

A REGRESSION-BASED APPROACH TO SHORT-TERM SYSTEM LOAD FORECASTING

Alex D. Papalexopoulos

Timothy C. Hesterberg

Systems Engineering Group
Pacific Gas and Electric Company
San Francisco, California

- ABSTRACT

This paper describes a new, linear regression-based model for the calculation of short-term system load forecasts. The model's most significant new aspects fall into the following areas: innovative model building, including accurate holiday modeling by using binary variables; temperature modeling by using heating and cooling degree functions; robust parameter estimation and parameter estimation under heteroskedasticity by using weighted least squares linear regression techniques; the use of "reverse errors-in-variables" techniques to mitigate the effects on load forecasts of potential errors in the explanatory variables; and distinction between time-independent daily peak load forecasts and the maximum of the hourly load forecasts in order to prevent peak forecasts from being negatively biased. The significant impact of these issues on the accuracy of a model's results was established through testing of an existing load forecasting algorithm. The new model has been tested under a wide variety of conditions and it is shown in this paper to produce excellent results. It is also sufficiently general to be used by other electric power utilities.

Keywords: system load forecasting, regression analysis, ARIMA models, parameter estimation.

I. INTRODUCTION

An accurate System Load Forecasting (SLF) function, used to calculate short-term electric load forecasts, is an essential component of any Energy Management System (EMS). These load forecasts, which can be forecasts of either peak MW requirements during a period (typically 1 hour) or total MWH requirements for that period, are used by system dispatchers and operations analysts to control and to plan power system operations.

Forecasts of hourly loads for up to one week ahead are necessary for scheduling functions such as Hydro-Thermal Coordination and Transaction Evaluation and for short-term analysis functions such as Dispatcher Power Flow and Optimal Power Flow. The Hydro-Thermal Coordination function requires hourly system load forecasts for the next day or the next week in order to establish the hourly schedules for generation resources that will minimize the system operating cost, subject to reliability requirements, operational constraints and other limitations imposed on system operation. The Transaction Evaluation function requires the hourly system load profile for its study period in order to determine the costs and effects of proposed capacity and energy exchanges. The

This paper was sponsored by the IEEE Power Engineering Society for presentation at the IEEE Power Industry Computer Application Conference, Seattle, Washington, May 1 - 5, 1989. Manuscript was published in the 1989 PICA Conference Record.

Dispatcher Power Flow and Optimal Power Flow require load forecasts for their study dates and hours in order to examine the steady-state conditions of the system under a variety of postulated conditions. Based on this information and depending on the operational security objectives of a utility, system operators would then undertake the preventive or corrective actions needed to operate the power system securely and economically.

The accuracy of short-term load forecasts can have significant effects on power system operations, as the economy of operation and the control of the power system may be quite sensitive to forecasting errors. Significant forecasting errors can lead to either overly-conservative or overly-risky scheduling, which can in turn induce heavy economic penalties. Forecasts that are too high may result in the start-up of too many units and unnecessarily high levels of reserves. On the other hand, forecasts that are too low may result in failure to provide the necessary spinning and operating reserves required by power pool agreements. In both cases, forecast errors could result in increased operating costs. Consequently, efforts aimed at the development and implementation of forecasting techniques which reduce the magnitudes of forecast errors can often be justified on the basis of large savings. For the predominantly thermal British power system, it was estimated in 1985 that the increase in operating costs associated with a 1% increase in the forecasting error was 10 million pounds per year [1].

A wide variety of procedures for short term load forecasting have been reported in the literature. These procedures typically make use of two basic models: peak load models and load shape models [2]. In the peak load models, the daily or weekly peak consists of two parts: a weather-independent base load and a weather-dependent component added to the base load [3],[4]. Despite their structural simplicity, these models are not widely used since they do not provide any information about the shape of the load curve. Load shape models describe the load as a discrete time series over the forecasted interval. They can be categorized into two groups: static models and dynamic models. In the static models, the load is considered to be a linear combination of explicit time functions, usually sinusoids, exponentials or polynomials. The coefficients of these functions (model parameters) are estimated through linear regression or exponential smoothing techniques applied to a set of load data for the recent past [5],[6]. A special case of static models is the spectral decomposition model, in which the time functions represent the eigenfunctions corresponding to the autocorrelation function of the load time series [7]. Even though such time series models are structurally simple, they are of limited use since they do not accurately represent weather effects and the stochastically correlated nature of the system load process.

Dynamic load shape models incorporate in their predictions the cumulative effects of such factors as recent load behavior, weather, and random effects. They are of two basic types: autoregressive moving average (ARMA) models [16] and state space models. In ARMA models the load value is composed of two basic components: a deterministic component that represents the periodic behavior of the load curve, and a random

component that represents deviations from the periodic behavior due to deviations of weather from normal and due to random correlation effects [8],[9]. The deterministic component is usually a periodic non-linear function that depends only on the time of the day. The random component is modeled by an ARMA process: its autoregressive terms AR(p) represent the contribution to the current hour's random component as a weighted average of p past values of the random component. Its moving average terms MA(q) represent the contributions to the random component as functions of weighted averages of q deviations from normal conditions. ARMA models can be used to model stationary processes with finite variances. Non-stationary processes can be modeled by differencing the original process. If d is the order of the differencing operation, the resulting process is modeled by an autoregressive integrated moving average model ARIMA (p,d,q). ARIMA models can be converted into state space models (and vice versa [10]), in which the current and relevant past behavior is included in the current state of the system. Load and weather states are updated using Kalman filtering. With state space formulations it is possible to make new forecasts based on results from the previous hour, rather than recomputing the effects of the same past behavior in several previous hours.

The state of SLF technology has advanced considerably over the past decade. Physically meaningful models have been developed that are capable of describing influences due to weather pattern deviations from normal and random correlation effects using few explanatory variables and having modest computational requirements. Despite this progress in SLF algorithms, little work has been published on experience with such algorithms, using actual load and weather data over an extended period of time. Since all SLF models are empirical in nature, extensive experimental testing is the only practical way to objectively evaluate the performance of different SLF algorithms. For an evaluation to be meaningful, such testing should be carried out using data for a sufficiently long period of time under conditions that approximate as closely as possible actual conditions under which the model is used.

In order to assess the adequacy of its existing SLF program, PG&E recently performed a detailed study which addressed the following issues:

- o accuracy of the existing SLF model
- o ease of use and ease of maintenance of the existing SLF program
- o algorithmic changes to improve performance of the SLF model.

As discussed earlier, the first issue is of paramount importance, as an SLF program must produce accurate forecasts if it is to be used to guide the decision-making of power system operators. This requires that the SLF algorithm perform well during all seasons (and especially for peak hours), for periods of unusual weather conditions (such as warm fronts in winter and cool fronts in summer), and also respond accurately and consistently to system changes. These changes include events that are known a priori to influence the load pattern, such as shifts to and from daylight saving time, changes in rate structures, load effects associated with holidays, and special television broadcasts. They also include large random unpredictable disturbances such as the startups or shutdowns of steel mills and synchrotrons.

The second issue is of particular importance in an EMS environment, where system operations staff rather than engineers and analysts are typically responsible

for using the SLF model and updating its parameters. With respect to the third issue, PG&E's evaluation of its existing SLF program yielded insights which provided the basis for the substantial improvements in SLF modeling that are described in this paper.

Performing such an evaluation required the establishment of a database of historical data that was free of both errors and anomalous data. Therefore procedures had to be developed to identify and correct bad data. Anomalous data also had to be identified, and excluded from the forecasting database or be properly modified by taking into account its distinct nature.

This paper describes both PG&E's testing of its existing SLF program and the structure of the significantly improved SLF algorithm which was subsequently developed. Section II briefly describes the methodology of the existing SLF algorithm. Section III describes the procedures developed to evaluate its performance and identify its weaknesses, and outlines the major algorithmic changes which were motivated by the results of the evaluation. Section IV describes the new SLF algorithm, and Section V compares and discusses the results of the two models. Section VI summarizes the major contributions of this paper.

II. EXISTING SLF ALGORITHM

This section briefly describes PG&E's existing SLF algorithm. The algorithm produces forecasts for system loads for the current day and the following day using three models: a peak load forecast model, an hourly load forecast model and a load update model. The peak and hourly models contain separate submodels for the summer period (May through October) and the winter period (November through April). Figure 1 illustrates how the models are used to produce peak and hourly load forecasts.

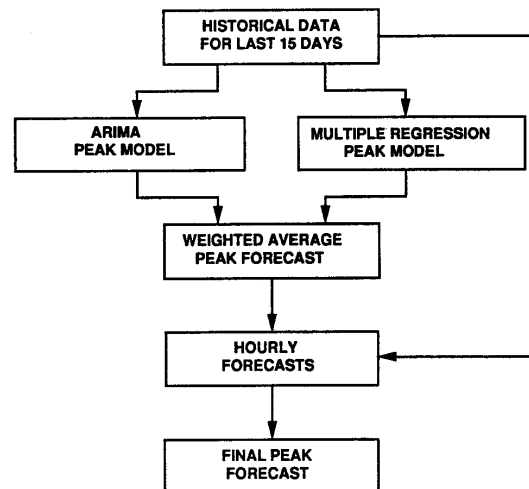


Figure 1: Structure of the Existing SLF Model.

The peak load forecast is computed using two different models: a Multiple Regression (MR) model and an ARIMA model. The results from these two models are combined using a weighted average scheme. The MR model for the winter period determines the peak load as a linear combination of historical peak loads at lags of 1, 2, 6 and 7 days. The MR model for the summer period determines the peak load for the next day using a quadratic function of the current day's actual and next

day's predicted system-weighted average maximum temperatures, the current day's actual peak load, and the peak load 6 days ago. For both periods, the coefficients are determined annually by least squares estimation. The day-of-the-week effect is treated by estimating a different set of coefficients for each day of the week.

The ARIMA ($p=1$, $d=7$, $q=9$) model for the winter period expresses the daily peak load as a linear combination of historical load data and the errors from the previous forecasts. The day-of-the-week effect is eliminated and the data is made stationary by taking the seventh-order difference. The ARIMA ($p=0$, $d=1$, $q=14$) model for the summer period expresses the daily peak load as a function of past daily peak loads, past and forecasted values of the system-weighted average maximum temperatures, and past forecast errors.

The hourly load model forecasts the 24 hourly loads for a given day. These forecasts are based on the forecasted peak and the previous day's hourly loads. The summer model also includes temperature data for the forecasted day and previous day. Once the hourly forecasts have been computed, however, the forecast from the peak load models is discarded and the maximum of these hourly load forecasts is considered to be the peak load forecasted for the next day. For both summer and winter hourly loads there is a separate equation for each of the 24 hours of the day, and for each of the seven days of the week (i.e., 168 equations per model).

The load update model is used to update and improve the load forecast for the remainder of the current day by utilizing the latest available actual load data. The adjusting factor is a function of actual and forecasted hourly loads for the four most recent hours.

III. SLF TESTING

System Characteristics

PC&E's service area covers approximately 94,000 square miles and comprises a wide range of geographical zones, including coastal, inland valley and mountain areas with a large amount of climatic diversity. The inland valley areas contain significant quantities of weather-dependent loads, such as pumping, refrigeration, air conditioning and agricultural irrigation, while the coastal areas are influenced by coastal fog. The yearly profile of daily peak loads contains significant peaks in both the summer and winter, with the annual peak occurring in the summer. The 1985 summer peak of 16,500 MW exceeded the winter peak by approximately 4000 MW. The 1985 ratio of peak to minimum MW load was almost four to one.

The effects of winter weather on load demand are less than those of summer weather in most of the PC&E's service area, as electric heating load constitutes only a small portion of total load. Temperature is the most important weather variable. Other meteorological factors, such as humidity, do not seem to affect PC&E's system load significantly. When temperatures exceed 90°F in inland valley regions, a 10°F change in temperature results in an approximate 200 MW change in peak load. Due to non-uniform geography and climate, temperatures from six areas are combined for use in prediction of the aggregate system load.

Test Procedure

The performance of the existing SLF algorithm was evaluated by using it to reforecast 1985 loads from actual load and temperature data. First, the database of historical loads and temperatures was cleaned to

eliminate obvious data errors. Next, the model's parameters were estimated using load and temperature data from 1979 through 1984. Finally, the SLF function was simulated to run each day at midnight to predict the next day's peak and hourly loads, for all of 1985.

Statistics were developed for the following variables: daily peak load forecast error (forecasted daily peak load minus actual daily peak load) in MW, daily peak load relative forecast error, hourly load forecast error (forecasted hourly load minus actual hourly load) in MW, and hourly load relative forecast error. The mean, standard deviation, minimum value, maximum value, and number of "large" errors were calculated for each of the variables above, for the entire period, and for each season, day type and hour. From a practical point of view, it is "large" errors that undermine a system operator's confidence in an SLF algorithm. PC&E's operations personnel consider a 400 MW peak load forecast error to be the rule-of-thumb dividing line between accurate and inaccurate forecasts, as errors exceeding this threshold may cause substantial problems in load dispatching, reserve allocation, security assessment, and scheduling of large steam-driven generating units.

Results

Table 1 displays the most significant results of the statistical analysis. Examination of the results led to the following conclusions:

- o The summer model is more accurate than the winter model, even though the weather-sensitive load component for the summer period is greater than that of the winter period. This is due to the fact that the winter model does not incorporate weather information.
- o Forecasts (particularly daily peak load forecasts) are consistently low. This negative bias stems from the use of the maximum of the hourly forecasts for the next day as the daily peak load forecast for the next day; such an approach produces a negatively biased daily peak load estimator that is significantly less accurate than the original forecast. (Even though the actual daily peak load is the maximum of the actual hourly loads, the forecasted daily peak load should be larger than the maximum of the hourly load forecasts. When comparing any given hour to the peak, it is clear that the daily peak load can be larger than the hourly load, but not vice-versa. Furthermore, if there is uncertainty as to the hour in which the daily peak will occur, then there is a non-zero probability that the daily peak will be larger than the load at a given hour, as the peak may not occur at that hour.)

Table 1: Error Statistics For Existing SLF Program

	MEAN ERROR MW(%)		STANDARD DEVIATION MW(%)	
	HOURLY	PEAK	HOURLY	PEAK
WHOLE YEAR	-20(-0.15)	-56(-0.44)	346(3.6)	380(3.5)
SUMMER	-30(-0.30)	-90(-0.69)	316(3.2)	346(3.0)
WINTER	-2(-0.02)	-23(-0.17)	375(4.0)	410(3.9)
HOLIDAYS	1253(13.8)	1593(15.2)	820(8.8)	849(8.6)

- o The SLF model performs equally well for all days, except holidays, for which the standard deviation of the relative error is as high as 8.8% for the hourly forecasts and 8.6% for the daily peak forecasts. This is due to the fact that the SLF model does not have a special model for holidays, but simply uses the previous year's loads as forecasts.

Further examination of results not presented in this paper revealed that the model performs poorly during changing weather, which occurs mainly during spring and fall periods, because the summer model does not model the effects of heating loads due to cold fronts and the winter model does not model the effects of cooling loads due to warm fronts. Overall, based on accuracy criteria suggested in the literature [14], the performance of the existing SLF program can be characterized as "fair". According to those criteria, less than 2% error standard deviation can be considered to be "good" performance.

Conclusions

Based on the results obtained from the testing, an effort was undertaken to develop an improved SLF model. This model is described in Section IV. The most significant improvements made include:

- o Modeling the effect of cold temperatures on heating loads by using heating degree functions. These functions are based on physical considerations and an analysis of the historical relationship between loads and temperatures.
- o A more accurate representation of the effect of warm temperatures on cooling loads by using cooling degree functions. These functions are computed in the same way as heating degree functions.
- o Modeling of holiday effects by using binary variables (i.e., variables which take on values of either 0 or 1), for major and minor holidays and appropriate days surrounding those holidays.
- o Use of a single combined model instead of separate summer and winter models. In this model differences between summer and winter load patterns are represented using sinusoidal terms and a daylight saving time variable.
- o Preserving the distinction between the daily peak forecast and the maximum of hourly load forecasts, in order to reduce the negative bias due to the uncertainty as to when the daily peak will occur.
- o Capturing the growth and seasonal variations in loads by the inclusion of trend, daylight saving time and time-of-year terms in the regression model.
- o Correcting for consistent under- or over-estimation by adjusting the forecasts based on an exponential smooth of past errors.
- o Elimination of the ARIMA model. Analysis of the results showed that the regression and ARIMA models largely duplicated each other. Although the weighted average of their results produced a more accurate forecast than either model separately, the amount of improvement was so small that the ARIMA model could be eliminated with no significant deterioration in forecast quality. This change substantially increased the simplicity of the SLF algorithm.

Taken together, the above improvements resulted in an SLF algorithm that is accurate, robust and adaptive to changing conditions. This new algorithm constitutes a step forward in the SLF technology, as many of the concepts utilized are innovative. It is therefore fully described below.

IV. NEW SLF ALGORITHM

The model's major new aspects fall into the following areas: innovative model building, including

temperature modeling and holiday modeling, robust parameter estimation including "reverse errors-in-variables" techniques and parameter estimation under heteroskedasticity, and preserving the distinction between daily peak load forecasts and the maximum of the hourly load forecasts.

The model is based on a linear regression formulation. A number of statistical reasons motivated this choice, including ease of modeling holidays, holiday weekends and other special effects; ease of modeling different days of the week with different error variances (heteroskedasticity); ease of handling missing data; and reduction of the effects of outliers (i.e., large load forecast errors) that adversely impact the quality of the parameter estimates. All of these issues were deemed to be significant for improving the SLF performance as indicated by the detailed analysis of the test results. In particular, with a regression-based approach:

- o Holidays and other special effects can be modeled using binary variables.
- o Days of the week having different error variances can be handled by using appropriate independent variables for each day of the week and using weighted regression.
- o Observations for which data is missing can be deleted (as is done in the new SLF model), or the missing values can be estimated.
- o The impact of outliers can be reduced by utilizing a weighted regression method with weights depending on forecasted errors.

Other models, particularly ARIMA models, are less flexible in these aspects. In addition to these statistical reasons, there is also a practical reason for choosing linear regression models: linear regression capabilities are well-integrated parts of the statistical packages that are required for effective model selection and parameter estimation.

Overview

The improved model produces an initial daily peak forecast and then uses this initial peak forecast to produce initial hourly forecasts. It then uses the initial daily peak forecast, the maximum of the initial hourly forecasts, the most recent initial peak forecast error, and an exponential smooth of errors as variables in a regression model to produce an adjusted peak forecast. Through this adjustment, the autocorrelation structure of random effects, which is not modeled in a pure autoregressive approach, is partially captured.

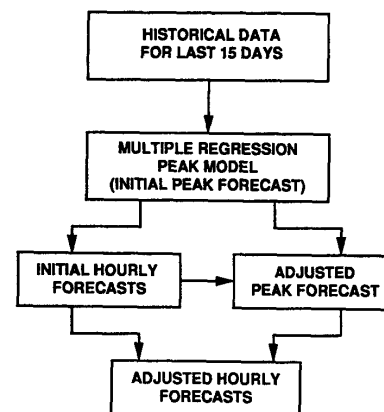


Figure 2: Structure of the New SLF Model.

Finally, any hourly forecasts that are larger than the adjusted peak forecast are reduced to the value of the adjusted peak forecast. Figure 2 illustrates the structure of the new SLF model, which is described in detail below.

Initial Daily Peak Forecast

The initial peak forecast model produces a peak forecast for tomorrow based on the peak loads today and six days ago (and three days ago, for Monday forecasts), the temperatures for those days, the day of the week, tomorrow's forecasted temperature, the time of year, and holiday information. This model is a regression model of the form:

$$\hat{P}(i) = a(d) + b_1(d) \cdot (1 + 365 \cdot Y) + \sum_{k=1}^2 [b_2(d, k) \cdot \sin(k\theta) + b_3(d, k) \cdot \cos(k\theta)] + \sum_{\ell=1,3,7} b_4(d, \ell) \cdot P(i-\ell) + \sum_{\ell=0,1,7} [b_5(d, \ell) \cdot F_{d-\ell}^C(T_{i-\ell}) + b_6(d, \ell) \cdot F_{d-\ell}^H(T_{i-\ell})] + \sum_{k=1}^{60} b_7(k) \cdot I\{i=\text{hol}(k)\} \quad (1)$$

where:

i = day to be forecasted, $i=1,2,\dots,365$

$\hat{P}(i)$ = initial peak forecast for day i

$P(i)$ = actual peak load for day i

Y = year, ($Y=1986, 1987, \dots$, or $Y=1, 2, \dots$)

d = day of week (Monday = 1, Tuesday = 2, ...)

$\theta = 2\pi/365$

ℓ = number of days to lag

$F_d^C(T_i)$ = Cooling degree function for temperature T_i and day of week d (i = day to be forecasted)

$F_d^H(T_i)$ = Heating degree function for temperature T_i and day of week d (i = day to be forecasted)

$\text{hol}(k)$ = day of year for holiday k (this depends on the year)

$I\{\cdot\}$ = indicator function, defined by:

$$I\{i = \text{hol}(k)\} = \begin{cases} 1, & i = \text{hol}(k) \\ 0, & \text{otherwise} \end{cases}$$

$a(d)$ = estimated constant term

$b_j(d, m)$ = estimated parameters ($j=1, 2, \dots, 7$), and

m an additional index for b_2, \dots, b_6 .

Parameter Estimation

The parameters $a(d)$ and $b_j(d, m)$ are chosen annually to minimize the weighted sum of squares of the differences between the actual historical loads and the forecasts (i.e., the forecast errors).

$$J = \sum_i W(i) \cdot [\hat{P}(i) - P(i)]^2 \quad (2)$$

subject to linear constraints on the parameters. These constraints are an important part of the model. The number of degrees of freedom in the regression equation is far fewer than implied by equation (1), because many of the parameters are constrained. For example, the only coefficients which differ for $d=2, 3, 4, 5$ (Tuesday-Friday) are the intercept terms $a(d)$; the temperature, time-of-year, and previous load effects are modeled to be equal for these days. Other constraints are imposed based on either physical considerations or numerical tests. If, for example, the numerical difference between two parameters was

judged to be insignificant, that difference was eliminated from the model using an equality constraint. This is equivalent to reducing the number of independent variables in the model. The model is mathematically (though not notationally) parsimonious.

The weighted least squares linear regression [15] for parameter estimation is done using a three-step regression process:

1. Unweighted least squares regression with the most recent six years of load and temperature data, (in our case, from 1979 through 1984)
2. Weighted least squares regression with weights based on step 1, using the same data
3. Weighted least squares regression with weights based on step 1 and holiday parameters fixed from step 2, using the most recent three years of data (in our case, from 1982 through 1984).

The decision regarding how much data to use to estimate parameters involves a tradeoff between bias and variance. The more data used the more accurate parameter estimates can be, but this implies that older data will be used, including data which does not reflect changes in load patterns over time. The optimal tradeoff is different for holiday loads (where a small amount of data is available in any given year, since any given holiday occurs only once) than it is for non-holidays.

This three-step process makes parameter estimation robust by minimizing the effect of outliers, and improves holiday parameter estimation. The weights are computed using the forecasted errors for each day (the difference between the actual load and the forecasted load based on a regression with that day deleted), using the formula:

$$W(i) = \min(1, 400/|e(i)|), \quad e(i) = \hat{P}(i) - P(i) \quad (3)$$

Any day with a forecasted error larger than 400 MW is de-weighted, with weight inversely proportional to the size of the error.

These weights implement a one-step approximation to regression using Huber's M-estimate [11], [12]. Huber's estimate has the property that the influence of outliers is minimized. Other M-estimates can be used for robust regression, using either one-step or iterative parameter estimation methods. M-estimates can be used to provide optimal estimates for certain error distributions [11].

The second regression step uses six years of data to estimate all parameters, but saves only the holiday parameters. The third step uses the three most recent years of data only, with the relative magnitudes of holiday parameters fixed at the values estimated in the second step. This is done by defining a new "holiday effect" variable as the weighted sum of the holiday binary variables:

$$\text{HOL}(i) = \sum_{k=1}^{60} b_7(k) \cdot I\{i=\text{hol}(k)\} \quad (4)$$

This variable is used as an independent variable in subsequent regression models (all of which use only three years of data), in lieu of other holiday variables.

Temperature Forecast Parameter Estimation

The historical data used for the parameter estimation typically include actual recorded

temperatures. In practice, of course, the actual temperatures for the next day are not available. Parameters estimated using actual temperatures when applied in actual operation, produce load forecasts as if temperature forecasts were perfectly precise. This results in an undue reliance of load predictions on the temperature forecasts. A technique that reduces the sensitivity of load predictions to temperature forecast errors is described next.

Temperature forecasts can be interpreted as the actual temperature plus a noise component v (the temperature forecast error). This suggests that in the parameter estimation process it is appropriate to add similar noise, and to replace the temperature for the day to be forecasted, T_i , with $T_i + v_i$. This noise can be modeled as a random variable with zero mean and the same standard deviation σ_v as actual temperature forecasts. The probability density function of the added noise is of the form:

$$f_v(v) = \frac{1}{2} \cdot \delta(v - \sigma_v) + \frac{1}{2} \cdot \delta(v + \sigma_v) \quad (5)$$

where $\delta(\cdot)$ is the impulse function, i.e., the noise takes the values $\pm\sigma_v$ with equal probability. This concept is implemented deterministically by creating two pseudo-observations out of each original observation, i.e., two days out of every day of data; one with the temperature adjusted upwards by σ_v and the other adjusted downwards by the same amount. Parameters are then estimated based on this doubled data set, using weighted least squares as described above. The result is that cooling and heating terms which use the forecasted temperature have smaller parameters, so the model is less sensitive to temperature forecast errors.

This doubling technique minimizes a weighted sum of squares of the form:

$$J = \sum_i W(i) \cdot \{ [\hat{P}(i, T + \sigma_v) - P(i)]^2 + [\hat{P}(i, T - \sigma_v) - P(i)]^2 \} \quad (6)$$

This technique is similar to an implementation method for ridge regression, though here the relationship between temperature and load is nonlinear. This parameter estimation problem is the converse of the "errors-in-variables" problem. There, the available data for the independent variables has noise, and the problem is to estimate the regression coefficients that are valid for independent variables without noise; those coefficients would be best for forecasting, if future values of the independent variables would not be noisy. But in our case, the past data for independent variables does not have noise, while the future data will.

Initial Hourly Forecast

The initial hourly forecast model produces hourly forecasts for tomorrow based on the initial peak forecast for tomorrow, the hourly loads of today, the error history for the initial peak forecasts, and a subset of the variables used for the peak forecast. The regression model is of the following form:

$$\begin{aligned} \hat{P}(i, h) = & a(d, h) + b_1(d, h) \cdot (1 + 365 \cdot Y) + \sum_{k=1}^2 [b_2(d, h, k) \cdot \sin(k\theta) + b_3(d, h, k) \cdot \cos(k\theta)] + \\ & \sum_{\ell=1}^2 b_4(d, h, \ell) \cdot P(i - \ell) + \sum_{\ell=0}^1 [b_5(d, h, \ell) \cdot F_{d-\ell}^c(T_{i-\ell}) + b_6(d, h, \ell) \cdot F_{d-\ell}^h(T_{i-\ell})] + \\ & b_7(h) \cdot \text{HOL}(i) + b_8(h) \cdot \text{DL}(i) + b_9(h) \cdot \text{HOL}(i) \cdot \text{DL}(i) + b_{10}(h) \cdot P(i - 1, h) + \\ & b_{11}(h) \cdot P_s(i - 1, h) + b_{12}(h) \cdot \hat{P}(i) + b_{13}(h) \cdot e(i - 1) + b_{14}(h) \cdot E_{es}(i - 1) \quad (7) \end{aligned}$$

where:

$\hat{P}(i, h)$ = initial hourly forecast for hour h of day i
 $\text{HOL}(i)$ = estimated holiday effect for the peak demand for the day i (computed from equation (4))
 $\text{DL}(i)$ = daylight saving binary variable for day i
 $P(i, h)$ = actual load for hour h of day i
 $P_s(i, h)$ = smoothed hourly load, given by:

$$P_s(i, h) = (P(i, h - 1) + 2 \cdot P(i, h) + P(i, h + 1)) / 4$$

 $E_{es}(i)$ = exponential smooth of errors, given by:

$$E_{es}(i) = \lambda \cdot e(i) + (1 - \lambda) \cdot E_{es}(i - 1)$$

 $b_j(d, h, m)$ = estimated parameters ($j = 1, 2, \dots, 14$),
 and m an additional index for b_2, \dots, b_6 .

The "optimal" value of λ was determined through testing to be 0.98. The other variables are defined in the initial peak forecast model. Coefficients are estimated using weighted least squares regression. Additional linear constraints are imposed based on either physical considerations or numerical results to make the model parsimonious. The use of a single holiday variable reduces computational requirements by eliminating many variables and by reducing the number of years of data required from six to three. Holiday effects are relatively strong during peak load periods (summer afternoons, and winter evenings); this is captured using a holiday/daylight saving time interaction.

Adjusted Peak and Hourly Forecasts

The adjusted peak model produces a new peak forecast, incorporating the initial peak forecast, the maximum of the hourly forecasts, and past errors in the initial peak forecasts. The forecasting equation is:

$$\begin{aligned} P(i) = & a + b_1 \cdot \hat{P}(i) + b_2 \cdot \max\{\hat{P}(i, h); h = 1, \dots, 24\} + \\ & b_3 \cdot e(i - 1) + b_4 \cdot E_{es}(i - 1) \quad (8) \end{aligned}$$

where, the peak forecasted error, $e(\cdot)$, is defined in (3) and all other variables are defined in (7).

The coefficients are estimated using weighted least squares regression. This model performs four functions. It chooses a weighted average of the initial peak and the maximum hourly forecasts to use as a new peak forecast. The constant term adjusts the forecast upward to correct for the bias introduced by the use of the maximum of the hourly forecasts. The term for the previous day's error provides a correction based on the first-order autocorrelation of forecasted errors from the initial peak regression. Finally, the use of the exponential smooth corrects for forecasts which have been consistently too high or low over the recent past (roughly the last two months). With this adjustment the recent load behavior is reflected in the forecasts; the model reacts quickly to changing overall load levels but does not attempt to adjust in real-time to changes in relative contributions of specific types of loads.

The adjusted hourly model produces the final hourly forecasts. These are the same as the initial hourly forecasts, excepts that any forecasts which exceed the adjusted peak forecast are reduced to the value of the adjusted peak forecast.

Temperature Modeling

Temperature effects are modeled using heating and cooling degree functions which are based on physical considerations and on graphical analyses of the

historical relationships between loads and temperatures. The cooling degree function is zero until a threshold temperature $T_{1c}(d)$ is reached, increases quadratically until a second threshold $T_{2c}(d)$ is reached, and finally increases linearly. The heating effect function has the same form, but increases as temperatures decrease. I.e.,

$$F_d^c(T) = \begin{cases} 0 & , T < T_{1c}(d) \\ (T - T_{1c}(d))^2 / 2(T_{2c}(d) - T_{1c}(d)) & , T_{1c}(d) \leq T < T_{2c}(d) \\ T - (T_{1c}(d) + T_{2c}(d)) / 2 & , T \geq T_{2c}(d) \end{cases}$$

$$F_d^h(T) = \begin{cases} 0 & , T > T_{1h}(d) \\ (T_{1h}(d) - T)^2 / 2(T_{1h}(d) - T_{2h}(d)) & , T_{2h}(d) < T \leq T_{1h}(d) \\ (T_{1h}(d) + T_{2h}(d)) / 2 - T & , T \leq T_{2h}(d) \end{cases}$$

These functions reflect a physical model with the following characteristics:

- Below some lower threshold there is negligible cooling load.
- As temperatures increase, more individuals and offices turn on air conditioners, trying to keep temperatures at their own individual thresholds. Thus the air conditioning load grows, and the rate of growth increases as well.
- Finally, there is some upper threshold beyond which the load increase is linear, once all individual thresholds are met.

The reverse occurs for heating loads.

The thresholds were chosen based on graphical analysis of historical data using a spline smooth of loads and temperatures, disaggregated by weekday and weekend [13]. Graphical and numerical results showed that the air conditioning thresholds should be lower during the week than on the weekend.

The temperature variable used in the degree day calculations is a weighted average of the daily maximum temperatures at six locations in the PG&E service area.

Holiday Modeling

Holiday loads are generally lower than other loads, and this effect spills into surrounding days as well. A good SLF algorithm must model these effects as accurately as possible. It must also ensure that low loads on holidays do not adversely affect forecasts on subsequent days which use the holiday load data as independent variables in those forecasts.

Holidays are modeled in the new SLF model using binary variables. Binary variables are defined for major holidays (New Year's Day, Presidents' Day, Easter Sunday, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day), for minor holidays (Super Bowl Sunday, Columbus Day, Veterans Day), and for appropriate days surrounding those holidays. The coefficients associated with these binary variables are used to adjust forecasts for the same days in succeeding years. These effects are added to create a single variable which measures the "holiday effect"; this effect is negative on holidays and holiday weekends and is positive on the first and sixth days after holidays. For example, the coefficients estimated for Memorial Day and the Tuesday following Memorial Day are -1537 and 1106, respectively, adjusting the Memorial Day forecast downward by 1537 MW, and the following day forecast upward by 1106 MW. The latter adjustment compensates for the difference

between a Memorial Day load and a normal Monday load when producing the forecast for the Tuesday following Memorial Day.

Linear constraints are used to reduce the number of parameters to be estimated, by combining terms judged to be similar. For example, the coefficients for 7 days after Memorial Day and 7 days after Labor Day are constrained to be the same.

V. COMPARISON OF RESULTS

This section compares some of the results of the old and the improved SLF algorithms.

Table 2 presents the results of the statistical analysis performed on the 1985 peak load forecast errors produced by the two models.

Table 2: Peak Load Forecast Error Statistics for Both Models

	EXISTING MODEL		IMPROVED MODEL	
	MEAN ERROR MW(%)	STD DVN MW(%)	MEAN ERROR MW(%)	STD DVN MW(%)
WHOLE YEAR	-56 (-0.44)	380 (3.50)	-12 (-0.10)	193 (1.62)
SUMMER	-90 (-0.69)	346 (3.00)	-12 (-0.10)	205 (1.66)
WINTER	-23 (-0.17)	410 (3.90)	-11 (-0.09)	181 (1.57)
MONDAY	-32 (-0.25)	641 (6.05)	-5 (-0.04)	241 (1.94)
TUES-FRIDAY	-41 (-0.32)	262 (2.16)	11 (0.09)	171 (1.37)
SAT-SUN	-99 (-0.89)	278 (2.59)	-61 (-0.51)	203 (1.83)
LARGE ERRORS (ERROR > 400 MW)	60		14	

As can be seen, performance improved substantially for all seasons and day types. The standard deviations of the peak load forecasting errors for the entire year, for the summer and for the winter were reduced by 49%, 40% and 56%, respectively. The number of large peak forecast errors (>400 MW) was reduced by 77%; on the average, large errors were reduced from one per week to one per month. Also the negative bias was virtually eliminated from the forecasts, with the average forecast error reduced drastically for all seasons and day types. For the summer forecasts, for example, it was reduced from -90 MW to -12 MW, which amounts to an improvement of 87%.

Figure 3 demonstrates the distinction discussed earlier, between a time-independent peak load forecast and the maximum of the day's forecasted hourly loads.

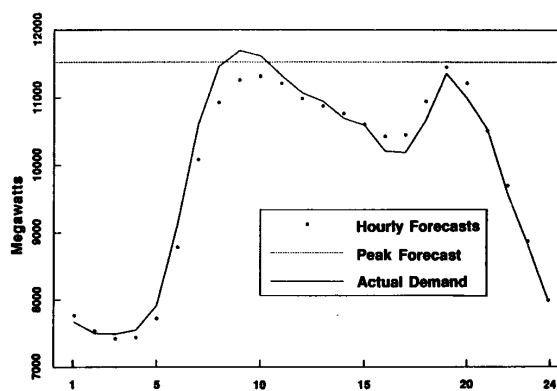


Figure 3: Time-Independent Peak Load Forecast vs Maximum Hourly load Forecast.

It illustrates the time-independent forecasted peak load, the forecasted hourly loads and the actual demand for the day of March 8, 1985. The maximum of the hourly forecasts is smaller than the peak forecast by 53 MW; using the maximum hourly forecast as a peak forecast (as the existing SLF model does) would cause a negative bias of the same amount. (This negative bias would be compounded to much larger values if the 1-day model were called iteratively to produce multiple-day forecasts. For example for weekly forecasts, a negative daily bias of 53 MW could be compounded to as much as 371 MW, unless the problem was properly addressed.)

Substantial accuracy improvements were achieved in the hourly forecasts as well. Figure 4 presents the standard deviations of the hourly (and peak) relative forecast errors (excluding holidays) as a function of the hour of the day for the 1985 forecasts. The differences between the two models are especially pronounced during peak hours. This is important, since operators are much more concerned with the forecasts for the peak hours than those for the off-peak hours.

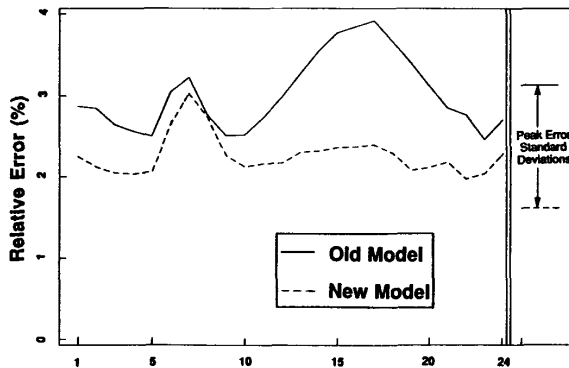


Figure 4: Standard Deviations of Errors By Hour.

To illustrate the type of forecasts that are obtained, the actual and forecasted daily peak loads for the month of April of 1985 are presented in Figure 5. This month was characterized by rapidly changing weather (especially for the days of April 12, 13, 14, 15 and 29) and by the presence of holidays and holiday weekends (April 7 was Easter Sunday). As can be seen, the old model did not perform well under these

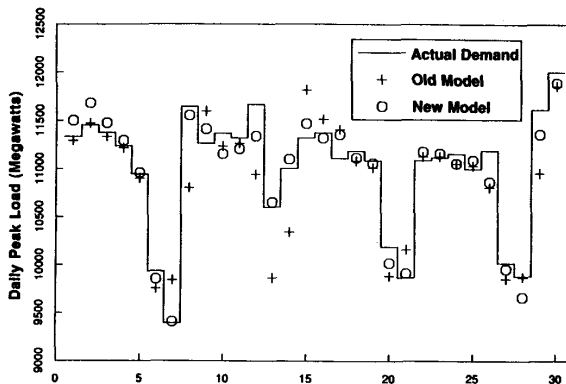


Figure 5: Actual and Forecasted Daily Peak Loads for April of 1985.

circumstances but the improved model's forecasts followed the actual load pattern very accurately and reliably throughout the forecasted period. This high level of performance is due mainly to the more accurate holiday modeling and temperature modeling.

Discussion

The results presented above and similar results obtained with data for 1982 provide compelling evidence of the excellent performance of the new SLF model. Although some parameter tuning is required, as well as some judgment specific to a utility's unique situation, in order to set up the constraints, the techniques presented in this paper are generally applicable and can be used by other electric power utilities to improve their SLF algorithms. The major advances in system load forecast modeling incorporated in PG&E's new SLF algorithm include the following:

- o Accurate modeling of special events, including holidays and holiday weekends. The effects of special events spill into surrounding days and can adversely impact the quality of the future forecasts unless they are explicitly handled by the SLF algorithm. These effects are modeled using binary variables which can be easily incorporated in regression-based models.
- o Accurate temperature modeling. For electric power systems with a substantial weather sensitive load component, effective temperature modeling is very important. Extensive data analysis generally needs to be undertaken in order to develop models that can effectively capture the historical relationship between loads and temperatures. Temperature effects are incorporated in PG&E's new SLF model using heating and cooling degree functions.
- o Robust parameter estimation. Outliers can adversely impact the quality of parameter estimates. Weighted least squares linear regression techniques are effective in minimizing the effects of such outliers. (The weights should depend on what each utility considers to be the dividing line between accurate and inaccurate forecasts.) Weighted regression can also minimize the impact of heteroskedasticity by using appropriate independent variables for each day of the week.
- o Effective use of "reverse errors-in-variables" techniques. The results presented above were obtained using actual temperatures. In practice, of course, the actual temperatures for the future are not available. To shield the model from any large temperature forecast errors it is important (especially for electric power utilities with climatic diversity and rapidly changing weather conditions) to effectively use "reverse errors-in-variables" techniques. These techniques reduce the sensitivity of load forecasts to temperature forecast errors.
- o Distinction between time-independent daily peak load forecasts and the maximum of the hourly load forecasts in order to prevent peak forecasts from being negatively biased. This is especially important for applications that rely on accurate daily peak forecasts.
- o Finding that available ARIMA models are not flexible enough to handle SLF modeling requirements, in particular: modeling holidays and other special events, treatment of actual temperatures to mimic forecasted temperatures, handling different behavior on different days of the week, modeling heteroskedasticity, handling missing data and minimizing the

effects of outliers. (Incorporation of robust ARIMA methodology into statistical packages would meet the latter requirements.)

- o Accurate model fitting. It is more important for good performance that the algorithm selected fit the historical data well, than that it be theoretically attractive. Accurate and efficient model fitting is aided by the use of statistical packages which provide integrated data management, model building and numerical and graphical model analysis capabilities. It also requires procedures to identify and correct bad, anomalous, and missing data.

PG&E's new SLF model is also suitable for future extensions that will be required in the short-term load forecasting area. Two needs that have already been identified are for multiple day load forecasting capabilities and multiple area load forecasting capabilities.

Some EMS scheduling and analysis functions (for example Hydro-Thermal coordination functions) require multiple day (typically weekly) forecasts. The new SLF algorithm described in this paper has the potential to be used for this purpose: it can be called iteratively to produce these multi-day forecasts because the negative bias in the peak load forecasts has been virtually eliminated. Also, for some electric power utilities with a wide range of geographical zones and climatic diversity, multiple area forecasts might be needed in order to adequately use EMS operations planning functions. For example, power flow-based contingency studies require more than the system load - they require bus load forecasts. With a single system load forecast as input to such studies, all conforming bus loads must increase or decrease together. With multiple area forecasts, flexibility is provided so that conforming loads in each geographical zone can follow their own distinct load patterns. In this case the service area can be divided into multiple areas, each with its own load forecast. PG&E's new SLF algorithm has been kept sufficiently simple, so that migration to area load forecast, is feasible using the same basic algorithms described in this paper.

VI. CONCLUSION

The System Load Forecasting model is a critically important decision support tool for operating the electric power system securely and economically. Conducting in-depth tests of the performance of an SLF can lead to the identification of model weaknesses and areas for improvement, and consequently to large savings. Those areas that need improvement can be associated with system-dependent peculiarities or with more general algorithmic aspects of an SLF. In PG&E's case, the testing of its existing SLF program provided insights that led to the development of a significantly improved SLF algorithm. The testing provided evidence that the modeling of holidays, holiday weekends and other special effects, the handling of heteroskedasticity and missing data, and the reduction of the effects of outliers and large temperature forecast errors are quite important for accurate load forecasts. Since ARIMA models are not flexible in many of these areas and in addition are complicated and difficult to develop, PG&E's new SLF model is based on a linear regression formulation. This improved simpler model was tested extensively using two years of real data. It was found to produce excellent load predictions under a wide variety of power system operating conditions. It is parsimonious, robust, adaptive to changing conditions, consistent with physical intuition and capable of incorporating in its forecasts the cumulative effects of such factors as

recent load behavior as well as weather and random effects.

Several important modeling issues that have been discussed are of general applicability and can be used by other electric power utilities. They include:

- o Accurate modeling of special events using binary variables.
- o Accurate temperature modeling using heating and cooling degree functions.
- o Robust parameter estimation using weighted least squares linear regression techniques to minimize the effects of outliers and reduce the impacts of heteroskedasticity.
- o Effective use of "reverse errors-in-variables" techniques to make the load predictions less sensitive to large temperature forecast errors.
- o Distinction between time-independent daily peak load forecasts and the maximum of the hourly load forecasts to prevent peak forecasts from being negatively biased.

ACKNOWLEDGEMENTS

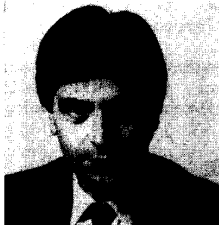
The authors acknowledge the many useful discussions they had with Mr. Carl Imparato, Mr. Andrew Bell, Professor Felix Wu and Mr. Robert Harshbarger.

REFERENCES

1. D.W. Bunn and E.D. Farmer, *Comparative Models for Electrical Load Forecasting*, John Wiley & Son, NY, 1985.
2. G. Gross, F.D. Galiana, "Short-Term Load Forecasting," *Proceedings of IEEE*, December 1987.
3. T.N. Goh, H.L. Long, and Y.O. Lee, "A New Approach to Statistical Forecasting of Daily Peak Power Demand," *Elec. Power Syst. Res.*, Vol. 10, No. 2, pp. 145-148, March 1986.
4. P.C. Gupta, "A Stochastic Approach to Peak Power Demand Forecasting in Electric Utility Systems," *IEEE Trans. Power App. Syst.*, Vol. PAS-90, March / April 1971.
5. R.P. Thompson, "Weather Sensitive Electric Demand and Energy Analysis on a Large Geographically Diverse Power System - Application to Short Term Hourly Electric Demand Forecasting," *IEEE Trans. Power App. Syst.*, Vol. PAS-95, No. 1, pp. 385-393, January/February 1976.
6. W.R. Christiaanse, "Short-Term Load Forecasting Using General Exponential Smoothing," *IEEE Trans. Power App. Syst.*, Vol. PAS-90, No. 2, pp. 900-911, March/April 1971.
7. W.D. Laing, "Time Series Methods for Predicting the CEGB Demand in Comparative Models for Electrical Load Forecasting," D.W. Bunn and E. D. Farmer, Eds. New York, NY: Wiley, 1985, pp. 69-85.
8. S. Vemuri, B. Hoveida, and S. Mohebbi, "Short-Term Load Forecasting Based on Weather Load Models," in *Proc. IFAC Conf. on Power Systems and Power Plant Control* (Beijing, People's Rep. China, August 1986), pp. 565-570.
9. S. Vemuri, W. L. Huang, and D.J. Nelson, "On-Line Algorithms for Forecasting Hourly Loads of an Electric Utility," *IEEE Trans. Power App. Syst.*, Vol. PAS-100, No. 8, pp. 3775-3784, August 1981.
10. L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA:MIT Press, 1983.
11. P.J. Huber, *Robust Statistics*, John Wiley & Sons, New York, 1981.

12. D.F. Andrews, et. al., Robust Estimates of Location: Survey and Advances. Princeton, New Jersey, Princeton University Press, 1972.
13. C.H. Reinsch, Smoothing by Spline Functions, Numerische Mathematik 10, 177-183.
14. F.D. Galiana, "Short Term Load Forecasting," in Proc. Engineering Foundation Conf. - Systems Engineering for Power: Status and Prospects (Henniker, NH, 1975), ERDA 76-66, Conf. 750867, pp. 17-22.
15. N.R. Draper and H. Smith, Applied Regression Analysis, Second Edition, John Wiley & Sons, New York, 1981.
16. G.E.P. Box and G.M. Jenkins, Time Series Analysis Forecasting and Control, Revised Edition, Holden-Day, Oakland, 1976.

BIOGRAPHIES



ALEX D. PAPALEXOPOULOS received the Electrical and Mechanical Engineering Diploma from the National Technical University of Athens, Greece in 1980, the M.S. degree in Electrical Engineering in 1982 and the Ph.D. degree in Electrical Engineering in 1985, from the Georgia Institute of Technology, Atlanta Georgia. Since October 1985, he has been a member of the Pacific Gas and

Electric Company's Systems Engineering Group, working on the development of advanced applications for PG&E's new Energy Management System including network equivalencing, optimal power flow, transmission constrained economic dispatch and system load forecasting. His primary research interests include applications of large-scale systems theory to the real-time control of power systems, dynamic simulation of power systems, and electromagnetic transient analysis. Dr. Papalexopoulos is member of Sigma Xi, IEEE and the Technical Chamber of Greece.



TIMOTHY C. HESTERBERG received his B.A. degree in Mathematics from St. Olaf College in 1980, his M.S. in Statistics from Stanford University in 1983, and his Ph.D. in Statistics from Stanford in 1988. He joined the Pacific Gas & Electric Company's Systems Engineering Group in 1985 as a statistical consultant. His primary

interests include bootstrap analysis and Monte Carlo simulation methodologies. Dr. Hesterberg is a member of the American Statistical Association and the American Mathematical Association.

Discussion

JAMES A. LARSON, Northern States Power Co., Minneapolis, MN.: I was impressed with the very good accuracy of their forecasting algorithm, particularly in a service area as large and as climatically diverse as the PG&E system.

We have used regression equations to predict future values from past historical information. We have not considered adding noise to our historic data. Therefore, I was surprised to see a suggestion in this paper to add noise terms to one's historic data set before doing the regression analysis. In particular, the authors added noise terms to the historic temperature values. Their reason for doing so is that temperature forecasts for tomorrow can be interpreted as the actual temperature plus a noise component. The noise component is the temperature forecast error. Therefore, in the regression parameter estimating process, the authors state that it is appropriate to add similar noise to the historic temperature data. The resulting regression equations will reduce the sensitivity of load predictions to temperature forecast errors.

It is the purpose of my discussion to challenge the above logic. In particular, I will demonstrate that adding noise to the historical temperature adds additional error to the regression equation, and results in poorer load forecasts.

I use a simple linear regression example to show the problem most clearly. Imagine that one has a historic data set of load vs. temperature, and wishes to fit it with a regression equation of the form $L(T) = aT + b$. T is the temperature in degrees Fahrenheit, and L is the load in MW. An example made-up historic data set, and the best fit linear regression equation to that data set, is shown in Fig. 1 below. This regression line is labelled "historic". The data set ranged from 70 to 110 degrees, but Fig. 1 shows only a zoomed in portion ranging from 82 to 94 degrees.

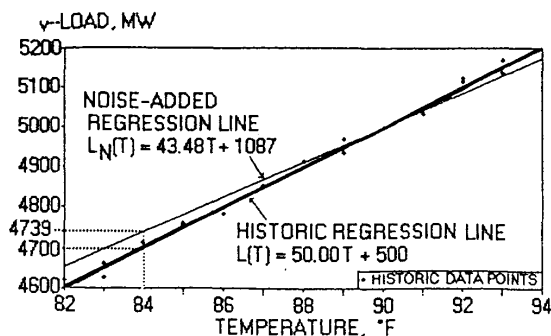


Fig. 1 System Load Vs. Temperature -- Example Historical Data, And Regression Equations With And Without Noise Added

Then, as recommended in the paper, I added and subtracted a noise component v from each temperature in the historic data base, and performed a linear regression fit to this new noise-added data. This regression line is labelled as "noise - added" in Fig. 1. I used $v = 5$ degrees.

If the temperature forecast for tomorrow is probabilistically distributed with a mean of T and a standard deviation (s.d.) of σ degrees, then, using the regression equation, the prediction for tomorrow's load would have a mean of $\bar{L} = a\bar{T} + b$, and a standard deviation of $\sigma_L = a\sigma_T$.

Assume the forecasted temperature for tomorrow is 84 degrees, with a s.d. of 0, 5, and 10 degrees. Then

by using the above equations with the historic regression equation, and the noise-added regression equation, I obtain the following predictions of tomorrow's load:

Case	Temperature, Degrees F.		MW Load Predicted By The Historic Regression Equation		MW Load Predicted By The Noise-Added Regression Equation	
	mean	s.d.	mean	s.d.	mean	s.d.
1	84	0	4700	0	4739	0
2	84	5	4700	250	4739	217
3	84	10	4700	500	4739	435

The historic data suggests that if tomorrow's temperature turns out to be 84 degrees, tomorrow's load will most probably be 4700 MW, not 4739 MW. For days when the temperature is forecasted to be 84 degrees, the long run average load is most likely to be 4700 MW. The historic regression equation gives a better mean value of predicted load (4700 MW) than does the noise-added regression equation (4739 MW). In this example, adding noise to the temperature data causes, in the long run, an average 39 MW error to be added to the predicted load on days when the temperature is forecasted to be 84 degrees. There is no statistical justification for watering down coefficients because of uncertainty.

Manuscript received June 1, 1989.

B. Buchenel, K. Imhof (Asea Brown Boveri, Network Control, Switzerland): We wish to congratulate the authors for their comprehensive paper about a regression based algorithm for short term load forecasting.

Quite often different models used for forecasting show a close relationship. As mentioned by the authors, ARIMA models can be transformed into state space models. There is also a relationship between regression based and ARIMA based models. As a consequence the choice of the used method becomes almost a matter of taste. More important is the question of how the algorithm is made. In the method proposed by the authors the tools provided by regression theory are skillfully applied. From the author's starting point which is often called old fashioned, good forecasts may be obtained, as shown.

The present model however can not be understood as a pure regression model. Among other terms the inclusion of autoregressive and moving average terms in the formulas (1), (7), (8) are characteristic of ARIMA models.

Parameter estimation: The very large number of parameters (more than 1000?) to be determined per time sequence require a complex program for parameter estimation. If the program has to be generalized to meet the special requirement of other utilities, the estimation will have to be done by an expert usually not available on site. We guess that reformulating some parts of the regression based model into an ARIMA based model could significantly reduce the number of parameters. The fact that many parameters have almost the same values, as mentioned by the authors, indicates that the model is far from being parsimonious.

Another objection for the regression method is that it requires pretty large sets of past data to be provided. What happens in case of changing cir-

cumstances, e.g. strong economic growth? The forecasting parameters in such a case are oriented too much to the past and do not take into account the future. Based on the fact that parameter estimation is only done on an annual basis, we guess that the model has only a limited adaptivity. In stable conditions this does not need to be a disadvantage. The formula (1) e.g. suggests as a trend over the year (or half year) a set of linear functions (parameters $a(d)$, $b1(d)$). How about the model accuracy after a year without reestimation? Does the error variance increase during the year?

Fig. 4 suggests that the regression based forecasting method has a better behaviour especially for larger lead times. How is the behaviour for lead times up to 1 week?

How are special events like TV broadcastings treated? The problem is that these are exceptional load behaviours. It has to be avoided that this untypical load behaviour is propagated to the next hours and days. How is the forecasting error determined if the forecast has been modified previously by the operator (input for (7) and (8))?

Manuscript received June 1, 1989.

Mo-yuen Chow (North Carolina State University): Congratulations for the fine work on the paper "A Regression-Based Approach to Short-Term System Load Forecasting" that combines two expertise fields—power engineering, which gives valuable insights of the practical problems, and statistics, which gives a theoretical base on the formulated problem. I am very impressed with the results obtained from the new SLF algorithm mentioned in the paper. Nevertheless, would you please further comment on the following questions:

1. Power system load is a slow time varying function and you have used data collected from the past six years to estimate the parameters of the model. The long historical data will very likely give biases on the parameters. How does it compare with the model by using only, say data from one or two years?
2. The model developed in the paper contains more than 70 terms and can be used for a whole year including different season and day type (weekday, and weekend). How does it compare with some simpler models containing less terms and which are easier to be handled for individual seasons and day types, say the whole year model contains four season models and two day types?

Manuscript received May 22, 1989.

A. D. Papalexopoulos and T. Hesterberg: We thank the discussers for their interest in our paper. Their insightful questions and thoughtful comments add perspective and complement the paper and raise several interesting points regarding SLF algorithms. We respond to each discussor separately:

Drs. Imhof and Buchenel:

- 1) The regression model proposed in the paper does have much in common with ARIMA models. ARIMA formulations generally would not reduce the number of parameters, except for holiday modeling. We chose a linear regression-based approach for practical considerations. Specifically, the availability of SAS regression procedures and data manipulation facilities made it easier to meet the diverse modeling requirements that deemed to be critical for accurate SLF forecasts. These requirements include: handling of holidays and other special events, handling of different behavior on different days of the week, treatment of heteroskedasticity and missing data, modeling temperature and seasonal changes in load shapes, minimizing the effects of outliers (robust parameter estimation), model building diagnostics, etc. No commercial program meets these requirements within a robust ARIMA framework, and incorporation of these features into a custom package would be impractical. On the other hand, statistical packages such as

SAS and S provide integrated data management, model building and numerical graphical model analysis capabilities. They meet the SLF requirements mentioned above in a regression framework, and programs for these packages can be easily modified by other electric power utilities.

- 2) The model is mathematically (though not notationally) parsimonious, since the linear constraints on the parameters reduce the number of degrees of freedom. For the six years of data and for the initial daily peak forecast model there are 74 degrees of freedom. After the definition of the "holiday effect" variable there are 36 degrees of freedom for the initial daily peak forecast model, of which 24 are specific to certain days of the week. There are only four degrees of freedom for the adjusted peak forecast model.
- 3) Load growth and seasonal variations in loads are captured by the inclusion of trend, daylight saving time and time-of-year terms in the regression model. In addition the exponential smooth of previous errors (as a term in the initial hourly and adjusted peak forecasts, equations (7) and (8) of the paper) corrects for forecasts which have been consistently too high or low over the recent past (roughly the last two months). This allows the model to adapt quickly to changes in the overall load levels caused by changing conditions, e.g. strong economic growth. In effect, the intercept coefficient changes quickly based on changing conditions. The other parameters are more stable (e.g. the coefficient for the previous day's load does not change, if all loads increase by the same factor) and are modified only at the annual parameter estimation. Parameters could be updated more often than annually and less than three years of data could also be used for parameter estimation (except that the six years of data should still be used for holiday parameter estimation since only a small amount of holiday data is available each year). The error variance does not seem to increase during the year, so more frequent parameter estimation should produce only a small improvement in accuracy.
- 4) Figure 4 of the paper illustrates error standard deviations for different times of the day for forecasts made at midnight, but does not indicate how error variances would increase for multiple day forecasts. Work is currently underway to develop a one-week model that will be used in conjunction with the Hydro-Thermal Coordination function of the PG&E's new EMS.
- 5) Special events can be handled by the operator in the on-line SLF function by adjusting the "actual" load which is used as input for later forecasts. In this way, usual load behavior is not propagated to later hours and days.

Mr. Larson

We thank Mr. Larson for his challenging observation. There are two models one might use to model temperature forecasts errors:

$$\text{Model 1: } F = T + \epsilon$$

$$\text{Model 2: } T = F + \epsilon$$

where F is the forecast temperature, T is the actual temperature and ϵ is the temperature forecast error which has zero mean and standard deviation σ . In the first model the error ϵ is uncorrelated with the actual temperature T and $\text{Var}(F) = \text{Var}(T) + \sigma^2$. In the second model the error ϵ is uncorrelated with the forecast F and $\text{Var}(T) = \text{Var}(F) + \sigma^2$. Note that in the second model the forecast is the expected value of the distribution of the temperature variable.

In general, the doubling idea can be used with either model (by utilizing a 2-point distribution) on approximate the conditional distribution of the temperature forecast given tomorrow's actual temperature and other independent variables. The doubling procedure outlined in the paper is based on the first model. For this model and using Mr. Larson's example, doubling reduces the error standard deviation from 250 MW to 217 MW. If the forecast is 84 degrees, the expected value of the distribution of the temperature is 84.78 degrees (assuming that F and T are jointly normal random variables), and 4739 MW is the best load forecast. If the second model is valid, the expected value of the distribution of the temperature is 84 degrees, and 4700 MW is the best load forecast.

A doubling technique for the second model is more complicated and can be described as follows: let F_T be the best temperature forecast based on other independent variables (today's temperature and time of the year but not current meteorological information) and replace the actual historical temperature T with a "forecast" temperature F given by:

$$F = F_T + a(T - F_T) \pm b$$

The parameters a and b are chosen so that the "forecasts" have the proper

covariance with T and F_T , i.e.,

$$a = \frac{\text{Var}(F - F_T)}{\text{Var}(T - F_T)}$$

$$b^2 = \frac{\text{Var}(F - F_T) \cdot \text{Var}(T - F)}{\text{Var}(T - F_T)}$$

If the load/temperature relationship is assumed to be linear the resulting regression model with doubling will be exactly the same as the one obtained using historical data without doubling. With a nonlinear relationship however, the resulting regression equation tends to counteract the bias due to nonlinearity. Approximately:

$$L_2(F) \cong E[L(T) | F]$$

where:

$L_2(X)$ is the load forecast given temperature X estimated with the doubling technique

$L(X)$ is the load forecast given temperature X estimated without the doubling technique and,

$E[Y | Z]$ is the conditional expected value of Y given Z .

A simpler way to estimate $E[L(T) | F]$ is to use the quantity:

$$\frac{L(F + \sigma) + L(F - \sigma)}{2}$$
 which does not require doubling in parameter estimation.

We are currently investigating doubling techniques for the second model.

Dr. Chow

- 1) As mentioned in the paper the decision regarding how much data to use to estimate parameters involves a tradeoff between bias and variance. In the three-step weighted least squares linear regression process, proposed with this model, three years of data were used. The six years of data were used only to provide the weights given by equation (3) of the paper, and to estimate holiday coefficients. These weights were used in the third step of the regression process to minimize the effects of outliers. Less than three years of data could be used, except for holiday parameter estimation.
- 2) Four day types are necessary, because the previous day's loads are used as independent variables in the regression formulation (e.g. the coefficient for the previous day loads should be smaller for Saturday than for Sunday load forecasts). Seasonal models would not be significantly simpler than the single combined model in which the difference between seasonal load patterns are represented using sinusoidal terms and a daylight saving time variable; for good performance Spring and Fall models would still need to model the effects of both heating loads due to cold fronts and cooling loads due to warm fronts.

Finally, the authors wish to thank the discussers again for their insightful questions and interesting comments.

Manuscript received August 7, 1989.

Discussion

T. J. Hammons (Glasgow University, Glasgow G12 8QQ, Scotland, UK): The authors have provided useful information on torsional interaction between non-identical turbine-generators in a two-machine configuration. They have based their analysis on a modal model which enables independent variation of mode damping, inertia and frequency. Independent variation of these parameters is difficult to achieve using conventional spring-mass simulations of turbine-generators. They have shown that the torsional characteristics of the non-identical generators with a coincident mode which they have investigated are very dependent, both in relative amplitude of generator responses as well as their phase relationship, on the degree of series capacitor compensation employed in the transmission system. They have found that a practically realistic difference in mode damping of 0.5 1/s can result in a 40% phase displacement and a 40% difference in amplitude between the two predominant torsional oscillators, and that differences in mode frequency can affect torsional interaction significantly. For example, a 0.1 Hz difference in frequency may result in a difference in phase displacement of up to 30° and a difference in amplitude of up to 55% in torsional oscillation of the two-machine system they have investigated.

This discussor notes that for all results the generators are unloaded, that automatic voltage regulator (avr) action has been ignored, that the turbine-generator-excitation system has been simulated using a six discrete spring-mass system, and that generator sub-transient phenomena (which can affect generator electrical damping for the lower torsional modes) has been neglected.

For realistic assessments of torsional interaction between non-identical turbine-generators the above phenomena should be simulated. This is particularly true if precise predictions are to be evaluated.

Investigations made by this discussor depict that in real machines theoretical values of damping of shaft torsional vibrations due to viscous damping varies with turbine load for the turbine, and with both frequency of modal vibration and generator load angle for the generator. Time constants for decay of predominant shaft torsional vibrations, or equivalent logarithmic decrements, vary significantly from vibration to vibration on account of mode shape and magnitude of the viscous damping which acts on each cell or section of the shaft.

Steam viscous damping, it will be recalled, is based on the constant power relationship of the characteristic which relates turbine speed with torque. This assumption leads to steam viscous damping which acts at the HP, IP and LP turbines which is proportional to the machine load.

Electrical viscous damping is dependent on frequency of modal vibration of the shaft and on effective load angle at which the generator operates. Effectiveness of each cell or section of the generator rotor in damping predominant shaft torsional vibrations can vary significantly from cell to cell or section to section depending on mode shape. This is particularly true of the higher frequency shaft torsional vibrations, whether subsynchronous or supersynchronous, where modal displacement can vary significantly along the active length of the generator rotor. Damping effects are accumulative should modal displacement change sign; excitation of torsional vibration effects cancel should modal displacement reverse along the generator rotor.

It should be noted that electrical damping at the higher torsional frequencies is effected by eddy currents which flow in the solid iron sections of the generator rotor and associated damper bars. Very little damping is effected by currents which flow in the generator field winding on account of the deep bar effect. Emfs induced in the generator field winding due to generator rotor torsional vibrations will cancel should modal displacement along the active length of generator rotor change its sign. The assumption that electrical damping is accumulative when modal displacement changes sign is valid for the higher modes of torsional vibration of the shaft.

Logarithmic decrements or equivalent damping expressed by time constants due to steam viscous damping and electrical phenomena may be calculated using steam and electrical viscous damping coefficients, the inertias, and the eigenvectors for continuum or reduced-mass models corresponding to each mode of torsional vibration of the shaft. They may also be calculated using classical techniques. Continuum and reduced shaft models in evaluating damping of predominant torsional vibrations in turbine-generator shafts are discussed in Reference [A].

At light loads, when the load angle to effective infinite busbar is small, electrical damping is effected by damper (which includes skin effect) currents which flow in the quadrature-axis dampers of the synchronous machine. Conversely, when the generator operates at a high load angle, currents which flow in the direct-axis damper circuits are effective for electrical damping of torsional vibrations in turbine-generator shafts.

Effect of generator load angle on electrical damping of predominant shaft torsional vibrations is discussed in Reference [B]. Effect damper modelling has on damping shaft torsional vibrations is discussed also in Reference [C].

Reduced shaft models deduced from continuum models should be tuned to the phenomenon that is being investigated if precise assessments are to be evaluated. Different shaft models may be necessary to evaluate different torsional phenomena such as subsynchronous resonance which may occur in an exciter shaft or shaft torsional phenomena following a severe supply system disturbance. If modal displacement varies significantly along the length of the generator rotor, two or more masses may be required to model the generator rotor. For similar reasons, multi-mass simulation of the exciter system may be necessary for faithful simulation of exciter dynamics in the generator/exciter shaft.

Avr action is often employed as a countermeasure for subsynchronous resonance. As avr action can affect torsional response, it should be simulated if precise assessments are to be made, even if the avr is not adapted specifically as a countermeasure for subsynchronous resonance.

Classical equivalent circuits with one damper on each axis have long been recognized as inadequate for investigating subsynchronous resonance and torsional interaction between turbine-generators phenomena. This is particularly true for the lower frequencies of torsional vibration of the shaft. Additional circuits are essential to match the operational inductances of the synchronous-generator rotor when shaft torsional phenomena is involved. Rotor-circuit fits to standstill frequency response data depicted by equivalent circuits given by Hammons and Canay (1985) in Reference [C] are adequate to match the operational inductances of generators in the analysis of torsional interaction between turbine-generators.

It should be noted that such equivalent circuits can be derived also from design data. Since effective transmission line resistance and inductance is frequency dependent, appropriate values corresponding to each torsional oscillation of interest should be used when real system analyses are made.

Could the authors comment on the importance of including the above phenomena in calculations if precise assessment of torsional interaction between real non-identical turbine-generators are to be assessed. Their views on the damping of resonances, the effect varying modal displacement with a change of sign along the generator rotor might have in real predictions, the use of different spring-mass models to study interactions at different modal frequencies where more than one modal vibration is significant, and use of the avr to control torsional interactions would be welcome. The authors' comments on the above observations would enhance practical applications of the studies they have reported to date.

References

- [A] T. J. Hammons, M. Istin, and A. Crocqueville. "Analysis of Continuum and Reduced Shaft Models in Evaluating Turbine-Generator Shaft Torsional Response following Severe Disturbances on the System Supply". *Electric Machines and Power Systems*, Vol. 13, (6), 1987, pp. 387-408.
- [B] T. J. Hammons. "Effect of Fault Clearing and Damper Modelling on Excitation and Decay of Vibrations in Generator Shafts following Severe Disturbances on the System Supply". *IEEE Transactions on Energy Conversion*, Vol. EC-2, (2), 1987, pp. 308-320.
- [C] T. J. Hammons and I. M. Canay. "Effect of Damper Modelling and the Fault Clearing Process on Response Torque and Stressing of Turbine-Generator Shafts". *IEEE Transactions on Energy Conversion*, Vol. EC-1, (1), 1986, pp. 113-121.

Manuscript received November 8, 1989.

G D Jennings and R G Harley : The authors would like to thank Dr Hammons for his interest shown in the paper and for his comments.

Dr Hammons has commented on the importance of including AVR action, sub-subtransient effects and frequency dependency of transmission system inductance and resistance when making precise assessments of torsional interaction in a real system. The above factors influence the electrical damping of the two generators and it has been shown that differences in electrical damping between two generators will result in phase shifts between their oscillations (Ref [5] in paper). Fig 1 below shows the amplitude ratio and phase deviation calculated for the two generators in the paper with equal mechanical parameters (including mechanical damping) but at different

power levels. These results thus show the effects on torsional interaction of electrical damping differences caused only by a difference in load angle. The effects of electrical damping are clearly important in determining torsional interaction characteristics.

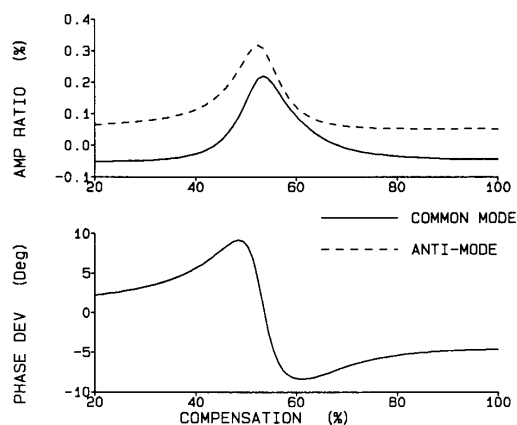


Fig 1 Amplitude ratio and phase deviation for $P_{g1}=0.05$ p.u. and $P_{g2}=0.9$ p.u.; for both generators $\sigma_3=0.0$ s⁻¹.

In this paper we were mainly concerned with the effects of *mechanical* parameter differences on the torsional interaction of the two machines and therefore purposefully chose the generators to be electrically identical and at the same operating point for all results. If precise calculations are wanted in a real study then any factors influencing the torsional interaction characteristics should be included and accurately modelled.

However it should be noted that the inclusion of AVR action and sub-transient effects would not in themselves cause an amplitude difference or phase deviation between generator oscillations unless it could be identified that the AVR action on each generator was different or that the sub-transient characteristics of each generator were different and these differences were included in the model. Frequency dependence of transmission line resistance and inductance cannot cause amplitude and phase variations between generator oscillations. All three effects will influence the amount of electrical damping present in the system and this would affect the results for mechanical damping differences which depend on the total system damping. The less the system damping, the greater the predicted phase and

amplitude differences will be for a given difference in mechanical damping.

On the damping of resonances, the effect on this of differences between generator parameters is not an easy question that can be nicely summarized in a few statements. Initial studies we have done show that a small variation in mechanical parameters tends to cause the torsional modes to be destabilized to a smaller extent (less damping needed to be added to stabilize system) as compared to identical parameters, but may cause instability to be in effect over a wider range of compensation level. This is however under further study.

Concerning the effect of a varying modal displacement along the generator rotor, we agree with Dr Hammons' comments on how this influences electrical damping. In our experience though we have not encountered modes in the subsynchronous range that have a large change in modal displacement along the length of the rotor and certainly none that have a change in the sign of the modal displacement. If however, for a particular mode the modal displacement does change sign along the rotor, then it will be necessary to model this if the resulting change to the electrical damping is significant with respect to the system damping.

With regard to the use of different spring-mass models, the choice of model is certainly determined by the torsional phenomenon being investigated. This paper has considered a model that includes only one torsional mode; additional torsional modes could be added if necessary. The effect of using reduced order models and neglecting significant modes on SSR calculations is discussed in Ref [a].

The AVR affects the electrical damping of torsional modes so it can in theory be used to control torsional instabilities. The effectiveness of this depends very much on the nature of the exciter, particularly time constants and voltage ceiling limits. Our experience with a slow acting rotating diode exciter has shown this form of excitation system to be ineffectual in damping torsional oscillations, particularly for large disturbances (Ref [b]). The problem of damping torsional oscillations in neighboring non-identical units is complicated by the resulting amplitude differences and more importantly phase shifts between generator oscillations and is a subject that we feel needs further investigation.

- [a] G D Jennings, R G Harley and D C Levy, "Sensitivity of subsynchronous resonance predictions to turbo-generator modal parameter values and to omitting certain active subsynchronous modes", IEEE Transactions on Energy Conversion, Vol EC-2, No 3, Sept 1987, pp 470-479.
- [b] D J N Limebeer, R G Harley and M A Lahoud, "The suppression of subsynchronous resonance with the aid of an auxiliary excitation control signal", Transactions of the South African IEE, Vol 74, Part 8, Aug 1983, pp 198-209.

Manuscript received November 8, 1989.

Discussion

A. Monticelli (UNICAMP, Campinas, Brazil) and L. Radu (Consolidated Edison, New York, NY): This well written paper, as well as the related paper [A], have ingeniously combined some previously known techniques (matrix blocking, Hachtel's tableau method) in order to overcome numerical problems in state estimation.

We note that, in fact only the injection measurements should be handled in Hachtel's way. The flow measurements (as well as the voltage measurements) could be directly included in the authors reduced matrix by equivalent algebraic expressions (corresponding to normal equations) with no need for a partial factorization. Also, we note that when a bus is observable only from a neighboring injection (zero or measurement), a situation quite common in practical systems, it is not possible to pair the injection with a local state variable. Moreover, for some low redundant measurement structures the ordering of equations has to be constrained (some injections should be eliminated first as suggested in [A]). This being the case, we would like to make the following general comment:

Apparently there are two different approaches to the factorization of indefinite matrices such as the ones that are considered in the paper. In the first approach factorization is performed taking into account both matrix sparsity and numerical conditioning; as examples one can mention the delayed pivoting scheme and the mixed 1×1 , 2×2 pivoting used in the Harwell routines. The other approach is based on some kind of prearrangement/preordering that usually allows a positive definite like factorization: the blocking scheme of the paper is a representative example of this approach. As a matter of fact, in general this has been a popular choice since the inception of computer methods for power system analysis, not only in state estimation but also in related problems such as power flow and optimal power flow calculations. A typical example is the use of topological observability analysis in connection with normal equations

method for state estimation: the idea is to determine in advance parts of the network whose state can be estimated. Under ideal conditions an observable network leads to a positive definite matrix which can be factorized without problems. Similar approaches are used regarding the solvability of power flow and optimal power flow problems where a previous definition of bus types and blocking arrangements usually lead to nonsingular (or even positive definite) problem matrices. Usually, but not always. As we know, numerical problems happen. The bottom line is that though qualitative methods may help, one may still need some numerical checking to make sure the problem we are trying to solve is really solvable with the algorithms, parameters and computers we have at our disposal at a given point in time. That is to say, problem solvability is not always guaranteed by structural analysis such as the ones implicit in topological observability and blocking schemes.

This being the case, we suggest that some flexibility would be gained by using an on the fly ordering-factorization with a numerical observability algorithm invoked, when numerical problems are encountered during the factorization process, in order to adaptively change the scope of the internal state estimation.

For a discussion on a closely related problem, see [B] which deals with numerical problems associated to the Newton OPF approach.

References

- [A] R. Nucera and M. Gilles, "A Blocked Sparse Matrix Formulation for the Solution of Equality-Constrained State Estimation", Paper 90WM 234-5 PWRS, IEEE/PES 1990 Winter Meeting, Atlanta, GA.
- [B] A. Monticelli and W-H. E. Liu, "Adaptive Movement Penalty Method for the Newton Optimal Power Flow", Paper 90 WM 252-9 PWRS, IEEE/PES 1990 Winter Meeting, Atlanta, GA.

Correction to "State Estimation Using Augmented Blocked Matrices"¹

In the above paper, 1, the following discussion should have appeared.

¹F. Alvarado and W. Tinney, *IEEE Trans. Power Systems*, paper 90 WM 241-0, vol. 5, pp. 911-917, August 1990.