Linear and Mixed-Integer Programming Models of Investment in the Electric Power Industry

Peter Schaeffer and Lou Cherene

I am making this paper, which first appeared as chapter 4 of a report to the Electric Power Research Institute (EPRI), available online. It is a survey of linear programming (LP) and mixed integer linear programming (MILP) model and applications to the electric power industry. Dr. Richard H. Day was the principal investigator and project director of the EPRI-funded research. To show the scope of the whole project, I included Dr. Day’s introductory statement and the report’s table of contents.

There are a few differences between this version and the original version. I corrected typographical and spelling errors that had gone undetected in the original document. The page numbers are also a little different, even if it is not by very much. The page numbers in the table of contents have been adjusted accordingly. I recreated most of the figures and retyped table 4-4 for what I hope is easier reading of its contents.

Although the electric power industry has undergone fundamental changes since we conducted the research reported here, the paper still has some relevance. It provides a summary of the state-of-the-art at a time when the desktop computer revolution had barely begun and when computer capacities imposed limitations on the size and complexity of model. Thus, at a minimum, it is of some historical interest. In addition, the survey shows the ingenuity of scholars to overcome computing limitations and analyze important questions concerning investments into electricity generating capacity.

The issue addressed by the authors discussed in this survey are still relevant. What is the optimal mix of different plant types when changing demand must be met at a moment’s notice? The growing availability of solar and wind power are raising questions about system’s reliability with new twists as such energy is not always available when needed most, and when it is available, it may not be needed. This problem will persist until we find relatively inexpensive ways of storing energy.

Therefore, there are practical reasons to study past contributions and learn from the insights they generated. If you use this paper in your research, please reference it as follows:


Peter V. Schaeffer, May 17, 2017
ABSTRACT
As part of a program of research sponsored by the Electric Power Research Institute, Economic Dynamics has developed a new, dynamic, quantitative, simulation model of technology adoption in the thermal electric power generation industry. The model, which incorporates engineering data, economic analysis and operational rules of managerial behavior provides an improved understanding of how technological, economic and behavioral considerations interact at the micro level to produce temporal patterns of industry behavior and performance. In addition to providing a better description of industry dynamics the model should provide a useful tool for studying likely economic responses by electric utilities to external conditions. In particular, the model should be applicable to the problem of helping to identify suitable technologies for the efficient and environmentally acceptable production of electricity that will meet future needs while at the same time satisfying the physical, regulatory and financial constraints that impinge on industry decision-making.

The detailed project description is divided into two parts. The first, which is the subject of this part of the final report, deals with load curves and the capacity expansion decision. The reason for singling out this aspect prior to the more comprehensive analysis of industry planning and behavior is because of its unique importance in the economics of electric power generation. Traditional economic theory has not incorporated the temporal distribution of demand. Mathematical programming models of investment in electric power generation have also failed in this regard, replacing the demand load distribution curve with the simpler cumulative load duration curve. In several chapters of the present volume, Gordon C. Winston reexamines the traditional theory of production and investment. He identifies the major potential effects of load curves on the economics of capital utilization and capacity expansion. This is followed by a review of the literature dealing with capacity expansion. This review is then followed with a description, authored by Lou Cherene and Peter Schaeffer, of a new mixed-integer programming model that, unlike its predecessors, incorporates the load curve explicitly in a computationally economical manner. While this model incorporates the long horizon required for capital intensive investment, it treats uncertainty in a somewhat unsatisfactory way by
means of what are, in effect, safety-first constraints. In order to study the effects of uncertainty on technology choice in a theoretically more satisfying way, a dynamic programming model that incorporates probabilistic demand forecasts is reviewed. Experimental computations illustrate the advantage which flexible, short lead time plants possess in an uncertain environment. The report concludes with an appendix which summarizes the general features of alternative generation technologies.

The second part of the study, which is described in the third volume of this final report, incorporates a model of capacity expansion into a dynamic, adaptive simulation framework which is used to study the behavior of production, technology diffusion, and capacity expansion over time. In principle, this model could incorporate as modules or components either of the mathematical programming models described in Volume 2 of the study; or, a third module that incorporated the best features of both could be used. For purposes of the present experiment, however, a much simplified version of the load curve-capacity expansion-technology choice module was used. The primary product of the resulting analysis is an improved understanding of how demand forecasting, technology choice, capacity expansion, financing, pricing and regulatory activities interact to produce specific patterns of industrial evolution. In particular, the emergence after a prolonged period of efficacious growth of financial difficulties related to rising input prices and the resort to less capital intensive, shorter-lead time technologies is brought out in a quite realistic manner.
EPRI PERSPECTIVE

(no text)
CONTENTS
Volume 2

Abstract ii
EPRI Perspective iv
Contents v
Detailed Contents vi

Section

1 Introduction and Summary by Richard H. Day 1 – 1

2 The Effects of Rhythmic Demand on Investment and Choice of Technology in Electric Power Generation: A Time-Specific Analysis by Gordon C. Winston 2 – 1

3 Electricity, Load Curves and Technology: The State of Empirical Knowledge by Gordon C. Winston 3 – 1

4 Linear and Mixed-Integer Programming Models of Investment in the Electric Power Industry by Peter Schaeffer and Lou Cherene 4 – 1

5 Load Curves and Investment in Electrical Generation: A Mixed-Integer Programming Model by Lou Cherene and Peter Schaeffer 5 – 1

6 Construction Lags, Uncertainty and the Economics of Capacity Expansion by Robert Boyd and Roderick Thompson 6 – 1

7 Appendix: An Overview of Technological Aspects of Electric Power by M. Hussain Toofaninejad 7 – 1
# DETAILED CONTENTS

## Volume 2

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction and Summary by Richard H. Day</td>
<td>1–1</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>1–1</td>
</tr>
<tr>
<td></td>
<td>Summary of Volume 2</td>
<td>1–1</td>
</tr>
<tr>
<td>2</td>
<td>The Effects of Rhythmic Demand on Investment and Choice of Technology in Electric Power Generation: A Time-Specific Analysis by Gordon C. Winston</td>
<td>2–1</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>2–1</td>
</tr>
<tr>
<td></td>
<td>Summary and Implications</td>
<td>2–2</td>
</tr>
<tr>
<td></td>
<td>Load Curves: Rhythmic Demand</td>
<td>2–7</td>
</tr>
<tr>
<td></td>
<td>Time-of-Day (Year) Analysis: Production as a Process in Time</td>
<td>2–9</td>
</tr>
<tr>
<td></td>
<td>Tracking Load Curves I – Varying Utilization</td>
<td>2–13</td>
</tr>
<tr>
<td></td>
<td>Tracking Load Curves II – Varying Speed</td>
<td>2–22</td>
</tr>
<tr>
<td></td>
<td>Changing Relative Input Prices</td>
<td>2–30</td>
</tr>
<tr>
<td></td>
<td>The Stochastic Component of Demand</td>
<td>2–31</td>
</tr>
<tr>
<td></td>
<td>Load Curve Modification</td>
<td>2–33</td>
</tr>
<tr>
<td></td>
<td>Modelling Load Curves and Technology</td>
<td>2–35</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
<td>2–37</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>2–38</td>
</tr>
<tr>
<td>3</td>
<td>Electricity, Load Curves and Technology: The State of Empirical Knowledge by Gordon C. Winston</td>
<td>3–1</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>3–1</td>
</tr>
<tr>
<td></td>
<td>Models of Generating Technology and System Load Information Requirements</td>
<td>3–2</td>
</tr>
<tr>
<td></td>
<td>How Load Curves are Modelled in Economic Studies</td>
<td>3–9</td>
</tr>
<tr>
<td></td>
<td>Available and Proposed Information: Load Curve Studies</td>
<td>3–10</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>3–17</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>3–17</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Pages</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>4</td>
<td>Linear and Mixed-Integer Programming Models of Investment in the</td>
<td>4 – 1</td>
</tr>
<tr>
<td></td>
<td>Electric Power Industry by Peter Schaeffer and Lou Cherene</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>4 – 1</td>
</tr>
<tr>
<td></td>
<td>Mathematical Preliminaries</td>
<td>4 – 11</td>
</tr>
<tr>
<td></td>
<td>Survey of Linear Programming Models</td>
<td>4 – 15</td>
</tr>
<tr>
<td></td>
<td>Summary and Comparison</td>
<td>4 – 71</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>4 – 75</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>4 – 76</td>
</tr>
<tr>
<td>5</td>
<td>Load Curves and Investment in Electrical Generation: A Mixed-Integer</td>
<td>5 – 1</td>
</tr>
<tr>
<td></td>
<td>Programming Model by Lou Cherene and Peter Schaeffer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>5 – 1</td>
</tr>
<tr>
<td></td>
<td>Load Curves and the Reckoning of Time</td>
<td>5 – 6</td>
</tr>
<tr>
<td></td>
<td>Generating Equipment and Production Activities</td>
<td>5 – 7</td>
</tr>
<tr>
<td></td>
<td>The Objective Functional</td>
<td>5 – 10</td>
</tr>
<tr>
<td></td>
<td>Utilization Constraints</td>
<td>5 – 10</td>
</tr>
<tr>
<td></td>
<td>Demand and Generating Capacity Constraints</td>
<td>5 – 12</td>
</tr>
<tr>
<td></td>
<td>Investment and Maintenance Constraints</td>
<td>5 – 14</td>
</tr>
<tr>
<td></td>
<td>What is Not Constrained</td>
<td>5 – 15</td>
</tr>
<tr>
<td></td>
<td>The Size of the Model</td>
<td>5 – 15</td>
</tr>
<tr>
<td></td>
<td>The Solution</td>
<td>5 – 16</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>5 – 18</td>
</tr>
<tr>
<td>6</td>
<td>Construction Lags, Uncertainty and the Economics of Capacity</td>
<td>6 – 1</td>
</tr>
<tr>
<td></td>
<td>Expansion by Robert Boyd and Roderick Thompson</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>6 – 1</td>
</tr>
<tr>
<td></td>
<td>The Taxonomy of Demand Uncertainties</td>
<td>6 – 2</td>
</tr>
<tr>
<td></td>
<td>The Reliability Constraint and the Planning Reserve Margin</td>
<td>6 – 6</td>
</tr>
<tr>
<td></td>
<td>The Model</td>
<td>6 – 7</td>
</tr>
<tr>
<td></td>
<td>The Results</td>
<td>6 – 13</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>6 – 25</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>6 – 26</td>
</tr>
<tr>
<td>7</td>
<td>Appendix: An Overview of Technological Aspects of Electric Power</td>
<td>7 – 1</td>
</tr>
<tr>
<td></td>
<td>by M. Hussain Toofaninejad</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>7 – 1</td>
</tr>
<tr>
<td></td>
<td>The Electric Power Industry</td>
<td>7 – 1</td>
</tr>
<tr>
<td></td>
<td>Electricity Generation Technologies</td>
<td>7 – 2</td>
</tr>
<tr>
<td></td>
<td>Technical Aspects of Electricity Generation</td>
<td>7 – 6</td>
</tr>
<tr>
<td></td>
<td>Economic Aspects of Electricity Generation</td>
<td>7 – 7</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>7 – 27</td>
</tr>
</tbody>
</table>
INTRODUCTION

The pioneering application of linear programming techniques to the problem of electricity generation in 1957 (Massé and Gibrat, 1957) has given rise to many similar modelling efforts. Process analysis has been applied to solve problems of short term dispatching of generators to hourly variations in demand to seasonal scheduling of maintenance to long term investment plans for capital equipment. Of course, other modelling methods, econometrics, for instance, have also flourished on data generated by electrical utilities, but the importance of more detailed results and the availability of technical engineering data makes linear programming a viable and inviting approach.

The existence of many large electrical utility firms, either government owned or regulated, indicates the importance of this industry to the economies of the developed world, and, hence, to the desirability of building and maintaining detailed models for their efficient operation and expansion. Indeed, the generation of electricity is considered a prerequisite to industrial growth by less developed countries and a key factor to the maintenance of living standards among industrialized nations. Thus, the expected returns to the allocation of effort in optimizing the supply of electric power are high.

Certain characteristics of the electricity market have given rise to central recurring themes of these models. High capital intensity and long construction lead times make accurate demand forecasting essential for avoiding over-investment and excess capacity on one side, power outages and "brown outs" on the other. The issue of returns to scale is unresolved, for larger generators, longer transmission lines, and fluctuations in demand all interplay in the determination of average costs. The oil embargo of the early 1970’s introduced uncertainties of fuel prices and fuel availability into decisions regarding selection of plants. Pollution and siting restrictions are becoming more important as the size of the capital stock expands and the most
efficient locations are exhausted. Regular changes in demand with daily peaks and troughs in consumption of this non-storable commodity creates complicated problems of dispatching and pricing. These issues arise in many of the models presented in this paper, requiring special treatment of variables and constraints in order to capture their essential effects upon optimal investment plans and operating policies.

This survey shall start with a review of these issues and their ramifications for the results of these models, followed by a short mathematical section dealing with integer programming and its uses in costing generation activities. The survey of the models then commences. The models surveyed are only those of the linear programming or mixed-integer-linear programming (MILP) type that have been published for public scrutiny. The list of models in Table 4-4 is by no means exhaustive, but represents an array of approaches within the linear programming framework as broad as the specific issues each model attempts to deal with in a realistic and computationally feasible manner.

**Economies of Scale**

In 1930 the per capita consumption of electrical energy was 740 kilowatt hours (KWH). The total production of electrical power was 91 billion of KWH, most of which was consumed by industrial users. 70.9 percent of the total went to industrial, 16.3 percent to commercial customers, and the remaining 12.8 percent was used by households. By 1972 total production had increased to 1,748 billion of KWH, and of this the industrial sector used 41.5 percent, the commercial sector 23.4 percent, and the residential sector 35.1 percent. The per capita consumption for that year was 8,394 KWH (Scott, 1976, pp. 8-9).

The electric utilities have successfully met the challenge of quickly growing and changing demand patterns. The average unit and plant size grew considerably. In 1930 the largest steam electric unit in the United States had a capacity of about 200 megawatts (MW), and the average size was 20 MW. Forty years later, the largest unit was rated at 1,150 MW and three 1,300 MW units had been ordered. The units under construction had an average size of 450 MW, the average of all units was about 160 MW (Federal Power Commission, 1970, p. I-5-2).
The large increase in the unit size leads one to suspect that considerable economies of scale do exist. However, the empirical evidence on this matter is ambiguous, and several reservations must be made. First, one needs to distinguish between economies of scale and technical progress. This is particularly important when time series data are being used, where the economies of new technology grow simultaneously with demand. An excellent survey can be found in Galatin’s book (1968). The newer econometric research is reviewed in Cowing and Smith (1978). The second important consideration regards the level at which economies of scale are being measured. Nerlove (1963, as quoted in Cowing and Smith, 1978) presented an example which illustrates the possibility of experiencing economies of scale at the plant level and diseconomies of centralization at the level of the firm due to transmission costs. The complex relationships between the activities at different levels of the production-distribution process to the flow of demand over time leaves the issue of returns to scale verifiable only by simulation of the process as a whole at a given level of demand and stock of equipment. Lumpiness of utilization and surface area to volume ratios tend to increase efficiency as plant size grows, constant alterations of output rates, decreasing reliability, and long transmission lines tend to decrease these efficiencies.

For simulation, there is also the question of how to allow for scale effects in a linear model. As we shall see, this can only be done if integer variables are introduced. It is of interest to see how different authors deal with the question of scale economies.

**Uncertainty**

Data illustrating the magnitude of the change the industry went through over the past 40 to 50 years have already been given above. During most of this time the industry had no troubles adjusting to higher levels of demand. The increased efficiency of newer and larger plants made it possible to keep the price of electricity low. Scott (1976, Figure 3-1) shows that the price paid by the consumers decreased significantly when expressed in “1968-dollars.” This period ended around 1970 after which consumer prices started to rise. Several factors contributed to this change. Rapidly rising fuel costs played their part. According to the 1970 National Power Survey (Part I, Figure 5-6), approximately 40 percent of the costs of producing electrical power in 1969
were attributed to expenditures on fuels. Total costs consisted of a fixed charge which included an allowance for construction costs, fuel costs, and operation and maintenance costs. The 40 percent represented an average over a large number of privately owned utilities and did not refer to a particular plant size. Given the importance of the fuel costs, it is clear that uncertainty about future prices and the availability of fuels makes long term investment planning difficult and risky.

There are several ways for the utilities to deal with these problems. They can try to increase their flexibility by investing in equipment that can burn different fuels; The price to be paid is a decrease in efficiency when compared to more specialized equipment. If the efficiency loss is considerable it may be advantageous to build more but smaller plants burning different fuels. In this case a loss in economies of scale is offset by a gain in flexibility. Another alternative is to buy deposits of the fuel or make long term contract for fuel supplies. This has the disadvantage of binding the utility for a long time. Unfortunately, the most desirable alternative, namely improved forecasts, seems not to be feasible. So long as the political situation in the Near and Middle East remains unstable, developments in the world oil market will remain highly uncertain.

There are several other sources of uncertainty to deal with. As already mentioned, demand has to be met instantly. Therefore, planned capacity must be sufficient to satisfy expected future demand. But since it is not known for certain what future demand levels will be, uncertainty about the level of demand enters into the decision-making process. This problem is aggravated by the existence of long lead times necessary to build some power plants. For nuclear and conventional steam thermal plants, the lead times are between 8 – 15 years. Boyd and Thompson (1979) refer to this as Type II uncertainty. Type I uncertainty causes demand and supply to be uncertain on a time scale substantially shorter than that required to build a power plant usually due to variations in weather. Causes of Type I uncertainty also include forced outages, which is related to the problem of the reliability of the equipment.

This deserves a short comment because it is related to economies of scale. The newer and bigger plants are more efficient. But generally, they are less reliable than the smaller plants. Even in the case where there is no difference in reliability between plants of different sizes, the system’s reliability will decrease as the average plant size increases. Suppose that one 1,000 MW plant
replaces four identical 250 MW plants, and that the mathematical expectation of forced outage is the same for all plants, say 1/10. The probability that all four small plants, which together have the same capacity as the new one, are down together at the same time, is 1/1000. Compare this to 1/10 for the large plant. Thus, the mathematical expectation that generating capacity equivalent to that of the large plant must be backed up at a moment’s notice, has increased greatly. Since capacity is expensive, the necessary reserves are important in considering size and technology for a new plant.

During times of technological advances there will always exist a tendency to wait to the last moment with ordering a new plant. Currently, however, due to the markedly increased costs of electricity and also due to the problems associated with nuclear power plants, a whole range of new options is being investigated. An assessment of many new technologies can be found in Hottel and Howard (1971). For the utilities, there is the question as to what new technologies will become available during their planning period. If they invest heavily using technologies existing today, they run the danger that the equipment may become obsolete in the near future.

The capital requirements of the electric utilities exceeded 15 percent of the new personal savings in the economy in recent years. The utilities’ ability to finance themselves internally has decreased from 48.4 percent in 1966 to 33.8 percent in 1973 (Weidenbaum, 1975). Given this situation, the industry faces at least two sources of uncertainty. The first is due to regulation. How will the regulating agencies act in the future? Will they allow rate increases to help generate internal funds for investments? The second source is due to increased competition for funds. Other industries in the energy business also have adjustments to make, and need outside capital to do this (Weidenbaum, 1975). For the electric utilities, this, together with a decreased credit rating, will result in higher rates and more difficulties to get the money to finance their projects.

Virtually all uncertainties discussed above are affected by changes in regulatory practices. Future demand depends on price regulation. The choice of fuels depends on environmental requirements. The connection between credit rating, credit availability and regulation has already been mentioned. Examples of the effect of government regulations on construction lead times and costs can be found in Gilbert and Stiglitz (1978) and in Boris (1978). “Regulatory lag,”
where changes in prices occur long after operating costs have increased, affects the financial capabilities of the industry to invest and expand.

**Demand for Electricity**

The demand for electricity fluctuates regularly and periodically. Figure 4-1 provides an illustration of a representative weekly load curve, showing the differences in demand loads between hours of the day and days of the week. The seasonal as well as the regional differences are illustrated in Figure 4-2. The shape of the load curve and the duration of the peak demands are vital information for investment planners. The information contained in the load curves is usually condensed into the so-called load duration curve. An illustration of a plausible annual load duration curve is given in Figure 4-3. The load duration curve provides the information for how long a given level of demand has to be met. If the intention is to use linear programming techniques to determine optimal investment plans, the load and load duration curves can be approximated by discrete blocks.

The same load duration curve can result from different load curves. This may constitute a problem when load duration curves are being used. The need to constantly adjust to changing demand levels requires substantial flexibility from the electric utility. The load duration curve captures this need only partially. Consider a utility facing a daily load curve with a single peak of short duration. It may be optimal for this utility to operate cheap equipment burning expensive fuels to satisfy peak demand. Another utility faces a daily load curve with two peaks of short duration. The time between peaks is such that it is not economical to turn some equipment off between peaks. Under this set of conditions, a plant type using less expensive fuel may prove to be more cost efficient.

The possible importance of the shape of the load curves is discussed extensively in Winston (1979a, b). Most of the studies reviewed below use load duration curves in their models.
Optimizing Behavior

The neoclassical theory of the firm assumes that profits are being maximized. In an industry where the rate of return is fixed by regulatory agencies, this assumption may be invalid. Averch and Johnson (1962) maintained the hypothesis that profit is being maximized. If the permitted rate of return is less than the profit-maximizing rate of return, it has been shown that the firm will employ more capital than would have been the case in the absence of regulation. Baumol and Klevorick (1970) point out some of the shortcomings of the Averch-Johnson hypothesis, which shows that it is of limited use for understanding the investment behavior of electric utilities. The assumption that the firm is free to choose any price-output combination to maximize profit subject to the regulatory constraint is particularly important because it does not hold for electric utilities. They are required to meet demand at a regulated price. This point is important because it can be shown that firms obeying the Averch-Johnson assumptions will not choose cost minimizing input proportions to produce a given output (Baumol and Klevorick, 1970).

It has often been asserted that large companies where management and owners do not coincide maximize revenue rather than profit. Again, it can be demonstrated that a sales-maximizing firm does not choose the cost minimizing input combination if the rate of return is set by the regulators (Baumol and Klevorick; 1970).

Suppose that the management follows a strategy to maximize the rate of return. If the rate of return is fixed by regulation, then the labor-capital ratio will be indeterminate (Baumol and Klevorick, 1970). All input combinations which result in the regulated rate of return are equally preferred.

Given that electrical utilities must apply for price changes before the regulatory commission, we can argue that cost minimization is a rational objective for the management. If it can be demonstrated that electricity is produced at the lowest possible cost, it will be much easier to convince the commission of the necessity to raise the price.
Revenue requirements minimization is related to cost minimization. The two procedures are identical if we optimize over an infinite time horizon. Revenue minimization, where only capital costs that accrue to the utility prior to the end of the planning horizon are taken into account, may be a reasonable objective when financial constraints are of great importance.

The objective does not necessarily have to be in monetary terms. There may be projects where the main concern is to minimize the environmental impact. Financial considerations may enter as constraints. Other possibilities arise if the firm were viewed as optimizing more than one objective, as is generally the case. In the traditional linear programming approach, other objectives can enter as additional constraints or as penalties in the objective function. This may not always constitute a satisfactory treatment.

Figure 4-1. Weekly Load Curve. Source: Federal Power Commission, *The 1970 Power Survey*, Part I, Figure 3-1, page I-3-2.
Figure 4-2. 1968 - 69 Peak Day Load Patterns for Selected Systems (by FPC Region
Figure 16, page IV-4-22.
Figure 4-3. Annual Load Duration Curve
MATHEMATICAL PRELIMINARIES

Practically all the studies reviewed in Galatin (1968) and Cowing and Smith (1978) conclude that there exist economies of scale in the electric power industry. The only unresolved discussion is about the plant size at which all scale economies are exhausted. Therefore, it is important that ways to incorporate scale economies into linear investment models are found.

One solution is obvious. We can choose, for each plant type, typical sizes with their associated cost data. This requires that all investment variables be specified as integers. Such “lumpiness” of investment is a traditional source of economies of scale cited in economic theory.

There are two such approaches that have been developed. One is the fixed charge approach. Cost increases linearly with plant size, but total average cost decreases, due to the presence of warming up, cooling down, and maintenance – the fixed costs. A second method, proposed by Markowitz and Manne (1957), is the piecewise linear approximation approach, where the true production function exhibits increasing returns to scale.

Indivisibilities introduce additional difficulties into the solution process of linear programs. It is tempting to solve the program without paying attention to the integer constraints, and then round the solution to the next feasible integer. It does not require much time to demonstrate that this procedure may result in a suboptimal solution. Figure 4-4 illustrates a maximization problem with two constraints. Point (a) represents the optimal solution. If we round off to the nearest integer, the solution is (b). However, (c) is superior to (b). This proves that in general it is necessary to introduce integer constraints explicitly.

The first practical algorithms for integer linear programs were developed by Gomory (1963). Further references can be found in an extensive bibliography published by Springer Verlag (Kastning, 1976; Hausmann, 1978). An excellent survey article has been written by Balinski (1965). Geoffrion and Marsten (1972) provide a more recent review of available solution methods. Also of practical and theoretical interest is the study by Baumol and Gomory (1960). They discuss the interpretation of the dual to the optimal solution of a linear program with integer constraints. Under certain conditions the duals still have the interpretation of shadow prices.
Figure 4-4. Approximation of the Solution to an Integral Linear Program by the Solution of a Continuous Linear Program

**The Fixed Charge Problem**

We follow Dantzig (1963) for the discussion of the fixed-charge problem. The method is straightforward. Let $X$ be the level of output, $K$ the marginal cost, and $B$ the fixed charge. Suppose that cost, $C$, is characterized by the following relationship.

$$C = \begin{cases} KX + B & \text{if } X > 0 \\ 0 & \text{if } X = 0 \end{cases}$$

(4-1)

This can be transformed into a mixed-integer problem. Let $S$ be a 0,1 integer variable. Choose an upper limit of $X$, say $U$. Since $U$ can be made arbitrarily large, this does not restrict the original problem. Then (4-1) can be written in a new form.

$$C = KX + SB \text{ and } X \leq SU; S = 0 \text{ or } 1$$

(4-2)

If $S = 0, X = 0$, therefore, $C = 0$, and if $S = 1$, then $X > 0$ and $C = KX + B$. Clearly, (4-1) and (4-2) are equivalent formulations.
**Markowitz and Manne (1957)**

The approach developed by Markowitz and Manne is quite different. They deal with a production function displaying increasing returns to scale. The idea is to approximate the true function by a piecewise linear function. In an ordinary linear programming model, the solution algorithm would pick the most efficient activity first, then the second most efficient, and so on. This is equivalent to saying that we have decreasing returns to scale, the opposite effect of what is being modelled. To overcome this difficulty, new integer variables and additional constraints are introduced.

To better understand the working of the method, the reader is referred to Figure 4-5. It is the same as Figure 2 in Markowitz and Manne (1957) on page 86, except that the axes are interchanged to get a more familiar looking graph.

$F(R)$ is the true production function and $F'(R)$ is its approximation. Suppose that the goal is to maximize total output, $X$.

$$\text{Max. } X = \sum_{i=1}^{K} X_i$$  \hspace{1cm} (4-3)
\( K \) is the number of linear segments used to get the approximation of \( F(R) \). The larger \( K \), the better the approximation. To complete the problem, constraints are needed which ensure that not more than the available amount of resources, \( R \), is used, and that the least efficient activity is used first, the next worst second, etc. Let \( a_i \) be the input requirement per unit of output of activity \( i \), \( v_i \) the upper bound of output level \( i \), and the \( q_i \) the auxiliary 0,1 integer variables. Now consider the following system of constraints.

\[
\sum_{i=1}^{K} a_i X_i \leq \bar{R}; \quad X_i \geq 0 \quad (4-4)
\]

\[
q_i \geq \left( \frac{1}{v_i - v_{i-1}} \right) X_i; \quad q_i = 0 \text{ or } 1; \quad i = 1, \ldots, K \quad (4-5)
\]

\[
q_{i+1} \leq \left( \frac{1}{v_i - v_{i-1}} \right) X_i; \quad i = 1, \ldots, K - 1 \quad (4-6)
\]

Equation (4-4) is the resource constraint. To understand how (4-5) and (4-6) work towards the goal of approximating a production function with increasing returns to scale, notice that \( v_i \) increases with the subscript, so that \( v_i - v_{i-1} > 0 \). Therefore (4-5) implies that if \( q_i = 0 \), then \( X_i \leq 0 \). Together with the nonnegativity constraint, this forces \( X_i = 0 \). But \( X_i = 0 \) forces \( q_{i+1} = 0 \) in (4-6). Thus, the lower output activities must be used to take advantage of the increasing efficiency of higher output activities.

Both the Dantzig and Markovitz-Manne methods of linearizing decreasing costs or increasing returns to scale involve the use of integer variables. While integer and mixed-integer programming models have many desirable qualities, it should be kept in mind that the introduction of integer variables often greatly increases the size of a model, and makes the use of fast and efficient linear programming algorithms impossible.
SURVEY OF LINEAR PROGRAMMING MODELS

In this section, we will discuss models of the electric power industry in detail. It should be interesting to compare different approaches to the same or similar problems. To facilitate comparison, it is desirable to unify the notation. We will try to use the symbols and notation presented in Table 4-1. Where this is not possible, we will introduce the notation with the model.

Dispatching of Generating Equipment: Brauer (1973)

One of the key points in the pioneering study of Massé and Gibrat (1957) is that electric utilities do not produce a homogeneous good. Peak load electricity is different from base load electricity. All generating capacity can be used to produce either for peak, intermediate, or base demand. But it will generally be profitable to use different plant types for different tasks. For this reason, investment and production (dispatching) decisions must be made simultaneously. The purpose of a dispatching model is to decide how the available capacity shall be optimally used to satisfy a fluctuating demand. The time horizon is very short, not more than a day, and does not allow for any changes in capacity. Investment models either must contain at least elements of a dispatching model, or it must be given exogenously which plants produce for base, intermediate, and peak load.

Because of the interactions between the uses of the currently installed capacity and planned investments, it is useful to study a dispatching model before going on to investment models. When one uses load curves instead of load duration curves, a dispatching model can be integrated into the investment model. Therefore, we present Brauer's (1973) linear programming dispatching model.

Brauer did a study for the electric utility of West Berlin, Germany. Due to its isolated location, this system is a closed one. Reliability gains additional weight, since there is no neighboring utility to rely on in the case of an emergency. The model is a linear programming model with all variables 0,1 integers. Total operating costs are to be minimized. The time horizon is extremely short run. Brauer mentions that the program has been run up to 18 times a day to determine the optimal allocation of available resources as the conditions changed. This approach presumes that the daily
load curve is not known in advance. Only the general shape is given, especially the time of the occurrence of the peak. A numerical example is presented in the paper, applying the model to a fictitious system with two power plants.

Table 4-1

NOTATION AND SYMBOLS

Remarks: It is not easy to unify notation among different models. There are some sources from which confusion could arise. We try to prevent this by explaining differences in the text. The same symbols always have the same meaning. However, one time it may be an activity and another time it may be a parameter. Dimensions may also change.

Subscripts:

- $h$: block of approximations to load and load duration curves, $h = 1, ..., H$
- $j$: plant type/technology, $j = 1, ..., J$
- $s$: season, $s = 1, ..., S$
- $v$: vintage of plant (except in Scherer (1977))
- $-V$: age of oldest plant in system
- $x$: hydro year, $x = 1, ..., s$
- $y$: planning “year;” need not be identical to calendar year, $y = 1, ..., Y$
- $Y$: planning horizon expressed in planning “years”

Variables:

- $FC(.)$: fixed charge 0,1 integer variable associated with setup costs. Where appropriate we use $GFC(.)$ for fixed costs in generation and $IFC(.)$ for fixed investment costs
- $G(.)$: generation level (MW or MWh)
- $I(.)$: investment level (# of plants; MW or MWh)
- $K(.)$: choice of capacity level (MW or MWh)
- $KT(.)$: choice of transmission capacity level
- $ON(.)$: switching on (# of plants)
- $R(.)$: retained capital (# of plants, MW or MWh)
- $SPIN(.)$: spinning (integer number of plants)
- $STOR(.)$: activity of storing water for generation (MWh)
- $T(\alpha, \beta, .)$: transmission of electricity from $\alpha$ to $\beta$
Table 4-1 (continued)

Parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CFC(.)$</td>
<td>fixed charge or fixed costs. Where appropriate we use $CGFC(.)$ or $CIFC(.)$</td>
</tr>
<tr>
<td>$CG(.)$</td>
<td>cost of generation</td>
</tr>
<tr>
<td>$CI(.)$</td>
<td>investment cost</td>
</tr>
<tr>
<td>$CK(.)$</td>
<td>annual capacity cost for generating equipment</td>
</tr>
<tr>
<td>$CKT(.)$</td>
<td>annual capacity cost for transmission equipment</td>
</tr>
<tr>
<td>$CON(.)$</td>
<td>cost of switching on and turning off a plant</td>
</tr>
<tr>
<td>$CR(.)$</td>
<td>maintenance costs</td>
</tr>
<tr>
<td>$CSPIN(.)$</td>
<td>spinning cost</td>
</tr>
<tr>
<td>$CSTOR(.)$</td>
<td>cost of storing water</td>
</tr>
<tr>
<td>$GMIN(.)/GMAX(.)$</td>
<td>bounds on generation level</td>
</tr>
<tr>
<td>$IMIN(.)/IMAX(.)$</td>
<td>bounds on investment level</td>
</tr>
<tr>
<td>$K(.)$</td>
<td>capacity parameter</td>
</tr>
<tr>
<td>$KMIN(.)/KMAX(.)$</td>
<td>bounds on capacity decision</td>
</tr>
<tr>
<td>$LI(.)$</td>
<td>construction delivery lead time</td>
</tr>
<tr>
<td>$LS(.)$</td>
<td>warm-up lead time</td>
</tr>
<tr>
<td>$WFLOW(.)$</td>
<td>water flow</td>
</tr>
<tr>
<td>$WSTOR(.)$</td>
<td>amount of stored water</td>
</tr>
<tr>
<td>$D(.)$</td>
<td>estimate of rate of demand</td>
</tr>
<tr>
<td>$m(.)$</td>
<td>reserve requirement coefficient (usually % of $D$)</td>
</tr>
<tr>
<td>$ND(s)$</td>
<td>number of days in season</td>
</tr>
<tr>
<td>$\theta(h)$</td>
<td>duration of $h^{th}$ load level</td>
</tr>
</tbody>
</table>

One of the plants has two, the other has three units. The results from the example are not representative. Before we present the model we quickly introduce the necessary notation.

Subscripts:

<table>
<thead>
<tr>
<th>Subscript</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td>period of day, $h = 1, \ldots, H$</td>
</tr>
<tr>
<td>$j$</td>
<td>plant, $j = 1, \ldots, J$</td>
</tr>
<tr>
<td>$u$</td>
<td>unit, $u = 1, \ldots, U$</td>
</tr>
<tr>
<td>$m$</td>
<td>mode of operation, $m = 1, \ldots, M$</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>level of output, $\varphi = #$ of MW</td>
</tr>
</tbody>
</table>
Activities:

\[ G(j, u, m, \varphi, h) \]

- generation variable
  \[ = 1, \text{if unit } u \text{ in plant } j, \text{working in mode } m, \text{produces } 1 \text{ KWH of electricity in period } h \]
  \[ = 0, \text{otherwise} \]

\[ ON(j, u, m, h) \]
- turning up
  \[ = 1, \text{if in period } h, \text{unit } u \text{ in plant } j \text{ is turned up from mode } m - 1 \text{ to } m \]
  \[ = 0, \text{otherwise turning down} \]

\[ OFF(j, u, m, h) \]
- turning down
  \[ = 1, \text{if in period } h, \text{unit } u \text{ in plant } j \text{ is turned down from mode } m \text{ to mode } m - 1 \]
  \[ = 0, \text{otherwise} \]

Parameters:

\[ \theta(h) \]
- length of period \( h \)

\[ D(h) \]
- total output required from the system in period \( h \) (MW)

\[ DMIN(j, h) \]
- minimum output required of plant \( j \) in period \( h \) (MW)

\[ \varepsilon(j, u, h) \]
- readiness parameter
  \[ = 1, \text{if in period } h \text{ unit } u \text{ in plant } j \text{ is available for generating electricity} \]
  \[ = 0, \text{otherwise} \]

\[ \beta(j, u, m, h) \]
- state parameter
  \[ = 1, \text{if in period } h \text{ unit } u \text{ in plant } j \text{ is available for generating electricity} \]
  \[ = 0, \text{otherwise} \]

\[ CG(j, u, m, \varphi) \]
- cost of generating \( \varphi \) units of output from unit \( u \) in plant \( j \), working in mode \( m \) ($/hr)

\[ CON(j, u, m) \]
- cost of turning unit \( u \) in plant \( j \) up from mode \( m - 1 \) to \( m \), ($)

\[ COFF(j, u, m) \]
- cost of turning unit \( u \) in plant \( j \) down from mode \( m \) to \( m - 1 \), ($)

\[ r \]
- reserve requirement (% of total energy required)

The objective function given in equation (4-7) consists of two terms and gives the total operating cost for period \( h \). The first one shows the total costs of generating a given output. The second one gives the costs of changing the modes of production of some generators.

\[
\min_{G_{ON, OFF}} \sum_{j=1}^{J} \sum_{u=1}^{U} \sum_{m=1}^{M} \left[ \sum_{\varphi=\varphi(Min)}^{\varphi(Max)} \theta(h) CG(j, u, m, \varphi) G(j, u, m, \varphi) + \theta(h) CON(j, u, m) ON(j, u, m, h) + COFF(j, u, m) OFF(j, u, m, h) \right]
\]

The first constraint is a demand constraint requiring that the total output of the system is equal
to required output.

\[
\phi \sum_{j=1}^{J} \sum_{u=1}^{U} \sum_{m=1}^{M} \phi^{(\text{Max})} G(j,u,m,\varphi) = D(h)
\]  \hspace{1cm} (4-8)

Every plant must produce at least a predetermined minimum output of electrical energy.

\[
\phi \sum_{u=1}^{U} \sum_{m=1}^{M} \phi^{(\text{Max})} G(j,u,m,\varphi) \geq D \text{MIN}(j,h)
\]  \hspace{1cm} (4-9)

If a generator is not ready for production, it cannot be used

\[
\phi \sum_{m=1}^{M} \phi^{(\text{Max})} G(j,u,m,\varphi) \leq \varepsilon(j,u,h); \quad j = 1,\ldots,J; \quad u = 1,\ldots,U
\]  \hspace{1cm} (4-10)

Recall that \(\varepsilon(\cdot)\) is a 0,1 parameter. When a unit is not available for production because it is scheduled for maintenance, or for any other reason, \(\varepsilon(\cdot) = 0\), and constraint (4-10) makes sure that no production activity is assigned to that unit.

We need a constraint to ensure that \(ON(.)\) and \(OFF(.)\) are both 0 if we do not change the mode of production, or that only one is equal to 1 if we do.

\[
\left[ \phi \sum_{n=m}^{M} \phi^{(\text{Max})} G(j,u,n,\varphi, h) \right] - ON(j,u,m,h) + OFF(j,u,m,h) = \beta(j,u,m,h); \quad j = 1,\ldots,J; \quad u = 1,\ldots,U; \quad m = 1,\ldots,M
\]  \hspace{1cm} (4-11)

To understand the working of these last constraints, notice that constraint (4-10) allows each unit only one mode of operation and output level at any given time. If at the beginning of period \(h\) the generator in question is operating in mode \(m\), and \(G(j,u,m,\varphi,h) = 1\), then both \(ON(j,u,m,h)\) and \(OFF(j,u,m,h)\) must be zero and there are no adjustment costs. Conversely, if the exogenous parameter \(\beta(j,u,m,h) = 0\), either the generator has not been operated during the last period and will also not be operated during the current period, in which case no adjustment is necessary, or the generator has been operated during the last period, and the mode must be changed. Then the constraint forces one of the adjustment variables to be equal to 1, since \(1 - 1 = 0\).

We have already pointed out that reliability is of great concern in this study. The next constraint
imposes reserve requirements on the system. Constraint (4-12) requires that if any one unit breaks down, there must be enough capacity ready to make up for that loss at a moment's notice. The loss of a unit is not allowed to be made up at the expense of the "regular" reserve margin \( rD(h) \).

\[
\sum_{j=1}^{J} \sum_{u=1}^{U} \sum_{m=1}^{M} \phi(\text{Max}) \left( \varphi(\text{Max}) - \varphi \right) G(j, u, n, \varphi, h) \geq rD(h) + \sum_{\varphi=\varphi(\text{Min})} \phi G(j, u, n, \varphi, h);
\]

\( j = 1, \ldots, J; \quad \bar{u} = 1, \ldots, U; \quad \bar{m} = 1, \ldots, M \)  

The left-hand side of (4-12) represents the total idle capacity ready to produce instantly, and the right-hand side shows the total reserve requirement. This formulation has two interesting properties. In the introduction, we discussed the problems of reliability that arise as larger and larger units/plants are introduced into a system; (4-12) penalizes larger units by requiring larger standby reserves. Brauer's approach to reliability also distinguishes between idle capacity and idle capacity ready to produce at a moment's notice. The existence of idle capacity cannot prevent a loss of load if the start-up time is too long.

As formulated, the model is designed to minimize total costs from period to period, without looking ahead. If we have information about the expected loads in each period, the daily total costs can be lowered by planning for more than one period. To do this, \( \beta(j, u, m, h) \) is replaced by a binary \((0,1)\) variable \( B(j, u, m, h) \). The interpretation of \( B(\cdot) \) is the same as for \( \beta(\cdot) \), except that \( B(j, u, m, h), h \neq 1 \), is determined by the model. \( B(j, u, m, 1) \) is still given exogenously. To determine \( B(\cdot) \) from period to period, additional constraints are necessary.

\[
B(j, u, m, h) + ON(j, u, m, h) - OFF(j, u, m, h) - B(j, u, m, h + 1) = 0;
\]

\( j = 1, \ldots, J; \quad \bar{u} = 1, \ldots, U; \quad m = 1, \ldots, M; \quad h = 1, \ldots, H - 1 \)  

Recall that \( B(j, u, m, h) \) is 1 if unit \( u \) in plant \( j \) is ready to produce at mode \( m \) at the beginning of period \( h \). Therefore, \( B(j, u, m, h) \) and \( ON(j, u, m, h) \) cannot both be equal to 1. There are other logical impossibilities so that only four combinations of \( B(\cdot, h), ON(\cdot, h), OFF(\cdot, h), \) and \( B(\cdot, h + 1) \) have to be considered. They are summarized in Table 4-2.
### Table 4-2

Possible Combinations of $B(., h), ON(., h), OFF(., h),$ and $B(., h + 1)$ in (4-13)

<table>
<thead>
<tr>
<th>Unit read at the beginning of period $h$</th>
<th>At the beginning of period $h$, unit is turned up</th>
<th>At the beginning of period $h$, unit is turned down</th>
<th>Unit ready at the beginning of period $h + 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes: 1 / No: 0</td>
<td>Yes: 1 / No: 0</td>
<td>Yes: 1 / No: 0</td>
<td>Yes: 1 / No: 0</td>
</tr>
<tr>
<td>0 =</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 =</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0 =</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 =</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Brauer (1973), Table 4, page 433

Obviously, Brauer's model is highly detailed, designed for very short-term dispatching scheduling. All variables are binary, so none of the efficiencies of linear programming may be employed in its solution. The number of variables is also very large, further complicating its solution. The solution to this program yields a schedule of generation levels, and turning up and turning down activities for each unit of each plant. Transmission is not considered in this model, nor any investment decisions for the expansion or capacity or reduction of costs through changes in the portfolio of generating units.

The stage is now set for the investigation of investment models. We shall start with the seminal article by Massé and Gibrat (1957) of Electricité de France. The influence of this paper will easily be recognized in later studies.

**Linear Programming Investment Model: Massé and Gibrat (1957)**

The electric power industry in France is nationalized, investment decisions having to be approved by the government. In the fifties, Electricité de France did not obtain the permission to build a reservoir for a hydroelectric power plant in the French Alps because the cost per kilowatt hour in an average year was estimated to be about three times higher than for a hydroelectric plant without reservoir. “It is clear to the discerning that this difference in cost is accompanied by a difference in value – for the energy accumulated in the reservoirs can render much more service than can the fleeting energy of rivers. But this difference in value, due to the flexibility in exploitation of reservoirs, was insufficiently understood.” (Massé and Gibrat, 1957, page 149).

The authors felt it would be possible to demonstrate the benefits of the added flexibility of reservoirs over river plants by using a linear programming model. In the model, five plant types
were considered: steam plants, hydroelectric plants without reservoir, with a small reservoir, with a large reservoir, and tidal plants. The purpose of the research was to demonstrate that even if an unlimited quantity of “good” river plants and “good” steam plants (“good” meaning cheap) could be constructed, it would still be optimal to have some reservoir and tidal plants, when total discounted costs had to be minimized. Two objective functions had been considered: maximization of profits given all prices, and minimization of costs subject to given demand requirements. Cost minimization was chosen because of the monopoly position of Electricité de France in the market for electric power, while, on the other hand, this company was considered too small to significantly affect the prices of the inputs.

$C_G(j)$ is the total capitalized unit cost of operation of plant $j$, and $CI(j)$ is the total unit investment cost. The model is static in the sense that only one period is considered. During this period, it is possible to vary capacity arbitrarily. There are no construction lead times. To impose capacity restrictions, three demand types are introduced: base load, $A_0(MW)$, peak load, $B_0(MW)$, and total annual demand, $C_0(MWh)$. For each plant type, the relative contributions to these demands per MW of capacity installed, $a(j)$, $b(j)$, $c(j)$, was predetermined.

The problem can be stated as follows,

$$\text{MIN} \sum_{j=1}^{5} [CI(j) + CG(j)]K(j),$$  \hspace{1cm} (4-14)

where $K(j)$ is the is the nonlumpy installed capacity of plant type $j$, and $CG(j)$ includes all operating and maintenance costs. The demand constraints are straight-forward.

$$\sum_{j=1}^{5} a(j)K(j) \geq A_0; \sum_{j=1}^{5} b(j)K(j) \geq B_0; \sum_{j=1}^{5} c(j)K(j) \geq C_0$$  \hspace{1cm} (4-15)

Since $a(j)$, $b(j)$, $c(j)$ are predetermined, the dispatching problem has already been implicitly solved. The optimal plant mix does not change the optimal use of each plant type – a very strong assumption.

Equations (4-14) and (4-15) together form a complete investment model. In the empirical part of their study, Massé and Gibrat included a financial constraint as well. Investment expenditures could be no larger than a given amount, CIMAX.
\[ \sum_{j=1}^{5} CI(j)K(j) \leq CIMAX \] (4-16)

The question of uncertainty received a lot of attention in the introduction. “In the matter of electric production, the uncertainty of the future affects at once the supply and the demand; it is necessary to take account of the joint variability of the consumption, of the water levels and of machine availability” (Massé and Gibrat, 1957, page 150). The authors considered the reliability of plants, availability of water and inputs (fuel or others), and cost of investment. The problem of uncertainty is handled through the introduction of safety margins; 5 percent is added to demand estimates. Production of a plant is evaluated for the middle of the winter, when the annual demand peak occurs, while the water levels reach their minimum. The output from steam plants is assumed to be 15 percent less than the expected output. Investment cost estimates are increased over and above what was thought to be reasonable. The same procedure is still used in most linear programming models of the electric industry.

The data for the empirical part of the paper comes from studies done within Electricité de France. By the time the paper was written, it was already obsolete and was used only to illustrate the workings of the model. The financial constraint proved to be of crucial importance to the solution. With investment funds below 195 billion francs, the demand requirements cannot be met. As the investment funds are increased, the present value of total costs decreases from 43 billion \((CIMAX = 195\text{ billion francs})\) to 375 billion \((CIMAX = 621\text{ billion francs})\). Above 621 billion francs, the financial constraint is no longer binding. The plant mix changes with the investment expenditures. At 195 billion, only steam plants are built. As available funds are increased, tidal plants are introduced first, and then hydro plants with large reservoirs. Above investments of 227 billion francs, no steam plants will be built and at 360 billion francs, only tidal plants will be used in the system. Above that amount, tidal plants will be partially replaced by river plants. From 621 billion francs on, no further changes take place; the system uses only river and tidal plants. The results show that the trade-off between investment and operating costs is critically affected by the availability of funds.
Notice that this model is very simple compared to Brauer's. All variables are continuous, allowing the linear programming formulation. Demand is expressed as a three-step load duration curve, eliminating the necessity to deal with time of day considerations of load management – warm-up lead times, capacity range, and fixed operations costs. In fact, this model does not deal with numbers of generators, but with output capacities of different generator types. Thus, important constraints on hour-to-hour operations and the engineering design of electricity generating equipment play out a small role in the investment decision. Neither is there jumpiness of investment nor increasing returns to scale.

Since the work done by Massé and Gibrat, linear programming has been widely used to model the electric industry. In what follows, we will distinguish five main types of models. Investments models in the tradition of Massé and Gibrat (1957) will be investigated first. Models that allow for increasing returns to scale will then be treated separately. Then Scherer's (1977) marginal cost model will be presented. The estimation of marginal costs is closely related to the question of cost minimization. Third, we will look at a model that focuses on the effects of environmental constraints. The fourth model type deals with the introduction of new technology into an existing system of plants. In all but one case, the new technology is completely known. The models intend to shed light on the role the new technology will play in the future. In the model presented by Manne (1974), the date of the commercial availability of nuclear breeder plants is assumed to be a random variable. The last model considers the loss of load probability (LOLP) as a criterion for holding reserves, instead of fixed reserve margins.

**Models in the Tradition of Massé and Gibrat (1957): a) Non-increasing Returns to Scale**

McNamara (1976). To evaluate different investment projects, a great amount of information has to be analyzed. This difficult task is made even harder by many interdependencies existing between different variables. McNamara sees the purpose of his model not in predicting optimal future capacities, but to determine the major implications of changes in trends. He designed the model with the planning departments in electric utilities as potential users.

The model is a single region capacity expansion model. Transmission costs are assumed to be fixed. This allows one to neglect spatial aspects. Minimization of the present value of total costs is the
objective of the program. The demand side is represented by a load duration curve, which is roughly approximated by two blocks (Figure 4-6).

![Figure 4-6. Approximation of Load Duration Curve.](image)

Source: McNamara (1976), Figure 1, page 229.

The approximation must be chosen such that the load factor is the same as for the true load duration curve. In this example, the load factor ($L.F.$) is given by

$$L.F. = \frac{\text{Area under the curve}}{\text{Peak demand} \times 100\%} = \frac{4000 \text{ MWh} \times 0.15 + 2000 \text{ MWh} \times 0.85}{4000 \text{ MWh} \times 1.00} = 0.575.$$

The denominator shows the energy required if demand were uniformly at the peak level, while the numerator represents the amount of energy needed over the whole period and is equal to the area under the load duration curve. The planning horizon is ten years, divided into two planning periods ($y = 1, 2$) containing five years each ($t = 1, 2, \ldots, 5$). Three generating technologies are considered ($j = 1, 2, 3$): nuclear, fossil-steam, and combustion turbine. Total costs are made up of investment and operating costs. Operating costs, denoted $CG(\cdot)$, include all costs necessary to maintain and operate a plant. (4-17) represents the objective function.

$$\min_{\{K(j, y, h)\}} \sum_{j=1}^{3} \sum_{y=1}^{2} \sum_{h=1}^{2} \left( CI(j, y) + CG(j, y, h)K(j, y, h) \right)$$

(4-17)
\( K(\cdot) \) is the installed capacity. Suppose a plant has a nameplate capacity of 1000 MW and is available 80 percent of the time. The effective capacity in this case is 800 MW. \( CI(\cdot) \) has the dimension (\$/effective MW/planning period). \( CG(\cdot) \) has the dimension (\$/MW/planning period), and its value changes as we move along the load duration curve. McNamara assumes that both cost components increase exponentially with time.

The demand constraints are easy to interpret. Effective capacity has to be sufficient to meet demand. Demand is also assumed to increase exponentially with time.

\[
\sum_{j=1}^{3} \sum_{y=1}^{y^*} \sum_{h=1}^{h^*} K(j,y,h) \geq D(y^*, h^*); y^* = 1, 2; h^* = 1, 2
\]  

(4-18)

Let \( D(\cdot, 1), D(\cdot, 2) \) denote base and peak load respectively. Then (4-18) says that installed effective base load capacity has to be sufficient to satisfy base load demand, and the sum of effective base and peak load capacity must be large enough to meet peak demand.

The utility has inherited capacity. This is expressed by constraints (4-19).

\[
\sum_{h=1}^{3} K(j,1,h) = K(j,1); j = 1, 2, 3
\]  

(4-19)

Notice that it is left to the model to determine what proportions of the inherited capacity are to be used for peak and base load, respectively.

What are some of the properties of this formulation? The main concerns, as outlined in the introduction, are, scale economies, uncertainty, load versus load distribution curves, and the objective or objectives of the decision makers. Most of these concerns are dealt with in a traditional way. Because the model is continuous and linear, there can be no economies of scale. Uncertainty on the demand side can be handled using safety margins. Uncertainty about the availability of installed capacity is approached by using effective or expected capacities. Other kinds of uncertainty cannot be introduced without major changes. An approximation to the load duration curve is used to model the demand side. Finally, the objective is cost minimization, as in most other studies.
It should be noted that the variable $K(\cdot)$ stands for capacity and generation, at the same time. This implicitly assumes that a plant of type $j$ is always used to the same degree during load period $h$. Capacity is distinguished into peak load and base load capacity. Peak load is defined as the demand over and above base load. The total required capacity of plants of type $j$ is equal to $\sum_{h=1}^{2} K(j, y, h)$, that is, the sum of peak and base load capacity. By using the same variable for capacity and generation, the size of the model can be kept small. If this approach would result in fixed load factors for the different plant types, the realism of the model would have to be questioned. But since the allocation of total capacity to peak and base capacity can change over time, the load factor can vary. Therefore, it is possible to obtain the result that old capacity moves up on the load curve. Thus, when a plant first comes on line, it is the most efficient technology available, and is therefore used for base load generation. As even more efficient technologies become available, its use would shift toward intermediate and peak load, until its economic lifetime is over.

Using this model, an empirical analysis has been conducted. No source is given for the data used. The figures appear to be realistic for the time of the study. It is assumed that fuel and investment costs increase at the same rate for all plant types and fuels. Initially, only conventional steam and combustion turbines are available. Nuclear capacity is introduced mainly to satisfy base load demand. Long lead times of construction are accounted for by making investments in the first years very expensive.

The numerical example underlines the importance of the tradeoff between operating and investment costs for the choice of equipment. This confirms the result obtained by Massé and Gibrat (1957). It would be interesting to see if the addition of a financial constraint would produce results similar to those of Massé and Gibrat.

McNamara briefly outlines the extension of his model to a two region model, neglecting non-linearities in transmission costs. The discussion of transmission is discussed below.

Anderson (1972); Turvey and Anderson (1977). Anderson (1972) provides an excellent survey of cost minimization models of the electric industry, with emphasis on linear programming models. His study is reprinted with minor changes as Chapter 13 in Turvey and Anderson (1977).
Electrification is an important condition for modern economic development. In Chapter 8, Turvey and Anderson present a linear programming model for investment planning in Turkey. In that country the demand for electricity has been growing very rapidly in the past and shall continue for some years to come. The model draws from Anderson (1972). The work was supported by the World Bank, with the overall goal of developing models that can help to design “investment policies in the developing countries which could promote economic efficiency ... and in providing service to low-income groups in rural and urban areas” (Anderson, 1977, page xvi).

The planning horizon of this study is 35 years. The model uses five investment periods of 5 years each, and one end condition of 10 years. On the demand side, a block wise approximation to the load duration curve is being used. Four demand levels are considered: peak demand, two intermediate levels, and base load. In the empirical part of the study, the median demand for each period and demand block is used. Future supplies will come from a mix of plants using different technologies. Turkey still has considerable untapped hydro- electric sites. It can also make use of its lignite deposits to support steam-electric plants. All together nineteen plant types are considered, among them, thirteen different hydro plants. The remaining six are gas turbines, oil burning steam plants, three lignite burning plant types, and nuclear power stations. Demand, cost, and, and technical data had been supplied by the Turkish Electricity Authority.

In the tradition of Massé and Gibrat (1957), minimization of total discounted costs is the objective. The objective function consists of two parts: present value of total investment costs, and total discounted operation costs. Operating costs include expenditures for fuel, maintenance, etc. The first item in (4-20) only includes new plants. Old plants only contribute to operating and maintenance costs.

\[
\text{MIN} \sum_{j=1}^{I} \sum_{v=1}^{V} \sum_{h=1}^{H} CI(j,v)I(j,v) + \sum_{j=1}^{I} \sum_{y=V}^{V} \sum_{h=1}^{H} CG(j,y,v,h)G(j,y,v,h)\theta(h) \tag{4-20}
\]

\(\theta(h)\) is the duration of demand level \(h\). Plants are distinguished not only by type, but also by vintage. \(-V\) is the age of the oldest plant in the system, and investment and operating costs, \(\{CI(\cdot)\} \text{ and } CG(\cdot)\},\) vary with the vintage, \(v\), of each plant.
Installed capacity $K(j, y) = \sum_{v=-V}^{y} I(j, v)$. The first constraint requires that total available capacity is sufficient to meet demand plus a reserve requirement to provide for unforeseen events.

$$\sum_{j=1}^{J} \sum_{v=-V}^{y} a(j,v) I(j,v) \geq (1+r) D(y,h); y=1,...,Y; h = peak$$  \hspace{1cm} (4-21)

At any time, total output has to be sufficient to satisfy demand.

$$\sum_{j=1}^{J} \sum_{v=-V}^{y} G(j,y,v,h) \geq D(y,h); y=1,...,Y; h = 1,...,H$$  \hspace{1cm} (4-22)

For each plant type, the output cannot exceed effective capacity. Effective capacity is less than installed capacity, to account for maintenance and forced outage.

$$G(j,y,v,h) \leq a(j,v) I(j,v); j=1,...,J; y=1,...,Y; v=-V,...,y; h = 1,...,H$$  \hspace{1cm} (4-23)

The next two sets of constraints limit the use of hydro plants. Define $b(j)$ to be the load factor, that is, the ratio of average and maximum output. (4-24) ensures that hydro plants do not require more water than what is available.

$$\sum_{h=1}^{H} G(j,y,v,h) \theta(h) \leq b(j) I(j,v); j=hydro; y=1,...,Y; v=-V,...,y$$  \hspace{1cm} (4-24)

During the dry seasons of the year, the total output from hydro plants is limited by the storage capacity. If total demand is above that capacity, non-hydro backup plants are needed. As a result, the amount of hydro power capacity in the system has to be limited. It would have been desirable to have the model determine this limit. However, the available seasonal data did not permit this. The limit is expressed as a fraction, $f$, of total peak demand. This fraction was arrived at from studies of water flows and storage capacities. It was found to be around 50 percent for Turkey.

$$\sum_{j=1}^{J} \sum_{v=-V}^{y} I(j,v) \leq fD(y,h); j=hydro; y=1,...,Y; h = peak$$  \hspace{1cm} (4-25)

Total capacity installed in a given technology may be limited by the availability of resources. Hydro power is limited by the number and characteristics of available sites. Denote this limit by $KMAX(j)$ for technology $j$. 
\[
\sum_{v=1}^{V} I(j, v) \leq KMAX(j); \quad j = 1, \ldots, J
\] (4-26)

For technical reasons, it may be necessary that investment in a plant of type \( j \) be no less than \( IMIN(j, v) \) and no bigger than \( IMAX(j, v) \). Such a constraint can be written in the following form:

\[
IMIN(j, v) \leq I(j, v) \leq IMAX(j, v)
\] (4-27)

A similar constraint could also be written for generation. The problem with (4-27) is that it requires a minimum level of investment, even if it would be most cost efficient not to use technology \( j \) at all. A constraint which imposes that \( I(j, v) \) must either be zero or greater than \( IMIN(j, v) \), would be more satisfactory, but would require the use of integer variables.

The empirical results of the study yield the optimal capacity expansion program the plant operating schedules, and total discounted costs. The dual variables to the constraint set (4-22) have the interpretation of marginal costs. Not unexpectedly, the optimal mix for Turkey is a balance of hydro and lignite plants during the first ten years. In later years, nuclear and oil burning power plants are introduced into the system. The time of the transition will depend on demand growth and relative price changes. Lignite initially provides the bulk of the base load. It is gradually replaced by nuclear power. Oil fired plants provide peak load energy. Hydro power plants provide over the range from intermediate to peak output, depending on the load factor.

Turvey and Anderson handle long run uncertainty on the demand side in the traditional way of including a reserve margin over and above expected future peak demand levels. Other uncertainties are not dealt with explicitly.

In the present application, uncertainty about water availability is an important factor on the supply side. It would be desirable to include these aspects explicitly in the model. Such an expansion is proposed by Turvey and Anderson (1977) in Chapter 13. The water inflows, \( WFLOWS(\cdot) \), change with the seasons, \( s \). The objective function (4-20) has to be changed to include storage capital costs.
The constraints to the revised problem are in principle the same as the ones presented above. We add the restriction that water stored at the end of a season, expressed in MWh, plus the amount used for generation during that season, cannot exceed the water inflow during the season, plus the amount stored at the beginning of the season.

\[
\begin{align*}
\text{MIN}_{I,\text{STOR},G} & \left\{ \sum_{j=1}^{J} \sum_{v=1}^{V} CI(j,v)I(j,v) + \sum_{j=\text{hydro}}^{J} \sum_{v=1}^{V} CSTOR(j,v)STOR(j,v) \right. \\
+ & \left. \sum_{j=1}^{J} \sum_{v=1}^{V} \sum_{y=V}^{Y} \sum_{s=1}^{S} \sum_{h=1}^{H} CG(j,y,v,s,h)G(j,y,v,s,h)\theta(h) \right\} \\
\end{align*}
\]

\[(4-20')\]

Over long planning horizons, it may be unrealistic to neglect retirement of plants. Especially if the focus is on the introduction of new technologies, it is interesting to see whether old plants will eventually be replaced. We alter the objective function (4-20) to include retirement.

\[
\begin{align*}
WSTOR(j,y,v,s+1) + \sum_{h=1}^{H} G(j,y,v,s,h) \leq WSTOR(j,y,v,s) \\
+ WFLOW(j,y,v,s); s = 1, ..., S-1; h = \text{hydro} \\
WSTOR(j,y,v,1) + \sum_{h=1}^{H} G(j,y,v,S,h) \leq WSTOR(j,y,v,S) \\
+ WFLOW(j,y,v,S) \\
\end{align*}
\]

\[(4-28')\]

In (4-20), \(CG(\cdot)\) includes maintenance costs. Here, maintenance costs, \(CR(\cdot)\), are separate. Capacity has also to be redefined, \(K(j,y) = \sum_{v=V}^{Y} R(j,y,v)\). Therefore, constraints (4-21) have to be rewritten:

\[
\sum_{j=1}^{J} \sum_{y=V}^{Y} R(j,y,V) \geq (1+r)D(y,h); y = 1, ..., Y; h = \text{peak} \\
\]

\[(4-21'')\]

It is also required that retained capacity of a type cannot exceed the capacity of that type and vintage initially installed.
\[ R(j, y, v) \leq I(j, v); j = 1, ..., J; y = 1, ..., Y; v = 1, ..., y \]  \hspace{1cm} (4-28'')

Also, the capacity of the remaining plants of type \( j \) and vintage \( v \) can never increase.

\[ R(j, y + 1, v) \leq R(j, y, v); j = 1, ..., J; y = 1, ..., Y; v = 1, ..., y \]  \hspace{1cm} (4-29''')

Constraints (4-22) - (4-27) still hold. The original problem has successfully been transformed into one that allows retirement of capital. The formulation implies that there is no cost to scratch a plant.

As easily as the original model changed to include water storage or retirement of plants, transmission of electricity could be included. In Scherer (1977), transmission is an integral part of the model. His treatment is very similar to that suggested by Turvey and Anderson (1977). Therefore, this topic is postponed until Scherer’s model is discussed.

With each expansion of the model, its size increases. This is the trade-off between realism and simplicity. In mathematical programming, the model size is limited by the computational ability of the computer. Anderson (1972) tried to solve this problem by making use of the load duration curve. Along the load duration curve, demand decreases. Therefore, the assumption is made that the required output from any plant does not increase along the load duration curve. Define a new variable, which Anderson calls \( Z(j, y, v, h) \).

\[
Z(j, y, v, h) = G(j, y, v, h) - G(j, y, v, h + 1) \geq 0
\]

\[
h = 1, 2, ..., H - 1
\]

\[
Z(j, y, v, H) = G(j, y, v, H) \geq 0
\]  \hspace{1cm} (4-30''')

Replace the constraint set (4-23) by a new set of restrictions.

\[
\sum_{h=1}^{H} Z(j, y, v, h) \leq a(j, v) I(j, v); j = 1, ..., J; y = 1, ..., Y; v = 1, ..., y
\]  \hspace{1cm} (4-23''')

(4-30''') simply says that along the load duration curve, output never increases. (4-23''') is (4-23) written in different form, because
\[ \sum_{h=1}^{H} Z(j, y, v, h) = \sum_{h=1}^{H} (G(j, y, v, h) - G(j, y, v, h+1)) = G(j, y, v, 1). \]

Since \( G(j, y, v, 1) \geq G(j, y, v, h) \) for all \( h \) (by assumption), we can decrease the number of constraints in (4-23) by the factor \( H \). The procedure is simple. Assumption (4-30′′′) may not always be justified, however. The annual peak demand may coincide with the most severe meteorological conditions for the operation of hydroelectric plants. In such a case, using \( Z(\cdot) \) substitutes is not feasible.

Anderson and Turvey (1977) applied their model to Turkey. It could also be applied to a single utility. The California Energy Resources Conservation and Development Commission has obtained a version of the model and a description of a computer program needed to apply it is available (Thompson and Graham, 1977).

Remarks. The models presented so far do not allow increasing returns to scale. All of them have the objective of minimizing total discounted costs. Their features are representative of the state-of-the-art in this area. The paper by Cavalieri et al. (1971) is a special case. It considers the optimal operation of a single thermal plant, when fuel use has to be minimized. 0,1 integers are used to guarantee that a unit cannot operate below a minimum level. The next step is to minimize total expected costs associated with the expansion of the plant. The possible capacity extension involved the evaluation of twenty additional combinations of boilers and turbines. It is an interesting paper because it stresses engineering aspects more than other models. However, because of its limited applications it shall not be reviewed here in depth.

Models in the Tradition of Massé and Gibrat (1957): b) Increasing Returns to Scale

Gately (1970). Gately addressed three main questions: What type of plants should be built? How large should the plants be? When should they be built? The model was developed for the State of Madras, India. In spite of the large size of the planning area, roughly one-fourth the size of France, the spatial dimensions of the model are neglected.

The problems of plant size and timing are related to the question of returns to scale. If there are constant returns to scale, plants will be built when demand requires them. When economies of
scale can be realized, it may be more efficient to construct a large plant, even if this means that it will not be fully utilized immediately (Chenery, 1952). Gately approximates returns to scale using the fixed charge approach discussed above. For this purpose, he has to introduce 0,1 integer variables.

The planning horizon is 15 years, divided into 5 planning periods of 3 years each \((y = 1,2,3,4,5)\). Four generating technologies are considered: hydroelectric power, conventional thermal, lignite, and nuclear power generation. Demand requirements are represented by a load duration curve which is approximated by three blocks \((h = 1,2,3)\). \(h = 1\) is non-peak, \(h = 2\) is normal-peak, \(h = 3\) is super peak demand. Since hydroelectric power is an important source of energy in the State of Madras, three different hydrological years are distinguished, and a probability of occurrence is assigned to each. The subscript \(x\) is used for the hydro-years. Within a year, \(s = 1,2\) stands for the dry and the wet season.

The realistic inclusion of investment into hydroelectric capacity causes some difficulties not usually encountered with other technologies. While a large degree of standardization can be assumed for the various types of thermal plants (including nuclear), the unit costs of hydro projects vary widely among sites. To overcome these difficulties Gately enumerates all hydro projects and represents them by 0,1 integer variables. Investments in hydroelectric power plants are only undertaken during the first three periods.

The objective is again to minimize total discounted costs. The objective function consists of three parts: construction costs, fixed operating and maintenance costs, and variable operating costs. The discounting procedure enters the objective function explicitly. Hydroelectric plants take one period to be completed. The construction costs are divided equally between period \(y\), when the project is started, and period \(y + 1\), when it is completed. The interest rate used for discounting is 10 percent. \(P(x)\) is the probability that hydro-year \(x\) occurs.
\[ \begin{align*}
\text{MIN} & \quad I_G \\
& \quad \sum_{j=\text{hydro}}^{3} \sum_{y=1}^{2} \frac{1}{2} \frac{1}{1.10^{3y-1}} + \frac{1}{1.10^{3y+2}} CI(j, y) I(j, y) \\
& \quad \sum_{j=2,3}^{4} \sum_{y=1}^{2} \frac{1}{1.10^{3y-1}} (CIFC(j, y) IFC(j, y) + CI(j, y) K(j, y)) \\
& \quad \sum_{j=\text{hydro}}^{3} \sum_{y=2}^{y'=1}^{y'=2} \frac{1}{1.10^{3y-1}} CGFC(j, y') K(j, y') I(j, y') \\
& \quad \sum_{j=2,3}^{4} \sum_{y=2}^{y'=2} \sum_{y'=3}^{y'=4} \frac{1}{1.10^{3y-1}} CGFC(j, y') K(j, y') \\
& \quad \sum_{j=1}^{5} \sum_{y=1}^{3} \sum_{x=1}^{2} \sum_{s=1}^{3} \frac{1}{1.10^{3y-1}} CG(j, y) G(j, y, x, s, h) P(x) \\
& \quad \sum_{j=1}^{5} \sum_{y=1}^{3} \sum_{x=1}^{2} \sum_{s=1}^{3} \frac{1}{1.10^{3y-1}} CG(j, y) G(j, 5, x, s, h) P(x)
\end{align*} \] (4-31)

\(j = 1\) stands for conventional thermal plants. Notice that the formulation prevents any plant of this type from being built. The first two terms give us the total discounted investment costs. \(I(j = \text{hydro}, y)\) is an integer variable and \(K(j = \text{hydro}, y)\) is a parameter, while \(K(j \neq \text{hydro}, y)\) are continuous variables. The next two terms express the fixed costs of operation and maintenance. \(CGFC(j = \text{hydro}, y')\) is the total fixed charge ($/plant), while \(CGFC(j \neq \text{hydro}, y')\) has the dimension ($/MW). The last two terms represent total variable operation costs.

The model contains seven constraint sets. One of them is the non-negativity requirement which we do not list explicitly. In each period it is required that generation is sufficient to satisfy demand.

\[ \sum_{j=1}^{5} G(j, y, x, h, s) \geq D(y, x, s); y = 1, 2, 3, 4; x = 1, 2, 3; h = 1, 2, 3; s = 1, 2 \] (4-32)

The next two constraint sets restrict capacity. No more power can be produced than what the installed capacity allows.
\[
\sum_{h=1}^{3} G(j, y, x, h, s) \leq \sum_{y' = 1}^{y = 1} K(j, y'), \text{ with } K(j, 0) = \text{ inherited capacity}
\]

\[
\sum_{y = 1, 2, 3, 4, 5} G(j, y, x, h, s) \leq \sum_{j = \text{hydro}} \left( \sum_{y' = 1}^{y = 1} K(j, y', s) I(j, y') + K(j, 0, s) \right)
\]

\[
y = 1, 2, 3, 4, 5; x = 1, 2, 3; s = 1, 2
\]  

(4-33)

To prevent confusion, notice that \(K(j, y)\) as used in the objective function is the nameplate capacity of each hydro plant. \(K(j, y, s)\) is the actual capacity available in season \(s\), expressed in MW. In the next constraint set, we are requiring that not more hydroelectric energy may be produced than compatible with the size of the reservoirs. \(K(j, y, x, s)\) is the maximum capacity (in MWh) of a reservoir.

\[
\sum_{j = \text{hydro}} \sum_{h = 1}^{3} G(j, y, x, h, s) \leq \sum_{j = \text{hydro}} \left( \sum_{y' = 1}^{y = 1} K(j, y', x, s) I(j, y') + K(j, 0, x, s) \right)
\]

\[
y = 1, 2, 3, 4; x = 1, 2, 3; s = 1, 2
\]  

(4-34)

For lignite and nuclear plants, the fixed-charge variable \(FC(\cdot)\) should be positive if and only if the corresponding \(K(\cdot)\) is positive, zero otherwise. For mathematical reasons, an upper limit of 5000 MW is imposed on the capacity of lignite and nuclear plants. This limit is chosen large enough that it is not binding.

\[
5000FC(j, y) - K(j, y) \geq 0; j = 2, 3 \text{ (lignite and nuclear); } y = 1, 2, 3, 4
\]  

(4-35)

In the case of hydroelectric projects, we have to prevent the same site from being used more than once.

\[
\sum_{y = 1}^{3} I(j, y) \geq 1; j = \text{hydro}
\]  

(4-36)

Finally, recall that

\[
I(j, y) = 0 \text{ or } 1; j = \text{hydro}; IFC(j, y) = 0 \text{ or } 1; j = 2, 3
\]  

(4-37)
The complete model has 237 rows, 629 columns, and 2774 matrix entries. Of the variables, 38 are of the 0,1 type, the other are continuous variables. Gately remarks that it took about 2 minutes to solve the model on an IBM 360/85 computer. A branch-and-bound algorithm was used to obtain a solution.

The results from this study show that conventional thermal plants will be used as little as possible, due to high operating costs. They are generating only during the dry season. Nuclear plants are used exclusively for base load generation. Hydro power plants operate at full capacity over the peak and intermediate loads. The remaining capacity is used to contribute to base load demand. Lignite plants carry the burden caused by the great variability of hydroelectric power output. Their load varies inversely with that of hydro plants. The investment schedule reflects the desirable properties of hydroelectric power. During the first period, the construction of a total of 425 MW of hydroelectric capacity is started. In addition, a 215 MW lignite plant is being built. In the second three-year period, only one hydroelectric plant of 5.0 MW capacity is constructed. In the third period, an 890 MW lignite plant is constructed, and in the fourth period, another lignite plant of 700 MW plus a 540 MW nuclear power plant are planned.

Demand forecasts over a 15-year horizon incorporate a large degree of uncertainty. To find out how variations in future demand loads affect the investment and operation schedules, Gately performed runs under different demand growth assumptions. In the original version, demand grows 33 percent per period. The two alternatives considered were growth rates of 24 percent and 42 percent, respectively. There were changes in the size and the timing of investment. The basic pattern was not disturbed, though. Under the assumption of slower demand growth, the same hydroelectric plants were constructed, but some of them a period later. The lignite and nuclear plants were planned at a smaller scale. Faster demand growth caused more hydroelectric capacity to be developed at an earlier stage, and nuclear and lignite plant sizes were increased. Both results confirm the comparative advantage of hydroelectric power. Nuclear and especially lignite plants are built to supplement the hydroelectric power stations.
Variations in the discount rate have a more pronounced effect. While an increase from 10 percent to 12.5 percent does not cause a change in the pattern of investment, at 15 percent investment shifts away from nuclear and hydro plants to lignite plants. This reflects the increased importance of construction over operating costs at high costs of capital.

Cherene and Schaeffer (1979). The model proposed by Cherene and Schaeffer integrates dispatching and the capacity expansion problems explicitly. As long as the load duration curves are used to represent the demand side, optimal operation planning must remain rudimentary. In Cherene and Schaeffer, the load curve of the “engineering design day” is used. Different curves can be used for different seasons, so seasonality of demand can be easily included.

Demand is assumed to vary in a predictable and identical fashion for each day of the season, numbering $ND(s)$ for season $s$. In the presence of high storing costs of the output, the utilities will look for alternatives to varying generation rates. They could vary the number of plants in operation, which is efficient for plants that are quickly and cheaply started up and shut down. An alternative is to use many plants, and vary their production rate. This is economical for plants with long start-up times and a wide range of efficient production. The importance of production flexibility is the major aspect of electricity generation focused upon by this model.

The model is formulated as a mixed-integer linear program. There are five activities; investment, maintenance, turning a plant on, letting a plant spin, and generation. Revenue requirements are to be minimized. The objective function consists of five terms; annualized investment costs, maintenance costs, turning on and spinning costs, and generation costs. All costs are expressed in present value. Some costs vary “hour” by “hour” due to the non-uniformity of the chronological length of the demand “hour” over the load curve.
The first set of constraints deals with the flexibility of each plant. A plant can spin at a given hour if and only if it was either spinning throughout the previous hour or has been warming up for the requisite warm-up lead time, \(LS(j, y, s, h)\).

The first inequality in (4-39) is straightforward. The second one “turns the corner” at the end of the day. It describes the case where equipment must be warmed up the day before. Since it is assumed that the operating strategy is the same from day to day, a plant can only be spun in the morning if it is warmed up that night! Obviously, these constraints must be meticulously tailored for each plant type and each load curve.

The assumption of a fixed daily operating procedure implies another simplifying property: any plant that is warmed up during the day is cooled down during the same day. Thus the cost parameter \(CON(j, y, s, h)\) is, without loss of generality, the sum of warm-up and cool-down costs.

The sum of the output of all plants has to be sufficient to meet demand.

\[
\sum_{j=1}^{J} G(j, y, s, h) \geq D(y, s, h); y = 1,...,Y; s = 1,...,S; h = 1,...,H
\] (4-40)
ready to generate electricity at a moment's notice is large enough to make up for the loss of the largest unit in the system, without endangering the regular reserve margin. This formulation requires more reserve capacity the larger the size of the largest plant. By doing so it captures the cost of decreased reliability of large plants.

\[
\sum_{j=1}^{J} \left[ SPIN(j, y, s, h)K(j) - G(j, y, s, h) \right] + \sum_{j \in M} \left[ R(j, y) - SPIN(j, y, s, h) \right] \\
\geq MD(y, s, h) + 2K^*; y = 1, ..., Y; s = 1, ..., S; h = 1, ..., H
\] (4-41)

The first term on the left-hand side denotes the excess capacity of the turned on plants. The second term adds the capacity of those plants that are not turned on presently, but can start operating almost instantly. \( M \) is the set of these plants. \( K^* \) represents the capacity of the largest plant in the system.

Electricity generation cannot exceed the capacity ready to produce. \( SPIN \) is the only variable required to be of the integer type. The dimension is (\# of plants).

\[
G(j, y, s, h) \leq SPIN(j, y, s, h)K(j); \forall j, y, s, h
\] (4-42)

For technical reasons, generation may have to be zero or above a given minimum level, say \( KL(j) \).

\[
G(j, y, s, h) \geq SPIN(j, y, s, h)KL(j); \forall j, y, s, h
\] (4-43)

Of course, one cannot spin more plants than one has at a given time:

\[
SPIN(j, y, s, h) \leq R(j, y); \forall j, y, s, h
\] (4-44)

Notice that the activity of maintaining plant-type \( j \) in year \( y \) has its costs as expressed in the objective functional. Thus no more plants will be maintained than are necessary for current or future utilization. But utilization is integer valued, so a least cost strategy will only retain an integer number of plants.

The number of plants maintained is bounded by the number built:
\[ \begin{align*}
R(j, y) &\leq R(j, y-1) + I^*(j, y); y < LI(j) \\
R(j, y) &\leq R(j, y-1) + I(j, y); y \geq LI(j)
\end{align*} \]  \tag{4-45}

\( I^*(j, y) \) denotes inherited investment plans. Since \( y \) is less than the construction lead time, \( LI(j) \), these projects are already in progress. The inequality indicates that any of these projects may be cancelled should other cost considerations deem it optimal, for it allowed the possibility of retiring plants, thus saving maintenance costs.

This model incorporates several new features. Most important among them is the use of load curves in a capacity expansion model. This allows a comparison of plants not only with respect to heat rates, capital and maintenance costs, but also with regards to flexibility as described by warm up lead times and range of output. One will notice the influence of the studies of Brauer (1973) and Anderson (1972), and Turvey and Anderson (1977) on the model.

The price to be paid for the innovations is in terms of the model size. Suppose there are ten technologies \( (J = 10) \) and a ten “years” planning horizon \( (Y = 10) \), six seasons in each planning “year” \( (S = 6) \), and each load curve is approximated by four blocks \( (H = 4) \). In this case the model has the following number of variables:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G(j, y, s, h) )</td>
<td>2400</td>
</tr>
<tr>
<td>( I(j, y) )</td>
<td>100</td>
</tr>
<tr>
<td>( ON(j, y, s, h) )</td>
<td>2400</td>
</tr>
<tr>
<td>( R(j, y) )</td>
<td>100</td>
</tr>
<tr>
<td>Continuous Variables</td>
<td>5000</td>
</tr>
<tr>
<td>( SPIN(j, y, s, h) )</td>
<td>2400</td>
</tr>
<tr>
<td>Total Activity Variables</td>
<td>7400</td>
</tr>
</tbody>
</table>

The number of constraints would be 10,180. To arrive at a manageable size, the model must be simplified. For example, if \( S = 2 \), then the total number of variables reduces to 2400. To date, the model has not been used for empirical analysis.
One feature of the model is worth stressing. By making SPIN an integer variable, \( I \) and \( R \) are also forced to be integers. Because of this, economies of scale can be handled. A plant of type \( j \) is characterized by its technology and capacity.

Like others, Cherene and Schaeffer (1979) are able to deal only with some of the problems discussed in the introduction. Scale economies can be handled, though the model quickly becomes very large. With respect to uncertainty, a little progress has been made. Large plants require a larger reserve capacity. This reflects their poorer reliability. Otherwise, the treatment is comparable to that in earlier studies.

de la Garza, Manne, and Valencia (1973). Another model making use of integer variables to capture the indivisibility of investments, has been developed by these three authors. Their study is part of a development plan for Mexico. Theirs is a spatial model with eight regions. Eight plant types and two kinds of transmission lines are considered. Demand is represented by a peak and an off-peak level for each of the three seasons.

The constraints are straightforward. There are capacity and demand constraints for each region. For inherited plants, peak power capacity is a datum. Fossil and nuclear plants cannot be operated below a minimum level. Hydroelectric output is bounded from above. Also, there are restrictions concerning reservoir limits. Transmission is handled in a fashion very similar to Scherer’s (1977) model. We postpone the discussion until we review Scherer’s contribution.

The data for the empirical part of the study come from the Mexican Comision Federal de Electricidad. Regionalization of the model was important. The annual peak for the country as a whole occurs during the time from November to February. The three northern regions of Torreon-Chihuahua, Falcon-Monterrey, and Mazatlan-Sonora constitute an exception to the national trend. Irrigation and air conditioning loads shift the annual peak into the season lasting from July to October. Once transmission links are established between north and south, the July-October peak demand constraints are no longer binding.
Marginal Cost Models (Scherer (1977)).

Traditionally, electric utilities based their prices on average costs. In recent years, the interest shifted towards marginal cost pricing. The research on peak load pricing underlines this. Scherer's intention was to empirically estimate marginal costs. The customer buys his electricity from the system, not from a particular plant. Therefore, he should be charged the cost he causes the system as a whole. Marginal costs depend on the particular time-of-day (year) and the location of the customer. In order to get these results, Scherer uses a spatial model. Since capacity is allowed to vary, the marginal costs will include both capital and operating costs. In other words, the ex-ante marginal costs are estimated. Scherer defines system average cost as the total annual cost divided by the actual amount of energy produced. The dimension is ($/KWh). Of course, this figure varies with the load factor. Hence, Scherer keeps the load factor constant as system output is increased. This is achieved by maintaining a constant ratio between loads in each demand period and at each load center.

Operating costs arise as a direct consequence of generation. Capital is consumed slowly over the useful lifetime of a plant. In order to arrive at meaningful cost figures, operating and capital costs must have the same dimension. Scherer solves this problem by using the capital recovery factor (CRF). The CRF, multiplied by the original investment costs, gives us the amount that would have to be paid each year in order to pay for the equipment during its lifetime. The total annual cost of a plant is equal to this figure plus the operation costs.

With the exception of Massé and Gibrat (1957), the models discussed so far are dynamic in the sense that capacity can be expanded sequentially. In these models, an increase in capacity from one period to the next requires investment to expand the stock of equipment. Like Massé and Gibrat, Scherer proposes a static model where capacity is continuously variable. In this fashion the long run cost curve for a given output can be found. Only one year is considered. The demand side is represented by a load duration curve. Scherer discusses the possibility of using several seasonal load duration curves. He does not use the idea in his work due to computational constraints. The load duration curve is approximated by demand blocks.
The objective of the model is to minimize total system’s cost. Operating costs are assumed to be linear. For new transmission lines, new steam turbine-generator plants, and existing plants, there is a fixed charge with an associated 0,1 variable. This binary variable also makes it possible to consider the retirement of equipment.

Transmission costs do not enter the objective function directly. They influence the optimal solution indirectly, by requiring a larger output as transmission losses increase. Transmission losses are nonlinear, increasing at an increasing rate. The loss curve can be approximated by a piecewise linear curve. This is illustrated in Figure 4-7.

![Figure 4-7. Transmission Loss Curve for Transmission from α to β.](image)

Source: Scherer (1977), Figure 4-4, page 80

The objective function can be written as the sum of the total generation curve capital costs, and transmission costs $CK(j, \alpha)$ denotes the total annual capacity cost per MW of a plant of type $j$ at location $\alpha$, except for the fixed charge, where applicable.

Units in an existing plant often have significantly different heat rates. Therefore, the cost curve has to be represented by a piecewise linear function with increasing slope. Subscript $\nu$ denotes the
linear segments. Note that here \( v \) does not stand for the vintage of the units, but for the number of units.

The objective function is the sum of investment operating and transmission costs.

\[
\text{MINIMIZ}E : \quad \sum_{\alpha = \text{all permissible locations}} \left[ CFC(j, \alpha)FC(j, \alpha) + CK(j, \alpha)K(j, \alpha) + \sum_{h=1}^{H} CG(j, \alpha, h)G(j, \alpha, h) \right];
\]

\( j \) = new steam turbine-generator plants

\[
+ \sum_{\alpha = \text{all permissible locations}} \left[ CK(j, \alpha)K(j, \alpha) + \sum_{h=1}^{H} CG(j, \alpha, h)G(j, \alpha, h) \right]; \quad j = \text{gas turbines}
\]

\[
+ \sum_{\alpha = \text{all permissible locations}} \left[ CFC(j, \alpha)FC(j, \alpha) + \sum_{h=1}^{H} \sum_{\nu=1}^{V} CG(j, \alpha, h, \nu)G(j, \alpha, h, \nu) \right];
\]

\( j \) = existing plants

\[
+ \sum_{\alpha = \text{all permissible locations}} \left[ CK(j, \alpha)K(j, \alpha) + \sum_{h=1}^{H} CG(j, \alpha, h)G(j, \alpha, h) \right]; \quad j = \text{pumped storage}
\]

\[
+ \sum_{\alpha} \sum_{\beta} CKT(\alpha, \beta)KT(\alpha, \beta); \text{all locations with existing transmission lines}
\]

\[
+ \sum_{\alpha} \sum_{\beta} \left[ CFCT(\alpha, \beta)FCT(\alpha, \beta) + CKT(\alpha, \beta)KT(\alpha, \beta) \right];
\]

all locations w/o existing transmission lines

There are no operation costs associated with transmission, and no capital costs with existing plants. Retirement of existing plants is assumed to be costless.

The nonlinearity of the cost curve of new steam generators is approximated using the fixed-charge approach. This approximation is good only over a range of plant sizes. The first constraint puts an upper bound on the size of new steam plants.

\[
K(j, \alpha) \leq FC(j, \alpha)KMAX(j) \quad (4-47)
\]

\[
FC(j, \alpha) = 0 \text{ or } 1 \quad (4-48)
\]
\[ G(j, \alpha, h) \leq K(j, \alpha) \] (4-49)

\[ j = \text{new steam turbine generator plants}; \]
\[ \alpha = \text{all permissible locations}; h = 1, \ldots, H. \]

(4-49) simply states that actual output cannot exceed the capacity of a plant. A constraint equivalent to (4-49) also holds for new gas turbines. No other restrictions are imposed on the construction and use of this technology.

\[ G(j, \alpha, h) \leq K(j, \alpha); j = \text{new gas tubines}; \alpha = \text{all permissible locations}; h = 1, \ldots, H \] (4-50)

The total output from an existing plant is obtained by summing over all units, \( V \). Total output may not exceed the nameplate capacity of a plant. The output from each unit may not exceed the capacity rating of that unit.

\[ \sum_{v=1}^{V} G(j, \alpha, h, v) \leq FC(j, \alpha) K(j, \alpha) \] (4-51)

\[ FC(j, \alpha) = 0 \text{ or } 1 \] (4-52)

\[ G(j, \alpha, h, v) \leq GMAX(j, \alpha, h, v) \] (4-53)

\[ j = \text{existing plants}; \alpha = \text{all permissible locations}; h = 1, \ldots, H; v = 1, \ldots, V \]

Since \( CG(j, \alpha, h, 1) < CG(j, \alpha, h, 2) \), etc., an optimal program will choose \( G(j, \alpha, h, v) \) such that the most efficient unit is used to capacity first, then the second one, etc.

The capacity of pumped storage at any location may not exceed a certain size, as dictated by site characteristics. The first constraint expresses this restriction. Equation (4-55) restricts the amount of electricity generated in any demand period to less than the capacity installed. Pumped storage plants consume electricity during off peak hours and return it during peak hours. A pumped storage plant cannot produce more electricity than it has previously stored. Constraint (4-56) expresses this relationship and also takes into account the loss of 1/3 of the electricity stored that occurs in the process.

\[ K(j, \alpha) \leq KMAX(j, \alpha) \] (4-54)
\[ G(j, \alpha, h) \leq K(j, \alpha) \]  \hspace{1cm} (4-55)

\[
\left(\frac{2}{3}\right) \sum_{h=\text{hours when electricity is consumed}} \theta(h) G(j, \alpha, h) - \sum_{h=\text{hours when electricity is generated}} \theta(h) G(j, \alpha, h) = 0
\]  \hspace{1cm} (4-56)

\( j = \text{pumped storage}; \alpha = \text{all permissible locations}; h = 1, ..., H \)

Recall that \( T(\alpha, \beta, h, u) \) is the amount of electricity transmitted from \( \alpha \) to \( \beta \) during demand period \( h \). The subscript \( u \) allows us to approximate the nonlinear relationship between the amount transmitted and the losses, as shown in Figure 4-7. In an optimizing model, during the same demand period, electricity will only flow in one direction. This allows us to write the capacity constraint in the following form.

\[
\sum_{u=1}^{U} T(\alpha, \beta, h, u) + \sum_{u=1}^{U} T(\beta, \alpha, h, u) \leq KT(\alpha, \beta)
\]  \hspace{1cm} (4-57)

Notice that \( KT(\alpha, \beta) = KT(\beta, \alpha) \). The next constraint is straightforward and follows from our discussion of Figure 4-7.

\[ T(\alpha, \beta, h, u) \leq TMAX(\alpha, \beta, h, u) \]  \hspace{1cm} (4-58)

For (4-57) and (4-58): all \( \alpha, \beta; h = 1, ..., H; u = 1, ..., U \). New transmission lines cannot exceed a predetermined maximum size.

\[ K(\alpha, \beta) \leq FCT(\alpha, \beta) KMAX(\alpha, \beta) \]  \hspace{1cm} (4-59)

\[ FTC(\alpha, \beta) = 0 \text{ or } 1 \]  \hspace{1cm} (4-60)

for all \( \alpha, \beta \) with new transmission lines.

There are several demand constraints that have to be satisfied. Demand at each location has to be satisfied. This is the case if generation at location \( \alpha \), plus the amount of energy received from adjacent nodes, minus the amount of energy sent to adjacent nodes, is sufficient to satisfy demand in \( \alpha \).
\[
G(\alpha, \beta) + \sum_{b=\text{all adjacent models}}^{U} \sum_{u=1}^{T} T(\alpha, \beta, h, u) - \sum_{b=\text{all adjacent models}}^{U} \sum_{u=1}^{T} (1 + TLOSS(\alpha, \beta, u)) T(\alpha, \beta, h, u) = D(\alpha, h)
\]  

(4-61)

for all \( \alpha; h = 1, \ldots, H \)

At locations with pumped storage we must also make sure that demand plus the energy required by that plant will always be met.

\[
G(\alpha, \beta) + \sum_{b=\text{all adjacent models}}^{U} \sum_{u=1}^{T} T(\alpha, \beta, h, u) - \sum_{b=\text{all adjacent models}}^{U} \sum_{u=1}^{T} (1 + TLOSS(\alpha, \beta, u)) T(\alpha, \beta, h, u) = D(\alpha, h)
\]  

(4-62a)

all \( \alpha \) with pumped storage plants; \( h \) = hours when pumped storage plants generate electricity.

\[
-G(\alpha, \beta) + \sum_{b=\text{all adjacent models}}^{U} \sum_{u=1}^{T} T(\alpha, \beta, h, u) - \sum_{b=\text{all adjacent models}}^{U} \sum_{u=1}^{T} (1 + TLOSS(\alpha, \beta, u)) T(\alpha, \beta, h, u) = D(\alpha, h)
\]  

(4-61)

all \( \alpha \) with pumped storage plants; \( h \) = hours when pumped storage plants consume electricity.

Equations (4-62a,b) are equivalent to (4-61), except that (4-62b) allows storage of energy through pumping, while (4-62a) is the case when the energy is recovered.

The reserve requirements demand that total installed capacity minus transmission losses is large enough to satisfy peak demand plus a prespecified amount of reserves, \( RES \). \( RES \) is the capacity of the largest plant with matched units.
\[
\sum_{j=1}^{J} \sum_{a \in \alpha} K(j, \alpha) - \sum_{a \in \alpha} \sum_{\beta \in \beta} \sum_{u=1}^{U} TLOSS(\alpha, \beta, u) T(\alpha, \beta, h, u) \geq \sum_{a \in \alpha} D(\alpha, h) + RE\$; h = \text{peak}
\]

The model was applied to the control area of the New York State Electric and Gas Corporation. Data on inherited capacity and loads came from that corporation. Cost figures were extracted from statistical works published by the Federal Power Commission, except for transmission. For transmission Scherer relied on information obtained from the transmission engineers of the corporation.

Scherer is able to derive cost curves from his study. The results are summarized in Chapter 7 of his book. He includes environmental constraints in his model. Pollution standards show a significant effect on the cost of using coal fired steam plants.

The matrix of the empirical model contained approximately 160 rows by 450 variables, including slack variables. On an IBM 360/145 computer the model, with six integer variables, took about three minutes per solution. Scherer used the IBM MPSX-MIP package. An increase in the number of integer variables from three to ten increased computer time to four minutes. On an IBM 360/19 computer, solutions were obtained about ten times as fast.

**Capacity Expansion and the Environment. Thompson, Calloway, and Nawalanic (1977).**

The production of electricity has a strong impact on the environment. Most of the concerns are with regards to the influence of generating facilities on air and water quality. As a consequence, a set of environmental regulations faces the electric utilities, and contributes to the increasing costs of capacity expansion and operation. Thompson, Calloway, and Nawalanic attempt to analyze the effect of environmental controls on the optimal plant mix.

Environmental constraints can be met in different ways. Old “dirty” plants can be scrapped, or their output limited. Low sulfur coal can be used in place of high sulfur coal. The production process may be changed, cleaning devices can be installed. In summary, there exist several alternative ways to achieve the same goal. Process analysis is a tool well suited to this kind of problem. Three basic
inputs are needed for electricity production: fuel, water, and capital. Besides electrical energy, there are several other, undesired, outputs. Among them are sulfur oxides, heat, and ash. The relative outputs can be varied by changing the input and technology mix. This interaction is graphically represented in Figure 4-8.

![Figure 4-8. Input-Output Structure of the Electric Power Generation Model](source)

The goal of this study is in many ways similar to that of Scherer (1977); to see the effect of environmental controls and constraints on resource availability, marginal costs are derived. The objective of the linear programming formulation is the minimization of total costs. There are three basic plant types: hydro-electric, nuclear, and conventional thermal plants. Coal gasification is considered as a possibility. Conventional thermal plants can be subdivided according to the fuel they burn: natural gas, medium-sulfur coal, low-sulfur coal, and high and low-sulfur oil. Plants can also be distinguished by the kind of environmental control they use, so that there really is a great variety of plant types.
The model is static; only one planning period is considered. All new capacity will come from non-nuclear steam plants. New nuclear plants are not included because it was felt that their characteristics could not be estimated reliably enough to give meaningful results. All variables are continuous, avoiding issues of increasing returns to scale or lumpy investment.

The notation in Thompson et al. (1977) is quite different from that of the models discussed so far. Technical factors play the key role in their model. Instead of writing out all relationships, we present the model in a concise form in Table 4-2, which has been taken from the appendix of Thompson et al. (1977). The constraints will be discussed, as long as they are not self-explanatory. This should give a sufficiently good idea about the characteristics of the model.

The first set of constraints limits the availability of resources. This can include capital constraints. Demand must always be met. There is a minimum output requirement for electricity from new thermal plants. The resource balance constraints are self-explanatory as are the primary residual balance equations. The last constraint set requires that discharge of pollutants does not exceed a tolerated maximum. Finally, the objective function consists of the costs of exogenous supplies, the costs of production, costs of treatment of the residuals, and effluent charges. An optimal solution is a mix of plants and technologies which minimizes total costs.

Most of the data for the empirical application came from the Federal Power Commission. The results show that the plant mix is highly sensitive to environmental constraints. As the discharge of pollutants becomes more restricted, investment into base load capacity shifts away from high-sulfur coal burning plants to coal gasification combined cycle power plants. Constraints on capital availability are also important. The results confirm the observations made by Massé and Gibrat (1957) twenty years earlier that there is a trade-off between capital and operating costs. Demand is estimated and treated in a simplified manner. Plants are categorized into peak, mid-range, and base load plants, according to the number of hours they are operated in one year, base load plants generating electricity during at least 5000 hours a year, mid-range plants between 1000 to 5000 hours, and peak plants for 1000 hours or less. Different plant types were prespecified as either peak, mid-range, or base load plants. It would be possible to relax these assumptions and use a load
duration curve, allowing the model to assign the plants to different loads.

Table 4-2
Mathematical Description of the Thompson, Calloway, Nawalanic Electric Power Model

<table>
<thead>
<tr>
<th>Exogenous Resources</th>
<th>Exogenous Supplies $X_1, ..., X_S$</th>
<th>Exogenous Supplies $P_1, ..., P_J$</th>
<th>Exogenous Supplies $T_1, ..., T_L$</th>
<th>Exogenous Supplies $D_1, ..., D_V$</th>
<th>Right-Hand Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exogenous Supplies</td>
<td>$Y_i$</td>
<td>$Y_j$</td>
<td>$e_i, ..., e_j$</td>
<td>$K_i$</td>
<td>$K_j$</td>
</tr>
<tr>
<td>Electricity Demand</td>
<td>$Q_i$</td>
<td>$Q_j$</td>
<td>$E_i$</td>
<td>$E_j$</td>
<td>$F_i$</td>
</tr>
<tr>
<td>Demand Resource</td>
<td>$E_i$</td>
<td>$E_j$</td>
<td>$E_i$</td>
<td>$E_j$</td>
<td>$E_i$</td>
</tr>
<tr>
<td>Resource Material</td>
<td>$M_i$</td>
<td>$M_j$</td>
<td>$M_i$</td>
<td>$M_j$</td>
<td>$M_i$</td>
</tr>
<tr>
<td>Balances</td>
<td>$M_i$</td>
<td>$M_j$</td>
<td>$M_i$</td>
<td>$M_j$</td>
<td>$M_i$</td>
</tr>
<tr>
<td>Primary Residual</td>
<td>$r_i$</td>
<td>$r_j$</td>
<td>$a_i$</td>
<td>$a_j$</td>
<td>$a_i$</td>
</tr>
<tr>
<td>Residual Balances</td>
<td>$r_i$</td>
<td>$r_j$</td>
<td>$a_i$</td>
<td>$a_j$</td>
<td>$a_i$</td>
</tr>
<tr>
<td>Residual Discharge</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
</tbody>
</table>

Source: Thompson, Callaway, Nawalanic (1977, page 87)

In Thompson et al. (1978), the electricity model is put in the context of the energy sector in general. The model developed by Thompson, Calloway, Nawalanic, and their collaborators is interesting because of the careful discussion of the techniques and their influence on the environment. Nowhere else are the production processes included as explicitly as here.

Until recently, nuclear power stations were not widely used. According to Scott (1976, p. 24), in 1969, only one percent of the U.S. electric energy came from nuclear plants. This figure increased to 4.5 percent in 1973, and further increases are expected. When a new technology becomes commercially available, the utilities have to decide what role this new technology is going to play in the system. In the case of nuclear power, it was self-evident that this technology would not provide economical peaking capacity. But it still was conceivable that nuclear power could provide relatively cheap intermediate capacity. The papers by Deonigi (1971), Deonigi and Engel (1973), Frankowski (1971), de Boer, Leclercq, and de Haan (1971), and Pozar and Udovicic (1971) investigate the role of new nuclear plants. Some of the properties of the models may also be applicable to other cases.

Deonigi (1971) and Deonigi and Engel (1973). Deonigi did his research for the Atomic Energy Commission (AEC). The goal was to develop a simulation model “of the U.S. power economy which is capable both of evaluating the role of nuclear power and fossil fuel power and of determining the proper course for developing reactor systems that could improve resource utilization. This simulation is based on linear programming techniques and is readily adaptable to represent any geographic region or set of national constraints” (Deonigi, 1971, p. 521). Deonigi (1971) contains a verbal discussion of the model and its results. We rely on the version presented in Deonigi and Engel (1973).

As mentioned, the main purpose is the evaluation of the role of nuclear power. While various nuclear technologies are modelled in detail, conventional thermal plants are treated as one single technology. Eight different constraint sets are introduced, mostly technical in nature and applying to nuclear plants. Instead of adopting the notation developed in earlier sections, it seems more convenient to use the notation of the original article. This is reasonable since the nature of this model is different enough from those of the models above, to make a direct comparison largely irrelevant. We now introduce all the symbols that will occur in the objective function.
Amount of material (isotopes) $k$ bought externally at the beginning of period $i$

Amount of material $k$ sold externally at the beginning of period $i$

Balance of material $k$ bought at the end of period $i$

The selling, buying, and storage coefficients, respectively, for material $k$ in period $i$

Amount of fuel purchased in period $i$ at the $j^{th}$ price step

The $j^{th}$ price step for fuel

Amount of fuel processing in period $i$

Unit variable cost of fuel processing

Set of all plants built in time interval $i$

Total discounted cost coefficient of plant $r_i$

The cost coefficient $C_{r_i}^t$ of plant $r$ includes investment and operating costs. To be able to do this, the lifetime of each plant has to be known and is assumed to be 14 years. The extent to which the plant capacity is being used in each period is also specified, though it would be preferable to let the model solve for the information. Deonigi and Engel report that in an earlier version this was the case. The procedure used up a lot of computer time without yielding significantly different results for the cases considered. Therefore, the model was changed. Now the user specifies the load factor.

The objective function has four components: the net cost of buying and storing materials (isotopes) for nuclear plants, the total cost of fuels used in production, the cost of fuel processing, and the costs of investment and operation, as far as they have not already been captured by other components.

\[
MIN \sum_{b,s,JU,W,r} \left\{ \sum_{k=1}^{K} \left( p_k^b b_k^i - p_k^s s_k^i + p_k^r I_k^i \right) + \sum_{j=1}^{J} p_j^u U_{j,i} + C_j^i W_i + \sum_{r_j \in R_j} C_{r_j}^t r_j \right\}
\]

(4-64)

$P$ is the total number of time intervals considered, $J$ is the number of uranium price steps, and $K$ is the number of isotope stockpiles. It is assumed that fuel is priced along a step function, with only limited amounts available at each price. The price increasing with the amount used. The demand power constraints consist of two equalities for each load history $i$ and each period $j$. By load history is meant the utilization of a plant in each period. The first equation calculates the total nuclear capacity with load history $i$ added in period $j$, while the second constraint splits total electric power generating capacity among nuclear and other plants.
We now present the notation and the constraints.

\( N_{ijk} \)  Total capacity of nuclear plants with load history \( i \), installed in region \( k \), during time period \( j \) (gigawatts)

\( K \)  Number of regions

\( R_{ij} \)  Set of all nuclear plants with load history \( i \), that may be installed in time period \( j \)

\( c_\gamma \)  Capacity (gigawatt) of a power plant; used for both fossil or nuclear power plants

\( F_{ijk} \)  Set of fossil plants with load history \( i \), available in region \( k \) during time period \( j \)

\( T_{ijk} \)  An input constant representing the total capacity (gigawatts) of new plants with load history \( i \), installed in region \( k \), during time period \( j \)

\( r_\gamma \)  Nuclear reactor \( \gamma \)

\( f_\gamma \)  Fossil plant \( \gamma \)

The constraint looks as follows.

\[
\sum_{k=1}^{K} N_{ijk} = \sum_{c_\gamma \in R_{ij}} c_\gamma f_{c_\gamma} \quad \forall i, j \\
N_{ijk} + \sum_{f_\gamma \in F_{ijk}} c_\gamma f_{f_\gamma} = T_{ijk} \quad \forall i, j, k
\]

Unfortunately, the article is not very clear at this point. We presume that if you take the sum of \( T_{ijk} \) over \( i \) and add to this the capacity already available, we get a capacity big enough to satisfy demand in region \( k \) in period \( j \). However, this is our interpretation and does not follow from the article.

The fuel constraints allow for preprocessing of fuels, and limit the amount that can be bought at a given price. The price for uranium is \( P_{ji}^{uu} \). The first constraint says that the total amount of fuel purchased in period \( j \) is equal to the deficit in the stockpile, plus the stock of preprocessed fuel at the end of period \( j \), minus the stock of preprocessed fuel brought out from last period, plus the total amount of fuel required for production. The second constraint makes sure that a solution to the program does not require more fuel at a given price than is available.

\( M \)  Number of price steps

\( U_{ij} \)  Amount of fuel purchased in time interval \( j \) at price \( i \)

\( D_j \)  Deficit in the fuel processing residual stockpile

\( W_j \)  Preprocessed fuel stockpile at the end of period \( j \)

\( R_j \)  Set of all reactors that affect requirements in time interval \( j \)
Amount of fuel required by $r_{\gamma}$ in period $j$

Number of time intervals

Total amount of fuel available at price $i$

For $j = 1, 2, \ldots, P$

$$\sum_{j=1}^{M} U_{ij} = D_{j} + 1.34 W_{j} - 1.34 W_{j-1} + \sum_{r_{\gamma} \in R_{j}} r_{\gamma} u_{\gamma}$$

and for $i = 1, 2, \ldots, M = 1$

$$\sum_{j=1}^{P} U_{ij} \leq A_{i}$$

The 1.34 coefficient expresses the fuel requirement (in kg) for the fuel treatment process, at an assumed four percent enrichment.

The third type of restrictions are the material balance constraints. For each isotope stockpile in each interval we need one equation. Its interpretation is straightforward – the material balance for a given isotope at the end of interval $j$ is equal to the previous balance multiplied by a decay constant, plus purchases from outside, minus sales to the cubicle, plus the net material discharge by the nuclear power plants.

Material balance at the end of time interval $j$

Decay constant for the length of the time period

Set of all reactors that charge or discharge the material in time period $j$

Amount of material output, r.s.p., sold outside the system at the end of time interval $j$

Net amount of material discharged by reactor $r_{\gamma}$ in period $j$

$$I_{j} = dI_{j-1} + b_{j} - s_{j} + \sum_{r_{\gamma} \in R_{j}} a_{\gamma} r_{\gamma}$$

The fuel processing residual constraints are optional in this model. They can be of interest when a by-product of a fuel processing plant can be used as fuel by some power plants. For each time period we require one constraint. The equation says that the residual stockpile minus the deficit in the residual stockpile at the end of period $j_{r}$ are equal to the residual stockpile at the end of the previous period plus the residual discharge from the reactors.
Material balance at the end of time interval $j$

Deficit in residual stockpile at the end of the $j^{th}$ time interval. This variable enters the fuel usage constraint for the $j^{th}$ time interval as a fuel requirement.

Set of all reactors that affect the residual stockpile in time period $j$

Amount of residual discharge by reactor $r_\gamma$ in period $j$

The constraint for each period $j$ is then

$$T_j - D_j = T_{j-1} + \sum_{r_\gamma \in R_j} t_\gamma r_\gamma$$  \hspace{1cm} (4-68)

There are two fuel processing constraints for each time period $j$. The first says the change in the stock of preprocessed fuel is equal to the amount processed minus the amount used up by the power plants, during a period. The second is a capacity constraint on the quantity of fuel that can be processed at a given time.

Stockpile of preprocessed fuel at the end of the $j^{th}$ time interval. These variables enter the fuel usage constraints.

Amount of fuel processing performed in the $j^{th}$ time interval.

Set of all reactors requiring or getting credit for fuel processing during time period $j$. A plant gets a credit if the plant fuel is not totally consumed, such as nuclear fuel.

Net fuel processing require by reactor $r_\gamma$.

Total amount of processing that can be performed during period $j$.

Then,

$$W_j - W_{j-1} = S_j - \sum_{r_\gamma \in R_j} s_\gamma r_\gamma$$  \hspace{1cm} (4-69a)

and

$$S_j \leq L_j$$  \hspace{1cm} (4-69b)

The technology introduction constraints are of great interest because utilities are expected to carefully evaluate new technologies. A new procedure will first be tried out in one “experimental” plant, and only if the result is satisfactory will the introduction of the new process gain momentum. The technology adoption constraints simulate this behavior.
\( R_j \) Set of reactors in the class that can be built in time period \( j \)

\( P_j \) Set of reactors in the class that can be built prior to time period \( j \)

\( K \) Fixed growth factor

Then, for each time period \( j \),

\[
\sum_{r_x \in R_j} r_x = \sum_{r_x \in P_j} r_x + k
\]

To get some feeling for how the constraint works, suppose \( k = 2 \). In the first period only two plants may be built. In the second period we can build four plants, two plus the number built in the previous period. In the third period we can construct two, plus four for the ones built in the second period, plus two for those built in the first period, making a total of eight new plants in period three.

The present commitment constraints also have the effect of limiting the introduction of new nuclear capacity. This set of constraints fixes the number of nuclear plants built in the early period of the model to these plants that are presently committed to construction.

\( K \) Number of load histories

\( N_{kij} \) Number of nuclear plants with the \( k^{th} \) load history built in period \( j \) in region \( i \)

\( B_{ij} \) Number of nuclear plants committed in region \( i \) in period \( j \)

\[
\sum_{k=1}^{K} N_{kij} = B_{ij}
\]

Obviously we could use this constraint for all periods and constrain the nuclear capacity to be added exogenously.

The last set of constraints is introduced to simulate the unwillingness to stop the production using a given technology, and the caution a company displays when entering a new technology. As we mentioned at the beginning of the discussion of this model, the lifetime of construction capacity is assumed to be 14 years. The following symbols are being used.
For simplicity, let $t_i = 1$. Then the constraint states that the annual new construction capacity is equal to capacity growth plus the retired construction capacity,

$$C_i = \left( \frac{1}{t_i} \right) R_i - R_{i-1} + \sum_{j=(S-14)}^{(f-14)} C_j$$

In the case of $t_i = 1$, this becomes

$$C_i = R_i - R_{i-1} + C_{i-14}.$$ 

To understand the constraint, suppose $C_i = 0$. Then $R_i = R_{i-1} - \sum C_j$. The new capacity built can only decrease as fast as construction capacity expires.

One point about the model is unclear. The variables $r_Y$ are referred to as numbers of plants. This would imply that these are integer variables. However, this seems not to be the case. Since this is not a mixed-integer programming approach, it cannot exhibit increasing returns to scale.

The simulation shows the increasing role that nuclear plants are expected to play in the future of electricity generation. The planning period starts in 1970. After the year 2000, the amount of generating capacity provided by fossil plants declines. This result depends on the availability of more advanced reactors, e.g., Liquid Metal Fast Breeder. Were they not to become commercially available, fossil plants would continue to play a major role. Light water reactors will only satisfy a relatively small part of the total energy requirements.

De Boer, Leclercq, and de Haan (1971). de Boer, Leclercq, and de Haan address the same questions Deonigi (1971) and Deonigi and Engel (1973). They present two models that have been used in Europe. Both have similar features. One includes nonlinear relationships, but the other is a linear programming model and will be discussed here.
The three authors point to the problems one encounters when a new technology is introduced. Due to lack of experience with the new equipment, it is difficult to assess the true characteristics of a plant using the modern technology. Because of this the new technology will be adopted gradually. A linear program that does not take into account these considerations, will suggest a faster rate of investment in new equipment than will actually occur.

The model presented seems to have been developed by Memmert and Wellmann (n.d., as quoted in de Boer et al. (1971)) for the European Euratom organization. The authors use the familiar approximation to the load-duration curve. The length of a load period, $\theta(h)$, is expressed in percent of the time, so that $\sum_{h=1}^{U} \theta(h) = 1$. For each load period, the load factor, $LF(h)$, is defined.

$$LF(h) = \frac{\text{Average load supplied during load period } h}{\text{Peak Load}}$$

The cost function is to be minimized. $CK(j, y, h)$ is the total discounted annual cost of operating plant $j$ during the load period $h$, in year $y$.

$$MIN \sum_{j=1}^{J} \sum_{y=1}^{Y} \sum_{h=1}^{H} CK(j, y, h) K(j, y, h)$$

The total installed capacity of plant type $j$ is

$$K(j, y) = \sum_{h=1}^{H} K(j, y, h).$$

The claim is made that the program solves the problem of assigning plant types to load levels.

“From the form of this object function it is clear that in theory every power plant type, characterized by its own annual costs $CK(j, y, h)$ is admitted in every load zone. It is only through the optimization calculation that the share of a particular type in a load factor zone and hence its mean load factor in the system is determined. Hence, the distribution of the power plant types over the reserve capacity in particular (load factor = 0) is a result of the optimization” (de Boer, et al., 1971, page 515). This is an ingenious procedure that saves many variables. Of course, one has to assume that all costs are strictly linear.
The capacity constraint is the same as usual. Enough capacity has to be available to satisfy demand.

\[ \sum_{j=1}^{J} K(j, y, h) = \theta(h) D(y); y = 1, \ldots, Y; h = 1, \ldots, H \quad (4-73) \]

The interesting constraints are those which put restrictions on the introduction of new and the withdrawal of old equipment. As mentioned above, the introduction constraint simulates the behavior of a utility in the face of uncertainty about the true qualities of the new technology.

Let \( y^*(j, \varphi) \) be the year of the first introduction of the type characterized by \( \alpha(j, \varphi) = 1; \varphi = 1, \ldots, L_1 \leq J \). \( K(\varphi) \) is the initial capacity at the starting time, \( T(\varphi) \) is the doubling time.

\[ \sum_{j=1}^{J} \sum_{h=1}^{H} \alpha(j, \varphi) K(j, y, h) \leq K(\varphi) 2^{y-y^*/T(\varphi)} Z(y, \varphi) \]

\[ Z(y, \varphi) = \begin{cases} 0 & \text{for } y < y^* \\ 1 & \text{for } y \geq y^* \end{cases} \quad (4-74) \]

Unfortunately, the constraint is not well explained in the article. Particularly, the significance of the new subscript \( \varphi \) and the parameter \( \alpha \) is not clear. It still shows though that in the periods after the new technology has been introduced, a faster rate of adoption will be allowed.

The withdrawal restrictions will not be presented. The constraint requires that a plant may be retired only after a prespecified number of years. Retirement has already been discussed in connection with Anderson (1972) and Anderson and Turvey (1977). Like Deonigi and Engel (1973), a material balance equation is included in the model and will not be discussed here. No mention is made of an empirical application of the model.

Frankowski (1971). Frankowski studied the role that nuclear energy could play in Poland. The delivery of nuclear capacity is limited by the ability of the nuclear industry. Given the current expectations about demand growth for electricity, this led to the conclusion that nuclear power stations would be used for base load generation only, for the next twenty years. Frankowski calls this the early stage of development. In later stages, nuclear power is used for generation of other loads as well.
These special characteristics of the Polish situation allow the problem to be divided into two parts. First, it is determined what share nuclear power should contribute to base load generation. Then, the nuclear plants are studied alone to find an optimal mix of nuclear technologies.

The model itself presents no new elements. It is a linear programming model of the nuclear capacity of the electricity supply system. As usual, we have cost minimization, capacity constraints, and material balance constraints; The data used are estimates by the author.

Because the lifetime of the plants was not specified, a plant, once built, is available over the whole planning horizon. This leads to results that differ significantly as the planning horizon is changed. Since forecasts twenty years ahead can be assumed to be more reliable than those forty years into the future, the difference in results must be interpreted with this increasing uncertainty about the future in mind. Frankowski proposes to use the results for planning periods of twenty to forty years as indicators of the possible range within which the solutions might fall, rather than interpreting one optimal solution in a deterministic way.

Pozar and Udovicic (1971). The Socialist Republics of Slovenia and Croatia in Yugoslavia rely to a great extent on hydroelectric power. Existing hydroelectric power stations are the cheapest source of electric energy. To ensure enough capacity to satisfy a growing demand, new plants have to be built, involving choosing from among eight types of hydroelectric power and four different thermal technologies, including nuclear power plants. The planned capacity with the lowest total discounted cost is considered optimal. The costs are based on the assumptions that hydroelectric plants have a useful lifetime of fifty years and thermal plants one of twenty-five years. To make the discounted investment costs comparable it is assumed that after the first twenty-five years, a thermal plant is replaced by an identical plant. The projects are financed with loans at 6.5 percent interest, to be repaid within fifteen years. Hydroelectric plants are characterized by their maximal possible capacity or instantaneous output (MW) and the total possible annual production (MWh).

The demand requirements are modeled by an approximation of the load duration curve. The model can take account of several seasons with different load duration curves in each season. The approximated load duration curve is piecewise linear, differing from the block wise approximation.
usually encountered.

Because existing hydroelectric power plants are the cheapest energy source, they should be used to the full extent of their capacity. The new capacity is needed to provide the balance of the capacity requirements. The interesting question in this context is how the new plants, especially the thermal plants, will fit into the existing system consisting almost exclusively of hydroelectric plants.

The objective function consists of three terms. The first denotes the total discounted cost of new hydroelectric plants. The second term represents the cost of holding reserve capacity, expressed as a fraction of thermal capacity. They are chosen such that during the critical period, when hydroelectric capacity is at its lowest level, the probability of loss of load does not exceed an acceptable level. Finally, the third term denotes the variable operating and maintenance costs of thermal capacity. In the case of hydro capacity, total annual generation is equal to total capacity expressed in MWh:
The unit costs for hydro have the dimension ($/MWh). For thermal plants the dimensions are ($/MW) for capacity costs, and ($/MWh) for operating and maintenance costs. Capacity is distinguished into "constant" (load) and "variable" (non-load) capacity; $h = 1, 2$. The assumption that hydro plants are always used at full capacity fixes the load factor of hydro plants to always equal one. In the case of variable thermal power, the extent of the use has to be determined. The load factor also influences the variable unit costs. The duration of the engagement of thermal plants during the "variable" load period will be denoted by $\theta(t, s, 2); t = \text{thermal}$. As a first step,

$$\theta(t, s, 2) = \frac{1 + \alpha}{2} WD \times \theta(s, 2)$$

(4-76)

This allows generation of electrical power from thermal plants as long as

$$\sum_{j=\text{thermal}} G(j, s, 2) \leq \beta(s) K(s, 2)$$

(4-77)

Figure 4-9 should help to understand the above two relationships. $\theta(t, s, 2)$ can be revised after the program has been solved the first time. From the results, one can arrive at a new utilization figure for thermal plants. As the assumed load factor of thermal plants is changed, the unit variable cost coefficients of these plants has to be recalculated. With these new figures, the program is solved again. New load factors and cost coefficients can be calculated, etc. The iterative procedure is reported to converge rapidly. The objective function is total cost:

$$\min_{K, G} \left\{ \sum_{j=\text{new hydro}} CK(j) K(j) + m \sum_{j=\text{thermal}} CK(j) K(j) + \sum_{j=\text{thermal}} \sum_{s=1}^{S} \sum_{h=1}^{H} CG(j, s) G(j, s, h) \theta(j, s, h) \right\}$$

(4-78)

The constraint that justifies the inclusion of this model among those that deal with the introduction of new technologies is the one that requires new investments to compensate for that part of the electric energy requirements that existing hydro plants are unable to provide. Let $G(0, s, h)$ denote
the energy coming from the inherited hydroelectric stations. The constraint can be written as follows.

\[
D(s,h) - G(0,s,h) \leq \sum_{j=\text{new hydro}} G(j,s,h) + \theta(t,s,h) \sum_{j=\text{thermal}} G(j,s,h); \ s = 1, ..., S; h = 1, 2
\]  (4-79)

In order to keep the number of variables down:

\[
G(j,s,h) = \gamma(j,s,h) K(j); \ j = \text{new hydro}
\]  (4-80)

where \( \gamma(j,s,h) \) is the percentage of the total annual production delivered in season \( s \), period \( h \). This simplifying procedure works because of the assumption that hydroelectric plants, inherited and new, will always be used to capacity; second guessing the solution always simplifies the structure of the problem.

The other constraints are similar to those of models already reviewed. They are the capacity constraints, assuring that thermal generation cannot exceed installed capacity, and technical minimum generation requirements. The installation of new hydro capacity is limited by the availability of feasible sites. Finally, investment costs may not exceed a predetermined amount.

The model is static in the sense that there is only one planning period over which the optimum solution is found. In the empirical application, the length of the period was 24 seasons of one-month length each. The model yielded an optimal construction program. The results show that hydroelectric plants can compete with nuclear power stations. Coal fired plants are not economically attractive and as thermal capacity assumes a larger share, oil fired plants become attractive because of low capital costs that compensate for higher fuel costs.

(b) Introduction of a New Technology when the Date of Commercial Availability is a Random Variable: Manne (1974)

Long lead times force utilities to plan far ahead (see Table 4-3). The reliability of forecasts is inversely related to the time that lies between the date when the forecast is made and the date to which it applies. Manne's (1974) paper deals with one particular kind of uncertainty resulting from imperfect
knowledge about the future. Everything is assumed to be known, except the exact date, \( s \), when nuclear breeder becomes commercially available. Breeders are defined to be commercially available when their costs are such that they are able to compete with light water reactors (LWR). Using Manne’s notation, \( s \) is a random variable, and the utilities are able to form subjective probabilities, \( p(s) \). After a certain date, the new technology will be available with certainty. The characteristics of all plant types, including the new one, are assumed to be known.

Table 4-3
Lead Times for Selected Technologies (years)

<table>
<thead>
<tr>
<th>Technology</th>
<th>From Purchase Order to Delivery</th>
<th>From Planning Stage to Delivery (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear (LWR)</td>
<td>10</td>
<td>15.5</td>
</tr>
<tr>
<td>Coal</td>
<td>7</td>
<td>11.5</td>
</tr>
<tr>
<td>Conventional Oil-fired</td>
<td>6.5</td>
<td>11</td>
</tr>
<tr>
<td>Combined Cycle</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

(*) includes the following stages: advance planning, financial, regulatory, purchase order, construction

Source: Summarized from California Energy Commission (1979a)

The objective of the model is to minimize expected discounted costs over a 45 year planning horizon. Each of the nine planning periods contains five years. 1985, 1990, 1995, up to 2025 are the representative years in each period, \( t \). The period from 1983 through 1987 is taken as a starting point because for the years between 1973, when the paper was written, and 1983, the plant mix is already determined by past decisions. It is known with certainty that the breeder will not be available during the first planning period, \( p(2) = 0.2, p(3) = 0.4, \) and \( p(4) = 0.4 \). After period 4, the breeder will be available with certainty.

The demand requirements are represented by a load duration curve, approximated by three blocks. There are six alternative technologies considered: LWRs, breeders, pumped storage, and three types of fossil fuel plants. To incorporate some flexibility in the use of the plants, six modes of operation, \( k \), are defined, three for the five thermal and three for the pumped storage hydro plants. Since
convex combinations are possible in linear programming; other modes of operation are defined implicitly.

A very interesting feature of the model is that peak demand requirements can be varied endogenously. It is assumed that peak demand has a price elasticity \( \eta = -0.5 \). The demand curve is extrapolated from the reference value \( \bar{p} \) and \( \bar{q} \), price and quantity, which are used for comparison in the case when the requirements are fixed. Let \( q \) be the future peak demand. It is assumed that the money value of total benefits is an isoelastic function of \( q \): \( u(q) = aq^b + c \). The constants \( a, b, c \) are to be estimated from the demand curve. Equating marginal benefits and market price, \( p \), it follows that \( p = abq^{b-1} \). Hence, \( \eta = 1/(b-1) \), or \( b = -1 \). The constant term, \( c \), is chosen such that \( u(\bar{q}) = 0 \). This is necessary if the case with fixed demand requirements shall be comparable to the one with flexible demand. Once \( a, b, c \) are estimated, a piecewise linear approximation to the benefit function is chosen. The index \( \varphi = 1, ... , 9 \), and the approximation refers to the following nine points:

\[
\begin{align*}
\bar{q}(1,t) &= \bar{q}(t) \\
\bar{q}(2,t) &= \bar{q}(t)[0.95] \\
& \vdots \\
\bar{q}(\varphi,t) &= \bar{q}(t)[1-0.05(\varphi-1)] \\
& \vdots \\
\bar{q}(9,t) &= \bar{q}(t)[0.60]
\end{align*}
\]

This means the calculations are based on the range from 100 percent to 60 percent of \( \bar{q}_t \). The loss in benefits has to be added to the cost function and can be interpreted as a penalty for failing to meet peak demand. The rest of the notation required to present the objective function is introduced below.

- \( CP(i,1) \) Capacity increments, first available in period \( t \) (MW)
- \( CP(i,t,s) \) Plant types
- \( j = 1, 2, 3 \) Energy blocks of load duration curve
- \( t = 1, ... , 9 \) Planning periods
- \( s = 2, 3, 4 \) Date of commercial availability of breeder
- \( C(i) \) Initial investment costs ($/MW) of plant type \( i \)
\( \text{UT}(i,k,1) \) \quad \text{Peak capacity utilization in the representative year of period } t \text{ (MW)}

\( k = 1, \ldots, 6 \) \quad \text{Modes of operation}

\( \text{OM}(i) \) \quad \text{Operating and maintenance costs (\$/MWh) type } i

\( F(i,t) \) \quad \text{Fuel costs (\$/MWh) type } i \text{, mode } k

\( H(k) \) \quad \text{Hours operated per year, mode } k

\( U(\tilde{q}(\varphi,1)) \) \quad \text{Shortage cost for failure to meet peak demand levels}

\( \beta \) \quad 5-year discount factor

The objective is to minimize total discounted capital, operation, maintenance, and fuel costs, and a penalty in case of failure to meet peak demand levels.

\[
\begin{align*}
\text{MIN}_{\text{CP,UT,WT}} & \quad \left[ \sum_{i=1}^{6} C(i) \beta^{0.5} \text{CP}(i,1) + \sum_{i=1}^{6} \sum_{k=1}^{q} \sum_{s=2}^{l} p(s)C(i) \beta^{0.5} \text{CP}(i,1) \\
& \quad + \sum_{i=1}^{6} \sum_{k=1}^{6} [\text{OM}(i) + H(k)F(i,1)] 5\beta \text{UT}(i,k,1) \\
& \quad + \sum_{i=1}^{6} \sum_{k=1}^{6} \sum_{r=2}^{q} \sum_{s=2}^{l} p(s)[\text{OM}(i) + H(k)F(i,1)] 5\beta \text{UT}(i,k,1) \\
& \quad - \sum_{\varphi=1}^{5} u(\tilde{q}(\varphi,1)) 5\beta WT(\varphi,1) \\
& \quad - \sum_{\varphi=1}^{5} \sum_{r=2}^{q} \sum_{s=2}^{l} p(s)[u(\tilde{q}(\varphi,1)) 5\beta^r WT(\varphi,1)] \right]
\end{align*}
\]

(4-82)

The fixed demand requirement constraints are the usual ones. Let \( A(j,k) \) be the availability factor and \( j = 1,2,3 \) the index for the demand blocks. Then \( \sum_{i} \sum_{k} A(j,k) \text{UT}(i,k,t,s) \) has to be at least as large as fixed energy demands minus the exogenously given hydroelectric energy supplies. In the case of flexible peak demand, the fixed energy demand is replaced by the following weighted average: \( \sum_{\varphi} \tilde{q}(\varphi,t)WT(\varphi,t,s) \), where \( WT(\varphi,t,s) \geq 0 \) are the peak demand interpolation weights with \( \sum_{\varphi} WT(\varphi,t,s) = 1 \). Further, initial and new capacities together have to be sufficient to satisfy peak capacity requirements plus a reserve requirement.

If in period 2, the breeder is not yet available, the investments made during that period should be the same, whether the breeder will become available in period 3 or 4, i.e., \( \text{CP}(i,2,3) = \text{CP}(i,2,4) \).
The demand data for an application of the model is taken from the 1970 National Power Survey. Cost coefficients are chosen as to be “reasonable.” The coefficients for the breeder reactor are chosen to be economically competitive with LWRs.

The results for the United States show that a fast development of breeder capacity results in considerable savings, and these savings are relatively insensitive to the assumed annual fossil fuel price increase and variations in peak demand. The discounted costs would be in the order of $4 to $5 billion. The savings from peak-load pricing would be even more substantial: $30 to $40 billion.

To test for possible aggregation bias when applying the model to the whole United States, the country has been subdivided into six regions. Then the calculations were repeated. The results indicate that the aggregation error in the future demand for fossil fuels was around 30 percent.

Planning with Explicit Stochastic Reserves Constraint: Scherer and Joe (1977). Most utilities use loss of load probabilities (LOLP) as a criterion for system reliability. The models considered so far implicitly or explicitly employ restrictions which express the reserves as a fixed percentage of the capacity requirement. The reliability of a system planned in this fashion cannot be expressed easily. Therefore, the result is somewhat arbitrary. It would be desirable to incorporate the LOLP concept into linear models of the kind discussed above.

Scherer and Joe propose a simple mixed integer model. The important assumptions are that each plant has only one generating unit, and the operating state (up or down) for each plant is independent of that of all other plants. Demand is determined exogenously and assumed to be price insensitive. The true concave cost function is approximated using the fixed charge approach. Finally, a Bernoulli outage distribution is assumed, where \( p, q \) are the probabilities that a plant is up or down and are assumed not to depend on plant size. Having \( p, q \) given for each plant, every possible configuration of up and down plants can be assigned a probability \( P(n) \), where \( n \) is the index of each configuration or state \( p \) is the required reliability level.

The objective is to minimize total capacity costs, consisting of a variable part and a fixed charge:

\[
CK(j) = \frac{\$/MW}{year}, \ K(j) = MW
\]
\[
\text{MIN} \sum_{j=1}^{J} (CK(j)K(j) + CF(j)FC(j)); \quad FC(j) = 0 \text{ or } 1; \quad j = 1, ..., J
\]

Since there are \(J\) plants, there are \(2^J\) possible states. Let the 0,1 integer variable \(X(n)\) be associated with each state. Let \(S(n)\) denote the set of plants that are up in state \(n\). If \(X(n) = 1\), then the capacity of all the plants up in state \(n\) has to be sufficient to meet demand. However, the capacity constraint (4-84) cannot be met with certainty. Therefore, the reliability constraint (4-85) requires that the system must be able to satisfy demand with a probability of at least \(P\). \(1 - P\) is the loss-of-load probability (LOLP).

\[
\sum_{j \in S(n)} K(j) \geq DX(n); \quad n = 1, ..., 2^J
\]

(4-84)

\[
\sum_{n=1}^{2^J} P(n)X(n) \geq P
\]

(4-85)

The approximation to the cost curve is only good within a limited range. This motivates the next constraint.

\[
0 \leq K(j) \leq KMAX(j)FC(j); \quad j = 1, ..., J
\]

(4-86)

Finally, the state variables are binary.

\[
X(n) = 0 \text{ or } 1; \quad n = 1, ..., 2^J
\]

(4-87)

The calculation of the \(P(n)\) values implies that all \(J\) plants will always be built, only their size has to be determined. However, the solution can be \(K(j) = 0\), so that the procedure is not restrictive in this respect. Can a solution in which not all plants are built be feasible? Scherer and Joe (1977, page 981) give an affirmative answer to this question.

As expected, numerical computations show that the number of plants increases with demand, but system capacity reserves decrease as a percentage of demand. This shows that the rule-of-thumb of constant reserve margins may lead to significantly different results, especially in a growing system. The costs of computation of the Scherer-Joe model are increased due to the introduction of \(2^J\) demand constraints instead of the usual one.
SUMMARY AN COMPARISON

Linear programming models of investment have been used in the electric industry with varied purposes and results. The main properties of the models surveyed are summarized in Table 4-4. This table outlines the similarities between the models discussed. All models, except one, minimize total discounted costs. Minimization of revenue requirements is very closely related. The load duration curve is used in most of the models to represent fluctuating demand requirements. Only one model has been able to incorporate uncertainty about future demand levels other than by imposing fixed reserve margins. Manne’s (1974) model incorporates uncertainty about the date of the commercial availability in a more sophisticated manner.

Linear programming is a convenient and flexible mathematical tool, having a large number of users and many efficient computer algorithms. The detail incorporated in the models discussed above could not easily have been handled using dynamic or nonlinear programming. Many insights have been gained from linear programming investment models since the pioneering study of Massé and Gibrat (1957).

But there is one area where the weaknesses of the method cannot be overlooked. Linear programming models have not been able to incorporate the effect of uncertainty on decision makers in a satisfactory manner. Until about 1970, planning cautiously, i.e., constructing too much rather than too little capacity, carried a small risk. Demand grew at a quick and steady pace, so that unplanned overcapacity was soon absorbed. After 1970 this has not been the case. As a consequence, one expects to observe a tendency among planners of utilities to favor plants with short construction lead times over those with long lead times, even if the former cost a bit more. Linear programming models are unable to capture this effect. This inability has motivated research using other mathematical tools. The most popular model, the Wien Automatic System Planning Package (WASP; see Jenkins and Joy; 1974), uses dynamic programming coupled with user intervention. The model consists of six modules. The first describes the existing system. In the next module, the possible expansion candidates are defined. Plants of the same type but with different capacities will be treated as different expansion candidates. Then the demand requirements are described for each planning year. Load duration curves are used. In the fourth module, the program user can specify minimum and maximum reserve requirements. The minimum and maximum
number of units of a given type can also be restricted. The module then finds all feasible configurations among the expansion candidates. The next module then generates the operating costs for each of the feasible plans. This is done with a probabilistic simulation model which allows the incorporation of random forced outages. In the last module, a dynamic programming model is used to find the optimal expansion schedule. An optimal plan must not exceed a predetermined loss-of-load probability. This is an attractive and powerful model. Among the disadvantages is the fact that dispatching and investment decisions have to be made separately.
<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>General Description</th>
<th>General Purpose</th>
<th>Objective</th>
<th>Treatment of Demand Requirements</th>
<th>Economies of Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massé and Gibrat (1957)</td>
<td>One-period LP investment model</td>
<td>Demonstrate effect of fluctuating demand on optimal plant mix</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>McNamara (1976)</td>
<td>Multi-period LP investment model</td>
<td>Provide easy to use model to simulate effects of changes in parameters</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>Anderson (1972); Turvey and Anderson (1977)</td>
<td>Multi-period MIP investment model</td>
<td>Normative model to determine optimal plant mix given many constraints</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>Gately (1970)</td>
<td>Multi-period MIP investment model</td>
<td>Normative planning model to determine optimal plant mix, investment schedule</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>Yes, fixed charge approach</td>
</tr>
<tr>
<td>Cherene and Schaeffer (1979)</td>
<td>Multi-period MILP investment model</td>
<td>Determine optimal plant mix when dispatching problem in integrated into model</td>
<td>Revenue requirement</td>
<td>“Typical” daily load curve for each season</td>
<td>Yes, plant size fixed</td>
</tr>
<tr>
<td>de la Garza, Manne, and Valencia (1973)</td>
<td>Multi-period MILP investment model</td>
<td>Normative regionalized planning model to determine optimal plant mix</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>Yes, fixed charge approach</td>
</tr>
<tr>
<td>Scherer (1977)</td>
<td>One-period MILP investment model</td>
<td>Determination of marginal costs</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>Yes, fixed charge approach</td>
</tr>
<tr>
<td>Thompson, Calloway, and Nawalanic (1977)</td>
<td>One-period LP investment model</td>
<td>Finding marginal costs, especially as a consequence of stricter environmental constraints</td>
<td>Cost minimization</td>
<td>Total demand</td>
<td>No</td>
</tr>
<tr>
<td>Deonigi (1971) and Deonigi and Engel (1973)</td>
<td>Multi-period LP investment model</td>
<td>Simulate role of nuclear power plants in existing system. Use for cost-benefit analysis</td>
<td>Cost minimization</td>
<td>Total demand</td>
<td>No</td>
</tr>
<tr>
<td>de Boer, Leclercq, and de Haan (1971)</td>
<td>Multi-period LP investment model</td>
<td>Determine role of nuclear plants in existing system of non-nuclear plants</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>Frankowski (1971)</td>
<td>Multi-period LP investment model</td>
<td>Determine role of thermal (incl. nuclear) plants in existing system that is predominately hydro</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>Pozar and Udovicic (1971)</td>
<td>One-period LP investment model</td>
<td>Determine role of nuclear plants in existing system of non-nuclear plants</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>Manne (1974)</td>
<td>Multi-period probabilistic sequential LP investment model</td>
<td>Optimal investment strategy when commercial availability date of new technology is uncertain</td>
<td>Cost minimization</td>
<td>Approximate to load duration curve</td>
<td>No</td>
</tr>
<tr>
<td>Scherer and Joe (1977)</td>
<td>Single-period MILP investment model</td>
<td>Comparing LOLP and fixed margins as criteria for reliability reserves</td>
<td>Cost minimization</td>
<td>Aggregate demand</td>
<td>Yes, fixed charge approach</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Treatment of Uncertainty</td>
<td>Spatial Model, including Transmission</td>
<td>Financial Constraints</td>
<td>Lead Times</td>
<td>Integrated Dispatching and Investment Decision</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------</td>
<td>---------------------------------------</td>
<td>-----------------------</td>
<td>------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Massé and Gibrat (1957)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>Limit on investment spending</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>McNamara (1976)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>No</td>
<td>Yes. Investments in early period are very expensive</td>
<td>No</td>
</tr>
<tr>
<td>Anderson (1972); Turvey and Anderson (1977)</td>
<td>Reserve requirements</td>
<td>One version, yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Gately (1970)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Cherene and Schaeffer (1979)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>No</td>
<td>Yes. Explicitly incorporated in model</td>
<td>Yes</td>
</tr>
<tr>
<td>de la Garza, Manne, and Valencia (1973)</td>
<td>Reserve requirements</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Scherer (1977)</td>
<td>Reserve requirements</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Thompson, Calloway, and Nawalanic (1977)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>Limit on investment spending</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Deonigi (1971) and Deonigi and Engel (1973)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>de Boer, Leclercq, and de Haan (1971)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Frankowski (1971)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Pozar and Udovicic (1971)</td>
<td>Reserve requirements</td>
<td>No</td>
<td>Yes. Limit on investment spending</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Manne (1974)</td>
<td>Uncertainty about availability of new technology</td>
<td>No</td>
<td>No</td>
<td>Yes. Planning for periods after all inherited orders have been delivered</td>
<td>No</td>
</tr>
<tr>
<td>Scherer and Joe (1977)</td>
<td>Incorporation of LOLP</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
CONCLUSIONS

Despite the shortfalls in handling uncertainty, Table 4-4 shows that the other main concerns listed in the introduction can be incorporated into linear programming models. This optimism is dampened by problems with computational efficiency, which occur especially when integer and mixed-integer methods are employed. Given these limitations, one could use different methods and models to answer different questions. It would be interesting to investigate the relationships among the various models. Possibly, the result of one model could be used as input for another. Naturally, this is theoretically less appealing than the simultaneous determination of all variables, but it may just be the way utilities plan in reality, and such a model may yield good forecasts.

Finances are another area that has become important in planning of utilities. In only a few models are financial constraints considered. Massé and Gibrat (1957) demonstrate the presence of a trade-off between operating and maintenance costs on one side, and capital costs on the other. Given the increasing difficulties of utilities to obtain investment capital, they may want to or be forced to exploit this possibility. An investment planning model for the next twenty or thirty years should be able to capture this effect or it probably will yield unrealistic results.
REFERENCES


---

1 A later version of this paper with a one-period empirical application was published as PV Schaeffer, LJ Cherene, 1989. “The Inclusion of ‘Spinning Reserves’ in Investment and Simulation Models for Electricity Generation.” European Journal of Operational Research 42(2): 178-189


