IEEE Transactions on Power Apparatus and Systems, Vol. PAS-100, No. 7 July 1981

STOCHASTIC OPTIMAL ENERGY DISPATCH

G. L. Viviani, Member, IEEE
E. I. DU PONT DE NEMOURS
Engineering Department
Wilmington, DE 19898

G. T. Heydt, Member, IEEE
PURDUE UNIVERSITY
School of Electrical Engineering
W. Lafayette, IN 47907

Abstract - This paper presents a useful algorithm to incorporate the effects of uncertain system parameters into optimal power dispatch. This method employs the multivariate Gram-Charlier series as means of modeling the probability density function (p.d.f.) which characterizes the uncertain parameters. The proposed method is a direct extension of existing techniques; therefore, minimal additional computation effort is necessary. Potential applications include the calculation of probabilistic parameters (eg., probability, expectation) relating to power dispatch. An example of such a parameter the probability of a bus voltage out of range during optimal dispatch. A practical example is pursued to demonstrate the usefulness of the method.

NOMENCLATURE

ARG	angle of a phasor quantity
B_{0i}, B_{1i}, B_{2i}	cost coefficients
COV	covariance matrix
DET	determinant
E( )	expectation
F(x,u)	cost function
G(x,u,p)	equality constraints
$H_r(x)$	rth Hermite polynomial
L(x.,u,λ)	Lagrangian
m_i	ith statistical moment
MGC	multivariate Gram-Charlier (series)
p	parameter vector
P_j	real injected power at bus j
Q_j	reactive injected power at bus j
u	control vector
V_k	complex voltage at bus k
x	state vector
Y_{ij}	i,j th element of Y-bus matrix
$( )_G$	pertaining to a generator
$( )_L$	pertaining to a load
λ	Lagrange multiplier

1. INTRODUCTION

The economic dispatch problem of energy systems maintains increasing importance as operating costs continue to escalate. The aims of the techniques described in the paper are to identify sources of uncertainty in the optimal dispatch problem and to calculate the effects of the uncertainties in dispatch algorithms. Uncertainty is introduced into system analysis from various sources including:

1. long and short term forecast error
2. measurement or telemetering errors
3. system configuration error.

Therefore, in this paper it is propossed to consider uncertainty in optimal dispatch for the purpose of both system operation and system planning. In the case of system operation, the appropriateness of future dispatch schedules may be assessed. In the case of system planning, alternative plans may be evaluated on the basis of their dispatch. In both cases, the probability of occurance of operating scenarios may be evaluated.

The various sources of error are grouped into a single parameter vector, p,

$$p = \left[P_{Li}\ ,\ Q_{Li}\ ,\ ARG(Y_{jk})\ ,\ |Y_{jk}|\right]^T$$

where i represents load busses. Figure 1 illustrates standardized (zero mean and unit variance) variations of a typical load versus time (days). Superimposed is a curve which is intended to represent a typical approximation to the actual load variations. Since optimal dispatch considers only a static (as compared to dynamic) optimization of the objective function, the dispatch is evaluated at times such as those indicated by a,b, or c. Clearly, there is some uncertainty in the actual value of the demand at points a,b, or c. Neglecting the effects of this uncertainty introduces error into the deterministic optimal dispatch algorithm which may lead to unnecessary additional operation cost.

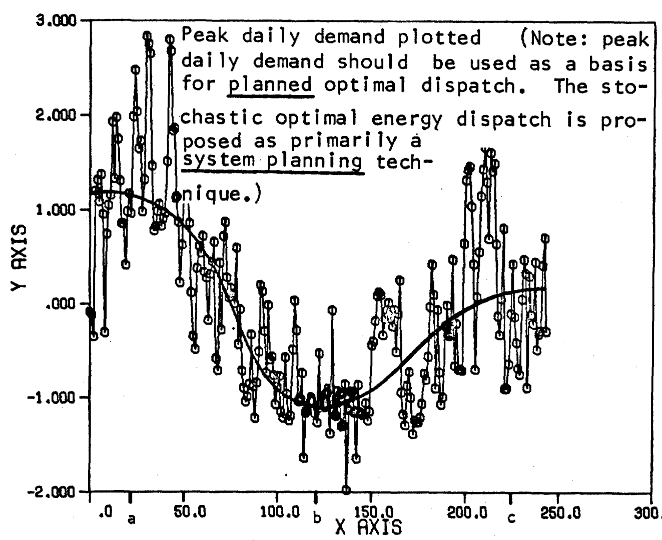
To make further analysis more convenient, the remaining energy system parameters are grouped into the vectors x (state vector), and u (control vector), where x is,

$$x = \left[ARG(V_k)\ ,\ |V_l|\right]^T$$

and u is,

$$u = \left[P_{Gj}\ ,|V_m|\right]^T.$$

In the above vectors, k, l, j, and m represent all load busses except the slack, all load busses, all generator busses except the slack, and all generator busses respectively.



X AXIS = time (days)
Y AXIS = standardized load data

Figure 1 Typical loading of a power system

81 WM 062-9 A paper recommended and approved by the IEEE Power System Engineering Committee of the IEEE Power Engineering Society for presentation at the IEEE PES Winter Meeting, Atlanta, Georgia, February 1-6, 1981. Manuscript submitted May 23, 1980; made available for printing November 5, 1980.

To account for the random behavior of the loading conditions (or another condition), it is assumed that the components of the p vector may be modeled as,

$$p_i(n) = E(p_i(n)) + \Delta p_i(n) \quad n = a,b,c.$$

The statistics of the Δp variations will be studied assuming strict sense stationarity [1] in any given time interval. The fluctuations $\Delta p_i(n)$ are represented by a probability density function (p.d.f.) which is known or may be approximated by random sampling. The stochastic optimal energy dispatch (SOED) problem concerns the minimization of a cost function, F(x,u), subject to the random fluctuations of p. In the deterministic optimal dispatch problem, a deterministic optimal control vector u^* is found; the stochastic optimal dispatch problem solution results in a p.d.f. representation of the changes in the optimal control vector Δu^* about an operating point. The random fluctuations in p are approximately linearly related to u^* allowing the p.d.f. of Δu^* to be calculated from the p.d.f. of Δp. The p.d.f. of Δu^* provides useful information about the likelihood of certain events associated with u^*. The probability of events pertaining to u^* are valuable in assessing the ability of a system to perform at minimal cost; further, this formulation is a more accurate and realistic representation of the optimal dispatch problem.

2. STATISTICAL THEORY

The solution of the SOED problem requires theoretical development in several statistical areas. These areas include the multivariate Gram-Charlier series and transformations to normality. These topics are briefly discussed in this section; however, more detail is presented in the appendices. The multivariate Gram-Charlier series is utilized to represent an n dimensional p.d.f. [2]. The MGC series is an infinite series composed of Hermite polynomials (see APPENDIX A). The p.d.f. characterizes a data set of independent identically distributed random observations found by substituting the sample moments of the observations into the MGC series expansion. (Note: the variates of the multidimensional random vector modeled by the MGC series need not be independent and, in fact, are usually correllated). This procedure is done for only a finite number of terms, resulting in a truncated approximation to the actual p.d.f. Accurate results are usually available from approximations involving up to all permutations of fifth order terms. The MGC series is primarily limited in application due to storage limitations. However, p.d.f.'s of order $n \leq 10$ are readily accommodated.

The MGC series converges more rapidly when the standardized random vector being modeled, z, is approximately multivariate Gaussian,

$$f(z) = \frac{EXP(\frac{-1}{2} z^T (COV(z))^{-1} z)}{(\sqrt{2\pi})^n \sqrt{DET(COV(z))}}$$

where n is the dimension of z. Transformations to normality [3,4] have application to this study as a method for reduction of the approximation error introduced by truncating the series. The required number of statistical moments to be calculated for use in the MGC series rises sharply with the number of terms retained. Therefore, any method which permits early truncation of the series is highly desirable. The conditions under which improved accuracy in the truncated MGC series is attained are outlined in [5]. Various methods for enhancing the joint normality of a general p.d.f. are available but this paper relies on a more novel polynomial transformation (APPENDIX B). Each of the components, z_i, of an n dimensional random vector z are transformed to enhance the normality of its associated marginal density. While marginal normality does not insure joint normality (except if all the variates are independent), referenc [3] has shown that "enhanced" joint normality will result from marginal transformations. The usefulness of these transformations are more apparent if the reader observes that the MGC series representation of an n dimensional Gaussian density reduces to one non-zero term. For densities of practical concern, enhancing the normality corresponds to reducing the number of terms necessary to accurately represent the p.d.f.

3. OPTIMIZATION THEORY

This section considers the deterministic optimal dispatch problem and an associated linearization. A comprehensive survey of optimal dispatch techniques has been given by Happ [8].

The objective of optimal dispatch is to minimize some cost function. Often, this function is a quadratic generation cost function,

$$F(x,u) = \sum_{i=1}^{N_G} B_{0i} + B_{1i} P_{Gi} + B_{2i} P_{Gi}^2$$

where N_G is the number of generators in the system. This particular expression is a function of x only at the slack bus and a function of u only at all other generator busses in the system. The LaGrangian is formulated as,

$$L(x,u,\lambda) = F(x,u) + \lambda^T G(x,u,p)$$

where G(x,u,p) are the standard equality constraints derived from Kirchhoff's Laws. At the optimum, the following conditions are satisfied,

$$L_x = \frac{\partial F}{\partial x} + [\frac{\partial G}{\partial x}]^T \lambda = 0$$

$$L_u = \frac{\partial F}{\partial u} + [\frac{\partial G}{\partial u}]^T \lambda = 0$$

$$L_\lambda = G(x,u,p) = 0.$$

The differential form of these equations yield,

$$-L_x = L_{xx} \Delta x + L_{xu} \Delta u + J^T \Delta\lambda$$

$$-L_u = L_{uu} \Delta u + L_{ux} \Delta x + K^T \Delta\lambda$$

$$-L_\lambda = J \Delta x + K \Delta u$$

$$J = \frac{\partial G}{\partial x}$$

$$K = \frac{\partial G}{\partial u}.$$

Simultaneous solution of the above differential equations results in [6],

$$\Delta x = -J^{-1} G - S \Delta u$$

$$\Delta\boldsymbol{\lambda} = -[J^T]^{-1}(L_x + L_{xx} \Delta x + L_{xu} \Delta u)$$

$$S = J^{-1} K$$

and

$$\Delta u = -H_{RED} G_E$$

where

$$H_{RED} = L_{uu} - L_{ux} S - S^T L_{xu} + S^T L_{xx} S$$

$$G_E = L_u - S^T L_x - [L_{ux} - S^T L_{xx}] J^{-1} G.$$

The above equations are solved iteratively until subsequent improvements are within the desired tolerances.

A linearized approximation relating u^* and p is derived by noting that at an optimum solution,

$$L_U(U^*, p) = 0$$

where,

$$U = (x, u, \lambda)^T.$$

For small perturbations this is expressed in differential form as,

$$L_{UU}(U^*, p)dU^* + L_{Up}(U^*, p)dp = 0,$$

which is solved to yield

$$dU^* = -(L_{UU})^{-1} L_{Up}\, dp.$$

Expressing this in more detail,

$$\begin{bmatrix} dx \\ du \\ d\lambda \end{bmatrix} = - \begin{bmatrix} L_{xx} & L_{xu} & J^T \\ L_{ux} & L_{uu} & K^T \\ J & K & 0 \end{bmatrix}^{-1} L_{Up}\, dp,$$

it is apparent that all the submatrices are already available from the optimal dispatch algorithm above. Hence, small changes Δu^* are easily calculated from small changes Δp. For convenience the matrix R is defined as,

$$R = -[L_{UU}]^{-1} L_{Up}.$$

4. DESCRIPTION OF THE ALGORITHM

The solution of the stochastic optimal dispatch problem incorporates the results of APPENDICES A and B to formulate this algorithm:

1. Obtain an initial optimal operating point by a second order optimization method retaining the various second order partial derivatives of the LaGrangian for future analysis.
2. Identify and tabulate parameter variations, Δp.
3. Obtain and model constraints on controls and system states.
4. Calculate the linear relation between Δu and Δp,

$$\Delta U = R \Delta p$$

5. Transform Δu to approximate.normality (APPENDIX B).
6. Represent the p.d.f. of Δu via the MGC series (APPENDIX A).
7. Calculate desired probabilities of events utilizing the integral form of the MGC series.

The software for this algorithm has been implemented and an illustrative example is pursued.

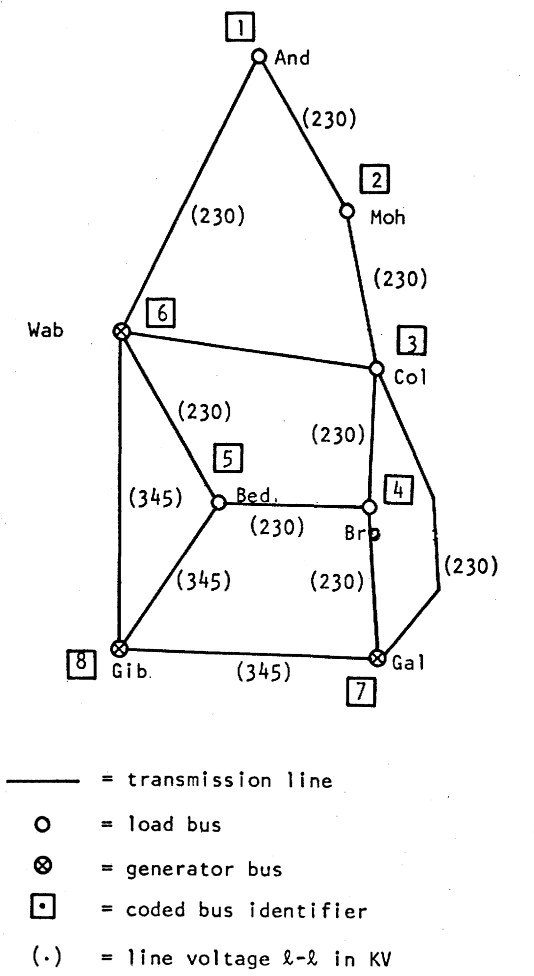


Figure 2 Example Midwestern power system

5. PRACTICAL EXAMPLE

This section presents an example application of the proposed techniques to a practical power system. The system chosen is similar to a reduced equivalent of the Public Service Indiana (PSI) system. This system contains five load busses whose demands were obtained from actual PSI data which have been magnitude scaled upward. An eight bus system was chosen to illustrate the technique for purposes of this paper. Much larger systems are treated in the same way. (note that the transformations of the variates of Δu^* permit handling of large systems without extensive computational problems associated with the MGC series.)

The proposed network for study is shown in Figure 2, with corresponding line impedances found in Table 1. The line impedances were calculated for 230 KV and 345 KV lines as the following per-unit values (using a 100 MVA base):

$$\overline{Z}_{(230KV)} = 9.86 \, x \, 10^{-5} + j6.45 \, x \, 10^{-4}$$

$$\overline{Z}_{(345KV)} = 2.21 \, x \, 10^{-5} + j3.16 \, x \, 10^{-4}.$$

The system shown is a scaled equivalent of the PSI system.

Table 1

Example power system line impedances

Line From	Line To	Resistance (pu)	Reactance (pu)
1	6	0.02500	0.16340
1	2	0.01625	0.10621
2	3	0.00875	0.05719
3	6	0.01625	0.10621
3	4	0.00438	0.02860
4	5	0.00500	0.03268
5	6	0.01125	0.07353
6	8	0.00364	0.05213
8	7	0.00448	0.06416
8	5	0.00308	0.04411
4	7	0.00813	0.05311
3	7	0.01250	0.08170

With the system configuration known, attention is directed to determining appropriate loading. Available information includes the hourly real and reactive power, P_L and Q_L, at busses 1 through 5. The intent is to accurately determine representative values for the stochastic variations , Δp, about a proposed operating point of the system. Representative values are found by modeling variations with an auto regressive (AR) [9] equation of the form,

$$y_i(t) = A_1 y_i(t-1) + A_7 y_i(t-7) + \omega_o(t)$$

where

1. $y_i(t)$ = process changing as a function of time, with time measured in discrete daily increments
2. A_1, A_7 = constant depending on y(t)
3. ω_o = Gaussian standard measure random noise.

The form of the model indicates that the future value of the loading process is dependent on the most recent daily value, as well as the value which was obtained a week earlier. This is very reasonable for a power system demand since it is likely that a demand is a function of the weather during a particular week as well as the particular day of the week. The coefficients were calculated as,

Bus	$A_1(P_L)$	$A_7(P_L)$
1	2.003024E-01	6.640934E-01
2	9.472535E-01	1.825297E-02
3	7.813616E-01	1.212142E-01
4	4.146909E-01	5.190445E-01
5	7.395055E-01	1.986481E-01

Bus	$A_1(Q_L)$	$A_7(Q_L)$
1	2.880077E-01	4.372310E-01
2	8.974314E-01	4.620967E-02
3	7.001010E-01	1.245421E-01
4	3.930626E-01	4.172327E-01
5	7.519240E-01	1.410025E-01

Here (P_L) and (Q_L) denote coefficients pertaining to the real and reactive components of the demands of the p vector. The expected value of the AR equation above is successively calculated in order to evaluate one step ahead forecasts.

The resulting forecast values are shown by the dotted line in Figure 3, for time greater than 100. Labelling the forecast values of y_i, as y^*_i, and the actual value of the process as y_a, the variations, Δp_i are found as

$$(y^*_i - y_a) = \Delta p_i.$$

The last component in the system under study is the generators which are responsive to the load demands for the system. Associated with each generator is a quadratic cost formula

$$C_i = B_{0i} + B_{1i} P_{Gi} + B_{2i} (P_{Gi})^2.$$

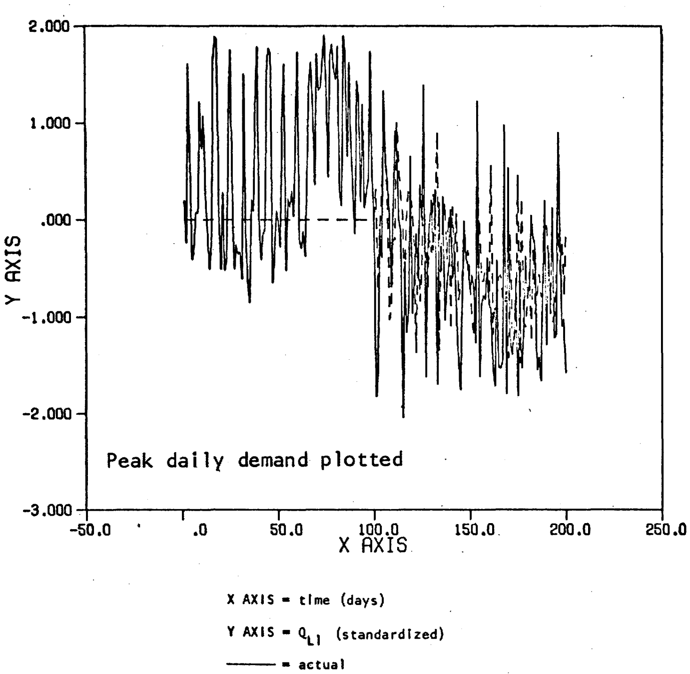


Figure 3 Prediction of Q_{L1} versus time

The coefficients are:

Bus	B_{0i}	B_{1i}	B_{2i}
6	1000.0	200.0	25.0
7	250.0	175.0	10.0
8	1000.0	250.0	40.0

A second order optimization method is employed to determine the initial optimal settings of the control vector

$$u = \begin{bmatrix} P_{G6} \\ P_{G7} \\ |V_6| \\ |V_7| \end{bmatrix}$$

(note: the slack bus has been eliminated from the control vector for convenience) based on the particular loading conditions presented. The initial optimal dispatch is shown in Table 2.

To test the accuracy of the SOED algorithm, the Δu variations, which are readily calculated from the linearization involving the R matrix, are determined. Utilizing the polynomial transformation, Δu_i is transformed to the more normal $\Delta u'_i$. For this example, only the first two variates of the Δu vector are transformed. For a standard measure (zero mean and unit variance) normal variate, all the cumulants [4] except the second are zero. The second cumulant is unity. Therefore, a measure of normality is how close a gen-

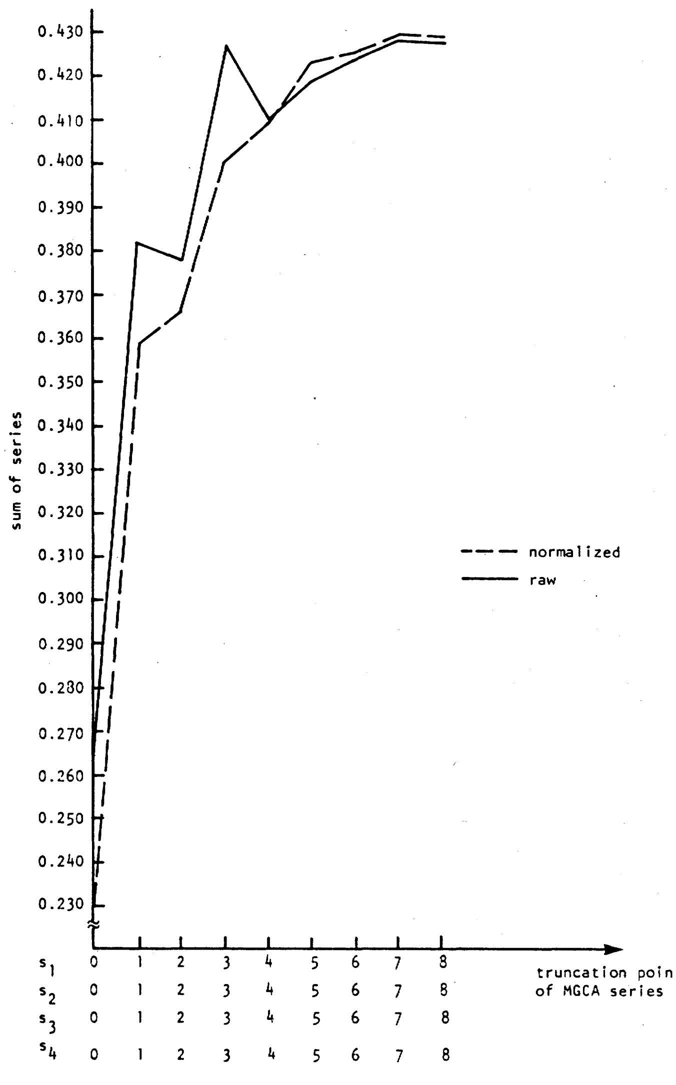


Figure 4 Truncated MGC series versus truncation point for case 1

Table 2

Initial optimal dispatch

BUS	|V| (pu)	ARG (V) (deg)
1	0.9831	-16.650
2	0.9889	-14.998
3	1.0391	-9.610
4	1.0442	-7.656
5	1.0167	-8.676
6	1.1270	-2.348
7	1.1837	5.497
8	1.0000	0.000
	P (pu)	**Q (pu)**
1	-2.000	-0.4000
2	-1.500	-0.5000
3	-5.000	-0.5000
4	-3.500	-0.5000
5	-6.000	-1.0000
6	4.586	6.0141
7	11.956	8.1451
8	2.214	-5.4656

Injection notation used. 100 MVA base.

eral random variables' statistics approximate those of a normal random variable. The first and second cumulants of both Δu and $\Delta u'$ are zero and unity respectively, as a result of the standardization. Relative improvements in normality are evidenced by reduction of the magnitude of higher order cumulants as shown:

cumulant	Δu_1	Δu_2	Δu_3	Δu_4
1	0.0000 E+00	0.0000 E+00	0.0000 E+00	0.0000 E+00
2	1.0000 E+00	1.0000 E+00	1.0000 E+00	1.0000 E+00
3	4.7703 E -01	4.7974 E -01	1.1020 E -02	2.2476 E -01
4	7.9165 E -01	8.5183 E -01	-1.2466 E -01	7.2908 E -02
5	1.0893 E+00	1.1662 E+00	5.6033 E -01	5.5781 E -01
6	-1.5203 E+00	-1.4143 E+00	4.1288 E -01	4.9251 E -01
7	-1.9120 E+01	-1.8965 E+01	-5.1265 E+00	-7.5227 E+0[illegible]
8	-8.2611 E+01	-8.7134 E+01	-1.2604 E+01	-2.8550 E+01
9	-1.2691 E+02	-1.6268 E+02	4.0710 E+01	1.6456 E+01
10	1.0281 E+03	9.2500 E+02	1.9035 E+02	3.7734 E+0[illegible]

cumulant	$\Delta u_1'$	$\Delta u_2'$
1	0.0000 E+00	0.0000 E+00
2	1.0000 E+00	1.0000 E+00
3	0.0000 E+00	0.0000 E+00
4	0.0000 E+00	0.0000 E+00
5	0.0000 E+00	0.0000 E+00
6	-5.5651 E -01	-6.5554 E -01
7	7.4838 E -02	3.5384 E -01
8	-2.2850 E+00	-7.6124 E -01
9	-6.3074 E+00	-1.1495 E+01
10	9.1345 E+01	7.3532 E+01

The statistics for the $\Delta u'$ variations about an operating point are used to calculate various probabilities of occurance. These probabilities are obtained by direct application of the integral form of the MGC series. Two sample cases are shown below.

Case 1.

This example considers the system such that all generators are operating at their derated maximum value. It is desired to evaluate the probability that under usual operating conditions (voltages are within limits, no faults, no outages) the generators would be forced to operate in the region above their derated maximum yet below their absolute maximum. In general terms, the above situation may be referred to as over-stressing the generators. Specifically, the statistics of p at the known operating point were found and the integrated MGCA was employed to evaluate the joint probability,

$$Prob \begin{bmatrix} 4.587 & \leq & P_{G6} & \leq & 7.587 \\ 11.956 & \leq & P_{G7} & \leq & 13.956 \\ 1.0270 & \leq & |V_6| & \leq & 1.2270 \\ 1.0837 & \leq & |V_7| & \leq & 1.2870 \end{bmatrix}.$$

This calculation illustrates the capability of the proposed SOED algorithm to calculate probabilities of specific operating scenarios under optimal dispatch.

The resulting probability of occurrence versus the number of terms in the truncated series is shown in Figure 4. The transformed MGC series representation has smaller oscillations as it approaches its limiting value.

To assess the accuracy of the MGCA, one hundred Monte Carlo simulations of the operating conditions were run. After considerably more computational effort, 44/100 simulations resulted in the control vector u assuming a value between the specified limits. This agrees well with the predicted value of the MGCA series considering the reliability of Monte Carlo simulations.

Case 2.

In this example the probability that the generators are operating within a range ±2 p.u. about their derated maximum and that the voltages at this bus is ±0.1 p.u. is assessed. This situation translates to a scenario where the system is termed secure. The results of the Monte Carlo simulation indicate that this situation occurred 98/100 times. The MGCA series results for the transformed Δu variations are shown in Figure 5. Mathematically, the probabil-

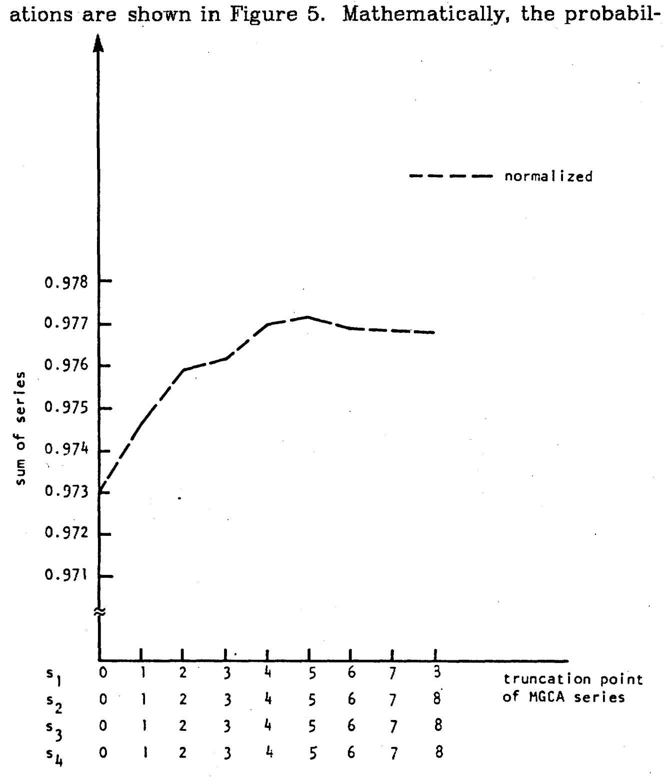


Figure 5 Truncated MGC series versus truncation point for case 2

ity for Case 2 is expressed as

$$Prob\begin{bmatrix} 2.587 \le P_{G6} \le 7.587 \\ 9.956 \le P_{G7} \le 13.956 \\ 1.027 \le |V_G| \le 1.2270 \\ 1.0837 \le |V_7| \le 1.2870 \end{bmatrix}$$

Again, the Monte Carlo simulations confirm the accuracy of the results obtained by the MGCA series.

Note that this technique is primarily a system planning tool rather than an operating procedure since computation time, while not prohibitive, is not fast enough to be on-line.

6. CONCLUSIONS

The uncertainties present in electric energy system loads and configuration effect optimal power dispatch. These uncertainties are effectively accounted for through the use of a sensitivity technique which requires minimal calculation above that of a second order optimal dispatch formulation. This paper has presented the computational details of the stochastic optimal energy dispatch problem. The SOED algorithm employs the multivariate Gram-Charlier series to statistically model the p.d.f. of the control vector. The applicability of this series has been limited (in the past) by the high computational requirements of calculating high order statistical moments. The method described obviates some of the difficulty through a polynomial transformation of the variates to be modelled in order to enhance normality. The transformation process permits truncation of the series early, resulting in high accuracy with relatively little computation. The techniques proposed have been tested with an eight bus equivalent system.

ACKNOWLEDGEMENT

This research was made possible by grant EC-77 S-02-4345 from the Department of Energy. This support is gratefully acknowledged.

REFERENCES

[1] Papoulis, W. D., **Probability, Random Variables and Stochastic Processes,** McGraw Hill Book Co., New York, 1965.

[2] Sauer,P. and Heydt, G.,"A Convenient Multivariate Gram Charlier Type A Series", IEEE Trans. on Communications, Vol. CM-27, No. 1, January 1979,pp. 247-248.

[3] Andrews,D. ,Gnanadesikan, R., and Warner, J., "Transformations of Multivariate Data",BIOMETRICS, Vol. 27, 1971, pp. 825-840.

[4] Kendall, M. and Stuart, A., **The Advanced Theory of Statistics,** Volume I, 4th edition, MacMilliam Publishing, NY, 1977.

[5] Viviani, G.L., "Stochastic Optimal Energy Dispatch", Ph.D. Thesis, Purdue University, May 1980.

[6] Bala, J.L., and Thanikachalam, A., "An Improved Second Order Method for Optimal Load Flow", IEEE, PAS-97, July/August 1978, pp. 1239-1244.

[7] Williams, J.O., "An Approximation to the Probability Integral", Annals of Mathematical Statistics, Vol. 17, 1946, pp. 363-365.

[8] Happ, H.H., "Optimal Power Dispatch - A Comprehensive Survey", IEEE Trans. on Power Appartatucs and Systems, PAS-96, no. 3 May/June, 1977, pp. 841-854.

[9] Kashyap, R.L., Rao, A.R., **Dynamic Stochastic Models from Empirical Data,** Academic Press, New York, 1976.

APPENDIX A. THE MULTIVARIATE GRAM-CHARLIER SERIES

The MGC series is an infinite series representation of a p.d.f. in terms of the statistical moments of the variate being statistically modelled. The folowing is a presentation of the MGC series and its application to the solution of the SOED problem. For simplicity the MGC series is presented without derivation (for further detail see [5]). Defining $x = (x_1, x_2, \ldots, x_n)^T$ and $dx = (dx_1\ dx_2 \ldots dx_n)$, the MGC series is

$$f(x) = \sum_{s_1 \ldots s_n = 0}^{\infty} E\left[\prod_{i=1}^{n} H_{s_i}(x_i)\right]\left(\prod_{p=1}^{n} \frac{H_{s_p}(x_p)G(x_p)}{s_p!}\right)$$

where

$$G(x_p) = \frac{1}{\sqrt{2\pi}} EXP\left[-\frac{x^2}{2}\right]$$

$$E\left[\prod_{i=1}^{n} H_{s_i}(x_i)\right] = \int_{-\infty}^{\infty}\left[\prod_{i=1}^{n} H_{s_i}(x_i)\right] f(x)\, dx$$

and H_{s_i} is termed a Hermite polynomial. The Hermite polynomials are defined implicitly by

$$EXP\left[tx - \frac{t^2}{2}\right] = \sum_{i=0}^{\infty} t^i \frac{H_i(x)}{i!}.$$

The salient property of these polynomials is that they are othogonal. The first five Hermites are,

$H_0(x) = 1$
$H_1(x) = x$
$H_2(x) = x^2 - 1$
$H_3(x) = x^3 - 3x$
$H_4(x) = x^4 - 6x^2 + 3$
$H_5(x) = x^5 - 10x^3 + 15x.$

The Hermite polynomials are polynomials in x which implies that the expected value of the Hermites are functions of the moments of x. Substituting the sample moments of x into the various expectations of products of Hermites in the MGC series provides the means to represent a multi-dimensional density when only sample observations are available. The expectations of the Hermite polynomials are evaluated more directly by evaluating sample Hermite expectations. This is accomplished by substituting the observation data directly into the expressions for the Hermite polynomials.

Often, the actual multivariate p.d.f. is not as important as the integral of the p.d.f., which represents the probability of certain events. The integral expression of the MGC series is easily evaluated due to the properties of the Hermite polynomials. Specifically,

$$\int H_l(x)G(x)dx = -H_{l-1}(x)G(x).$$

The error function (ERF) is defined as,

$$ERF(a) = \int_{a}^{\infty} G(x)dx,$$

which for computer implementation is approximated by [7],

$$ERF(a) = \frac{1}{2} \pm \frac{1}{2}\left[1 - EXP\left(\frac{-2a^2}{\pi}\right)\right]^{\frac{1}{2}} \begin{cases} + & a<0 \\ - & a>0 \end{cases}$$

Recognizing the uniform convergence of the MGC series [5], each term of the series is integrated separately and the integral expression becomes,

$$I = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \sum_{s_1 \ldots s_n = 0}^{\infty} E\left[\prod_{i=1}^{n} H_{s_i}(x_i)\right] * \left(\prod_{p=1}^{n} \frac{H_{s_p}(x_p)G(x_p)}{s_p!}\right) dx$$

$$= \sum_{s_1 \ldots s_n = 0}^{\infty} E\left[\prod_{l=1}^{n} \frac{H_{s_l}(x_l)}{s_l!}\right] * \prod_{p=1}^{n} H_{s_p-1}(a_p)G(a_p) - H_{s_p-1}(b_p)G(b_p).$$

This formula provides an easy means for assessing desired probabilities.

APPENDIX B. TRANSFORMATIONS TO NORMALITY

A polynomial transformation to normality is considered in this appendix. The transformation described enhances joint normality by transforming each of the variates of a random vector to approximate normality. All variates considered are in standard measure (zero mean and unit variance). For a standard measure Gaussian density, G(z),

$$m_i^o = E[z^i] = 1*3*5*....*n-1, \quad i = 2,4,6,...n$$

$$m_i^o = E[z^i] = 0 \quad i = 1,3,5,....$$

Consider the moments, m_i, of a general sample marginal density. These moments are compared to those of G(z) as a measure of normality as,

$$Q = \sum_{i=1}^{l} (m_i - m_i^o),$$

where l is the order of the polynomial transformation operating on the untransformed data, y, such that

$$z_i = a_0 + a_1 y_i + ... + a_l y_i^l.$$

To approximate normality, the a_i's$[i=1,2,....l]$ must be determined to minimize Q. This is accomplished by a gradient technique:

1. Estimate values for the a_i coefficients in the polynomial transformation.
2. Calculate the sample moments of z_i for i ≤ l. Label these moments m_i^*,$[i=1,2,....l]$.
3. Perturb each of the a_i coefficients by a small amount ε, such that $a_i = a_i + \varepsilon$.
4. Recalculate the sample moments m_i^*,$[i=1,2,...l]$, utilizing the a_i's.
5. Label the vector $a = (a_1, a_2, ..., a_l)^T$ and the vector $m = (m_1, m_2, ..., m_l)^T$. Then by Euler's approximation

$$\Delta m = \frac{\partial m}{\partial a} \Delta a$$

where $\Delta m_i = m_i^* - m_i$ and $\Delta a_i = \varepsilon_i$,

$$\frac{\partial m}{\partial a} = \begin{bmatrix} \frac{m_1^* - m_1}{\varepsilon_1} \cdots\cdots\cdots \frac{m_1^* - m_1}{\varepsilon_l} \\ \frac{m_2^* - m_2}{\varepsilon_1} \\ \frac{m_l^* - m_l}{\varepsilon_1} \cdots\cdots\cdots \frac{m_l^* - m_l}{\varepsilon_l} \end{bmatrix}$$

6. Successively calculate values of vector a by

$$a^{new} = [\frac{\partial m}{\partial a}]^{-1}(m^o - m) + a^{old}$$

where m^o is the vector fo moments for a standard measure random variable.

7. Return to step 3 if a^{new} is not equal to a^{old}, otherwise the solution vector has been found.

This procedure is repeated for each of the y_i variates resulting in approximate joint normality of z. This is useful in improving the truncation error of APPENDIX A. (Note: limits of integration must also be transformed.)

Gary L. Viviani (S'77-M'80) was born in Mc Allen, TX, on 1 June 1955. He received the B.S.E.E., M.S.E.E., and Ph. D. (electrical engineering), all from Purdue University, W. Lafayette, IN, in 1977, 1978 and 1980, respectively.

In 1980 he joined the E.I. DU PONT DE NEMOURS Company, Wilmington, DE. Currently, his interests include energy system analysis and control, as well as load management. Previously, he has served as a consultant to various private and public agencies.

Dr. Viviani is a member of IEEE and Eta Kappa Nu.

Gerald Thomas Heydt was born in New York State on October 1, 1943. He received the B.E.E.E from Cooper Union, New York, in 1965. Subsequently he worked with the United States Atomic Energy Commission and E.G. and G., Inc. in Mercury and Las Vegas, Nevada. He received the M.S.E.E. and Ph. D. degrees for Purdue University, West LaFayette, Indiana in 1967 and 1970 respectively. Since 1970, he has had industrial experience with the Commonwealth Edison Company, Chicago and in various consulting positions in the United States and Latin America. Dr. Heydt is presently Associate Professor of Electrical Engineering at Purdue University where he also serves as Chairman of the Purdue Electric Power Center. Dr. Heydt's interests are mainly in reliability of power systems and large system analysis.

Dr. Heydt is an avid amateur radio operator operating under advanced license WB9SJW. He is also a private pilot. Among his professional society memberships are IEEE, CIGRE, ACM, Sigma Xi and Eta Kappa Nu.

Discussion

M. E. El-Hawary (Memorial University of Newfoundland, St. John's, Newfoundland, Canada): The authors have presented a very interesting and pioneering paper in the area of application of probabilistic methods to optimal power flow. This discusser wishes to commend the authors for a clear and well written paper.

The proposed procedure is essentially an optimal power flow followed by evaluation of the statistical properties of the control variables given statistics of the parameters. A possible alternative that may be of interest is to perform a sensitivity analysis of the optimal power flow. Combining this with the parameter statistical properties would then result in the required statistical properties for the control variables. The authors' comments on the advantages of their method over this simpler procedure would be appreciated.

A major source of uncertainty in optimal dispatch is that associated with the cost coefficients B_{0i}, B_{1i}, B_{2i}. Is the proposed procedure capable of handling this situation? Would the authors comment on the reasons for taking a deterministic cost function instead of a stochastic one involving the expected cost.

A minor point is related to the terminology employed in labelling the procedure as an energy dispatch. It seems more appropriate to denote the procedure in terms of optimal power flow. This does not in any way distract from the clear contribution this paper makes to power systems engineering.

Manuscript received February 19, 1981.

P. W. Sauer (University of Illinois, Urbana, Il): Viviani and Heydt have made several interesting contributions in this work. They have successfully incorporated uncertainties into the optimal energy dispatch algorithm and illustrated a novel approach to the computation of multivariate proabilities. The methods appear to be applicable to many other optimal load flow routines as well. The transformations to normality presented in Appendix B made the multivariate Gram-Charlier series a practical method for computing output statistics. In this approach, an *l*th order polynomial transformation was used to minimize the difference between normal and non normal variable moments. Would the authors please comment on the choice of the order *l*? Does the transformation to normality improve as *l* is increased? If so it would appear that there is a trade off between the benefits gained in the Gram-Charlier series calculation and the effort required to minimize Q. Would the authors also please comment on their selection of Q? It seems that minimizing a weighted sum of squared errors would be more logical since the Q shown may have positive or negative terms.

The cost function used in the paper reflects the total cost of operation for a given generation dispatch. With the statistics of all P_{G_i} available from the SOED algorithm, the expected range of the minimum cost could be obtained easily. Have the authors considered the significance of computing this range as an indication of the possible variation in operating costs and therefore a projection of profit variation? Clearly even the mean value of this minimum cost would be different from the base case solution due to the quadratic variation in P_G.

The methods presented in this paper may have further application in other optimization routines. The minimum time restoration problem must dispatch both load and generation in an optimal sequence. The uncertainty of load on a bus following an outage must be considered in this algorithm, and the results of this paper may prove very useful in that regard. The authors are congratulated for another fine addition to the stochastic power flow area.

Manuscript received March 3, 1981.

G. L. Viviani and **G. T. Heydt:** The authors appreciate the valuable comments of the discussers. We shall consider their questions independently.

Dr. Sauer has raised an interesting point concerning the order of the polynomial transformation to normality. Consider the case that all the moments determine the p.d.f. For a particular univariate normal density all the moments are unique. If the quadratic cost function, Q, is chosen such that the m_0's are those of a standard measure normal variable, increasing the order of *l* toward infinity will assure normality in the limit. For finite *l* it is difficult to determine a functional relationship between *l* and the degree of normality. The authors chose *l* based on its relative merits: it provided significant improvement in normality without undue computational effort. A more conclusive relationship would be desirable. Practically, the marginal transformations to normality are a useful technique for enhancing the accuracy of a series representation of a p.d.f. Dr. Sauer is correct in noting the typographical error in the equation for Q. The equation for Q is,

$$Q = \sum_{i=1}^{\ell} (m_i - m_i^o)^2 .$$

Dr. El-Hawary indicates the similarity between sensitivity analysis and the approach taken in this paper. The approach suggested by Dr. El-Hawary is essentially the same as the one chosen by the authors with the exception that the authors' analysis includes a transformation to normality to enhance the accuracy of the series representation of the p.d.f. We are in agreement with Dr. El-Hawary concerning the uncertainty associated with the cost coefficients. Unfortunately, this research does not address that important problem.

Drs. El-Hawary and Sauer have both raised the question of a stochastic cost function. The aim of this work was to produce a tool which would be useful from an operational standpoint, i.e., to devise a methodology for determining the proper settings on control variables with respect to uncertainty in the parameter vector, p. Characterization of the cost function as a random variable would be a useful and important extension of this existing analysis.

Manuscript received March 3, 1981.