이 학 석 사 학 위 논 문

Comparison Study of Neural Network Methods for Electricity Forecasting

전력량 예측에 대한 신경망 방법론 비교

2017년 2월

서울대학교 대학원

통계학과

김 정 애

Comparison Study of Neural Network Methods for Electricity Forecasting

지도교수 이 상 열

이 논문을 이학석사 학위논문으로 제출함
2016년 10월

서울대학교 자연과학대학원
통계학과
김 정 애

김정애의 이학석사 학위논문을 인준함
2016년 12월

위 원 장	오 희 석	(인)
부위원장	이 상 열	(인)
위 원	이 재 용	(인)

Comparison Study of Neural Network Methods for Electricity Forecasting

by

Kim Jeongae

A Thesis
submitted in fulfillment of the requirement
for the degree of
Master of Science
in
Statistics

The Department of Statistics
College of Natural Sciences
Seoul National University
February, 2017

Abstract

Load forecasting has an important meaning in economic and secure operation of power systems. So, numerous methods are proposed to enhance the accuracy of load forecasting. In this paper, we introduce these methods and proposes two methods that combines neural networks and time series model. First one is neural networks with global ARMA model and the other one is neural networks with local ARMA model which uses moving window. And then we compares the ordinary artificial neural network model and our proposed models by simulation studies and load demand data from France. Since load demand data has trend and seasonality, we use differenced data to fit the ARMA model. Finally, we compares the forecasted values up to 24 hours to see the accuracy of each models.

Keywords: Load forecasting, Load management, Neural networks, ARMA model.

Student Number: 2015-20294

Contents

List of Figures

Chapter 1

Introduction

The estimation of future demand for electric load is an important task because it can be used to manage power systems more efficiently. Demand forecasts can be used to control the generation and distribution of electricity more efficiently. From an operational point of view, the key question is whether there will be problems in meeting the peak demand; a failure to meet this peak demand could result in some blackouts. In addition, forecasts of demand have become even more important because they are also required for estimating future electricity spot prices.

Therefore, various attemps are made to predict exact values such as time series modeling[1], semiparametric regression [9] [4], neural networks [7], Bayesian statistics [8], time-varying splines[12], judgmental forecasting [14], grey dynamic models [18], transfer functions [19], exponential smoothing [20] and decomposition techniques [22]. These forecasts can be characterized by their forecasting lead time: a) Long-term forecasts of the peak demand, needed for capacity planning and maintenance scheduling [17], six month to ten years ahead; b) Medium-term demand forecasts, required for power system operation and planning [10], one week to six months ahead; c) Short-term load

forecasting: control and scheduling of power generation system, five minutes to one week ahead. Short-term forecasts are also has security and economic aspects. Hence, error in predicting electricity demand has significant cost implications for companies operating in competitive power markets [3].

In the short run, the electricity demand has a strong deterministic component such as meteorological conditions, seasonal effects and special events. Particulary, the strongest one is weather related variation, when predicting electricity demand for lead times beyond a day-ahead [5] [21]. However, in the case of short term load forecasting, univariate models are considered to be more practical than multivariate models [3]. Thus, in this study, we exclude an weather variables, although we accept that a differently specified neural network with meteorological variables may be useful for load modeling. We would hope that the better performing methods in our study can serve as benchmarks in future studies with ANNs.

In recent study [15], a novel hybrid model of artificial neural networks using ARIMA is proposed, and empirically show that this model can be an effective way to improve forecasting accuracy. Since it has not been considered for electricity forecasting, we would compare this method using two kinds of ARIMA models, global and locally fitted models. Using hourly electricity data from France, we compare not only the methods identified in [15] and but also an artificial neural networks with preprocessed data[7]. All the methods are based on the differences of data to deal with the seasonality that typically arises in demand data, intra day and intra week seasonal cycles. We consider lead times up to 24 hour ahead because the prediction for lead times less than six hours ahead is of particular interest.

We first consider an artificial neural network with preprocessed data. In a second stage, we turn our attention to two kinds of hybrid neural network with ARIMA models in section 2 . Then we compare such methods using

simulation data in section 3. In section 4 compares our models using load demand data. Finally section 5 provides concluding remarks.

Chapter 2

Models

2.1 Nerual networks

Neural network models have been recieved much attentions in load forecasting problems. Although various designs of neural network models have been proposed, the most widely used one for modeling and forecasting time series is single hidden layer feedforeward network. Fig 2.1.1 shows a typical single hidden layer feedforward model used for forecasting purposes. The input nodes are the previous lagged observations while the output provides the forecast for the future value. Hidden nodes with appropriate nonlinear transfer functions are used to process the information received by the input nodes. The model has the following mathematical representation :

$$f(z_t, v, w) = g_2\left(\sum_{j=0}^{m} v_j g_1\left(\sum_{i=1}^{k} w_{ji} z_{t-i}\right)\right) \tag{2.1.1}$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are activation functions, which we chose it as relu and linear, respectively. The w_{ji} and v_j are the weights (parameters); $k$is the number of input nodes; and m is the number of hidden nodes.

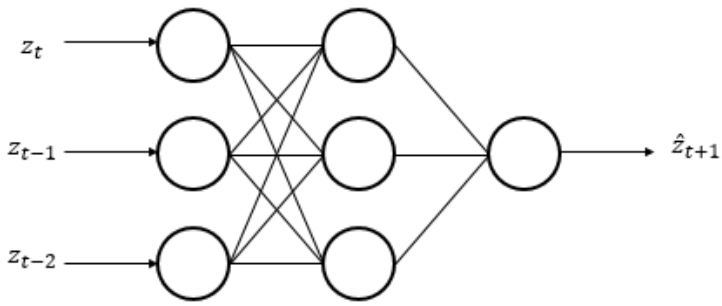

Figure 2.1.1: Neural network model

Functionally, the neural network expressed in (2.1.1) is equivalent to a nonlinear AR model. This kind of neural network predict well in almost any cases of time series, but someone suggest that when time series has seasonal and trend component, data preprocessing improve the neural network modeling and forecasting performance [24]. In light of this facts, we use raw data when data is stationary like simulation studies in section 3 and preprocessed data when data is non-stationay like load demand data in section 4.

2.2 Neural network with global ARMA models

There exist few guidelines for building a neural network model for time series. One of them considered time series as non linear function of several past observations and random errors [15]. Since electricity data is known as non linear time series data, we selected benchmark of this method to forecast the load demand. So, the final model is :

$$y_t = f[(z_{t-1}, z_{t-2}, \cdots, z_{t-m}), (e_{t-1}, e_{t-2}, \cdots, e_{t-n})] \tag{2.2.1}$$

where f is a nonlinear function determined by neural networks, $z_t = (1-B)^d y_t$, e_t are residuals at time t and m and n are integers. So in the first stage, an ARIMA model is fitted in order to generate the residuals e_t. Then, a neural network is used in order to model the nonlinear and linear relationships existing in residuals and original data. That is,

$$z_t = w_0 + \sum_{j=1}^{Q} w_j \cdot g\left(w_{0j} + \sum_{i=1}^{p} w_{ij} \cdot z_{t-i} + \sum_{i=p+1}^{p+q} w_{ij} \cdot e_{t+p-i}\right) + \epsilon_t \quad (2.2.2)$$

where, $w_{ij}(i = 0, 1, 2, \cdots, p+q, j = 1, 2, \cdots, Q)$ and $w_j(j = 0, 1, 2, \cdots, Q)$ are connection weights and the p, q, Q are integers which should be determined in a design process of neural networks.

The number of p and q are determined by the underlying properties of data. That is, if the data only consist of nonlinear sturcture, then q can be 0 since ARIMA is a linear model and does not able to model nonlinear relationship.

It can be noted that suboptical models can be used in hybrid model. However, suboptimality may not affect the usefulness of the hybrid model. Granger pointed out that for a hybrid model to produce superior forecasts, the component model should be suboptimal [11].

2.3 Neural network with local ARMA models

Our final model to introduce is the benchmark version of the model in section 2.2. These two models are almost same but the only difference is the way to get residuals, e_t. In this version of hybrid model, residuals are obtained from locally fitted ARMA models. To explain this, we introduce the concept of a moving window with size w and interval b: see Figure 2.3.1,

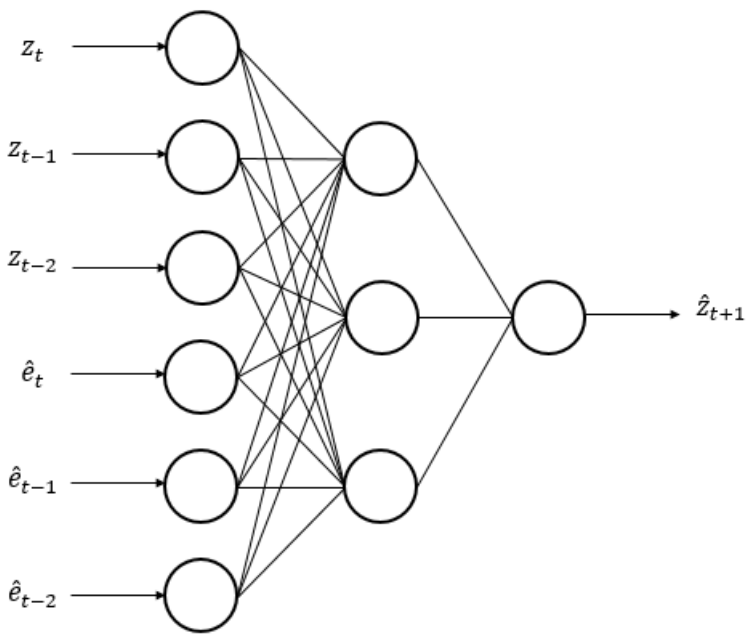

Figure 2.2.1: hybrid neural network model

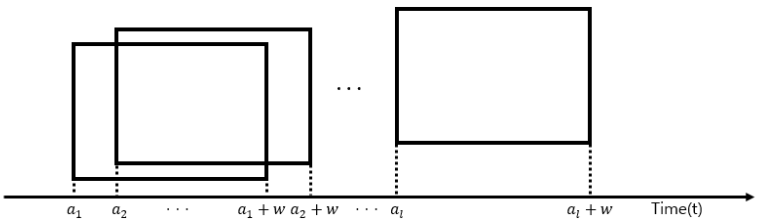

Figure 2.3.1: Example of moving window

where $a_{t+1} = a_t + b$. The rectangles represent the time segments over which the ARMA model fitted. Then we get the residuals from differently fitted ARIMA models. After that, we feed this residuals and original data to train and test neural network models.

To determine the best neural network structure, we compare the test errors. Because of the nature of the autocorrelations in time series, the number of input nodes or the lagged observations used in the neural networks is often a more important factor than the number of hidden nodes[23]. We here use R to conduct automatic ARMA and neural model building.

Chapter 3

Simulation Studies

In this section we perform a simulation study, using random coefficient autoregressive processes to investigate the accuracy of the methods we have described. In [2], the random coefficient autoregressive model of order 1 is given by the equation:

$$X_t = (\varphi + b_t)X_{t-1} + e_t \tag{3.0.1}$$

where φ is a parameter and $\{b_t\} \sim \text{iid}N(0, \omega^2)$ and $\{e_t\} \sim \text{iid}N(0, \sigma^2)$ are independent, where $\omega^2 + \sigma^2 < 1$.

To see the different forecasting accuracy of various models, we use the different levels of coefficients and variances. For each $j + 1 \leq t \leq j + 400$, $j = 0, 400, 800, \ldots, 3200$, we generate X_t from model 3.0.1 with $\sigma^2 = 0.5$ and randomly chosen φ and ω^2 from the uniform distributions $U[-0.3, 0.3]$, $U[0, 0.1]$: see the plot 3.0.1.

The first 2520 observations are used for the model selection and parameter estimation and the last 1080 points are reserved as the test set for forecasting evaluation and comparison. And the forecasts above the 1 horizon are obtained by iteration: the current output is used as an input for predicting the next

output.

We use mean absolute error (MAPE) to measure the performance, because MAPE is the most widely used to measure the accuracy of load forecasting. It usually expresses accuracy as a percentage, and is defined by the formula:

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{3.0.2}$$

where n is the number of points in the test set and is about 940 in this study. The y_t is the actual value and $\hat{y}_t$ is the forecast value.

While MAPE is one of the most popular measures for forecasting error, there are many studies on shortcomings and misleading results from MAPE[13]. To overcome these issues, there are some other measures proposed in literature, such as mean absolute scaled Error (MASE), max absolute percentage error and symmetric mean absolute percentage Error (sMAPE). We also calculated these measures, but we do not report these results here because the relative performances of the methods for these measures are very similar to those for the MAPE..

Figure 3.0.2 show the forecasting accuracy of the three methods for lead times up to 24 lag which dose not multiplied with 100. No differences between three methods exist and thus, all the lines are overlapped and looked as one line. Also, error rates with all leading times are above 27%, this means that the neural network is working poorly for non-stationary data like as in [6]. Furthermore, ARMA models are not fitted well because of its nonstationarity, so the residuals does not helpful to enhance to forecasting accuracy.

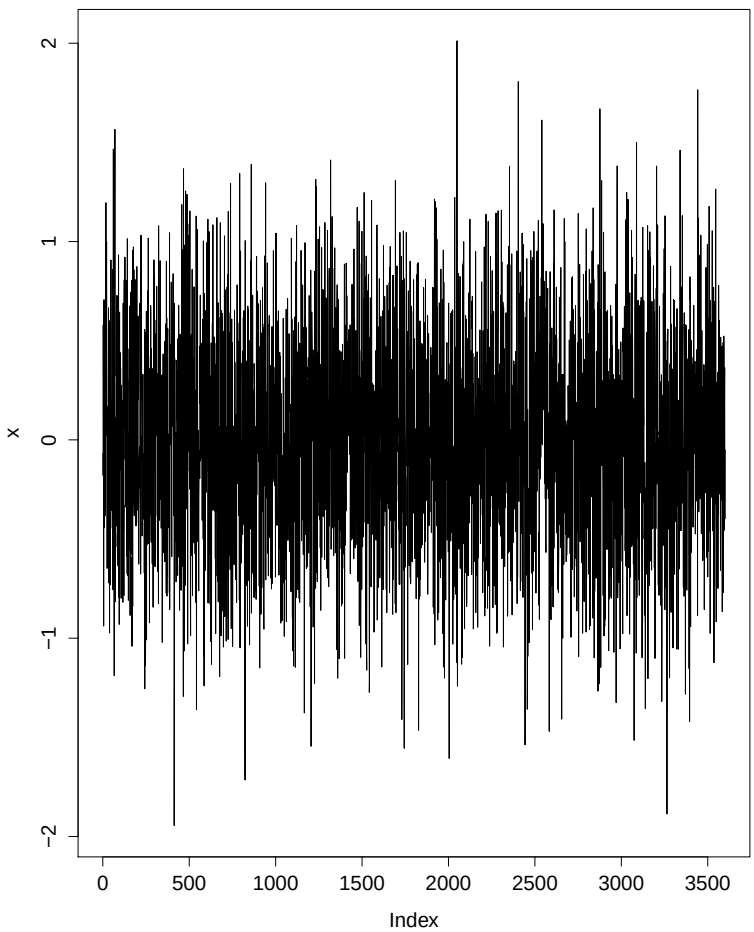

Figure 3.0.1: Simulated data

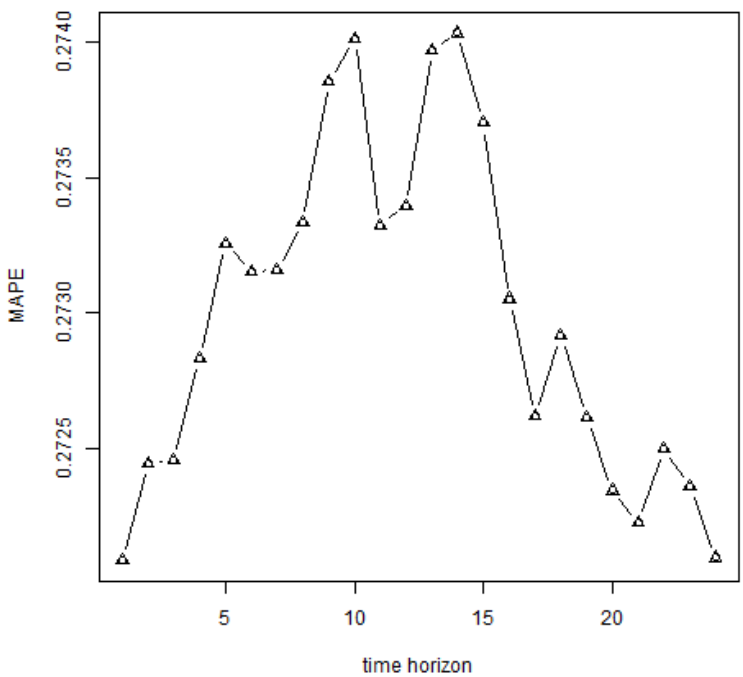

Figure 3.0.2: MAPE of the simulation data

Chapter 4

Post-sample Forecasting

We consider the electric load time series of France for the 39 week period from Monday, 30 March, 2015 to Thursday, 31 December, 2015. The data is obtained from European network of transmission system operators for electricity (https://www.entsoe.eu), which provides hourly electricity data of 35 countries across Europe. We first use 30 weeks of the series (training sample) to estimate ANN parameters and 8 weeks (testing sample) to evaluate the post-sample accuracy of forecasts up to 24 hours ahead (24 number of forecasts). This actually implies that the first 5133 observations are used for the estimation and the remaining 1344 observations are used for evaluation.

Since electricity demand data is non-stationary as shown in figure 4.0.1 and neural network shows better results when using deseasonaliztion and detrending [24], we use the differenced data as following:

$$z_t = (1 - B)(1 - B^{24})(1 - B^{168})y_t \tag{4.0.1}$$

to make the dataset stationary. Then, we can get stationary dataset as shown in Figure 4.0.2. All the methods specified in section 2 are applied to this dataset.

Figure 4.0.3 show that the post-sample forecasting accuracy of 8 weeks for lead time up to a day. We use MAPE as our error summary measure and choose the number of input nodes as three for the neural model, because ANN with more than three input nodes shows not very different results. And we use four input nodes (three of them are z_t and the others are e_t) for hybrid neural models. We also try to fit the model with more than four nodes, but the case with four input nodes shows the best results.

We can see that there are no big differences between these three methods. Especially, performance of ANN and ANN with global ARMA model is almost same except for lead time 6, 15, 16. And ANN with local ARMA model is little higher than other models when lead time up to 7 and over 19. However, the difference is very small, smaller than 0.01%, thus it can be said that all models achieve fairly equivalent results regarding the MAPE. It implies that there is no significant influence of the residuals on the data analysis.

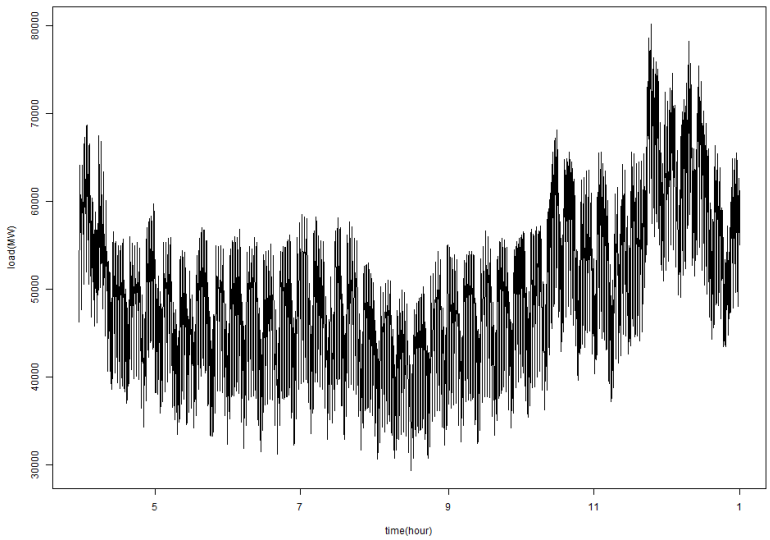

Figure 4.0.1: The time series plot of the hourly load data

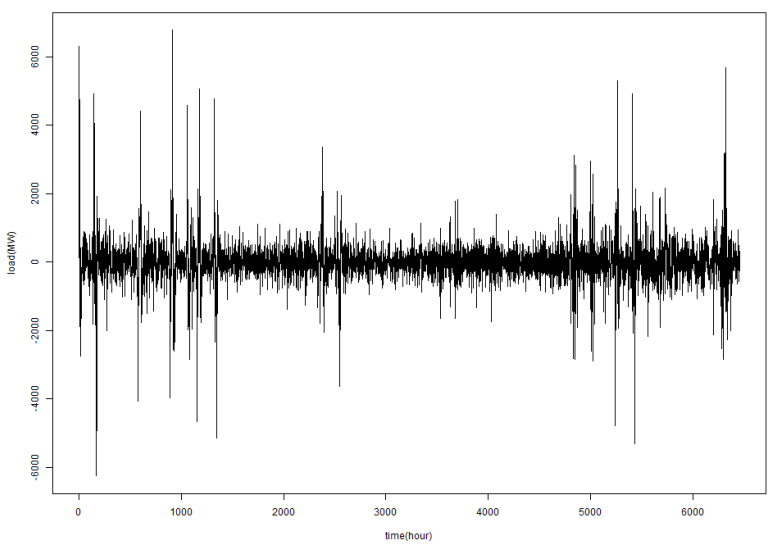

Figure 4.0.2: The time series plot of the differenced load data

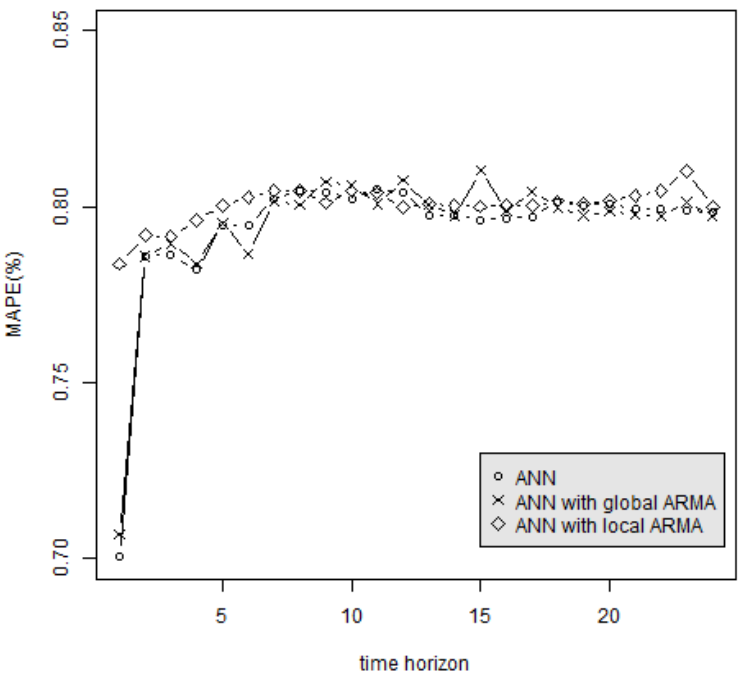

Figure 4.0.3: MAPE results of three methods

Chapter 5

Concluding remarks

In this study, we compared the performance of the three neural network methods to predict electricity demand. We proposed to use two kinds of hybrid models that combine ARMA model and neural network model to forecasting electricity demand. In our proposed model, a time series is considered as a nonlinear function of several past observations and random errors and therefore, ARMA models are used to generate the residuals. We first examined the accuracy of these three models using a simulated data and then the electricity data. In the simulation study, we used random coefficient autoregressive models to see the forecasting accuracy. All of our methods yielded almost the same values regardless of leading time. In real data analysis, we used hourly time series recorded in France. This time series has seasonality at 168 and 24, and trend, thus we used the differenced data to remove the seasonality and trend. However, the MAPE of all three models appeared to be similar to each other. This result indicates that the hybrid neural networks does not perform better than the ANN. If we consider the training time, the traditional neural network is a better choice for the electricity load forecast.

Bibliography

[1] Amjady, N. (2001). Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Transactions on Power Systems*, 16(3):498–505.

[2] Aue, A., Horváth, L., and Steinebach, J. (2006). Estimation in random coefficient autoregressive models. *Journal of Time Series Analysis*, 27(1):61–76.

[3] Bunn, D. W. (1982). Short-term forecasting: a review of procedures in the electricity supply industry. *Journal of the Operational Research Society*, 33(6):533–545.

[4] Bunn, D. W. (2000). Forecasting loads and prices in competitive power markets: the technology of power system competition. *Proceedings of the IEEE*, 88(2):163–169.

[5] Chow, T. and Leung, C. (1996). Neural network based short-term load forecasting using weather compensation. *IEEE Transactions on Power Systems*, 11(4):1736–1742.

[6] Cottrell, M., Girard, B., Girard, Y., Mangeas, M., and Muller, C. (1995). Neural modeling for time series: a statistical stepwise method for weight elimination. *IEEE Transactions on Neural Networks*, 6(6):1355–1364.

[7] Darbellay, G. A. and Slama, M. (2000). Forecasting the short-term demand for electricity: Do neural networks stand a better chance? *International Journal of Forecasting*, 16(1):71–83.

[8] Douglas, A. P., Breipohl, A. M., Lee, F. N., and Adapa, R. (1998). The impacts of temperature forecast uncertainty on bayesian load forecasting. *IEEE Transactions on Power Systems*, 13(4):1507–1513.

[9] Engle, R. F., Granger, C. W., Rice, J., and Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American statistical Association*, 81(394):310–320.

[10] Gonzalez-Romera, E., Jaramillo-Moran, M. A., and Carmona-Fernandez, D. (2006). Monthly electric energy demand forecasting based on trend extraction. *IEEE Transactions on power systems*, 21(4):1946–1953.

[11] Granger, C. W. (1989). Invited review combining forecasts—twenty years later. *Journal of Forecasting*, 8(3):167–173.

[12] Harvey, A. and Koopman, S. J. (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association*, 88(424):1228–1236.

[13] Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688.

[14] Kandil, M., El-Debeiky, S., and Hasanien, N. (2001). Overview and comparison of long-term forecasting techniques for a fast developing utility: part i. *Electric Power Systems Research*, 58(1):11–17.

[15] Khashei, M. and Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with applications*, 37(1):479–489.

[16] Makridakis, S. (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, 9(4):527–529.

[17] McSharry, P. E., Bouwman, S., and Bloemhof, G. (2005). Probabilistic forecasts of the magnitude and timing of peak electricity demand. *IEEE Transactions on Power Systems*, 20(2):1166–1172.

[18] Morita, H., Kase, T., Tamura, Y., and Iwamoto, S. (1996). Interval prediction of annual maximum demand using grey dynamic model. *International Journal of Electrical Power & Energy Systems*, 18(7):409–413.

[19] Pardo, A., Meneu, V., and Valor, E. (2002). Temperature and seasonality influences on spanish electricity load. *Energy Economics*, 24(1):55–70.

[20] Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8):799–805.

[21] Taylor, J. W. and Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, 19(1):57–70.

[22] Temraz, H., Salama, M., and Quintana, V. (1996). Application of the decomposition technique for forecasting the load of a large electric power network. *IEE Proceedings-Generation, Transmission and Distribution*, 143(1):13–18.

[23] Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1):35–62.

[24] Zhang, G. P. and Qi, M. (2005). Neural network forecasting for seasonal

and trend time series. *European journal of operational research*, 160(2):501–514.

국문초록

전력량 예측은 전력 장치를 안전하고 경제적으로 운영하기 위해 중요한 의미를 지닌다. 따라서 예측률을 향상 시키기 위한 다양하 방법들이 제안되었다. 본 논문에서는 이러한 분석방법들을 소개하고 인공 신경망 모형과 시계열 모형을 결합한 새로운 방법을 제안한다. 총 두가지 방법을 제안하며, 첫번째는 전체 데이터에 대한 시계열 모형과 인공 신경망 모형을 이용하는 방식, 두번째는 무빙 윈도우를 이용한 부분적인 시계열모형과 인공 신경망 모형을 이용하는 방식이다. 더불어 시뮬레이션과 프랑스의 전력 수요 자료 분석을 통해 새로운 방법들과 기존 인공신경망 모형을 비교한다. 시계열 모형을 적합하기 위해 차분된 전력 수요량 자료를 이용했으며, 1시간에서 24시간에 이르기까지의 전력량 예측을 통해서 모형의 성능을 비교한다.

주 요 어 : 전력량 예측, 인공신경망 모형, ARIMA 모형

학 번 : 2015-20294