Load Pocket Forecasting Software

E. A. Feinberg, D. Genethliou, J.T. Hajagos, B.G. Irrgang, and R. J. Rossin

Abstract—In this paper we describe the load pocket forecasting software that can be used by electric utilities to estimate and forecast the load growth in different service areas. The software builds statistical load models for various service areas (load pockets), estimates weather-normalized loads, estimates the ratios between the actual peak loads and the loads that would happened on designed days (weather normalized factors), and estimates the next year peak load. In particular, the software can be used to calculate the probability distributions of the next year peak loads for different load pockets. The software can be used for area planning purposes. The software contains user's tools to design new load pockets and to modify the existing ones.

Index Terms—Load forecasting, power system planning

I. INTRODUCTION

THIS paper introduces a load pocket forecasting software system developed at State University of New York at Stony Brook and by GMLI Corporation. The software analyzes and predicts peak loads in particular areas called load pockets. This software can be used to improve area planning for electric utilities. It provides the tools to analyze the dynamics of changes of electrical loads at different areas and to estimate the weather-normalized load trends. The utilities spend significant resources to maintain and upgrade their infrastructure and the described system improves decision-making capabilities relevant to capital expenditures. The appropriate investments in the company infrastructure lead to the increase of the system reliability and reduce the probability of blackouts.

The software analyzes the historical weather and load data, estimates weather-normalized electric loads, computes design-day parameters, computes weather normalized factors, estimates trends, and calculates the probability distribution of

the peak load for the next year. The software takes into account that weather conditions in different areas can vary from one area to another. In addition, the weather conditions change from one year to another in the same area. The software estimates the weather-normalized load trends and estimates the deviation from these trends for particular years due to specific weather conditions. The software can be used for one or several load pockets. It demonstrated a high level of accuracy of produced models. The software can be used for standard 50-50% planning scenarios traditionally used by utility companies or it can be used to make decisions to satisfy electric demand with higher probabilities.

The developed software is designed for PCs. It was coded in SAS programming language and contains a convenient graphical user interface.

Most forecasting methods described in the literature fall into one of the two categories: land-use simulation and trending. Land-use simulation method distinguishes customers according to the type of land use [1]-[4]. Trending methods work with historical load data, extrapolating past load growth patterns into the future. The most common trending method is polynomial curve fitting [1], [4]-[6]. The accuracy of this trending method will be improved if horizon year loads are included [1].

One of the major problems with trending methods is that it cannot establish any trend in a vacant area with no load history. A method called Vacant Area Interface (VAI) is used to resolve this problem [1], [4], and [7]. In order to resolve the load switching problem two methods called load transfer coupling (LTC) and LTC Classified Trending (LTCCT) are presented in [1], [4], and [5].

In addition to versions of trend and simulation methods, we use a multiplicative statistical model for the load described in [8], [9] and learning algorithms to estimate the model parameters. This allows us to estimate the probability distributions for the next year peak loads for various areas. These distributions are used to calculate the peak load forecasts.

II. SOFTWARE DESCRIPTION

A. Model Description

The software builds a multiplicative statistical model that takes into account time factors such as the day of the week and the hour of the day, as well as weather factors such as temperature and humidity.

The input data required by the software are hourly weather and load data. Hourly weather data for a five-month period

This research was partially supported by a research grant from NYSTAR (New York State Office of Science Technology and Academic Research).

E. A. Feinberg is with the Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11105 USA (e-mail: Eugene.Feinberg@sunysb.edu).

D. Genethliou is with the Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11105 USA.

J. T. Hajagos is with Long Island Power Authority, Hicksville, NY 11801 USA.

B. G. Irrgang is with Long Island Power Authority, Hicksville, NY 11801 USA.

R. J. Rossin is with Long Island Power Authority, Hicksville, NY 11801 USA.

May through September could be provided by the NRCC (Northeast Regional Climate Center at Cornell University) and by NCDC (National Climatic Data Center). Hourly load data (at the substation/bank/feeder levels) were provided from the SCADA monitoring system.

The developed multiplicative model has the following form:

$$L(t) = L(d(t), h(t))f(w(t)) + R(t) \quad (1)$$

where $L(t)$ is the original load at time t , $d(t)$ is the day of the week, $h(t)$ is the hour of the day, $L(d, h)$ is the daily and hourly component, $w(t)$ is the weather data that include the temperature and humidity, $f(w)$ is the weather factor, and $R(t)$ is the random error. The above multiplicative model was compared with other methods such as time series and conventional regression analysis and gave more adequate results.

In fact, $w(t)$ is a vector that consists of the current and lagged weather variables. This reflects the fact that electric loads depend not only on the current weather conditions but also on the weather during previous hours and days. In particular, the well-known effect of the so-called heat waves is that the use of air conditioners increases when the hot weather continues for several days.

To estimate the weather factor $f(w)$, we used the regression model

$$f(w) = \beta_0 + \sum \beta_j X_j$$

where X_j are explanatory variables, which are nonlinear functions of current and past weather parameters and β_0, β_j are the regression coefficients. The parameters of the model can be calculated iteratively. The convergence of iteration procedures was explained in [8], [9].

Taking into account the weather influence on electric loads improves the accuracy of the forecasts. There are several weather variables that are responsible for changes in load such as dry bulb temperature, wet bulb temperature, dew point, humidity, wind speed, wind direction, sky cover, and sunshine. In the literature every load-forecasting model has a different selection of weather variables. For example a study in Saudi Arabia [10] proposed a mathematical model that focuses on the impact of the temperature and humidity variations on the air-condition load model characteristics. Also an electric load prediction survey published in [11] indicated that of 22 research reports considered 13 made use of temperature only, 3 made use of temperature and humidity, 3 utilized additional weather parameters, and 3 used only load parameters.

An important question in electric load forecasting is which of the weather variables to use in forecasting. A general approach for the weather variable selection is to experiment with models of different complexities and use the simplest model achieving the desired prediction accuracy.

Originally our model included five weather variables, dry bulb temperature (we will refer to it as temperature), dew point, wind speed, sky cover, and sunshine (this variable

indicates when the sun is up and when the sun is down). The use of too many variables though can lead to “overfitting” of the model. To avoid this problem, the final selection of variables used in the model was done by the automatic variable selection procedure known as stepwise regression to get the best linear model.

B. Computational Results

The above model was evaluated on several load pockets in the Northeastern part of the USA. The real load and weather data, for a five-month period May through September, were used for several consequent years. The results shown below are for a particular load pocket for one year.

To illustrate the adequacy of the model we calculate the correlation, the R^2 , and the normalized distance D , of the actual load versus the model. The normalized distance was calculated from the following formula:

$$D = [Var(X - Y) / Var(Y)]^{1/2}$$

that measures relative quadratic distance between two random variables with the same averages. These results are shown below in table I. In this table we show the results of the original model (that means with the five weather variables mentioned before) and the results after stepwise regression. With stepwise regression we were able to reduce the number of parameters into the model and the model remain adequate

TABLE I
STATISTICAL CHARACTERISTICS OF THE MODEL

# of parameters	R	R^2	D
67	0.9755	0.9486	0.2199
34	0.9756	0.9483	0.2196

As shown in the table in both cases the model gave similar results with the main difference in the second case the number of variables used in the model were significantly reduced. The correlation (R) and the normalized distance (D) also indicate good fit of the model. The R^2 is close to 0.95 meaning that our model explained 95% of the variability of the data. The statistical characteristics R , R^2 , and D for stepwise regression are shown in Fig. 1-3.

Fig. 1. Correlation between the actual load and the model using stepwise regression.

Fig. 2. R^2 between the actual load and the model using stepwise regression.

Fig. 3. Normalized distance between the actual load and the model using stepwise regression.

C. Software Capabilities

The software has a graphical user interface. An example of a partition of an area into different load pockets is shown in fig. 4. Using the software the user can design and redesign the load pockets. In addition to the load pockets into which the area was initially partitioned, see Fig. 4, the user can also create temporary load pockets and run the model described above for both permanent and temporary load pockets. The system is enabled in such a way that many users can create custom load areas and run experiments and simulations without making permanent changes to the system. A special user called the system administrator must approve all permanent changes. The administrator may also merge experiments with the rest of the system if the results are good.

For each load pocket, the software calculates the model described in (1). A scatter plot of the actual load versus the model for one of the load pockets is shown in Fig. 5. The scatter plot illustrates how well the model fits the data. The model demonstrated a high level of accuracy in tests for several years of real data. The software also estimates the weather-normalized loads.

Based on the historical weather data and the current model, the software simulates the current load model on the historical weather data (table II). An operator can select which of the past years should be used for the design-day parameter

calculations.

The software analyzes and forecasts two types of peaks for each load pocket. The first type is the so-called pocket peak. The pocket peak is the maximal hourly load during the year. The second characteristic is the pocket load during the hour when the system experienced its peak.

One of the software inputs is historical system peak dates and hours. The software estimates historical peak days for load pockets (table II). These are called pocket peak days and hours. The software then calculates the design-day parameters. For each load pocket, the design-day parameters are calculated for pocket peak days and for system peak days. Different load pocket may have different design days. An example of design-day parameters is presented in tables III and IV.

Once design-day parameters are calculated, the system calculates the ratio of the estimated load on the design day to the actual peak load. This ratio is known as the weather-normalization factor (WNF). The WNFs are useful in explaining what part of the annual load pocket peak should be attributed to the specific weather conditions for that particular year. In addition, the software calculates weather-normalized trends as shown in tables III and IV.

The software contains the probability distribution calculator of the peak load for the next year (fig. 6). The user has two input parameters: the load and the probability. The user may see the probability that the load will exceed a particular value. For any probability, the user can see the load value that the peak demand will not exceed with the given probability. This peak distribution analysis is available for both pocket and system peak data.

The system can generate custom reports after the analysis is completed. The reports can be customized to include model data (fig. 5), load curves, algorithm convergence, historical data (table II), design days, weather normalized factors (tables III and IV), and other software outputs.

Fig. 4. Map showing the partition of the area

Fig. 5. Scatter plot of the actual load versus the model

TABLE II
POCKET PEAK DATES

HH	DATE	LOAD	STRING	TEMP	DP	THI	THI DAILYMAX	THI DAILYMAX TIME	THI_4	THI_24	TEMP DAILYMAX	TEMP DAILYMAX TIME
19	7/8/71	334	0	81	65	75	81	13	78	78	95	95
16	7/19/72	355	1	85	75	80	83	13	81	78	90	90
16	8/28/73	413	2	97	70	85	86	15	85	78	98	98
16	7/9/74	407	3	95	69	84	84	13	84	79	96	96
14	6/19/75	372	0	90	70	81	81	14	80	75	90	90
16	6/25/76	359	1	90	66	80	80	16	79	74	90	90
13	7/19/77	378	4	95	69	84	84	12	83	78	95	95
15	8/17/78	391	4	93	66	82	82	14	82	78	94	94
16	9/6/79	339	0	85	67	78	78	16	77	74	85	85
12	7/21/80	391	2	96	75	86	86	12	83	79	96	96
16	7/9/81	435	3	96	72	85	86	13	85	81	97	97
13	7/19/82	412	4	96	75	86	86	13	84	79	96	96
16	7/15/83	404	3	93	73	87	87	15	86	77	93	93
13	6/11/84	385	4	93	73	84	84	13	82	77	93	93
15	9/6/85	388	3	91	71	82	82	13	82	78	92	92
16	7/7/86	414	2	96	71	85	85	14	85	77	98	98
16	8/18/87	384	2	92	59	80	82	12	81	78	93	93

TABLE III
NORMAL WEATHER AND WNF
POCKET WNF

2002 Design Day	
TEMP	92
TEMP_DAILYMAX	94
TEMP_DAILYMAX_HOUR	14
STRING	3
THI	83
THI_DAILYMAX	83
THI_4	82
THI_24	78

YEAR	MODEL RESULT	LOAD PEAK	WNF	Date	Hour
1998	284	303	0.9373	7/22/98	16
1999	320	352	0.9091	7/6/99	15
2000	381	381	1	6/26/00	16
2001	390	419	0.9308	8/8/01	16

Year	Trend
1998 - 1999	1.1268
1999 - 2000	1.1906
2000 - 2001	1.0184
2001 - 2002	1.0613

TABLE IV
NORMAL WEATHER AND WNF
SYSTEM WNF

2002 Design Day	
TEMP	89
TEMP_DAILYMAX	93
TEMP_DAILYMAX_HOUR	13
STRING	3
THI	81
THI_DAILYMAX	83
THI_4	82
THI_24	78

YEAR	MODEL RESULT	LOAD PEAK	WNF	Date	Hour
1998	282	302	0.9338	7/22/98	17
1999	318	352	0.9034	7/6/99	15
2000	364	314	1.1592	8/8/00	17
2001	389	404	0.9629	8/8/01	18

Year	Trend
1998 - 1999	1.1277
1999 - 2000	1.1447
2000 - 2001	1.0687
2001 - 2002	1.1031

Fig. 6. Peak distribution screen

III. CONCLUSION

This paper describes the software that allows a utility to estimate load trends at different locations called load pockets, to estimate next-year peaks for various load pockets, to calculate weather normalized factors, and to estimate the probabilities of the peak load for next summer in various load pockets. The use of this software improves the decision-making capabilities of the area planning group, improves the utility's capabilities to allocate and manage their capital expenditures, and to improve the reliability of service to their customers.

IV. REFERENCES

- [1] H. L. Willis, Spatial Electric Load Forecasting, New York: Marcel Dekker, 1996.
- [2] R. E. Brown, A. P. Hanson, and D. L. Hagan, "Long range spatial load forecasting using non-uniform areas," in Proc. 1999 IEEE Power Engineering Society Transmission and Distribution Conf., pp. 369-373.
- [3] H. L. Willis, M. V. Engel, and M. J. Buri, "Spatial load forecasting," IEEE Computer Applications in Power, vol. 8, pp. 40-43, Apr. 1995
- [4] H. K. Temraz and V. H. Quintana, "Analytic spatial electric load forecasting methods: a survey," Canadian Journal of Electrical and Computer Engineering, v 17, pp. 34-42, Jan. 1992.
- [5] H. L. Willis, H. N. Tram, and G.B. Rackliffe, "Short-range load forecasting for distribution system planning – an improved method for extrapolating feeder load growth," IEEE Trans. Power Systems, vol.7, pp.1306-1312, Aug. 1992.
- [6] S. Blagajac, S. Krajcar, and D. Skrlec, "A computational tool to support spatial electric load forecasting," in Proc. 2000 of the Mediterranean Electrotechnical Conf., pp. 1116-1119.
- [7] H.L. Willis and H. Tram, "Cluster based V.A.I. method for distribution load forecasting," IEEE Transactions on Power Apparatus and Systems, v PAS-102, pp. 2677-2684, Aug. 1983.
- [8] E.A. Feinberg, J.T. Hajagos, and D. Genethliou, "Load pocket modeling," in Proc. of the Second IASTED International Conf. On Power and Energy Systems, pp. 50-54.
- [9] E.A. Feinberg, J.T. Hajagos, and D. Genethliou, "Statistical load modeling," in Proc. of the Seventh IASTED International Conf. On Power and Energy Systems, pp. 88-91.
- [10] C. A. Belhadj and R.B. Mansour, "Temperature and Humidity Impact on Power System Load Modeling," in Proc. of the IASTED International Conference, Power and Energy Systems, pp. 77-82.
- [11] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation," IEEE Trans. Power Systems, vol. 16, pp. 44 –55, Feb. 2001.

V. BIOGRAPHIES

Eugene A. Feinberg was born in Moscow on April 28, 1954. He received his MS in Applied Mathematics from Moscow University of Transportation in 1976 and Ph.D. in Probability and Statistics from Vilnius University, Lithuania, in 1979. He immigrated to the United States in 1988 and, after holding a faculty position at Yale University, he joined the State University of New York at Stony Brook, where he is currently a Professor at the

Department of Applied Mathematics and Statistics. His research interests include stochastic models of operations research, Markov Decision Processes, and industrial applications of Operations Research and Statistics. He has published more than 60 papers and edited the Handbook on Markov Decision Processes. His research is partially supported by the National Science Foundation, New York State Office of Science, Technology and Academic Research, and private industry. In 1996 he founded Geo Management & Logistics, Inc, a company that has developed several software packages for electric utilities.

Dora Genethliou was born in Cyprus, on January 7, 1978. She received her bachelor degree from the University of Cyprus, in 1999 and her Master degree from Stony Brook University, in 2001. She is now working towards her Ph.D on electric load forecasting.

reactive resource management.

Janos T. Hajagos PE. was born on February 25, 1941, Budapest Hungary. He has a Bachelors of Electrical Engineering from Newark College of Engineering, and a Master degree in Business Administration from Adelphi University. He has 40 years of experience in the Electric Utility business. Currently, he is the Supervisor of Area Planning for the LIPA T&D system. His research interest includes power quality, load forecasting, and

Bryan Irrgang is a Group Leader in the Electric Planning & Forecasting area at KeySpan Corporation where his responsibilities are developing forecasted electric sales, peak demands & revenues for the Long Island Power Authority. His interests include system losses, residential energy use and accuracy assessments.

Richard Rossin was born on Staten Island, New York in the USA on October 13, 1954. He graduated from Brooklyn Technical High School, has a Bachelors of Mechanical Engineering and Masters of Mechanical Engineering from Stevens Institute of Technology, and a Masters in Business Administration from Hofstra University. His employment experience includes Stone & Webster Engineering Corporation and KeySpan. His current position is Program Manager in the Electric Research & Development Department and supports the Long Island Power Authority for electric utility research efforts.