Abstract - This paper presents a new approach for the optimal long-range generation planning. A completely new analytical approach for the production costing model and reliability measure is developed under the assumption of Gaussian probabilistic distribution of random load fluctuations and plant outages. The long-range generation investment problem is formulated as an optimal control problem to determine optimally the annual investment in new generation capacity. The Hamiltonian minimization is performed by using a version of the gradient projection method.

INTRODUCTION

The problem of long-range generation expansion planning has received a considerable attention, and the motivation for more efficient and more sophisticated techniques of evaluating utility expansion policies has been increased during the past decade [1,4,7,11,12]. In these studies, there has always been an issue of trade-off between computational effort and accuracy of the desired solution.

The basic idea of new approach is similar to the probabilistic production costing of Booth and Balaleraux [12,19], where the load is modeled as a random variable and generating units are modeled as power sources with available capacity represented also as a random variable. However, this paper presents a new approach for the long-range generation expansion planning, which provides a comparable accuracy with much less computational effort than the conventional one. A completely new analytical approach for the production costing model was developed under the assumption of Gaussian probabilistic distribution of random load fluctuations and plant outages. Analytic formulas were derived not only for operating costs and reliability measures (unserved energy, loss-of-load probability, etc.), but also for their marginal values.

The problem considered is to determine the most economical and reliable generation expansion plan in order to meet forecasted loads over a long-range horizon (usually five to thirty years), subject to a multitude of technical, economical and social constraints. The method presented here uses the discrete optimal control theory to determine optimally the annual investment in new generating capacity, while the Hamiltonian minimization problem within the dynamic optimization process is solved by using a version of the gradient projection method which has been developed specially for this purpose.

In the new algorithm, system reliability measures such as loss-of-load probability or expected unserved energy can be considered as operating constraints rather than supply-shortage cost. Also, a variety of constraints (financial, social, etc.) can be introduced similarly as operating constraints.

The long-range generation planning problem is decomposed naturally into one master problem and a set of subproblems. The master problem is to determine optimally the yearly investment in new generation capacity. The subproblems are used to determine the minimum operating cost of generation systems and annual reliability measures. Because of the Gaussian assumption on load fluctuations and plant outages, the subproblems are completely solved analytically. Associated with the solution of the subproblems is a set of marginal values which are changes in system operating costs and reliability caused by marginal changes in installed capacities. These marginal values not only provide a planner technical and managerial information, but also are the essential values returned to the master problem, which is then modified and re-solved to determine a new investment plan. The long-range generation planning problem is solved iteratively, by alternately solving the master problem and the subproblems, until the optimal solution is found.

The organization of this paper is as follows. The analytic formulas for expected annual cost and reliability measures are first developed and followed by the master problem of optimal yearly investment. The accuracy of analytic formulas developed in the subproblems are then verified numerically. Finally, a case study is presented to demonstrate the usefulness of the proposed generation investment planning algorithm.

BASIC SUBPROBLEMS FOR OPTIMAL OPERATION

This section develops analytic formulas required to compute expected annual cost and reliability measures. One of the primary features of the presented approach is the direct use of load curves rather than the equivalent inverted load duration curves. It is well known that the computational burden of most planning packages is primarily due to the numerical convolution processes involving load duration curves. The numerical error accumulates as the convolution is repeated each year. By avoiding the use of load duration curves, not only is the computational requirement reduced in the order of magnitudes, but also the accuracy achieved as will be demonstrated in later sections.

In order to reduce the computational requirement, an effort is made to develop analytic formulas rather than numerical (convolution) approaches. This is achieved by assuming Gaussian probability distribution on the random load fluctuations and plant outages.

1. Representation of Random Load Fluctuations

The load at a particular time of the day of a week in a given season fluctuates randomly and it is reasonable to assume that it behaves with Gaussian distribution since its value is obtained from large


0018-9510/85/0002-0300$01.00 ©1985 IEEE
number of previous historical data. When a weekly load curve is used, there are several options to represent it as follows:

1) use 7 daily load curves, each with T time-bands, or
2) use one equivalent weekday load curve with duration of 5 days and T time-bands, and one equivalent weekend load curve with duration of 2 days and T time-bands, or
3) use one equivalent load curve with duration of 7 days and T time-bands.

It is important to remember that in planning problem, although the overall pattern of the load on an electric utility is predictable, it is difficult to make hourly prediction due to its randomness. For this reason, in the conventional method, the time-independent load curve is converted into a load duration curve. Even in generation planning with limited energy and storage plants the model of storage is an approximation using load duration curves, which ignores the inherent chronological nature of storage operations since a chronological model would require too much detail and computation to be suitable for use in a long range generation planning model.

In the present, season, the load curve can be either time-dependent or time-independent. For example, the option 1) above mentioned is time-dependent and the options 2) and 3) are time-independent. The options 2) and 3) would be more suitable for long range generation planning problem. The study shows that the difference among the results obtained by using different options of load curve is insignificant.

The statistical means and variances for past years can be easily computed following standard statistical analysis. However, they cannot be computed for future years and must be extrapolated on the basis of past values as well as other economic and social considerations. This is normally done in load forecasting, and the long-range planning problem is to determine the most economical and reliable generation expansion plan to meet these forecasted loads.

The random load \( L_{i,t} \) [MW] in each time-band \( t \) can be represented by a normal distribution with its statistical mean \( \overline{L}_{i,t} \) [MW] and variance \( \sigma_{i,t}^{L^2} \) [MW²]. Given the statistical values of future loads on a basis of one equivalent load curve and load forecasting, the expected annual energy demand for year \( i \), \( D_i \), can be accounted as

\[
D_i = \sum_{s=1}^{S} \sum_{t=1}^{T} n_t \overline{L}_{i,s,t} \]  

where \( \Delta = i,s,t \): index of year, season, and time-band, respectively,

\( \overline{L}_{i,s,t} \): the statistical mean of the random load \( L_{i,s,t} \) for years, season, and time-band \( t \) [MW],

\( n_t \): number of time-bands in a load cycle,

\( S \): number of load cycles in season \( s \),

\( T \): number of seasons in a year.

Here, the number of time-bands, \( T \), represents the number of intervals that a load cycle can be broken into, and \( t \) is the length of each time-band, not necessarily of the same length, resulting from the division. Similarly, the number of seasons, \( S \), can be chosen to reflect how many distinct seasons one wishes to use to represent a year.

The generating plants are grouped into several types, such as nuclear, base-oil, intermediate-oil, base coal, intermediate coal, peaking gas turbine, etc. All units of the same plant type are assumed to have the same economic and technical characteristics. This makes it possible to represent each type with a number of identical single units of same size. The unit capacity, of course, will differ from type to type.

We will first consider the installed capacity for each type, and then represent the available generation capacity by incorporating random plant outages, maintenance requirements, energy resource availabilities, and the teething factors. The teething factor is defined to be a discount factor to reflect the nature of newly installed unit which cannot operate at its full capacity during the first year of operation.

The installed capacity of plant type \( j \) in year \( i \), \( g_j \), is the sum of the previously installed capacity and newly added capacity, i.e.,

\[
g_j = x_j + u_j \]  

where \( x_{i,j} \): previously installed capacity of plant type \( j \) in year \( i \) [MW] (state variables),

\( u_{i,j} \): newly added capacity of plant type \( j \) in year \( i \) [MW] (decision variables).

The probabilistic nature of plant outages of units in each plant type \( j \) makes the available generation capacity \( y_{i,j} \) [MW] as a random variable. Each unit is considered to have two stages of operation: full-capacity operation with availability \( p \), and full outage with failure rate \( 1-p \). Each plant type has many such single units, which are assumed to be mutually independent. Therefore, according to the law of large numbers, the total available generation capacity \( y_{i,j} \) for type \( j \) can be represented by a normal distribution

with its statistical mean \( \overline{y}_{j} \) and variance \( \sigma_{j}^{2} \), which are derived as

\[
\overline{y}_{j} = p\overline{y}_{i} + (1-p)\overline{x}_{i,j} \]  

\[
\sigma_{j}^{2} = \sigma_{i}^{2} \]  

where \( \Delta = i,s,t \): index of year, season, and time-band, respectively,

\( \overline{y}_{i} \): previously installed capacity of plant type \( j \) in year \( i \) [MW],

\( u_{i,j} \): newly added capacity of plant type \( j \) in year \( i \) [MW],

\( \overline{x}_{i,j} \): teething factor for units in type \( j \) [p.u.],

\( \sigma_{i}^{2} \): energy resource distribution factor of plant type \( j \) in time-band \( t \) of season \( s \) of year \( i \) [MW],

\( \overline{p}_{i} \): maintenance rate of plant type \( j \) in season \( s \) of year \( i \) [MW],

\( a_{i} \): availability of units in plant type \( j \) [p.u.],

\( \overline{a}_{i} \): unit capacity of plant type \( j \) [MW].

The energy resource distribution factors \( \overline{p}_{i} \)'s are fixed to one for non-hydro plants. But they can be determined optimally for hydro plants to represent the optimal use of hydro resources. The maintenance rates \( \overline{p}_{i} \)'s are also fixed in our problem, but can be determined optimally from maintenance scheduling subproblem.

2. Available Generation Capacities

3. Expected Plant Outputs and Annual Energy Generation
Comparing forecasted loads and available generation capacities, expected plant outputs can be determined optimally in the production scheduling. If the generating plant types are confined to the non-hydro's, the optimal solution to this operating problem can be based upon the conventional loading-order concept. The loading-order is defined as the economic merit order in which plant type \( j \) is loaded in the order of increasing operating costs. Let \( j = 1, 2, \ldots, J \) be the indices of plant types already ordered in the order of increasing operating costs. The total power output from plant type \( j \) up to type \( j \), \( J^p_j \), does not exceed the loads, and can be expressed as

\[
J^p_j = \min \{ \bar{L}, \bar{Y} \}^k_{k=1} \]  

(5)

where \( \bar{L} \) is the random load and \( \bar{Y}^k \) is the random available generation capacity of plant type \( k \). Since the load and the available generation capacities have the Gaussian probability distributions, the total power output \( J^p_j \) is also a random variable with the Gaussian distribution, whose statistical mean (expected value) and variance are computed as [see Appendix A]:

\[
\bar{J}^p_j = \bar{L} - J^2_{\Delta} \cdot (0.5 \cdot \text{erf}(J^2_{\Delta}/\sigma_{\Delta})) \cdot (J^2_{\Delta}/\sqrt{\pi}) \exp(-0.5J^2_{\Delta}/\sigma_{\Delta}) \]  

[\text{MW}]  

(6)

\[
\sigma^2_{J^p_j} = \sigma^2_{\Delta} + \frac{1}{2} \sum_{k=1}^{J} \sigma^2_{k} \]  

[\text{MW}^2]  

(7)

where

\[
J^2_{\Delta} = \frac{L}{\Delta} - \frac{1}{2} \sum_{k=1}^{J} Y^k \]  

[\text{MW}]  

(8)

\[
\text{erf}(\cdot): \text{error function} \]  

\[
\exp(\cdot): \text{exponential function}. \]  

Here \( J^2_{\Delta} \) represents the difference between the expected load and the sum of the available generation capacities from type 1 to type \( j \). Thus the expected power output of each plant type, \( J^p_j \), can be simply computed as

\[
J^p_j = \left[ \begin{array}{c} \bar{L} - J^2_{\Delta} \cdot (0.5 \cdot \text{erf}(J^2_{\Delta}/\sigma_{\Delta})) \cdot (J^2_{\Delta}/\sqrt{\pi}) \exp(-0.5J^2_{\Delta}/\sigma_{\Delta}) \\ \sigma^2_{\Delta} + \frac{1}{2} \sum_{k=1}^{J} \sigma^2_{k} \end{array} \right] \]  

[\text{MW}, \text{MW}^2]  

(9)

where \( i, s, t \): index of year, season, and time-band, respectively.

\( \bar{L} \): the expected load [MW],

\( J^p_j \): the expected total power output from all plant types, i.e., from type 1 to type \( J \) [MW].

\( \bar{Y}^k \): the expected total power output from all plant types, i.e., from type 1 to type \( J \) [MW].

Associated with this unserved power, is the loss-of-load probability which is derived as [see Appendix A]

\[
LOLP_{\Delta} = 0.5 \cdot \text{erf}(J^2_{\Delta}/\sigma_{\Delta}) \]  

[p.u.]  

(13)

where \( J^2_{\Delta} \) and \( \sigma_{\Delta} \) are defined in (8) and (7).

Consequently, the expected annual unserved energy and the annual loss-of-load probability are integrated over a year as

\[
R_i = \sum_{s=1}^{S} \sum_{t=1}^{T} n_s t \bar{Y}_{\Delta} \]  

[\text{MWH}]  

(14)

\[
LOLP_i = \frac{\sum_{s=1}^{S} \sum_{t=1}^{T} n_s t \cdot \text{LOLP}_{\Delta}}{\text{w}} \]  

[p.u.]  

(15)

where \( w \) is the total hour in a year.

5. Expected Annual Cost

Having computed the generation capacities and the expected total energy generated, the economic annual cost can be computed in a straightforward manner.

The annual cost \( G_i \) is made of three cost terms: the capital cost \( c_i \), fuel cost \( f_i \), and the other non-fuel maintenance cost term \( m_i \). All costs are, of course, present-values, and are escalated and discounted during the projected period. Moreover, in the case of capital cost, a credit is given as a salvage value for the unused portion of the plant life. In summary, the expected annual cost is

\[
G_i = c_i + f_i + m_i \]  

[p.u.]

(16)

where

\[
\text{pc}_i, \text{pf}_i, \text{pm}_j: \text{factors to levelize the resultant present-value unit capacity cost, unit fuel price and unit non-fuel maintenance cost, respectively, of plant type } j \text{ in year } i \text{, [p.u.]} \]

\[
c_i, f_i, m_i: \text{unit capacity cost } [\$/\text{MWH}], \text{ unit fuel price } [\$/\text{MWH}], \text{ and unit non-fuel maintenance cost } [\$/\text{MWH}], \text{ respectively, of plant type } j \text{ in year } i \text{, [p.u.]} \]

\[
x_i: \text{factor for the salvage value of capital cost for type } j \text{ installed in year } i \text{, [p.u.]} \]

\[
r_i: \text{factor for the salvage value of capital cost for type } j \text{ installed in year } i \text{, [p.u.]} \]

\[
u_i: \text{previously installed capacity of plant type } j \text{ in year } i \text{, [MW]} \text{ (state variables)} \]

\[
u_i: \text{newly added capacity of plant type } j \text{ in year } i \text{, [MW]} \text{ (decision variables)} \]

\[
\bar{E}^i_j: \text{the expected total energy generation of plant type } j \text{ in year } i \text{, [MWH]} \text{ defined in (10).} \]

The decision variables are the newly added capacities \( u_i^j \) for each year and are determined optimally in the master problem.
MASTER PROBLEM FOR OPTIMAL PLANNING

The long-range problem for a system of generating plants is to minimize the present value of total costs subject to reliability constraints, and is now stated formally as follows:

\[
\text{minimize } G = \sum_{i=1}^{I} G_i(x_i^j, u_i^j) \\
\text{subject to } R_i(x_i^j, u_i^j) \leq c_i \\
u_i^j \geq 0
\]

where
\( i = 1, 2, \ldots, I \): index of years in planning horizon,
\( j = 1, 2, \ldots, J \): index for plant types,
\( x_i^j \): previously installed capacity of plant type \( j \) in year \( i \) [MW] (state variables),
\( u_i^j \): newly added capacity of plant type \( j \) in year \( i \) [MW] (decision variables),
\( G_i \): present-value expected annual cost [\$] defined in \( (16) \),
\( R_i \): expected annual unserved energy [MWh],
\( c_i \): desired reliability level of annual unserved energy [MWh].

The equation \( (18) \) is the dynamic state equation describing the state of installed capacities each year, that is, the sum of previously installed capacity and newly added capacity is the installed capacity for the next year. Considering the newly added capacities \( u_i^j \) as decision or control variables, the master problem \( (17)-(20) \) is now formulated as an optimal control problem.

According to Pontryagin's minimum principle of modern optimal control theory [17], the optimal control must satisfy the necessary conditions:
\[
x_i^{j+1} = x_i^j + u_i^j, \ x_i^j = \text{specified (state equation)} \\
\lambda_i^j = \frac{\partial G_i}{\partial x_i^j} + \lambda_i^j + \lambda_i^{j+1} = 0 \quad \text{(costate equation)}
\]

and \( u_i^j \) minimizes the Hamiltonian
\[
H_i(x_i^j, u_i^j, \lambda_i^{j+1}) = G_i(x_i^j, u_i^j) + \sum_{j=1}^{J} \lambda_i^j (x_i^j - x_i^{j+1})
\]

subject to
\[
R_i(x_i^j, u_i^j) \leq c_i \quad \text{(reliability constraint)}
\]

\[
u_i^j \geq 0 \quad \text{(control constraint)}
\]

where \( \lambda_i^j \) is the costate (adjoint) of plant type \( j \) in year \( i \).

Note that the minimization of Hamiltonian \( H_i \) is to be done for one year at a time while the minimization of the total cost \( G \) in \( (17) \) needs to be performed for all years in the planning horizon simultaneously. This reduces the number of decision variables \( u_i^j \) to \( J \times I \) in the minimization algorithm.

The minimization of Hamiltonian \( (23) \) subject to inequality constraints \( (24)-(25) \) is a typical nonlinear programming problem. It is solved by using a version of the Gradient Projection Method (GPM) [18], which has been developed specially for this purpose. The results from several sample studies show remarkable advantages of this method in computational efficiency and reliability.

The reliability constraints \( (24) \) was linearized using the gradients or marginal values \( \partial G_i/\partial u_i^j \). The marginal values of expected annual costs \( \partial G_i/\partial x_i^j \) and \( \partial G_i/\partial u_i^j \) with respect to marginal changes in the previously installed capacity and the newly added capacity, respectively, are also required in the GPM. Since these marginal values are not only required to find the optimal solution, but also provide a planner useful technical and managerial informations, they are computed in the basic subproblems. Again, complete analytic formulas are developed and given in Appendix B.

The computational steps of solving the optimal control problem \( (21)-(25) \) are summarized as follows:

1) Initially guess \( u_i^j, \ i = 1, 2, \ldots, J \).
2) Solve the state equation \( (21) \) forwards to determine \( x_i^j \), where \( j = 1, 2, \ldots, J \) and \( i = 1, 2, \ldots, I \).
3) Solve the costate equation \( (22) \) backwards to determine \( \lambda_i^{j+1} \), where \( j = 1, 2, \ldots, J \) and \( i = 1, 2, \ldots, I \).
4) Set \( i = I \).
5) Solve the subproblems to determine the optimal annual operation and get other necessary information for year \( i \).
6) Minimize the Hamiltonian \( H_i \) in \( (23) \) subject to the inequality constraints \( (24)-(25) \) to update \( u_i^j \) where \( j = 1, 2, \ldots, J \).
7) Solve \( x_i^j \) forwards by \( (21) \).
8) Increase \( i = i + 1 \).
9) If \( i \) is greater than \( I \), go to Step 10. Otherwise go to Step 5.
10) If the cost function in \( (17) \) is sufficiently converged, terminate the calculation. Otherwise go to Step 3.

COMPUTATIONAL EXAMPLES

The accuracy and efficiency of presented method were tested in several different systems. Among the results, two of them are shown in this section: the one for the accuracy test of the analytic formulas developed in the section of Basic Subproblems for Optimal Operation and the other for the convergence test of the computational algorithm developed in the section of Master Problem for Optimal Investments.

1. Test for Accuracy of Formulas

The load cycle of Scenario B and the full scale system of Scenario D from EPRI Synthetic Utility Systems [15] are used as load and generation system data in order to compare the accuracy of solutions by the presented formulas and by the conventional method which is based on the use of equivalent load duration curves and their recursive convolutions.

The result of using the conventional method was obtained from EPRI Report EA-1411 [16]. It is shown in Table 1 and compared with the result computed using the presented formulas.

Simulation period, peak load and minimum load are 728 [hrs], 26,000 [MW] and 13,842 [MW], respectively. According to given data, the total installed capacity up to the 5th plant type is 12,600 [MW]. It is less than the minimum load, thus the capacity factor of each plant type from 1 to 5 should be the same as its availability.

The result by the conventional method, however,
as shown in Table 1, exhibits inconsistencies as the number of plant types increases. For example, for the 5th plant type, the capacity factor is 0.8339 instead of 0.870. This is due to the recursive computation of EPRI's which cause accumulated error. It can be observed from Table 1 that up to the 5th plant type each capacity factor obtained by using the presented analytic formulas is in perfect match with its corresponding availability. For example, for the 5th plant type, both capacity factor and availability are 0.870.

The overall accuracy was also evaluated by comparing the energy balance in both methods. The sum of total generated energy and unserved energy in each case is compared with the total load demand of 15,017.24 [GWH] found directly from the load curve unit 1). The result from the conventional method is 14,293.74 [GWH] or 4.82% error in energy balance, while the result from the analytic formulas is 15,017.20 [GWH] or 0.002% error.

The LOLP calculated by the conventional method is normally inaccurate since it is obtained from the final equivalent load duration curve after all units (great in number) are convolved. Thus, in view of the above discussions, the LOLP calculated by simple analytical expression, Eq. (15), is believed to be more accurate.

In conclusion, the results of the presented analytic formulas are shown to be more accurate inspite of the Gaussian probabilistic approximation of load fluctuations and plant outages in the analytical formulation.

The CPU time requirement for this example by the presented method is 1.30 sec.

2. Test for Solution Convergence and Computing Efficiency

The generation expansion planning problem based on a full scale system of Scenario D from Synthetic Utility Systems [15] is considered in order to test the convergence characteristics of the optimal solution and to evaluate the computing efficiency of the presented method. The system has 174 existing units and is described briefly in Table 2.

All three weekly load cycles of Scenario D corresponding to spring (or fall), summer and winter, respectively, are used in order to represent more realistic energy demand in a year since the analytical nature of the presented method allows for the use of different load curves in each season.

The planning horizon (I) of 15 years beginning in 1988, 4 seasons (S) for each year, 4 time-bands (T) for each day and 1980 as base year for costs and loads are used in planning, and the growth in energy demand in subsequent years is set at an annual rate of 5.3%. Peak load and energy in the base year are taken to be 15,774 [MW] and 129,142 [GWH], respectively. The reliability standard, expected unserved energy, is set at 0.5% of the energy demand. Future costs escalate at an annual rate of 10%. A discount rate of 15% is used in computing present values. The costs are based on projections for 1980 escalated to the beginning of the planning horizon. The energy resource distribution factor β is 1 since all units are non-hydro type. Other necessary data are based on EPRI Report EL2561 [1].

Five types of alternative units are considered to be available for installation in each year, as described in Table 3. The teething factor for each plant type is 0.8.

The computed optimal solution is shown in Table 4. The 3rd plant type (Coal II) has more added capacity during the planning period than the 2nd plant type (Coal I). This is because, though the Coal I units have slightly lower maintenance cost, the Coal II units have higher availability and lower maintenance rate but the same fuel cost. By varying the planning horizon, it was noticed that each run gives almost identical capacity additions compare to the 15 year run. The most of combustion turbine units were installed during the last years of the planning period. This is reasonable solution path since the fuel savings associated with the base nuclear units are less pronounced during the last years of the planning period. If a salvage value assigned to the capital stock along with associated costs is not considered, there will be a tendency of installing more capacities to the lower capital cost units as the year increases.

Table 1. Comparison between conventional and presented methods

<table>
<thead>
<tr>
<th>Plant Type</th>
<th>Total Capacity [MW]</th>
<th>Availability [%]</th>
<th>Energy Generated [GWH]</th>
<th>Capacity Factor [%]</th>
<th>Energy Generated [GWH]</th>
<th>Capacity Factor [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear</td>
<td>7,200</td>
<td>0.850</td>
<td>4,455.36</td>
<td>85.00</td>
<td>4,455.36</td>
<td>85.00</td>
</tr>
<tr>
<td>Nuclear</td>
<td>800</td>
<td>0.850</td>
<td>495.06</td>
<td>85.00</td>
<td>495.06</td>
<td>85.00</td>
</tr>
<tr>
<td>Coal</td>
<td>800</td>
<td>0.760</td>
<td>442.16</td>
<td>75.92</td>
<td>442.16</td>
<td>75.92</td>
</tr>
<tr>
<td>Coal</td>
<td>1,800</td>
<td>0.790</td>
<td>1,028.11</td>
<td>78.46</td>
<td>1,035.21</td>
<td>79.00</td>
</tr>
<tr>
<td>Coal</td>
<td>2,000</td>
<td>0.870</td>
<td>1,222.92</td>
<td>83.99</td>
<td>1,266.50</td>
<td>87.00</td>
</tr>
<tr>
<td>Coal</td>
<td>6,400</td>
<td>0.950</td>
<td>4,392.41</td>
<td>72.69</td>
<td>4,127.18</td>
<td>85.90</td>
</tr>
<tr>
<td>Oil</td>
<td>800</td>
<td>0.760</td>
<td>214.00</td>
<td>53.92</td>
<td>336.9311</td>
<td>57.51</td>
</tr>
<tr>
<td>Oil</td>
<td>1,800</td>
<td>0.790</td>
<td>635.32</td>
<td>48.50</td>
<td>733.3517</td>
<td>55.96</td>
</tr>
<tr>
<td>Oil</td>
<td>900</td>
<td>0.870</td>
<td>298.41</td>
<td>51.24</td>
<td>333.6385</td>
<td>57.28</td>
</tr>
<tr>
<td>Oil</td>
<td>4,600</td>
<td>0.926</td>
<td>1,597.30</td>
<td>47.70</td>
<td>1,455.9280</td>
<td>43.77</td>
</tr>
<tr>
<td>Combustion</td>
<td>4,800</td>
<td>0.760</td>
<td>298.19</td>
<td>8.53</td>
<td>309.3273</td>
<td>8.85</td>
</tr>
</tbody>
</table>

A. Total Energy Generated [GWH] 14,279.62 14,999.1187
B. Unserved Energy [GWH] 14.11 18.0835
C. Total (Aw) 14,293.73 15,017.2022
D. Total Load Demand 15,017.24
E. E Error in Energy Balance 4.82 0.00
F. LOLP (p=+0.0024 0.0256

Table 2. Data for existing units

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear</td>
<td>1,700</td>
<td>0.851</td>
<td>2.41</td>
<td>210</td>
<td>0.1342</td>
<td></td>
</tr>
<tr>
<td>Total (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coal</td>
<td>600</td>
<td>0.845-0.917</td>
<td>4.21</td>
<td>450</td>
<td>0.1253 – 0.5075</td>
<td></td>
</tr>
<tr>
<td>Total (14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oil</td>
<td>800</td>
<td>0.824-0.919</td>
<td>11.30</td>
<td>195</td>
<td>0.0999 – 0.0575</td>
<td></td>
</tr>
<tr>
<td>Total (20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combustion</td>
<td>50</td>
<td>0.895</td>
<td>12.16</td>
<td>235</td>
<td>0.0384</td>
<td></td>
</tr>
<tr>
<td>Total (94)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Data for candidate units, available each year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear</td>
<td>1,200</td>
<td>0.851</td>
<td>1.157</td>
<td>2.41</td>
<td>210</td>
<td>0.1342</td>
</tr>
<tr>
<td>Coal I</td>
<td>400</td>
<td>0.876</td>
<td>735</td>
<td>4.21</td>
<td>450</td>
<td>0.0959</td>
</tr>
<tr>
<td>Coal II</td>
<td>200</td>
<td>0.917</td>
<td>695</td>
<td>4.21</td>
<td>516</td>
<td>0.0575</td>
</tr>
<tr>
<td>Coal III</td>
<td>400</td>
<td>0.876</td>
<td>341</td>
<td>11.30</td>
<td>195</td>
<td>0.0975</td>
</tr>
<tr>
<td>Combustion</td>
<td>50</td>
<td>0.895</td>
<td>152</td>
<td>12.16</td>
<td>235</td>
<td>0.0384</td>
</tr>
</tbody>
</table>

Maintenance rate but the same fuel cost. By varying the planning horizon, it was noticed that each run gives almost identical capacity additions compare to the 15 year run. The most of combustion turbine units were installed during the last years of the planning period. This is reasonable solution path since the fuel savings associated with the base nuclear units are less pronounced during the last years of the planning period. If a salvage value assigned to the capital stock along with associated costs is not considered, there will be a tendency of installing more capacities to the lower capital cost units as the year increases.
In this example, the variation of capital costs showed to give a sensitive change to a configuration of added capacities in Table 4. For example, an increase of 10% in capital cost of nuclear units caused 36% reduction in their total added capacity while the most of this reduced capacity was deslocated to the Coal I and II units as expected. The grand total of added capacities was decreased by 3.2% since the units with higher availability and lower maintenance rate need less capacity to be installed in order to meet the reliability constraints and the load demands. The total cost, though the capital cost was increased for nuclear units, decreased by 0.7% due to the less amount of the total installed capacity. The optimization algorithm of the master routine was iterated 9 times by means of AS 9000/N computer system. The CPU time for this example is 22.32 sec., which shows that the presented method is significantly much faster than the conventional one due to its analytical nature. Table 5 shows CPU time requirements with respect to the number of different load curves or seasons used in a year. The conventional method normally uses one shape of load curve for whole planning period.

### Table 4. Computed result of 15-year generation expansion planning

<table>
<thead>
<tr>
<th>Plant Type</th>
<th>New Added Capacities [MW]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1988</td>
<td>1,021</td>
</tr>
<tr>
<td>1989</td>
<td>1,681</td>
</tr>
<tr>
<td>1990</td>
<td>1,559</td>
</tr>
<tr>
<td>1991</td>
<td>1,721</td>
</tr>
<tr>
<td>1992</td>
<td>1,942</td>
</tr>
<tr>
<td>1993</td>
<td>1,968</td>
</tr>
<tr>
<td>1994</td>
<td>1,984</td>
</tr>
<tr>
<td>1995</td>
<td>2,085</td>
</tr>
<tr>
<td>1996</td>
<td>2,633</td>
</tr>
<tr>
<td>1997</td>
<td>2,820</td>
</tr>
<tr>
<td>1998</td>
<td>3,037</td>
</tr>
<tr>
<td>1999</td>
<td>3,118</td>
</tr>
<tr>
<td>2000</td>
<td>3,156</td>
</tr>
<tr>
<td>2001</td>
<td>3,329</td>
</tr>
<tr>
<td>2002</td>
<td>3,137</td>
</tr>
</tbody>
</table>

Total (MW): 32,054
Grand total (MW): 48,138
Total Cost [$]: 3.88261 x 10^10

### Table 5. CPU time requirements

<table>
<thead>
<tr>
<th>No. of Load Curves</th>
<th>CPU time [sec.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.98</td>
</tr>
<tr>
<td>2</td>
<td>16.02</td>
</tr>
<tr>
<td>3</td>
<td>22.32</td>
</tr>
</tbody>
</table>

### CONCLUSIONS

This paper presents a new analytical approach to the optimal long-range generation planning with the following conclusions.

1. The analytic formulas are shown to be sufficiently accurate in spite of the assumed Gaussian distribution of loads and plant availability.
2. The investment optimization algorithm, based on the discrete maximum principle and the gradient projection method is shown to be reliable and much faster than the conventional one due to its analytical nature.
3. The new algorithm allows for the incorporation of any kind of constraints which are found necessary.
4. The extended studies on both the hydro and maintenance schedule subproblems will supplement and complete the research on the optimal long-range generation expansion planning. These will be reported in near future.

### ACKNOWLEDGEMENT

The authors gratefully acknowledge the Korea Electric Company, and the University of Houston who supported this project under the New Research Opportunity Program.

### REFERENCES

APPENDIX A: EXPECTED PLANT OUTPUTS

The total power output by the plants from type 1 up to j, \( J^\Delta \), is given by (5):

\[
J^\Delta = \min(L^\Delta, \frac{1}{j} \sum_{k=1}^{j} Y^\Delta_k)
\]  

(A1)

with the probability density function

\[
f_p(z) = (1-F_L(z))f_Y(z) + (1-F_Y(z))f_L(z)
\]  

(A2)

where \( f_L(\cdot) \), \( f_Y(\cdot) \), and \( f_L(\cdot) \), respectively, are the probability density functions of \( J^\Delta \), \( \sum_{k=1}^{j} Y^\Delta_k \) and \( L^\Delta \).

\( F_L(\cdot) \) and \( F_Y(\cdot) \) are corresponding distribution functions.

Accordingly, the expectation of (A1) is expressed by

\[
J^\Delta_p = \int_{-\infty}^{\infty} z f_p(z) dz
\]  

(A3)

On the other hand, the assumed Gaussian distributions of the loads and the available capacities, and mutual independence among them give the following probability density and distribution functions:

\[
f_L(z) = \frac{1}{(\sqrt{2\pi} \sigma_L^2)} \exp\left\{-0.5(z-\mu_L^2/\sigma_L^2)\right\}
\]  

(A4)

\[
f_Y(z) = \frac{1}{(\sqrt{2\pi} \sigma_Y^2)} \exp\left\{-0.5(z-\mu_Y^2/\sigma_Y^2)\right\}
\]  

(A5)

\[
F_L(z) = 0.5 + \text{erf}\left((z-\mu_L)/\sigma_L\right)
\]  

(A6)

\[
F_Y(z) = 0.5 + \text{erf}\left((z-\mu_Y)/\sigma_Y\right)
\]  

(A7)

Substituting Eq. (A4) through (A7) in Eq. (A2) and then using Eq. (A3), Eq. (6) is derived.

The positive value of random variable \( Z \) defined in Eq. (B) is the shortage of generation. Since \( f_L \) and \( f_Y \) are Gaussian, the probability density function of \( Z, f_Z^\Delta \), is also Gaussian. Thus, LOLP\(\Delta \) is

\[
\text{LOLP}^\Delta = \int_0^{\infty} f_Z^\Delta(z) dz = 0.5 + \text{erf}\left(\frac{Z^\Delta}{\sigma}\right)
\]  

(B8)

APPENDIX B: EXPECTED MARGINAL VALUES

The expected marginal values of annual costs and reliability measure are computed for the optimal solution to the operating subproblem, and then transferred to the master problem to be used by the Gradient Projection Algorithm to find the optimal planning solution.

The expected marginal values are computed analytically by taking the partial derivatives of the reliability measure \( R_i \) in (14) and the expected annual cost \( G_i \) in (16) with respect to the previously installed capacity \( x_i^k \) and the newly installed capacity \( u_i^k \). They are summarized as follows:

\[
g_Ri = k S T \sum_{s=1}^{S} \sum_{t=1}^{T} \left[-p^k(1-v^k_s)\beta^k \left\{ 0.5 + \text{erf}(\frac{J^\Delta}{\sigma}) \right\} + \left( p^k(1-p^k)(1-v^k_s) \beta^k k/(2\sqrt{2\pi} \sigma) \right) \exp\left\{-0.5 \left( J^\Delta - J^2 \right)^2/\sigma^2 \right\} \right]
\]  

(B1)

\[
g_G_i = \frac{a^f_i}{a^u_i} + \frac{a^m_i}{a^u_i} \frac{\sigma}{a_i^k}
\]  

(B2)

\[
g_x^k = \frac{a^c_i}{a^u_i} + \frac{a^f_i}{a^u_i} + \frac{a^m_i}{a^u_i} \frac{\sigma}{a_i^k}
\]  

(B3)

where

\[
h(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \\ \end{cases}
\]
Discussion
L. R. Noyes (Philadelphia Electric Co., Philadelphia, PA): The authors have presented an excellent framework for developing optimal long-range generation plans. However, I question the appropriateness of using a Gaussian distribution to model the skewed distributions of load and capacity in calculating production costs. Most researchers involved with applying statistical procedures to reliability and production cost analyses have used methods which recognize higher order moments [20-23] since they have found the use of merely the first and second order moments to be inadequate.

The authors purport to have tested the accuracy of the Gaussian model by comparing it with results of a recursive model [16]. Having observed differences between the results, the authors chose to attack the accuracy of their benchmark rather than questioning either the accuracy of their Gaussian model or the validity of their comparison.

I did not believe that the magnitude of apparent errors in the results of the recursive model could be due to numerical considerations as suggested by the authors. Therefore I investigated the load and capacity data in [15] and [16] and found that the load data given in Tables 4-2 and 4-5 of [16] had unfortunately been interchanged. While page 4-1 and the footnote of Table 4-2 of [16] indicate that the spring/fall load cycle of Scenario D was used to obtain the recursive results, the data in Table 4-2 are the summer load cycle of Scenario B (see also Tables 4-27 and 4-34 of [15]). Thus the recursive model results were undoubtedly based on the load cycle of Scenario D while the results of the authors' Gaussian model were based on the load cycle of Scenario B.

Since the recursive results were based on a load model with a minimum load of 11,994 MW, the observed capacity factor of the 5th plant type is justifiably lower than its availability factor. In addition, there is zero error in energy balance rather than the stated 4.82%.

Given the invalid comparison, the authors have no basis for attacking the accuracy of the recursive model or evaluating the accuracy of the Gaussian model. It would be helpful if the authors could report the results of the Gaussian model using the load cycle of Scenario D. If, as I suspect, the Gaussian model should prove to be inadequate, the authors should consider extending their work to include higher order moments through the use of cumulants and Gram-Charlier series approximations.

I also note that since the variance of available generating capacity of each plant type is proportional to the sum of the squares of each individual unit capacity, the teething factor in Eq. 4 should be squared.

REFERENCES

Manuscript received August 6, 1984.

Y. M. Park, K. Y. Lee and L. T. O. Youn: We appreciate Mr. Noyes's interest in the reported work and thank him for his kind comments.

As Mr. Noyes mentioned, we also discovered that Table 4-2 and 4-5 of [16] had unfortunately been mislabeled. Therefore, we tried both load curves given in Table 4-2 and 4-5 of [16] to find out which one was used to obtain the recursive result reported in Table 4-8 of [16]. The study showed that, though the spring/fall load cycle of Scenario D was mentioned, the summer load cycle of Scenario B had actually been used. It is very easy to verify that the total load demand of the spring/fall load cycle of Scenario D is 13,983.57 [GWH], which is the area under the load curve. However, the results in Table 4-8 of [16] as well as Table 1 of our paper show that the sum of the generation and unserved energy is 14,293.73 [GWH], which is more than the demand of Scenario D. Therefore, it is hard to believe that the Scenario D is used for the recursive model results in Table 4-8 of [16]. To reaffirm our conclusion, one can observe that the 3rd and 4th plants will actually generate more energy than those shown because their capacity factors are 0.08% and 0.54% less than the respective availabilities. This makes the actual sum of the generation and unserved energy greater than the reported value of 14,293.73 [GWH]. On page 4-1 of [16], it states that "a peak load of 26,000 [MW] is used to represent the system load." Thus, the minimum loads in both load data of Scenario B and D are 12,607 [MW] and 13,842 [MW], respectively, (obtained from the weekly-hourly loads). Since the total installed capacity from the plant type 1 up to 5 is 12,600 [MW], the capacity factors of these plants must be equal to their respective availabilities, no matter which load data could have been used.

Also, we would like to state that the total load demand of the summer load cycle of Scenario B is 15, 017.24 [GWH]. The reported figure of 14,293.44 [GWH] in Table 4-8 of [16] is neither for Scenario B nor for Scenario D since the demand of Scenario D is 13,983.57 [GWH]. Just for the 4th plant alone, there is a numerical error of 7.1 [GWH]. Thus, how can the recursive result have the zero error in energy balance? The analytical approach, in spite of Gaussian model, showed consistencies in the capacity factors and the energy balance as shown in Table 1 of our paper.

The error which can occur when the load curve is converted into the load duration is eliminated, and also the use of the concept of power in the original load curve rather than the concept of energy in the load duration curve will facilitate to solve the generation expansion planning problem when the hydro plants are included.

In IEEE Region 5 Conference 1983 [24], we also mentioned that although the Gaussian assumption gives sufficiently reliable results, further accuracy can be obtained by reflecting higher moments in the probability density functions in the model. The extended study including higher moments will be reported in the near future.

Regarding the Gaussian load model, we would like to emphasize that our analytical method is based on the original load curve rather than the load duration curve as shown in Fig. 1. The suggested use of cumulants and Gram-Charlier series is applicable for the equivalent load duration curve, but not for the original load curve.

Finally, there is no need to square the teething factor for the variance of the available generating capacity. This is because the identical multiple units are assumed for each plant and the variance of each plant is the sum of the variance of each individual unit.

Fig. 1. Gaussian random load fluctuations.

REFERENCE

Manuscript received September 4, 1984.