An Evaluation of Conventional and Computational Intelligence Methods for	Đánh Giá Các Phương Pháp Thông Thường và Trí Tuệ Tính Toán Trong Dự Báo Phụ

Medium and Long-Term Load Forecasting in Algeria checked	Tải Điện Dài Hạn Và Trung Hạn ở Algeria
Abstract— Electric load forecasting is one of the most important areas in electrical engineering, due to its main role for economic and reliable operation in power systems. In particular, accurate medium and long-term forecasts have significant effect on grid expansion planning and future generating capacity scheduling. This paper uses the Algerian electricity demand observations to evaluate methods for medium and long-term predictions. We consider methods designed to capture the trend and the seasonal cycle in the data as well as computational intelligence techniques. Among the variety of methods considered, satisfactory results were obtained by the adaptive neuro-fuzzy inference system and the autoregressive integrated moving average based approaches.	Tóm tắt-Dự báo phụ tải điện là một trong những lĩnh vực quan trọng nhất trong kỹ thuật điện nhằm đảm bảo tính kinh tế và vận hành đáng tin cậy của các hệ thống điện. Đặc biệt, dự báo dài hạn và trung hạn chính xác có ảnh hưởng đáng kể đến kế hoạch mở rộng lưới điện và hoạch định công suất phát điện trong tương lai. Bài báo này dựa trên nhu cầu điện năng ở Algerian để đánh giá các phương pháp dự đoán trung hạn và dài hạn. Chúng tôi xét các phương pháp được thiết kế để ghi nhận xu thế và chu kỳ theo mùa của dữ liệu cũng như các kỹ thuật trí tuệ tính toán. Trong số những phương pháp được xét, chúng tôi nhận thấy kết quả thu được tốt nhất đối với các phương pháp tiếp cận dựa trên hệ thống suy luận mạng nơron thích nghi mờ và trung bình trượt kết hợp tự hồi quy.
I. INTRODUCTION	[redacted]
In power generating industry, electric load forecasting plays a dominant part for achieving the goal of economic and	[redacted]

reliable operation in power systems. Power companies need load forecasting models, to accurately predict the amount of power that must be delivered to their customers and so as to adjust supply/demand balance at any time in the best conditions of cost and safety. Accurate load forecasting models are needed for a variety of time horizons: very short-term for load-frequency control and economic dispatch functions; short-term for the day-to day operation, scheduling and load-shedding plans of power utilities; medium-term for maintenance programs; and long-term for power system expansion planning. Very short-term (VSTLF), short-term (STLF), medium-term (MTLF) and long-term load forecasts (LTLF) are range from few minutes to an hour, an hour to one week, one week to one year, and one year to decades, respectively. As very-short and short-term load predictions are very important for the daily operation of generation and distribution facilities, it have been extensively studied in the literature of load forecasting [1-5].	[redacted]

Instead, only a few works can be found in the literature about MTLF and LTLF [6-8].	[redacted]
In the recent years, along with deregulation of electricity market and renewable energy integration, the issue of accurately electricity demand forecasting has received more attention for power system managers and consequently, the field of electric load forecasting becomes more and more important. Load forecasting accuracy has significant effects on power system operation and production costs: forecasting errors typically cause highly increasing operating costs [9]. Overestimation of future load results in excess supply, and it is also not welcome to the international energy network. In the contrast, underestimation of load leads to a failure in providing enough reserve and implies high costs in peaking unit.	[redacted]
However, it is difficult to predict electricity demand accurately over a distant planning period because the uncertainly related to a large number of exogenous factors that affect directly or	[redacted]

indirectly on load variation (climate change, social activities, and economic factors... etc.). Therefore, any MTLF and LTLF model is inaccurate by nature [10]. In the last few decades, many techniques have been developed to improve the accuracy of medium and long-term load forecasting. They can be typically divided into two broad categories: statistical approaches; and artificial intelligence based methods. Conventional methods include regression approaches [11], autoregressive time series approaches [12]. However, the linear modeling process of classical methods faces difficulty to determine empirically the correct complex relationship that exists between the load and the explanatory inputs that influence on it. Consequently, the forecasting accuracy is shown up to 20% error compared with real values [6]. In recent years, considerable attention has been given to developing electricity demand forecasting models based on intelligent techniques, such as artificial neural network (ANN) [13], fuzzy logic	[redacted]

systems [14], expert systems [15], data mining [16], adaptive neuro-fuzzy inference system (ANFIS) [17], and some hybrid models [18].	[redacted]
Due to the clear interest of MTLF and LTLF for power system operation and planning, and to the challenging task of accurate load predictions for such distant periods; the present work focuses on the application and the evaluation of different conventional and computational intelligence methods for performing medium and long-term load forecasts of the Algerian power system. We also aim to reduce the uncertainty related to the forecasted loads, by estimating an interval in which the future electric load observations will fall.	[redacted]
The rest of the paper is organized as follows. Section II designates the evolution of the electricity demand in Algeria and describes the used data. Section III describes the proposed estimation methods. Section IV provides and explains forecasting results. Finally, Section V concludes the paper.	[redacted]

II. MEDIUM AND LONG TERM LOAD FORECASTING DATA

In this study, we consider two load series; one is for MTLF and the other is for LTLF. The first; as shown in Fig. 1 consists of five years of monthly observations for electricity demand in Algeria: from January 2010 to December 2014. This series shows an intra-year seasonal cycle of duration 12 periods. Hence, the electricity consumption reaches its highest levels in winter when all inhabitants are all together at home and simultaneously use, lighting, electric heaters and other household appliances (televisions, computers ...etc.); while in summer, peaks are caused principally by the rise of using air conditioning appliances in latest few years. The second series records the Algerian national annual peak power consumption along fourteen years: from 2001 to 2014. As illustrated in Fig.2, the electric power consumption at the national level is increasing year after year due to demographic and urban development. The national annual peak

[redacted]

has multiplied by about 2.28 between 2001 and 2014 and rising also from 4791 MW to 10927 MW; which represents an average annual growth rate equal to 6.59%.

[redacted]

III. PROPOSED LOAD FORECASTING METHODS

[redacted]

The medium-term load forecasting models proposed in this study are designed to forecast the monthly peak load at the next year in Algeria, while the long-term load forecasting methods are designed to forecast the national annual peak load up to six years ahead. Holt-Winters exponential smoothing (HWEs) and seasonal autoregressive integrated moving average (sARIMA) methods are used for MTLF, due to its ability for modeling the seasonal cycle on the electric load time series. Holt’s exponential smoothing (HES) and autoregressive integrated moving average (ARIMA) methods are used for LTLF. While ANN and ANFIS models are used for both MTLF and LTLF.

[redacted]

A. Holt’s exponential smoothing technique	[redacted]
The Holt’s exponential smoothing method is used when data pattern contain a trend. The formulation for the Holt’s exponential smoothing is given in the following expressions:	[redacted]
St and Tt, are the smoothed level and trend; a and y are the smoothing parameters; yt is the actual value of the time series in period t; and yt(k) is the k step-ahead forecast made from forecast origin t.	[redacted]
Hence, smoothing parameters of any exponential smoothing technique have a great influence in the forecasting accuracy. In order to improve the forecasting accuracy, the optimization of the Holt’s method parameters has been considerably drawn our attention. In this paper, HES method parameters are estimated by minimizing the forecasting error of the one step ahead annual peak power demand forecasting over the period from 2004 to 2009. We first generated 2000 vectors of parameters from a uniform random	[redacted]

number generator between 0 and 1. For each of the vectors, we then evaluated forecasts over the previously discussed period. The one producing the lowest value of the mean absolute percentage forecasting error (MAPE) was chosen as the final parameter vector. The optimized values of method parameters were: a = 0.701, and y = 0.971.	[redacted]
B. Autoregressive integrated moving average model	[redacted]
Introduced by Box and Jenkins [19], the ARIMA model has been one of the most popular approaches in forecasting. In an ARIMA model, the future value of a variable is supposed to be a linear combination of past values and past errors, expressed as:	[redacted]
The model is generally referred to as an ARIMA(p, d,q) model where parameters p, d, and q are non-negative integers that refer to the order of the autoregressive (AR), integrated (I), and moving average (MA) polynomials of the model, respectively. The ARIMA (p, d, q) model	[redacted]

can be represented as:

0p(B)Vdyt = Co + 9q(B)et (5)

Where yt is the actual value, et is the random error at time t, 0i and 9j are the coefficients, B is back-shift operator and C0 is a nonzero constant. Basically, three phases are included in an ARIMA model: model identification, parameter estimation, and diagnostic checking [20]. The adopted ARIMA model in this paper can described as ARIMA{ 3,1,1).

C. Holt-Winters method

For seasonal time series, the Holt-Winters method is useful because it can capture both trend and seasonality in the historical data. The formulation for multiplicative seasonality is given in the following expressions:

St and Tt, are the smoothed level and trend; Dt and is the seasonal indices for the intra-year seasonal cycle si of duration twelve periods; a,y,S are the smoothing parameters; yt is the actual value of the time series in period t; and yt(k) is the k step-ahead forecast made from forecast origin t.

[redacted]

[redacted]

In this paper, HWES method parameters are estimated by minimizing the MAPE value of the monthly peak power demand forecasting. We first generated 2000 vectors of parameters from a uniform random number generator between 0 and 1. For each of the vectors, we then evaluated the MAPE on the monthly peak demand forecasting on the period between 2012 and 2013. The one producing the lowest MAPE value was chosen as the final parameter vector. The optimized values of method parameters were: a = 0.093; y = 0.253; and S = 0.748.

D. Seasonal autoregressive integrated moving average

Seasonal autoregressive integrated moving average is an extension of the autoregressive integrated moving average processes which has been introduced to model time series with trends, seasonal pattern and other non-stationary characteristics. In MTLF, the seasonality component comes from the intra-year seasonal cyclic of duration twelve periods (twelve months). A SARIMA, denoted by

[redacted]

SARIMA (p, d, q) x (P,D,Q)S , can be stated in the following form: (10) Where yt and £t are the actual value and random error at time t , respectively. B is the lag operator that satisfies: Btyt=Yt-1. ^p(Bs), 0Q(Bs) and (1 -BS)D are corresponding autoregressive, moving average and differencing parts for seasonal components. While VP(B), 9q(B) and (1 - B)d are corresponding autoregressive, moving average and differencing parts for the non-seasonal component. S is the length of the season. The adopted SARIMA model in this paper can described asSARIMA(1,1,1) x (1,1,1)12. E. Artificial neural network model ANNs are non-linear and nonparametric methods that have inherent capability to solve any complex nonlinear relationships between input and output through a learning process involving historical data trends. The objective of training is to find the set of weights between neurons that determine the global minimum of error function. In this regard, back propagation	[redacted] [redacted]

(BP) is a kind of learning algorithm which is both simple and applicable [21]. Therefore, a univariate feedforward back-propagation neural network is considered to be used in this work.	[redacted]
The proposed load forecasting ANN model consists of one hidden layer of three sigmoid neurons followed by an output layer of linear neuron. The load data from previous year is adopted as input, and load data for the next year is taken as the output of the network. The Levenberg-Marquardt algorithm is designed to train the network and the performance function criteria is the mean square error (MSE) criteria. The maximum number of epochs for training the network is limited to 20 epochs.	[redacted]
F. Adaptive neuro-fuzzy inference system	[redacted]
An adaptive Neuro-Fuzzy inference system is a fuzzy Takagi-Sugeno model put in the framework of adaptive systems to facilitate learning and adaptation [22]. An adaptive network is a multi-layer feed-forward network in which each node	[redacted]

(neuron) performs a particular function on incoming signals. The first layer executes the fuzzification process. The second layer performs the fuzzy AND of the antecedent part of fuzzy rules. The third layer normalizes the membership functions. The fourth layer executes the consequent part of the fuzzy rules, and finally the fifth layer calculates the output of fuzzy system by summing up the outputs of the fourth layer.	[redacted]
The proposed Neuro-Fuzzy inference structure use two sigmodal membership function, two fuzzy rules, and twenty epochs for identifying the fuzzy inference system (FIS) parameters and the optimal topology of the ANFlS. The peak load data from previous four years are adopted as inputs in LTLF model, while the number of inputs is reduced to one in the case of MTLF.	[redacted]
IV. RESULTS AND DISCUSSION	[redacted]
Proposed MTLF models are designed to forecast the Algerian monthly peak load for a lead time of a one year ahead on the period between 2014 and 2015. While the	[redacted]

lead time in the case of LTLF, is a one year ahead on the period from 2010 to 2014; and six years ahead on the period from 2015 to 2020. Forecasting results of MTLF and LTLF models are presented on Fig. 3 and Fig. 4, respectively. These figures illustrate that the proposed MTLF and LTLF models have successfully predict peak electricity demand, since the shape of predicted load curves follows almost the shape of actual load curves. However, one can observe that the accuracy of the proposed LTLF artificial neural network model is sometimes unsatisfactory. For example, an overestimation in the peak load forecasting in 2011 can be noted. Furthermore, the model was not able to provide several steps ahead forecasting in the period between 2015 and 2020, because the annual load-demand growth of forecasted loads in the cited period is relatively low. This decrease on forecasts precision is mainly related to the lack of a sufficient number data for training the network.	[redacted]

However, in order to evaluate the obtained results, two error measurement criteria have been calculated: mean absolute percentage error, and root mean square error (RMSE): Where, yt is the forecasted load at time t, yt is the real load at time t. n is the total number of forecasting cases. Detailed forecasting results, as indicated on Table. 1 and Table.2; show a superior accuracy of the proposed adaptive neuro-fuzzy inference system on LTLF; with a MAPE value of 3.017%. The best approach in MTLF is the seasonal autoregressive integrated moving average model, with a MAPE value of 4.620%. Furthermore, the best two methods in both MTLF and LTLF are the ANFIS model and a time series based approach (ARIMA or SARIMA). In addition to accurate load forecasts necessity, it is also important to estimate an interval in which future electric load observations will fall. In Fig.5, a simple combination of the forecasted loads of SARIMA and ANFlS models for 2015 is	[redacted] [redacted]

plotted; together with the predicted intervals. In Fig. 6, the combination between ARIMA and ANFIS models forecasts is also plotted together with the predicted intervals for the yearly peak load forecasting on the period between 2015 and 2020. Forecasted loads for 2015 and for the period from 2015 to 2020 are also presented on Table.3 and Table.4, respectively. Summer peak (August 2015) of the combined SARIMA- ANFIS is assumed to be 11423 MW; which is 4.539% higher than the peak of August 2014 (10927 MW). We also estimate an average annual peak load growth on the period from 2015 and 2020 equal to 5.826%; and the load for 2020 is supposed to be multiplied by about two if we take the electricity demand of 2010 as reference.	[redacted]
TABLE III. COMBINED SARIMA-ANFIS: RESULTS FOR THE MONTHLY PEAK LOAD FORECASTS OF 2015	[redacted]
TABLE IV. COMBINED ARIMA-ANFIS: RESULTS	[redacted]

FOR THE ANNUAL PEAK LOAD FORECASTING FROM 2015 TO 2020 V. CONCLUSION This paper uses Algerian electricity demand observations to evaluate methods predictions up to six years ahead. Post-sample forecasting results showed that the adaptive neuro- fuzzy inference system and the autoregressive integrated moving average based approaches compete well in medium and long-term electricity demand forecasting, since its mean absolute percentage error value has not exceed 5% in both medium and long-term load forecasting. To reduce the uncertainty related to the forecasted loads, we also estimated an interval in which the future electric load observations will fall.	[redacted]