this article, the development process of electric load forecasting articles was followed and a combined model of load forecasting for a period of one year was formed based on the data of the Iranian electricity market. The results show an improvement in forecasts compared to other short forecasting models. It is the load period.

REFERENCES

[1]. Amine Bellahsen, Hanane Dagdougui, Aggregated short-term load forecasting for heterogeneous buildings using machine learning with peak estimation, Energy and Buildings, Volume 237, 2021, 110742, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2021.110742.

[2]. Xuedong Wu, Yaonan Wang, Yingjie Bai, Zhiyu Zhu, Aiming Xia,Online short-term load forecasting methods using hybrids of single multiplicative neuron model, particle swarm optimization variants and nonlinear filters, Energy Reports, Volume 7, 2021, Pages 683-692, ISSN 2352-4847, https://doi.org/10.1016/j.egyr.2021.01.030.

[3]. Chen Li, A fuzzy theory-based machine learning method for workdays and weekends short-term load forecasting, Energy and Buildings, 2021, 111072, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2021.111072.

[4]. Yuntian Chen, Dongxiao Zhang, Theory-guided deep-learning for electrical load forecasting (TgDLF) via ensemble long short-term memory, Advances in Applied Energy, Volume 1, 2021, 100004, ISSN 2666-7924, https://doi.org/10.1016/j.adapen.2020.100004.

[5]. Yuanyuan Wang, Shanfeng Sun, Xiaoqiao Chen, Xiangjun Zeng, Yang Kong, Jun Chen, Yongsheng Guo, Tingyuan Wang, Short-term load forecasting of industrial customers based on SVMD and XGBoost, International Journal of Electrical Power & Energy Systems, Volume 129, 2021, 106830, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2021.106830.

[6]. Shouxiang Wang, Xinyu Deng, Haiwen Chen, Qingyuan Shi, Di Xu, A bottom-up short-term residential load forecasting approach based on appliance characteristic analysis and multi-task learning, Electric Power Systems Research, Volume 196, 2021, 107233, ISSN 0378-7796, https://doi.org/10.1016/j.epsr.2021.107233.

[7]. Mohamed Massaoudi, Shady S. Refaat, Ines Chihi, Mohamed Trabelsi, Fakhreddine S. Oueslati, Haitham Abu-Rub, A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for Short-Term Load Forecasting, Energy, Volume 214, 2021, 118874, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2020.118874.

[8]. Gholamreza Memarzadeh, Farshid Keynia, Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm, Electric Power Systems Research, Volume 192, 2021, 106995, ISSN 0378-7796, https://doi.org/10.1016/j.epsr.2020.106995

[9]. Ashutosh Kumar Dubey, Abhishek Kumar, Vicente García-Díaz, Arpit Kumar Sharma, Kishan Kanhaiya, Study and analysis of SARIMA and LSTM in forecasting time series data, Sustainable Energy Technologies and Assessments, Volume 47, 2021, 101474, ISSN 2213-1388, https://doi.org/10.1016/j.seta.2021.101474.

[10]. Gholamreza Memarzadeh, Farshid Keynia, Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm, Electric Power Systems Research, Volume 192, 2021, 106995, ISSN 0378-7796, https://doi.org/10.1016/j.epsr.2020.106995.

[11]. Haixiang Zang, Ruiqi Xu, Lilin Cheng, Tao Ding, Ling Liu, Zhinong Wei, Guoqiang Sun, Residential load forecasting based on LSTM fusing self-attention mechanism with pooling,

Energy, Volume 229, 2021, 120682, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2021.120682.

[12]. Salami, M.; Movahedi Sobhani, F.; Ghazizadeh, M.S. Short-Term Forecasting of Electricity Supply and Demand by Using the Wavelet-PSO-NNs-SO Technique for Searching in Big Data of Iran's Electricity Market. Data **2018**, 3, 43. https://doi.org/10.3390/data3040043

[13]. M. Salami, F. Movahedi Sobhani and M. Sadegh Ghazizadeh, "Evaluating Power Consumption Model and Load Deficit at Different Temperatures Using Clustering Techniques and Presenting a Strategy for Changing Production Management," 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 2020, pp. 341-345, doi: 10.1109/DSMP47368.2020.9204247.

[14]. Salami, Mesbaholdin, Sobhani, Farzad Movahedi and Ghazizadeh, Mohammad Sadegh. "Shared Subscribe Hyper Simulation Optimization (SUBHSO) Algorithm for Clustering Big Data – Using Big Databases of Iran Electricity Market" Applied Computer Systems, vol.24, no.1, 2019, pp.49-60. https://doi.org/10.2478/acss-2019-0007

[15]. M. Salami, M. R. Niakan and M. Hasani, "Electricity Load Forecasting: A New Hybrid Model Combining Wavelet Transform, Genetic Algorithm and LSTM Models," 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Mauritius, Mauritius, 2021, pp. 1-6, doi: 10.1109/ICECCME52200.2021.9591091

[16]. Salami, Mesbaholdin, Sobhani, Farzad Movahedi and Ghazizadeh, Mohammad Sadegh , Multi-Objective Model for Fair Pricing of Electricity Using the Parameters from the Iran Electricity Market Big Data Analysis, International Journal of Industrial Mathematics, 2018, 10,4,359-374,https://ijim.srbiau.ac.ir/article_13242.html

[17]. Fred Rieke; William Bialek & David Warland (1999). Spikes: Exploring the Neural Code (Computational Neuroscience). MIT Press. ISBN 978-0-262-68108-7.

[18]. Wang, J.; Zhang, T.; Lu, T.; Xue, Z. A Hybrid Forecast Model of EEMD-CNN-ILSTM for Crude Oil Futures Price. Electronics **2023**, 12, 2521. https://doi.org/10.3390/electronics12112521

[19]. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf, "Support vector machines," in IEEE Intelligent Systems and their Applications, vol. 13, no. 4, pp. 18-28, July-Aug. 1998, doi: 10.1109/5254.708428.

[20]. Harifi, S., Mohammadzadeh, J., Khalilian, M. et al. Giza Pyramids Construction: an ancient-inspired metaheuristic algorithm for optimization. Evol. Intel. **14**, 1743–1761 (2021). https://doi.org/10.1007/s12065-020-00451-3

[21]. C.J. Willmott, Some comments on the evaluation of model performance, Bull. Am. Meteorol. Soc. 63 (1982) 1309–1313.