EVALUATION AND ERROR MINIMIZATION OF DYNAMIC SHORT TIME LOAD FORECASTING MODEL WITH CONTROL CHARTS AND PROCESS CAPABILITY ANALYSIS IN THE PRESENCE OF DISTRIBUTED GENERATION

Bahman Jamshidi	Abbas Saghaei	Farhad Kianfar	Mojtaba Naghdi	Reza Vasigh
West Tehran PPDC	Islamic Azad University,	Sharif University	West Tehran PPDC	West Tehran PPDC
Sharif University of	Science and research	of Technology–Iran	Iran	Iran
Technology – Iran	branch –Iran			
BahmanJamshidi@gmail.com	a.saghaei@srbiau.ac.ir	fkianfar@sharif.edu	Naghdi@wtppdc.ir	Vasigh@wtppdc.ir

ABSTRACT

Short time load forecasting (STLF) is a pivotal concept in energy marketing, therefore, regulatory has defined penalty for load forecasting errors, which disturb energy market balance. Distributed generation (DG) has two effects on STLF models: first, in the presence of DG these models inevitably entail non-repeating data as well as load trends, and second, the share of DG in power generation is not constant. Therefore, the STLF model should be evaluated and improved continuously otherwise model accuracy will dwindle gradually.

A lot of STLF models have been developed but there isn't proper tool to assess their accuracy in the presence of DG. For controlling the impact of probabilistic behaviour of distributed generators on load forecasting, West Tehran province power distribution company (WTPPDC) combined dynamic model, statistical control chart and Process capability analysis for continual evaluation and monitoring of the STLF model. In this study WTPPDC have used process capability analysis for evaluation of forecasting capability of model and control charts for detecting out of control error and accumulative bias in prediction in the presence of DG. Quality approach to load forecasting error controlling can help distribution companies to improve their model before forecasting errors reduce their profit and business confidence.

INTRODUCTION

Short time load forecasting models estimating load for each hour of the following days are very important for economic and secure operation of network and energy market balance so the ability of accurate load forecasting is one of the main targets of modern distribution companies. STLF models have been fitted to previous load and future load will be predicted by fitted models. All models for shot time load forecasting which have been developed have their own advantages and constraints, therefore distribution companies should select the model that fit their requirement. The regulatory have defined considerable penalties for load forecasting errors therefore power distribution companies should control and evaluate their load forecasting model to prevent revenue loss.

A lot of industries especially in industrial province like Alborz have their own generators which are used whenever their owners want. Growth of distribution generation (DG) in industrial zone and probabilistic behavior of these generators will increase the level of short time load forecasting error. Some sophisticated models that use artificial intelligence for load forecasting can predict short time future load with acceptable accuracy in non-dynamic condition. But growth of distributed generation will change the load forecasting rules. DG share in total energy generation cannot be estimate very precisely. Many factors like weather change or DG owners decision change DG energy production. It is normal that distribution companies pay more forecasting error penalties.

West Tehran Province Power Distribution Company (WTPPDC) control network engineering team has developed relative simple methods for control of load forecasting errors. In this project, Minitab and MATLAB have been used for statistical analysis.

In this paper we present a short literature review on load forecasting models and tools that have been used for forecasting model evaluation in two following section. In third section we use capability analysis for estimating the rate of intolerable forecasting errors. Finally we introduce useful control charts that could be used for forecasting errors continual monitoring.

SHORT TIME LOAD FORECASTING MODELS

Many methods for load forecasting have been developed in previous years. They are based on various statistical methods such as regression, exponential smoothing, stochastic process, auto-regressive and moving average (ARMA) models, datamining models, and the widely used artificial neural networks (ANN) [1-2].

Each model has its own weakness. Linear regression does not properly represent the complex nonlinear relationships that exist between the load and parameters that influence it due to the lack of self-learning capability. Time series methods such as auto-regressive (AR), moving average (MA) and ARMA models which have been widely used for load forecasting, are vulnerable to dirty data. ANN models are sensitive to unsuitable training data. In all models historical data always is divided into pieces according to different days. After division, load curves of particular days are used for training model and model forecasts 24 hour next day load [3].

Statistical methods like regression can find mathematical relations in historical data for load forecasting but related method could be very complicated and sometimes the results is not satisfactory. Fitting the training data as precisely as possible cannot maximize total accuracy of regression models. Models that their parameters have been estimated by traditional linear regression have low bias but large variance. We can use methods like shrinkage and best subset selection for reducing the number of exogenous and delayed endogenous variables. This methods sacrifice a little bit of bias to reduce the variance of the load prediction, improve total accuracy and make models more understandable. Proper complexity reduction of model can improve model accuracy and stability.
Recently many hybrid methods have been developed for the load forecasting. For example Filik and Kurban have introduced a Short-Term Load Forecasting model with autoregressive and artificial neural network [4]. Khairul Hasan et al have used hybrid approach of Neural Network (NN) along with Particle Swarm Optimization (PSO) for load forecasting [5]. Jain and Satish have used support vector machine (SVM) and time series technique for Short Term Load Forecasting [6]. Shayeghi, Shayanfar and Azimi have combined Continuous Genetic Algorithm (CGA) and optimum large neural networks structure for one-day ahead electric load forecasting [7].
Many distribution companies like West Tehran Province Distribution Company (WTPPDC) have adopted autoregressive model for load forecasting.

FORECASTING MODEL CONTROL AND EVALUATION TOOLS

Several tools have been used for error analysis in forecasting models. In this section we have summarized four important and popular tools:

1- Mean Absolute Percentage Error (MAPE)

MAPE expresses accuracy as a percentage of the error. Because this number is a percentage, it may be easier to understand than the other statistics.

2- Mean Square Error (MSE)

MSE represents the total effect of bias and variance. Outliers have more influence on MSD than MAPE.

3- Coefficient of determination

Coefficient of determination is the proportion of the total variance that is explained by regression, Thus Coefficient of determination is a measure of the explanatory of the model.

4- Durbin-Watson statistics

Durbin-Watson test estimates autocorrelation in residuals. Autocorrelation means that adjacent observations are correlated. In the case of positive first order serial correlation, Durbin- Watson statistics lies significantly below 2. An example of this is the case of estimation of a linear relationship when the actual relationship is nonlinear.
The mean squared error (MSE) which is the second moment of the error and incorporates both the variance of the estimator and its bias is used as criterion of model accuracy. With MSE and MAPE we can measure the total error of forecasting models. Models with lower MSE and MAPE can forecast future load more accurately. These models usually have higher coefficient of determination therefore, these models can explain the variation of forecasted load. MSE and MAPE can't show the trend and distribution of forecasting error and more effective tools for evaluation and control of forecasting error are needed.
Forecasting models can be considered as a process and forecasted load is product of this process. Industrial engineers have developed very powerful tools for process control like process capability analysis and control charts. These methods can evaluate the STLF models accuracy and researcher can use them for any kind of load forecasting models including fuzzy and neural network models.

FORECASTING CAPABILITY OF MODELS

Defined penalties for load forecasting error stimulate distribution companies to minimize load forecasting error. In most cases there are acceptable band with no or little penalty but if load forecasting error exceed band limits, imposed penalty are considerable. In these energy marketing balance strategies, companies need tools that can estimate penalty risk. MSE is only capable to estimate penalty in quadratic and MAPE is only useful in linear penalty strategy. In more complicated load forecasting penalty strategies like the strategy which is used in Iran companies need to estimate penalty risk by statistical methods.
Process capability indices are measure relation between actual performance of a process and its specified performance [8]. These indices are used for estimating capability of process to make products with predefined specification and load forecasting can be treated as a conventional process.
Distribution companies can estimate forecasting capability of selected model in specified condition (for example weekend midnight or weekday morning) as well as overall forecasting capability.
Estimating the process capability of forecasting or forecasting capability divided in following steps:
1- Collect real data and forecasted data of load and measure the error (difference between estimated and distributed energy). If there are different penalty policies for different times (For example peak time or normal time, weekdays or weekend …) error data should be classified and separate forecasting capability

indices should be calculated
2- Check the normality of error distribution. (Whether error probability density is normal or not)
3- Calculate the average and variance of error and fit proper probability density function.
4- Set upper and lower limit for acceptable forecasting error. If penalty policy include many bands or relation between penalty and error can be fitted by separated curves depend on error level, more than one limit pair is needed.
5- Calculate CPL, CPU, CPK and expected overall performance. The method for calculating all parameters could be found in reference 8.
If CPK value is less than one, model is not capable and there will be non-conforming forecast, If CPK is more than one but less than two forecasting model is not really capable and few unacceptable errors will be happened, finally if CPK is more than two the forecasting model is capable and forecasting error is often in acceptable band. Overall performance could be calculated by probability density function integration. Note that stages 3, 4 and 6 are performed by Minitab software in automatic procedure.

MONITORING OF FORECASTING MODEL ERRORS WITH CONTROL CHARTS

In statistical process control theory variation in any process like load forecasting is caused by two sources:
1- Random causes
2- Assignable causes
Random variation is the sum of the multitude of effects of a complex interaction of random causes. When only random variations exist, usually it will not be possible to trace their causes. [8] Assignable variations are normally large in magnitude and its cause should be traced and eliminated. When assignable causes of variation are present, the process is classified as 'unstable', 'out of statistical control' or beyond the expected random variations [8]. Control charts are special tools that can detect assignable variation. There are many types of control charts.
In load forecasting process control there are two dangerous conditions that distribution companies should be aware of them:
1-When distribution companies select a forecasting model for future load estimation they select models that fit their previous load precisely but load demand behavior will change in presence of DG and forecasting model accuracy may degrade over time. Some control chart like CUSUM (cumulative sum) and EWMA (exponentially weighted moving average) can find degradation trend.
2- If some factors that can affect load demand are not included in selected forecasting model especially non repetitive factors like different holidays or social events model cannot estimate load of related time. Some control charts like Shewhart can separate these assignable variations from random variation.
Distribution companies can use control chart for continual monitoring of their load forecasting models. When any assignable variation in load forecasting is found these companies should change or modified forecasting model and reduce their future forecasting error penalties. WTPPDC have used one-sided CUSUM and Zone control chart in this project for monitoring its load forecasting model.
In Zone Control chart, which is modified version of Shewhart control chart -3σ, -2σ, -σ, 0, σ, 2σ, 3σ boundary lines and actual level of forecasting errors are drawn. Any region made by these boundaries has its weight. If any error is placed in specified region we add its weight to predefined variable. If value of this variable exceeds specified number or error level is beyond -3σ and 3σ boundaries the model is out of control and forecasting model should be modified.
One-sided CUSUM control chart can detect accumulated overestimated and underestimated forecasting error. If these accumulated errors exceed predefined boundaries, the model is out of control. CUSUM detail procedure can be found in reference 8.

RESULT AND DISCUTION

In this part of article dynamic behavior of autoregressive model have been studied. Selected data is forecasting error of 12 noon in October and acceptable error is assumed 3.5%. In Iran the acceptable forecasting error for any distribution company is between %2 and %5 depend on other companies forecasting errors therefore calculate forecasting capability have been calculated with different acceptable error bands.
In table 1 result of forecasting capability, in figure 1 histogram of error and fitted normal probability distribution function, in figure 2 Zone control chart and in figure 3 CUSUM control chart of forecasting error have been illustrated. Out of control points in figure 2 have been marked by numbers more than 7 and in figure 3 with small squares and triangles.

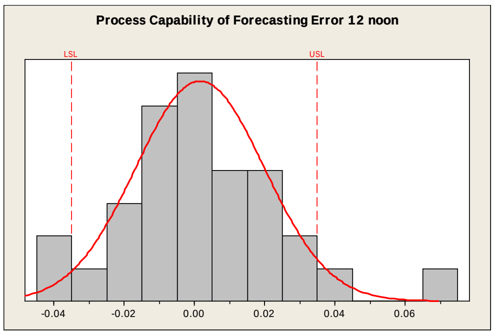

Figure 1: histogram of error and fitted normal probability distribution

Table1: result of forecasting capability

LSL	-%3.5	USL	%3.5
Sample Mean	0.00156	Standard Deviation	0.0183
CPL	0.67	CPU	0.61
CPK	0.61	ZScore	1.58
Overestimation risk	%3.40	Underestimation risk	%2.30

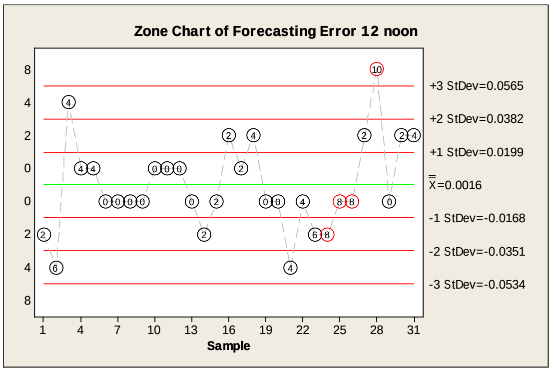

Figure 2: Zone control chart

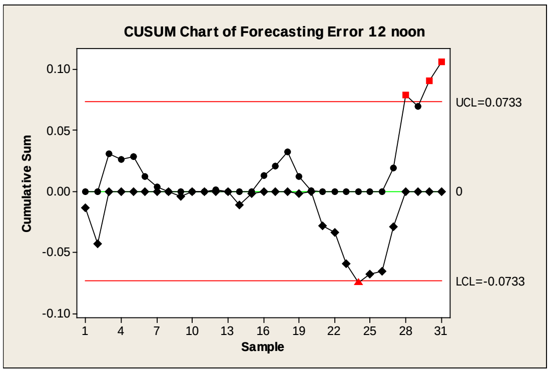

Figure 3: CUSUM control chart

It is obvious that load of four days have been estimated with unacceptable error. The risk of load overestimation is %3.4 and the risk of load overestimation is %2.3.

In figure 2 and 3 it seems that forecasting model has estimated load with acceptable error and adjacent errors have not been correlated until 20 OCT. after 21 OCT out of control points have been occurred in both Zone and CUSUM control charts. Especially in 28 OCT this model error has been more than %6.

Distribution companies engineering teams have investigated all of out of control points and detect the cause of forecasting errors. In this case it is found that 28 OCT is not only weekend but also special day that a lot of religious people have been mourn for holy martyr Imam and some of them left province for pilgrimage. In 20-31 Oct weather temperature has been changed dramatically and model accuracy in this dynamic condition is not acceptable. Engineer team in Alborz Distribution Company can use this information for their model continual improvement. Although first cause is not related to distribution generation, this method can use for any DG related load forecasting errors investigation.

CONCLUSION

Short time load forecasting is one of the most important aspects of distribution system that may affected by DG, therefore distribution companies should evaluate and monitor their load forecasting models over time. Process capability and control charts are very powerful tools for both evaluation and monitoring of short time load forecasting errors. Engineering teams in distribution companies can use these tools for forecasting model selection and improvement. In this paper practical approach of quality control in load forecasting error in presence of DG is introduced.

REFERENCES

[1] Asbury, C., 1975, “Weather Load Model for Electric Demand Energy Forecasting ", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-941111-1116

[2] Papalexopoulos, A.D., Hesterberg, T.C., 1989, "A regression based approach to short-term system load forecasting ", *Proceedings of PICA Conference*, 414-423

[3] Yuhang Yang, Yao Meng, Yingju Xia, Yingliang Lu, and Hao Yu, 2011, " An Efficient Approach for Short Term Load Forecasting", *International Multiconference of engineers and computer scestists*, IMECS, Hong Kong

[4] Ummuhan Basaran Filik, Mehmet Kurban, 2007, " A New Approach for the Short-Term Load Forecasting with Autoregressive and artificial Neural Network Models", *International Journal of Computational Intelligence Research*, ISSN 0973-1873 Vol.3, No.1, pp. 66–71

[5] Md. Khairul Hasan, Mohammad Asif A Khan , Suman Ahmmed and Ahmed Yousuf Saber, 2010, " An Efficient Hybrid Model to Load Forecasting", *IJCSNS International Journal of Computer Science and Network Security*, VOL.10, No.8, August 2010

[6] Amit Jain and B. Satish, 2008, " An Hybrid Approach for Short Term Load Forecasting using SVM and Time Series Technique", *Fifteenth National Power Systems Conference (NPSC)*, IIT Bombay, December 2008

[7] H. Shayeghi, H. A. Shayanfar, G. Azimi, 2009, " Intelligent Neural Network Based STLF", *International Journal of Computer Systems Science and Engineering*, 4:1 pp17-27

[8] John S. Oakland, 2008, *Statistical Process Control*, Elsevier publication, 4:1 pp257-272