Algerian Journal of Environmental Science and Technology
March edition. Vol.9. Nº1. (2023)
ISSN : 2437-1114
www.aljest.org

Load Energy Forecasting based on a Hybrid PSO LSTM-AE Model

A. Saoud[1], A. Recioui*

[1]Signals and Systems Research Laboratory (LSS), Institute of Electrical and Electronic Engineering, University M'hamed Bougara of Boumerdes, 35000 Boumerdes–Algeria

*Corresponding author: a_recioui@univ-boumerdes.dz; Tel.: +213247955 73; Fax: +213 248183 33

ARTICLE INFO

Article History :

Received : 03/12/2020
Accepted : 23/05/2021

***Key Words**:*

Long short-term Memory(LSTM);
Autoencoder;
Particle swarm optimization;
Hyperparameters selection;
Smart meter data;
Energy forecasting.

ABSTRACT/RESUME

Abstract:*In the smart grid, the data collected from the smart meters can be used to develop an accurate energy consumption forecasting. A close prediction is beneficial in providing a good energy scheduling, making balance between demand and generation of power which results in reducing the production costs of the energy. Several models are used to give an accurate energy consumption forecast. One of these models is long short-term memory (LSTM). LSTM model may be combined with other models to give better results. On the other hand, LSTM has a drawback of selecting the hyperparameters values. In this paper, we optimize a long short-term memory autoencoder (LSTM-AE) model with the metahuristic algorithm Particle Swarm Optimization (PSO) in order to obtain the optimal parameters for the model to give better results in forecasting and then compared to other forecasting models. The evaluation metrics used for the comparison are mean squareerror (MSE) and root mean square error (RMSE).*

I. Introduction

With the fast technology development in smart grids and smart buildings, an efficient load forecasting has an important role to give a better concept of smart cities [1].
The role of the load forecasting is to predict the future energy load consumption for better scheduling and management of the power system. Many research activitieshave been done in the purpose to find accurate load forecasting models that affect the economy and the reliable power system operations. An efficient forecasting model will avoid the blackouts that may occur in supplying power to the consumers [2]. Moreover, the load forecast allows the good scheduling and planning, maintenance, contractevaluation and cost adjustment [3] [4]. In addition to the power system operations management, a good load forecasting has a great impact on the economic side since the power generation will be managed and the production cost will be reduced.
For an efficient load forecasting, several works have been done to find an appropriate and efficient model for forecasting. As traditional linear statistical methods, the autoregressive (AR), moving average (MA) and the autoregressive integrated moving average (ARIMA) were investigated for time series forecast due to their assumptions that data are stationary, linear with particular statistical distribution [5.6]. The need to find an accurate load forecasting model is always present. Lately, researchers focused on the deep learning approaches. The artificial neural networks ANN are able to capture data behavior in nonlinear complex patterns and large data [7-9] deep learning can learn efficiently complex input-output relations due to the large number of hidden layers. In [10] stacked AE has been used for prediction to reduce noise disturbance from the energy data. Recurrent Neural Networks (RNN) especially Long Short –term Memory (LSTM) , developed by Hochreiter et Schmidhuber [11], are used lately for forecasting due to their internal memory that allows the model to learn long-term dependencies in sequential data as energy load and demand[12].
Recently, some studies have introduced a hybridation of methods to forecast the load energy

consumption. In [13], the combination of convolutional neural network CNN and LSTM to predict the energy consumption where in [14] the CNN was integrated with Bi-directional LSTM for forecasting. All the models have given the best results but still the error rate high to be used for a good forecasting.

As all the other models, LSTM has different hyper-parameters that should be set. Some of these parameters are the number of layers, the time lag or window size and LSTM hidden unit per layer. The window size has direct effect on the accuracy of the LSTM model since it eliminates the redundant features [15]. Yet, setting these parameters to the optimal value is quiet difficult regarding time and computation limitations to investigate all the parameter space. Obtaining the optimal parameters for LSTM has been the objective of several studies.

From these studies, we have focused on the hybridation and optimization of the LSTM for its good performance in energy consumption prediction. We introduced an optimized hybrid model of PSO LSTM-AE for load forecasting. The Particle Swarm Optimization (PSO)is a metaheuristic algorithm that minimizes of maximizes the cost functions [16], to get an optimized model for load forecasting. We emphasize our work objective on optimizing the window size and hidden units in the LSTM layers that are related to detect temporal patterns of the dataset. Setting the time lags, window size, is important in designing LSTM method. When it is small, significant information may be ignored and when it is big, the unnecessary signals will act as noise.

The remainder of the paper will be as follows: section II will be about the related works on the load forecasting. Section III introduces the background for LSTM, LSTM-AE, PSO and the evaluation metrics. Section IV presents the model methodology of our work and the dataset that are used. Section V provides the experimental results and discussions. Finally, section VI concludes the results and proposes a future research.

II. Related works

Energy load forecasting is based on dataset generated from smart meters that may have redundancy, missed values, and uncertainties [17, 18]. Researchers have investigated several models in order to get an accurate forecasting model. Load characteristics and capacities of micro-grids define the accuracy of the forecasting models. [19]. there are three categories for load forecast based on time that would be predicted : short-term when the prediction is from hours to week ahead, medium-term with a prediction of months ahead and long-term where the prediction is for years [20].

Lately, many works have been done on short term load forecasting for its advantages in the energy consumption, peak load anticipation and customer management [21]. Several models have been used for the short term load forecast: starting with traditional model analysis until the machine learning models. [22–26] [17].

Traditional models such as linear regression, multiple-linear regression, have been used to energy forecasting [27] [28]. In addition to these models, autoregressive integrated moving average (ARIMA) ,autoregressive (AR), moving average (MA) and autoregressive moving-average were used also for the energy consumption prediction[5-6-29].In [30] and [31] another model is used for short term forecasting which is the exponential smoothing method. These traditional methods are frequently used in energy load prediction since the temporal domain is directly modeled. Like all the methods, the statistical methods have some drawbacks that may prevent them to be the accurate model for forecasting and we can list some of the drawbacks as the dropping accuracy,dimentionality,the linearity of data and hard response to unprecedented changes when extending the building [12].

An alternative solution to the statistical methods is the machine learning methods.In these methods, a non linear correlation between electric load and related factors can be extracted which lead to reach a higher accuracy comparing to the other methods [32] Among the machine learning methods that are used in load forecasting, there are the artificial neural networks (ANNs), support vector regression (SVR) and support vector machine [33-39].But the ANNs are mostly prefered in energy consumption prediction especially multilayer percepetron since the non linearity modeling is easier. ANN has several applications in load forecasting as mentioned in [17]. However, they have the disadvatange of assuming the independency between successive energy consumptions that lead to delete the predective potential that is present in the dependency of sequential data.[bouktifpso]and when the network gets complex , it has many layers and many parameters,the training becomes difficult [40].

Recently, the deep learning techniques have been used as an alternative way for forecasting. Deep Neural Networks DNNs are ANNs with several hidden layers. The DNNs solve the sequential dependence in the data sequences and simultaneously are suitable with nonlinearity and periodicity of data. Since they are excellent in predictions, the DNNs are widely used in Natural Language Processing (NLP) and image classifications. The methods that may represent the DNNs are: Convolution Neural Network (CNN), Deep Belief Network (DBN) and Recurrent Neural Network (RNN). The RNNs are mostly used in

Algerian Journal of Environmental Science and Technology
March edition. Vol.9. Nº1. (2023)
ISSN : 2437-1114
www.aljest.org

time series analysis due to their architecture that has the feedback connections permitting the persistence of the past information, nonlinear and time series forecasting abilities. Unlike the ANNs that do not take into consideration the effect of past information [41], the RNNs do which makes them more suitable in processing sequential data like in energy consumption prediction. In this study the RNN will be applied and its model, the long short-term memory (LSTM) will be investigated. The LSTMs are the most developed DNN models. The high performance of LSTM is based on memorizing long term dependencies due to its internal memory. The LSTM internal memory allows the network to be appropriate to deal with sequence dependent pattern like in demand and load energy[12,42].
Load forecasting is based on the past information which makes the fact of choosing the right time steps crucial. Many studies have been done to solve this issue. In this paper, we propose a hybrid PSO LSTM-AE for load energy forecasting. The contributions of this paper are:

- ✓ Data preprocessing stage to eliminate the outlier, missing and redundant values and normalizing of data
- ✓ The hybrid model PSO LSTM-AE is implemented. The PSO algorithm is used to find the optimal time lags and hidden units of LSTM-AE layers and then the optimal model is used for energy consumption prediction
- ✓ The evaluation of the hybrid model using the evaluation metrics RMSE and MAE and comparing the results with those of other models.

III. Background

III.1. From RNN to LSTM

The recurrent neural network, developed in 1980s [43-45], is composed of three types of layers: the input layer, one or more hidden layers and the output layer. Its structure, chain-like, of repeating modules allows to save the information from the precedent processing steps. the neural network deals with a sequence of inputs due to the feedback loop included in RNNs, i.e. , the feedback of the output at t-1 into the network is used to produce the output at t, for each step. However, learning long time dependencies for more than few time steps is difficult for RNNs because of the vanishing and exploding gradient issues [46]. Figure 1 demonstrates the sequential processing in RNN.

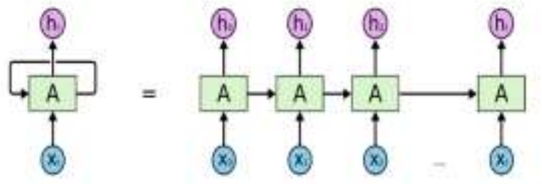

Figure 1. *Recurrent Neural Network (RNN) and the Sequential processing [47]*

The LSTMs were introduced by Hochreiter and Schmidhuber[11] developed from the RNNs by inserting new modules to solve and deal with the difficulties met with the RNNs concerning the long-range dependencies and reminding information for extended time periods. The LSTM technique has the chain structure form with the repeating module having another construction. It is composed of four interacting layers and one communication method [47]. Figure 2 illustrates the LSTM neural network structure.

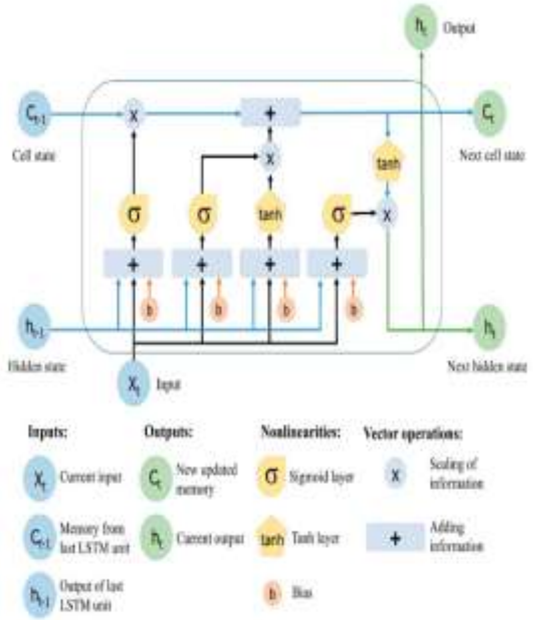

Figure 2. *Long Short-Term Memory (LSTM) neural network structure [48]*

The LSTM network is composed of cells. These cells are the memory blocks. The two states, cell and hidden states, are transmitted to the next cell. The unchanged data flows forward due to the cell state. The sigmoid gates allow the addition or the remove of data from the cell state. The memory cell and the gates are crucial parts in the LSTM. The long-term dependency problem is solved with the LSTMs by using gates, layer that compromises diverse individual weights, for the memorizing process control [48]. The input sequence {x1, x2, ...,xn} of the RNN are modeled recurrently:

$$h_t = f(h_{t-1}, x_t) \quad (1)$$

Where h_t is the hidden state and X_t is the input at t. the gradient vanishing and explosion problems are solved by the gates into the recurrence function f. the LSTM cells states are calculated as bellow:

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i] \quad (2)$$

$$f_t = \sigma\left(W_f.[h_{t-1}, x_t] + b_f\right] \quad (3)$$

$$o_t = \sigma(W_0.[h_{t-1}, x_t] + b_0] \quad (4)$$

$$\tilde{C} = tanh(W_c.[h_{t-1}, x_t] + b_c] \quad (5)$$

$$C_t = f_t \Theta C_{t-1} + i_t \Theta \widetilde{C_t} \quad (6)$$

$$h_t = o_t \Theta \tanh(C_t) \quad (7)$$

Where i_t, f_t and o_t are, respectively, the input, forget and output gates. The LSTM unit parameters are W's and b's, the actual cell state is C_t, Ce is the new values of the cell state. The sigmoid functions of i_t, f_t and o_t change the output to range between 0 and 1 as in equations (2)-(4). The results of these gates are dependent on X_t and h_{t-1}. The signal is blocked if the gate is 0. f_t determines the amount of the precedent state h_{t-1} permitted to pass. The decision of which new information from the input to be updated and added to the state cell is done by the i_t gate whereas determining which information to be the output based on the cell state is performed by the O_t. C is the memory cell which operates as a state information accumulator. Equation (6) defines how the update from C_{t-1} into $C_t$is done. C_e , the memory cell new values, and h_t, the actual output of the LSTM block are calculated in Equations (5) and (7). A transfer of the two states cell state and the hidden state to the next cell is performed for all t. this process will be repeated.

III.2. LSTM AE

The autoencoder AE is a feedforward neural network that is composed of three layers the input layer, the hidden layer, and the output layer. These layers are sequentially connected and work in an unsupervised learning.as a basic, the input size will be the same at the output thus they are used in generating data from the training dataset. These methods are called self-supervised too. AE is based on two stages: encoding stage, where it maps the input data into the hidden layer, and decodingstage, where it reconstructs the input data from the hidden layer. The encoder and decoder functions are explained by equations 8 and 9, respectively [49]:

$$Q : x(t) \longrightarrow F \quad (8)$$

$$D = F \longrightarrow X \quad (9)$$

Where X(t) is the input features and F is the feature space Q and D are the encoder and the decoder function, respectively[9]. The AE has an objective function which is the reconstruction error, the difference between the input and the output data that aims to minimize it. With several AE layers, minimizing this error will be considerable [50].

The LSTM AE is the change of the feed forward neural network of AE architecture to the LSTM recurrent neural network. LSTM AE is shown in figure 3. The LSTM algorithm is composed of two layers that act like encoder and decoder of AE. Learning from temporal dependencies from one sequence to another is possible with the LSTM cells [51]. The input sequence ($x1$, $x2$, ...,xn) are in encoded into the vector in the encoder layer. Afterthat, i twill be reconstructed into ($x\hat{}1$, $x\hat{}2$, ...,$x\hat{}n$) in the decoder layer. The root mean squared error that illustrates the error between the input and the output, reconstructed, data is the objective function.

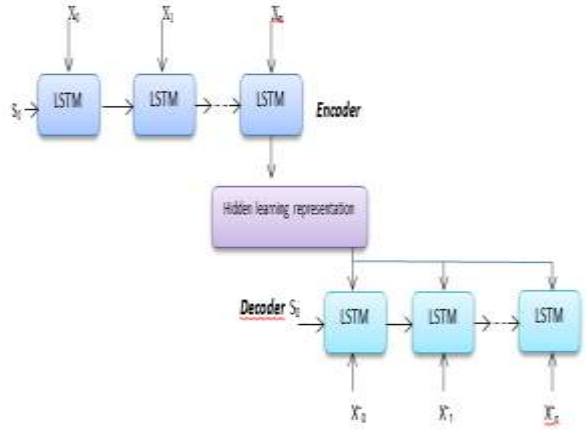

Figure 3.*The architecture of LSTM autoencoder.*

IV. Particle swarm optimization Algorithm

Optimization has gained much attention in various fields as computers came to an advanced stage. Many optimization techniques have been proposed to handle the more difficult problems that cannot be dealt with based on analytic or using classical techniques. Examples might include GeneticAlgorithms [51-52], FireflyAlgorithm [53], Differential Search Algorithm [54], Wind Driven Optimization [55],Teaching–Learning-Based Optimization [56-59], Grey Wolf Optimization [60], Biogeography based Optimization [61], Drone Squadron Optimization [62], Galaxy based search Algorithm [63], Spiral optimization technique[64] and Taguchi method [65].

In 1995 Kennedy and Eberhart have proposed the Particle Swarm Algorithm PSO [16]. It has been developed on the concept of the behavior of swarm of fishes or birds that search together for the best locations and move close to an optimum fitness function.

The PSO algorithm behaves like the other evolutionary computational algorithms. The swarm of particles, the population of solutions, explores the space for a suitable solution, particle, for a specific problem. Particles are distinguished by two properties: speed (velocity) and location (position).The PSO algorithm is illustrated in figure 4.

Algerian Journal of Environmental Science and Technology
March edition. Vol.9. Nº1. (2023)
ISSN : 2437-1114
www.aljest.org

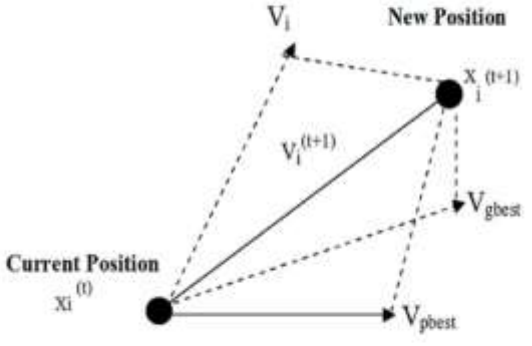

Figure 4*. PSO particlespositionchange [42]*

In two dimensional space, each element has a position expressed by x and y in the solutions space and a velocity that permits its exploration in the space. For each solution, the velocity is adjusted at eachiteration as in eq10 when the position is updated by adding the current position to the updated velocity as in eq11[66]

$$v_{k+1}^{i} = w.v_k^i + c_1.rand.\left(\frac{p_{best}^i - x_k^i}{\Delta t}\right) + c_2.rand.\left(\frac{p_{gbest}^i - x_k^i}{\Delta t}\right) \quad \textbf{(10)}$$

$$x_{k+1}^{i} = x_k^i + v_{k+1}^i.\Delta t \quad \textbf{(11)}$$

While v_{ik} is the previous velocity, $c_1.rand.\left(\frac{p_{best}^i - x_k^i}{\Delta t}\right)$ is knowledge, the cognitive part, and$c_2.rand.\left(\frac{p_{gbest}^i - x_k^i}{\Delta t}\right)$ is the collaboration between elements, the social part.

$(p_{best}^i - x_k^i)$and$(p_{gbest}^i - x_k^i)$are called acceleration by distance. The first term computes the distance between the best position found and the current position for the particle i while the second term calculates the distance between the best position of the particle i found by the swarm and its current location. The factors c1 and c2, the individual and social learning coefficients, are updated to c1+c2=4. The exploration capability of the algorithm is controlled by W; the particle inertia factor. It ensures the balance between the global and local explorations. Its values range from 0.2 to 0.8 which allows PSO algorithm to enhance its performance in the end of iterations and give its best convergence

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}}.iter \quad \textbf{(12)}$$

This process of updating the velocity and the position of particles is repeated until it meets the stopping criteria.The PSO algorithm steps are expressed in the flowchart of figure 5.

V. Results and Discussions

V.1. Model performance metrics

The forecast accuracy is evaluated by using some metrics like the root mean squared error (RMSE) and the mean absolute error (MAE) [68].

The RMSE is the square root of the average of the squared difference between the actual and the predicted values. In case of undesirable large errors, the RMSE is preferable. It is computed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad \textbf{(13)}$$

Where y_i is the desired output, y_i^ is the predicted output of the i^{th} observation, and n is the number of samples.

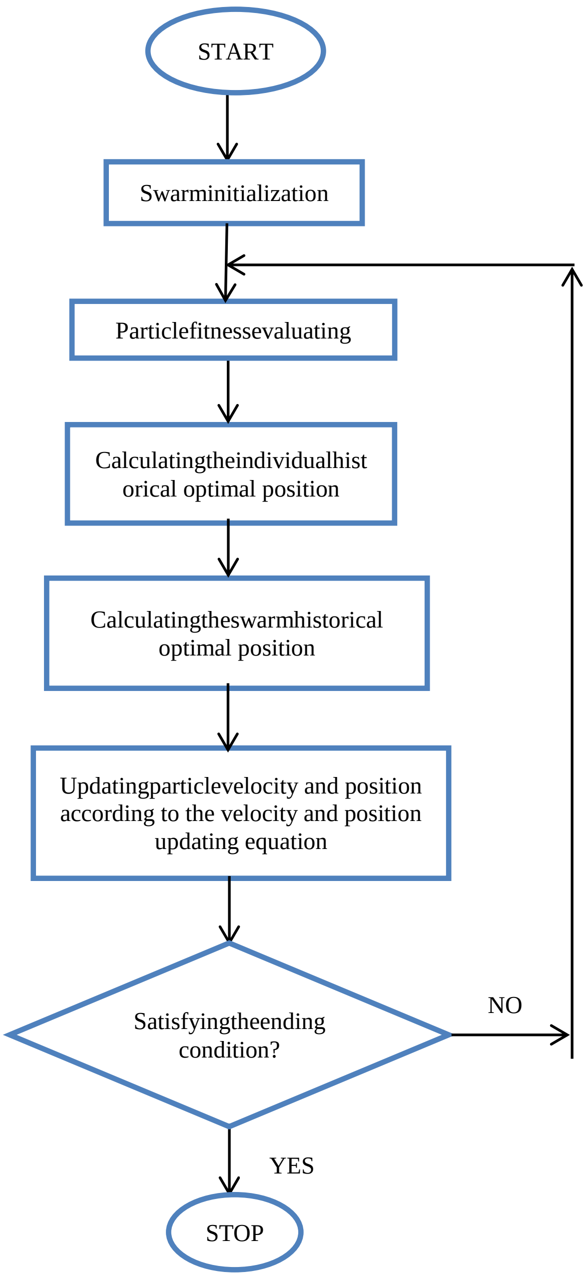

Figure 5*. Particle Swarm Optimization Algorithm Flowschart [53]*

While the MAE is the average of the absolute difference between the actual and the predicted values. The MAE is calculated as follows:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \quad \textbf{(14)}$$

V.2. Methodology

This section describes our suggested model to perform the energy load forecasting. The model is composed of some steps: the data preprocessing and preparation, searching the optimal hyperparameters time steps and the number of the hidden units in layers for the LSTM-AE model, validating the final optimized model and then comparing its performance with other models, deep and machine learning models, based on the evaluation metrics. Figure 6 portrays the framework of our proposed model.

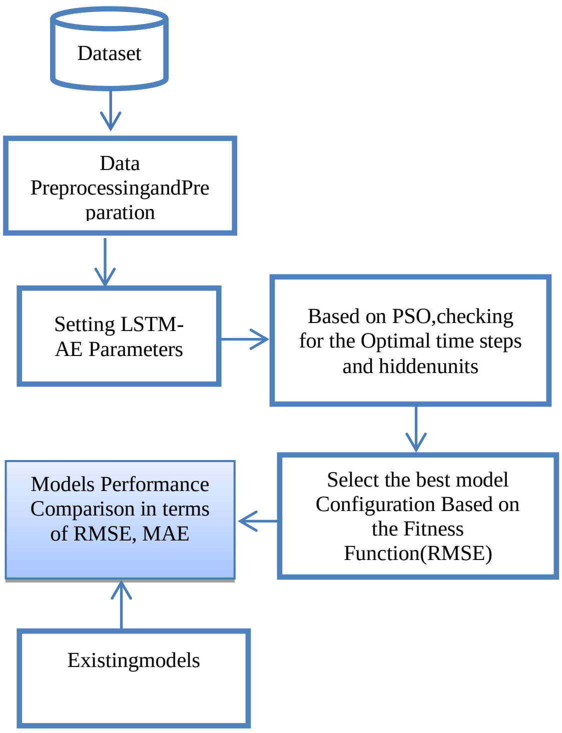

Figure 6. *Load energy forecasting proposed model.*

V.3. Data preparation and preprocessing

In this step, the preprocessing data is analyzed in details. The dataset used in this study is the American Electric Power coming from PJM's website and are in megawatts (MW). The data frequency used in this article is hourly and it was measured from 2004–10–01 to 2018–08–03. The total number of raw data points is 121271.
The dataset is treated before the training to eliminate the outlier, redundant and null values. The dataset is normalized also since the LSTM model is sensitive to the scale of the inputs. After that, we split the dataset into test and train set. We use the test data to evaluate accuracy of our model.

V.4. Model training

After the data is preprocessed, it passes to the training step. In this step we deal with the LSTM-AE model. In our study, the model consists of two encoder LSTM layers, a single repeated vector layer, two decoder LSTM layers and one dense layer. The LSTM-AE has several parameters to be set carefully in order to have an accurate model. Among the hyperparameters that must be there are the epoch number, batch size, time steps and number of neurons in LSTM layers.
There are also a various activation and optimization functions that must be selected to give better results. Setting these parameters manually is time consuming especially for the time steps where this parameter is important in prediction. For this reason, we use the Particle Swarm Optimization algorithm to find the optimal time steps and the hidden units, neurons, in layers numbers. As it is known, the metaheuristic algorithms do not give global optimum results but they are effective in finding the near optimal solutions. From this, we get several LSTM-AE configurations based on PSO. Based on the RMSE values, fitness function for PSO, of each configuration, the best model configuration is selected. Lately the performance of the selected optimized model is compared to other machine and deep learning models in terms of RMSE and MAE. The models that will be used in the comparison in our study are: Artificial Neural Networks (ANN) Convolution Neural Network (CNN), SVR, ARIMA, and Random Forest Regression.

V.4. Results and Discussions

a- Parameters setting

In this section, we set the parameters of the LSTM-AE model. Using the PSO algorithm, the best configuration model with optimal time steps and number of neurons in the two LSTM layer sis selected. The other parameters are set by testing the performance of the model for each one:

- ✓ The train and test size was tested for 50/50, 60/40, 70/30 and 80/20
- ✓ The number of epochs was varied from 50 to 400 while the batch size was varied from 8 to 256.
- ✓ The activation function selected for the model was the rectified linear unit (Relu) where the « ADAM » optimizer was selected for the optimization function.
- ✓ The optimal time steps and hidden neurons in LSTM layers are set using the PSO algorithm.

The best results that have been got were found with the model configuration with train and test split of

Algerian Journal of Environmental Science and Technology
March edition. Vol.9. Nº1. (2023)
ISSN : 2437-1114
www.aljest.org

70/30, 250 of epochs and 80 batch size. The optimal parameters found based on PSO are 25 for time steps and 60, 56 for the hidden units of the two encoder LSTM layers.

The experimental results

The actual data and the predicted data of our proposed model are plotted in the Figure 7.

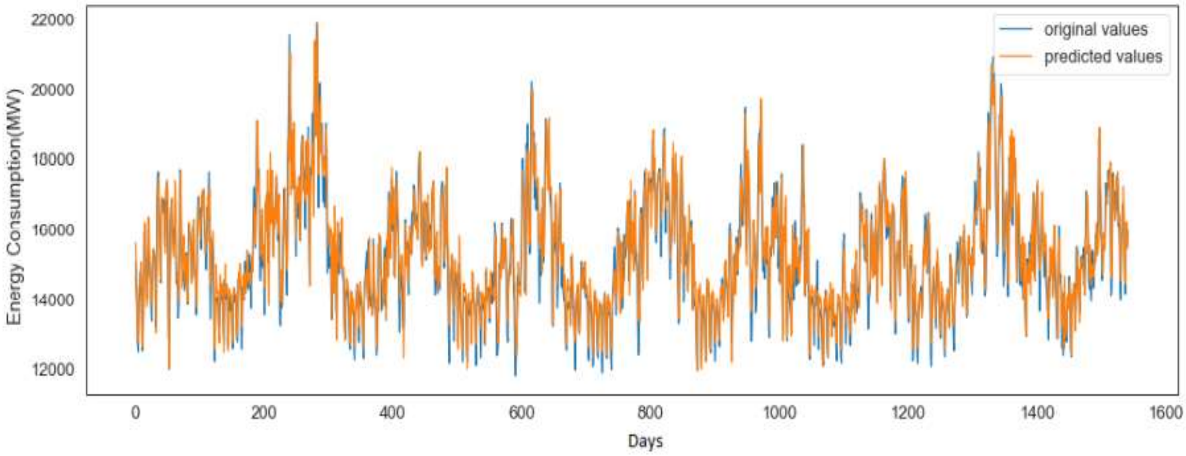

Figure 7. *Results of the forecasted energy compared to the raw measured data*

Figure 7 shows the predicted values fit the actual values in shape with a little difference but the performance of our model is good. In order to evaluate the performance of our proposed model, we compared it with the performance of other models in terms of RMSE and MAE. Table1 shows the RMSE and MAE results for each model.

Table 1.*Comparison of the proposed model to the state of art techniques*

model	RMSE	MAE
LSTM	720.67	593.07
CNN	790.45	574.63
Random Forest	915.6	794.5
ANN	743.8	531.21
Proposed model	**680.89**	**486.28**

From the results of table1, we can compare our proposed model with LSTM first. The table1 results show that the latter model outperforms the LSTM model in terms of RMSE and MAE. The PSO-LSTM-AE has RMSE and MAE of 680.89 and 486.28 respectively where LSTM has RMSE of 720.67and MAE of 593.07. These results show the effect of combining the LSTM with the autoencoder that gives a better performance than the usual LSTM. To evaluate the performance of our proposed model, table1 illustrates the performance of the other baseline models on the same dataset. The models are: CNN, Random Forest and ANN. CNN model has 790.45 as RMSE and 574.63 as MAE. Random forest scored the highest rate of errors with 915.6 RMSE and 794.5 MAE while ANN has 743.8 RMSE and 531.21 MAE. Comparing these results to those of the proposed model, it is obvious our proposed model scored better results in terms the specific metrics recording the smallest errors: 680.89 as RMSE and 486.28 as MAE.

IV. Conclusion

In this paper, we introduced an optimized model for load energy forecasting based on smart metering dataset. Firstly, we preprocessed the data to eliminate the null, outliers and redundant values. Moreover, we implemented our hybrid PSO with LSTM-AE method. The PSO algorithm was used to optimize the time steps number and the neurons in the LSTM layers. In addition, the hyperparameters of the model were set after multiple values testing in order to get better performance results. The experimental results of this hybrid model were compared to other results of other existing models to evaluate its performance. Finally, our hybrid model showed better results comparing to other models in terms of the evaluation metrics RMSE and MAE.

Conflicts of interest

No conflict of interest was declared by the authors.

V. Acknowledgements

This study has been conducted within the national research project N°:A01L07UN350120180002 Supported by “la Direction Générale de la Recherche Scientifique et du Développement Technologique (**DGRSDT**)” in Algeria.

VI. References

1. Borges,C. E.; Penya, Y. K.; Fernández, I.; Prieto, J.; Bretos, O. Assessing tolerance based robust short-term load forecasting in buildings. *Energies*(2013) 6:2110–2129. DOI: 10.3390/en6042110
2. Methaprayoon, K.; Lee, W.; Didsayabutra, P.; Liao, J.; Ross, R. Neural network-based short term load forecasting for unit commitment scheduling. *Proceedings of IEEE technical conference on industrial and commercial power systems. IEEE* (2003) 138–143.
3. Ruzic, S.; Vuckovic, A.; Nikolic, N. Weather sensitive method for short term load forecasting in electric power utility of Serbia. *IEEE Trans Power Syst*18 (2003) 1581–1586.
4. Kamel, N.; Baharudin, Z. Short term load forecast using Burg autoregressive technique. *Proceedings of international conference on intelligent and advanced systems ICIAS* (2007) 912–916.
5. Ahmed, N. K.; Atiya, A. F.; Gayar, N. E.; El-Shishiny, H. An empirical comparison of machine learning models for time series forecasting. *Econom. Rev*. 29 (2010) 594–621.
6. Khashei, M.; Bijari, M. An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Syst. Appl.* 37 (2010) 479–489.
7. Graves, A.; Jaitly, N. Towards End-To-End Speech Recognition with Recurrent Neural Networks. *Proceedings of the International Conference on Machine Learning*; Beijing, China (2014) 1764–1772.
8. Mao, J.; Xu, W.; Yang,Y.;Wang,J.; Huang, Z.; Yuille, A. Deep captioning with multimodal recurrent neural networks (m-RNN). *arXiv preprint* :1412.6632 (2014).
9. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *Proceedingsof the 27th International Conference on Neural Information Processing Systems* Montreal, QC, Canada (2014) 3104–3112.
10. Li, C.; Ding, Z.; Zhao, D.; Yi, J.; Zhang, G. Building energy consumption prediction: An extreme deep learning approach. *Energies* 10 (2017) DOI: 10.3390/en10101525
11. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* 9(1997) 1735–1780
12. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies* 11(7) (2018). https://doi.org/10.3390/en11071636
13. Kim, T. Y.; Cho, S. B. Predicting Residential Energy Consumption using CNN-LSTM Neural Networks. *Energy*182 (2019) 72-81.
14. Le, T.; Vo, M. T.; Vo, B.; Hwang, E.; Rho, S.; Baik, S.W.J.A.S. Improving electric energy consumption prediction using CNN and Bi-LSTM. *Appl. Sci*. 9(20) (2019); https://doi.org/10.3390/app9204237
15. Ribeiro, G.H.; Neto, P.S.D.M.; Cavalcanti, G.D.; Tsang, R. Lag selection for time series forecasting using particle swarm optimization. *Poceedings of the IEEE 2011 International Joint Conference on Neural Networks (IJCNN)* San Jose, CA, USA (2011) 2437–2444.
16. Kennedy, J., Eberhart, R.C. Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks* 4 Perth, Australia 1942– 1948
17. Deb, C.; Zhang, F.; Yang, J.; Lee, S.E.; Shah, K.W. A review on time series forecasting techniques for building energy consumption. *Renew. Sustain. Energy Rev.* 74 (2017) 902–924.
18. Kim, K. H.; Cho, S. B. Modular Bayesian Networks with Low-Power Wearable Sensors for Recognizing Eating Activities. *Sensors* 17(12) (2017) https://doi.org/10.3390/s17122877.
19. Liu, N.; Tang, Q.; Zhang, J.; Fan, W.; Liu, J. A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids. *Appl. Energy* 129 (2014) 336–345.
20. Hippert, H.S.; Pedreira, C.E.; Souza, C.R. Neural Networks for Short-Term Load Forecasting: A review and Evaluation. *IEEE Trans. Power Syst.* 16 (2001) 44–51.
21. Ryu, S.; Noh, J.; Kim, H. Deep Neural Network Based Demand Side Short Term Load Forecasting. *Energies* (2017) DOI: 10.3390/en10010003
22. Hagan, M.T.; Behr, S.M. The time series approach to short term load forecasting. *IEEE Trans. Power Syst.*2 (1987) 785–791.
23. Taylor, J.W. Short-term electricity demand forecasting using double seasonal exponential smoothing. *J. Oper.Res. Soc.* 54 (2003) 799–805.
24. Taylor, J.W.; de Menezes, L.M.; McSharry, P.E. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *Int. J. Forecast.*22 (2006) 1–16.
25. Park, D.C.; El-Sharkawi, M.; Marks, R.; Atlas, L.; Damborg, M. Electric load forecasting using an artificial neural network. *IEEE Trans. Power Syst.* 6 (1991) 442–449.
26. Hernandez, L.; Baladrón, C.; Aguiar, J. M.; Carro, B.; Esguevillas, S.A.J.; Lloret, J. Short-term load forecasting for microgrids based on artificial neural networks. *Energies* 6 (2013) 1385–1408.
27. Fumo, N.; Biswas, M. R. Regression analysis for prediction of residential energy consumption. *Renew. Sustain.Energy Rev*. 47 (2015) 332–343.
28. Amber, K.; Aslam, M.; Hussain, S. Electricity consumption forecasting models for administration buildings of the UK higher education sector. *Energy Build.*90 (2015) 127–136.
29. Chen, J. F.; Wang, W. M.; Huang, C. M. Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting. *Electr. Power Syst. Res.* 34 (1995) 187–196.
30. El-Keib, A.A.; Ma, X.; Ma, H. Advancement of statistical based modeling techniques for short-term load forecasting. *Electr Power Syst Res* 35 (1995) 51–58.
31. Taylor, J. W. Short-term load forecasting with exponentially weighted methods. *IEEE Transactions on Power Systems* 27(1) (2011) 458–464.
32. Ahmad, A.; Javaid, N.; Mateen, A.; Awais, M.; Khan, Z. A. Short-term load forecasting in smart grids: An intelligent modular approach. *Energies* 12(1) (2019) https://doi.org/10.3390/en12010164.
33. Barman, M.; Choudhury, N. B. D. Season specific approach for short-term load forecastingbased on hybrid FA-SVM and similarity concept. *Energy* 174 (2019) 886–896.
34. Bashir, Z. A. EL-Hawary, M. Applying wavelets to short-term load forecasting using PSO-based neural networks. *IEEE Transactions on Power Systems* 24(1) (2009): 20–27.
35. Chen, Y.;Luh, P. B.; Guan, C. ; Zhao, Y. ; Laurent, D. M.; Coolbeth, M. A. ; Friendland, P. B. ; Rourke ,S. J. Short-term load forecasting: Similar day-based wavelet neural networks. *IEEE Transactions on Power Systems* 25(1) (2009) 322–330.
36. Hong, W. C.; Fan, G. F. Hybrid empirical mode decomposition with support vector regressionmodel for short term load forecasting. *Energies* 12(6) (2019) https://doi.org/10.3390/en12061093.
37. Lee, S.; Jung, S.; Lee, J. Prediction model based on an artificial neural network for user-basedbuilding

Algerian Journal of Environmental Science and Technology
March edition. Vol.9. Nº1. (2023)
ISSN : 2437-1114
www.aljest.org

energy consumption in South Korea. *Energies* 12(4) (2019) https://doi.org/10.3390/en12040608.

38. Taylor, J.W.; Buizza, R. Neural network load forecasting with weather ensemble predictions. *IEEE Transactions on Power Systems* 17(3) (2002) 626–632.
39. Foucquier, A.; Robert, S.; Suard, F.; Stephan, L.; Jay, A. State of the art in building modelling and energy performances prediction: A review. *Renew. Sustain. Energy Rev.* 23 (2013) 272–288.
40. Sagheer, A.; Kotb, M. Unsupervised Pre-training of a Deep LSTM-based Stacked Autoencoder for Multivariate Time Series Forecasting Problems. *Sci Rep* 9 (2019). https://doi.org/10.1038/s41598-019-55320-6
41. Yoshihara, A.; Fujikawa, K.; Seki, K.; Uehara, K. Predicting stock market trends by recurrent deep neural networks. *Proceedings of the Pacific Rim International Conference on Artificial Intelligence* Gold Coast, Australia (2014) 759–769.
42. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M. Single and multi-sequence deep learning models for short and medium term electric load forecasting. *Energies* 12 (2019) https://doi.org/10.3390/en12010149.
43. Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagating errors. *Nature* 323 (1986) 533–536.
44. Werbos, P. J. Generalization of backpropagation with application to a recurrent gas market model. *Neural Netw.*1 (1988) 339–356.
45. Elman, J. L. Finding structure in time. *Cogn. Sci.* 14 (1990) 179–211.
46. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.*5 (1994) 157–166.
47. Xuan-Hien, L.; Hung, Viet, H.; Giha, L.; Sungho, J. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water* 11(7)(2019) https://doi.org/10.3390/w11071387
48. Mostofa, A. ; Kendall, N. Convolutional Neural Networks with LSTM for Intrusion Detection. *Proceedings of 35th International Conference on Computers and Their Applications* 69 (2020) 69-79.
49. Bao, W.;Yue, J.; Rao, Y. A deep learning framework for financial time series using stacked autoencoders and long short term memory. PLoSONE 12 (2017) 1–24.
50. Zulfiqar, A. K.; Tanveer, H.; Amin, U.; Seungmin, R.; Miyoung, L.; Sung, W. B. Towards Efficient Electricity Forecasting in Residential and Commercial Buildings: A Novel Hybrid CNN with a LSTM-AE based Framework. *Sensors* 20(5) (2020) https://doi.org/10.3390/s20051399.
51. Recioui, A. Azrar, A.; Bentarzi, H.; Dehmas, M.; Challal, M. Synthesis of Linear Arrays with Sidelobe Level Reduction Constraint using Genetic Algorithms. *International Journal of Microwave and Optical Technology* 3(5) 2008.
52. Recioui, A.; Azrar, A.; Dehmas, M.; Challal, M.; Bentarzi, H.Traveling wave antennas optimization using genetic algorithms. *3rd International Conference on Information and Communication* Damascus, Syria (2008).
53. Recioui, A.; Grainat, Y. 3D Antenna Array Design using Firefly Optimization Algorithm. *Algerian Journal of Signals and Systems* 4(2) (2019) 61-70.
54. Recioui**,** A. Optimization of Circular Antenna Arrays Using a Differential Search Algorithm. *ACTA PHISICA POLONICA A* 128 (2.B) (2015).
55. Recioui, A.; Benabid, M.; Djilani, N. Rectangular Antenna Array Optimization using Wind Driven Optimization. *Algerian Journal of Signals and Systems*1(2) (2016) 109-120.
56. Bouaraki, M.; Recioui, A. Optimal placement of power factor correction capacitors in power systems using Teaching Learning Based Optimization. *Algerian Journal of Signals and Systems* 2(2) (2017) 102-109.
57. Recioui, A.; Dassa, K. Design of Standalone Micro-Grid Systems Using Teaching Learning Based Optimization. *Algerian Journal of Signals and Systems*2(2) (2017) 75-85.
58. Recioui, A. Application of Teaching Learning-Based Optimization to the Optimal Placement of Phasor Measurement Units. Handbook of Research on Emergent Applications of Optimization Algorithms *IGI Global* (2018).
59. Azzougui, Y.; Recioui, A.; Mansouri, A. PMU Optimal Placement in Wide Area Monitoring Systemsusing Grey Wolf Optimization technique. *Algerian Journal of Signals and Systems* 4(1) (2019) 1-7.
60. Recioui, A.; Traikia, O. ; Grainat, Y. Elliptic and Concentric-Elliptic Antenna Array Design using Grey Wolf Optimization. *Algerian Journal of Signals and Systems* 4(1) (2019) 32-38.
61. Recioui, A.; Arhab, N.; Zeghad, I. Pattern Design of 2D Antenna Arrays using Biogeography Based Optimization. *International Journal of Computer Science, Communication & Information Technology (CSCIT)* 6 (2) (2019) 38-43.
62. Dekhandji, F. Z.; Azzougui, Y.; Recioui, A. PMU Optimal Placement for FaultDetection in Wide Area MonitoringSystemsbased on the Drone Squadron Optimization. *Algerian Journal of Signals and Systems* 4(1) (2019) 25-31.
63. Recioui, A. Thinning of Concentric Circular Arrays using Galaxy Based Search Algorithm. *IEEE fifth International conference on Electrical engineering (*2017) Boumerdes Algeria.
64. Recioui, A. Concentric Ring Arrays Optimization Using the Spiral Inspired Technique. *Algerian Journal of Signals and Systems* 3(1) (2018) 10-21.
65. Recioui, A.; Bentarzi, H. Null Steering of Dolph-Chebychev Arrays using Taguchi Method. *WSEAS Transactions on Communications* 9(11) (2010) 720-729.
66. Eberhart, R. C.; Simpson, P.; Dobbins, R. 1996. Computational Intelligence PC Tools. *Academic Press Professional*, Massachussets, EUA
67. Dongshu, W.; Dapei, T.; Lei, L. Particle swarm optimization algorithm: an overview. *Springer-Verlag Berlin Heidelberg* (2017).
68. Hyndman, R.; Koehler, A. Another look at measures of forecast accuracy. *Int. J. Forecast.* 22 (2006) 679–688.

Please cite this Article as:
Saoud A., Recioui A., Load Energy Forecasting based on a Hybrid PSO LSTM-AE Model., ***Algerian J. Env. Sc. Technology*, 9:1 (2023) 2938-2946**