IEEE Transactions on Power Apparatus and Systems, vol. PAS-94, no. 2, March/April 1975

STOCHASTIC LOAD FLOWS

J. F. Dopazo O. A. Klitin A. M. Sasson

American Electric Power Service Corp.

N. Y., N. Y.

Abstract:

The load flow study has been at the center of studies made for designing and operating power systems for many years. It is well known that forecasted data used in load flow studies contain errors that affect the solution, as can be evidenced by running many cases perturbing the input data.

This paper presents a method for calculating the effect of the propagation of data inaccuracies through the load flow calculations, thus obtaining a range of values for each output quantity that, to a high degree of probability, encloses the operating conditions of the system.

The method is efficient and can be added to any existing load flow program. Results of cases run on the AEP system are included.

Introduction

The load flow study has been at the center of analysis made to design power systems and delineate operating practices. From the studies performed in network analyzer days to the sophisticated digital computer programs of today, conceptually speaking, not much alteration has taken place to the definition of the load flow problem. The fast methods available now, with their efficient utilization of computer core and their capability of access to large data files, have made possible for the engineer to run a large number of cases, starting from some base system condition, to study his design or his proposed plan of operation under many alternatives. Important decisions evolve from these studies.

For each load flow case that is solved, it is necessary for the engineer to provide the required data that define the conditions for the case. This data usually include real and reactive loadings at so-called load busses and real power generation and voltage magnitude at generator busses, as well as the electrical characteristics of the system model.

The formulation of the load flow problem assumes that the data provided is absolutely precise and provides results totally compatible with the given data apart from round-off errors. However, in practice, it can be readily appreciated that load flow data can only be known within some finite precision, this being more the case as the study represents conditions that are more distant into the future. As a normal screening process, the engineer looks at the range of possible values for a particular piece of data and selects an average value as the number to be used in the load flow study.

Given the importance of the decisions that evolve from load flow studies, it appears that it is important to know the possible ranges of result quantities corresponding to the known range of data quantities. In other words, it is of interest to determine the effect on load flow results of our ignorance of input data values.

This paper will address itself to the problem of processing the expected errors in the data of the load flow problem, assuming that the electrical characteristics of the model are known sufficiently well that the effect of their inaccuracies is negligible.

Essentially, the method converts the load flow problem formulation from a deterministic one to a stochastic one. The results obtained can be considered to define the range of variation of result quantities, thus determining the worst conditions for each. Similar results to those obtained with the stochastic load flow can be arrived at by repeating a large number of conventional deterministic load flow cases in which for each case the data is perturbed, such that the various cases represent possible sets of data within the precision that the data is known. Once all these studies have been carried out, the range of values for a specified result quantity can be identified. What the stochastic load flow does is to obtain these ranges in one direct calculation. An important practical characteristic of the method is that it can be added to any existing load flow program since the process requires first the solution of a conventional load flow, followed then by a series of steps that determines the propagation of the errors of input data. The method is not only applicable to the load flow problem but to any problem where the model is a system of a sufficient number of equations.

MATHEMATICAL BACKGROUND

The method is based on the principles of statistical least squares estimation for linear systems. Some of the initial ideas of the stochastic load flow came about from the experience obtained with simulation of and testing on the AEP monitoring project [1-6]. The derivation of some of the equations to be presented can be found in these references or in advanced statistical texts.

Given a non-linear set of equations,

$$y' = f(x') + \epsilon \tag{1}$$

a Taylor's series expansion can be used to linearize (1),

$$y = Jx + \epsilon \tag{2}$$

where

$y = y' - f(x)|^{\circ}$
$x = x' - x'_{0}$
J = Jacobian of $f(x')$
ϵ = vector of error random variables

Equation (1) is interpreted in the following way.

There are data quantities y' which represent the average value of the range of possible values the piece of data may have according to some statistical distribution defined from our physical knowledge of the problem and the methods that were used to forecast the data. There are problem variables x' from which all quantities can be computed. These variables have a true, although unknown, value which represent conditions as they will exist at some future date. The vector ϵ represents the error between y' and $f(x')$. As x' will not be known until the future conditions are encountered, ϵ can

Paper T 74 308-3, recommended and approved by the IEEE Power System Engineering Committee of the IEEE Power Engineering Society for presentation at the IEEE PES Summer Meeting & Energy Resources Conf., Anaheim, Cal., July 14-19, 1974. Manuscript submitted August 29, 1973; made available for printing April 3, 1974.

only be described statistically as a random variable that has some mean and variance that represent our expectations of the way ϵ could vary. The meanings of y and x are similar to those of y' and x'. The following will assume the linearization presented in Equation 2.

The statistics of the vector ϵ can thus be specified as

$$E(\epsilon) = 0 \qquad E(\epsilon\epsilon^t) = V \tag{3}$$

where V is a diagonal matrix of data variances and E stands for expected value.

The least squares approach to solving the sort of equations given in Equation (2) correspond to the minimization of the function

$$F(x) = (y-Jx)^t V^{-1} (y-Jx) \tag{4}$$

and the value of x that minimizes $F(x)$ is called $\hat{x}$ and is given by [6]

$$\hat{x} = (J^tV^{-1}J)^{-1} J^tV^{-1} y \tag{5}$$

which reduces to

$$\hat{x} = J^{-1} y \tag{6}$$

for the case that J is a square matrix, implying the same number of equations as unknowns. $F(\hat{x})$ is equal to zero for this case.

As the original Equation (1) was nonlinear, $\hat{x}$ should be interpreted as the converged value after some iterations in which J is recomputed in each iteration.

Corresponding to $\hat{x}$ there is a $\hat{y}$ computed directly from Equation (2),

$$\hat{y} = J\hat{x} \tag{7}$$

and corresponding to the true, but unknown value of x, called x_t there is a true, and also unknown, value of y, called y_t,

$$y_t = J x_t \tag{8}$$

Statistical Properties: Expected Values and Variances

The quantities $\hat{x}$, x_t, y, $\hat{y}$ and y_t have known statistical properties that are of interest,

$$E(\hat{x}) = x_t \tag{9}$$

$$E\{(\hat{x}-x_t)(\hat{x}-x_t)^t\} = (J^tV^{-1}J)^{-1} \tag{10}$$

$$E(\hat{y}) = y_t \tag{11}$$

$$E(\hat{y}-y) = 0 \tag{12}$$

$$E\{(\hat{y}-y)(\hat{y}-y)^t\} = V-J(J^tV^{-1}J)^{-1}J^t \tag{13}$$

which reduces to the following if J is a square matrix,

$$E\{(\hat{y}-y)(\hat{y}-y)^t\} = 0 \tag{14}$$

Equations (9) and (11) imply that if ϵ has the statistical properties given in Equation (3), then, if $\hat{x}$ and $\hat{y}$ had been determined many times from different values of y, the average of $\hat{x}$ and $\hat{y}$ tends to the true values of x and y. In statistical terms this is called an unbiased process.

Equation (10) presents the variance of $\hat{x}$ which represents, the deviation of $\hat{x}$ from x_t. Equation (13) presents the variance of $(\hat{y}-y)$ and Equation (14) says that this variance is zero if there are as many equations as unknows. This simply means that $\hat{y}$ is equal to y, or that the actual solution of Equation (2) has been obtained, a fact also reflected in that $F(\hat{x}) = 0$.

Another statistical property is that

$$E\{(\hat{y}-y_t)(\hat{y}-y_t)^t\} = J(J^tV^{-1}J)^{-1}J^t \tag{15}$$

which reduces to

$$E\{(\hat{y}-y_t)(\hat{y}-y_t)^t\} = V \tag{16}$$

if J is a square matrix. Equation (16) implies that it is not possible to obtain a y closer to y_t than the actual original data y.

Assuming that there are other quantities z' related to x' as follows,

$$z' = g(x') \tag{17}$$

a Taylor's series expansion of $g(x')$ linearizes (17) into

$$z = Kx \tag{18}$$

from which $\hat{z}$ can be obtained from $\hat{x}$,

$$\hat{z} = K\hat{x} \tag{19}$$

and z_t from x_t

$$z_t = K x_t \tag{20}$$

The statistical properties of z can be summarized as follows:

$$E(\hat{z}) = z_t \tag{21}$$

$$E\{(\hat{z}-z_t)(\hat{z}-z_t)^t\} = K(J^tV^{-1}J)^{-1}K^t \tag{22}$$

Equation (22) presents the variance of $\hat{z}$.

Statistical Properties: Confidence Limits

It is now necessary to mention the probability distribution characteristics of the various quantities. The input quantity y can have any probability distribution function, which is known from the knowledge of the physical problem at hand. Its statistical properties given in Equation (3) can then be computed. The output quantities x or z are linear combinations of y and by the Central Limit Theorem can be taken as normally distributed random variables. It can then be said that $\hat{x}$ and $\hat{z}$ are $N(x_t, \sigma_x^2)$ and $N(z_t, \sigma_z^2)$ respectively.

A normally distributed random variable can always be transformed into a unit normal, thus

$$\frac{\hat{x} - x_t}{\sigma_x} = N(0, 1) \tag{23}$$

$$\frac{\hat{z} - z_t}{\sigma_z} = N(0, 1) \tag{24}$$

where

$$\sigma_x^2 = \mathrm{diag}\ \{(J^tV^{-1}J)^{-1}\} \tag{25}$$

$$\sigma_z^2 = \mathrm{diag}\ \{K(J^tV^{-1}J)^{-1}K^t\} \tag{26}$$

Although x_t and z_t are not known, a statement can be made as to a range which encloses them with some probability of being correct. Within a range of ± 3 times the unit standard deviation of the random variables defined by (23) and (24), it can be said that there is approximately a 99% probability of enclosing the true values.

Thus, it can be stated that

$$x_t = \hat{x} \pm 3\,\sigma_x \qquad (27)$$

$$z_t = \hat{z} \pm 3\,\sigma_z \qquad (28)$$

with a 99% probability of being correct.

THE LOAD FLOW PROBLEM

The previous section presented the general theory required for the stochastic load flow. This section will relate that theory to the load flow problem. The various quantities referred to in the previous section will now be defined in terms of the variables familiar to the load flow.

- y refers to load flow data, that is, load bus P and Q and generator bus P and E.
- x refers to bus state variables E and δ
- V refers to the variances of the errors assumed on the y input data.
- J refers to the Jacobian of the load flow equation, that is the partials of P, Q and E input quantities y with respect to E and δ load flow variables x.
- z refers to output quantities of the load flow computed from x, such as, line flow P and Q and generator bus Q.
- K refers to the Jacobian of output quantities z with respect to load flow variables x.

Having related the variables used in the previous section to those of the load flow problem, they shall be used interchangeably in what follows. The various steps of the process will now be presented.

Step 1 - A Load Flow

The first step is to compute $\hat{x}$ from y as shown in Equation (6) with an iterative process to take into account the non-linear nature of the load flow equations, Equation (1). Equation (6) shows that for the case of equal number of equations as unknowns, the statistical properties of the errors of y drop out of the equations. This converts Equation (6) to imply the solution of a load flow problem with average numerical values assigned to the load flow data quantities y. Any existing load flow program can be used in this first step, thus making this step equivalent to the solution of the problem as is done in practice, with average values for the forecasted input data.

Step 2 - Computation of Covariance Matrix of $\hat{x}$

The covariance matrix of $\hat{x}$ was presented in Equation (10) as

$$\mathrm{Cov}(\hat{x}) = (J^tV^{-1}J)^{-1} \qquad (10)$$

To calculate the $\mathrm{Cov}(\hat{x})$, the Jacobian of the load flow equation must be first formed and evaluated at the solution point of the load flow of Step 1. Sparsity techniques should be used as both J and $J^tV^{-1}J$ are sparse matrices. An ordering scheme should be used to order $J^tV^{-1}J$. In the same way that the structure of the admittance matrix Y can be used to order J in a Newton's load flow, the structure of Y^tY can be used to order $J^tV^{-1}J$. Triangular factorization techniques can then be used to triangularize this matrix and repeated back substitutions will yield the elements of the covariance matrix. It is only necessary to save those terms that structurally are in the same positions as the original $J^tV^{-1}J$ and then only a symmetrical part. This is an important characteristic since, in general, the Cov (x) is a full matrix.

Step 3 - Computation of Covariance Matrix of $\hat{z}$

The covariance matrix of $\hat{z}$ was given in Equation (22) as

$$\mathrm{Cov}(\hat{z}) = K(J^tV^{-1}J)^{-1}K^t \qquad (22)$$

To calculate the $\mathrm{Cov}(\hat{z})$ the Jacobian of output quantities z must be calculated. The structure of this matrix is also dependent on that of the admittance matrix and is thus related to that of J. Matrix K is thus a sparse matrix and should be calculated as such. Because of its structural relationship to J, in the triple product indicated in Equation (22) only $J^tV^{-1}J$ structured elements of $(J^tV^{-1}J)^{-1}$ are needed. This is the reason for saving only those elements in Step 2. In the case of Equation (22), only diagonal elements need be computed and saved.

Step 4 - Computation of Confidence Limits

Confidence Limits for $\hat{x}$ and $\hat{z}$ were given in Equations (27) and (28) as

$$x_t = \hat{x} \pm 3\,\sigma_x \qquad (27)$$

$$z_t = \hat{z} \pm 3\,\sigma_z \qquad (28)$$

The σ_x^2 vector are the diagonals of the $\mathrm{Cov}(\hat{x})$ while the σ_z^2 vector are the diagonals of the $\mathrm{Cov}(\hat{z})$. Those elements of these two matrices have been already calculated and saved in Steps 2 and 3.

AREA CONSTRAINTS

The stochastic load flow, as presented in the previous section, amounts to summarizing the results of many load flows, as far as the quantities obtained in the confidence limits are concerned. These load flows could have been run perturbing all input quantities according to the statistics of the uncertainty of these values as derived from the knowledge of the physical problem. The results would reflect cases in which total load was increased with maybe generation decreased, forcing the slack generator to have wide limits. This situation might be improved if there is additional information available. In addition to the forecast made on individual loads and generators, it is often the case that total load and generation of given areas of the system are known to a greater precision. If conventional load flows were run perturbing the data, area load and generation constraints would be used to restrict the data.

Area load and generation constraints can be included in the stochastic load flow. An area constraint acts as an additional equation in Equations (1) and (2). The additional equation is obtained by adding some of the other equations already included in (1) and (2).

For example, if the loads of nodes 3, 5 and 7 are to be constrained to a finer degree, new equations are formed,

$$P_{n+1} = P_3 + P_5 + P_7 + \epsilon_{p,n+1} \qquad (29)$$

$$Q_{n+1} = Q_3 + Q_5 + Q_7 + \epsilon_{q,n+1}$$

This new equation appears as the equation of a new imaginary node, connected to 3, 5 and 7 and to all nodes to which these are connected to. However, the new node does not contribute any new variable x to the problem, thus making J a non-square matrix. It is important to realize that even if J is non-square in Steps 2, 3 and 4, this does not alter the fact that a conventional load flow program can be used in Step 1. This comes from the condition that the solution of the original n equations in (2) automatically solve the additional equations introduced by the area constraints.

The inclusion of area constraints as presented in Eq. (29) considers the input covariance matrix to be diagonal making the additional equations independent from the original set. Considering an area of k nodes, Eqs. (1) become

$$y_i = f_i(x) + \epsilon_i,\ i = 1, \dots k \tag{30}$$

which Eq. (29) has the form

$$y'_{n+1} = f_{n+1}(x) + \epsilon_{n+1} \tag{31}$$

The relation between Eqs. (30) and (31) is the following,

$$f_{n+1}(x) = \sum_i^k f_i(x) = \sum_i^k (y'_i - \epsilon_i) \tag{32}$$

Substituting in Eq. (31),

$$y'_{n+1} = \sum_i^k y'_i - \sum_i^k \epsilon_i + \epsilon_{n+1} \tag{33}$$

$$y'_{n+1} = \sum_i^k y'_i + \epsilon'_{n+1} \tag{34}$$

It is of interest to compare Eqs. (31) and (34). The area constraint formulation considers

$$E(\epsilon_{n+1}) = 0 \tag{35}$$

$$E(\epsilon_{n+1}\,\epsilon_i) = 0,\ i = 1, \dots, n \tag{36}$$

$$E(\epsilon_{n+1}\,\epsilon_{n+1}) = V_{n+1} \tag{37}$$

However, had y'_{n+1} been defined as in Eq. (34), the statistics of ϵ'_{n+1} would have been the following, assuming independence,

$$E(\epsilon'_{n+1}\,\epsilon'_{n+1}) = \sum_i^k V_i + V_{n+1} \tag{38}$$

Eqs. (37) and (38) show that there is a distinct difference between modelling the area constraint as a function of a stationary true value x' and as the sum of random variables.

CONSTRAINT ON TOTAL SYSTEM LOAD

The previous sections have shown that the stochastic load flow represents the combined results of many deterministic load flows. Two types of load flow runs have been modelled:

1. A set of load flows in which each piece of data is randomly varied. It was discussed that this should produce large confidence limits for the slack generation and for lines near that bus due to some cases representing changes in total load and in generator outputs.

2. A set of load flows in which each piece of data is randomly varied but construing the variations such that the load and generation of areas within the system are allowed to vary by small amounts.

The area constraints would substantially reduce slack bus confidence limits but has two basic limitations:

1. Area constraints deteriorate the sparsity of the state inverse covariance matrix as their effect is similar to the introduction of a node connected to all other nodes in the area. In the limit, considering the entire system as one single area, this matrix becomes full.

2. Confidence limit results would, in general, be somewhat sensitive to slack bus location.

The above limitations can be easily overcome by modelling the load flow slack bus as any other generator bus, with P and E equations but allowing only E to be a state variable, keeping δ fixed. The forecasted error in total system load plus losses can be distributed among generator P equations, thus making all generators compensate for forecasting errors. It is not necessary to divide the total expected error evenly among generators. No area constraints are used.

The changing of position for the angle reference bus has no effect on confidence limits other than those of voltage angles themselves. As angles are really angle differences between busses and the reference bus, these values and their confidence limits are dependent on the reference bus location. However, the confidence limits of angular differences between any two busses are independent of reference bus location.

As in the case of area constraints, the reference bus equation is considered to be an independent equation. The modelling of the reference bus equation is made as a function of the system state variables and not by adding all other inputs. This model also reflects the nature of the distribution of uncertainties throughout the system to be one of non-simultaneity of occurrence of individual uncertainties and not one of correlation.

Of negligible effect is the condition of having one extra equation, producing an overdetermined set of input equations and a filtering effect.

SOLUTION OF PRACTICAL CASES

Several test cases are presented and summarized in Table 1. The AEP system data, as described in Reference 7, was used in all the cases. Reference 7 presented results of 50 runs of a conventional load flow in which the load flow data was randomly perturbed assuming it behaved in a rectangular distributed manner within some error bounds that, in general, were in the order of 10 to 20 percent of the mean data. That same reference presented the maximum error of each result quantity. The maximum error of one quantity does not occur in the same load flow run as the maximum error of another quantity, so in that sense they are not simultaneous. However, as it is not known what are the true inputs in a given run, any of the maximums could occur. Table I summarizes the average values (results of the base case load flow or Step 1 of the stochastic load flow) together with the maximum errors as determined from Reference 7.

The variances of the input data reflect the rectangular distributed error assumed in the relatively few cases of Reference 7 by computing V from,

$$V = (\text{Reference 7 error bound}/2.5)^2$$

The output quantities can be considered to be normally distributed from the Central Limit Theorem. Thus, Reference 7 results are considered to be approximately three sigma values and are compared in Table I with the corresponding results of the stochastic load flow. It is to be considered that while the 50 cases run in Reference 7 are statistically significant they are less so than the results of say 100 or 500 cases. In essence, the stochastic load flow summarizes the results of a greater number of cases than the 50 of Reference 7.

Table I presents results of some typical output quantities. In particular, it is of interest to look at the flows in lines near the slack bus, bus number 1. These are presented in the first 8 rows in Table I. The next two groups of 3 rows each refer to lines in two different sections of the system, both removed from the slack bus. The following group of 3 rows refer to generation results at the slack bus and two other generators. Since the real powers of the last two generators are data quantities they are not included in Table I. The last group of 3 rows presents the voltage magnitude and angle at three different busses, bus 21 being a generator bus.

The six cases presented in Table I are now discussed:

Case 1

Case 1 is referred to as the Fixed Voltages case in Table I. In this case generator bus voltage magnitudes were considered to be fixed, all other quantities having the bounds given in Reference 7. A comparison with the results of Reference 7 shows that the real line flow Confidence Limits are of the same order of magnitude, actually slightly larger. Reactive flow for lines near the slack bus are smaller in Case 1, the same being true for the results of reactive generation.

Case 2

Case 2 is referred to as the Variable Voltages case in Table I. For this case generator voltage magnitudes were considered both as input quantities and as variables of the stochastic load flow. A comparison with Case 1 shows that results are very similar except for reactive generation and reactive flow of lines near the slack bus. An error bound of 1.2 percent was assumed at all generator busses, this being the only error bound data not being identical to that of Reference 7, although it is of the same order of magnitude.

Cases 1 and 2 appear to draw the following two conclusions:

1. The linearization implied in going from Equation (1) to Equation (2) is valid for the level of error bounds considered.

2. The decoupling of P-δ and Q-E quantities in the load flow is visible in the comparison of the two cases. The addition of voltage magnitude variations had a considerable effect on reactive flow and generation, and a negligible one on real power flow and angles.

Case 3

Case 3 includes additional constraints, in both load and generation areas, such that the total load or generation of each area has an error bound of 5% of the average total load or generation for the area. All other input quantities have individual bounds as in Case 2.

The results of Case 3 show a general attenuating effect on the Confidence Limits of all quantities, especially those near the slack bus. This is as expected. In reference 7, and in the equivalent results of Cases 1 and 2, the data used for the load flows implied situations where the total load was increased or decreased from the average value, the slack bus having to adjust the differences. As a

TABLE 1 - STOCHASTIC LOAD FLOW TEST CASES

			CONFIDENCE LIMITS						
FLOW BETWEEN BUS TO BUS		AVERAGE VALUES	REFERENCE 7	CASE 1 FIXED VOLTAGES	CASE 2 VARIABLE VOLTAGES	CASE 3 5% AREA CONSTRAINTS(1)	CASE 4 2% AREA CONSTRAINTS	CASE 5 5% AREA CONSTRAINTS(2)	CASE 6 3% TOTAL LOAD CONSTRAINT
1	38	-118-J 68	411+J245	505+J 5	505+J178	180+J148	93+J148	214+J148	54+J206
1	2	515+J 51	78+J 91	130+J 19	130+J 76	62+J 63	51+J 63	49+J 62	55+J 88
1	3	237-J 0	162+J 63	191+J 4	191+J 48	72+J 40	43+J 40	100+J 40	23+J 56
38	39	680-J695	163+J 51	218+J 48	218+J 56	97+J 32	76+J 29	178+J 46	76+J 30
38	40	1-J380	272+J174	321+J 14	321+J166	131+J139	99+J138	222+J139	76+J138
40	43	733-J130	208+J163	250+J102	249+J201	109+J148	94+J147	198+J147	68+J167
3	7	-101+J 39	175+J 99	205+J 25	206+J102	94+J 83	79+J 83	186+J 85	64+J 83
7	12	-158+J 8	152+J163	219+J 51	219+J138	110+J113	98+J113	225+J112	77+J114
16	20	-96-J 63	81+J 35	76+J 50	76+J 51	40+J 29	34+J 28	81+J 32	53+J 41
16	15	-102-J150	50+J 31	55+J 44	56+J 44	40+J 36	39+J 36	61+J 45	40+J 36
15	9	-184+J 51	39+J 33	53+J 42	53+J 43	34+J 32	32+J 32	59+J 28	36+J 36
33	32	-260-J 34	42+J 23	62+J 24	62+J 28	50+J 21	50+J 21	87+J 18	17+J 22
32	37	-222-J 15	47+J 19	60+J 13	60+J 19	48+J 14	47+J 14	86+J 14	9+J 15
37	36	229-J 28	53+J 28	61+J 22	61+J 26	49+J 19	49+J 19	89+J 19	9+J 20
GENERATION AT BUS									
1		634-J 17	609+J334	773+J 21	773+J200	264+J166	117+J166	290+J166	15+J295
21		-35+J 59	J377	J117	J347	J272	J272	J272	J289
40		800-J200	J540	J105	J506	J413	J413	J413	J420
VOLTAGE AT BUS									
10		96∠-21.8	2.42∠4.24	3.60∠5.34	3.80∠5.35	1.70∠2.17	1.60∠1.36	1.80∠2.99	0.70∠1.61
21		1.02∠-8.4	1.21∠4.84	∠6.37	1.40∠6.37	1.20∠2.41	1.20∠1.55	1.20∠3.49	1.20∠0.87
39		98∠-12.6	0.85∠4.91	0.30∠6.36	.90∠6.36	0.70∠2.48	0.70∠1.63	0.70∠3.84	0.70∠1.61

(1) Individual data errors are as in Reference 7.

(2) Individual data errors are 40% of Average Values.

consequence, not only the slack generator but also the lines near the slack bus were found to have large Confidence Limits. By constraining the load and generation of areas of the system to smaller variations than the individual loads and generations, the total load variation implied is reduced, producing the attenuating effect.

Case 4

Case 4 is similar to Case 3, the area constraints reduced from 5% to 2%. Some further attenuation occurs, especially to the slack bus generation. However, most output quantities did not exhibit significant Confidence Limit variations, indicating that they were more due to the individual data error bounds than to the area error effect.

Case 5

This case presents the effect of increasing the error bounds of the individual data quantities beyond those of Reference 7. A normally distributed error bound of 40% of the average data value was assumed for each quantity, constraining at the same time the areas to within 5% of total average values.

A comparison of Cases 3 and 5, both with 5% area constraints, shows that while the Confidence Limits of real power flows and voltage angles increase, those of reactive flow and voltage magnitudes are negligibly affected. The implication of these results appear to be that real power is a much more sensitive quantity to variations in load and generation data error bounds than reactive power. Reactive power produced by transmission lines appears to be the cause. The same voltage magnitude error bounds were used in Cases 3 and 5. However, Cases 1 and 2 showed that reactive power was very sensitive to voltage magnitude errors. Thus, larger voltage magnitude errors in Case 5 would have increased the Confidence Limits of reactive power.

Case 6

This case uses the final formulation of the stochastic load flow modelling the slack bus equation as any other generator and converting its role only to that of voltage angle reference. A total system load plus losses inaccuracy of about 3% resulted in a 210 MW error bound which was distributed evenly among the 14 generators. Thus, the difference between Reference 7 and Case 6 data are only in generator real power error bounds. Apart from the confidence limits of voltage angles, all other confidence limits were invariant as the reference bus position was altered.

It is interesting to note that the research program used throughout this investigation calculated the confidence limit of the reference bus real power as an output quantity. In Case 6, however, the error bound of P, was given a value of 15 MW as an input quantity. This same numerical value was produced by the program as a calculated quantity, confirming that there was no filtering effect caused by the redundancy of the slack real power equation.

CONCLUSIONS

This paper has discussed the extension of the conventional load flow problem to include the calculation of the effects of inaccuracies in input data on all output quantities. The authors have recently become aware of a paper [8] addressing the same problem but with a different solution approach.

Three models were presented in the paper:

1. A model in which all input data was assigned error bounds.
2. A second model including constraints on load and generation of areas within a system.
3. A final model in which a constraint on total system load plus losses was placed by including a real power equation for the slack bus.

As a conclusion of the theoretical discussions and numerical solutions using the three models, the third model is recommended for implementation for being conservative in case requirements, for being independent of voltage angle reference bus location and for handling the important practical constraint on total system load plus losses. An attractive characteristic of the method is that it involves a series of non-iterative calculations to be carried out after the solution of a conventional load flow by any method.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the suggestion made by Mr. A. F. Gabrielle, Head of the Computer Application Division of AEP, to study the effect of area constraints in the formulation of the stochastic load flow. The authors' appreciation to Mr. F. Aboytes of Imperial College, London for his private discussions on the statistical significance of some of the models considered is also gratefully acknowledged.

REFERENCES

1. J. F. Dopazo, O. A. Klitin, G. W. Stagg and L. S. Van Slyck, "State Calculation of Power Systems from Line Flow Measurements," IEEE PAS-89, pp. 1698-1708, September/October, 1970.

2. J. F. Dopazo, O. A. Klitin, and L. S. Van Slyck "State Calculation of Power Systems from Line Flow Measurement, Part II," IEEE PAS-91, pp. 145-151, January/February, 1972.

3. J. Dopazo, O. Klitin, A. Sasson, L. S. Van Slyck, "Real-Time Load Flow for the AEP System," Paper No. 3.3/8, 4th Power Systems Computation Conference Proceedings, Grenoble, France, September, 1972.

4. J. F. Dopazo, S. T. Ehrmann, O. A. Klitin and A. M. Sasson, "Justification of the AEP Real-Time Load Flow Project," IEEE Paper No. T 73 108-8, Winter Power Meeting, New York, 1973.

5. J. F. Dopazo and A. M. Sasson, "AEP Real-Time Monitoring Computer System," Symposium on Implementation of Real-Time Power System Control by Digital Computer, Imperial College of Science and Technology, London, September, 1973.

6. J. F. Dopazo, O. A. Klitin and A. M. Sasson, "State Estimation for Power Systems: Detection and Identification of Gross Measurement Errors," Proceedings of the 8th IEEE PICA Conference, June, 1973.

7. L. S. Van Slyck and J. F. Dopazo, "Conventional Load Flow Not Suited for Real Time Power System Monitoring," Proceedings of the 8th IEEE PICA Conference, June, 1973.

8. B. Borkowska, "Probabilistic Load Flow," IEEE Paper No, T 73 485-0, presented at the Summer Meeting, Vancouver, 1973.

Discussion

A. Semlyen (University of Toronto, Toronto, Ontario, Canada): This is a very timely paper. Probabilistic methods are of increasing significance in many power system studies due to the prohibitively large computer requirements for handling a huge number of individual deterministic problems resulting from many different combinations and magnitudes of the input variables. Fortunately, linear functions of independent random variables, with Gaussian distribution, are also Gaussian and their statistical characteristics are easy to correlate with those of the input variables. The merit of the paper consists in applying this fact to the load flow problem where the forecasted data are only estimated.

I would appreciate clarifications on the following details.

1) The load flow problem is basically non-linear and its linearization produces some inaccuracy in the simple relationship between statistical characteristics. This may be insignificant in many cases but may have importance in long range planning. Some variables may be more affected than others. Would the authors comment on the effect of linearizing the load flow solution?

2) Some input variables, like P and Q at the same bus, are apparently not uncorrelated. How would this affect the general theory and would it significantly alter the results based on the approach adopted in the paper? I feel that even if the matrix V is not strictly diagonal (say, block-diagonal) and the inputs close to Gaussian (which, probably, is in general a reasonable assumption) the Central Limit Theorem will still apply for practical evaluations.

My belief is that the application of stochastic methods to this important power system problem will prove to be stimulating to engineers engaged in research in some other areas of power systems where direct statistical results are more meaningful and practical than very large numbers of deterministic calculations.

Manuscript received August 5, 1974.

H. Duran (University of the Andes, Bogota): The authors should be commended on their effort to formulate and solve a problem whose importance has not yet been fully recognized.

As described in the introduction, the stochastic load flow problem is concerned with finding the probability distribution, and in particular the expected value and the variance of the solution of a load flow problem. As such, it is a problem of probability calculus and not one of statistical estimation. Hence the approach that the authors take to solve the problem using statistical principles appears to be misleading and unnecessarily complicated.

A more direct approach would be as follows. Let $\bar{y}$ and V be the expected value and the covariance matrix, respectively, of the input data. Let $\hat{x}$ be the solution of the load flow problem using $\bar{y}$ as data, that is

$$\bar{y} = f(\hat{x})$$

Using a Taylor's series expansion around $\hat{x}$, and neglecting second and above order terms gives:

$$y - \bar{y} = f(x) - f(\hat{x}) = J(x - \hat{x})$$

Finally,

$$E(y - \bar{y}) = 0 = J\left(E(x) - E(\hat{x})\right) \quad \text{or} \quad E(x) = \hat{x}$$

and,

$$\mathrm{cov}(x) = E\left((x - \hat{x})(x - \hat{x})^t\right) = J^{-1} E\left((y - \bar{y})(y - \bar{y})^t\right) J^{-1t} = J^{-1} V J^{-1t}$$

One important question that the authors do not consider in their paper relates to the accuracy of the formulas. It should be borne in mind that due to the linearization introduced in the Taylor's expansion $\hat{x}$ is not the actual expected value of x, and equation (10) does not give the exact value of cov (x). To see this, consider the following two-node example. A generator supplies a load of P MW at unity power factor and unit voltage magnitude through a transmission line of resistance R and negligible capacitance. The Generation G is then given by,

$$G = P + RP^2$$

Let us assume that P has a normal distribution with mean $\bar{P}$ and variance σ_p^2. What is the distribution, mean and variance of G? It can easily be shown that:

$$\bar{G} = E(G) = \bar{P} + R(\bar{P}^2 + \sigma_p^2)$$

and

$$\sigma_g^2 = \sigma_p^2 + 4R\bar{P}\sigma_p^2 + 4R^2\bar{P}^2\sigma_p^2 + 2R^2\sigma_p^2$$

while, the formulas in the paper give instead,

$$\bar{G} = \bar{P} + R\bar{P}^2$$

and

$$\sigma_g^2 = (1 + 2R\bar{P})^2\sigma_p^2 = \sigma_p^2 + 4R\bar{P}\sigma_p^2 + 4R^2\bar{P}\sigma_p^2$$

Hence the errors in the expected value and the variance of G, are $R\sigma_p^2$ and $2R^2\sigma_p^2$ respectively. Are these errors significant? and, if they are, how can they be calculated or estimated in the general case? If the error introduced in calculating the expected value of a variable is of the same order of magnitude of its standard deviation, what degree of confidence could we have in the Confidence Limits? Regarding the probability distribution of G, one thing that can be said is that it does not have a normal distribution since the P^2 term gives rise to a χ^2 pattern. I could not follow the author's use of the Central Limit Theorem to conclude that the output quantities can be taken as normally distributed. Would they like to comment on the assumptions underlying its use here?

As a final remark I would like to compliment the authors again for pointing at a very important, interesting and difficult problem. They have given one of the first bites to a hard bone. I hope this discussion will encourage them to continue their work since the problem has not been solved yet.

Manuscript received August 7, 1974

R. N. Allan and C. H. Grigg (Univ. of Manchester Inst. of Science and Technology, Manchester, England): We would first like to state how much we agree with the authors for the need to treat the power flow problem probabilistically. We also feel that, because the variables involved vary statistically and are forecasted statistically, deterministic calculations can lead to erroneous planning and operating conditions and are, at best, only subjective assessments.

We would, however, like to comment on the author's use of the Central Limit Theorem to assume that $\pm 3\sigma$ predicts satisfactory confidence limits. We, in a mutual collaborative effort between UMIST in England and the Institute of Power in Warsaw, have been currently investigating the same problem. We, however, not only characterise the power flows by expected values, standard deviations and confidence limits similar to that proposed by the authors, but also calculate the complete probability density curves of the power flows of interest. It was in order to calculate these density curves that the initial formulation, as described by Borkowskia,[1]. was limited to the d.c. case since these calculations are inherently complex.

One of our typical calculations produced the density curve shown in Fig. 1. This is clearly not a normal distribution and suggests that the Central Limit Theorem does not apply. The difficulties arise because the Theorem is only applicable for a large number of random variables which is not necessarily the case when analysing typical power systems. One consequence is illustrated in Fig. 1 where the value of the density function at a power level of E + σ is nearly three times greater than at a power level of E.

We can confirm that all our results to date indicate that the $\pm 3\sigma$ limits enclose more than 99% of the probabilities of occurrence as assumed by the authors. We feel, however, that this could still lead to erroneous decisions if power flows of the type shown in Fig. 1 occur. In this example, the $\pm 3\sigma$ confidence limits enclose power levels between 100 and 1100 MW. Our results show, however, that the probability of power flows greater than 800 MW is negligible. Therefore, with the authors' confidence limits, the line could be almost 40% overdesigned.

We accept our results were obtained using a d.c. representation and objections may be raised concerning their accuracy. Preliminary work using an a.c. model, however, indicates that the same trend prevails.

In ending we would like to add our weight to the authors' belief that probabilistic assessment of system behaviour is necessary to enable

Manuscript received May 6, 1974.

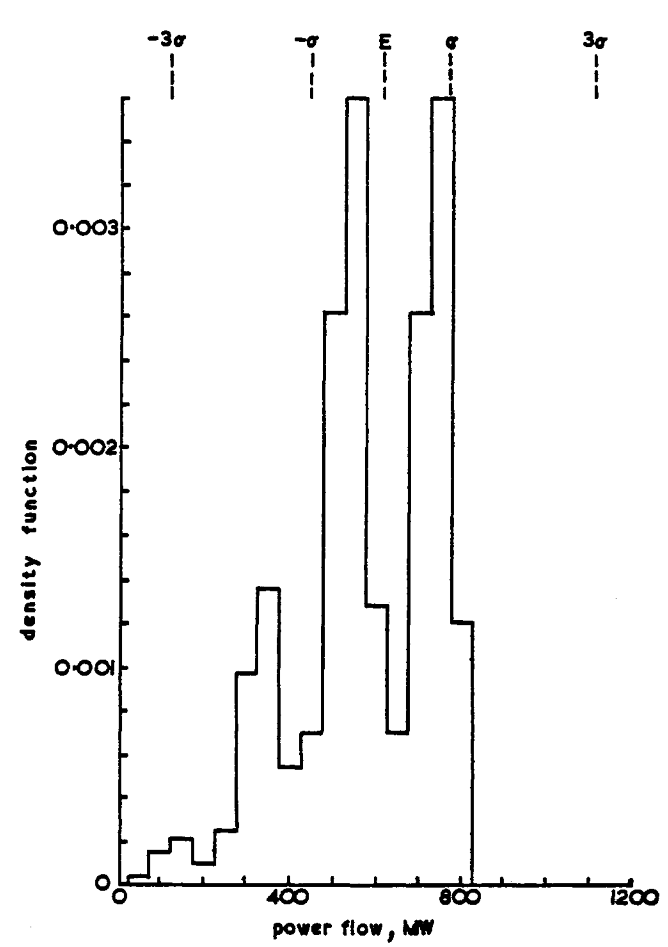

Fig. 1 – Typical probability density curve

accurate security prediction of a present system and a balance between cost and security of a future planned system to be treated objectively and not, as at present, subjectively. We consider, however, that for probabilistic methods to supercede presently used deterministic methods, the full potential of probability assessment must be utilised as we are currently pursuing.

REFERENCE

[1] Borkowska, B. "Probabilistic Load Flow", IEEE paper no. T73 485-0, presented at Summer Power Meeting, Vancouver, 1973.

Barbara Borkowska (Instytut Energetyki, Warsaw, Poland): I regard the method presented by J. F. Dopazo, O. A. Klitin and A. M. Sasson as very interesting. The method seems to be useful and efficient in numerical calculations. The uncertainty of the input data can be considered from two points of view: "a posteriori" [uncertainty of the past or present state of system] and "a priori"[uncertainty of the future state of system]. The uncertainty a posteriori has given the impulse to elaborating the various methods known as "state estimation". The formulation of the load flow problem used in a posteriori study was taken by the authors as the starting point to elaborate the method of a priori analysis. This is right since the redundancy of data takes place in a priori study too. The use of global forecast for areas in order to limit the variance of output quantities is an interesting idea. However, the uncertainty of a posteriori data is at least in two points different from this of a priori.

1. The correlations between the measurement errors often are negligible while the errors of the forecasted input data may be strong correlated because of nodal input powers being the function of the some random parameter [temperature, the configuration of lower voltage network etc.]. Therefore the matrix V defined by the equation 3 is not diagonal. It complicates the computations but not the mathematical model. It should be noted that taking into account the correlation diminishes the variance of the output quantities.

2. The variance of the input data a posteriori is constant while the variance a priori is variable according to the time horizon. The further in future the forecasted moment [time horizon] the greater is the variance of input data. It leads to great uncertainty of the results. Taking advantage of the additional information makes the space of the possible system states limited. This result was achieved by the authors by introducing additional Equation [29] which is to be combined with these [35] to [37]. Particularly interesting seems to be the Equation [36] which means that the errors of individual data are not correlated to the error of global forecast. I regard such assumption as fully acceptable because the methods and the data used for individual and global forecasting are different.

As to Central Limit Theorem some comment seems to be necessary. Since Jacobian J is sparse, the variances of the input data are different, and the density curves of the errors of the input data are not defined, then it is not obvious that the CLT holds. It is known that, especially for the further time horizon, the density curves of the power generation (of the power stations with few large generating units) differs substantially from the Gaussian curve. The some might apply to the density curves of the power flows in branches connected to generation nodes. In such a case the Eq. [27], [28] are not true. However in cases in which the density curves of errors of input data can be properly defined [a.g. as a normal] the equations hold.

It would be interesting to know to what level of error bounds the linearisation applied in going from Equation [1] to Equation [2] is valid.

Concluding the method seems to be very useful in power system operational problems and in some problems of power system planning.

Manuscript received August 13, 1974.

G. T. Heydt (Purdue University, West LaFayette, Indiana): This paper presents a new formulation of the power flow problem in stochastic terms in which the load/generation schedule is a random vector. The solution vector to the power flow problem is written in a linear approximation to the nonlinear power flow problem, and in such a formulation, the solution is a linear transformation of the input (where the "input" is interpreted as the load/generation schedule). The mathematical formulation involves the usual decoupling process between bus voltage magnitudes and angles, and therefore in the Newton-Raphson solution, only real numbers appear. The central limit theorem prescribes that the distribution of the solution variables will be Gaussian since these variables are linear combinations of a large number of random variables. Messrs. Dopazo, Klitin and Sasson have made a valuable contribution in their observations of the various intricacies of the stochastic power flow formulation. I have worked on this problem from a different point of view, and I would like to briefly present this alternate approach in order to elucidate certain generalizations of the methods used in this paper. The alternate approach also raises some interesting questions which I would like to pose to the authors.

An alternate formulation is obtained by ignoring bus voltages as problem variables; instead, consider line complex power flows as "output" variables and bus demand/generation as "inputs." In linear formulation, the complex factors relating these variables are known[1] and the output is simply a linear transformation of the input. Unlike the formulation in this paper, the variables so formulated are not real but complex. Data used for load and generation schedules are random complex vectors of known distribution (or at least known mean and covariance). The covariance matrix should not be considered diagonal since there is no reason to believe that bus demands at several different busses are independent. The covariance matrix is Hermitian, however, as may be observed from the definition of element i, j of the covariance matrix V,

$$(V)_{ij} = E((x_i - \mu_i)(x_j - \mu_j)^*) \tag{1}$$

where the μ entries are the means of the random variables, x, and $(.)^*$ denotes complex conjugation. In vector notation,

$$V = E((X - M)(X - M)^H) \tag{2}$$

where M is the mean vector and $(.)^H$ denotes the Hermitian operation (complex conjugation followed by transposition). Comparison of Eq. (2) of this discussion with Eq. (3) of the paper shows the generalization necessary to approach this problem as I suggest. When the problem is formulated with line power flows rather than bus voltages, not only does

Manuscript received August 13, 1974.

the mathematical treatment change, for example, Eq. (5) of the paper becomes

$$\hat{x} = (J^H V^{-1} J)^{-1} J^H V^{-1} y,$$

but also the information derived changes. Line loading is obtained directly, and many of the power flow statistical inferences carefully described in this paper are also available. The advantage of the method of Messrs. Dopazo, Klitin and Sasson primarily lies in the availability of the Jacobian matrix. The advantages of the alternate formulation which I suggest primarily lie in the speed of computation (no inversion is required) and certain other computational factors.

I would like to raise some questions of the authors in order to clarify some points. Why is covariance matrix V,

$$V = E(\epsilon, \epsilon^t)$$

treated as diagonal? Presumably, if the statistics of ϵ are available, even cases where independence does not hold could be handled. Secondly, I would like to ask a difficult question concerning the accuracy of the calculation of the line flows (labelled z in the paper) by the method given by the authors. When the method described in the paper is used, a linearization is required to obtain bus voltage information (Eq. (1) of the paper) and a second linearization is required in order to obtain the line flows (Eq. (18) of the paper). Would more accurate statistics of z be obtained by formulating the load flow problem directly with z as the independent variable rather than x? When this is done, only one linearization is needed.

This paper presents a well written account of stochastic calculations in power flow studies. I have no doubt that the results and techniques will be used by others to obtain many other useful statistical inferences of electric power system operation.

REFERENCE

[1] G. T. Heydt, B. M. Katz, "A Stochastic Model for Simultaneous Interchange Capacity Calculations," to be published, IEEE Trans. on Power Apparatus and Systems.

K. F. Schenk and **K. Singh** (University of Ottawa, Ottawa, Ontario, Canada): This paper is another one of a series of papers which the authors have written in their systematic efforts to develop practical applications of statistical notions to the analysis of power system problems. This is a natural development when one recognizes that deterministic models and solutions do not fit well with real-type problems. Although deterministic results have a greater appeal than stochastic results, the latter recognizes the impossibility of exact parameters, data and models. This is, in our view, a step forward in the right direction. The authors are to be congratulated for their timely efforts.

As is pointed out in the paper, the vector of 'error' random variables, ϵ, may be considered to be normally distributed as a consequence of the Central Limit Theorem. This being the case, the estimators for x and $\sigma_x{}^2$ (and for z and $\sigma_z{}^2$) may be obtained by maximizing the likelihood function. This is, of course, a well known fact, but it gives an additional insight into the formulation of the problem. Moreover, maximum likelihood estimators (m.ℓ.e.) which are unbiased (not all of them are!) also exhibit some other properties which make them very desirable estimators: minimum variance, consistent, efficient, sufficient, complete and independent. Furthermore, under quite general regularity conditions on the density function[1], any function u $(\hat{x})$ of a m.ℓ.e. $\hat{x}$ of $\acute{x}$ is the m.ℓ.e. of u(x).

In addition, it may be worthwhile to point out that useful and meaningful results can be extracted from the method if it is properly applied. Only those cases which are considered normal with small data variations about an operating point will give useful results. Extreme variations in the input data (which are allowed by a normal distribution) may give uninterpretable results.

REFERENCE

[1] F. A. Graybill, "An Introduction to Linear Statistical Models", Volume 1, McGraw-Hill, 1961.

Manuscript received July 2, 1974.

S. T. Despotovic (Research Institute Nikola Tesla, Beograd, Yugoslavia): The authors have presented an interesting, useful and simple method for calculating the effect of the propagation of forecasted data inaccuracies through the conventional load flow. The method, using the principles of statistical least squares estimation for linear systems and addressing itself to the problem of processing the expected errors in the input data, that is, load bus P and Q, and generator bus P and E, converts the load flow problem formulation from a deterministic one to a stochastic one. The proposed method, through the extension of the conventional load flow problem, has included the calculation of the effects of inaccuracies in input data on all output quantities, namely on all bus state variables E and δ or on line flow P and Q and generator bus Q. The stochastic load flow represents, in this way, the combined results of many deterministic load flows in which for each flow the data are perturbed, such that the various flows represent possible sets of data within the precision that the input data is known.

The model of the load flow in which a constraint on total system load plus losses has been placed by including a real power equation for the slack bus can be considered as a realistic one, since the forecasted errors are then distributed among generator P equations. It is to be believed, that further experience in using the developed method will make it simpler and more attractive.

The authors should be complimented for this nice paper.

Manuscript received August 13, 1974.

A. Petroianu (National Center of Romanian Power System, Bucharest, Romania): The authors are to be commended for a thought provoking paper related to an interdisciplinary in character problem.

The paper is devoted to the research problem of influence of initial data random errors upon the results of load flow calculation.

Such an investigation is very timely for the comparison of different mathematical models suitable for "off" and "on" line load flow calculations.

The authors have a probabilistic approach to the problem of estimation of the results of calculation. From the probabilistic point of view the initial data should be considered as random variables with multidimensional distribution law.

We can underline two distinct direction of research in the frame of the probabilistic approach:

1–**Initial data errors are of stochastically definite nature**

In this case the initial data are given by their mean value (mathematical expectations) and by their extreme limits of deviations (errors) from mean values. The law of distribution of initial data errors is supposed to be known.

It seems to the discusser that such an approach was used in[1].

2–**Initial data errors are of stochastically indefinite nature**

In this case only the extreme values of initial data errors are known.

I believe that considering "a model in which all input data was assigned error bounds" the authors are approaching a more practical way since the law of distribution of errors is often unknown.

In the frame of the probabilistic approach, considering that the deformation of the initial data is small, the knowledge of the expectations and of the variance-covariance matrix sufficiently characterizes the distribution law of the calculated values which constitutes our aim.

I would like to remark that a geometrical approach could introduce some fresh point of view that corresponds to the necessity of reducing the great output data streams, generated by the computer, to a small number of pertinent, synthetic and intuitive information [2 ÷ 5].

The geometrical theory, which can be conceived in parallel to the probabilistic approach, permits us to find and to describe an area of possible errors and the guaranteed region to which the calculated values belong.

$$x = \varphi(y) \tag{1}$$

where x, y are column vectors in m– and n– dimensional Euclidean spaces

$$y = \bar{y} + \delta y \tag{2}$$

where δy– small random vector with variance-covariance matrix Δy.

In accordance with the principle of linearization:

$$x = \varphi(\bar{y}) + \delta x \tag{3}$$

$$\delta x = J \delta y \tag{4}$$

Manuscript received July 30, 1974.

where J Jacobian matrix for φ taken at the point $\bar{y}$.

The variance-covariance matrix Δx for δ x can be calculated accordingly,

$$\Delta x = E\left[(\delta x)(\delta x)^T\right] =$$

$$= \left[J(\delta y)(\delta y)^T J^T\right] = J \Delta y J^T \qquad (5)$$

and characterizes the errors in the solution vector x due to the non-exactness in giving initial vector data – y.

Formulae (5)* corresponds to the general matrix expression of a quadratic form: $x_T^* Ax$

In the realm of an n-dimensional space with which is our present concern this quadratic form can be geometrically interpreted as representing the area of the error hyperellipsoid.

Even under the simplifying assumptions that the admissible errors of x and y are sufficiently small to be related by the linear transformation (4), the small region in which δ y is contained is stretched in one direction and compressed in the other.

If the estimate of errors is made by the spectral vector norm, then the admissible regions for δy will be an hyperellipsoid.

If in turn we estimate δx by the spectral norm we shall be forced to take as admissible the hypersphere with radius equal to the largest half-axis of the hyperellipsoid. This hypersphere will be larger than the hyperellipsoid; this replacement of the hyperellipsoid by the hypersphere involves the loss of information due to the uncertainty and errors in input data.

Formulae (5) expresses a hyperellipsoid which is in a skew position with respect to the oblique coordinate axes; by means of a suitable transformation to the principal axes of the hyperellipsoid the equation can be reduced to the canonic form

$$\underline{\lambda}_1 \underline{w}_1^2 + \ldots + \underline{\lambda}_n \underline{w}_n^2 = \Delta x \qquad (6)$$

where $\underline{\lambda}_1 \ldots \underline{\lambda}_n$ are the eigenvalues.

That is equivalent to consider only the variance of components (i.e. ignoring the correlation coefficient); in this case we shall obtain the confidence hyperellipsoids with the axes being directed along the coordinate axes.

The degree in which the uncertainty of basic information is "conditioning" the uncertainty of the steady state solution is expressed by the square root of the ratio of the eigenvalue of maximal modulus to the eigenvalue of minimal modulus.

Formulae (7) which gives the elliptic norm (conditionality number),

$$H = \sqrt{\frac{\lambda_{MAX}}{\lambda_{min}}} \qquad (7)$$

can be geometrically interpreted as the ratio of the hyperellipsoid semiaxes.

By virtue of the above formulae it is evident that the precision, convergence and stability properties of the load flow solution are optimum for

$$H = 1,$$

that corresponds to the perfect symmetry of a hypersphere.

The discusser's comments, regarding the geometrical approach, were intended to underline the importance of the authors' contribution to a new emerging research problem – the probabilistic load flow problem – and to add some bits of information in an effort to relate various existent schools of thought.

REFERENCES

[1] Power System Computer Feasibility Study – 1968 – vol. I, IBM Research Division, San Jose, California.

[2] C. Lanczos, "Applied Analysis" Prentice Hall, 1956.

[3] A. Petroianu "A geometrical approach to the steady state problem of electrical networks". Rev. Roum. Sci. Techn.–Electrotechn. et Energ., 14, 4 p. 623-630, Bucharest, 1969.

[4] V. I. Idelchik "About the error influence of initial data upon the result of power flows calculations in power systems" Izvestia, the USSR A.S. Energetics and Transport, 1968, No. 2, p. 9-15.

[5] D. K. Faddeev, V. N. Faddeeva "Stability in Linear Algebra Problems" A.F.I.P.S. Congr. 1972. Ljubliana.

B. F. Wollenberg (Leeds and Northrup Company, North Wales, Pa.): This is a very important paper. The techniques used in calculating the covariance matrices of system states and quantities computable from the states are definitely useful in all analytical studies requiring a load flow solution. I agree that any time a load flow is run, whether for real time studies or planning studies, a recognition of the data inaccuracies is always useful – and necessary.

The authors have been careful to develop a technique which is used in conjunction with a conventional load flow solution and not a least squares state estimator type solution. The primary contribution of the paper is in showing how to compute the covariance matrices when adjustments such as area load or system total load, known to some specified accuracy, are placed on the solution. The authors correctly point out that such adjustments do not imply anything more than a conventional load flow solution. The covariance matrix calculations used to model the effects of the area or system load adjustments are correct for a least squares solution to the load flow with the area or system load values acting as redundant information. When a load flow is solved with the loads preadjusted to meet area or system conditions, the number of degrees of freedom remains the same and there is no filtering effect. A least squares solution with area or system load values as redundant information will produce filtering. The statement in the description of the authors' Case 6 that "... there was no filtering effect caused by the redundancy of the slack real power equation." is somewhat misleading. It should have read "... no *apparent* filtering...", that is, there was no filtering which could be measured. In part this is due to the fact that the error bound of 210 MW was allocated to the 14 generators by dividing 210 MW by 14 to give a 15 MW error bound. This calculation should have divided 210 MW by the square root of 14 which would have given a 56.1248 MW error bound for the generators. The 15 MW used by the authors was so much smaller than the 40% error band on the loads that any filtering effect would be minimal. This result is probably true for the 56.1248 MW error bound also and therefore shows that the covariance matrix calculation used is quite adequate.

Manuscript received August 2, 1974.

J. F. Dopazo, O. A. Klitin, and A. M. Sasson: We are very pleased to the response of so many discussers to our paper and acknowledge their contributions to the subject. We are encouraged that in all cases the discussers emphasize the need for treating the load flow stochastically.

Messrs. Allan and Grigg and Ms. Borkowska present approaches in which they calculate actual density curves with simplified models. We consider that the distribution of input quantities, which is needed in their method, is generally not known. For instance, consider the generation of a large plant with several units which Ms. Borkowska mentions. Over a period of a day the output of the plant can vary substantially and a non-normal distribution could be determined. However, the distribution of plant output at peak system conditions at some future year can be considered as a normal distribution with a small variance that depends significantly on what the distribution of the system peak is. The same discussors and Dr. Duran, express reservations upon the applicability of the Central Limit Theorem. The reasoning given for the distribution of input quantities also apply here. Line flows near generating plants, which is the example mentioned by Ms. Borkowska, can thus be considered as normal. Dr. Duran finds some chi-squared component in the output when no linearizations are made. We agree that results are only approximately normal in the sense Dr. Duran points out. However, the chi-squared distribution rapidly approaches a normal as the degrees of freedom increases which justifies our use of the CLT for practical purposes. On the other hand, we are really after the bounds of output quantities and not their distribution. The unknown but bounded theory (9) applies here and our procedures are further justified.

The arguments given above apply to a large extent to the comments made by Messrs. Heydt and Semlyen and Ms. Borkowska on the use of a diagonal input convariance matrix. Even in the case of P, Q values at the same bus which Dr. Semlyen refers to, while the variations during a day are strongly correlated our expectation of P and Q at some future year for system peak conditions are considerably less so.

Messrs. Schenk, Singh, Semlyen and Duran and Ms. Borkowska ask on the accuracy of results for large input variances, given the linearizations involved in our method. Obviously, one can not expect to increase input variances indefinitely without serious deterioration in the accuracy. We would advise that tests similar to the ones we performed be made to answer this question for a specific system. If Monte Carlo load flows results can be reasonably duplicated by the stochastic load flow, then the latter applies for that level of input variances.

Manuscript received November 4, 1974.

We were pleased to find the comments made by Ms. Borkowska and Dr. Despotovic on our differentiation between global and individual forecasts and the inclusion of total load error as variances over all generators thus eliminating the need for defining a slack bus. On this point, Mr. Wollenberg feels that our calculation of individual generator variances is in error. Considering a total load error bound of 210 to be spread over 14 generators he considers that the sum of the variances of generators should be equal to that of the total. Thus,

$$210 = 3\sqrt{\sum_{1}^{14}(a/3)^2} = \sqrt{14}\ a$$

What we do is to divide the 210 evenly among the 14 generators. With Mr. Wollenberg's suggestion we would be considering load flow situations in which generation exceeds the bounds of total system load thus being unrealistic.

Dr. Heydt briefly presents a method he has developed in which he states there are some computational advantages. We understand that he will be presenting his method in a paper that will be presented shortly and we would like to reserve our comments until we read his forthcoming paper and understand his approach better. We will only point out that the use of sparsity techniques makes our method computationally attractive.

We appreciate Dr. Petroianu's comments on our paper and his geometric interpretation of the mapping of input errors to output quantities.

Dr. Duran's derivation of the paper's equations are the same as ours. We "complicated" matters by trying to show the relation between the theory of state estimation and that of stochastic load flows to emphasize that historically the idea of the latter came from the former.

We find Dr. Duran's error analysis quite enlightning. However, the last term of the exact formula for the variance of G should be $2\ R^2\sigma_P{}^4$ and not $2\ R^2\sigma_P{}^2$. This makes the errors in the mean and variance to be $R\sigma_\rho{}^2$ and $2\ R^2\sigma_P{}^4$ which are second order effects in their respective equations. For instance, consider as an example that per unit R and P are equal to 1 and an error bound for P equal to 0.3. This makes $\sigma_\rho{}^2 = (0.3/3)^2 = 0.01$. Then, the error in the mean and variance of G are 0.01 and 0.0002 while our computed mean and variance are 2 and 0.9 respectively.

REFERENCE

[9] F. C. Schweppe "Uncertain Dynamic Systems," Prentice-Hall Inc., Englewood Cliffs, N. J., 1973.