

Archive

Energy Laboratory
in association with
Electric Power Systems Engineering Laboratory
Department of Civil Engineering

THE ECONOMICS OF RELIABILITY FOR
ELECTRIC GENERATION SYSTEMS

by

M.L. Telson

MIT Energy Laboratory Report No. MIT-EL 73-016
May 1973

ENERGY LABORATORY

The Energy Laboratory was established by the Massachusetts Institute of Technology as a Special Laboratory of the Institute for research on the complex societal and technological problems of the supply, demand and consumption of energy. Its full-time staff assists in focusing the diverse research at the Institute to permit undertaking of long term interdisciplinary projects of considerable magnitude. For any specific program, the relative roles of the Energy Laboratory, other special laboratories, academic departments and laboratories depend upon the technologies and issues involved. Because close coupling with the normal academic teaching and research activities of the Institute is an important feature of the Energy Laboratory, its principal activities are conducted on the Institute's Cambridge Campus.

This study was done in association with the Electric Power Systems Engineering Laboratory and the Department of Civil Engineering (Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics and the Civil Engineering Systems Laboratory).

ABSTRACT

THE ECONOMICS OF RELIABILITY FOR
ELECTRIC GENERATION SYSTEMS

by

MICHAEL LAWRENCE TELSON

Providing excess generation capability for reliability purposes costs a utility money. It is also true, that providing higher reliability adds value to electric service. After some point, however, the additional benefits do not warrant the additional cost. This work deals with the questions of how reliable should generation capability be for meeting system loads, what models should be used to measure this reliability and what bases should be used for answering the above two questions.

We critique the measures that have been used and suggest that an energy shortage related measure is probably the best one to use. In addition, we specify what load and what benefit measurement models should be used. The resulting procedures are then used to develop techniques for 1) creating long term expansions of electric utility systems at various reliability levels and for 2) analyzing the costs and benefits of plant additions to arbitrary system expansions.

From representative data, we conclude that present generation reliability levels for operation are probably too high. We discuss the magnitude of possible savings and we find that although only a few % of total system costs, they may have substantial profit impact.

PREFACE

This report is based on the thesis which bears the same name; however, it is mercifully shorter although still too long and involved to be easily understood. For this reason, I decided it would be better to include a road map to the report while referring to it as a preface.

I am relieved that the table of contents is a useful point of departure. The Introduction-Chapter I is an attempt to place the subject of the reliability of bulk electric generation systems in its overall context. We discuss why it has become important and what is--and what is not--being done about it in both the private and public spheres. Chapter II describes the reasons why a measure of reliability is needed for bulk electric generation systems and discusses the role of reliability criteria in system planning.

In Chapter III we begin to get rather involved in the nitty-gritty details of generation system reliability measurement. We first give an overview of the measurement problem. We discuss the fact that reliability measurements must somehow probabilistically model the capacity available to a given electric system, as well as the demand (load) on that capacity. We then introduce and discuss the more common measures of generation reliability: the Loss of Load Probability (LOLP), the Loss of Energy Probability (LOEP), the Expected Loss of Load (XLOL), and the Frequency and Duration approach (FAD).

After presenting these traditional approaches, we introduce and discuss a technique which allows for calculation of system expected production costs called Probabilistic Simulation (PROSIM). PROSIM also calculates the LOLP and LOEP measures as a byproduct and it is rather simple to program. It will be used to explore the costs of reliability in Chapter IV.

At this point, we critique the various measures and explore their inherent shortcomings; i.e., what they can and can not do. Also there is great confusion regarding the proper formulation and interpretation of the measures and we explore these issues. We conclude that a type of LOEP measure--an expected energy deficit related measure--is our best hope of measuring economically relevant phenomena. This conclusion is not made in vacuo, but is based mainly on our thoughts in Chapter V.

Chapter IV discusses in depth the questions relating to the cost of providing reliability for electric power networks. We construct a framework which allows us to think of (1) the cost of expanding systems at different reliability levels or (2) the cost of making marginal additions to systems expanded at given reliability levels. We find that there is theoretical justification for finding that $c(\text{LOEP})$ is logarithmic in $(1/\text{LOEP})$. We also find that the second approach above, the marginal cost approach, yields interesting and useful results regarding the costs and benefits of making further additions to given systems.

Chapter V constructs an analytical framework for thinking of the benefits of reliability. It also sets up a rough measurement scheme which can be defended over reasonable ranges of generation reliability.

In Chapter VI we put together the results from Chapters IV and V. These results show that (1) "optimal" LOEP levels are two orders of magnitude less reliable than present ones and that (2) customers are implicitly paying a rate two orders of magnitude higher than their regular rates for their peak power (simply because peaking units are not expected to generate much energy). The results suggest that if the assumptions behind the models are believed, present levels for generation reliability may be too conservative. It is worthwhile to point out that there have been very few instances of

generation induced blackouts and that most of the recent ones are due to distribution and transmission system failure. The results also discuss the potential impact on savings on generation expenditures.

We suggest that the reader not get bogged down in Chapter III. In particular, sections 3.2.4 and 3.2.5 can be skipped without impeding understanding of the results. Chapters IV and V are necessarily quite detailed and may contain more material than any given reader may find useful. Each reader should decide what is of interest by glancing at the table of contents and then perusing the text to find what is most useful. Chapter VI will probably be of general interest because it is easy, for it contains "the results". Because of this, I should remind the reader of the maxim "easy come, easy go" and wish him Godspeed in his journey.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	2
PREFACE	3
TABLE OF CONTENTS	6
LIST OF FIGURES	10
CHAPTER 1 - INTRODUCTION	12
1.1. - The Rapidly Changing Electric Utility Milieu	12
1.2. - Public Concern and Reaction	17
CHAPTER 2 - THE ROLE OF RELIABILITY IN ELECTRIC SYSTEM OPERATION AND PLANNING: THE ROLE OF EXCESS CAPACITY	28
2.1. - Spinning Reserves	28
2.2. - Operating Reserves	33
2.3. - The Role of Reliability Criteria in System Planning	34
CHAPTER 3 - THE MEASURES OF GENERATION RELIABILITY . . .	38
3.1. - Overview	38
3.2. - Calculation of the Measures of Reliability	46
3.2.1 - Scope of the Measures	48
3.2.2 - The Loss of Load Probability Method	52
3.2.2.1 - The Capacity Model	52
3.2.2.2 - The LOLP Model	58
3.2.2.3 - Hourly Load Duration Curve	59
3.2.2.4 - Daily Peak Load Duration Curve	65

	<u>Page</u>
3.2.2.5 - Other Methods	66
3.2.3 - Loss of Energy Probability Measure	69
3.2.4 - The Expected Loss of Load Measure	73
3.2.5 - The Frequency and Duration Approach	79
3.2.5.1 - FAD Capacity Model	81
3.2.5.2 - The One Machine Capacity Model	82
3.2.5.3 - The Two Machine Capacity Model	86
3.2.5.4 - Generalized n-Machine Problem	89
3.2.5.5 - Procedure to Obtain FAD Measures for the n-Machine Problem	90
3.2.5.6 - The HRW Recursive Algorithm	94
3.2.5.7 - Proof of HRW Recursion Algorithm	99
3.2.5.8 - A Load Model for the FAD Method	101
3.2.5.9 - The Combined Load and Capacity Measure	103
3.3. - Probabilistic Simulation	103
3.3.1 - Another Technique for Reliability Calculations	103
3.3.2 - Other Outputs of PROSIM	108
3.4. - A Critical Review of Reliability Technique	115
3.4.1 - Critique of the Measures	115
3.4.2 - A Critique of Technique Application and Interpretation	118
3.4.3 - Interpretation of LOLP and LOEP as Expectations Rather than Probabilities	131
3.4.4 - The Loss of Load Event	133
3.5. - Conclusions	136

	<u>Page</u>
CHAPTER 4 - THE COSTS OF ADEQUATE GENERATION	140
4.1. - The Function $C(r)$	142
4.2. - Expansions at Two Different Levels of Reliability	143
4.3. - Treatment of Marginal Changes to an Expansion Plan	148
4.4. - Marginal Change Cost Analysis	150
4.5. - Construction of $C(LOEP)$	157
4.5.1 - The Iterative Search Procedure for Determining an Expansion Strategy	157
4.5.2 - Example of Iterative Search Procedure for Expansion Strategy	162
4.6. - An Example of a $C(LOLP)$ Study	168
4.7. - Theoretical Justification for $C(LOEP)$ Form	170
CHAPTER 5 - THE BENEFITS THAT RESULT FROM ELECTRIC POWER SYSTEM RELIABILITY	173
5.1. - Framework for Analysis	173
5.2. - Evaluation of Benefits	184
5.2.1 - Benefits are in the Eyes of the Beholder	184
5.2.2 - Examples of Losses from Lower Reliability of Service	185
5.2.3 - Basic Variables that Affect the Valuation of the Benefits of Increased Reliability	189
5.2.4 - Long Run vs. Short Run; Marginal vs. Substantial Changes in Reliability Levels . .	199
5.2.5 - Distribution Reliability is a Limit on Customer Reliability and Provides an Opportunity for Observing Customer Reaction to Sharply Lower Reliability Levels	202
5.2.6 - Customer Options: Interruptible Power Contracts and Backup Supplies	203
5.3. - Construction of a Benefit Function	214

	<u>Page</u>
5.3.1 - The Issue of Acquisition of Backup Capability	216
5.3.2 - Total Losses Will Depend on How the Unserved Energy is Apportioned	221
5.3.3 - By Assumption Losses are Linearly Related to Unserved Energy	223
5.3.4 - The L(LOEP) Expression	227
5.3.5 - The L(EU) Expression	229
CHAPTER 6 - DETERMINATION OF DESIRABLE LEVELS OF GENERATION RELIABILITY	233
6.1. - Determination of LOEP* for Expansion Studies	233
6.2. - The Costs and Benefits of Marginal Additions to Specific Expansion Plans . . .	236
6.2.1 - Use of the Probabilistic Simulation Technique	236
6.2.2 - Theoretical Analysis	240
6.3. - Analysis of the Results	247
6.4. - What is the Magnitude of the Cost Differentials Involved	253
6.5. - Conclusions and Suggestions for Further Research	255
BIBLIOGRAPHY	264

LIST OF FIGURES

CHAPTER 3	<u>Page</u>
3.1 The Probability Density of \tilde{C}_{tot} , a Sum of Independent Random Variables, \tilde{C}_i	55
3.2 The Probability Density of \tilde{C}_{tot} , a Specific Example	57
3.3 Hourly Load Duration Curve	61
3.4 Calculation of the LOLP Measure	64
3.5 LOEP Measure Calculation	71
3.6 The Basic Markov Model	82
3.7 The Curves $\{F_i(x)\}$ and $F_n(x)$	106
3.8 The Effect of Using DPLDC's vs. LDC's	125
3.9 An Equivalent Load Distribution Curve and its Significant Margin States	129
CHAPTER 4	
4.1 Iterative Search for Expansion Strategy	160
4.2 Output for Initial ϵ Strategies	164
4.3 LP Output for the Six Attempted ϵ Strategies	166
4.4 PROSIM Output for the Six Attempted ϵ Strategies	167

CHAPTER 5	<u>Page</u>
5.1 Comparison of Two Systems Expanded at Different Levels of Reliability	174
5.2 Effect of Adding Another Generator to a System Design	174
5.3 Benefits and Costs as Functions of Reliability	176
5.4 Marginal Benefits and Costs	178
5.5 Benefit Evaluation Process	184
5.6 Customer - Loss of Load Incident Impact Matrix	193
5.7 The Effect on Unserved Energy of Adding Gas Turbine to Two Different Expansion Plans	232

CHAPTER 6

6.1 Marginal Changes in \bar{EU}	243
6.2 Sensitivity of Expansion Strategies 1 and 6 Forecasting Errors and Delays in Capacity Installation	252
6.3 A Possible Formulation for $L(\bar{EU})$	260



CHAPTER I

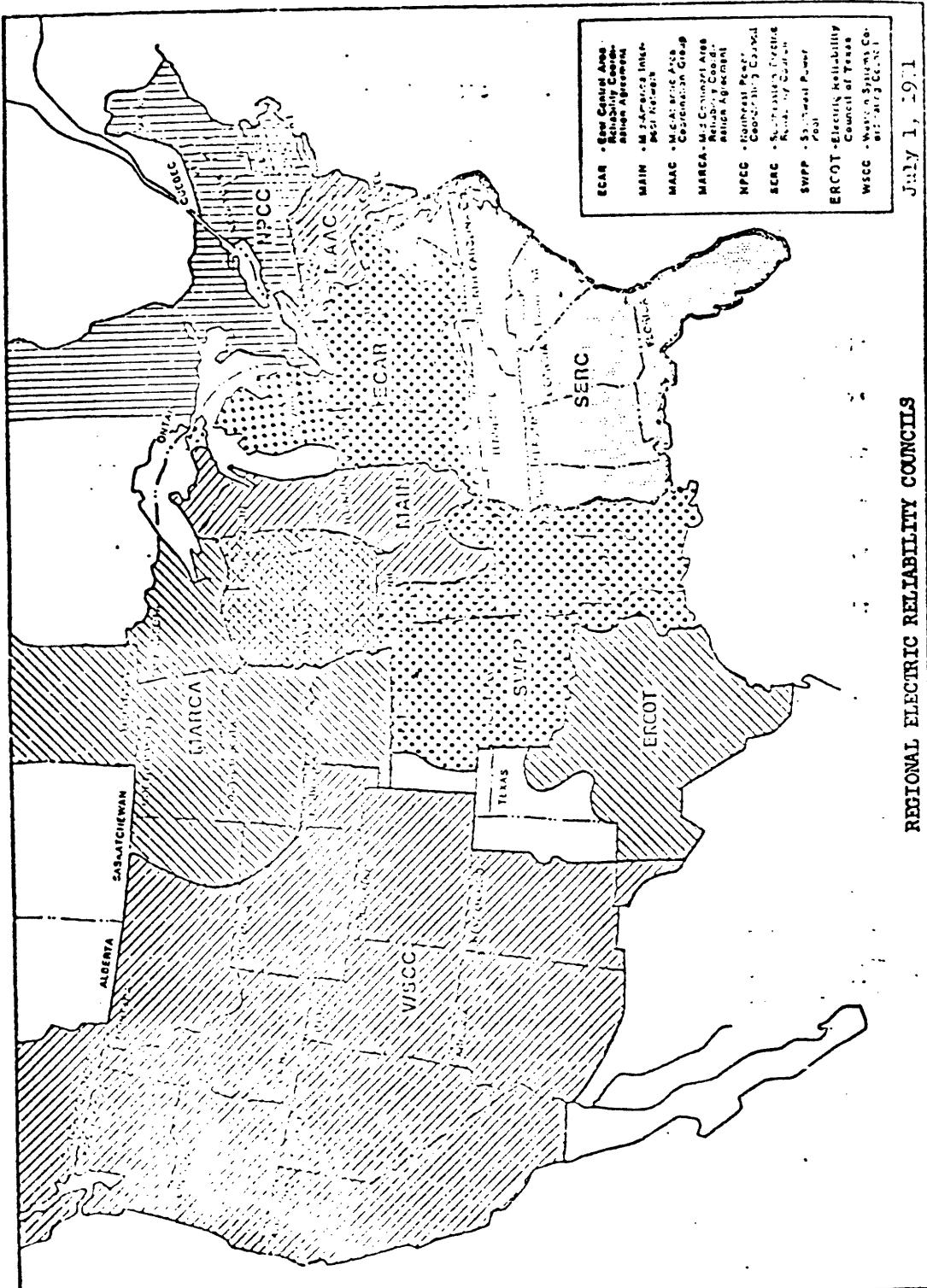
INTRODUCTION

1.1 The Rapidly Changing Electric Utility Milieu

The issue of bulk electric power supply reliability has become increasingly important over the past seven years beginning with the Northeast Blackout of November, 1965. One of the initial industry reactions to this concern was the creation of the National Electric Reliability Council¹ (NERC) to plan effectively the power supply to ensure that the type of technical failure that led to the 1965 incident not reoccur. The Council set up regional organizations (shown on the next page) that would encourage effective planning to support future loads on an individual and collective system basis. [See Northeast Power Coordinating Council (NPCC) report. NPCC is a member of NERC]. The Council's efficacy in achieving its objectives is strongly disputed by people inside and outside of the electric industry, though it is generally agreed that it has been of some help as a device for communication.

At about the same time that we discovered our vulnerability, national concern for environmental quality began to

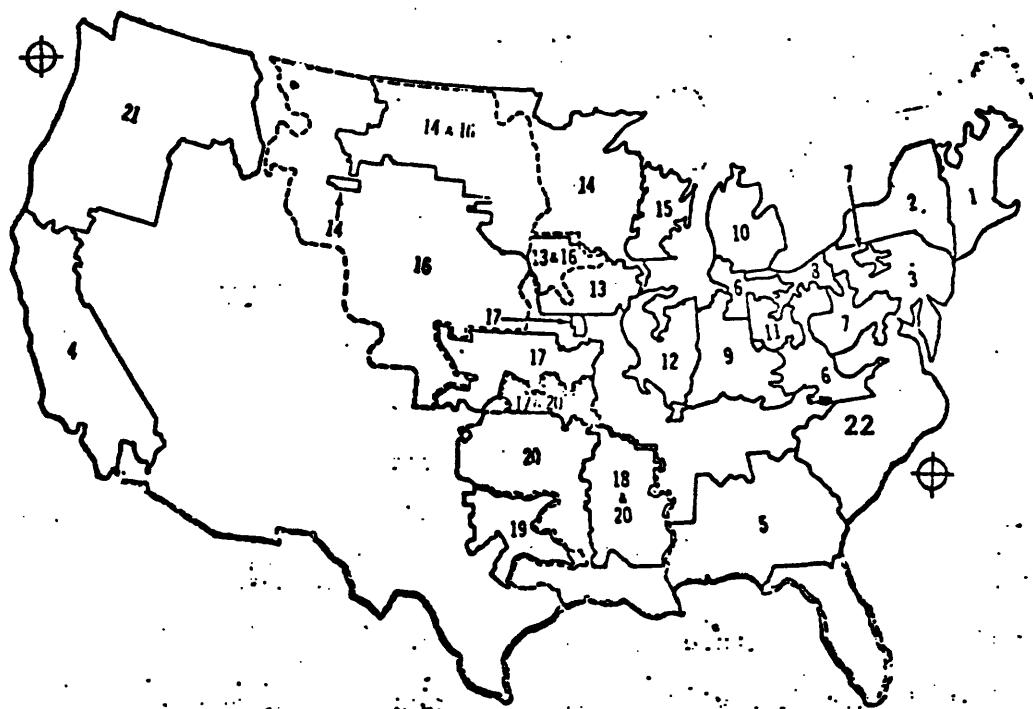
¹The National Electric Reliability Council itself has a small staff but is composed of utility task forces called together by the NERC Chairman. The area wide councils are depicted in the accompanying chart.



July 1, 1971

REGIONAL ELECTRIC RELIABILITY COUNCILS

FORMAL POWER POOLS



- 1. New England
- 2. New York
- 3. P-J-M Interconnection
- 4. California
- 5. The Southern Company System
- 6. American Electric Power System
- 7. Allegheny Power System
- 8. Central Area Power Coordination
- 9. Kentucky - Indiana
- 10. Michigan
- 11. Cincinnati, Columbus, Dayton
- 12. Illinois - Missouri
- 13. Iowa
- 14. Upper Mississippi Valley
- 15. Wisconsin
- 16. Missouri Basin Systems Group
- 17. Missouri - Kansas
- 18. Middle South Utilities System
- 19. Texas Utilities
- 20. South Central Electric Companies
- 21. Pacific Northwest Coordination
- 22. Carolinas

Formal Power Pools

be expressed in increasingly strong fashion. One of the results of the developing national environmental interest was the passage of the National Environmental Policy Act of 1969 (NEPA). NEPA required that an "environmental impact statement" (section 102c) be submitted before the initiation of a project; the meaning and content of this requirement has been strengthened by subsequent court decisions, among them the Calvert Cliffs, Department of Interior and Kalur decisions. As a result of this feature of NEPA and of other law, and because of plain, simple schedule slippage, the U.S. is experiencing great delay in the completion of plants under construction and in getting the approvals for new construction. Because of the long lead time (expected to require 8 to 10 years now) for planning, approving and constructing new electric plants, these delays have left the utilities utterly helpless with recourse in the short run only to expensive, rapidly obtained and installed, gas turbine plant capacity. The fact that during this period much of the new generator capacity was nuclear, and therefore most vulnerable to technologic and legal delay, just aggravated the situation.

A second problem the utilities are beginning to be confronted with involves their prerogatives to use their generating equipment in the same manner as in the past. A specific example involves the modifications in plant operation specified by city regulation whenever air pollution emergencies

occur in New York City.¹

In effect, the utility can not count on being able to use its equipment with as much private discretion as they used to use. Previously, the criteria for operation of a plant, if it was capable of functioning, were mainly economic. The criteria will probably now have to be modified by new environmentally related ones. Eventually, these new operational criteria will have to influence the planning process itself in order to make it possible to design long term electric system architectures which are economically and environmentally improved.

System planners will not only have to make assumptions about future fuel and equipment prices but also have to project what the area's future environmental conditions and regulations might be in that they may influence conditions for future system operation. Also, there is widespread recognition that future fuel prices and supply situations will not be nearly as stable as they once were; this would be yet another area of utility planning calling for new imaginative approaches on the part of management.

The industry has yet to adjust to these new and rapidly changing situations. Among other problems, the reliability

¹See New York City Environmental Protection Administration, Emergency Control Board, "Regulations Pertaining to the Air Pollution Warning System," June 9, 1971.

problem has become of increasing concern. In a February, 1972, publication the National Electric Reliability Council¹ warned that ". . . In five of the nine regions . . . generation reserves would drop to 10% or less in the summer of 1972, reaching levels so low that periodic load curtailment would be expected. . . ." The reliability mentioned here is not the same that would have prevented the 1965 Northeast Blackout. Reliability in this context refers to the adequacy of the industry's generating capacity to accomodate:

1. The demand system load.
2. Forced outages.
3. Capacity out on preventive scheduled maintenance.
4. The needs for retaining operational flexibility in case of unexpected contingency.

1.2 Public Concern and Reaction

The Federal Government's concern is reflected in several legislative attempts in the Congress as well as in the activities of the Federal Power Commission (FPC). Under revised Order No. 331-1 in June of 1970, the FPC required that all significant service interruptions (the lower of 100 megawatts or one half the annual system peak load) be reported;

¹NERC Report "Impact of a Twelve Month Delay of New Nuclear and Fossil-fired Steam Generating Units on the Adequacy of Electric Power in the United States," February 1972.

and that all instances of voltage reductions and utility public requests for reduction of power consumption be reported as well. In January of 1972 in Order No. 445, the FPC issued a policy statement that urged the utilities that had not done so, to develop contingency plans for emergency or short supply situations; in the Order they also suggested a few priorities for steps to be followed during a contingency.

The reliability problem has become particularly serious in the New York City-Westchester County Consolidated Edison Company, (Con Ed) area. Even so, it is of interest to note that the NPCC-Con Ed's NERC Council-is not one of the five serious regions mentioned in the NERC report previously referenced. We will concentrate on the New York City problems because of the greater availability of data on the situation and because as the NERC report shows, it is indicative of what may occur in other regions of the country if present trends continue.

Although New York City has experienced few prolonged blackouts in recent years (September 22, 1970 Richmond-Staten Island, and July, 1972 throughout the City), it has had countless incidents of voltage reductions. Furthermore, supporting the operation of the Con Ed system has placed great strain on the entire New York Power Pool, the systems tied to the P.J.M. intertie, the New England Power pool and the Ontario system; indeed, the entire electric system in the Northeast U.S. These systems have bailed Con Ed out in times of stress

and have in this fashion exposed themselves to risk of experiencing serious system difficulties; needless to say, this is not a satisfactory method of operation for the U.S. Power System.

One of New York City's more serious incidents occurred in the subway system on July 28, 1970. As a consequence of placing the system on "series operation" subway speeds were reduced by approximately 50%; since the NYC subway system teeters on the verge of congestion anyway, the speed reduction was enough to cause serious buildups of commuters in stations. A New York City spokesman later commented that ". . . there was some concern by the Police Department that you may have riots and people may be injured or thrown off the platforms."¹

The area's problems have been aggravated by unfortunate forced outage² experience with Big Allis (Ravenswood Unit #3), a 1000 megawatt fossil fueled unit. Con Ed has also incurred great delays in bringing on line most if not all of their baseload planned completions. Con Ed was not able to bring any baseload capacity on line for almost three years until one of the Bowline plant units (jointly owned with Orange and Rockland utilities and, significantly, outside the

¹Carmine G. Novis, Director of the Emergency Control Board of the City of New York in the testimony on case 25937 conducted by the Public Service Commission of the State of N.Y.

²Forced outage--a machine condition occurs that requires that the machine be shut down, i.e., forced out, without prior warning.

area jurisdiction of Con Ed) came on line in July of 1972. It has instead acquired significant amounts of gas turbine capacity to the point where it represents almost 25%¹ of its total installed capacity. This is an undesirable situation from at least two standpoints:

1. Gas turbines consume premium fuels and
2. they are a very inefficient source of power.

It also turns out that gas turbines are very unreliable and therefore one must allow for greater forced outage rates and thus install more capacity than would otherwise be necessary for units with better reliability. Furthermore, it is suspected that because gas turbines are located at ground level, they contribute much to ground level air pollution levels despite their relative exhaust cleanliness.

The electric power situation in the area has deteriorated to the point where both New York City and New York State have addressed themselves in public fora to these shortage questions. The New York State Public Service Commission (NYPSC) proceedings on load shedding² have been of most interest. In early 1971, the NYPSC started hearings on what should be done about managing power shortages if one should occur.

¹Source: N.Y. State Public Service Commission.

²Case 25937. "Proceeding on Motion of the Commission as to the plans and procedures of electric corporations for load shedding in time of emergency."

The great value of these proceedings was that they pointed out some of the social costs of not having electric power available when demanded, given the present utility organization for implementing a load curtailment. As part of this thesis we will expand this question to cover the long run effects of possible curtailment schemes as well as the short term effects.

A very clear lesson of those hearings so far is that there are very high social costs to an indiscriminate load curtailment; examples include the previously mentioned subway problem, and area blackouts that could result in people being trapped in elevators and hot unsafe buildings, as well as having the effect of rendering helpless persons who rely on life support equipment such as iron lungs. There are many more such examples.

The hearings participants took for granted that voltage reductions would be effective in shedding load; the long term effectiveness of this approach is disputed, although at present roughly 5% of the Con Ed peak load can be shed by means of an 8% voltage reduction. Many types of electric consuming equipment will not reduce their power consumption when the voltage is lowered by the maximum 8% prescribed, and on the contrary become less efficient and may even consume more power to accomplish a given task. A good example of this is that of voltage regulators that strive to keep the voltage constant independent of voltage fluctuations, if the voltage

should fall the device becomes less efficient, and thereby uses more power to produce the same appointed task; another example is that of a thermostat controlled air conditioner, which when exposed to a lower voltage will simply become less efficient and keep its compressor on longer than otherwise necessary to accomplish its task. The only loads that would clearly be reduced would be those resistive loads such as lights, irons, and electric stoves. In any case, the NYPSC was also concerned that the load shedding procedures be equitable to all groups concerned in addition to minimizing total social costs. In particular, this meant that a certain group of consumers could not be singled out for carrying the burden of such blackouts at all times unless voluntarily done through offering them some quid pro quo.

One of the reasons for the seriousness of load shedding incidents is that present systems have not been built to cope with load shedding on a rational basis. For example, almost all of New York City is impossible to selectively blackout except on a building by building basis, and this would require installing special types of switches at each building. In particular, load in New York City can presently only be shed on higher feeder network bases leaving 10 square block areas without power at a time.

However, it should be possible in the future to build in some load shedding capacity in new system construction. Yet installing this capacity will not be cheap, and providing

for its management in times of dropping load and in the process of coming back up will be still more expensive.

Even if the planning were done, load shedding will still have a:

1. Social cost (a monetary cost, and an inconvenience and anxiety cost.)
2. A monetary cost to the utilities (sales cost, increased system management and purchases of "expensive" power from other utilities which cuts into present profits since rates are fixed in the short term),
3. and over the long run perhaps an environmental cost in the sense that some more environmentally degrading activities may replace the central station electric generating system.

On the other hand, providing reliability costs money, and it also incurs environmental costs. It should be noted that marginal additions of capacity are expensive in relative terms. In particular, because 80% of revenues go to nonfuel expenses¹ it can be seen that most expenses are incurred whether or not the plant generates electricity. This fact will only become more true as time goes on and utilities acquire more nuclear capacity which has an even greater capital

¹Vennard, E., The Electric Power Business, McGraw-Hill, N.Y., 1970, p. 52.

to fuel cost ratio.

Second, building extra capacity for reliability purposes has environmental costs in the form of a greater number of degraded sites and transmission corridors. However, it should be noted that environmental quality factors that relate to the amount of energy produced rather than relating to the capacity installed or the differential locational impact of such activities--such as gross amount of air emissions, thermal discharges and fuel resources consumed--are not affected by these higher capacity levels, and depend only on the gross energy sales realized.

Each utility has faced this dilemma in determining its proper level of reliability. There are also strong inter-utility pressures to keep these levels high for fear that one system may bring down others. Most utilities have aimed toward having almost perfectly reliable service. In practice, their system planners usually use the "loss of load probability" (LOLP) method and set a "one day in 10 years" loss of load probability objective.¹ This target has been used because over the years it has seemed to yield good performance. Other utilities have used the so called "frequency and duration" (FAD) measures, and still others have used the "loss of energy probability" (LOEP) method to plan their system additions.

¹See First Report Northeast Power Coordinating Council, September 1968, p. 22.

Typically the reliability criterion is taken as a given that is to be met by the system plan regardless of its cost; i.e., it is not a variable parameter in the plan.

The reliability planners' job will in the future become more difficult and more important in at least two ways. First, there will be a need for development of sensible reliability measures which will facilitate the development of reasonable governmental policy with regard to pollution control regulations. For instance, there will be increasing interest in allowing for intermittent emission controls set on the basis of attaining the ambient air quality standards rather than keeping emission standards fixed at minimum levels that ensure that the quality standards are met even during the worst of atmospheric conditions. Also, present regulations pertaining to the implementation of the Clean Air Act will require retrofitting of older plants with sulfur oxide scrubbers thus causing longer than usual shutdowns and higher than expected forced outage rates partially because of shakedown periods for the new equipment and because of the fact that there are more things that can go wrong. Both of these pollution control regulations should force a general rethinking of the reliability measure problem.

Second, in the future there will be a greater penalty on overbuilding than there has been up to now. In the past, overbuilding has not been a serious problem because of two simple economic facts:

- 1) plant efficiencies have been increasing rapidly
- 2) the capital cost per kw. of installed capacity had been decreasing until about 1966, aided by the impressive gains afforded by economies of scale.

Because of these facts, any overbuilding of capacity in past years had a bright side to it since it allowed for recouping fuel and overhead savings through earlier retirement of, and/or less reliance on, older plants whose heat rates (20,000 BTU/kwhr) could be up to twice as large as those of the newer base load capacity being installed, as well as higher than those of currently available peaking capacity (at 15,000 BTU/kwhr or less).

These two facts have changed in the past few years. From now on, efficiency improvements do not promise to be significant, and rapid inflation has dwarfed any economies of scale and made the capital costs of new capacity soar thus creating real incentive not to retire present older capacity. Before moving on, we would like to point out that postponement of retirement could be used as a partial buffer to offset unexpected delays in plant construction.

We will explain the methods that are being used for making reliability calculations later in the text; however, none of these has been analyzed carefully enough to see whether it is a satisfactory measure for dealing with the new problems the utility is being faced with. In particular, it has become clear that the new public concerns are not being

satisfactorily handled by the older established planning methods. This is evidenced by the increasing role the public (the governmental and non-utility private sectors such as private citizens and environmental groups) is playing in determining operating and construction priorities. It is our intent to help build a bridge between these concerns.

CHAPTER 2

THE ROLE OF RELIABILITY IN ELECTRIC SYSTEM OPERATION

AND PLANNING: THE ROLE OF EXCESS CAPACITY

It is not generally understood that extra capacity, above that strictly needed to serve the peak load, is an indispensable factor in the planning and operation of electric power systems. On a second by second basis it is necessary for a system to have quick starting reserves available in case of sudden unforeseen contingency; these reserves are called spinning reserves. An electric power system also needs reserves on a longer term basis, to replace the forced outages that are expected to occur, to fill in for the units on scheduled maintenance or to provide a protection margin for load forecasting errors. Note that the scheduled maintenance is the only controllable part in the above and that, in fact, it is usually managed to coincide with periods of slack demand.

2.1 Spinning Reserves

An electric power system generates exactly the amount of power that is consumed by the ultimate consumers plus the losses. It produces no more, no less, day-by-day and at every instant of time. Sudden occurrences and contingencies such as sudden demand surges, generator failures or transmission line

outages are facts of life in a business that operates day-in, day-out in an uncontrolled environment (demand, weather, forced outages, etc., are all uncontrolled).

There are intricate safeguards to protect system operation and the generators themselves in case of power shortage (say because of a sudden generator failure). In an impending shortage situation it must be possible for the remaining generators to automatically adjust to the new situation. At first the required energy will be provided by the inertia of the generators which will start to slow down as a result, this will cause system frequency to drop. If these remaining generators are not being operated at maximum output, i.e., if there is some slack, these generators may be able, by means of governors, to pick up the excess demand and serve it while keeping the system "stable" on the way.

Alternatively, the system could have some rapidly interruptible load available to shed and thus bring into balance the instantaneous supply-demand equation. Also, the system could have some rapidly starting reserve available which could come in rapidly to fill the gap before system frequency fell too low and started to trip out some system load, initiate voltage reductions or start damaging some of the other generating equipment. These three forms of relief are referred to as spinning reserves. However, there is strong controversy within the industry as to the propriety of arbitrarily lumping these together to satisfy electric pool generating rules.

Note that spinning reserve is a must item for safe system operation. As soon as there is an outage that causes some of the spinning reserve to be used, procedures are set in motion to restore reasonable amounts of spinning reserve to the system since the same contingency could happen again. This is usually done by starting some of the slower starting units that can be used to replace some of the temporarily used spinning reserve. For example, an intermediate fossil unit can be started which within a few hours will allow the gas turbines to be shut off and allow the other generators to return to lower output operation and thus restore some of the "slack" reserves.

If there is no spinning reserve when a contingency occurs, the only remedy is to automatically reduce load through a voltage reduction or blackout areas without warning. Because of this fact, power systems tend to initiate voltage reductions and blackouts before they run out of spinning reserve. The reason for this outwardly contradictory behavior is that by doing this before the "emergency" hits, system controllers can retain some measure of control over where and how the loss of load is administered, and prevent runaway total system breakdowns.

For example, when system load rises to within 5% of system capacity, a certain set of procedures is initiated. These procedures include:

- 1) Contacting other utilities for emergency purchases.

- 2) Pushing the machines on the system beyond the points deemed prudent for operation.
- 3) Intracompany load reduction.
- 4) Possible 3% voltage reduction (which usually brings some load relief, they are referred to as brown-outs).
- 5) Interruptible contract interruption.
- 6) Appeals for voluntary curtailment on radio and T.V.
- 7) Up to an 8% voltage reduction.
- 8) Actual load curtailment (blackouts) in selected areas (if possible) including central station triggered, as well as off-site manual, relay tripping.

All of the above policies serve to push load levels below what they would be without corrective action. It then follows that actual blackouts will start occurring when "uncorrected" loads might have been as much as 5% greater than the installed capacity.

This is a superficial treatment of the spinning reserve problem yet we hope it clears up the reasons that an electric power system has for instituting criteria for spinning reserves. Automatic load shedding devices have become increasingly popular in the last few years to ensure that load will be shed as system frequency drops before the system goes entirely down or before the generators endure serious damage. We should also mention that reconnecting a system--bringing it back up--also

presents major problems of timing and sequencing.

At least two problems remain. The need for keeping spinning reserves available to a system costs money; therefore, an important question is what their amount should be. Also, as interconnected pool operation becomes more common, there will be a need for prescribing for each member firm what their respective contribution should be. These are questions that do not have simple answers and we can at least surmise that their answers depend on the amount of capacity expected to be on line serving the system load, the kind and size of the individual units on line, and the particular dynamics of the units on and off line at any given time.

Some of the questions are answered in tentative guideline fashion in the section of the NPCC First Annual Report prepared by Stone and Webster. The NPCC agreed to ". . . carry enough spinning reserve, properly located, to replace capacity equivalent to the largest single contingency loss in each area within five minutes. . . . In addition, a very fast temporary response should be provided by allocating the spinning reserve to many machines. All machines should be operated, if possible, under governor control, and normally no more than 10 per cent of the spinning reserve in each area should be allocated to one thermal machine.

We recognize that these hydro and quick-start combustion units which can respond in 5 minutes or less may properly be considered as "spinning reserve" and this nonsynchronized

reserve shall not exceed 33.3% of the area's "spinning reserve" requirement."

Obviously, these criteria are based on reasonable assumptions. Note that if the system expansion size guideline calls for new units whose size is a fixed percentage of maximum load, we can expect spinning reserve criteria to stay approximately constant reflecting the system's need for protection against sudden outage of these large units. Presently, system planning reliability criteria do not accommodate the inclusion of spinning reserve as a factor. Present models presume a loss of load occurs only when available capacity goes under the forecast demand; however, we have seen that load curtailment measures may be started before that point. For example, load reductions are not a function of only the deficit of capacity, they also depend on what configuration of capacity available led to that deficit; specifically, a deficit caused by five 200 M.W. plants on outage versus that caused by a 1000 M.W. unit being out, in turn cause different spinning reserve requirements.

2.2 Operating Reserves

These reserves serve a different function than the spinning reserves; operating reserves are expected to be used over longer lead times of warning. These primarily serve to buffer a system:

- 1) against expected, but unpredictable, forced outages.

- 2) against load forecasting errors.
- 3) against tardiness in bringing new capacity on line.
- 4) against higher forced outage rates than expected.
- 5) while plants are placed on preventive scheduled maintenance and
- 6) to provide for operational flexibility.

Note that spinning reserves are a subset by definition of the operating reserves. Also note that preventive maintenance is discretionary with regard to when it is applied and it is usually scheduled in order to equalize the overall system's exposure to risk of loss of load throughout the year.

2.3 The Role of Reliability Criteria in System Planning

It would be worthwhile at this point to give an explanation of the process of system expansion planning and the role reliability criteria play in the process. A typical approach is to design a 30 year plan for system expansion that:

- 1) conforms to a prior mix guideline (that takes the form of target proportions of capacity to be served by each class of generation).
- 2) when its operation is simulated (according to economic loading criteria) the plan yields results that confirm the assumptions that were incorporated to determine the prior mix guidelines.
- 3) Now that the target mix has been determined, the next question involves the size of the next unit

to be brought on line. Because of economies of scale the tendency has been to acquire the largest possible under present technical capabilities moderated by the consideration that the larger the unit the more the required reserve for reliability purposes. One system has adopted 5% as the guideline for the size of new base load unit additions. The size and mix guidelines are combined to generate a unit "push down list" from which construction plans will be formed.

- 4) Last, the reliability criteria are used to trigger new construction from the push down list.

Two further points should be made:

- 1) The prior mix guidelines are designed to keep system expansion present value costs as low as possible without actually dealing with the complex optimization process that should be done. We will discuss this in the chapter on cost evaluation.
- 2) The reliability criteria are derived on the basis of measures that have been found to be historically adequate in their performance.

The previous explanation is only one example of the approaches that are taken to system expansion planning and it is illustrated in the "Interconnected New England Generation Study, Generation Task Force Report No. 4" of May 1971.

Lately, there have been attempts made to optimize

system expansion planning by using linear and dynamic programming approaches. Again, reliability targets are treated as hard constraints. In either approach there are problems concerning what the proper length of planning horizon should be. Also since it seems impossible to translate all the design considerations into these simple optimization techniques, the systems obtained from any of these approaches should be modified by judgement and experience before they are adopted and built.

An example of how this is done follows. The planning year is divided into 13 four week planning segments. The seasonal deratings of each unit, and the units placed on maintenance, are assumed known and constant over each of these periods. Since expected system demands (predicted on sensible weather assumptions) change over the year, a combined capacity and load model is designed for periods over which they are each expected to stay constant. Sometimes a week by week analysis is performed to assure reliability over specially critical summer weeks. Although any of these periods could be used for analysis, attention is usually focussed upon the "worst" week of the year where it is planned that no capacity will be placed on scheduled maintenance (if possible to schedule this way). The assumption is made that there will be sufficient slack at other times of the year to perform required system maintenance chores; because of seasonal load variations, this is usually a valid assumption. The

construction decisions for a given year are then based on the reliability analysis of this period. We will return to this later.

CHAPTER III

THE MEASURES OF GENERATION RELIABILITY

3.1 Overview

Presumably, expenditures to obtain adequate generation buy reliability for an electric system and what we are after is some measure of that extra reliability that is bought.

The first thing that strikes one is that reliability must be a probabilistically based concept; i.e., anything can happen, but an electric system is more reliable than another in the sense that objectionable situations or catastrophies are less likely to occur in the former than in the latter. A second notion of reliability is that it must in some sense be related to the objectionability of the contingencies that are likely to result; i.e., if a system is greatly protected against the worst accidents and allows for smaller insignificant difficulties to occur, most people would agree that it would then be more reliable than if it were protected against all the smaller difficulties and not against the graver ones, even though these may number few in comparison. This is another way of saying that if one buys something, the measure of what is bought should be related to its value in order to be useful for decision making. Our problem then becomes one of finding a measure of reliability such that by using it we can

measure the benefits that accrue due to that reliability.

Before we begin, we will repeat that we will only treat the issues relating to the reliability of the generation supply to serve the forecasted demanded load, the forced outages and the scheduled maintenance. Obviously, transmission system reliability and distribution system reliability play important roles in that the story is not complete without them; they are of ultimate import to the reliability of customer supply and the customer, naturally, does not care which part of the system is responsible for the mishap he experiences. We will not treat the transmission and distribution reliability issues in our costing section, however, we will discuss them in our benefits section. We believe that in spite of this partial limitation, that some of our results will be useful to those concerned with these issues in the sense that to some extent reliability can be treated separately at each level. Specifically, it is possible to plan separately for a reliable generation supply, a reliable transmission system and a reliable distribution system using some of the concepts we shall develop. In fact, these jobs are done separately now.

In the section where we treat customer benefits from increased reliability, customer distribution-system reliability will play an important role since it may, at times, be the limiting factor in increasing customer reliability. We will attempt to relate the importance of these other categories of expenditures in system reliability by their relative magnitude.

Over the years many approaches have been developed to treat system reliability issues; we will discuss a few non-mathematical ones and spend considerable time analyzing the ones that are mathematically based.

We will first discuss and critique the rule of thumb approaches that incorporate the following techniques:

- 1) A strict 20-30% extra reserve guideline.
- 2) A double-worst-fault style design.

(This approach would plan to design the system in such a way that it would survive the failure of its two most important components plus some residual protection.)

These are becoming less satisfactory now in that the stable industry conditions that led to such guidelines have changed because:

- 1) newer extremely large units are being planned and built.
- 2) higher forced outage rates are being experienced.
- 3) increasing interconnections with other systems (leading to centralized power dispatch) are changing the face of the industry.

In the face of these circumstances, new rules of thumb based on old sound, common sense, procedures are needed.

A double-worst-fault style approach would design the system in such a way that it could survive the failure of any two components plus a small percentage margin of protection.

It should be apparent that a percentage reserve

guideline need not bear any necessary relationship to the reliability or cost of a system. For example, a system which depends on two large plants for most of its reserve can, in a real sense, be less reliable than a system with a lesser but better distributed reserve margin. It also follows that if economies of scale are sufficiently large, that it may be cheaper to acquire larger amounts of reserve in large units rather than smaller reserves in a greater number of smaller units.

A recent unfortunate example points out the incorrectness of comparing systems by their percentage reserves. Consolidated Edison acquired the "Big Allis" (Ravenswood Unit #3 purchased from Allis-Chalmers, hence the name) unit at a time when it represented more than 15% of the company load. This unit has been an extremely poor performer, but even if it had not been such a poor performer, the satisfaction of spinning reserve criteria would have placed stringent conditions on Con Ed's operations, i.e., for every M.W. of Big Allis power the utility would have to find another of hot spinning reserve for protective purposes.

In summary, it should be apparent that a percentage reserve guideline offers too simple a way in which to rank order alternative system designs.

The most commonly used mathematical reliability measures include the LOLP, LOEP and FAD approaches. These are mostly used by system planners (very few in each utility) in

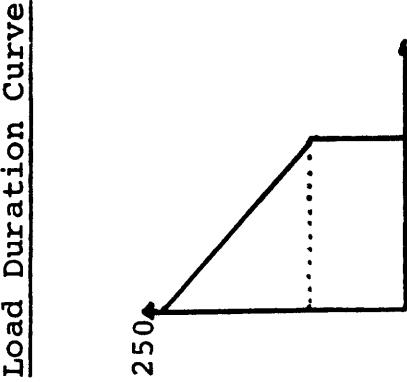
their design of the future system. These planners are supported technically in their work by G.E., Westinghouse and a few small consulting companies. It is well to point out that almost all the members of a company are not aware of this planning and if so, do not well understand the meaning of the typical company motto such as the "one day in 10 years" criterion for LOLP. Lest we be misunderstood, it should be noted that this planning is supported by less exact, rule of thumb heuristics at lower levels of an electric company which also do their planning. There is nothing wrong with this fractured pluralistic approach; the same situation prevails in most industries, strategic planners plan on one level, company salesmen and maintenance workers plan on another which hopefully accommodates their own problems in a more effective way. One of the goals of this work is to help each system's system planners understand how to relate these measures to their own load duration curves, customers, and reliability problems.

In our analysis we will examine each of the measures most commonly used (the loss of load probability, LOLP, frequency and duration, FAD, and loss of energy probability, LOEP, approaches) and will prescribe which ones, or combinations of them seem most adequate for measuring the increase of system "reliability". One measure may be deficient in certain respects; for example, if LOLP were the only measure used, great emphasis would be given to avoiding loss of load situations with stress on the fraction of time spent on outage but not on

how serious the outage situations may be in terms of load lost or energy not served when they occur. An excellent illustration of this occurs when one compares the LOLP and LOEP for the three systems (no maintenance) shown on the following page. LOLP says that the one generator system, System A, is preferred. Yet clearly, System B is to be preferred to System A; it is more reliable even though LOLP says it is not. The same reasoning shows that System C is also more reliable than System B. Note that the LOEP measures do capture this flavor.

The second question involves the rankings among alternatives that different measures yield. There is nothing to say that different measures, each sensitive to different problem parameters, will yield equal rankings among alternative systems. Specifically, if three systems are to be ranked, the measure m_1 may determine $S_1 > S_2 > S_3$; measure m_2 may rank them $S_3 > S_1 > S_2$.

Third, the different measures we will present use different models of the underlying physical situation. It would be surprising if they then yielded the same answers. In particular, the LOLP and FAD approaches use differing models to describe the loads demanded of, and capacities available to, the system. We cannot expect the same answers from them; however, suitably built models can be prepared so that the underlying physical situation models approximate each other in the relevant parameters. The measures may then be forced in some cases to yield directly comparable results.

	<u>Load Duration Curve</u>	<u>Capacity Model</u>	<u>LOLP</u>	<u>LOEP</u>
System A		one, 300 M.W. generator F.O.R. = .1 (forced outage rate)	.1	.1
System B		three, 100 M.W. generators F.O.R. = .1	$1 - (.9)^3 = .271$.025
System C		five, 60 M.W. generators F.O.R. = .1	$1 - (.9)^5 = .4$.011

A further complication is that in any given measure system, models can be built to describe the situation of greatest interest. For instance, LOLP models can be built for worst peak period behavior or for overall yearly behavior; they can also be built to incorporate load uncertainty or to leave this out. For instance, a LOLP figure calculated with regard only to the peakload forecast is unsatisfactory in that it disregards offpeak behavior; certainly one would be less concerned about a system whose daily peak load lasted one hour every day, than about a system whose daily peak load lasted 24 hours each day. The use of a daily peak load duration curve implies it is the latter situation that is of interest and which is being analyzed. The important thing is how to interpret the numbers obtained through any of these approaches.

The most useful reliability measure to a particular system would be one that would address those issues of greatest interest (or of broadly defined greatest cost) to the utility. From a societal point of view it would be best to use a reliability measure that would address itself to those issues of greatest (suitably defined) societal cost. In many instances, but not all, these concepts of cost might approximate each other. In this chapter we will attempt to prescribe what measure or combinations of measures seem to make sense for measuring system reliability in a way that optimizes the total costs to the society.

3.2 Calculation of the Measures of Reliability

There are four reliability measures that have been used for system planning purposes in the U.S. power industry:¹

- 1) The Loss of Load Probability (LOLP) measure.
- 2) The Loss of Energy Probability (LOEP) measure.
- 3) The Expected Loss of Load (XLOL) measure.
- 4) The Frequency and Duration (FAD) measure.

All of these measures are probabilistically based and yield differing, not strictly comparable, indices. The LOLP (called lollipop) method is the most commonly used and yields measures of the type "one day in 10 year loss of load probability", which is the criterion widely heralded by most of the industry as its goal.

The LOEP method is a variant of LOLP and gives an indication of what percent of system energy sales may not be met because of shortages; since the loss of energy probability is often close to zero it is convenient to refer to its ones complement, i.e., (1-LOEP), which usually is in the vicinity of 99+ percent.

The XLOL method, is again a refinement of LOLP and basically involves determining the mean outage size conditioned

¹Different kinds of models have been used for purposes other than long term planning; for instance, models other than the ones mentioned have been developed for short term contingency evaluation.

by the probability that an outage occurs, i.e., $[(\bar{L}-\bar{c})/(\bar{L}-\bar{c}) > 0]$.

The FAD method, on the other hand, yields measures such as "the mean recurrence time between outages will be seven years and when it occurs it will have a mean duration of three hours". In the following pages we will explain each method and later summarize their weaknesses.

Before we begin, we will point out that the first step in deriving these measures is to stipulate models that describe the situation of interest. It will be apparent that there are at least two situations that need to be modeled before determining the reliability of the available or future power supply over a period of time:

- 1) The available or future capacity situation.
(supply model)
- 2) The future load situation over the given period of time (demand model).

The reliability measure is obtained from combining both. The time unit usually considered for preparation of appropriate load and capacity models is the four week or 20 week-day period. We will discuss each of the methods in terms of these separate submodels for capacity and load and we will note that the models are in fact different and so should not be expected to yield the same values for their measures and, worse yet, need not necessarily produce equal rankings among a series of alternative power systems. Specifically, this means that if we consider the measures m_1 and m_2 , and the systems S_1 , S_2 and S_3 , then m_1 may for example produce a $S_1 > S_2 > S_3$ ranking and measure m_2 may on the other hand produce a $S_2 > S_3 > S_1$ ranking.

3.2.1 Scope of the Measures

All of the aforementioned measures provide probabilistic measures of the ability of the system's generating capacity to meet:

- 1) random forced outages.
- 2) random forecasted expected demand.

Unit maintenance is assigned during periods of low demand and we make the assumption for our planning purposes that there is sufficient off peak period slack to schedule all of the necessary maintenance. From this point on we will assume that we determine yearly expansion schedules by focusing on the behavior of the most critical maintenance period when the assumption is made that none is scheduled.

An inherent assumption in the models is that the capacity and load models are independent of each other, and that the generators are also independent of each other. It is then simple to see that the measures do not explicitly or implicitly deal with reliability considerations related to:

- 1) construction schedule slippage or other delay in bringing units on line.
- 2) incorrect model parameter specification such as specifying forced outage rates which may or may not be accurate.
- 3) incorrect description of the outage process as an independent one. For example, it may be that

capacity outages, especially in peak load periods, are correlated and not independent as the models assume; periods of stress may cause unwise usage of equipment which may lead to causally related failures.

We cannot expect to have these reliability measures deal with the issues above and therefore these considerations must be dealt with in other ways. For example, equipment delays may be handled by planning to have the equipment six to twelve months before it will be needed. Also the measures should be tested for their sensitivity to the forced outage rate assumptions.

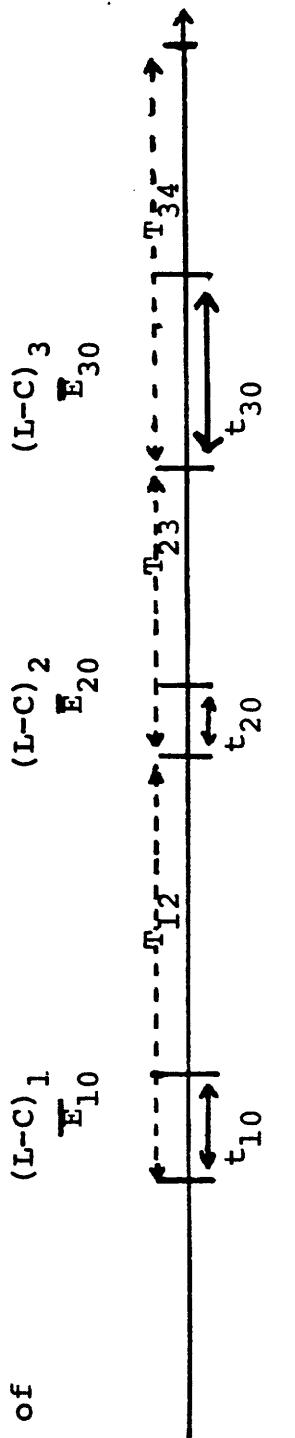
Another problem outside our scope is that of correct specification of generator forced outage rates; this is a tricky business in at least two respects. First, the assumption is made that there is sufficient operating history with a particular generator type to derive a reasonable forced outage rate that describes its behavior. In reality, design modifications and size increases combine to make these estimates less firm than otherwise possible. Second, forced outage rate determination is not as straight forward an exercise as might be thought. A naive assumption would be to equate a unit's F.O.R. with the fraction of time it is "down", but it should be apparent that this in turn depends on the resources the firm has to repair its units and on its need to repair it. For example, a base loaded unit will be repaired as quickly as

possible, whereas a gas turbine may not be repaired for months. What is really sought after is an index of non performance, i.e., the F.O.R. should be thought of as the fraction of time a unit was unavailable when it was expected to run. In the rest of our work we will assume they have been correctly given to us.

The measures are not well named; LOLP and LOEP are more expectations rather than probabilities. Roughly speaking, the measures we will discuss measure:

- 1) LOLP - the expected fraction of time the utility system will have a generation deficit with no consideration given to how large that deficit might be.
- 2) LOEP - the expected fraction of total energy sales the utility will not make due to generation shortages.
- 3) XLOL - the expected value of the generation deficit, given that one has occurred; however it does not give consideration to how long this deficit lasts.
- 4) FAD - the expected mean recurrence time between generation deficit events and their mean duration when they occur. Note that FAD as well as LOLP do not consider the size of the shortage when it occurs.

We will first discuss techniques for computing the



$LOLP \approx$ average fraction of time short $\approx \frac{1}{n} \sum_i \frac{t_{io}}{T}$

$LOEP \approx$ average fraction of energy unserved $\approx \frac{\sum_i E_{io}}{\sum_i E_{tot}}$

$XLOL \approx$ average shortage size given that a shortage has occurred $\approx \left(\frac{n}{\sum_{i=1}^n (L-C)_i} \right) \cdot 1/n$

FAD: $D \approx \left(\sum_{i=1}^n t_{io} \right) \cdot 1/n$, average t_{io}

$F = \frac{1}{\left[\sum_{i=1}^n T_{i,i+1} / n \right]} \approx 1/T$ average or a frequency of recurrence

$F.D \approx \frac{\sum_{i=1}^n T_{i,i+1}}{\sum_{i=1}^n T_{i,o}} \approx \frac{\text{time on shortage}}{\text{total time}}$

Figure 3 Reliability Measures at an Intuitive Glance

measures which are direct but computationally intractable. Later we will introduce a technique named probabilistic simulation which turns the computationally exhausting combinatorial problem into a recursive one.

3.2.2 The Loss of Load Probability Method

As previously mentioned, LOLP is the most widely used method in industry. The method was first in historical development and was a consequence of the work of Calabrese and Lyman, among others in the 1930's. The method is conceptually uncomplicated and yields answers with relative computational ease; we will proceed to discuss the capacity and load models respectively.

3.2.2.1 The Capacity Model

In its simplest form, the capacity model consists of modelling each of N generators, g_i , by its nameplate capacity, c_i , and characterizing the capacity it has in the system as a random variable \tilde{c}_i which has value 0 with probability equal to its forced outage rate, and value c_i with probability equal to the complement of its forced outage rate.

$$\tilde{c}_i = \begin{cases} 0, & P(\tilde{c}_i = 0) = F.O.R. = p_i \\ c_i, & P(\tilde{c}_i = c_i) = 1 - p_i \end{cases}$$

(capacity available to the system from generator, g_i)

The item of interest is of course the total capacity available to the system:

$$\tilde{c}_{\text{tot}} = \sum_{i=1}^N \tilde{c}_i$$

\tilde{c}_{tot} naturally is another random variable which describes the total capacity available to the system; if we make the further assumption of statistical independence between the random variables \tilde{c}_i describing each generator, then the probability distribution of \tilde{c}_{tot} is simply obtained by convolving the probability distributions for each of the \tilde{c}_i . (Reference: any standard introductory probability text.) The resulting distribution will extend between $\tilde{c}_{\text{tot}} = 0$ and $\tilde{c}_{\text{tot}} = \tilde{c}_i$ at its maximum point; the distribution will be nonzero only at a finite number of values in between this range which represent possible \tilde{c}_{tot} capacity states. Note that if all the c_i are equal, there will be exactly $(N+1)$ possible values for \tilde{c}_{tot} . Also note that at most there can only be 2^N values for \tilde{c}_{tot} and that this maximum need not be attained even if all the c_i are different.

Basically, at each feasible value of \tilde{c} , c_f , we compute the probability for all those events whose capacity sums to c_f and this involves a convolution of the present density function under consideration with that of the unit being added on. In this recursive fashion we develop the \tilde{c}_{tot} density function.

$$p(c) = \sum_{c_i=0}^{c_i=c} p(c-c_i)p(c_i)$$

This is in the form

$$f(c) = \int_0^c f(c-\tau) f_{c_i}(\tau) d\tau = f * f_{c_i}$$

Example 1 Equal Size and Forced Outage Rate Units.

system is 5 machines, $N = 5$.

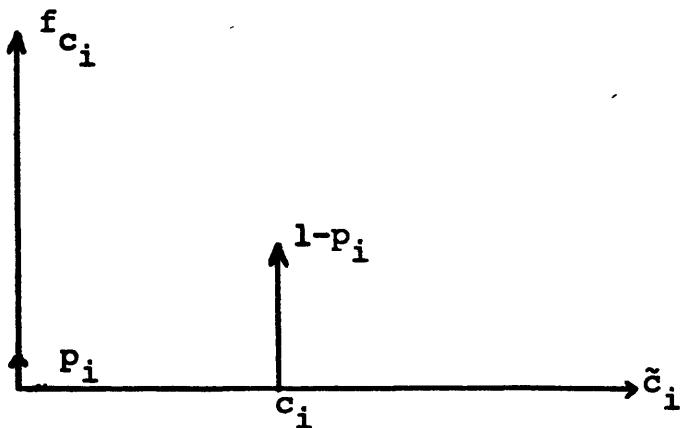
forced outage rate, $p_i = .01$, each generating unit.

capacity, $c_i = 40$ M.W. for each unit.

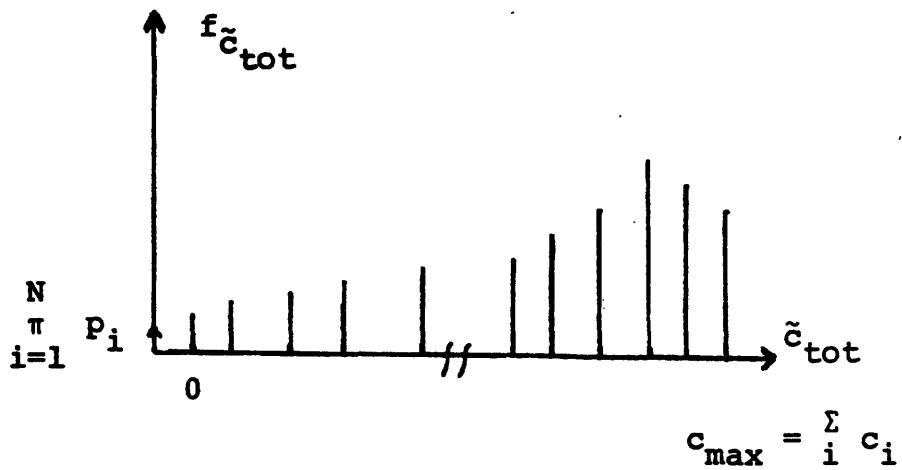
Compute probabilities for various outage events:

<u>c_o</u>	<u>$c_{in} = (c_{tot} - c_o)$</u>	<u>P_e (probability of exact event)</u>
0	200	$(.99)^5$, all are up $\approx .95$
40	160	$5 (.99)^4 (.01)$ $\approx .048$
80	120	$\binom{5}{2} (.99)^3 (.01)^2$ $\approx 10^{-3} = .00097$
120	80	$\binom{5}{3} (.99)^2 (.01)^3$ $\approx 10^{-5}$
160	40	$\binom{5}{4} (.99)^1 (.01)^4$ $\approx 10^{-8}$
200	0	$(.01)$ $\approx 10^{-10}$

$$\tilde{c}_{\text{tot}} = \sum_{i=1}^N \tilde{c}_i$$



$$f_{\tilde{c}_{\text{tot}}} = f_{c_1} * f_{c_2} * \dots * f_{c_N}$$



Note that if all the \tilde{c}_i are equal there are $(N+1)$ capacity values. Note that at most there are 2^N different values one for each of the possible on-off arrangements of the machines.

Figure 3.1

The Probability Density of \tilde{c}_{tot} , a Sum of Independent Random Variables, \tilde{c}_i .

We recognize that for equal capacity machines the problem is just computing the terms in a binomial expansion

$$(.99+.01)^5 = \binom{5}{0} (.99)^5 + \binom{5}{1} (.99)^4 (.01)^1 + \binom{5}{2} (.99)^3 (.01)^2 + \binom{5}{3} (.99)^2 (.01)^3 + \binom{5}{4} (.99)^1 (.01)^4 + \binom{5}{5} (.01)^5$$

For the general problem where all machines are not equal then there is no simple recourse. If some machines have equal capacity it is easy to form the density on capacity available from that group in the manner above and then convolve with all the other densities.

Each value of \tilde{c}_{tot} represents an outage event among the machines with its respective probability (or probabilities if more than one outage event results in the same capacity value; these would then be added since each outage event is an independent event). For instance, the 130 M.W. event occurs if and only if machines 1 and 2 are in service and machine 3 is out of service. The probability of this event is the product of the probabilities that machines 1 and 2 are up and machine 3 is down. See Figure 3.2.

It is possible to give better, more fine grained, representations of the available generators by representing their partially derated states through the use of suitable probability mass functions. We will not go into this in great detail except to sketch out the differences involved. Basically, the only change is to represent each generator, g_i , contributing capacity, \tilde{c}_i , a random variable describing each

Example 2 Differing Size and Forced Outage Rate Units

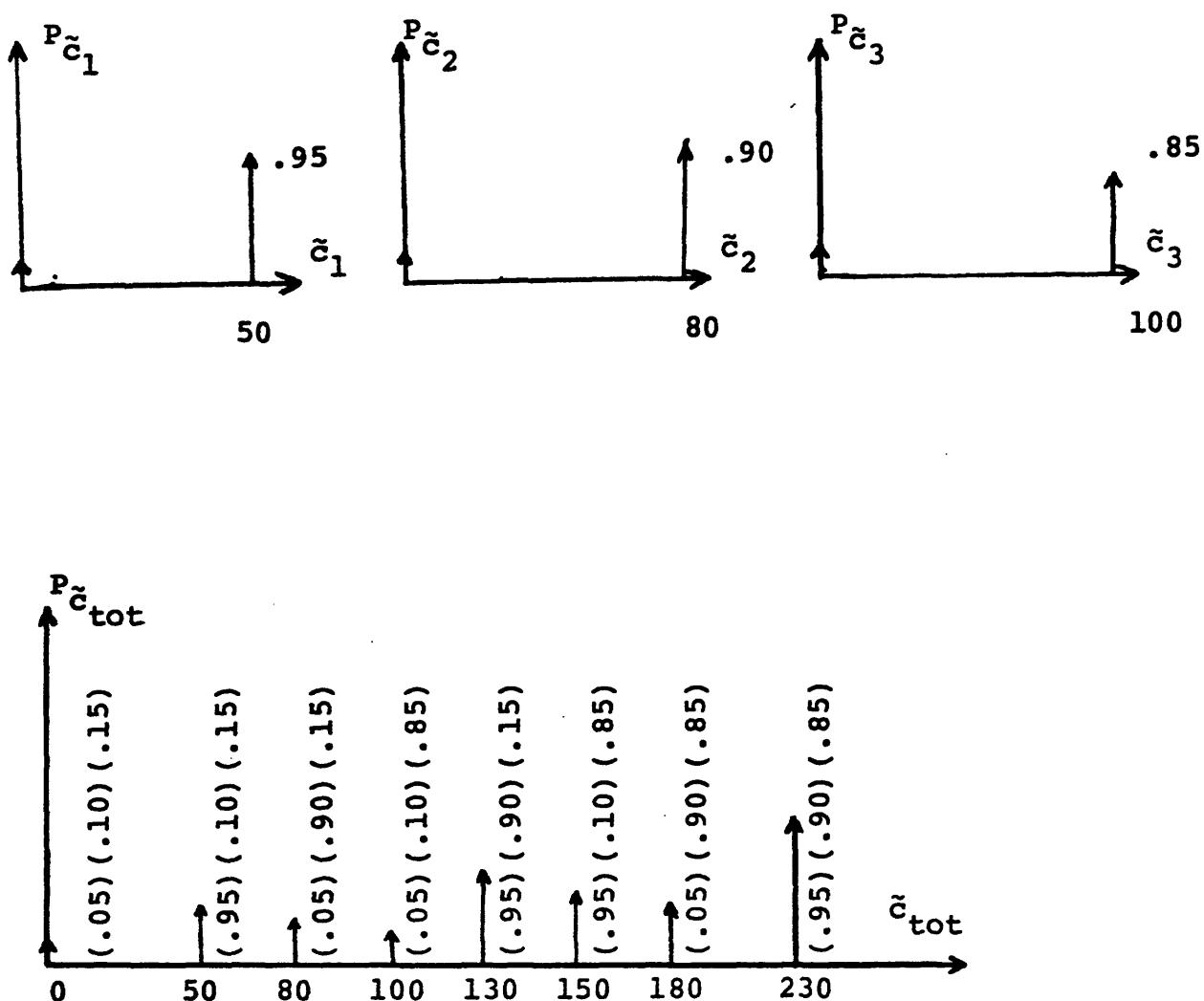


Figure 3.2

The Probability Density Function of \tilde{c}_{tot} , a Specific Example.
 [This could be obtained simply by forming $p_{\tilde{c}_1} * p_{\tilde{c}_2} * p_{\tilde{c}_3}$, where
 * is the convolution operation].

of its partially derated states. The probability mass function is nonzero at each of these possible derated capacity values and its value at each point is the long term fraction of time the generator spends in that derated state; this is obviously a suitable probability mass function since it sums to one.

3.2.2.2 The LOLP Model

There are several ways of modeling the load in an LOLP approach and people switch from one to another depending on the items of interest. Ultimately, the goal is to combine the capacity model previously derived with an appropriate load model to obtain a measure of the reliability of the system, i.e., the adequacy of the amount of generation available to serve:

- 1) an anticipated expected load.
- 2) to support the expected but unknown future forced outages.

Take the simplest case first. Let us assume that the load on the system is constant at L_o and is therefore perfectly predictable. The system loss of load probability would simply be the probability of all these outage events that could leave the system with effective capacity smaller than L_o . In other words, if $P_{cum}(0_k)$ is the cumulative probability of outages being greater than 0_k , and

$$M = c_{tot} - L_o$$

(M is the system margin by definition)

where L_o is the expected constant load to be served and c_{tot} is the total available capacity. Then LOLP is the probability that system outage is greater than system reserve margin.

$$LOLP = P_{cum}(M)$$

or

$$LOLP = \sum_{(L_k > M)} P(L_k)$$

However, the load is not usually known with such precision; in fact, the load could be any of a number of possibilities. There are two major ways to represent this uncertainty in the forecast period, through the:

- 1) hourly load duration curve (LDC) over the forecast period or
- 2) the daily peak load duration curve (DPLDC).

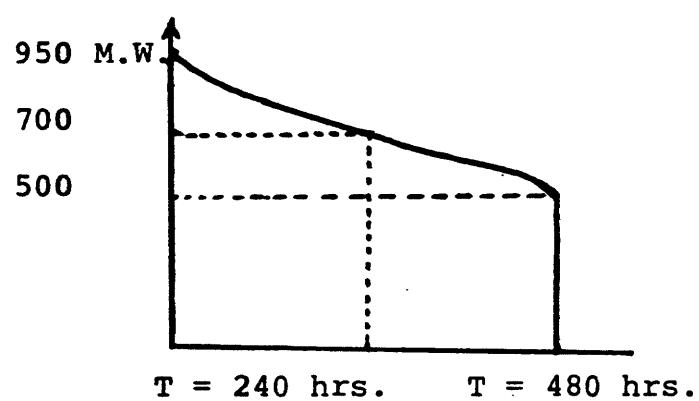
3.2.2.3 Hourly Load Duration Curve

An hourly load duration curve is obtained conceptually by first plotting on the vertical axis the M.W. demand forecasted for each hour in a planning period in chronological order along the horizontal axis. This simply is a

chronological hourly load forecast for the planning period. The LDC is obtained by organizing the previous chart in descending demand level order (see Figure 3.3). Assume the top forecasted demand occurs for one hour during each of the days in a 20 day planning period. Then we can say that the peak load occurred in $20 \text{ hrs} / 24 \cdot 20 \text{ days} = \text{fraction of the period}$. It then becomes a simple matter to combine load information with system outage information to obtain the system measure. For example, assume that there is an outage O_k which leads to system loss of load only when the peak load occurs; it is easy to see through use of the LDC that the outage will lead to a loss of load only $1/24$ of the time in the planning period.

Figure 3.3 expresses, among other things, the fact that the system's load was expected to be above 700 M.W. during 240 hours of the period, i.e., 50% of the time. If we normalize the horizontal axis to 1, the chart can be read as saying what percent of the time the load would be expected to be above a given level. So far, the LDC has been treated as a deterministic forecast. At this point it is simple to give the LDC a probabilistic interpretation and use it for determining LOLP measures; the curve can be read probabilistically to say that the load will be above 700 M.W. with probability equal to one half. If the LDC is turned counterclockwise 90° and then mirror imaged, the resulting curve gives the cumulative probability that the load is greater than or equal to the value on the abscissa; the load will be greater than 500 M.W.

M.W. of Load



(horizontal axis is normalized to 1 for LOLP calculation)

Figure 3.3 Hourly Load Duration Curve

with probability one and will be greater than 950 M.W. with probability 0.

The system LOLP is the probability of the event that capacity is smaller than load ($\tilde{c} < \tilde{L}$). One way to compute it is to compute for each value of \tilde{c} its contribution to LOLP. Since each \tilde{c} event is independent, and since in the aggregate they are collectively exhaustive, the sum of these terms is the system LOLP.

$$\text{LOLP} = \sum_{\substack{\text{over possible} \\ \text{values of } \tilde{c}}} P(\tilde{c}) \cdot P(\tilde{L} > \tilde{c})$$

For the example of Figure 3.3, the contribution to LOLP of the event capacity = 700 M.W. is just

$$\begin{aligned} P(\tilde{c} = 700 \text{ M.W.}) \cdot P(\tilde{L} > 700 \text{ M.W.}) \\ = (1/2) P(\tilde{c} = 700 \text{ M.W.}) \end{aligned}$$

Note that this is simply a type of convolution integral. The problem is that we are investigating the behavior of $(\tilde{c} - \tilde{L}) = \tilde{m}$, the margin random variable; we are interested in the probability that $\tilde{m} < 0$. Since \tilde{c} and \tilde{L} are independent r.v.'s, the density of $\tilde{c} - \tilde{L}$ is just the convolution of density of \tilde{c} with the density of $(-\tilde{L})$.

$$f_m = f_c * f_{(-L)}$$

This is the basic formulation of a LOLP load model and

of its use in calculating LOLP measures. We now formalize this treatment in terms of the system outage probabilities previously discussed in the capacity model section. Look at Figure 3.4 below. It is easy to see that any outage, 0_k will produce an outage during a fraction of $T(0_k)$ of the planning period. We plot $P(0_k)$ along the vertical axis of the LDC with $\bar{0}_k$, the complement of \bar{c} , increasing as it nears the horizontal axis. This value of 0_k represents the event in which the total installed capacity is out. In this formulation it is easy to see that:

$$LOLP = \sum_k P(0_k)T(0_k)$$

It is also possible to express LOLP in terms of the cumulative probability of outage, i.e., $P_{cum}(0_k) = \sum_{outages > 0_k} P(0_k)$. We can express the above as:

$$\begin{aligned} LOLP = & P(0_1)T(0_1) + P(0_2)[T(0_1) + T(0_2) - T(0_1)] \\ & + P(0_3)[T(0_1) + (T(0_2) - T(0_1)) + (T(0_3) - T(0_2))] \\ & \vdots \\ & + P(0_N)[T(0_1) + \dots + (T(0_N) - T(0_{N-1}))] \end{aligned}$$

if

$$D_{j-1,j} = T(0_j) - T(0_{j-1})$$

$$P(0_j) = p_j$$

$$T(0_j) = t_j$$

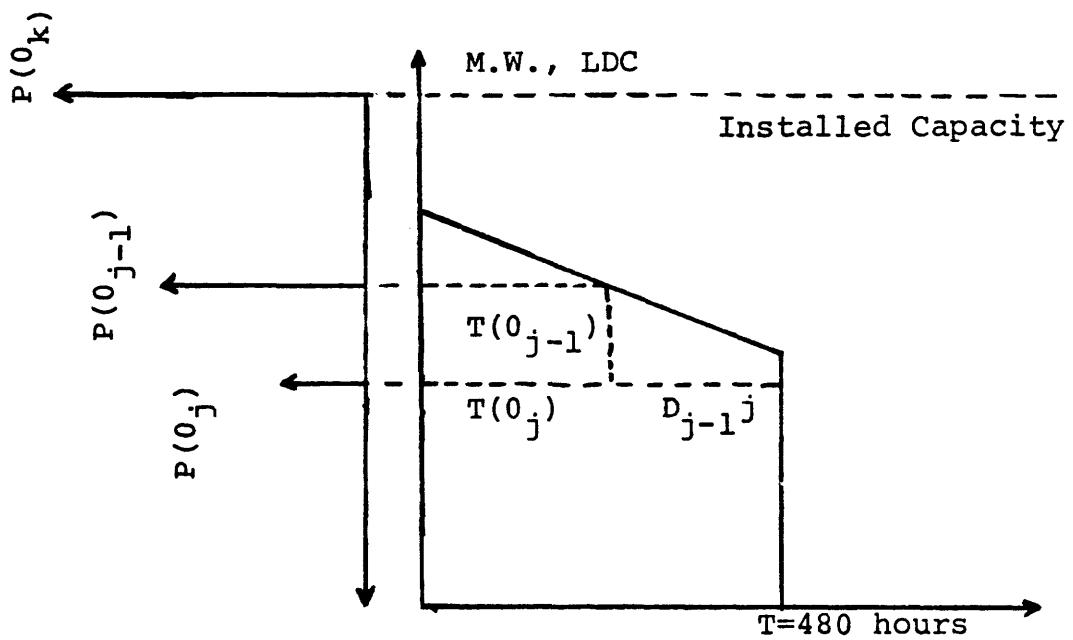


Figure 3.4

(horizontal axis normalized to 1 for LOLP calculation)

Calculation of the LOLP Measure

then

$$LOLP = P_1 T_1 + P_2 (T_1 + D_{12}) + P_3 (T_1 + D_{12} + D_{23}) + \dots + P_n (T_1 + \dots + D_{n-1,n})$$

since

$$P_{\text{cum}}(0_k) = \sum_{0 > 0_k} P(0_k)$$

$$LOLP = P_{\text{cum}}(0_1) T_1 + P_{\text{cum}}(0_2) D_{12} + \dots + P_{\text{cum}}(0_n) D_{n-1,n}$$

$$LOLP = \sum_{k=1}^{k=m} P_{\text{cum}}(0_k) D_{k-1,k} \text{ and } D_{0,1} = T_1$$

which we will find useful since it is in terms of the cumulative outage distribution function.

3.2.2.4 Daily Peak Load Duration Curve

Often it is of interest to concentrate attention on the behavior of the system during peak load hours. It is a simple matter to prepare what is called a daily peak load duration curve (DPLDC) which can be easily obtained from the hourly load duration curve. The procedures that are used to obtain system LOLP's are identical but care must be given to the interpretation of these measures vis a vis those obtained from an hourly load duration curve.

Specifically, let us assume that, $LOLP_{\text{DPLDC}} = f$. This means that a fraction f of the time represented in the DPLDC

of say 20 days and their respective peak loads, will be spent with a loss of load. If $f = 1/2600$, i.e., one day in 10 years, this means that on the average the system will be short in one peak load hour out of 10 years worth of peak load hours. Note that the same value for the hourly LDC LOLP means that the system will be short for one day (24 hours) in 10 years (2600 days with 24 hours each). We will return to this point in the critique section.

3.2.2.5 Other Methods

It is also possible to develop a load model which allows for expression of uncertainty in the forecast loads; note that the LDC so far discussed is essentially a deterministic forecast and that the probabilistic property of LOLP so far consists of describing the probability that a given probabilistic capacity event would combine with a given load event to cause a capacity shortage. As one can easily see, the DPLDC states that the maximum load will be $L_p^M.W.$ and there is no provision for describing the uncertainty in that forecast.

One way to handle this is to concentrate on the maximum daily peak load event, describe its value probabilistically \tilde{L}_p through the density function $f_{\tilde{L}_p}$, and convolve this with the capacity density function to obtain the probability that $\tilde{c} < \tilde{L}_p$. This procedure essentially equates LOLP with the system's largest LOLP in the planning horizon.

A slight variation on this approach is to model the density of the expected peak load as a normal curve. The peak expected demand in the period is modeled as the first order statistic for the number of sample peak loads generated for that period (in a 20 day period the number is 20). This information is combined with some assumptions about the standard deviation of the distribution to construct a parametrically based load model which attempts to describe peak load statistics. The number obtained from this approach is one closer to the numbers generated from assuming the peak load is deterministically known at a given level than they are close to numbers generated by DPLDC or LDC application.

Another possible way of incorporating this feature is to calculate an LOLP for each of a given possible set of system LDC's. The resulting LOLP's should be weighted by the probability of each LDC to obtain the overall system LOLP.

$$\text{LOLP} = \sum_{\text{LDC forecasts}} (\text{LOLP/LDC forecast}) P(\text{LDC forecast})$$

A third possible approach would be to designate densities around forecast peak loads for each day on the DPLDC, and then add up each day's contribution to system LOLP.

Note that each of these ways differs considerably from the rest and has its particular strengths and drawbacks. The optimal choice for a given system would depend on the system's particular features and its consumers' interests.

Note that we have so far disregarded maintenance scheduling and that we have assumed nameplate capacities to be the valid ones for LOLP calculations. It should be apparent that an LOLP calculation should be made for each period over which maintenance is fixed. Also, if because of some malfunction or seasonal effect, a partial derating is prescribed for a generating unit's operation, the LOLP calculation should take this into account in the capacity model and not simply use the unit's nameplate capacity.

Now, preventive maintenance scheduling can be done with reasonable latitude; a natural result of the LOLP analysis would be for the system planner to have as an objective to distribute maintenance scheduling in a way to keep the LOLP measure constant throughout the year. This would follow if:

- 1) the penalties on being short of capacity were independent of timing within the planning horizon; for example, if they were seasonally unaffected.
- 2) the penalties tended to escalate in a faster than linear relationship with respect to deviations from the target LOLP measure.

If property 2) held, then it would always be worse for the system to have one period carry a higher LOLP measure in order to allow lower LOLP measures during other periods (of course, this is strictly true if and only if capacity is available in continuous fashion and not in discrete chunks, i.e., it may physically be impossible to schedule maintenance to

keep a constant LOLP). If on the other hand, it could be demonstrated that there were larger penalties for being short during one time of the planning horizon rather than another, if for example summer shortages carried higher penalties than fall shortages, then it would probably be better to schedule maintenance to secure lower LOLP measures during the summer thus causing higher ones during the fall seasons.

3.2.3 Loss of Energy Probability Measure

The LOEP method is a variation of the LOLP method in that it uses essentially the same load and capacity models to derive a measure related to the expected fraction of energy sales lost due to capacity shortages. This measure is useful in that it can be related directly to the revenues collected by the utility and by extension to the purchases made by consumers.

For systems with reasonably high reliabilities, it should be obvious that even if load is shed from time to time the actual energy sales curtailed are a small percentage of total period sales; this is due to the fact that the energy sales lost are related to the time spent on outage and the amount of shortage and both of these are small related to total period time and average level of demand. Because of this smallness, one usually speaks of $(1 - \text{LOEP})$, and refers to it as the Energy Index of Reliability (EIR), the probable percent

of energy demand that will be fulfilled.

For the purposes of computing LOEP, a full system, hourly basis, LDC is used. The capacity model developed for LOLP is retained. The height of the LDC curve is the expected load demanded, the horizontal axis represents the time over which that load is demanded. It is easy to see that the area under the curve is in fact the expected energy demanded.

For any given outage event, 0_k , it is easy to see, by using the LDC, that a given amount of energy will not be served (see Figure 3.5).

The LOEP measure is obtained by adding for each possible outage event its contribution to LOEP:

$$\text{LOEP} = \sum_k P(0_k) \frac{E_k(0_k)}{E_{\text{tot}}}$$

alternatively we can calculate LOEP in terms of the cumulative outage distribution function

$$\text{LOEP} = \sum P_{\text{cum}}(0_k) \frac{DE_{k-1,k}}{E_{\text{tot}}}$$

Note that this is clearly a measure of expected energy not served and not a probability.

It is easy to give LOEP a simple physical interpretation; one can interpret the probability of an outage, 0_k , causing a loss of energy sales, $E(0_k)$, as the percent of time in

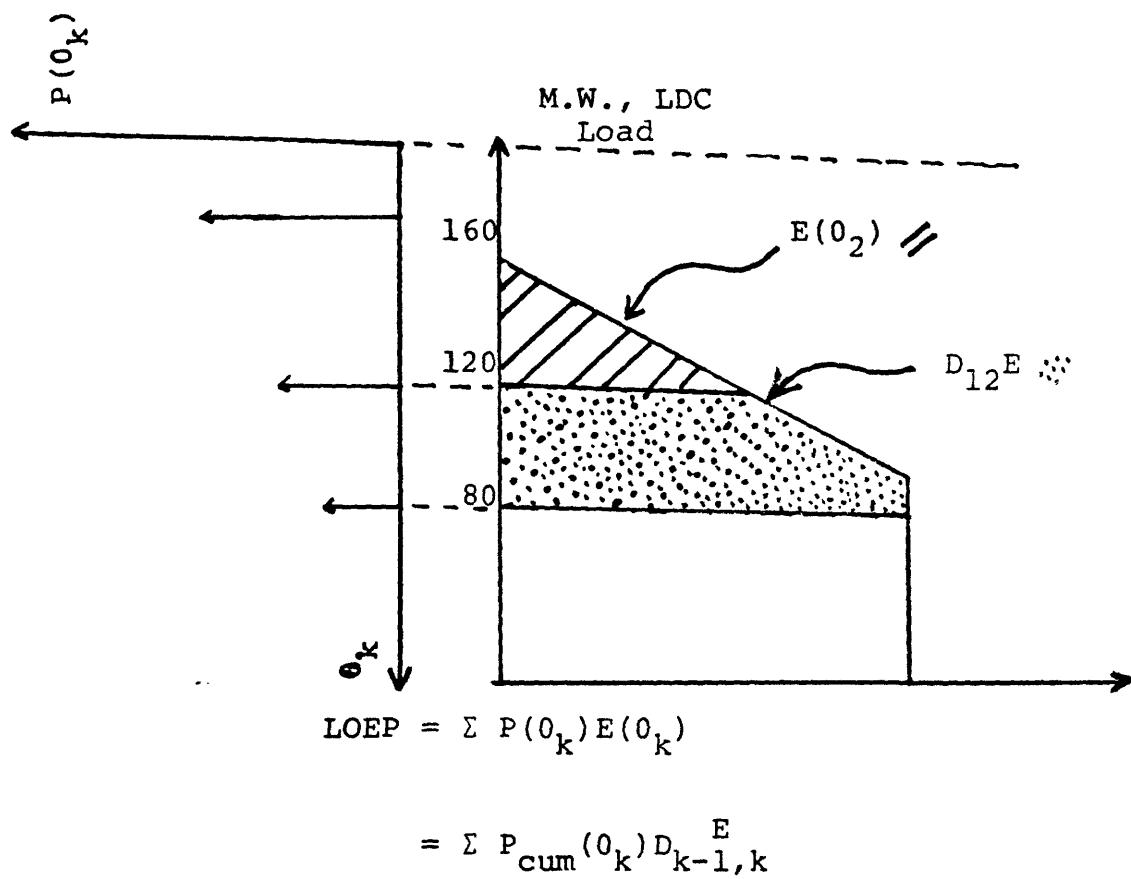


Figure 3.5 LOEP Measure Calculation

the planning period in which an expected fraction $E(O_k)/E_{tot}$ of total energy sales would be lost. By summing these terms over all possible outages we cover the total time period and can interpret LOEP as the fraction of energy sales unserved in the period.

To understand why LOEP is usually small, consider a typical loss of load situation. The energy sales lost during that loss of load are roughly proportional to the (loss of load - margin) and the period of time involved. Assume that the system's LOLP is .1 days per year; assume further that on the .1 days/year of loss of load, the loss of load is a fraction f of peak load (note that this is a reasonable assumption but no information that has any bearing on it is included in the LOLP index, however, such information can be devised from the LOLP load and capacity models). The total energy sales for the year are in the order of $260 \cdot 24 \cdot \bar{L}$ (average hourly load). If we assume the .1 days/year of loss of load occur at peak load, and that the loss is complete for that period, it follows that the energy sales lost are given by

$$E.S._{lost} = \left(\frac{.1}{260}\right)(260)(24)L_p$$

if, however, only a fraction f of L_p is lost during that period of load loss then

$$E.S._{lost} = \left(\frac{.1}{260}\right)(260)(24)fL_p$$

$$0 \leq f \leq 1$$

The LOEP = (lost energy sales)/(total energy sales) and we see that

$$\text{LOEP} = \frac{\text{E.S.}_{\text{lost}}}{\text{E}_{\text{tot}}} \approx \frac{(.1/260)(260)(24)f_L}{(260)(24)\bar{L}} p = (\text{LOLP}) \cdot f \cdot \frac{L_p}{\bar{L}}$$

Now, we don't expect to lose all of our peak load during the loss of load time period (i.e., $f \neq 1$ and usually $f < 20\%$) but even if this were the case, L_p/\bar{L} is seldom more than 2 and this would still leave a LOEP which is roughly .4 LOLP which can be seen to be small. If LOLP = .1 days/year then

$$\text{LOEP} \leq (\text{LOLP})(.2)(2) \rightarrow \text{LOEP} \leq (.4)\text{LOLP}$$

or

$$\begin{aligned} \text{LOEP} &\approx .4 \times 1/2600 \approx .4 \times 3.85 \times 10^{-4} \\ &\approx 1.54 \times 10^{-4} \end{aligned}$$

1-LOEP=EIR (Energy Index of Reliability)=.999846 or 99.9846%

3.2.4 The Expected Loss of Load Measure

The expected loss of load (XOL) measure has recently been developed in an effort to supplement the deficiencies in the LOLP measure, namely that while including information regarding the fraction of time loss of load events can be expected to occur, LOLP does not include information regarding the

possible size of system generation deficiency. XLOL is defined to be the expected value of generation deficiency given that a loss of load has occurred.

$$XLOL = \overline{[(\tilde{L}-\tilde{c}) / (\tilde{L}-\tilde{c}) > 0]}$$

The density for $(\tilde{L}-\tilde{c})$ conditioned on $[(\tilde{L}-\tilde{c}) > 0]$ is simply

$$f_{\tilde{L}-\tilde{c}/(\tilde{L}-\tilde{c}>0)} = \frac{f_{\tilde{L}-\tilde{c}}}{P(\tilde{L}-\tilde{c}>0)} \begin{cases} (\tilde{L}-\tilde{c}) > 0 \\ 0 \text{ elsewhere} \end{cases}$$

Then it follows that

$$XLOL = \frac{\int_{-\infty}^0 (\tilde{L}-\tilde{c}) f_{(\tilde{L}-\tilde{c})}}{LOLP}$$

There are certain problems with this measure taken as it is. First, it is an absolute magnitude and not related to other system parameters such as expected peak load; certainly an XLOL of 1000 M.W. is more significant in a 10,000 M.W. system than in a 30,000 M.W. system.

Second, XLOL can be shown to be a function of LOLP and LOEP, in particular

$$XLOL = \frac{LOEP}{LOLP} \cdot E_{tot}$$

We will shortly provide a proof of this result.

Third, it should be clear that XLOL alone cannot be a

relevant measure of system reliability (and its authors agree) since it yields no information regarding the extent of time this may be expected to occur. It can be useful though through providing another interpretation of the significance of the LOLP and LOEP measures, and through providing another technique for LOEP calculation.

Proof

$$XLOL = \frac{LOEP \cdot E_{tot}}{LOLP}$$

where

$$E_{tot} = \int_0^{MA} F_o(x) dx, F_o(x) = LDC$$

The "energy" to be produced over the load duration curve.

Note it is not true energy until multiplied by the time period involved, i.e., a 20 hour LDC will produce twice the energy of a 10 hour LDC but they will both have the same E_{tot} . This will be more intelligible once the reader has gotten through the Probabilistic Simulation Section.

By definition

$$XLOL = 1/LOLP \sum_{x<0} P_m(x) (-x)$$

$$= 1/LOLP \sum_{x>0} P_{-m}(x)x$$

The divisor, LOLP, is due to the fact that XLOL is the expected generation deficit conditioned on a loss of load event occurring, but since

$$[IC - (\tilde{\lambda} + \tilde{c}_o)] = \tilde{m}$$

$$-\tilde{m} = (\tilde{\lambda} + \tilde{c}_o) - IC$$

This can be written as

$$XLOL = 1/LOLP \sum_{x=IC}^{x=-\infty} (x-IC) P_{-m}(x-IC)$$

but

$$P_{-m}(x-IC) = -F_n(x)$$

where

$$F_n(x) \text{ is the } P(\lambda + c_o > x)$$

and therefore

$$-F_n(x) = P(\lambda + c_o = x)$$

since the event " $\lambda + c_o$ equal to x " is the same as the event " $-m$ equals $(x-IC)$ " changing to the continuous case

$$XLOL = \int_{x=IC}^{x=\infty} (x-IC) [-F_n(x)] dx$$

We integrate this expression by parts and let

$$(x-IC) = u$$

$$-\dot{F}_n = \dot{v}$$

$$\dot{(uv)} = \dot{uv} + v\dot{u}$$

$$\dot{vu} = \dot{uv} + (\dot{u}\dot{v})$$

$$(x-IC) \dot{(-F_n)} = +[(x-IC) \dot{(-F_n)}] - (-F_n) \dot{(1)}$$

$$XLOL = 1/LOLP \int_{x=IC}^{x=\infty} -(x-IC) \dot{F}_n(x) dx$$

$$= 1/LOLP \int_{x=IC}^{x=\infty} \{ F_n(x) - [(x-IC) \dot{F}_n(x)] \} dx$$

$$= 1/LOLP \int_{x=IC}^{x=\infty} F_n(x) - 1/LOLP [(x-IC) F_n(x)] \Big|_{IC}^{-\infty}$$

but the second term is 0 when evaluated at $x=IC$ and as x increases toward ∞ , $F_n(x)$ descends faster than x so that this too tends to 0 (it must be the case or else $F_n(x)$ would not be integrable) and therefore

$$X_{LOL} = 1/LOLP \int_{x=IC}^{x=\infty} F_n(x) = 1/LOLP (\bar{EU}_n)$$

$$= \frac{LOEP \cdot E_{tot}}{LOLP} \quad Q.E.D.$$

E_{tot} can be expressed as a load factor times the peak load MA

$$X_{LOL} = \left(\frac{LOEP}{LOLP} \right) (LF) (MA)$$

Another proof is given below

$$X_{LOL} = 1/LOLP \sum_{x<0} P_m(x)(-x) = 1/LOLP \sum_{x>0} P_{-m}(x)x$$

$$= 1/LOLP \sum_{x>0} P_{cum-m}(x)(\Delta x)$$

$$= 1/LOLP \int_0^{\infty} P_{cum-m}(x)dx$$

We now need to know the cumulative distribution function for negative margin.

$$\text{But by definition margin} = IC - (l+c_o)$$

$$-m = (l+c_o) - IC$$

and we see that the event

$$-m = x$$

implies the event

$$l+c_o - IC = x$$

$$l+c_o = x+IC$$

But we have the cumulative distribution function for $(l+c_o)$, it is $F_n(x)$ therefore

$$P_{\text{cum-m}}(x) = F_n(x+IC)$$

i.e., the cumulative on negative system margin is simply the cumulative on equivalent system load displaced to the left by IC.

Then

$$XLOL = 1/LOLP \int_0^{\infty} F_n(x+IC) dx$$

$$= 1/LOLP \int_{IC}^{\infty} F_n(\tau) d\tau = \frac{LOEP \cdot E_{\text{tot}}}{LOLP} \text{ Q.E.D.}$$

3.2.5 The Frequency and Duration Approach

The Frequency and Duration (FAD) Method incorporates further data about generators on a system and yields measures which differ from those given by LOLP. Basically, the FAD method uses a Markov model to characterize the capacity of the set of machines that are in service (in the up state) in the system.

By associating to each possible combination of machines in the up condition a state of the system, aggregating those with equal capacities, and characterizing the resulting state by the capacity available to the system, it is possible to derive a Markovian model for system capacity transitions. For example, it is possible to speak about the probability of any given capacity state (in this sense it doesn't differ from LOLP) and it is possible to discuss interstate transition statistics such as mean recurrence times.

The frequency and duration measures are simply related to the statistics associated with the probability of occurrence of negative margin states ($C-L = M$; $M < 0$) and their mean recurrence times. Note that the FAD method yields a measure of the type "on the average the system will encounter a negative margin state once every 5 years and its mean duration will be 4 hours". In other words, the measure yields information relating to the mean time between loss of load situations, and the mean duration of the contingency.

LOLP, on the other hand, yields a different piece of information, its measure is an indication of the long run fraction of time the system will not be able to serve 100% of demand. It should be apparent that given the FAD measures it is possible to define a quantity similar to LOLP for the given load and capacity models used to derive the FAD measures. These, however, may differ from those used to derive a given LOLP measure.

One can see that the LOLP, or the average long term fraction of time the load cannot be fully met, is just $(f \cdot d)$. A dimensional analysis may reassure the disbelievers: if f is in a per year dimension, and d is in a days dimension, then $(f \cdot d)$ has dimension of days per year.

3.2.5.1 FAD Capacity Model

There are five assumptions made to construct the FAD capacity model:

- 1) The generators' behaviors are statistically independent of each others'.
- 2) Each generator is characterized by its nameplate capacity.
- 3) Each generator can be found in an up ($\tilde{c} = c_i$) or down ($\tilde{c} = 0$) state. It is possible to handle multilevel deratings but for simplicity we'll stick to this model. Remember the \sim above a variable such as \tilde{c}_i indicates that we are speaking about the random variable \tilde{c}_i .
- 4) We assume that probabilities for transitions from up to down or vice versa are a function only of the present state they are in, and not of the machine history.
- 5) Transition times are described by exponential probability density functions which mean they are

independent of the time elapsed since a transition.

3.2.5.2 The One Machine Capacity Model

Let's set up the Markov Model for a one machine system

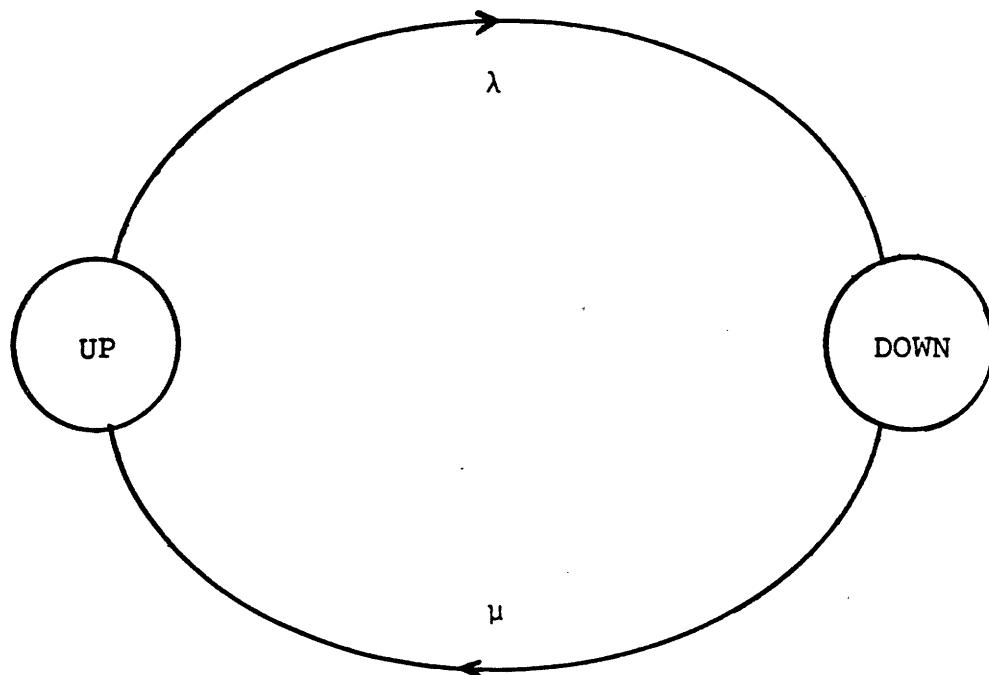


Figure 3.6 The Basic Markov Model

We assume exponential holding times in each state characterized by a departure rate.

λ for breakdowns $1/\lambda$ is the average holding time in state up.

μ for repairs $1/\mu$ is the average holding time in the down state.

The state equations then are

$$\begin{aligned}\dot{P}_1 &= -\lambda P_1(t) + \mu P_2(t) \\ \{ \quad & \\ [P_1 + P_2](t) &= 1\end{aligned}$$

at steady state

$$\begin{aligned}\dot{P}_1 &= \dot{P}_2 = 0 \\ P_1(t) &= P_1, P_2(t) = P_2\end{aligned}$$

Thus

$$\lambda P_1 = \mu P_2 \rightarrow P_1 = \mu / (\lambda + \mu)$$

$$P_1 + P_2 = 1 \rightarrow P_2 = \lambda / (\lambda + \mu)$$

What is the rate of transitions into state "up"? This will be the frequency of encountering state "up". This must be equal to

$$\sum_{j \neq 1} P_{j1} \lambda_{j1} = P_2 \mu$$

but this by means of the state equations above must equal

$$P_1 \lambda$$

another way to think of this relationship is that the transition rate into the up state equals

the rate of transitions out of state "up" = λ
times the long term probability of being in "up" state = P_1

This result is useful in that it allows expression of frequency of encounter of a state in terms of the steady state probabilities of being in that state, multiplied by the transition rates out of that state.

In general the Chapman Kolmogorov equations (from the previous state equations) state that:

$$\sum_{j \neq i} P_{ji} \lambda_{ji} = \sum_{k \neq i} P_{ik} \lambda_{ik} \text{ for } V \text{ states } i$$

This can be interpreted to say the average rate of transitions "in" must equal the average rate of transitions "out" which equals the frequency with which the state occurs. This is true for Markov processes which are connected and which have no trap states; happily, we will not encounter these exceptions in our problem formulation.

The exponential holding time model assumes that the time spent in the "up" state is a continuous random variable described by the probability density function $f(t) = \lambda e^{-\lambda t}$, this means

Prob. (transition in time t /system in "up" state) =

$$\int_0^t \lambda e^{-\lambda \tau} d\tau = 1 - e^{-\lambda t}$$

$$E(t) = \int_0^\infty \tau \lambda e^{-\lambda \tau} d\tau = 1/\lambda$$

Let's take a close look at the 1 machine example, if the

forced outage rate (for) = .02

and the mean repair time = $r = 2.04816$ days

let's then solve for μ and λ , and hence for the model parameters,

$\mu \equiv 1/r$ therefore $\mu = .49/\text{day}$

from Chapman Kolmogorov

$\lambda P_1 = \mu P_2$ remember P_2 is the probability of being down=for= $r/m+r$

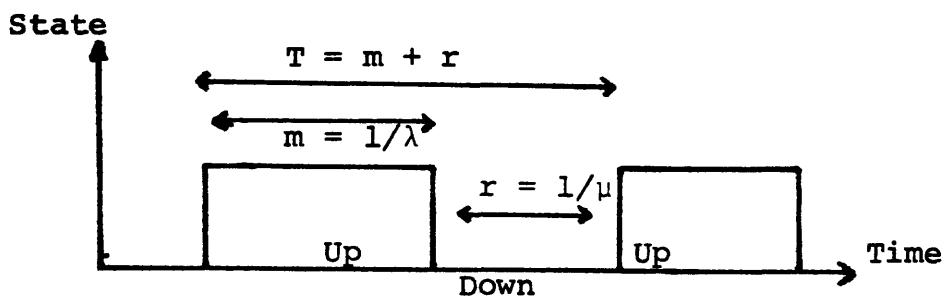
$$\lambda = \mu P_2/P_1 = (.49)(\frac{0.2}{0.98}) = .01 = \lambda$$

now $T = m + r = 1/f$ $m = 1/\lambda$ by definition of exponential holding times = $1/\lambda + r = 100 + 2.04816 \approx 102.05$

$f = 1/T = 1/102.05 \approx .01$ the frequency

of occurrence of down state, once every 100 days.

When it occurs its average duration is 2.04816 days.



$$\text{note that } f \cdot d \approx \frac{2.05}{102.05} \approx .02 \approx \text{f.o.r.}$$

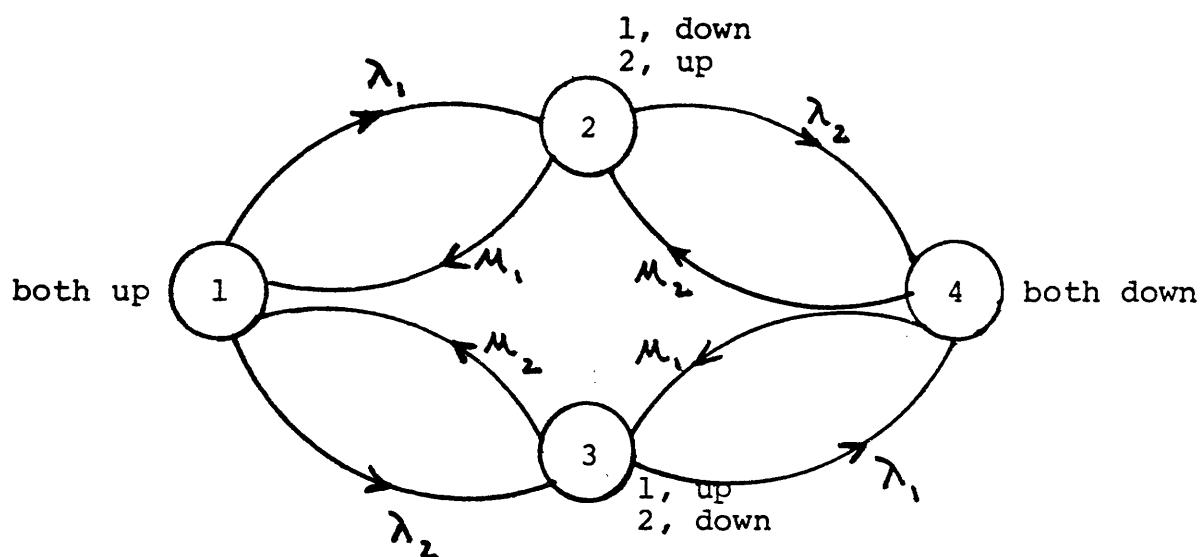
which is equal to LOLP if the load is always greater than 0.

3.2.5.3 The Two Machine Capacity Model

We have two generators in parallel.

Unit	Capacity	Availability	r (days)	u
1	20	.98	2.040816	.49 .01
2	30	.98	2.040816	.49 .01

<u>State numbers</u>	<u>Capacity avail.</u>	<u>A</u>
1	50	$P_1 = (.98)^2$
2	30	$P_2 = (.02)(.98)$
3	20	$P_3 = (.98)(.02)$
4	0	$P_4 = (.02)^2$



We know the rates of departure from each state are

<u>S</u>	<u>R.D.</u>	<u>f (frequency)</u>	<u>T=1/f</u>
1	$\lambda_1 + \lambda_2 = .02$	$P_1(\lambda_1 + \lambda_2)$	52.06
2	$\mu_1 + \lambda_2 = .5$	$P_2(\mu_1 + \lambda_2) = (.0196)(.49)$	102.04
3	$\mu_2 + \lambda_1 = .5$	$P_3(\mu_2 + \lambda_1)$	102.04
4	$\mu_1 + \mu_2 = .98$	$P_4(\mu_1 + \mu_2)$	3554.02

Therefore the frequency or cycle time between recurrence of any state has been solved. What about the average duration of a particular state?

The average duration of a particular state is just the reciprocal of the transition rate out of it.

$$d_{s_1} = 1/\lambda_1 + \lambda_2 = 50$$

$$d_{s_2} = 1/\mu_1 + \lambda_2 = 2$$

$$d_{s_3} = 1/\mu_2 + \lambda_1 = 2$$

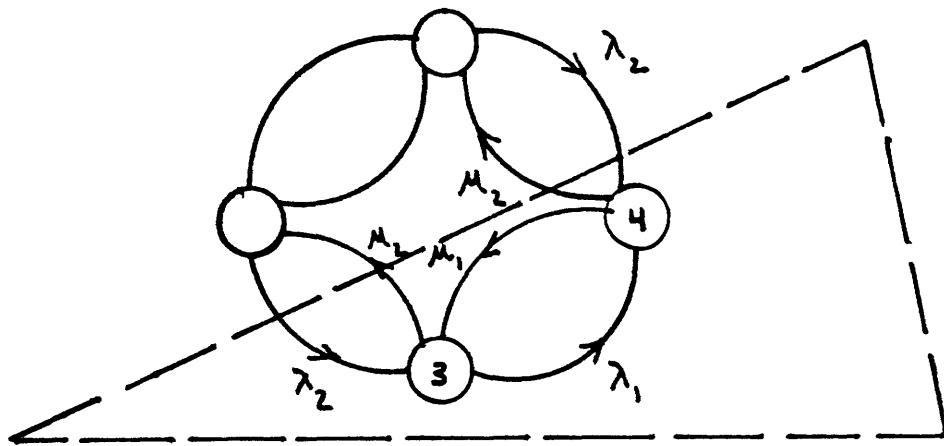
$$d_{s_4} = 1/\mu_2 + \mu_1 = 1.02$$

Let us assume our maximum load is smaller than 20 MW, then the average frequency of not being able to serve the load would be the frequency with which we encounter state 4, or once in every 2554.02 days. How long will the disturbance

last? Its average duration would be 1.02 days.

Let us assume our load is greater than 20 MW but smaller than 30. Then the system will fail whenever states 3 or 4 are encountered. What is the frequency of this occurrence? What is the average duration of it when it occurs?

Let's use the Markov process diagram:



If one looks at the combined states (3, 4) the frequency of encounter is given by

$$f = P_3 u_2 + P_4 u_2 \quad (\text{the two outgoing branches})$$

$$f = (.49)(.0196 + .0004) = (.49)(.02) = .0098$$

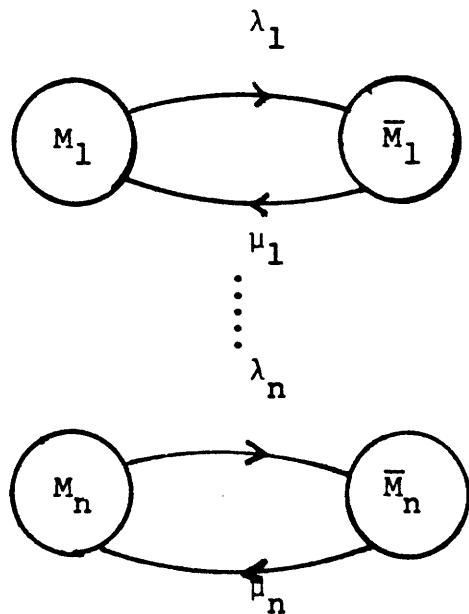
$$T = 102.04 \text{ days}$$

$$\bar{d} = \frac{\text{cumulative probability of states}}{\text{cumulative frequency}} =$$

average duration = $\frac{.02}{.0098} = 2.0408 = 1/u_2$ which represents the mean time for machine 2 repair.

3.2.5.4 Generalized n-Machine Problem

The process consists of n machines, each either up or down, therefore it is a combination of n two state Markov Processes.



We would like to ask questions about the ensemble behavior.

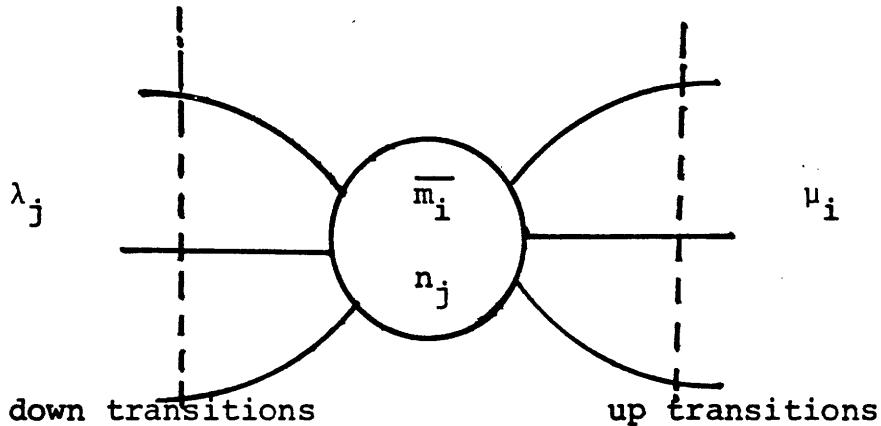
Since each process has 2 states and there are n processes we see there are 2^n possible states. Let's further assume that out changes cannot arrive at once (say a machine cannot fail and another be repaired in the same instant.) Then we see that each of these 2^n states can transit into any of n other states, representing one of the n machines changing state either up or down depending on their present status.

That is, of the machines that are down, \bar{M}_i , any can go up with u_i ; of the machines that are up M_j , any can go down

with λ_j .

$\{m_i, m_j\} = \{m\}$, the set of all machines

$n(m_i) + n(m_j) = n$, total # of machines



Of course, what we are interested in is the capacity available on the machines that are up $C(M_j) = \sum C_j = C(s)$, s , the state of the system. Note that if all machines have the same capacity that we will at most be interested in $(M+1)$ different capacity values, i.e., all states could be merged into one of these capacity values. At worst, we might be interested in 2^n capacity values. (If no partial sums of capacities = any other partial sums of capacities.)

3.2.5.5 Procedure to Obtain FAD Measures for the n-Machine Problem

We have seen that the FAD approach yields measures that relate to how often on the average will an outage event greater than or equal to 0_k occur, and on the average how long will it

last. If the load encountered were constant at L_k , say, and total capacity were equal to C_{tot} , then we would like to know the frequency and duration of outage events greater than $(C_{tot} - L_k = 0'_{k'})$ and these would be our reliability measures. There are several problems though: 1) the load is not constant (but we will postpone this discussion and present a more realistic model later), and 2) we have not presented a simple way to obtain the frequency and duration of any particular outage event, never mind that of the particular outage event mentioned above.

What we want is to split up the 2^n state Markov process diagram into two parts: 1) those states whose outage is greater than $0'_{k'}$ (or $C_k < L_k - 0'_{k'}$) and 2) the remaining ones which necessarily have outages less than or equal to $0'_{k'}$. To obtain the frequency of this cumulated state we would want to sum up over all the states i in the first group the terms $P_i \lambda_{ij}$ where the j refer to states in the 2nd group.

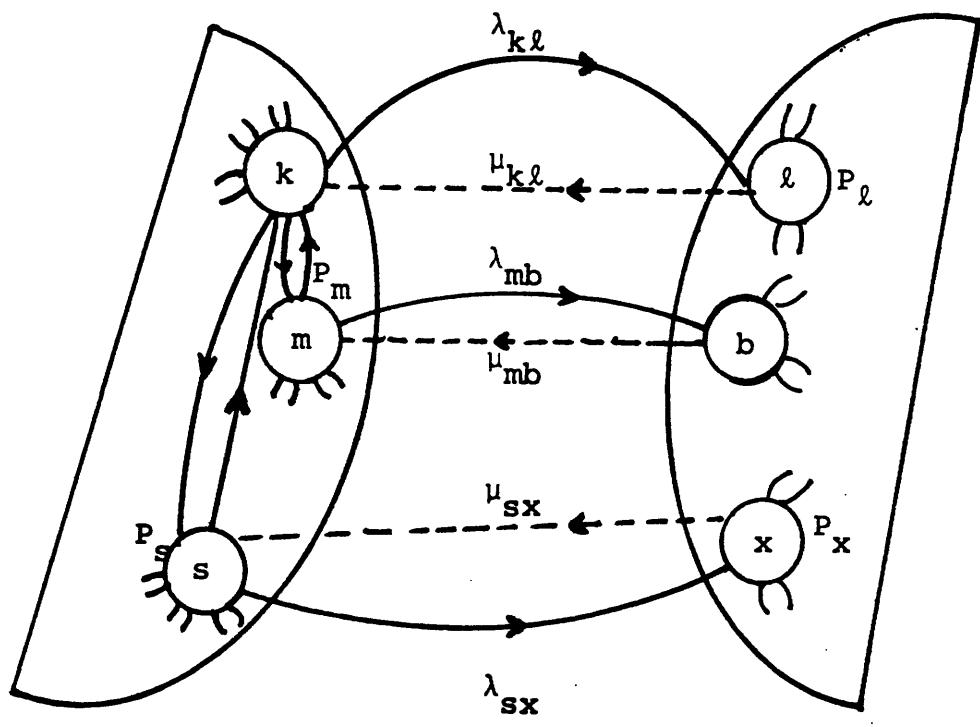
One could proceed to tabulate for each of the 2^n states, its respective frequency and duration of appearance. But note that if we summed these single state frequencies over the states in group 1 to obtain measures for the cumulated outage state 1, that we would be including in the sum some terms that are due to transitions within the cumulated state, i.e., intrastate transitions which should not be included to obtain mean recurrence time measures.

In the past this was used as an approximately correct

answer. We see that this method does not work for the computation of cumulated state frequencies of appearance even though it is possible so far to obtain probabilities for cumulated states by just adding up the single state probabilities.

For any given outage value, 0_k , it would be possible to divide the rates of transition out of states in cumulated state 1 into those that led to higher values of 0_k , i.e., transitions that stayed in cumulated state 1, and those that led to leaving cumulated state 1 and into cumulated state 2 by yielding lower outages (obviously any machine that was down in state 1 going up). It would then be possible to sum up the terms $P_i \lambda_{ij}^+$ and this would yield the frequency of the cumulated state 1. The duration would then equal the (cumulative probability/frequency) and we have computed both.

But note that this would require a new computation for each value of 0_k that would be exceedingly tedious to recompute. We will now present a simple procedure suggested by Hall, Ringlee and Wood (HRW) in their September 1968 paper in the IEEE PAS Transactions to compute FAD measures for generation systems. The benefit of the procedure is that it computes the relevant items once and for all and is easy to computerize. We will also present our own proofs of the procedure and our own suggested routine for computerization since they did neither (they did, however, provide a plausibility argument).



1

2

states with outages $> 0_k$

states with outages $< 0_k$

3.2.5.6 The HRW Recursive Algorithm

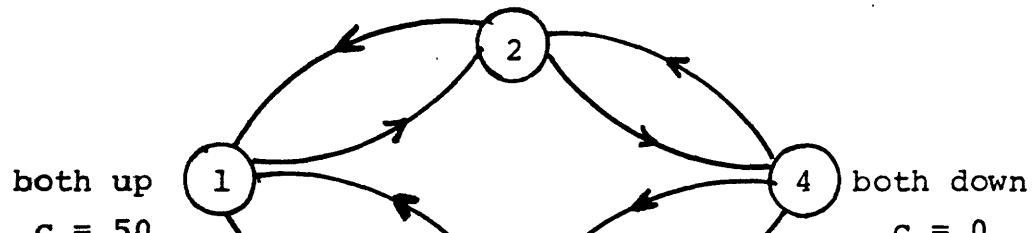
(r.d.)

<u>State</u>	<u>Available Capacity</u>	<u>P exact</u>	<u>Rate of Dept.</u>	<u>T=1/f, f=P_{ex} (r.d.)</u>
1	50	.9604	$\lambda_1 + \lambda_2 = .02$	52.06
2	30	.0196	$\mu_1 + \lambda_2 = .5$	102.04
3	20	.0196	$\mu_2 + \lambda_1 = .5$	102.04
4	0	.0004	$\mu_1 + \mu_2 = .98$	2551.02
		1.0000		

c = 30

2 = up

1 = down



both up
c = 50

both down
c = 0

1 = up

2 = down

c = 20

So far, we haven't done anything new. The previous

procedure has computed the exact probabilities of each state and using the relationship

$$f = P_s (\sum \lambda_{out} + \sum \mu_{out}) \\ \left(\text{s.s. probability of being in state } \right) \cdot \left(\text{total rate of departure from state} \right)$$

the procedure has derived the frequencies of encountering each state. Now, if we desired, we could compute the duration of each state by

$$\bar{d} = \frac{P_s}{F_s} = \frac{P_s}{P_s(\text{rate of dep.})} = \frac{1}{\text{rate of dep.}}$$

that is, the mean time in residence in any state is given by the inverse of the total rate of departure from that state.

The problem, however, is that we need expressions for the occurrence of an outage of a given magnitude or greater, the cumulative probabilities and frequencies. We note that obtaining the cumulative probabilities presents no problem; if the exact states are ordered in ascending order of capacity available, the cumulative probability is obtained by

$$P_n = P_{n-1} + P_k$$

So what does present a problem? How do we move from one capacity availability state to an equal or higher one in a way in which we can compute the cumulative frequencies for

these events? HRW present a technique which accomplishes this in a neat recursive algorithm that uses the probabilities P_k of the exact state that is being appended to the current cumulated state, as well as the state's λ_{-k} and λ_{+k} --the rates of departure to lower and higher capacity states.

Before we proceed: note that there may be some identical capacity states. We'll merge these before the algorithm is applied so that a perfect monotonically increasing order may be imposed. Merging rules will be:

- a) $P_k = P_i + P_j$ probability of merged capacity states c_i, c_j , with $c_i = c_j$, is just the sum of the probabilities.
- b) $f_k = f_i + f_j$ the frequency of encounter of the merged state is just the sum of the frequencies of each. This is natural and valid since they correspond to independent states which have no interstate transitions because any transition would lead to a different capacity state.
- c) $\lambda_{\text{up}, k} = \frac{P_i \lambda_{\text{up}, i} + P_j \lambda_{\text{up}, j}}{P_i + P_j}$

$$\lambda_{\text{down}, k} = \frac{P_i \lambda_{\text{down}, i} + P_j \lambda_{\text{down}, j}}{P_i + P_j}$$

the respective departure rates for the merged state are a weighted sum of the departure rates for the separate states.

Now we have a monotonically increasing set of capacity states. Let us start with the 0 capacity state.

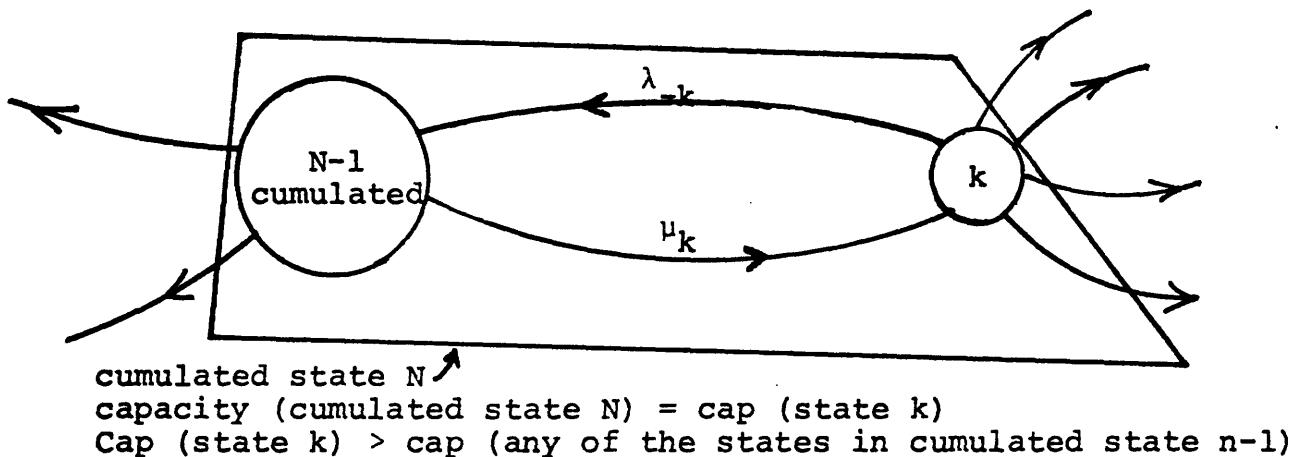
1) $P_1 = P_{c=0}$, $f_1 = f_{c=0}$ these are the initial values in our recursive formula, there is no trouble obtaining them off the exact state table. In our example

$$P_1 = .0004 \quad f_1 = (.0004)(.98) = 1/2551.02$$

2) Now let

$$f'_n = f'_{n-1} + P_k \lambda_{+k} - P_k \lambda_{-k}$$

note that this expression gives a way to calculate cumulative frequency of a higher capacity state starting with that of a lower capacity state and modifying it by statistics relating to the merged state k and departure rates from k to lower and higher states.



The formula above says that the frequency of cumulated state N is equal to that of cumulated state $(n-1)$ except you have to include the new possibility for upward transition out of state n into state $n+1$ and you have to exclude from consideration those transitions that occurred from cumulated state

(n-1) to state k.

(Note that because of ascending capacity order, the transition always corresponds to turning on one generator.)

Now we can show how this works in our example. (Denote by ' frequencies in the new cumulated state order.)

$$P'_1 = .0004 = P_4 \quad f'_1 = (.0004)(.98) = P_4(\mu_1 + \mu_2)$$

$$P'_2 = P'_1 + P_3 \quad f'_2 = f'_1 - P_3\lambda_1 + P_3\mu_2$$

$$P'_3 = P'_2 + P_2 \quad f'_3 = f'_2 - P_2\lambda_2 + P_2\mu_1$$

$$P'_4 = P'_3 + P_1 \equiv 1$$

↓

$$P'_1 = .0004 \quad f'_1 = .000392$$

$$P'_2 = .02 \quad f'_2 = .000392 + .48(.0196) = \\ .00392 + .0094 = .009792$$

$$P'_3 = .0396 \quad f'_3 = .009792 + .48(.0196) = \\ .009792 + .0094 = .019192$$

$$P'_4 = 1 \quad f'_4 = .019192$$

cycle $T'_1 = 2551.02$ $\frac{C_0}{50}$

times $T'_2 = 102.04$ 30

between $T'_3 = 52.1$ 20

outages

greater than or equal to C_0

What about the duration of the outages? These can be obtained from

$$\bar{d} = \frac{\text{cum. prob.}}{\text{cum. freq.}} = \text{cum. prob.} \times \text{cycle time}$$

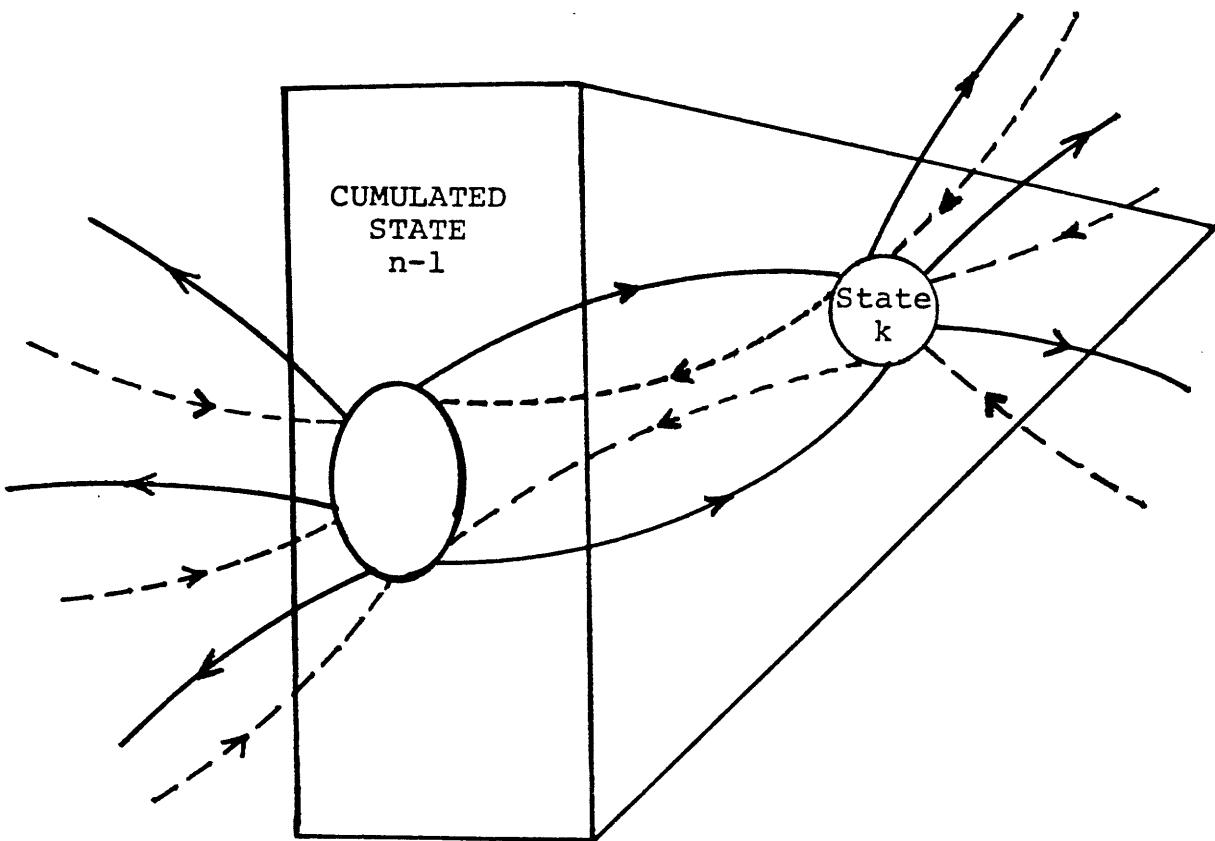
$$\bar{d}'_1 = (.0004)(2551.02) \approx 1 \text{ day}$$

$$\bar{d}'_2 = (.02)(102.04) \approx 2.04 \text{ days}$$

which agree with
previous answers

$$\bar{d}'_3 = (.0396)(52.1) \approx 2.05 \text{ days}$$

3.2.5.7 Proof of HRW Recursion Algorithm



$$1) \quad f'_n = f'_{n-1} + f_k \quad (\text{sum of encounters from } k \text{ to } n-1 \text{ and from } n-1 \text{ to } k)$$

$$f_k = p_k \lambda_k^+ + p_k \lambda_k^-$$

$$\text{encounters from } k \text{ to } n-1 = p_k \lambda_k^-$$

$$\text{encounters from } n-1 \text{ to } k = \sum_{j \in (n-1)} p_j \lambda_{jk}$$

$$f'_n = f'_{n-1} + p_k \lambda_k^+ + p_k \lambda_k^- - p_k \lambda_k^- - \sum_{j \in (n-1)} p_j \lambda_{jk}$$

$$f'_n = f'_{n-1} + p_k \lambda_k^+ - \sum_{j \in (n-1)} p_j \lambda_{jk}$$

$$2) \quad \text{Let's focus on the } \sum_{j \in n-1} p_j \lambda_{jk} \text{ terms}$$

The only difference between state j and state k is one machine that goes up on its way from j to k .

Let's compare the $P(j)$ and $P(k)$ the steady state probabilities of finding the system in states j, k .

Define $P(s) = \sum_{i \in s} \theta_i$ and $\theta_i = \begin{cases} p_i & \text{if machine in } i \text{ is up} \\ \bar{p}_i & \text{if machine in } i \text{ is down} \end{cases}$

where p_i is the steady state probability in the 2 state Markov process of finding machine i up, $\bar{p}_i = 1 - p_i$.

$$P(j) = \sum_{i \in (jnk)} \theta_i \cdot \bar{p}_l \quad \text{because these are statistically independent processes, } l \text{ is the machine that goes up}$$

$$P(k) = \sum_{i \in (jnk)} \theta_o \cdot p_l$$

since $P_\lambda \lambda = \bar{P}_\lambda \mu$ from 2 state Markov process

then $P(j)\mu = P(k)\lambda$

which says that each branch pair is equal so that

$$\sum_{j \in n-1} P_j \lambda_{jk} = P_k \lambda_k^-$$

it therefore follows that

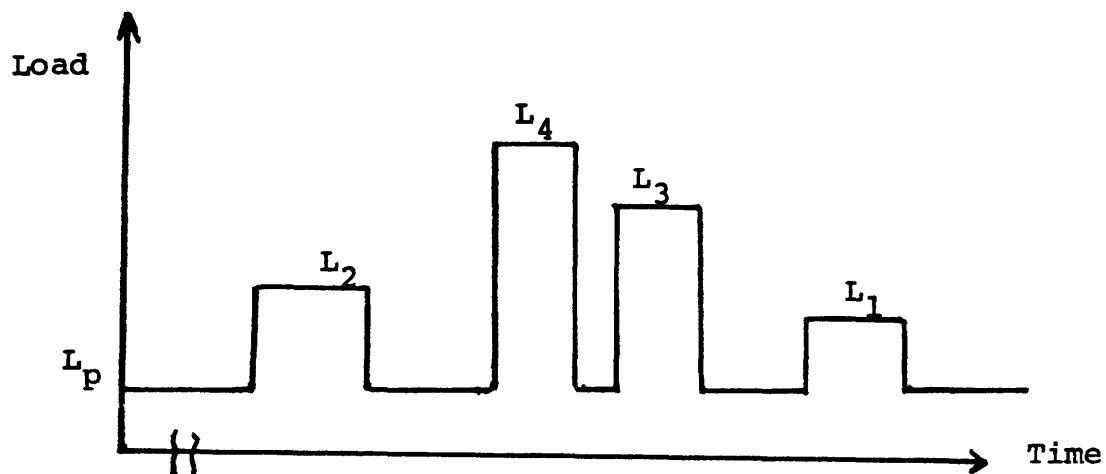
$$f'_n = f'_{n-1} + P_k \lambda_k^+ - P_k \lambda_k^-$$

and this is a useful result in that it only uses statistics of the next merged state.

3.2.5.8 A Load Model for the FAD Method

There are many of these, and they are still being improved upon, but they seem to be forced and lack the naturality of the LOLP load models. We will present one such load model.

The details of this model are as follows:



number of occurrences of L_i , n_i , $i=1 \rightarrow N$

interval length $D = \sum n_i$

expected peak duration = e

$$\text{probability of } L_i = P_i = \frac{n_i e}{D}$$

transition rate to higher load $\lambda_{+L_i} = 0$

transition rate to lower load $\lambda_{-L_i} = 1/e$

frequency of L_i $f_i = n_i/D$

for low load period

load state = L_0

$$P_0 = 1 - e$$

$$\lambda_{-L_0} = 0$$

$$\lambda_{+L_0} = 1/(1-e)$$

$$f_0 = 1$$

This model assumes load returns to L_0 before going to new peak

L_i . Now consider

$$M_k = C_n - L_i \quad \left\{ \begin{array}{l} \lambda_{+m} = \lambda_{+c} + \lambda_{-L} \\ \lambda_{-m} = \lambda_{-c} + \lambda_{+L} \end{array} \right.$$

$$P_k = P_n \cdot P_i$$

$$f_k = P_k (\lambda_{+k} + \lambda_{-k})$$

and this suggests a procedure for combined load and capacity model FAD measures.

3.2.5.9 The Combined Load and Capacity Measure

We have so far seen the separate developments of the load and capacity models for the FAD method. Combining these models for obtaining summary measures represents little problem except that of handling a much larger number of variables. Instead of solving for outage states as we did for the capacity model we now define a margin state model; potentially one for each value of C and L or $2^n \cdot (n_L)$. The problem then simply becomes that of finding the frequency and duration of the cumulated first negative margin state, i.e., we want to find the cumulative measures corresponding to situations that have margins more negative than the first situation in which $C < L$.

3.3 Probabilistic Simulation

3.3.1 Another Technique for Reliability Calculations

Our previous computational techniques for LOLP and LOEP have depended upon calculation of the probabilities of an exponentially rising number of possible outage states, i.e., as

the number of generators rises we face a maximum 2^n possible outage states. There is an alternative to these methods which transforms this combinatorial problem into a recursive one. This alternative technique is referred to as probabilistic simulation (PROSIM) and it is used for purposes beyond those of computing system reliability measures; in particular, the technique also yields the expected amount that each unit will be called upon to produce, a number which is extremely useful for determining system expansion plans. We will return to these other outputs later. Our work is based upon the original contributions of Baleriaux and Jamouille of Belgium, the later work of Booth in this country and most lately on work done by Deaton at M.I.T.

To calculate LOLP we have basically said that

$$LOLP = \sum_{\tilde{c}} P(\tilde{c}) P(L > \tilde{c})$$

By focusing on the capacity on outage (\tilde{c}_o) rather than on the capacity available (\tilde{c} , $\tilde{c}_o = IC - \tilde{c}$) we could have equivalently written this expression as

$$LOLP = \sum_{\tilde{c}_o} P(\tilde{c}_o) P(L > IC - \tilde{c}_o)$$

This result is rather interesting, especially if we look at the more general expression

$$F_n(x) = \sum_{\tilde{c}_o} P(\tilde{c}_o) P(L > x - \tilde{c}_o)$$

This latter expression is simply the convolution of the cumulative load distribution function (i.e., the load duration curve) with the density function for capacity on outage. In other words,

$$F_n(x) = F_o(x) * f_{c_o,1} * \dots * f_{c_o,n}$$

where $F_o(x)$ is the load duration curve. But we see that

$$F_n(IC) = LOLP$$

What else can we tell from this curve? We know from our previous work that

$$E_{tot} \cdot LOEP = \sum_{k=0}^{\infty} P(O_k) E(O_k)$$

or alternatively

$$\sum P(\tilde{c}_o) \int_{IC-c_o}^{MA} F_o(\tau) d\tau$$

where IC = installed capacity and MA = maximum forecasted load.

But we know that this is equal to

$$\sum P(\tilde{c}_o) \int_{IC-c_o}^{MA+IC-c_o} F_o(\tau) d\tau$$

because $F_o(\tau) = 0$ for $\tau > MA$ and $IC > c_o$ for all \tilde{c}_o .

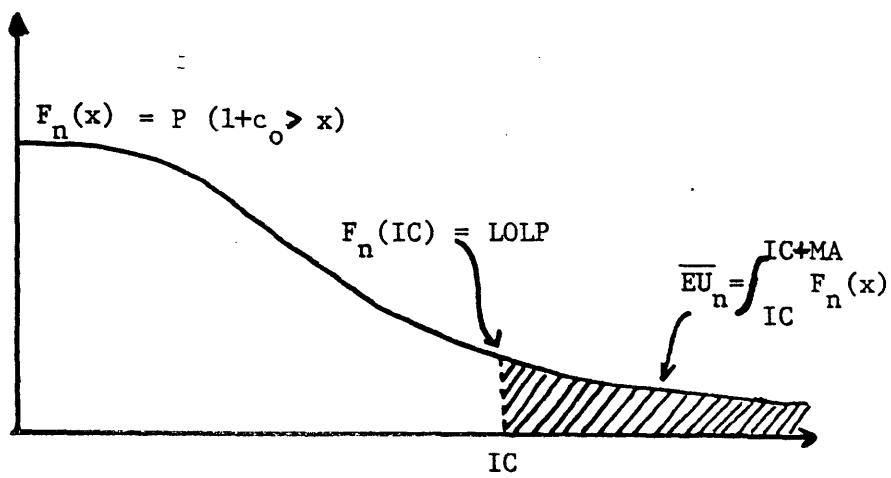
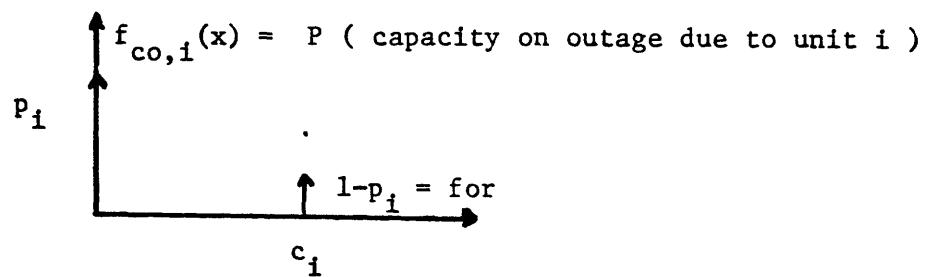
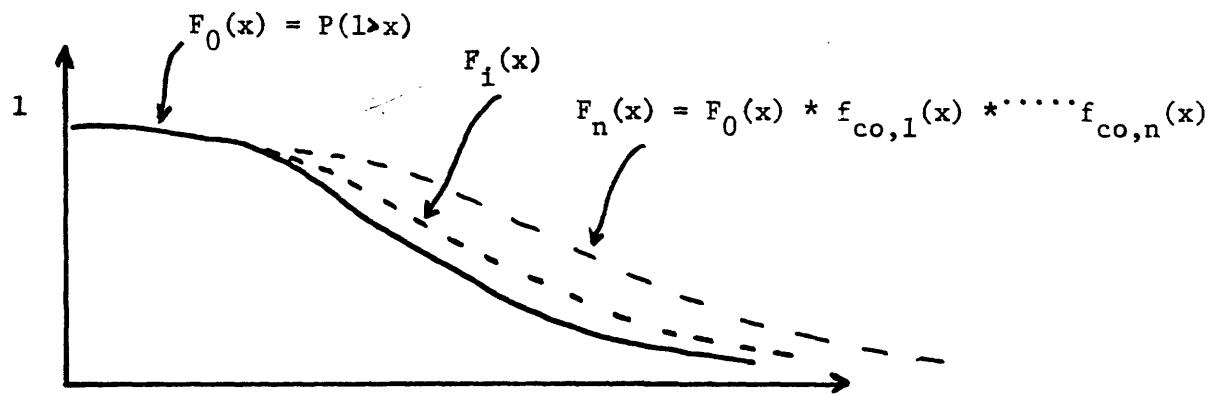


Fig. 3.7 The Curves ($F_i(x)$) and $F_n(x)$.

But this in turn

$$= \sum_{\tilde{c}_o} P(\tilde{c}_o) \int_{IC}^{IC+MA} F_o(x-\tilde{c}_o) dx$$

$$= \int_{IC}^{IC+MA} \sum P(\tilde{c}_o) F_o(x-\tilde{c}_o)$$

$$= \int_{IC}^{IC+MA} F_o(x) * P_{C_o}(x) dx$$

$$\text{LOEP} = \frac{\int_{IC}^{IC+MA} F_n(x) dx}{E_{tot}}$$

and the term

$$\int_{IC}^{IC+MA} F_n(x) dx$$

is proportional to the energy expected to remain unserved by the system (if EU is multiplied by the time in the period it acquires the dimensions of energy).

From the above analysis we see that it is possible to obtain these measures in a much simpler fashion consecutively convolving the load duration curve with the density for capacity on outage for each unit until all have been included to

form the equivalent load curve, $F_n(x)$. From $F_n(x)$ we can obtain:

$$1) \text{LOLP} = F_n(\text{IC})$$

$$2) \text{LOEP} = \int_{\text{IC}}^{\text{IC+MA}} F_n(x) dx \div \int_0^{\text{MA}} F_o(x) dx$$

3.3.2 Other Outputs of PROSIM

There are other useful facts that can be obtained from this approach; in particular, it is possible to use this technique to simulate future system operation in order to calculate future expected operating costs for each unit. Why is this method better than other deterministic methods used in the past? Simply because it seems to reasonably model the effects of random occurrences on system operation.

A bit of history before we proceed will illuminate this point. It should be obvious that:

- 1) if plant capacity is always larger than peak demand by 15% to 30% (the system margin) and
- 2) that in order to minimize operating costs only cheaper units will be operated (i.e., loading will be done in order of lowest incremental cost first; therefore the base loaded ones will be operated first, followed by intermediates and topped with peaking units which are characterized by lowest capital costs and highest operating costs).

- 3) it follows that the top part of the system's capacity--the system margin--in this deterministic loading scheme, will never be expected to generate power.

This is clearly a misleading result in that peaking units are operated throughout the year even though an effort is made to run them as little as possible. Their presence in the system can be explained through the system's effort to keep cheap capital cost but expensive operating cost equipment around to replace expensive capital cost but cheap operating cost equipment whenever the latter goes on forced or scheduled outage.

In other words, since they expect to use these replacement sources sparingly, they purchase the lowest capital cost equipment without heavy regard to its operating cost. These types of units are also purchased to provide for greater operational flexibility; for example, they are used for short time period generation for load following purposes where it is undesirable to run bigger, less flexible, machines.

Why is it necessary to simulate the units' production costs? The reason is obvious if one thinks about the generation expansion problem in which the simulation problem is embedded. In order to choose the lowest present value cost system to fulfill future electric demand while satisfying a minimum level of reliability it is necessary to pick a system whose investment as well as operating costs over time are minimum. However, the operating costs in period k, OC(k) are

obviously a function of the system that has been built up to date, $\underline{E}(k)$ ($\underline{E}(k)$ is a vector representing the different kinds of plants), which has an associated investment cost stream $I(k)$, and $OC(k)$ cannot be determined by treating each unit separately. Once the units are in hand, the system will take from each in such a way to minimize total production costs; by using PROSIM we can estimate total expected production costs. In particular, PROSIM does this by simulating each plant's loading in order of increasing incremental generating cost; i.e., the cheaper to operate, the more energy the plant will be expected to produce.

Even though it is not valid to associate an $OC(k)$ to each unit built before looking at system effects on that unit's expected production, this is the path we take in determining expansion patterns. Our main reason for doing this is that the programming problem for determining expansion patterns would become unwieldy without this simplification. We can afford to make this simplification because PROSIM gives us a method by which to check how good our $OC(k)$ assumptions have been and thereby to check the validity of the derived expansion pattern. If our assumptions are wrong, we modify them and recalculate a new expansion pattern and begin the process anew. In this fashion a solution is converged upon.

Another problem arises in our particular scheme for expansion determination. We use a linear programming approach and there is no way of effectively expressing a reliability criterion in the problem itself. Therefore, we use a reserve

percent margin guideline in the LP which is modified according to the calculations yielded by PROSIM and this information is combined with the simulated production history data to generate the modified expansion patterns.

First we shall expand a bit on the mechanics of probabilistic simulation. The idea is to figure out a way in which to accurately display the probabilistic nature of determining system operating costs. Clearly, we need to use a load duration curve-related concept for it tells us what will be expected from the capacity on line in terms of energy demanded. What is wrong with the deterministic approach for calculating fuel costs? Quite simply that when we load a unit, say the first one loaded, we assume it will always be available. This clearly is an incorrect assumption; if the unit is not available to us say 10% of the time when we want to use it (F.O.R. = .10), this means that at best the unit will produce 90% of the energy desired from it. This means that this extra energy will have to be produced by units further up the loading order. So what we need is a technique which will derive the energy produced by one unit and in some sense reapportion that portion which it will not produce because of its probability of being down.

Why can we approach this problem in recursive form and not worry that we may be losing something? Simply because the outage of a given plant will cause plants strictly above it in the loading order to produce more energy than otherwise would

have been the case. In no circumstances will that outage result in units below it changing their production pattern (this is not exactly the case for multiple valve point units but our discussion will be clearer if we postpone that discussion until later). Note that everytime we load a plant we must in some way modify the load duration curve faced by the later loaded generators because they may be called upon to produce some of the energy that the unit just loaded will not be able to. This means that one can proceed in the following way:

- 1) Schedule unit j on the present load duration curve $F_{j-1}(x)$, the modified load duration curve that has the information of all prior loaded units and their effect on all later units' output. Determine $E_{jp} = P_j E_{jd}$ (the p and d refer to probabilistic and deterministic).
- 2) Modify $F_{j-1}(x)$ to $F_j(x)$ to represent the new load duration curve with unit j scheduled. Stop if we have scheduled all units, i.e., if $j = N$, if not, advance the counter and return to 1).

We receive an unexpected benefit from this approach; at the end of our calculations we will also have calculated LOLP and LOEP as we discussed earlier. The key to our problem is the modification that must be made to $F_{j-1}(x)$ to get $F_j(x)$. We have already seen that we need that $F_j(x)$ reflect the probability that unit j will not be available to produce the energy that we thought it might produce when using $F_{j-1}(x)$ in

deterministic fashion. Let's look at $F_0(x)$ and how we get $F_1(x)$ from combining $F_0(x)$ and information from unit 1. What we would like is a curve $F_1(x)$ that at every point greater than c_1 would describe the probability that the load demanded plus the capacity on outage from unit 1 (which is 0 with $P=P_j$ and c_1 with $P = F.O.R. = 1-P_j$) would be greater than that x , because the integral of this function is what the next plant will produce. To see it another way, $F_1(x)$ should reflect the load duration curve faced by unit 2, i.e., if unit 1 is up with prob = P_1 , it will face $F_0(x)$ demand

$$\int_{c_1}^{c_1+c_2} F_0(x) dx$$

if unit 1 is down with prob = $1-p_1$, it will face $F_0(x)$ demand

$$\int_0^{c_2} F_0(x) dx$$

But this is equivalent to forming $F_1(x) = P_1 F_0(x) + (1-P_1) F_0(x-c_1)$ or alternatively $F_1(x) = F_0(x) * f_{c_{on}}(x)$. See Figure
In general

$$F_n(x) = F_{n-1}(x) * f_{c_{on}}(x)$$

and we see

$$\underline{E}_2 = P_2 \int_{c_1}^{c_1 + c_2} F_1(x) dx$$

But we can now see that what we have done is to form $F_n(x)$ which is nothing less than the cumulative probability distribution function for load plus capacity on outage due to units 1 through unit n. What else can we do with $F_n(x)$? We can derive the LOLP and LOEP measures and, better still, use the model to study the effect of different expansion policies in a more powerful analytical way than we could with our former techniques.

It should be apparent that one could proceed backward from $F_n(x)$ to determine unit production schedules and this is what Booth and Baleriaux-Jamouille advocate because it makes it possible to treat hydro and pumped hydro plants. The reason for this is that some nuclear generation will be used to feed the pumped hydro plant and this would not be possible to include in the forward scheduling technique. The backward technique is equivalent to the forward one otherwise and proceeds as follows:

- 1) First form $F_n(x)$.
- 2) Solve for LOEP and for LOLP = $F_n(1C)$.
- 3) Deconvolve $f_{c_o,n}$ the density for capacity on outage due to the nth unit on the loading order.

$$4) \quad \bar{E}_n = P_n \int_{IC-c_n}^{IC} F_{n-1}(x) dx$$

We will present several proofs and other useful results from PROSIM in an appendix.

3.4 A Critical Review of Reliability Techniques

3.4.1 Critique of the Measures

Some of these points we have mentioned before but we review them here for convenience. The measures can only produce measurements as good as the model assumptions; i.e.,

- 1) the models assume capacity outages are independent when in fact they might not be, in particular, if stringent climatic and operating conditions contribute to the probability of outage.
- 2) the models do not measure the system's protection against delay in installing units or against system parameter uncertainty (for example, F.O.R. variations from assumptions, load forecasting errors). A certain amount of protection can be acquired by studying system sensitivity to variations in input parameters and to delays in unit installation.

Each of the measures we have discussed also presents some problems, some in what they measure and some in the procedure used to derive the measures.

FAD

FAD's strengths include the information they give about a generation system regarding mean recurrence times between outage states and mean durations of these outage states. Note however, that no information is given about how serious the outages states are when they occur; we are much more likely to be concerned about a 1000 M.W. outage than one of 10 M.W. if they both present equal mean frequency and duration measures. Also, the FAD model requires more information about each generator, namely its mean outage frequency and outage duration. It is not clear that this detail of information from each unit can be had without greater uncertainty than in the simple forced outage rate. It also means that much more data must be collected and massaged to obtain measures.

Another problem with the FAD model is its load models' lack of naturality although it can be made to give approximately equal results to the LOLP formulation. In particular, if many load levels are assumed possible, FAD becomes much more intractable to work with.

It is also not clear that the additional information provided by FAD above LOLP and LOEP is significantly useful for design considerations.

LOLP

LOLP is a misnomer, it is really the expected fraction of time the system will spend on a loss of load incident. Note that it is related to FAD measures through the relation

$\text{LOLP} \approx \overline{F \cdot D}$, for appropriately defined load and capacity models. Note also that it does not say anything about how short the system is when a LOL occurs, i.e., all loss of load events count equally, except for their time duration, in LOLP. However, LOLP is exceedingly simple to calculate and is fairly simple to understand although there is little standardization to its application. Mindless application will yield answers that differ by orders of magnitude for equal systems. We will delve into this and show how this can occur in the next section.

LOEP

LOEP, again, is a misnomer. It is the expected fraction of system energy not served through loss of load incidents. Note that since LOEP uses the same LOLP models, and since LOEP weights each shortage by amounts proportional to the energy not served (and these fractions of energy unserved are always smaller than the fraction of time spent in outage for each outage) that this means $\text{LOEP} < \text{LOLP}$. It's been this feature of LOEP that has kept it from being a popularly applied measure, i.e., that it is too small for most American systems. But as we will discuss later, its smallness is of no object since the value of energy lost to consumers is certainly larger than the marginal revenues lost in a L.O.L. event, in other words, for value calculations LOEP must be multiplied or otherwise corrected by some factor larger than 1.

One more point. Although $\text{LOEP} < \text{LOLP}$ is always true for any particular system, it does not follow that $\text{LOLP} \approx \text{LOEP}$ for a

system S throughout time. That is, if $\text{LOLP}_{(S_1)} = k\text{LOLP}_{(S_2)}$ it does not follow that $\text{LOEP}_{(S_1)} = k\text{LOEP}_{(S_2)}$. The reason is simple, they measure different things. One way to see this clearly is to examine a typical LOEP-LOLP calculation and see what occurs as S grows through time.

3.4.2 A Critique of Technique Application and Interpretation

The field of reliability analysis has been plagued with major problems regarding standardization of application and of interpretation of the measurement techniques and of the adopted reliability criteria. Major pools, reliability councils and even companies within each of the above differ with respect to the measures they use, the target criteria they adopt and even with respect to the formulation and interpretation of the models used for determining the value of those measures. See the next page and note the diversity of target criteria across the country as of late 1971.

Because of intensive public pressure which led to the establishment of the reliability councils and because of the increasing need for interconnection in order to afford newer large units, much of the industry has been starting to say it is adopting the "one day in ten year" LOLP criterion. Why this should be the performance target no one seems to know. However, this is not the only problem since the different formulations

RESERVE CRITERIA CURRENTLY UTILIZED IN PLANNING STUDIES as of 10/31/71

Reliability Council

Reserve Criteria

I. NPCC

a. New England

b. NYPP

$LOLP = 0.1$

$LOLP = 0.1$ Days/Yr.
Loss of Load Occurs in Worst Month
of the Year

II. MAAC

PJM

$LOLP = 0.1$

III. ECAR

a. Michigan Pool

b. Allegheny Power System

c. CAPCO

d. Cincinnati, Dayton & Columbus

e. Kentucky-Indiana Pool

f. American Electric Power Company

g. Louisville Gas & Electric Company

h. NIPSCO

$LOLP = 0.2$

Being Studied

$LOLP = 0.2$

12% Through 1972, 15% Beyond 1972

15%

Judgement

15%

Being Studied

IV. SERC

- a. The Southern Company
 - b. Florida Power Corporation
 - c. Tampa Electric Company
 - d. TVA
 - e. Carolina Power & Light Company

V. MAIN

120

VI. MARCA

VII. SWPP

- | | | |
|----|--|-----------------------------|
| a. | Arkansas P&L, Louisiana P&L, Mississippi P&L, and New Orleans PS | One Occurrence in Ten Years |
| b. | Missouri Public Service | One Occurrence in Ten Years |
| c. | Southwestern Electric Power Company | $LOLP = 0.1$ |
| d. | Kansas City Power & Light Company | $LOLP = 0.1$ |

that are used are leading to results which can differ from one to two orders of magnitude when applied to identical system data. Old myths die hard, and it is easy to form a suspicion that present efforts at reliability planning are thinly veiled devices to justify decisions which are being made on the basis of maintaining a 20% to 30% reserve margin.

Lack of standardization enters into the problem of yearly calculation of LOLP. For example, one large power pool computes the LOLP during each maintenance period in terms of expected days of shortage for the period and then proceeds to add them to obtain the yearly LOLP.

$$\sum_{\substack{\text{periods } i \\ \text{in 1 year}}} (\text{LOLP})_i \cdot (\text{days in } i) = \text{LOLP}_{\text{year}} \text{ in days/year}$$

Another large power pool computes the yearly LOLP in terms of the worst maintenance period LOLP. The effort in this approach is to keep this period's LOLP at the target level, the assumption made is that there is enough slack throughout the rest of the year to accomplish scheduled maintenance tasks, etc., and keep the system LOLP below this targeted amount. Note that the previous approach, while keeping average LOLP at the target level, allows for great deviations over periods of the year so that both pools are in fact following very different criteria although they pretend to be doing the same.

Also it should be noted that the first pool's approach implicitly recognizes the expectation nature of LOLP. In

other words, by summing period LOLP's to obtain the yearly LOLP they are in fact summing expected values of times spent on outage rather than computing the probability of loss of load. If in fact they were interested in computing the probability of loss of load, and the assumptions were made that $(LOLP)_i$ represented the loss of load probability for period i and that the periods were independent of each other, then:

Probability (loss of load in periods 1 through n) = Probability (1 or more losses of load) = 1-P (no losses of load in periods 1 through n)

$$= 1 - \prod_{i=1}^n (1-LOLP_i)$$

where $(1-LOLP_i)$ = probability of no loss of load in period i. Note further that this probability quickly approaches 1 as the number of periods grows; this result follows naturally since it becomes progressively likely that there occur a loss of load the greater the number of periods involved.

A further set of comments should be made regarding the formulation and interpretation of these measures. First, the F.O.R. data are at best average data and do not accurately reflect the latest operating experience of the industry. Second, units take time to mature and this tends to cloud a reasonable forecast of future forced outage behavior. Third, the forced outage rate should reflect a unit's unavailability when required to perform rather than the time it spends in a

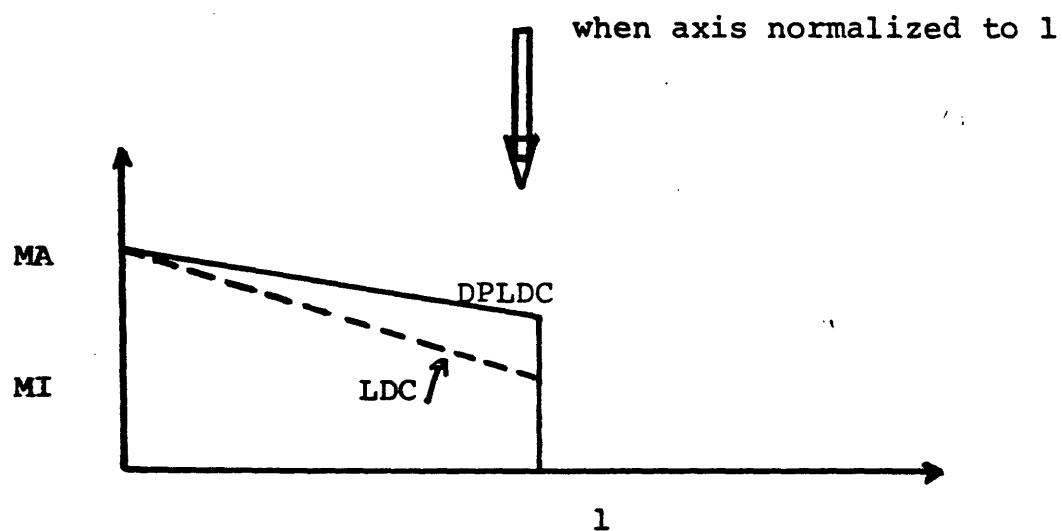
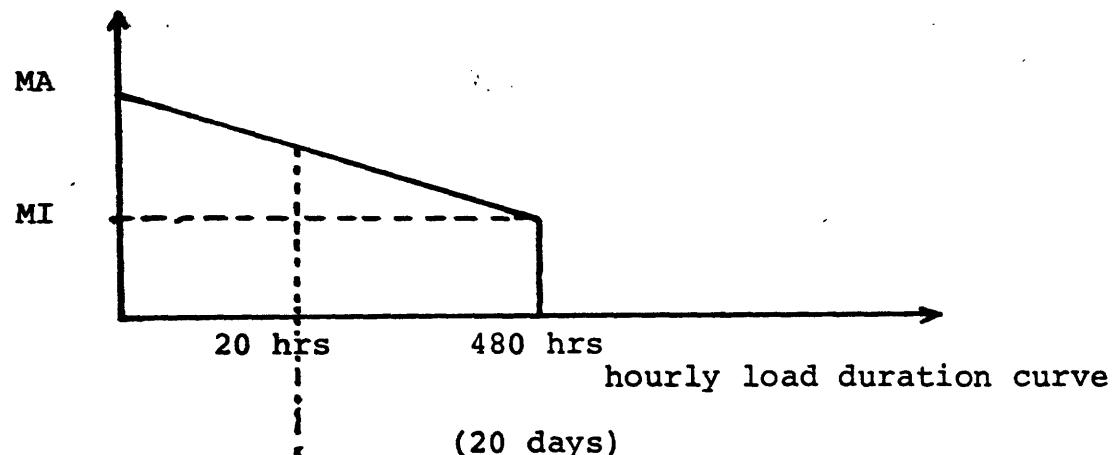
non-functioning state since the latter may depend on the resources committed to repairing units and on the urgency with which the particular unit is needed. Therefore great care must be used to ensure that proper data enters into a calculation.

A very common problem enters at this stage of problem formulation. What load model should be used: A daily peak load duration curve to thus concentrate analysis on peak load behavior, or alternatively an hourly load duration curve? Depending on which load model is used, the LOLP measure will be very different, and unless we are careful with its interpretation we may make incorrect decisions. Let's see why.

Consider the following problem: compute the LOLP for a given capacity model combined with an hourly load duration curve for a given period versus combining it with the daily peak load duration curve for the same period. See Figure 3.8. It is easy to see that because we have dropped out consideration of off peak hours the daily peak load duration curve is higher than the load duration curve, this in turn means that use of the DPLDC will yield a higher LOLP measure (as well as LOEP measure) than use of the LDC will.

$$LOLP_{DPLDC} = \sum P(0_k) T_D(0_k)$$

$$LOLP_{LDC} = \sum P(0_k) T_L(0_k)$$



Clearly the DPLDC is above the LDC and thus leads to higher LOLP and LOEP measures.

Figure 3.8
The Effect of Using DPLDC's vs. LDC's

Note that the $P(0_k)$ are still the same for both cases but in the normalization of the horizontal axis (the # of hours) to 1, each data point in the DPLDC now receives 24 times as high a fraction of the period. Thus the DPLDC approach is literally equivalent to assuming the daily peak load occurs throughout the whole day. Let's divide $\{0_k\}$ into those that cause losses of load when load is the minimum of the daily peak loads and those that do not and let us call the dividing line 0_D

$$LOLP_{DPLDC} = \sum_{0_k < 0_D} P(0_k) T_D(0_k) + \sum_{0_k > 0_D} P(0_k) T_D(0_k)$$

$$\text{for } 0_k < 0_D \quad T_D(0_k) = 24T_L(0_k)$$

$$0_k > 0_D \quad T_D(0_k) = 1 - T_L(0_k)$$

so clearly

$$LOLP_{DPLDC} > LOLP_{LDC}$$

also if most of $LOLP_{LDC}$ occurs for outages in the peak load range then

$$LOLP_{DPLDC} \approx 24 \cdot LOLP_{LDC}$$

(approximately equal to but smaller than)

This means that it makes a great difference in having a $LOLP = .1$ days/year whether one has used an LDC or a DPLDC for a load model. As we have seen, the latter could be 24

times as reliable as the former. Also, the interpretation made of the results must be carefully looked at. If $\text{LOLP} = .1$ days/year this really means $\text{LOLP} = 1/2600 = 3.85 \times 10^{-4}$ (using 260 week days/year as a conversion factor). If the hourly load duration curve is used, the $\text{LOLP} = 1/2600$ means that 1 hour in 2600 hours or one day (24 hours) in 10 years (2600 days \times 24 hours/day) will be the expected time of loss of load. If alternatively the DPLDC is used, the $\text{LOLP} = 1/2600$ will mean in 2600 hours of such peak load hours there will be on the average one with a loss of load. Now, as an extreme case, if the peak load hour is large relative to the rest of the day, we can say that it will take 2600 hours of such peak load hours, i.e., 10 years, to obtain one hour of loss of load. For this extreme case, $\text{LOLP}_{\text{DPLDC}} = .1$ day/year really means one hour/10 years and this is very different than one day in 10 years which is 24 hours/10 years, in fact, it is 24 times less expected outage time. The economic implications of these two numbers are very different indeed and the lack of careful analysis, formulation, and interpretation of results in the reliability field make our work more difficult than it should be.

Some planners justify their use of the DPLDC by saying that even if there is a one hour blackout they would count that as one day with a blackout. However, this approach does not allow for differentiation between sharply and flatly peaked load duration curves which should clearly be treated

differently.

Also it should now be obvious why previous results for LOEP were so much smaller than for LOLP. When using LOEP models LDC's were used and for LOLP models DPLDC models are used, this immediately introduces a factor of 20 into the normal LOLP-LP relationship.

Finally, industry has had much trouble with defining what the LOLP "one day in 10 year" criterion means. Some say it is the probability of having the system lose load. It is not. It really is the expected amount of time over a given time period that the system will have a generation deficit (without weighing how large the deficit). It also does not mean that the average frequency of occurrence will be once in ten years; LOLP is an expected fraction of time; 1/2600 means one day in 2600 days or 10 years. LOL events could occur very frequently for short time periods or infrequently for longer time periods.

In recent years there has been increasing realization that reliability targets for generation supply may have been set too high in contrast to those of transmission and distribution systems and there have been efforts to reduce such expenditures. Attention has focused around the definition of the loss of load event. Load is shed before available capacity goes under demanded load because satisfaction of the spinning reserve criteria require it. However, certain protective actions are usually taken that forestall actual customer

disconnection 'til \tilde{c} is actually five to ten percent below the demanded load. Let us illustrate this in the figure below.

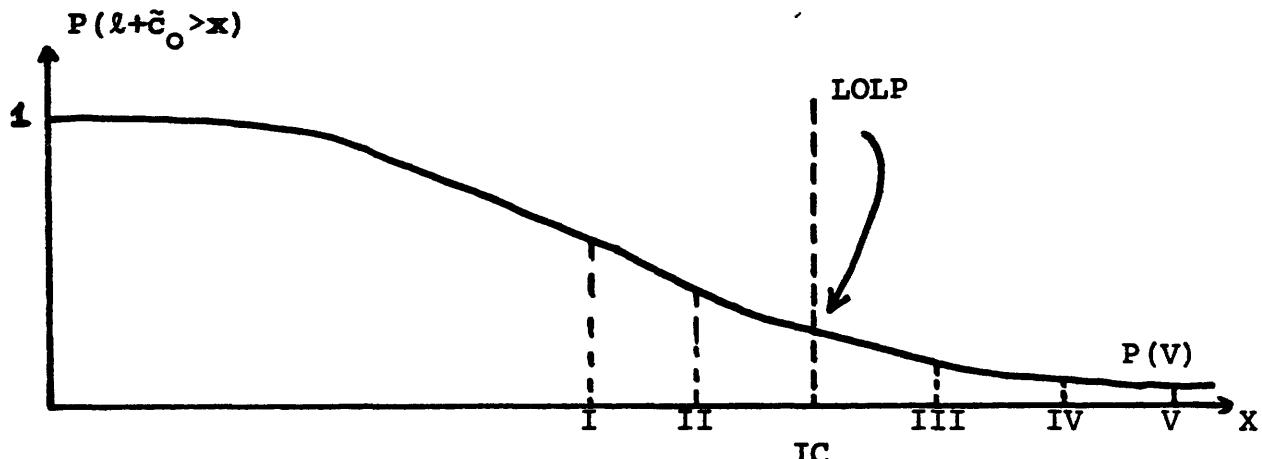


Figure 3.9 An Equivalent Load Distribution Curve and its Significant Margin States (not drawn to scale)

The figure depicts an equivalent-load distribution function which through a simple transformation ($m = IC - l - c_o$, or flipping it around its vertical axis and moving it IC to the left) can be thought of as the system margin distribution function. From this curve we can tell not only the system LOLP which is the "probability" that the system margin will be less than 0 but also the "probabilities" of all other margin states. This is important because different system loss of load procedures are initiated at different levels of available margin.

One utility has recently devised an innovative scheme which tries to better reflect the fact that customer loss of load occurs not for $m \leq 0$ but really for $m \leq -\theta < 0$.

Let us look at the figure again. The region of margin

to the left of I is a "normal operation" region, i.e., no loss of load procedures are initiated and normal economic criteria are the only ones followed. To the right of I any of a set of contingency loss of load operating procedures is initiated. For instance, for margins between I and II some unimportant initial actions are taken; the II to III region might include resort to a small voltage reduction; the III to IV region might call for some more drastic measure; the IV to V region might necessitate setting spinning reserves to 0 and finally the region past V calls for an actual disconnection. Clearly quality of service has suffered from the beginning of I and it suffers in progressively higher fashion until V is reached.

The specific idea of the utility is to plan in such a way that $P(V)$ --rather than LOLP--is targeted at a given reliability level, since it is at this point that a customer would actually lose load. We see that $P_{cum}(V) << P_{cum}(0) = LOLP$ which implies that the new criteria would lead to accepting higher levels of LOLP for generation adequacy than were previously deemed to be safe.

But if one gives this problem further thought, it should be apparent that the fundamental dilemma is still unresolved - what is so special about setting the expectation level for system outage time (newly defined at $M = V$ rather than $M = 0$) at $1/2600$ of the time in a given period? There is still no way to gauge how serious these outage events are when they occur, and clearly system outage time is only one of the

important parameters.

3.4.3 Interpretation of LOLP and LOEP as Expectations Rather than Probabilities

LOLP is a probability if and only if it is derived in the following way: it is the probability that on one trial of each of two random variables, \tilde{L} and \tilde{c} , that $\tilde{L} < \tilde{c}$, if c is described as a random variable with a density function and the same is done for \tilde{L} . However, \tilde{L} is not really a random variable when we describe it via an LDC, it is a series of expected values of the forecasted load and when we form LOLP based on that LDC we might as well have formed it by adding a component $LOLP_i$ for each load value in the LDC,

$$LOLP_{LDC} = \sum_{\text{each value of } \tilde{L}_i} LOLP_i(T_i)$$

where T_i is the fraction of total hours the $\tilde{L} = L_i$. But if we look closer, what we are doing is to calculate the probability of shortage for each hour (assuming trials generated as above) and multiplying the LOLP for each hour by 1, the time this

probability holds over, and summing over all L_i obtaining the expected number of hours on shortage and then dividing by the total hours to get LOLP-the % of total time in the period expected to be on shortage.

If we attempt to calculate LOLP for a given capacity model and two LDC's for different periods, it is equivalent to derive the LOLP for that capacity model and an LDC made up of the two previous LDC's

$$LOLP_1 = \sum_{0_k} P(0_k) T_1(0_k)$$

$$LOLP_2 = \sum_{0_k} P(0_k) T_2(0_k)$$

$$LOLP_1 \cdot H_1 = \sum_{0_k} P(0_k) [T_1(0_k) H_1] \quad \begin{matrix} H_i \\ \text{\# of hours} \\ \text{in period i} \end{matrix}$$

$$LOLP_2 \cdot H_2 = \sum_{0_k} P(0_k) [T_2(0_k) H_2]$$

$$LOLP_1 \cdot H_1 + LOLP_2 \cdot H_2 = \sum P(0_k) [T_1(0_k) H_1 + T_2(0_k) H_2]$$

but

$$T_1(0_k) H_1 + T_2(0_k) H_2 = T_{12}(0_k) (H_1 + H_2) \text{ by definition of LDC}_{12}$$

$$\frac{LOLP_1 H_1 + LOLP_2 H_2}{H_{tot}} = \sum P(0_k) T_{12}(0_k) = LOLP_{12}$$

so that the LOLP for the combination of two periods over which the capacity model remains constant is the weighted average of the separate period LOLP's. If $H_1 = H_2$ then LOLP_{12} is the arithmetic average, and if $\text{LOLP}_1 = \text{LOLP}_2$ then LOLP_{12} is also equal. It is much easier to think in terms of the $\overline{\text{TU}}$, the time expected unserved. It is clear that

$$\overline{\text{TU}}_{12} = T_{12}(0_k)(H_1 + H_2) = \overline{\text{TU}}_1 + \overline{\text{TU}}_2$$

So LOLP really yields the expected fraction of time the system will be in shortage. The same argument carries over for LOEP.

$$\text{LOEP}_1 \cdot E_{\text{tot}_1} + \text{LOEP}_2 \cdot E_{\text{tot}_2} = \text{LOEP}_{12} \cdot E_{\text{tot}_{12}}$$

$$\overline{\text{EU}}_1 + \overline{\text{EU}}_2 = \overline{\text{EU}}_{12}$$

$$\frac{\text{LOEP}_1 \cdot \frac{E_{\text{tot}_1}}{E_{\text{tot}_{12}}}}{E_{\text{tot}_{12}}} + \frac{\text{LOEP}_2 \cdot \frac{E_{\text{tot}_2}}{E_{\text{tot}_{12}}}}{E_{\text{tot}_{12}}} = \text{LOEP}_{12}$$

3.4.4 The Loss of Load Event

We have so far discussed the measures that exist, what they tend to measure, and the methods available for their computation. We have critiqued their application and interpretation and have settled upon LOEP as the more satisfactory measure of reliability for the purpose of measuring reliability benefits. At this point we would like to refocus attention on

the problem of definition of the loss of load event. We have so far naively assumed that a loss of load occurs when $\tilde{c} < \tilde{L}$, i.e., load must be shed when available on line capacity is smaller than demanded load.

There are at least two problems in this definition which work in opposite directions. First, load is shed before $\tilde{c} < \tilde{L}$, it is possible to shed load when in excess generation capacity mode in order to satisfy spinning reserve criteria. Second, load can be shed without necessarily causing partial blacking out of areas; it is possible to shed load through brownouts (voltage reductions) and through selective disconnection of customers who have backup such as aluminum companies on interruptible contracts and through disconnection of inessential in-house load such as utility company lights.

What is the effect of this realization of measure imperfection on our measure of system reliability? Assume a given system S, has been deemed to have reliability r_1 . By changing the definition of a loss of load event from $\tilde{c} < \tilde{L}$ to $\tilde{c} < \tilde{L} + S.R. (\tilde{L})$, we have effectively expanded the set of events that lead to losses of load, therefore $r_1' < r_1$ (or $LOEP' > LOEP$ for a given system S since LOEP decreases as system reliability increases).

The above fact that initial load shedding does not lead to serious social losses is not so much a problem with our measure of reliability, LOEP, as much as it is a problem with the measurement of the losses due to that level of LOEP.

To put it another way, the valuation function of the losses associated with LOEP should at first be smaller than that associated with larger losses of load. See Figure 6.3.

We will have more to say about this in the benefit evaluation section. From now on though, we will refer to:

- 1) LOEP as the measure of system reliability derived on the basis of defining a loss of load event as $\tilde{c} < \tilde{L}$ or

$$LOEP = \int_{IC}^{IC+MA} F_n(x)$$

$$F_n(x) = F_0(x) * f_1 * f_2 * \dots * f_n(x)$$

- 2) LOEP' as the measure of system reliability derived on the basis of defining a loss of load event as $\tilde{c} < \tilde{L} + S.R. (\tilde{L})$. If the spinning reserve criteria can be expressed as a fixed % of system load (one large utility does set S.R. = $.05\tilde{L}$, even though by doing this it disregards the problem of setting the criteria on the basis of units available to the system) then $\tilde{c} < 1.05\tilde{L}$ is the criterion and LOEP' can be calculated as before but

$$F'_n(x) = F_0(x/1.05) * f_1 * \dots * f_n$$

and

$$LOEP' = \int_{IC}^{IC+MA} F'_n(x) dx$$

and $F_0(x/1.05)$ is naturally "wider" than $F_0(x)$.

One way to view this is to assume the system load has grown by 5% above the previous one, i.e., it is as if a few months of load growth had gone on and the same capacity is still being used and naturally system reliability has decreased.

It should be apparent that $F'_n(x) > F_n(x) \forall x$, and in particular, system S is deemed to have $LOEP' > LOEP$.

3.5 Conclusions

Our problem has been to choose a measure of system reliability that will also be useful for the purposes of measuring the benefits due to that level of reliability to the system's consumers. As we have shown, FAD and LOLP concentrate on measuring the expected amount of time in a planning period that there will be a generation deficit without focusing on the seriousness of the events. We cannot begin to measure the value of reliability without knowing how bad expected outages are expected to be; the one day in 10 year LOLP measure could mean a 1 MWhr loss of energy or a 1000 MWhr loss of energy.

We will also argue that for reasonable ranges of outage duration, FAD information does not add much to benefit

valuation if the systems LOLP's are equal. The premise is that customer losses escalate in faster than linear fashion with respect to the length of the shortage duration and therefore the larger the mean shortage duration the worse the system. However, as long as the mean shortage duration and energy loss is not excessive, it is possible to rotate the shortage over the service area in such a way that consumers are not badly penalized. This is another way of saying that as long as total expected shortage time is the same for two system designs (equal LOLP measures), it does not really matter much if the outage time occurs in larger segments or in a greater number of smaller segments.

Because of the reasons discussed above and in our benefits section, we have chosen to use LOEP and LOEP' as suitable reliability measures for construction of reasonable benefit functions. There are a few drawbacks to using these, if there are identifiable fixed costs associated with each outage situation it might be better to use the frequency measure of FAD to develop that part of losses due to fixed costs incurred per shortage situation; however, typical systems' shortage frequencies are small and therefore there is little loss of accuracy by using the LOEP measure alone.

It is also worthwhile to point out that often times excess reserve margins are rationalized as effective protection against system wide shutdown. LOEP planning does not effectively reflect the losses incurred if this happens.

However, we would like to make the point that excess generation margins are:

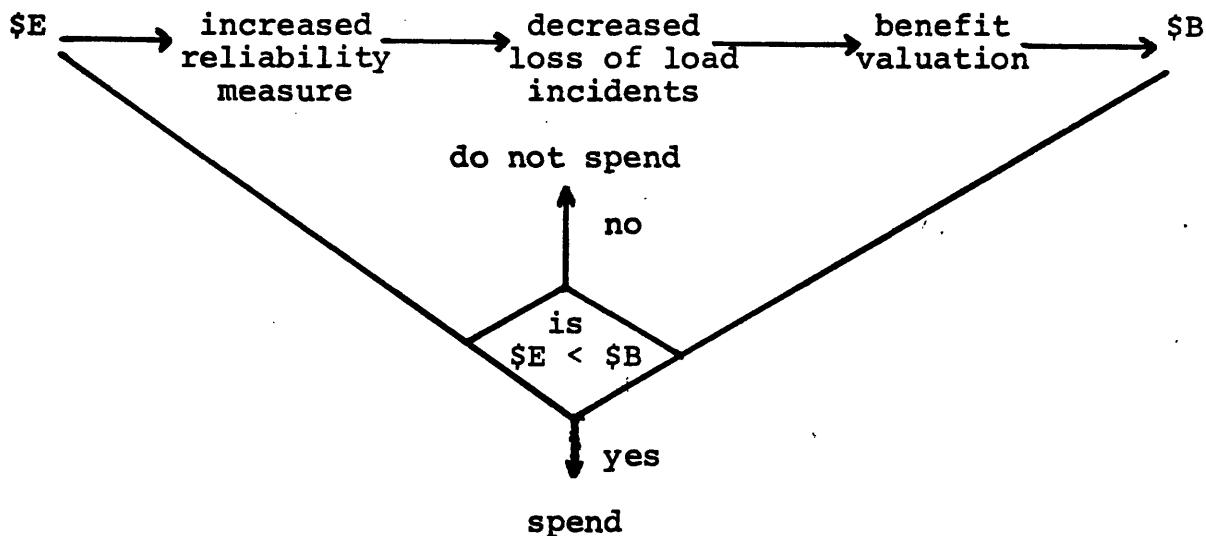
- 1) not an efficient method to accomplish protection against this phenomenon and
- 2) not wise in the sense that the only reasonable protection against this possibility is to devise effective counterstrategies (such as automatic load shedding devices, interconnections, etc.) to prevent this from occurring when a shortage situation occurs. Excess reserve margins only mean that shortage situations will probably occur less often, but when they occur the system may still be liable to wholesale shutdown in spite of the higher reserve margin if no effective counterstrategies have been devised.

Finally we should like to call attention to the problem of which load model seems most reasonable to use in order to calculate system LOEP measures. If most of the losses due to unreliable behavior occur during a portion of the day--for example the morning hours--it would be desirable to know what portion of these hours will be penalized through losses of load. Therefore it seems reasonable to use a load duration curve comprised of only the high use hours. An example will show what kinds of mistakes can occur if this is not done. Assume that on the basis of an hourly load duration curve a system is found to have a LOEP of 1/1000, but on the basis of

the high use morning hour LDC it has a LOEP of 1/500. It would not be accurate to associate the value of the societal output with the 1/1000 figure if most of it is produced at high load hours and is in fact subject to a higher (1/500) LOEP.

CHAPTER 4
THE COSTS OF ADEQUATE GENERATION

Our goal is to translate the dollars expended on reliability into reliability measures that can then in turn be related to possible incident reduction so that it becomes possible to properly evaluate its benefits; the purpose is to determine whether or not the expenditure is worthwhile. Below we display this concept in a flow chart.



By adopting this approach and focusing on incremental costs and benefits it is always possible for a system planner to discern whether or not there is a need for further expenditures on reliability at any point in time. Note that this is not equivalent to trying to decide whether total present

benefits are worth total reliability expenditures since we may or may not have reached the negative marginal improvement point.

Before we proceed to the first item, the dollars of expenditure, it is necessary to discuss certain issues. First, there is a need to identify what expenditures are related to what increases in reliability over what time period. There is a need to assign further out of pocket costs past the original capital expenditure.

Our approach will be to adopt a given expenditure, and proceed to levelize risk throughout the period of interest through proper maintenance scheduling. Note that optimally leveled risk of outage is not the objective as much as leveled risk of damages; the point is that damages at any given risk level during summer and winter extremes are likely to be more serious than those in the spring and fall, thus allowing for higher actual risk levels on off peak seasons. After leveling risk we will be able to compute the corresponding benefits.

There are several kinds of costs involved in our problem. First, the moneys can be expended on different types of generating equipment or on transmission equipment; obviously our interest is in the expenditure that minimizes the cost of the next increment of service reliability. Second, there are capital costs as well as operating costs, and furthermore the operating costs depend on the investments that have been made.

For instance, a higher capital cost expenditure may be undertaken to reduce operating costs sufficiently to make it worthwhile. We see that we need a discount rate to translate these flows of funds into commensurable units.

4.1 The Function C(r)

The problem we attempt to solve is 1) how to expand a given system throughout time (S) at 2) a given reliability target 3) given expected forecasted load to be served and 4) in such a way that total capital and operating costs are minimized in some sense (we minimize the total expected present value costs of purchasing and operating such a system). In other words, we would like to obtain the function $C(r)$ where C is the minimum present value cost of building and operating system S at reliability r .

How does this concept differ from what is presently done? At present, most utility expansion schemes are cost evaluators. In other words, a system is expanded throughout time subject to a generation mix guideline as well as to a unit size guideline while unit additions are triggered whenever the reliability criterion goes under a target level. The system (S) thus obtained is simulated to determine production costs and the expansion's capital and production costs are present valued and summed to determine $C(S)$. This routine is repeated several times for different assumptions regarding unit size guidelines, generation mix guidelines, inflation rates and

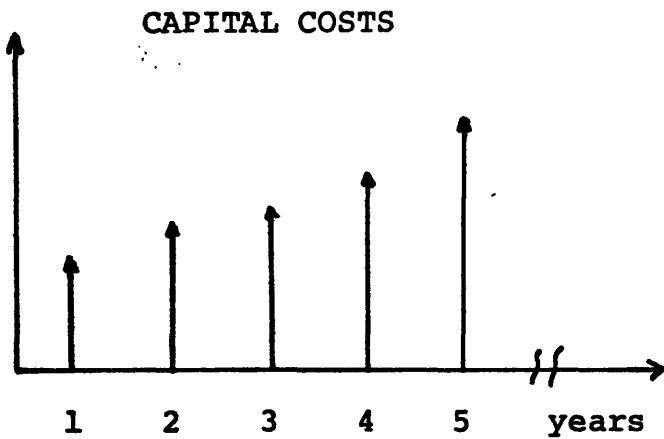
discount rates to determine the sensitivity of the resulting plan to these varying assumptions. Finally, by applying some skill and judgement it is usually possible to select a plan which seems to be close to lowest cost for the chosen reliability level, r .

Usually this is as far as most plans go, they do not seem to treat the reliability target as just as much of a parameter as the assumed discount rate. But this is precisely what we are interested in, $C(r)$. The approach we use substitutes a linear programming technique in place of the generation mix and size guideline and thus makes possible the formulation of several "least cost" expansions subject to all of the parameter assumptions that must be made.

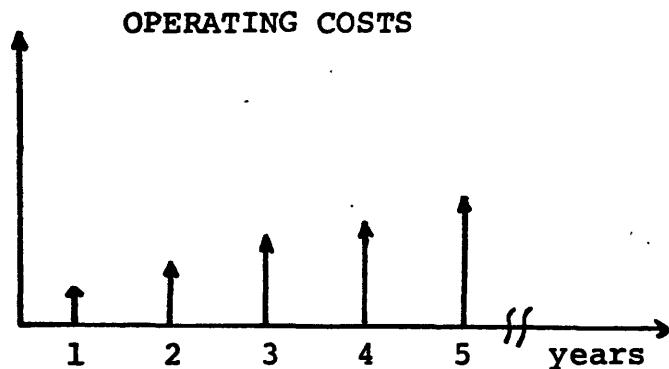
4.2 Expansions at Two Different Levels of Reliability

Let us assume that we have a technique that gives us a way to find $C(r_1)$ and $C(r_2)$ for two different levels of reliability, $r_2 > r_1$. Let us also graphically represent the capital and operating cost streams for the systems S_{r_1} and S_{r_2} . Let us refer to the present valued capital cost stream as CC_{r_i} and to the present valued operating costs stream as OC_{r_i} .

We derive the S_{r_1} expansion and its associated CC_{r_1} and OC_{r_1} cost streams. See the figures below.

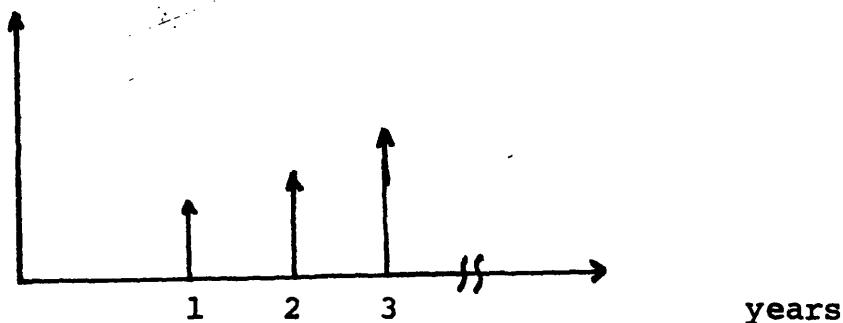


For S_{r_1} , if these streams are present valued they become CC_{r_1} and OC_{r_1} respectively and when summed, C_{r_1}



The system's losses, the benefits that are not obtained relative to a "perfectly reliable" system, are also portrayed in the figure below. They, of course, grow as the number of customers and energy delivered grows.

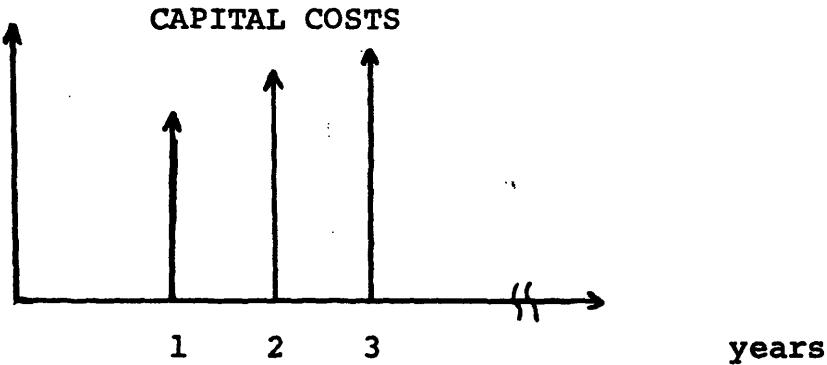
EXPECTED LOSSES FROM OPTIMAL



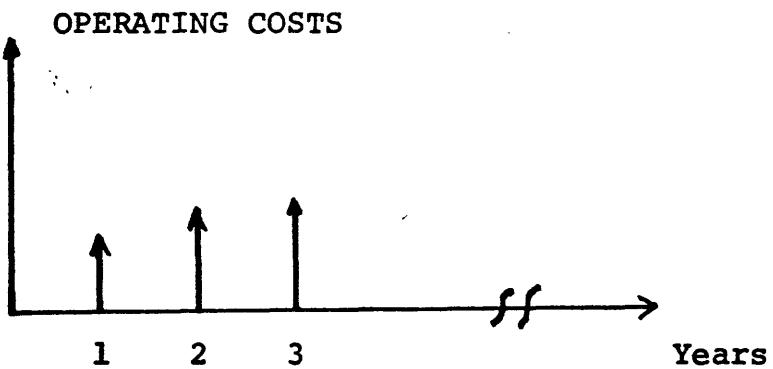
The present valued stream is L_{r_1}

Let us repeat the procedure and find a system S_{r_2} with reliability $r_2 > r_1$ and total present value cost C_{r_2} . Now $C_{r_2} > C_{r_1}$. If not, it would be possible to have higher reliability at lower cost thus violating the assumption of optimality for each system at its level of reliability. We portray its costs and benefits in the figures below.

CAPITAL COSTS

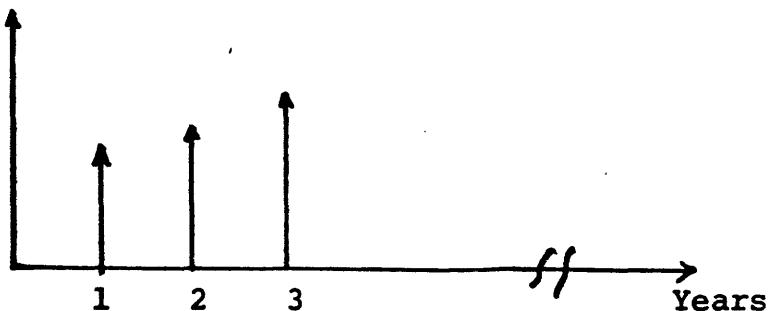


CC_{r_2} present value



OC_{r_2} present value

EXPECTED LOSSES FROM OPTIMAL



L_{r_2} present value

The procedure to find the optimal reliability is to increment r_2 in small steps until $C_{r_2} - C_{r_1} > L_{r_1} - L_{r_2}$. There are several things that can be said about the relationships between these streams of figures. Obviously $L_{r_1} > L_{r_2}$, because the more reliable the lesser the expected losses; furthermore, this must be true on a yearly basis.

Also, let us assume that by moving to a higher reliability level, the system composition does not vary significantly in its capital intensive vs. fuel intensive ratio during each year in the period. Specifically, assume that S_{r_2} contains

S_{r_1} , that is, at any point in time S_{r_2} has all that S_{r_1} had and then some. Then several other probable relationships can be deduced. For example, it is probably true that the system operation costs of S_{r_2} are no greater than those of S_{r_1} , year by year. Why? Because it would always be possible to operate the system as S_{r_1} was operated. This means that although $CC_{r_2} > CC_{r_1}$, $OC_{r_2} \leq OC_{r_1}$.

Also one can think of CC_{r_2} as simply consisting of the capital stream CC_{r_1} advanced in time. Both of these probable relationships make it possible to obtain estimates of $(C_{r_2} - C_{r_1})$ without going through the full blown calculation procedure to get S_{r_2} . For instance, if we think of CC_{r_2} as the stream CC_{r_1} advanced in time, then $CC_{r_2} = (1+d)^y CC_{r_1}$, where d is the applicable discount rate and y is the number of years advanced. We would also know that $OC_{r_2} \leq OC_{r_1}$ and we could probably resimulate system operation using the new schedule of plants in S_{r_2} against the expected loads in the planning period and thus compute OC_{r_2} . It might thus be possible to simplify the procedure computationally and avoid the long tedious calculations to derive S_{r_1} and S_{r_2} and their respective descriptions.

Another simplified approach for obtaining S_{r_2} from S_{r_1} is to focus on S_{r_1} and attempt to derive S_{r_2} from adding plants to S_{r_1} . Consider adding to S_{r_1} another plant, g_a .

If the basic expansion plan is fairly reliable, the cost implications of g_a will not be too difficult to estimate.

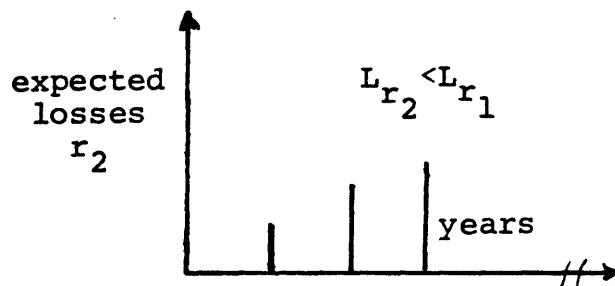
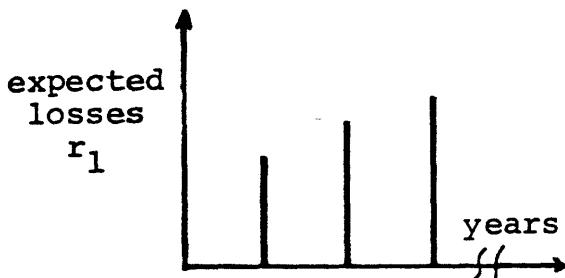
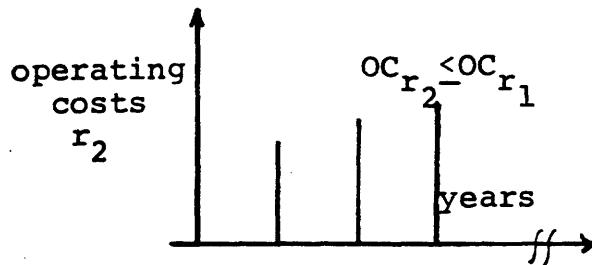
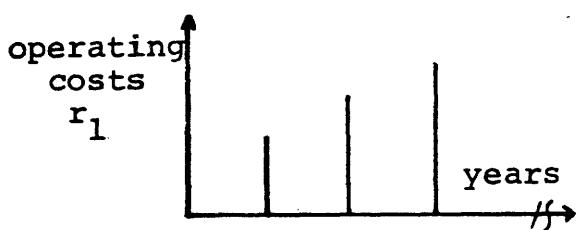
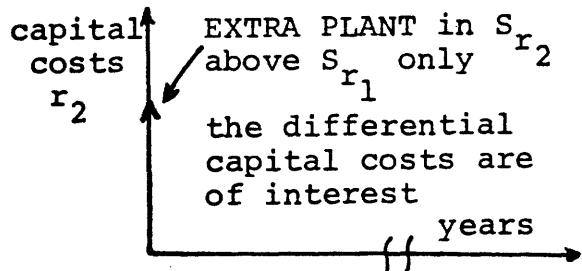
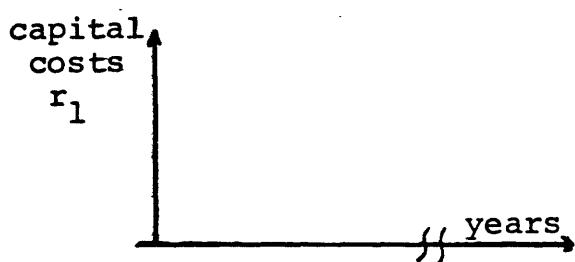
For reasons we will explain later, the differential operating cost implications will not be significant for marginal additions of this type. Therefore, the real differential costs will be those associated with the fixed and capital costs of a given project. The cheapest way to obtain increased reliability will be to purchase cheap capital cost, yet expensive to run, peaking gas turbines. (Some systems which are burdened with old, inefficient, base load capacity might find it convenient to shift these toward cycling peaking operation while acquiring efficient base load to take its place).

In other words, one way to get S_{r_2} from S_{r_1} is to add a few gas turbines to S_{r_1} till S_{r_2} is obtained; then calculate CC_{r_2} , compare it to CC_{r_1} and this becomes our cost differential.

4.3 Treatment of Marginal Changes to an Expansion Plan

To this point we have been comparing alternative system expansions at differing levels of reliability. It is valid to ask at any given point in an expansion plan, whether it would be wise to undertake a further project. If an additional plant were to be built, it will have both capital and operating cost implications, and the benefits redounding from this action should presumably be sufficient to offset the additional net present value cost. Later in the text we will present ways to approximate or calculate the net present value cost differentials associated with such actions. In the target level determination section we show how to compute the differential

reliability measures for each period and how to evaluate the resulting differential benefits before discounting them appropriately and summing them over the plant life. These differential benefits are compared to the differential costs to determine the advisability of a given action.



Several relationships are much easier to discern now; the capital cost difference ($CC_{r_2} - CC_{r_1}$) is simply the cost of the additional plant. Also, it clearly follows that $OC_{r_2} < OC_{r_1}$, and that $L_{r_2} < L_{r_1}$ during every year in the horizon. The

question simply is:

$$[C_{gen} + (OC_{r_2} - OC_{r_1})] \stackrel{?}{<} \sum_i [L_{r_1,i} - L_{r_2,i}] (1/(1+d))^i$$

years, i, in plan

But note carefully that this approach is much more powerful than one would think at first. For example, S_{r_1} and S_{r_2} need not be optimal expansion schemes, i.e., any two schemes can be compared in terms of their ΔC and of their ΔB . If $\Delta C < \Delta B$ then it pays to make the change, if not, it doesn't. This realization gives the analyst a powerful relationship which he can use even if he does not have a way of generating $C(r)$ as long as he has a method for evaluating ΔC and ΔB .

Further, this approach yields an interesting marginality condition for determining whether or not a given system expansion is close to optimal. If a given system is close to optimal, it should not be possible to undertake any further actions and receive benefits in excess of their cost.

4.4 Marginal Change Cost Analysis

So far we have discussed methods for evaluating the cost implications of strategies to expand power systems at targeted levels of reliability which are held constant throughout the period of interest. The question of interest has been "what is the total present value cost difference between a strategy of expanding the power system at $LOEP = 10^{-4}$ versus

the strategy of using $LOEP = 10^{-5}$ ".

Now we would like to explore the differential cost implications involved in making the decision whether to build one more unit above any particular expansion plan $S(t)$. Clearly there will be differential capital and operating costs. The capital costs will per force be higher than thos of $S(t)$ alone; in fact, higher by the cost of the unit g. (We assume all capital costs are appropriately discounted so that they coincide with the moment the plant is available for generation.)

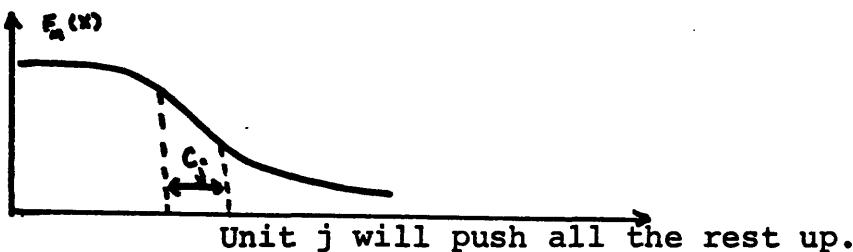
The expected operating costs, however, will be lower simply because it would always be possible to do at least as well as previously expected. Note that one should include as an operating cost, the costs of energy unserved valued at the highest of costs for generation; otherwise no generation at all would yield the cheapest operating cost system.

By how much will the expected operating costs be lower? In the PROSIM discussion we show that it is possible to solve exactly for this new expectation, however, it is usually quite expensive computationally to do this and total expected costs will likely not change much for small proposed generation changes.

It is possible, though, to set some limits to this difference. First we shall note that if the additional plant is a peaking plant which is naturally loaded at the top of the

loading order, its expected operating cost can be determined directly and the expected operating costs of plants further down the line will remain unchanged. In fact, the closer to the top of the loading order the plant is, the lesser the change in expected operating costs.

If the unit is a base loaded unit, then the expected operating cost will be lower. In effect, the base loaded unit will allow for displacement of more expensive to run capacity upward in the loading order providing savings throughout the loading order above itself, plus generating some extra energy that would otherwise be unserved.



We know how much new energy will be expected to be generated ΔEU , in total, = \bar{E}_{n+1} but the real savings come in because all the units above the new one in the loading order will be producing lesser amounts of energy than before, thus the new one will be producing a little bit instead of the other higher cost ones that was previously the case.

To see this, consider the calculations made if the unit is loaded at the end of the loading order. Energies produced will be as follows:

$$\bar{E}_1, \dots, \bar{E}_{j-1}, \bar{E}_j, \dots, \bar{E}_n, \bar{E}_{n+1}$$

(the last is the energy produced by the unit out of position)

If we recalculated everything the unit would be placed say at the jth position displacing every unit above it to a lesser producing situation.

$$\bar{E}_1, \dots, \bar{E}_{j-1}, \bar{E}'_j, \bar{E}_{j+1}, \dots, \bar{E}'_n, \bar{E}'_{n+1}$$

Clearly unit i corresponds to unit $(i+1)'$ in the new ordering. We are interested in computing the difference between

$$\sum_{k=1}^{n+1} C_0 k \bar{E}_k \text{ and } \sum_{k=1}^{n+1} C_0' k \bar{E}'_k$$

It is obvious that terms up to unit $j-1$ are the same for both equations, we can also see that for units j and above

$$\begin{aligned} \bar{E}_k &\geq \bar{E}'_{k+1} \\ C_0 k &= C_0' k+1 \end{aligned} \quad \text{for } k = j \rightarrow n$$

Then we can see that

$$\begin{aligned} \frac{\text{Operating Cost Savings}}{T} &= [C_0 - C_0'] = \left[\sum_{k=1}^{n+1} C_0 k \bar{E}_k - \sum_{k=1}^{n+1} C_0' k \bar{E}'_k \right] \\ &= C_0_{n+1} (\bar{E}_{n+1} - \bar{E}'_j) + \sum_{k=j}^n C_0 k (\bar{E}_k - \bar{E}'_{k+1}) \end{aligned}$$

It should be clear at this point that the first term is negative because in the new loading order the unit will be loaded as a base load rather than peak load unit and that all the terms in the second sum will be positive since all units above the newly inserted j' unit will be operated less than before. Note that it is not a simple matter to calculate these expressions; in order to calculate the first we must deconvolve all the units above unit $j-1$ out and then form

$$\bar{E}'_j = P'_j \int_{S_{j-1}}^{S_{j-1} + C'_j} F_{j-1}(x) dx$$

We might be lucky and get unit j' loaded first in which case no difficult work is involved since F_{j-1} becomes in fact $F_0(x)$. But barring this simplifying case it is not an easy chore.

However, we can approximate this term. If we know the place of unit j' in the loading order we know that

$$\begin{aligned} \bar{E}'_j &= P'_j \int_{S_{j-1}}^{S_{j-1} + C'_j} F_{j-1}(x) dx \\ &= P'_j \cdot C'_j \cdot F_{j-1}(\varepsilon) \text{ for } S_{j-1} < \varepsilon < S_{j-1} + C'_j \end{aligned}$$

Since $F_{j-1}(\varepsilon) \leq 1$ we can bound this first term

$$C_0_{n+1}(\bar{E}_{n+1} - \bar{E}'_j) \geq C_0_{n+1}(\bar{E}_{n+1} - P'_j C'_j)$$

We will use this result later.

The second term which consists of a series of sums can be bounded in a different way; since all the terms are positive clearly

$$\sum_{k=j}^n c_0_k (\bar{E}_k - \bar{E}'_{k+1}) \leq c_0_n \sum (\bar{E}_k - \bar{E}'_{k+1})$$

because $c_0_n \geq c_0_k$ $k = j \rightarrow n$ because of the loading order assumption.

Further, we have a very useful result from the conservation of energy produced principle, i.e.,

$$\sum_{k=1}^{n+1} \bar{E}_k = \sum_{k=1}^{n+1} \bar{E}'_k$$

and thus

$$(\bar{E}'_j - \bar{E}_{n+1}) = \sum_{k=j}^n (\bar{E}_k - \bar{E}'_{k+1})$$

then we can say that

$$\sum_{k=j}^n c_0_k (\bar{E}_k - \bar{E}'_{k+1}) \leq c_0_n (\bar{E}'_j - \bar{E}_{n+1})$$

Therefore we obtain

$$c_0 - c_0' = c_0_{n+1} (\bar{E}_{n+1} - \bar{E}'_j) + \sum_{k=j}^n c_0_k (\bar{E}_k - \bar{E}'_{k+1})$$

$$\begin{aligned} C_0 - C_0' &\leq C_{n+1}(\bar{E}_{n+1} - \bar{E}'_j) + C_n(\bar{E}'_j - \bar{E}_{n+1}) \\ &\leq (C_n - C_{n+1})(\bar{E}'_j - \bar{E}_{n+1}) \end{aligned}$$

Note both terms are greater than 0. $C_0 > C_{n+1}$ because C_{n+1} is assumed to be out of its correct loading order.

Now we can make use of the previously derived fact that

$$P_{n+1}C_{n+1} > \bar{E}'_j$$

$$C_0 - C_0' \leq (C_n - C_{n+1})(P_{n+1}C_{n+1} - \bar{E}_{n+1})$$

We can find \bar{E}_{n+1} ; we know P_{n+1} , C_{n+1} , C_0 and C_{n+1} . If need be, we can make a closer estimate to \bar{E}'_j than $P_{n+1}C_{n+1}$ by looking at the loading order position.

What are the quantities involved in a typical situation? Assume we have a unit, $n+1$, which will be loaded at an intermediate point on the LDC; assume that we derive an estimate of \bar{E}'_j as $(P_{n+1}C_{n+1})(1/2)(.5 = F_{j-1}(\varepsilon))$ for $s_{j-1} < \varepsilon < s_{j-1} + C'_j$. Assume further that although \bar{E}_{n+1} is found in a given problem, that it is negligible for our present purposes. Finally assume that $(C_0 - C_{n+1})$ is large, say it is approximately the system average production cost/kwhr. Then we see that

$$C_0 - C_0' \leq (\text{average cost/kwhr})(P_{n+1}C_{n+1})(1/2)$$

In terms of total production costs C_0 we see that

$$C_0 \approx (\text{average cost/kwhr}) \cdot (\text{total capacity}) \cdot (\text{system load factor})$$

$$\frac{C_0 - C_0^*}{C_0} \approx \frac{(P_{n+1} C_{n+1})^{1/2}}{IC \cdot (L.F.)} \approx \frac{P_{n+1} C_{n+1}}{IC}$$

If the unit is an intermediate unit, this will seldom be more than 3% and this bound vastly overestimates the cost difference. Therefore we can be sure that we are not making a bad assumption when we neglect operating cost differences for marginal comparisons.

4.5 Construction of C(LOEP)

4.5.1 The Iterative Search Procedure for Determining an Expansion Strategy

Our approach for obtaining an electric system expansion overtime, at a given level of LOEP, differs from the traditional one, and we outline it below. Given an arbitrary set of assumptions regarding system parameters such as capital costs of different types of generation, their forced outage rates, rates of inflation in their fuel and capital costs, etc., we allow a linear programming technique (LP) to pick the least present value cost system. The LP is given capacity and energy production constraints to meet and it then, in essence, substitutes for the work that goes into producing the size and mix guidelines in the traditional expansion approach.

In order to do this, we must give the LP some

indication of what energy production costs will be associated with the inclusion of a given unit into an expansion plan. However, we cannot know the various units' production histories *a priori* because they are really a function of the other units available to serve the load. In order to solve this problem, we will assume that we know, *a priori*, the production histories associated with each type of unit. If the LP solution is subject to meeting a given energy and capacity constraint in every period of the plan, the LP will be able to associate a total capital plus production cost to any given plant alternative it chooses to include in the expansion plan and it will then be able to proceed to derive an optimal expansion plan subject to these assumptions.

The LP has not explicitly dealt with the issue of reliability. There is no way to model the reliability constraint in the LP since it is an inherently non-linear problem. Because of this, we attempt to model this via the use of a peak load protection margin constraint. The constraint requires that the sum of the expected nameplate capacity in year k must be larger than, or equal to, the peak load plus a protection margin in each year.

$$\sum_{Vi} P_i C_i(k) \geq [1+\epsilon(k)] \text{ peakload}(k), \forall \text{ years } k \text{ in the horizon}$$

We do not know at this point whether the capacity factor history assumptions or the $\epsilon(k)$ assumptions that were

made to derive the LP results are in fact reasonable. Because of this, we must evaluate the operation of the LP solution expansion system with probabilistic simulation. PROSIM will tell us whether the assumed capacity factor histories were reasonable assumptions and whether the $\epsilon(k)$ assumptions led to the desired targeted LOEP levels. If not, then we must combine the new knowledge gained through the application of PROSIM with our previous assumptions to derive new assumptions which will provide the drive for a new LP expansion. This process will be repeated until we are satisfied that we have a near optimal plan. See Figure 4.1.

What kinds of modifications need to be made on the peak load constraints $\epsilon(k)$ and on the capacity factor histories capfac (Strictly speaking this is a matrix; for each type of plant there is an assumed time dependent capacity factor history.) from iteration to iteration? We will use the following terminology:

- 1) $\epsilon_m(k)$ is the peakload capacity constraint in the LP formulation for year k, on iteration m.
- 2) capfac is the assumed capacity factor history for the mth LP iteration.
- 3) $LOEP_t$ is the targeted LOEP reliability level.
- 4) $LOEP_{sim}(k)$ is the evaluated LOEP for the presently considered LP expansion, in period k.

What properties should the algorithm that enables us to move from the present assumed set of $\epsilon(k)$ inputs to those

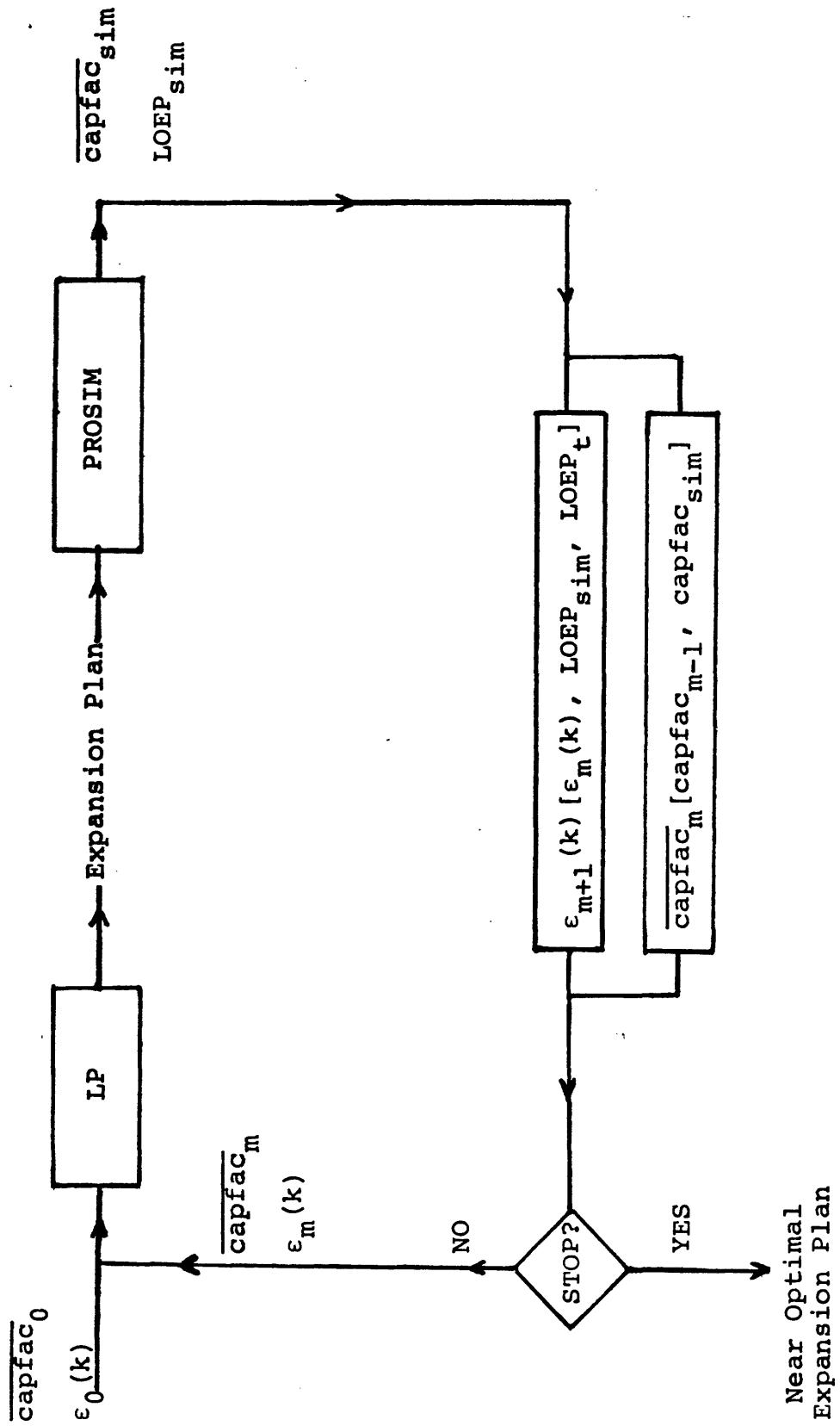


Figure 4.1 Iterative Search for Expansion Strategy

which will be made during the next iteration? Clearly, the farther $\text{LOEP}_{\text{sim}}(k)$ is from LOEP_t , the greater the correction we will want to make to $\epsilon_m(k)$. It should be easy to see that if system structure does not change much from iteration to iteration, a higher $\epsilon_{m+1}(k)$ will cause a lower $\text{LOEP}_{\text{sim},k}$ and vice versa. Therefore, a reasonable procedure to attempt is to set:

$$\epsilon_{m+1}(k) = \epsilon_m(k) [1 + C\Delta(k)]$$

where

$$\Delta(k) = (\text{LOEP}_{\text{sim}}(k) - \text{LOEP}_t)/\text{LOEP}_t$$

Do the above for all periods k , and stop modifying the ϵ for year k if

$$|\Delta(k)| < .05$$

In other words stop changing the ϵ for year n if the calculated LOEP for year n is within 5% of the target measure. Note that we will converge to a solution (when we engage the stopping rule) if C is small, but it may take long. The larger C is, the faster we will converge, but we may get into oscillations while finding the solution. Note also that whenever $\text{LOEP}_{\text{sim}}(k)$ is below LOEP_t , that is, the system is too reliable for period k , $\Delta(k)$ will be negative thus yielding as we expect

$$\epsilon_{m+1}(k) < \epsilon_m(k)$$

Since LOEP goes through an order of magnitude variation over changes in ϵ of a few percent, it would be wise to set $C < 1$; i.e., if $\Delta(k)$ is 100% we don't want to change $\epsilon(k)$ by that much when a smaller change will probably be sufficient.

4.5.2 Example of Iterative Search Procedure for Expansion Strategy

The material we have just explained is quite complicated and we have included the procedure we went through to develop a few system expansions at differing levels of reliability. All of the expansions start from the same initial system. Each starting system would of course lead to very different expansion strategies. Because of this, our purpose here is to show the feasibility of our approach in general terms.

We have simplified the problem of determining the optimal expansion strategy over a 30 year period into a 6 period one, each of which represents five years. The LP is asked to yield an optimal 10 period (50 year) expansion plan subject to the capacity factor history and the peak load constraint assumptions. The LP is allowed to choose from any of 10 types of plants and then for every period decides what kinds of plants it will build in that period and how many of each. Because of the structure of the problem, the LP chooses two types of plants in each period. A complete specification of

the LP model designed by Woodruff and Farrar can be found in Appendix 4.1. We pick a 10 period horizon even though we're only interested in 6 because we are trying to eliminate end effects.

When we obtain this 6 period plan we submit it to a probabilistic simulation run to check whether or not the capacity factors were correct and whether or not the peakload constraints had been correctly set to ensure LOEP target satisfaction. If the LOEP targets had not been realized we would focus on how to change the peakload constraints so as to ensure that we met them.

We first attempted system expansions at various constant $\epsilon(k)$; we tried LP runs for $\epsilon(k) = 6\%, 8\%, 10\%$ and 15% . We found that the basis, that is, the kinds of plants that were to be built in every period did not change over the ϵ range although their numbers did. The higher the ϵ , the more peak load type plants would be built and maybe one or two of the base loaded ones in each period would drop out. See Figure 4.2. Since the LP formulation was continuous, we then had to round off the number of plants in each category in preparation for resuming the PROSIM evaluation. The simulation showed that as the system grew from a peak load forecast of 20,000 MW at an 8% compounded growth rate (while keeping the same load duration curve shape), the LOEP decreased from period to period. This was not a surprising result because it is usually possible to reduce system margins as the system grows

		<u>ϵ STRATEGY</u>							
Period	Plant Type	No. of Plants	$\epsilon = 6\%$	$\epsilon = 8\%$	$\epsilon = 10\%$	$\epsilon = 15\%$	No. of Plants	No. of Plants	No. of Plants
1975-79	100 MW Gas Turbine	12		24		28		49	
	800 MW Fossil Peaker	11		10		10		9	
	1000 MW Fossil Peaker	4		4		5		6	
	1500 MW Fossil Base	5		5		5		4	
	100 MW Gas Turbine	48		50		54		62	
	1500 MW Fossil Base	8		8		8		8	
1985-89	100 MW Gas Turbine	80		85		90		116	
	1500 MW Fossil Base	12		12		12		11	
	1000 MW Fossil Peaker	18		20		21		26	
	2000 MW Nuclear	12		11		11		9	
1990-94	100 MW Gas Turbine	103		121		139		191	
	2000 MW Nuclear	23		23		23		22	

FIGURE 4.2 LP OUTPUT FOR INITIAL TRIAL ϵ STRATEGIES

and because our original system was quite small to begin with.

This initial set of runs gave us a feel for the range of $\epsilon(k)$ needed to achieve a given LOEP(k). Our next experiment involved setting different $\epsilon(k)$ within a single LP run to see whether we could achieve a LOEP expansion. We found a somewhat surprising result. Within reasonable limits, a given $\epsilon(k)$ would yield an expected LOEP(k) independent of what the $\epsilon(k)$ were in the other periods. This was an important result because it meant that we could modify all the $\epsilon(k)$ at once without having to pay attention to interactions between the $\epsilon(k)$. On the basis of our results we designed six different expansion strategies (six sets of $\epsilon(k)$). Three of these so called ϵ strategies yielded expansions at reasonably constant reliability levels. See Figures 4.3 and 4.4 for the summaries of the six expansions. We modified these strategies a bit further by inspection, i.e., without having to return to the LP and the resulting figures are included in Appendix 4.2.

Note that most of the time, for the particular LDC we used, there is a linear 20 to 1 relationship between LOLP and LOEP; this is not true in general but may be over certain ranges of LOLP and LOEP.

Once you have the desired expansion, it is a simple matter to compute its total present value cost by:

- 1) discounting the cost of the plants the LP instructs to build and summing in order to derive the capital cost component.

6 STRATEGY						
Period	Plant Type	No. of Plants				
100 MW	Gas Turbine	11	24	29	34	52
1 800 MW	Fossil Peaker	6	8	10	10	18
1000 MW	Fossil Peaker	11	10	10	9	65
2 1500 MW	Fossil Base	6	4	9	4	4
100 MW	Gas Turbine	5	5	5	5	6
3 1500 MW	Fossil Base	5	6	7	42	4
100 MW	Gas Turbine	43	40	7	8	44
4 1500 MW	Fossil Base	5	8	8	8	44
100 MW	Gas Turbine	74	76	78	77	75
5 2000 MW	Fossil Base	4.5	5.5	6.5	7	82
1000 MW	Fossil Peaker	12	12	12	12	10
5 2000 MW	Nuclear	15	15	17	18	18
100 MW	Gas Turbine	4	5	6	6.5	8
6 2000 MW	Nuclear	13	13	13	12	9
100 MW	Gas Turbine	85	98	111	120	135
6 2000 MW	Nuclear	3.5	4.5	5.5	6.0	7
1000 MW	Fossil Peaker	23	23	23	23	149
1000 MW	Nuclear	23	23	23	23	8
1000 MW	Gas Turbine	23	23	23	23	23

FIGURE 4.3 LP OUTPUT FOR THE SIX ATTEMPTED STRATEGIES

		ϵ STRATEGY										
		1			2			3				
Period	ϵ	LOLP	LOEP	ϵ	LOLP	LOEP	ϵ	LOLP	LOEP	ϵ	LOEP	
1	6	1.5×10^{-2}	9×10^{-4}	8	9×10^{-3}	5×10^{-4}	10	5×10^{-3}	3×10^{-4}			
2	6	1.3×10^{-2}	8×10^{-4}	8	9.5×10^{-3}	5.4×10^{-4}	9	6.6×10^{-3}	3.6×10^{-4}			
3	5	1×10^{-2}	5×10^{-4}	6	9.5×10^{-3}	4.6×10^{-4}	7	6.3×10^{-3}	2.9×10^{-4}			
4	4.5	4×10^{-3}	1.6×10^{-4}	5.5	2.1×10^{-2}	1×10^{-3}	6.5	2×10^{-3}	8×10^{-5}			
5	4.0	7×10^{-4}	2×10^{-5}	5	3×10^{-3}	1×10^{-4}	6	3×10^{-4}	8×10^{-6}			
6	3.5	2×10^{-2}	9×10^{-4}	4.5	3.2×10^{-2}	1.5×10^{-3}	5.5	9×10^{-3}	3×10^{-4}			
		4			5			6				
Period	ϵ	LOLP	LOEP	ϵ	LOLP	LOEP	ϵ	LOLP	LOEP	ϵ	LOEP	
1	12	3×10^{-3}	1.5×10^{-4}	16	8×10^{-4}	4×10^{-5}	18	3.8×10^{-4}	1.6×10^{-5}			
2	10	$4,4 \times 10^{-3}$	2.3×10^{-4}	13	2×10^{-3}	9×10^{-5}	16	6.6×10^{-4}	2.8×10^{-5}			
3	8	4.1×10^{-3}	1.8×10^{-4}	10	2×10^{-3}	9×10^{-5}	13	7.1×10^{-4}	2.6×10^{-5}			
4	7	1.1×10^{-3}	3.6×10^{-5}	9	6.6×10^{-4}	2×10^{-5}	10	3.9×10^{-4}	1.2×10^{-5}			
5	6.5	1.4×10^{-4}	3×10^{-6}	8	6×10^{-5}	1×10^{-6}	9	3.9×10^{-5}	1×10^{-6}			
6	6	4.6×10^{-3}	1.4×10^{-4}	7	1.9×10^{-3}	5×10^{-5}	8	9×10^{-4}	2.5×10^{-5}			

Figure 4.4 PROSIM Output for the Six Attempted $\epsilon(k)$ Strategies

2) discounting the fuel costs computed by PROSIM in each year and summing in order to derive the fuel cost component.

Note that ε strategies 1, 4 and 6 are reasonably close to being expansions attempted at

$$1 - 8 \times 10^{-4} = LOEP_1$$

$$4 - 2 \times 10^{-4} = LOEP_4$$

$$6 - 2 \times 10^{-5} = LOEP_6$$

Further modification of the expansion plan at this point, without returning to the LP can yield the desired system expansion.

4.6 An Example of a C(LOLP) Study

The New York Power Pool recently conducted a simplified expansion study in which they computed the cost of expanding a system over 20 years at different levels of LOLP criteria by using 600 MW unit additions which were triggered by violation of the LOLP criteria.*

We can express the relationship they found between the cumulative present worth of annual charges in 1981 dollars and LOLP as approximately:

$$C(LOLP) = [34.5 - .75 \log (260 \text{ LOLP})] 10^9^{**}$$

for $1/260 > \text{LOLP} > 1/26,000$

* Personal communication: Mr. G. D. Garcy for the N.Y.P.P., Dec. 1972.
**Curve fit by author

$$C(1/260) = 34.5 \times 10^9$$

$$C(1/26,000) = 36 \times 10^9$$

This can be rewritten as

$$C(LOLP) = (C_0 - .75 \log LOLP) 10^9$$

Note then that

$$\frac{d/dLOLP [C(LOLP)]}{LOLP} = - \frac{(2.3)(.75) 10^9}{LOLP}$$

If we assume that the LDC is such that LOEP is linearly related to LOLP, say $LOEP = a LOLP$ (and a was about $1/20$ in the sample runs we prepared), then

$$\frac{d/dLOEP [C(LOEP)]}{LOEP} = - \frac{(2.3)(.75) 10^9}{LOEP}$$

This expression simply says that as LOEP increases within the specified range, the derivative of C , i.e., the marginal cost for providing a unit of reliability goes down as $1/LOEP$. In other words, the lower the reliability level, the cheaper it is to increase the reliability by a given amount. As the reliability gets higher ($LOEP$ approaches 0) the marginal cost gets very large.

Let us rewrite the expression as

$$C(LOEP) = C_0 - C_1 \log a - C_1 \log LOEP$$

We can see that the only parameter of interest in the derivative is C_1 and we know from elsewhere that reasonable bounds can be placed on C_1 . Look at

$$C(LOEP) - C(10 LOEP) = C_1$$

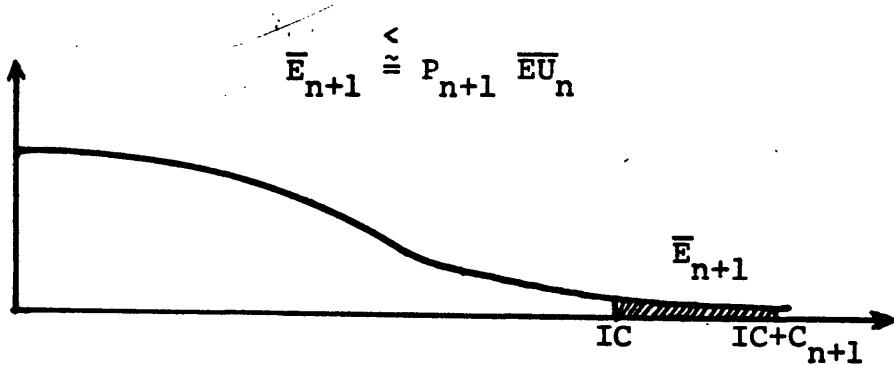
This means that C_1 is the cost difference between systems expanded at LOEP levels one order of magnitude apart. For reasonable levels of LOEP, and from experience, we know that the cost differences are within a few percent of each other

and the capital expenditure difference is only slightly greater. For example, the New York Power Pool example has C_1 set at $.75 \times 10^9$ out of a total of 36×10^9 . This knowledge effectively means that we have an independent order of magnitude check on $C(LOEP)$.

4.7 Theoretical Justification for $C(LOEP)$ Form

There is a reasonable theoretical basis to believe that $C(LOEP)$ is a logarithmic function of $1/LOEP$. We will explain why by reference to the figure below. Let us assume that a system has been expanded to reliability $LOEP_a$ where for all the periods in the plan \overline{EU}_n is not too large. Let us contemplate adding a unit which if perfectly reliable would drive \overline{EU}_{n+1} very close to 0; this can be done because by assumption \overline{EU}_n is small already. Note that the unit $n+1$ is not perfectly reliable, it has performance probability of P_{n+1}

and therefore



this implies that

$$\overline{EU}_{n+1} \approx (1-P_{n+1}) \overline{EU}_n$$

If we repeat the process, we see that

$$\overline{EU}_{n+2} \approx (1-P_{n+2})(1-P_{n+1}) \overline{EU}_n$$

What is this relationship telling us? It says that if we spend \$ x for unit n+1 we can reduce energy unserved, and therefore system LOEP, by $(1-P_{n+1})$ and that if we expend another \$ x for unit n+2 it is possible to reduce energy unserved and therefore system LOEP by another $(1-P_{n+2})$ fraction. In short, equal increments in the cost variable bring equal multiplicative fractions on the inverse of LOEP. We say that LOEP is an exponential function of expenditure or vice versa that cost is a logarithmic function of LOEP past a certain point.

$$LOEP \approx LOEP_a e^{-(m-R_0)/a}$$

Where m is system margin and every a increase in system margin, past margin equal R_0 , brings about an e^{-1} reduction in LOEP.

Therefore,

$$-(m-R_0)/a = \ln (\text{LOEP}/\text{LOEP}_a)$$

$$-(m-R_0) = a \ln (\text{LOEP}/\text{LOEP}_a)$$

Since system cost is related in rough linear fashion to system margin m

$$C(m) = C_0 + C \cdot (m-R_0)$$

$$C = C_0 + C[-a \ln (\text{LOEP}/\text{LOEP}_a)]$$

$$C(\text{LOEP}) = C_0 + C \ln \left(\frac{\text{LOEP}_a}{\text{LOEP}} \right)$$

CHAPTER 5

THE BENEFITS THAT RESULT FROM ELECTRIC POWER SYSTEM RELIABILITY

5.1 Framework for Analysis

Reliability of electric power supply is one of the quality indices of electric service. Presumably, the more reliable the service becomes the more money consumers are willing to pay for it. The converse of this proposition (i.e., the less reliable, the less people are willing to pay) is a fact and it is reflected in the lower fees that utilities charge for interruptible service contracts. In this chapter we will identify and discuss the kinds of reliability benefits that accrue to the various consumers of electricity. We will also show that some benefits also obtain to the producers, i.e., that the investment in reliability allows for greater sales of power, thus recouping some of the investment. In the remainder of the chapter, we will discuss the issues involved in evaluating the benefits discussed.

No one questions the fact that there are societal benefits to having a more reliable electric power system. The real problem is how to compare those benefits to the costs involved in producing the benefits. Before we begin to evaluate the benefits we need to develop a framework for identifying

what we shall call benefits and why, as well as what and who do they accrue to.

The first point to be made is that benefits accrue to actions that are taken with respect to the reliability of the system. In this work we want to focus on discussing the benefits due to two different types of actions.

- 1) The action of expanding an electric power system over time at reliability level r_2 versus expanding it at r_1 and
- 2) The action of adding another generator unit to any given system (see the figure below). Each of these actions has its respective costs and only if an action's benefits exceeds its costs will it be advantageous to go through with it.

<u>Systems</u>	<u>Reliability</u>	<u>Expected Total Present Value Cost (Capital Plus Operating)</u>	<u>Benefits for Reliability</u>
S_1	r_1	C_1	B_1
S_2	r_2	C_2	B_2

Figure 5.1 Comparison of Two Systems Expanded at Different Levels of Reliability

<u>Systems</u>	<u>Expected Total Outset Value Cost</u>	<u>Benefits</u>
S_1	C_1	B_1
$S_1 + g_1$	C'_1	B'_1

Figure 5.2 Effect of Adding Another Generator to a System Design

Second, we point out that the benefits accruing to a given action can be thought of as the avoided losses that would have been expected to occur had the action not been taken thus causing the reliability level to be lower. For example, we expect a higher reliability system at r_2 to have benefits larger than those of a system at r_1 ; the benefits consist precisely of the smaller losses that can be expected to accrue to S_2 .

$$B_2 - B_1 = L_1 - L_2 = -(L_2 - L_1)$$

$$\Delta B = -\Delta L$$

Third, we would like to comment on the properties of the measures of benefits or losses. Although all we need for the purposes of making a comparison of two actions is to have a relative measure, we will find it useful to develop a measure which deals with the whole range of possibilities. However, we will qualify how valid our approximations are and over what ranges they are most applicable. These measures would ideally display the following properties:

- 1) As the system becomes perfectly reliable, the loss measure should approach 0 asymptotically.
- 2) As the system becomes increasingly reliable, the benefits of a given increase become smaller, i.e., the marginal benefits of increased reliability fall monotonically as reliability increases.

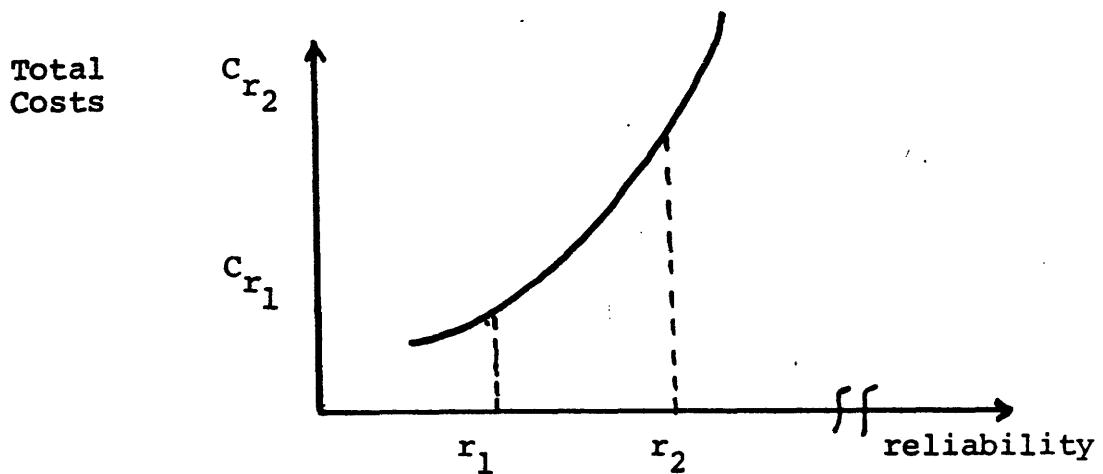
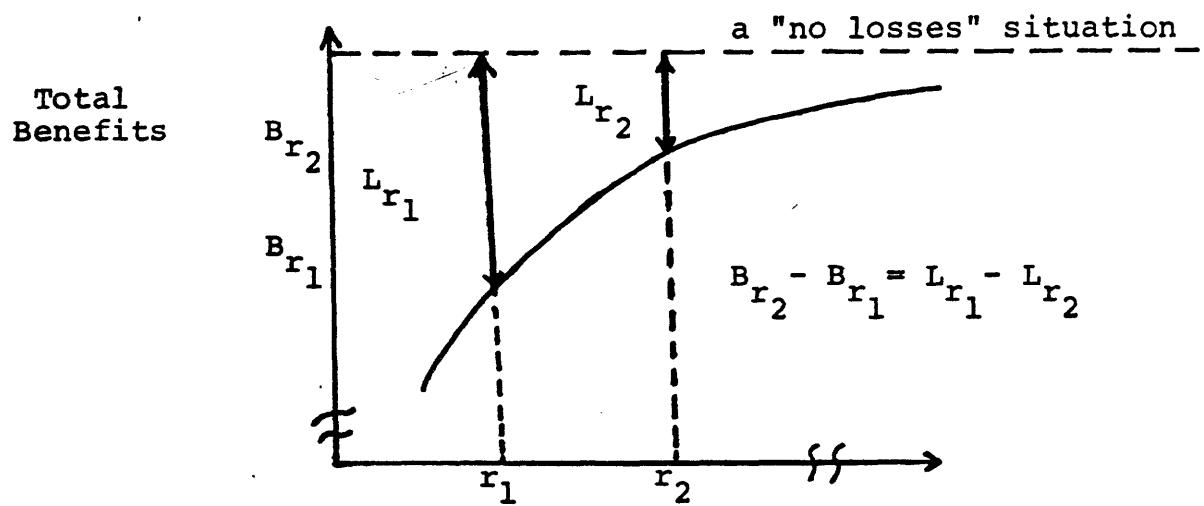


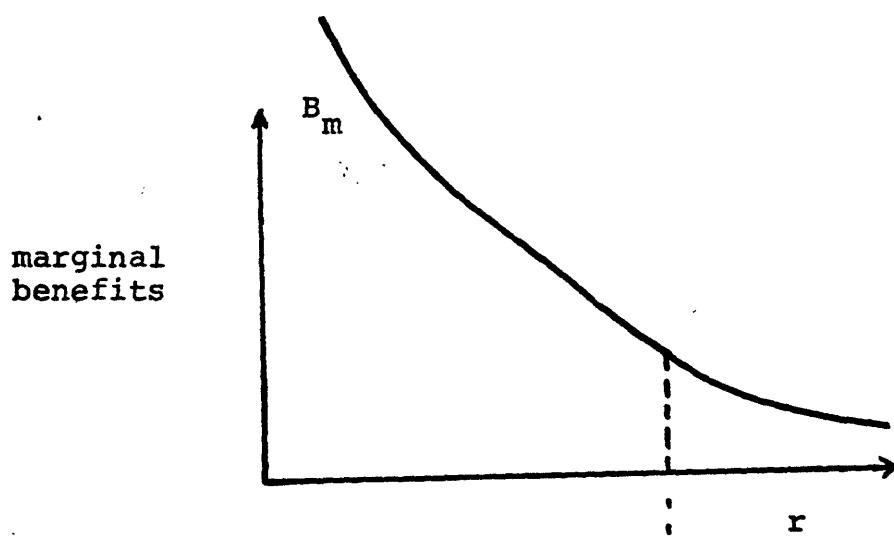
Figure 5.3
Benefits and Costs as Functions of Reliability

After a certain point increases in reliability do not substantially benefit the system performance, and we note that

- 3) marginal costs for added increments of reliability go up as reliability grows; i.e., there is always the possibility of available capacity being smaller than load and as expenditures continue to grow system reliability asymptotically approaches the perfectly reliable situation.

The argument for reliability expenditures has typically been couched in the following terms, "the total benefits of system reliability far outweigh those related to the additional costs." There is no question that total benefits do exceed total costs. The question is whether we have gone beyond the point of optimality and have spent some money that has not returned itself in benefits; or whether these benefits have been improperly perceived and we have not spent enough to reach optimality.

Fourth, the decision on what level of reliability to plan for, and what other societal arrangements to make, should be based on the economics of the situation as well as on questions of equity and overall societal intent. For instance, rural electrification would never have been done had it been based solely on an economic criterion. Another example that shows the need for incorporation of other than strictly



for optimal reliability, r_{opt}

$$\frac{dB}{dr} = + \frac{dC}{dr}, \text{ or}$$

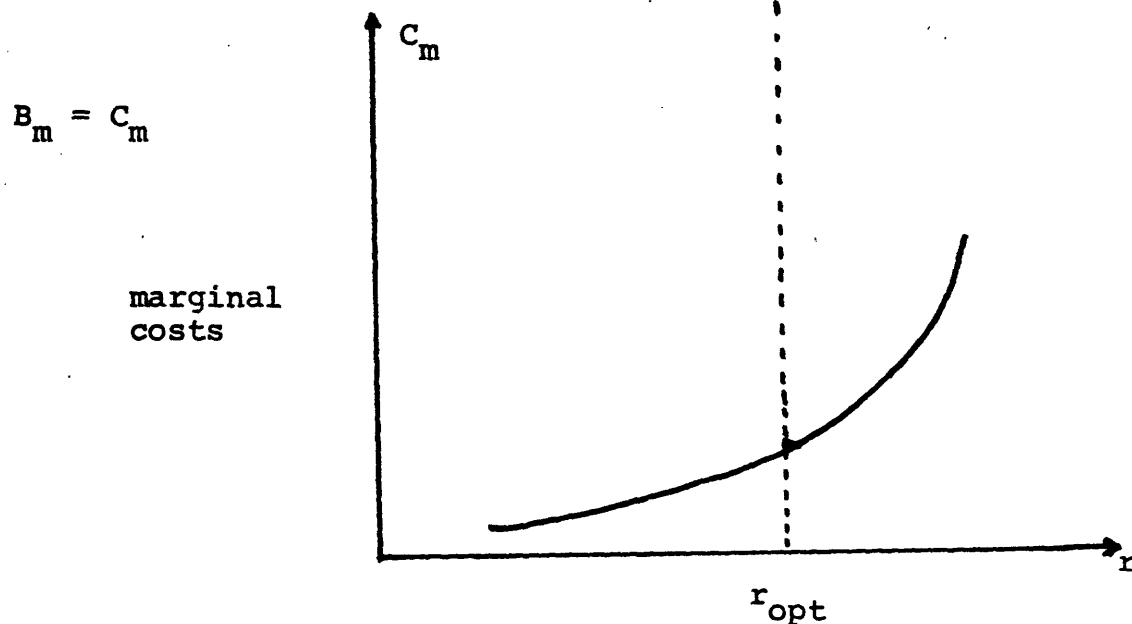


Figure 5.4
Marginal Benefits and Costs

economic criteria in policy formulation involves the mistake that would be made if residential consumers were singled out to solely bear the burden of blackouts because they seem to lose the least in economic terms from an occurrence. However, it is also true that were residential consumers to acquire backup generation for home use they would:

- 1) not be able to make good full time use of them thus not permitting efficient recovery of the fixed capital cost component.
- 2) have to purchase small units thus foregoing economies of scale.
- 3) could not afford to keep skilled staff around to maintain the machines and repair them.
- 4) and in a similar way not be able to defray all the other fixed costs over a large amount of electric production.

An aside here might be a useful illustration. Many run of the river plants which are smaller than 1 M.W. have been paid for for years and yet are being closed down. Why? There are no fuel costs and no capital costs. However, there are maintenance and operator costs which when allocated over the plant output make this power uneconomical.

- 5) Enforcement of environmental regulations would be made increasingly difficult by the growth in the number of emission sources. Because of the reasons above it would not make good sense to push

all residential consumers in the direction of acquiring backup. Instead it might make better sense to require those who could use backup efficiently (although they would lose most in blackout situations without backup) such as industrial consumers to install it and shed their load first.

In payment for this prerogative they might be given a break on their service rates.

Fifth, it is important to note that there are non-monetary as well as monetary benefits to increased reliable electric service. Monetary benefits would include as examples lower amounts of food spoilage due to blackouts and greater amounts of productive non-interrupted work, among others. Non-monetary benefits would include lesser amounts of social inconvenience and lesser social insecurity, as well as a greater flexibility of societal arrangement, than otherwise would be possible. For example, more secure power would allow for certain forms of commercial and other societal activity which would be economically marginal if they were forced to provide it for themselves. In other words, provision of highly reliable service has economic spillovers, it comes "free" as an additional bonus to those marginal types of enterprise. A specific example of this would include many small establishments such as butchershops, ice cream shops, etc., which would suffer serious damages if the power supply were unreliable and yet might not be able to survive if they had to provide their own

secure sources of power.

The Issue of Voltage Reductions

In our discussion of the measures of reliability we discussed the problem that one encounters when trying to define at what margin point does a loss of load event begin. As we have said, load is shed whenever the spinning reserve criterion of the system is violated, i.e., before load reaches available capacity, although no customers would actually be disconnected until the margin situation deteriorated further. Therefore, it should be clear that there is a range of seriousness to loss of load events.

The most effective technique for accomplishing a small amount of load shedding is the reduction of system voltage levels. Before we discuss specific examples of the losses involved in power disconnections we will make an aside to discuss the important issue of voltage reductions. There are two problems: first, there is considerable question regarding the efficacy of voltage reductions for reducing power load on electric systems over prolonged periods of time and second, voltage reductions affect different types of electric uses differently. In pure resistive applications such as incandescent bulbs, the power consumed becomes the square of the fractional reduction, i.e., power = V^2/R and if $V' = .95V_o$ then power' = $(.95)^2 V_o^2 \approx .9 P_o$. This usually means that lamps grow dimmer dryers don't dry and many such functions suffer decline in

quality. In motor uses, usually the motor performs at less efficiency at lower voltages, also more current is drawn and one gets more motor heating and perhaps electrical insulation deterioration; this becomes a real factor in curtailing lifetime in motors whose insulation is old and deteriorated.

Fluorescent lamp lifetime also deteriorates at lower voltages.

It is a fact that the Con Ed system today can effect roughly a 5% load reduction through an 8% voltage reduction. However, over long periods of time it is not clear that voltage reductions are a wise load reduction method. Many electric devices are keyed to performance of a function and will therefore perform the same function at a lower voltage except less efficiently while suffering higher resistive losses through enduring higher currents. The prime example of such machines includes the air conditioner; an air conditioner will work to keep a room under its present thermostat indicated temperature, if the voltage level is reduced the unit will stay on longer while performing its job slower and at a less advantageous and efficient design circumstance. Another example would be a blender or an iron that would take that much longer to accomplish its job and so prove ineffective in reducing total load and perhaps detrimental to the original objective if the job could not be done without expending greater amounts of energy than before over longer time periods of smaller power consumption.

Voltage reductions also impair TV set images and can

cause problems when operating machines that are designed to small voltage deviation tolerances. The problem can become serious when allowance is made for the fact that many buildings wirings already place consumers at significantly lower voltabe levels than nominal. Furthermore as voltage levels drop, constant load uses will demand more current [for these uses (current) X (voltage) is approximately constant] thus causing further voltage drops along highly resistive wiring, and worse yet, excessive heating up of conductors. Under certain adverse circumstances this higher distribution current load due to reduced voltages could lead to reduced distribution network cable lifetimes. It is believed that the excessive current loads carried in some distribution network cables because of a few cable failures was partially responsible, along with extreme hot weather, for starting the chain reactions that brought down several other distribution cables within the same bundles in NYC in the summer of 1972.

In any case, voltage drops can be gotten around of by many consumers. For instance, the smarter and wealthier corporate consumers, who cannot afford large voltage tolerances because of their computer systems for example, may purchase automatic voltage regulators that would automatically boost voltage to nominal levels whenever reduced by the utility. One unexpected result of our analysis is that in the long run voltage reductions themselves would become more inequitable in that they would impact on a specific societal group. We

should summarize our discussion by saying that up to 3% voltage reductions are probably harmless although it would probably not be wise to use them continually as a matter of policy (especially because of distribution cable deterioration).

5.2 Evaluation of Benefits

5.2.1 Benefits are in the Eyes of the Beholder

It is important to note that the benefit valuation process we are describing works in the following way (see figure below):

- 1) A more reliable electric network can be expected to result in fewer loss of load situations of less severe nature than one with lower reliability.
- 2) The valuation of this different expected physical situation depends on who is differentially affected and how by the new more or less reliable electric service. The valuation also depends on each consumer's alternatives and/or his personal tastes and attitudes toward risk.

if reliability different physical these new expectations
goes from r_1 → shortage events → have new values to the
to r_2 can be expected consumers of power

Figure 5.5
Benefit Evaluation Process

The point to be made is that the valuation of the benefits, indeed what can be considered as a benefit, depends on who is at the receiving end of the electric service, the consumer.

5.2.2 Examples of Losses from Lower Reliability of Service

Several specific examples of load shedding losses follow; please keep in mind the points made in our previous discussion about benefits:

- 1) Residential consumer in central city 10 hour blackout:

This fellow suffers a hot stuffy day without air conditioning (a/c) if he has had an a/c; his refrigerator goes off and there is partial food spoilage, also the freezer section might thaw and some further spoilage occurs. He is inconvenienced and may not be able to cook his food or iron, wash or dry his clothes. Also he may find it impossible to watch TV, listen to radio or do some work by table light. If he lives in a multifloor dwelling he may be endangered by the lack of lights in hallways, or staircases and the malfunctioning of elevators. Anyone with electricity powered life support equipment such as an iron lung would be endangered.

- 2) Residential consumer in the "boondocks";

The same things occur as to the fellow in the city, however, this one is used to it happening and has made provision

for the next time by having candles available, food ready when it may happen, etc. The only real large difference is that he knows what to do and is prepared to withstand it at an acceptably low cost while being inconvenienced. His counterpart in the city will be able to adapt similarly in most ways if present unreliable practices continue as indicated in an article on backup power supplies in the December 1971 issue of Popular Mechanics.

- 3) Tall building in central city, white collar workers,
a 5 hour disconnection:

The damages are much more extensive. First, much useful work is lost, although depending on how the residential customer values his leisure time and the inconveniences caused by a power loss, this component of total losses may be balanced out on a per person basis. If the situation lasts long the building may be evacuated and the whole remainder of the day lost for work.¹ If people return to work later, the duration of the situation is a lower limit to the time spent without because there are time delays for evacuating and returning to a building. Of course, the worker efficiency in the time spent on work may increase or decrease depending on whether

¹Studies by the Canadian National Research Council show that the evacuation of a fifty story building takes roughly two hours and eleven minutes. Source: National Observer, October 21, 1972.

they complete nearly the same amount of work they would have if nothing had occurred.

Second, a tall building without electricity is a very dangerous place. In order to evacuate, many stories must be descended and crowd hysteria may set in and some people may be harmed or asphyxiated on the way down. If anyone has a heart condition or some other sensitive condition, he may run the risk of aggravating it. Also, fire hazards become more serious, and water systems lose pressure and do not deliver water to the higher floors causing potentially dangerous situations and great inconveniences.

Third, because of the synergistic nature of city activity the losses may extend to other companies and people who interact with the buildings whose service has been interrupted. It should also be noted that it is very difficult to partially disconnect city networks and localize the blackout because of their highly interconnected nature. Because of these reasons, if service disconnections become much more common, there will be legislation created requiring the installation of backup reserves beyond presently required emergency lighting systems that generally run off batteries.

4) Regular commercial property:

There are several types of commercial properties that would be affected. One example is that of the Aqueduct race track in New York. On a recent blackout day, it had to close down and lose that day's revenues while still having to pay

its employees. The caterers had to throw out their food while still paying their employees. Another example was provided by the many butchers in New York City who complained that they had lost meat through spoilage.

5) Industrial load disconnected for 5 hours:

It may be that the industry has interruptible contracts and has thus planned for such situations; if they have no backup power they may suffer a loss of production during the period and hopefully the costs thus incurred are balanced by the lower rates they pay for their power. If they have their backup power they must expect that these losses will occur often enough to make the investment worthwhile. Whatever the situation, industries usually have the expertise to acquire the best combination of circumstances at the right costs so that this sector would probably be the least negatively affected. As we previously mentioned, they probably even have the resources to protect themselves against a voltage reduction, if the maintaining of nominal voltage levels is of importance. If electric power is important to proper operation, such as in the aluminum industry where molten metals would have to be discarded, work in progress losses rather than employee time loss would force them to install backup power.

6) Public sector disconnection:

Sewage and water supply plants may be dropped off line with consequent unprocessing of sewage and loss of water supply pressure.

7) General public effects:

Serious breaches of public order and safety may occur if traffic signals and other such systems, such as police and fire, do not properly function because of power loss.

Our goal in this section is to develop a clear notion of what could be called the benefits of a more reliable power system and to prepare a catalogue of what they are to the different consumers of power. When faced with the prospect of having to implement a load reduction, the electric company will probably be faced with disconnecting groups of these at a time and will not be able to easily segregate them out according to each situation; however, it will be possible to improve such area selections by thinking of each area disconnection in terms of the sets of activities within each option that are distributed.

5.2.3 Basic Variables that Affect the Valuation of the Benefits of Increased Reliability

The valuation of the benefits of increased electric power reliability depends on two basic issues:

- 1) The new expected loss of load situations; i.e., the new probabilistically described physical situation. These include shortage situation descriptors such as:
 - a) time of day
 - b) time of week

- c) time of year
 - d) the duration
 - e) the frequency of occurrence
 - f) the seriousness of the event, i.e., a 3% voltage reduction or a 50% blackout
 - g) the warning given beforehand.
- 2) The customers affected and their reactions to the loss of load situations. Their reactions depend on:
- a) for each customer his own expectations of the new loss of load situation
 - b) his valuation of these expectations
 1. the cost of his alternatives
 2. the preparations he has made
 3. his attitude toward risk
 4. his capitalization rate (the benefits occur over time)
 5. the inventory of specific tasks that are impaired and how severely
 - c) non-critical
 - d) critical (they may affect his, or overall public safety)
 1. communications (radio, TV, etc.)
 2. transportation (railroads, subways, bridges)
 3. street lights
 4. hospitals

5. water supply
6. sewage treatment plants
7. police services
8. fire services
9. high rise buildings (e.g., fire hazards)
10. life sustaining apparatus (e.g., iron lungs)
11. important record bearing computer systems.

Note that the benefits to each customer depend on his valuation of the new situation regarding uncertain loss of load events. It should be clear that there is no way of knowing what the actual benefits might have been; money is to be expended to improve the expected, but uncertain, future events.

Finally, overall societal valuation depends on the number of individuals in each class. A sum over all different classes affected, weighted by the amount of the effect will roughly give the desired result.

In order to facilitate this task, it will be necessary to classify the consumer sector into a reasonable set of groups each of which would tend to value a given reliability increase in the same fashion. This classification will also be useful in that it will be helpful in isolating which users are most hurt by lower reliability levels of service. If a loss of load situation occurs, a precondition for optimal handling is that those who are least hurt by a loss of load be disconnected first (if they are always to be lower in priority of

disconnection their rates should reflect this); the information so derived will then be useful to help determine priorities for a loss of load management system and to help provide long term guidelines for steering those users, who can least afford poor service and who can make best use of backup supplies, toward acquiring them.

The information could be displayed by constructing a matrix that would list the types of occurrences along its column headings, list the types of customers along the row headings and enter the numbers of people affected and the respective amounts in each entry. See the figure on the next page.

As an example, we will sketch out the differences between rural and central city consumers. Rural and outlying area consumers have always had low distribution system reliability. This fact means that these consumers have accommodated themselves to this fact of life and they have probably purchased lanterns, to prepare for the lack of power when it comes. It also means that they probably have gas or propane stoves and probably depend on oil or other in situ heating systems; if the winters tended to be extremely cold they would almost certainly avoid unreliable electric heat service. When a power blackout comes, they are inconvenienced and placed in some danger through the lack of neighborhood lights facilitating robberies, etc., and decreasing maneuverability. The question naturally is how come the electric company has given

	3% Voltage Reduction	8% Voltage Reduction	0-3 hr Blackout	3-6 hr ... Blackout
Residential consumer in central city				
Residential consumer in country				
Tall commercial property				
Small commercial property				
Industrial				
Public sector consumer				
General public				
⋮				

Figure 5.6 Customer Loss of Load Incident Impact Matrix

them this lower reliability? Simply, society has decided that the cost of upgrading reliability of service to outlying communities is not worth the benefits produced. There has been no formal analysis of this but it is easily seen that the cost of providing service to rural consumers (even at this lower level of reliability) is much higher than in a central city due to the ability to spread the costs to the latter group over much higher sales of power. This is true because lightly loaded lines are not fully used to their limits and because there are economies of scale at higher line voltage levels that can be obtained if many customers are served in a small area at small incremental transmission cost.

If this service reliability were increased, the benefits to this type of small rural consumer would include reduction of inconvenience and possibly some lower out of pocket costs due to avoiding expenditures on lanterns. Another advantage would be the greater flexibility provided to the consumer in the way of making electric appliances real alternatives to oil or gas fueled ones.

It should also be clear that facilities that "need" high reliability electric power would not settle in rural low reliability areas unless it were somehow economic to provide that reliability, e.g., an aluminum plant might settle near the power grid in a rural area so as to minimize the costs involved in transmitting from the grid to itself in order to obtain reliable service.

It might be safe to assume that lower bulk power reliability would affect these customers little since they have adapted as much as possible to the much more significant low reliability distribution network. It is also probably safe to assume that expenditures will not be made to significantly increase the distribution system reliability.

On the other hand, the central city presents quite a different situation. Intricately interconnected societal arrangements in central cities require highly reliable electrical power. The highly interconnected networks in cities reflect this need and as a result distribution system failures are much less frequent than in outlying areas. These areas are presently particularly incapable of successfully coping with power blackouts and there is evidence that it would be difficult to adapt over the long term to lower reliability situations except through procuring reliability through other means such as partial load emergency backup systems that would keep a minimum level of critical facilities running.

There are many reasons for this inability to adapt; they include issues involving the problems in tall buildings all the way to traffic signals in core city areas and the proper functioning of electrified mass transit systems. Elevators,

lights and space conditioning systems may fail in tall buildings leading to extremely uncomfortable and unsafe situations which may include lack of ventilation and air, and mass hysteria as crowds try to evacuate a building without quick exit facilities. If a fire broke out under these circumstances it would be impossible to quickly and safely evacuate the building. In addition water pumps that carry water to the higher floors may stop working thus leaving these floors without water.

At the minimum these buildings would have to be provided with emergency backup systems to operate these few vital functions such as emergency lights, and elevator egress systems. If the business deemed it worthwhile they may go all the way toward providing backup capacity for a couple of days worth. This decision would depend on the value the business placed on continuing operation rather than just providing for safe evacuation. This in turn would depend on how much the business' operation depended on others also functioning; it would be useless to stay in operation if one's customers did not come around.

There is some fear that such requirements would "make building uneconomical and drive the construction industry out of New York City". It should be pointed out that to provide safe evacuation would cost little as a percentage of total building costs, and provision of further backup would depend on the particular building's need to keep functioning throughout

the disturbance. For example, 50kw should be sufficient to keep an elevator running and this involves a capital cost expenditure of roughly \$10,000.

However, there is a significant question of public policy: to what extent should buildings be exposed to loss of load risk? A possible consequence of present policy is that most would be equipped with backup power systems. Would it be wise to require that this great an expenditure be made?

Other problems would also occur in the central city; among them, traffic signals may not function and traffic congestion (slowdowns in any event) will result. In addition, street conditions may become unsafe if the loss of power occurs during nighttime hours (very unlikely) and further damages are possible if the situation occurs during rush hours which are the peak hours of use and thus most vulnerable to shortages.

It is safe to say that there are users and uses of electric power who would require continued high reliability of service and there are others that will not. There are many in between that will react to change in reliability of electric service by modifying their practices; these users will weigh the pros and cons and decide whether or not they should obtain backup service, change their practices or simply endure the new situation. All of these actions will be taken based on some evaluation of the costs and benefits of each considered alternative. There is no a priori way of knowing how people will react to lower reliability electric service, we can only

make estimates based on reasoned arguments. We will attempt to lower and upper bound the value of these benefits, i.e., the benefits are at most worth what it would cost the customer to install reliable backup power (we assume the fuel costs of producing the same number of kwhrs with the backup is higher than the consumer's marginal electric bill cost and so we assume he would continue to purchase electricity). He may go as far as purchasing an interruptible contract from the utility if the discount were large enough to cover his hours spent interrupted while producing higher cost power and if the utility found it economical to provide interruptible service. The benefits will at least be worth the inconvenience and delay costs, the foregone purchases of power and some measure of losses due to improper appliance functioning, etc.

One of the contributions of this approach is that it helps to isolate those types of shortage situations which involve the least social cost and thus it may be helpful in suggesting approaches to designing least social cost blackout management systems. It will be worthwhile to remember that these management systems also cost money to build and operate and this should be a factor in determining the "least cost" system. This approach seems particularly valuable in that at present utilities are designing automatic load shedding systems to drop 25% of load in serial steps in order to protect system integrity against a cascading blackout. There are many technical problems involved in proper design of these systems

but very little attention has been paid to the serious questions of who should be dropped and how and when.

5.2.4 Long Run vs. Short Run; Marginal vs. Substantial Changes in Reliability Levels

We are all acquainted with stories about property losses which occur due to a power failure; commonly we find that business sales are lost, man hours worked drop (representing a loss to the employer or employee), food spoilages occur and many people are severely inconvenienced, in their homes or in their places of work. These losses occur because society has expected a certain level of reliability of service. If it were widely known a priori that the reliability of service, that which people had learned to live with, had suddenly become lower, then consumers of power would arrange their affairs trying to protect themselves so that losses were more acceptable when power failures occurred. This is just an expression of the fact that people tend to expect the service reliability they have received in the past, when the level suddenly changes windfall losses can be expected to occur.

If reliability of service were to drop, over the long run people would adapt to their new situation; for instance, fewer electric heating systems would be installed, more gas stoves would be bought instead of electric ones, people would purchase home lanterns and the sales of backup emergency

generators for stores and larger buildings may take off. It is also possible that new special purpose businesses will be formed, such as purveyors of dry ice to keep refrigerators cold and thus retard spoilage or insurers might go into business to protect small and large businesses against such losses.

Over the short run, people would not be expecting frequent power failures and they would therefore not deem it worthwhile to protect against the "once in a lifetime" prolonged blackout. Therefore, if a series of serious blackouts were to occur, they would be unprepared and thus suffer extensive damages. However, once it became apparent that this was the new state of affairs they would begin to adjust.

An example will help clarify what we are talking about and also expose some of the limitations of this analysis. Consider, for example, what occurred to the Aqueduct Racetrack in Yonkers, N.Y. in July 1972. Power to the racetrack was interrupted and it was claimed that they had sustained large losses; does that one-shot loss warrant installation of backup equipment? Obviously, the response depends on how often such a "one-shot" loss is expected to recur.

A reasonable approach would be to capitalize these expected losses and check whether they were worth the investment in backup equipment. We would assume the expected costs of its operation over the period of interest would be small relative to the capital costs and further that it would not be advisable for the Racetrack to obtain an interruptible contract

thereby exposing itself to having to use their backup equipment more often than desirable.

However, this reasonable approach may suffer from at least one drawback if it is generally applied to all such similar situations. Suppose, for instance, that the subways that would bring people to the Racetrack were disconnected at the time of their own system disconnection. It would then be futile for them to install a backup reserve of any consequence; their customers could not get there and at best a power reserve may only help them save some of the caterers' food from spoilage if sufficient storage facilities were available.

In calculating the damages to a consumer of power we conveniently neglected that just providing power for himself would not alleviate the synergistic societal component of his damages. This would indicate that this approach toward valuation would understate the consumer's damages, i.e., the damages to the Aqueduct Racetrack are greater than just the cost of providing emergency power backup.

In this work we will focus on ranges of reliability targets that will not cause massive changes in society. Later we will assume that all those who stand a lot to lose from a power interruption will protect themselves in some way. Therefore there will be no differential benefits because of these consumers over the range of reliability target levels we consider. In particular, we assume that in any case hospitals and tall buildings will acquire backup, that computer users

will install U.P.S. (uninterrupted power) systems, etc.

5.2.5 Distribution Reliability is a Limit on Customer Reliability and Provides an Opportunity for Observing Customer Reaction to Sharply Lower Reliability Levels

There is much experimental confirmation of what we have just discussed. People in rural and outlying suburban areas have much lower reliability service than those in the inner city. This occurs because distribution reliability per customer is terribly expensive to provide in low density, low power consuming areas; this has led electric companies to provide economically practicable levels of reliable service to these areas. Thus the companies have automatically limited the overall standard of customer service to these sectors; the distribution system reliability is the weakest link in the chain and thus limits the reliability of service perceived by the customer. If we neglect overlapping events which are very unlikely, generation unreliability will be in addition to distribution system unreliability so that the marginal impact will not be large compared to the marginal impact felt when distribution reliability is almost perfect, as in a central city area. Although it is true that bulk power interruptions tend to be of a different duration than local distribution failures, we shall consider this difference unimportant.

Consequently, the rural customer has developed very

different approaches to electricity use than the central city resident in response to his lower reliability of service. Most of these consumers have purchased lanterns for emergency use; they shy away from electric heating and electric stoves and electric water heaters. Other rural consumers who have a high need for electric reliability, such as chicken farmers who depend on electric heating for their chicken hatching rooms, have purchased backup generating equipment and some have obtained interruptible contracts from electricity companies that enable them to recoup some of their investment in the form of lower electric rates.

However, study of these limiting cases regarding system reliability will provide an interesting and invaluable source of information as a laboratory with respect to what people may do if their power service becomes less reliable.

5.2.6 Customer Options: Interruptible Power Contracts and Backup Supplies

Everything else equal, it would be optimal for each consumer to purchase the level of reliability that he requires; however, it is not possible to provide this option since the reliability of the system is available equally, at relatively low cost, to all those within a given area. But in the case of larger consumers, electric companies usually give these consumers a choice between regular and interruptible service.

Interruptibles carry lower rates per kilowatt hour delivered and are attractive to those who can tradeoff these savings against the costs of installing backup generation or against the costs of accepting the losses when they occur. It's important to note that as the electric system becomes less and less reliable the interruptibility option becomes less and less attractive since it will in fact be exercised more often. In any case, the contracts in existence for past expected reliability levels give us some indication of the relative benefits involved and the potential for expanding interruptible coverage although they reflect decisions made in times when different expectations about reliability were held. For this reason they should not be naively interpreted as a sign of what people would do if they expected to receive sharply different levels of reliability. A manager might find one or two possible yearly interruptions acceptable, and two or more probable ones unacceptable.

A utility could use different types of interruptible contracts to provide varying amounts of load relief. For example, a utility could sell a contract which would offer lower rates on the demand portion of the consumer contract which was placed on interruptible status. The utility would only find it worthwhile to offer this option to a certain class of customer because of the additional control gear involved. The customer's reaction to such an option would depend on his particular circumstances. What is the basic problem faced by

the consumer when he tries to decide whether or not to purchase an interruptible power contract? To make this decision he must decide whether or not the savings obtainable from his lower rates would be sufficient to offset the additional loss of load events he would encounter. It follows that in order to make such a decision he must also have some reasonable estimate of how often he would be called upon to shed load.

What are the considerations he must balance? Basically he must ask whether the rate savings offset the additional losses for his taking a one year contract.

If $(D - \bar{EU}_n)r - (D - \bar{EU}_i)i > (\bar{EU}_i - \bar{EU}_n)\ell$ then the customer will switch.

D = normal yearly demand in kwhrs.

r = normal rate per kwhr.

i = reduced interruptible rate per kwhr.

\bar{EU}_n = expected yearly loss of energy under the normal contract.

\bar{EU}_i = expected yearly loss of energy under the interruptible contract.

ℓ = losses per kwhr of loss of energy.

We note that this is a minimal requirement (based on the classical profit maximization approach) for the manager to find the interruptible contract worthwhile and thus incur the additional bother and draw on his time. Perhaps a more realistic approach would require that the manager find this "project" sufficiently attractive (in terms of the other

projects he could conceivably be involved in) to justify his undertaking it, i.e., he would require a certain premium before he undertook it.

Be that as it may, let us investigate what the equation implies for the desirability of interruptible contracts given certain parameter assumptions. Assume that:

$$\overline{EU}_i \approx .04D \text{ (say 10 days in the summer)}$$

$$\overline{EU}_n \approx .01D \text{ (say 2 days in the summer)}$$

$$l \approx 30r$$

what would i , the interruptible contract rate, have to be to make it desirable to switch?

$$(D - \overline{EU}_n)r - (D - \overline{EU}_i)i > (\overline{EU}_i - \overline{EU}_n)l$$

$$(.99D)r - (.96D)i > (.03D)(30r)$$

$$.09Dr > .96Di$$

$$r > \frac{.96}{.09} i \approx 10.7i$$

It is clearly not worthwhile to the utility to bill a group at one tenth the price. At this rate the utility would not cover its marginal generating cost and this has to be a lower limit on i .

If we assume $i \approx r/2$ which is probably the lowest it could be, what would this say about the largest value of l

which would make a consumer find the interruptible contract attractive?

$$.99Dr - .48Dr > .03D\ell$$

$$.51Dr > .03D\ell$$

$$\ell < 17r$$

Therefore if ℓ is greater than $17r$ and this is a reasonably low value for ℓ , a customer will remain with this regular contract.

Interruptible contracts, then, do not seem to be a very good source of load relief if consumers will not acquire backup and have to endure the actual losses of load. Things change if we consider the attractiveness of such a deal to consumers who will acquire backup in any case. There are two broadly defined kinds of consumers in this category: first, those, such as hospitals, who cannot afford (because of possible lawsuits due to criminal neglect) to be without power in any case and so will not want to run the risk of depending on their backup supply and second, those who will be able to afford to rely on their backup supply in order to save money on their normal electric power bills.

If we assume that the cost of the backup supplies are sunk, then the additional cost to the consumer during a power loss is the differential cost he incurs while generating his own more expensive power. If we represent the production cost

of the backup power as PCB per kwhr, the previous equation for desirability of interruptible status becomes

$$(D - \overline{EU}_n)r - (D - \overline{EU}_i)i > (\overline{EU}_i - \overline{EU}_n)PCB$$

and if

$$\overline{EU}_i = .04D$$

$$\overline{EU}_n = .01D$$

$$PCB = 3r$$

then

$$.99Dr - .96Di > (.03D)3r$$

$$r > 1.06i$$

This result means that for these customers who can afford to rely on their backup generation, acquiring an interruptible contract with a 20% rate break may be a reasonable thing to do. If the savings are large enough (say 20% of the annual electric bill) it might pay some consumers, who do not have backup power, to purchase some backup supplies.

While discussing these possible policies we should realize that there is a connection between the break in rates that can be given for interruptible contracts and the cost to the utility of providing the additional generation in house.

Both policies, that of creating interruptible contracts and that of buying additional generation capability, have costs, and that in order to ensure a non-wasteful use of societal resources much thought should be given to the level at which to set interruptible rates. Even if we neglect the effect of distribution reliability on the two broad approaches above, we believe that it will not be possible to give a clear cut answer to this problem because it is not possible to tailor fit the quality of service to each consumer, and therefore the marginal costs of interruptions to differing consumers will vary widely.

Instead of procuring peak load relief through offering lower rates for the right to interrupt supplies during peak load hours, we might consider obtaining the relief by imposing a peak load consumption surcharge. These are obviously symmetric approaches although they seem to have different side effects. We do not know too much about the effects of these policies and we will not deal directly with these problems to any extent in this work.

In the appendix we will treat a related question: if we can reduce the peak load forecast by Δ through peak load surcharges or through granting interruptible contracts, how much of a cost savings can be realized in total system operation costs given that we want to maintain a given reliability level? Clearly, the lower forecasts will now require lower expenditures to meet the same reliability target levels. To

the extent that most of system unreliability comes from peak loads, there will be significant possible savings.

If daily peak load duration curves have been used to measure system reliability a reduction in the forecast peak will have a stronger effect on system measures than if hourly load duration curves have been used.

We do not deal with the effectiveness of a given policy toward achieving a forecast change. This problem is outside the scope of our work. To give the reader a flavor for the complexities in answering these questions consider the effect of imposing a peak load consumption surcharge. Let us assume that there are no problems with regard to letting the consumer know when it is in effect. Clearly the consumer will now be forced to economize. However, it is altogether possible that his elasticity of consumption during peak load hours is very small compared to his elasticity during off peak hours; the result may be that he will continue consuming during peak hours and economize during off peak hours thus aggravating the "Peakingness" of the load duration curve. This is only one of the problems associated with gauging the effect of such policies and it is clear that we need further research, including controlled test experiments, in order to resolve these problems.

5.2.7 Historical Attempts at Valuing the Reliability of Electric Service

Historically, it seems to have been assumed that the

value of the reliability of electric service was much higher than the costs involved in providing it; we do not dispute this contention in the large, however, recent events have indicated that this conclusion has come into question, especially with regard to environmental questions, and it is thus worthwhile to examine precisely what are the benefits of reliable electric service and what would be the effects of a less reliable level of generation reserves, who would be affected, how and how much. It has also become clear that conditions of low excess capacity will continue for some time, that there may be efforts to curtail the historical growth of electric consumption and that new pollution control equipment will require more electric energy; this probably means that past trends are likely to change in unexpected fashion thus causing potentially dangerous situations if society does not properly adapt to the new conditions which are liable to be forecast incorrectly.

Our lack of proper preparation is clearly demonstrated by the almost universal lack of public urban contingency plans if power shortages occur as well as by a lack of efforts to determine what changes are to be made in planning processes given the apparent new equipment delays and equipment unavailability the NERC has reported. The FPC has asked that these plans be prepared but the only ones at present that have had public scrutiny are those of New York state.

As we previously discussed the NY PSC has initiated a

set of hearings to determine the public interest in designing the management of power shortages; in particular, the immediate objective was to determine a set of procedures the single electric utilities as well as the NY Power Pool would follow in the event of a loss of capacity and subsequent curtailment of load, the second objective was to discover and initiate changes in the planning process that would relieve the pressures on the future electric system. In connection with the latter objective, the examiner recommended that no new construction be allowed to connect to the Con Ed system and thus be obligated to provide for themselves; this was later modified by the Public Service Commission Staff's recommendation that no new building commercial establishment be allowed to connect into the power supply. Both of these recommendations were rejected by the Commission and they decided to allow management of load curtailment by rotating the blackouts among primarily residential areas.

In its testimony, the NY PSC was presented with two estimates of the losses involved in loss of load events; one of these was developed by the New York City Economic Development Administration (NYCEDA), the other by the chief economist, Olaf Hausgaard, of the NY PSC. Both of these estimates were approaches to quantifying the losses suffered if power were disconnected without great prior warning, with primary concern given to commercial areas. The NY EDA based its calculations on an assumed network disconnection in mid-Manhattan. It then

calculated the number of workers affected, multiplied by their salaries and assumed this worker output would be lost if there was a disconnection. This figure amounts to \$2.5 million per hour of disconnection but it disregards the fact that if this were to occur often enough, management would build their own backup plant and thus probably avoid the major part of the losses.

The second calculation provided by Hausgaard of the PSC, is more sophisticated but it too predicates its estimates primarily on the basis of manhours of work lost. This calculation yields a figure of \$2.17 million per hour of losses if 5 percent of the State's work force were impeded from fulfilling their jobs because of power reductions. Again, this figure has some merit in that it tries to compute the marginal impact of one more loss of load event at a given overall reliability level; however, this is not the calculation we are after to determine the cost of a step change in the reliability of service level, i.e., the question is how much will these blackouts cost in the long run after management and private parties adapt to the new conditions. The above computation has further difficulties in that the hours on the job may be less or more efficiently spent depending on the nature of the job, and the evacuation and return time may be much larger than the actual shortage duration. Also, the above calculation does not account for non-monetary losses such as danger, anxiety, etc., and Hausgaard acknowledges and emphasizes their importance.

Finally Shipley, Patton and Denison have published a paper "Power Reliability Cost vs. Worth" which was presented at the 1972 IEEE Winter Power Meeting. In the paper they suggest that the cost of interruptions is linearly related to the GNP, i.e.,

$$C = k(1-A)$$

where A is the availability, C is the cost and k is the GNP. The assumption is made that if availability is perfect the cost is 0 and if availability is 0, the cost is the total GNP. They then capitalize these costs (we call them losses) for comparison with the costs to provide a kw of capacity. Their optimal availability is given by that at which the derivative of the total costs (= system investment costs + capitalized interruption costs) is equal to 0. The paper is extremely illuminating in that it shows the issue is at least alive among engineers and secondly, because it suggests, controversially, that present availability levels are too high from an economic standpoint.

5.3 Construction of a Benefit Function

Up to this point we have discussed and analyzed the general problems that any benefit function would have to deal with. In Chapter 3 we showed that LOEP, an energy shortage related measure, was probably the most reasonable reliability

measure to use in spite of its shortcomings. In this section we will present techniques for relating the losses suffered because of loss of load incidents to the expected amount of energy sales not made because of the unavailability of generation capability.

We will speak of the benefits of one system over those of another as being equal to the difference in losses suffered by the latter over the first, i.e., $\Delta B = - \Delta L$. Also we will have to discount the differential flows over time in order to arrive at a net present value of benefits versus costs. The need for this comes about because the money that will be invested now in generating capability could be used to produce greater amounts of money in future periods.

What are the kinds of losses that accrue to poor system reliability? Let us examine this problem sector by sector and see what we are talking about. The residential sector does not really suffer financial loss most of the time, i.e., what this group suffers is inconvenience, discomforture, delays in performing certain chores. If the LOL is long enough they might also suffer food spoilage. This sector loses "leisure time" enjoyment if the enjoyment is intrinsically connected with the availability of electric power gadgets.

The commercial sector may suffer lost sales over a period of time if some of the sales that were not made during a blackout never materialize. There is some argument over the validity of this point. It often comes up in discussion of

state Sunday blue laws which curtail shop openings on Sundays. (It should be clear, though, that Sunday openings could allow better capital equipment use thus making the economy more efficient.) If sales are just delayed, but not lost, it would follow that this sector would not lose anything. If, on the other hand, delays result in customers switching their purchases to firms which are not blacked out, there will be no net loss but a real impact on the blacked out establishments. Also, commercial outlets may suffer inventory spoilage which in some cases may be rather significant (ice cream, meat, flowers, etc.).

The industrial sector might also suffer these inventory losses (either raw material or finished goods inventories) but they will all have to incur wage costs which will not be returned to them in the form of produced goods. In other words, a factory will have to pay its employees regardless of their production and in fact will have to pay them overtime to make up production targets.

5.3.1 The Issue of Acquisition of Backup Capability

Each consumer clearly values his power at least as highly as the cost that he would be willing to pay for it had it been available. It is also clear that many consumers will go so far as to install backup generation reserve to obtain the power, especially when the cost of power is a small

component of total costs and is indispensable for successful completion of a given task. It follows that, at the highest, the value of power to a consumer will be bounded by his cost of producing the power himself, i.e., the cost of producing it with backup generation supply.

retail cost of the power \leq value of benefits of x kwhrs \leq cost of least expensive option to produce the power on his own

Every consumer has the option of installing backup power, however, many will not do it because:

- 1) the backup expenditure must be made now, and the benefits will appear in the future and they may have a high discount rate.
- 2) they decide they can afford the uncertain losses they estimate they will incur relative to the certain expenditure they must make to avoid them. They are risk seekers or rather, not sufficiently risk averse, and would rather take their chances.

There may be other less obvious societal effects due to the reduction of levels of reliability of electric power. Consumers may start shifting their preferences away from electric appliances such as stoves, dryers and electric heaters to gas or oil operated ones; this is what has occurred to people who live in areas of low service reliability. There will also be a tendency to cut down on the feasibility of social arrangements; for instance, many marginal businesses which need

reliable service, say for refrigeration purposes, but who can not afford to buy backup will probably be unable to locate in such areas over the long term. This fact essentially means that in the past these have been spillover benefits given to these small consumers in urban areas which have been received without their paying extra for them since the marginal cost of providing the same reliability service to these consumers as well as to other larger ones was nearly zero.

What are the costs involved in acquiring backup and how would a consumer analyze his purchasing decision? Costs for backup diesel generators are in the range of \$200/kw. If we assume that reliability is sufficiently good so that a small amount of backup energy is generated throughout the equipment's lifetime, we can assume that the generating cost differential between backup power costs and the regular electric rates is not a significant factor in the decision. The dominant costs then consist of the capital costs plus maintenance charges that must be undertaken if the equipment is acquired. It is simple to see that for a small consumer, fixed overhead costs such as maintenance charges make the investment quite expensive in terms of a per mwhr charge. Also, it is easy to see that small load factor users will find the capital costs a rather expensive luxury. If we assume a fixed charge rate of 15% and maintenance costs of \$20/year we obtain a total yearly cost of \$80 for a 2kw generator (@ \$200/kw)

$$[\text{yearly cost} = (.15)(200)(\# \text{kw}) + 20]$$

which would be sufficient to keep essential lights, and televisions running (an air conditioner requires somewhat higher than 1.5kw to operate) in an average four member household. At present, this household, on the average, consumes about \$200/yr of electricity. It follows then that it would cost an average $80/200 \approx 40\%$ premium on electricity rates to obtain limited and imperfect protection.

On the other hand, a high load factor consumer with large consumption might wish to protect 50% of his capacity with backup. In this case we can assume fixed maintenance costs are small compared to the capital costs and the premium per kwhr consumption is

$$\frac{\text{yearly capital costs of backup}}{\text{total yearly consumption}} = \frac{(.15)(200)(.5)}{(8760)(20 \text{mils/kwhr})(\text{load/factor})}$$

and if the load factor is about 50%

$$\text{the premium} \approx \frac{15}{87.6} \approx 17\%$$

We learn from this that small consumers will probably not obtain backup but large ones with high load factors will, if interruptions are costly with respect to the cost of power itself.

If even a small number of loss of load events is expected, there will be many users who will find this cost

unacceptable and they will then have to provide for backup supply for the portion of their load which they consider critical. These customers include hospitals, police department communications networks, traffic signals, tall building critical egress facilities (emergency lighting and elevators at least), radio networks, electrified mass transit and petrochemical and aluminum manufacturing plants among others.

There are also many other users of electric power who need high quality electric power, i.e., power which is very close to the nominal 60 cycle-120 volt power on a single cycle basis. Computer installations need this kind of high quality electric service and it is clear that it would be senseless for the electric company to deliver this high quality power throughout its system, most users do not need it. The same argument can, and probably will, be made regarding interruptibility of electric service.

In order to protect themselves, users such as computer installations have purchased "uninterruptible power systems (UPS)". Each user has to decide how much of their load they want covered and for how long. For instance, they may keep the computer power on a UPS and the air conditioner that cools the computer on a backup diesel generator. There are two basic UPS arrangements: one is a solid state device and the other a mechanical device. The former consists of a rectifier-battery-inverter assembly which in effect buffers the computer from irregularities in the incoming power. The mechanical

device consists of a motor generator which drives a flywheel that stores energy, when the incoming power deviates from the standard, the flywheel releases some of its stored inertial energy.

5.3.2 Total Losses Will Depend on How the Unserved Energy is Apportioned

A consequence of basing our analysis of system losses on unserved energy is that the overall $B(LOEP)$ for a given system will be a function of how the $\overline{EU} = LOEP(E_{tot})$ is apportioned among the different types of consumers. For instance, it would be possible to always curtail service to only one group of consumers rather than spreading the loss around evenly. This is the crux of the loss of load management problem: who and how often should one group be called upon to curtail load? Since this involves taking from some and giving to others it has become a political problem as well as a technical one. New York State has held public hearings to determine satisfactory loss of load procedures and they have concluded that residential areas in Queens be the first to be actually disconnected for New York City curtailments.

What are the problems in determining the optimum conceptual treatment of this problem and what consequences does this have on our $B(LOEP)$ function? From an overall societal point of view it would be best to pick the loss of load

management scheme that minimized total losses. On first blush one would have to conclude that this would imply that residential consumers, those who have no loss of output to show for their disconnection apart from their anxiety and inconvenience, should be curtailed first. Things aren't so simple though, it may be that residential consumers' inconvenience should be valued at the equivalent value of their time had they used the time for working. It is a well known problem that GNP calculations include the value of maids' work while not counting the same value if it is provided by housewives who do not get paid for their work. To assume residential losses would be very small might incur the same bias.

One reason that many residences will not install backup is that it would be significantly more expensive for them to acquire it than it would be for commercial or industrial consumers. It would be reasonable to ask whether the latter two should shoulder more of this burden since it would be easier for them to purchase the backup. The upshot of this may be that although the effects of a curtailment are worse on industry than on the residential sector, it is much cheaper for industry to provide for backup than it is for the residential sector. It therefore seems that the decision on who to curtail and who should acquire backup should not depend only on whose losses are smaller but also on who can best afford to protect themselves.

The process of obtaining an overall societal optimum

will allow for later reapportioning the additional benefits that have been received; our first objective should be to make the pie larger (in the Pareto sense) so there is more to share. Many societal problems, though, get hung up on the redistribution question. For example, if the industrial and commercial sectors are going to have to bear a greater responsibility for absorbing curtailments, they should be rewarded. How much of a rate break should companies be offered for them to absorb the responsibility? It is extremely hard to get agreement to this question, even on strictly cost of service grounds, and because of this fact the overall approach is usually dropped, and each participant focuses on keeping their present pattern of service undisturbed.

5.3.3 By Assumption Losses are Linearly Related to Unserved Energy

Our basic assumption for valuation purposes is that customer losses are linearly related to the amounts of energy unserved because of generation capability insufficiency. Therefore what we seek is a valuation method based on:

- 1) how much energy will not be delivered to each customer because of insufficiency.
- 2) how does each customer value that energy which is not served.

Our linearity assumption is really only good over ranges

of energy curtailment which are small relative to consumer needs; i.e., if losses are great enough that it pays the consumer to obtain backup generation the loss function saturates at this point. This does not seem to be a problem over the ranges of LOEP we will investigate.

How much energy will each customer be docked depends on \overline{EU} and on the form in which \overline{EU} is distributed among the customers in the system; for instance, the bulk of the curtailment may be assigned to residential consumers. The system LOEP can be found from the E_{tot} forecast and the \overline{EU} for the period and vice versa.

What is the proportionality constant in the linear relationship between energy unserved and customer losses? We make the assumption that the percent of energy unserved causes a like percent of total wages to be lost and therefore the proportionality constant is the total amount of wages in the investigated period. For example, if we consider a factory which pays a yearly wage bill of one million dollars and consumes two million kwhrs, the loss per kwhr interrupted is \$.50/kwhr lost, and the losses for that year in terms of LOEP are:

- 1) Losses for the first year $L(LOEP) = LOEP \cdot 10^6 \$$.
- 2) Losses (\overline{EU}) = \$.50/kwhr unserved for year 1.

This is a computation of the losses to this factory if the LOEP is at this level. However, we know that because of load shedding priorities, different parties will receive

different quality of service. The question is how will the energy unserved, $\overline{E_U} = LOEP \cdot E_{tot}$, be distributed. At worst it will all fall upon the sector that suffers the most from loss of load events and at best on those who suffer the least, i.e., the proportionality constant must lie somewhere in between these values

$$k_L < k < k_U$$

We previously discussed that $L(LOEP)$ should saturate as the system approached $LOEP=0$; for the purposes of investigation of LOEP levels between 10^{-2} and 10^{-5} we will assume that L is a linear function of the LOEP level.

There are other problems with our approximation. On one hand, this function underestimates attendant losses because the effects of a loss of load event linger beyond the duration of the LOL event; for instance, people working in a building which has to be evacuated because of a loss of load will not be able to immediately resume their work. On the other hand, this function overestimates losses because it assumes that consumers will keep their average output per kWhr constant throughout time whereas they are bound to become more efficient when they are reconnected thus reducing the losses calculated when using the linearity assumption. Losses are also overestimated because loss of load events do not cause actual customer disconnection until system load has gone a few percent beyond

available capacity because of possible load relief tactics available to the system operator.

Note that we are relating losses to $\bar{E}U$, and that a problem arises if we carelessly relate losses to system LOEP or if we do not take care when computing the loss factor per kwhr interrupted. Most production takes place over the daytime hours; thus, it would be incorrect to associate all wages to all kwhrs produced. For the sake of convenience, and for the purposes of estimating an approximate relevant factor, we will assume that wages are related to that energy production which occurs during daytime hours, and this is how we will compute the proportionality constants for the $L(LOEP)$ and $L(\bar{E}U)$ functions.

Also note that the use of either an hourly load duration curve or of a twelve hour load duration curve (the hours between 7:30 a.m. and 7:30 p.m.) will yield approximately equal $\bar{E}U$ measures but probably twice as great LOEP measure in the latter than in the former. This is natural because the former only covers half the hours. Since most production occurs in the twelve hours of peak daily consumption, it is the percent of energy sales unserved for this LDC that determines the percent of wages lost. Since the losses occur over time, we use a discounting technique to compute their present values.

5.3.4 The L(LOEP) Expression

The losses in period i are equal to the percent of energy sales unserved in period i multiplied by the total wages in period i .

$$L_i \text{ (LOEP}_i\text{)} = \text{LOEP}_i \text{ (} w_i \text{)}$$

Then for a system expanded over thirty years at a given LOEP

$$L(\text{LOEP}) = \sum_{i=1}^{30} (\text{LOEP}) w_i (1/1+d)^i$$

If wages are rising at a yearly rate g , the total present value of system losses due to unserved energy can be expressed as

$$L(\text{LOEP}) = (\text{LOEP}) w_1 \sum_{i=1}^{20} \left(\frac{1+g}{1+d}\right)^i \quad (\text{eqn. 5.3.4.a})$$

we use 20 years because the N.Y. Power Pool Study covers a 20 year period.

Note that if $\text{LOEP}_2 = a \text{ LOEP}_1$, and system 2 is more reliable than system 1, i.e., $a < 1$, then $L_2 = a L_1$ and $L_2 < L_1$ as we expect.

Also note that instead of using wages as our proportionality constant, we could use some other indicator that we feel is correct. We can also turn the question around (as we will do in our optimality analysis section in Chapter 6) and

ask how large this proportionality constant should be before present reliability levels are justified.

Example

The Bureau of Labor Statistics in the U.S. Department of Labor publishes yearly statistics, by state and by industry division, on numbers of employees on non-agricultural payrolls and on average hours and earnings of production or non-supervisory workers on manufacturing and other payrolls.¹ It is possible to use these figures to derive estimates, imperfect as they may be, of wages paid in a given year in a given electric service area. Some states develop estimates of their own figures and we will make use of some data assembled by the New York State Department of Commerce in order to exemplify our approach to the problem. Note that if reasonable wage figures are not available it may be sufficient for the purposes of analysis to estimate how much of a multiple should be assigned to area kwhr sales to derive an alternative wage figure. For example, if sales are \$1 million, by assigning a multiple of 50 we may use a pseudo wage proportionality constant for the L(LOEP) function of \$50 million.

Our data from the N.Y. State Department of Commerce covers the year 1969 and indicates that wages and salaries paid in N.Y. State in 1969 were close to 55×10^9 \$. If we set $g=4\%$

¹"Employment and Earnings" states and areas 1939-1971. U.S. Department of Labor, Bureau of Labor Statistics.

and $d = 8\%$ in equation 5.3.4.a (the wage growth rate and the discount rate respectively) we find that in terms of 1981 dollars

$$\begin{aligned} L(LOEP) &= (1.04)^{12} \sum_{i=1}^{20} (.963)^i (55 \times 10^9) LOEP \\ &= (1.6) (13.59) (55 \times 10^9) LOEP \\ &\approx 1200 \times 10^9 LOEP \end{aligned}$$

5.3.5 The $L(\bar{EU})$ Expression

This formulation of the loss expression is especially useful for computing the reduction in losses that a marginal addition to a given system will yield. If we add an extra gas turbine unit in year 1 of the plan, it will reduce the energy unserved in each subsequent year of the plan by some amount, $\Delta \bar{EU}_i$. If L is the losses per mwhr unserved then

$$L(\Delta \bar{EU}_i) = \Delta \bar{EU}_i (L)$$

$$L(\text{marginal plan change}) = \sum_{i=1}^{30} (\Delta \bar{EU}_i) (L) \left(\frac{1}{1+d}\right)^i$$

where 30 years is the assumed life of the marginal additional plant. By using probabilistic simulation techniques it is simple to compute the $\Delta \bar{EU}_i$ for each period i . In the optimality analysis section we turn the question around and ask what

valuation constant L would make a given investment worthwhile?

Example

The critical question here is to associate to a given wage figure, the respective number of mwhrs. We have 1970 electric power consumption figures for the New York Power Pool¹ and they indicate that total commercial and industrial sales of power were 48.9×10^6 mwhrs out of a total of 89.2×10^6 mwhrs (corresponding to an uncorrected peak load of 17,500 mw which implies an approximate 57% LDC load factor). Since wages for 1969 were 55×10^9 \$, we can apply a 4% growth rate to find that 1970 wages were approximately 57.2×10^9 \$.

$$L = \frac{57.2 \times 10^9}{48.9 \times 10^6} \approx 1,170 \text{ \$/mwhr}$$

which is roughly a factor of 40 greater than the price of the energy at \$30/mwhr.

In our computer PROSIM runs we have calculated in detail the ΔEU_i for each period caused by adding one 100 mw gas turbine to two systems, one highly reliable and the other much less so. It is of interest to see what difference this makes to both systems in terms of energy unserved. Clearly, we should expect that the gas turbine will have more of an effect on reducing unserved energy in the less reliable system than

¹Source: New York State Department of Public Service Power Division. 1971 Report: Electric System Planning in New York State, p. 28.

on the highly reliable one and this is what we find.

In Chapter 4 we presented the results derived from attempting to expand electric power systems at given LOEP levels, we used 6 different expansions to check the differences that would be involved. We repeated the runs for the least and most reliable of these expansions, S_1 and S_6 , but this time with one additional 100 mw gas turbine; S_1 consisted of 320 units over the 30 year period and S_6 of 444 units over the same period.

The results due to the presence of the additional gas turbine are displayed in the figure below and the corresponding computer output is included in Appendix 5.2

We can see that the additional gas turbine produces a much greater reduction of unserved energy in system 1 than in system 6 which reflects proper behavior of a benefit function: as the system becomes more reliable, progressive expenditures produce progressively smaller returns.

The numbers derived are in terms of present value gigawatt hours. Their dollar value depends on how we value each gigawatt hour. We could use the factor of \$1170/mwhr we just derived and since we save 121 gwhrs., this would imply a total value of 140 million for this energy which only cost another 10 million, the price of the gas turbine, to provide. If the energy is valued at its revenue cost, then it is worth approximately 3.5 million. We will return to these arguments in Chapter 6.2.1.

Period	<u>S₁ Plus a 100 mw Gas Turbine</u>			<u>S₆ Plus a 100 mw Gas Turbine</u>				
	\overline{EU}_i	\overline{EU}'_i	$\Delta\overline{EU}_i$	Discount Factor at 8%	PV	\overline{EU}'_i	$\Delta\overline{EU}_i$	PV
1	494	443	51		51	9	7.7	1.3
2	620	574	46	.78	36	22.4	20	2.4
3	601	564	37	.61	22.5	30.5	28	2.5
4	281	266	15	.48	7.2	20.2	18.8	1.4
5	49	47	2	.38	.8	2.3	2.2	.1
6	3339	3261	78	.30	23.4	92.3	88.9	3.4
					141		141	6.5

(We multiply these numbers by .86 which represents a 2-1/2 year delay so losses are assumed to occur in the middle of an interval)

Total Present Value
for Adding Gas
Turbine:
to System 1 - 121 pvgwhrs
to System 6 - 5.6 pvgwhrs

FIGURE 5.7 The Effect on Unserved Energy of Adding Gas Turbine
to Two Different Expansion Plans.
(Unserved energy is in gigawatt hours)

CHAPTER 6

Determination of Desirable Levels of Generation Reliability

Our goal in this study has been to create techniques for long range electric power system reliability planning that would provide sensible bases for determining what target levels should be. Some have argued that generation reliability has in the past been too high (these people point out that most service interruptions have been caused by transmission and distribution, rather than generation, system problems) while others have taken the opposite position. We believe that the only resolution can come from an examination of the costs and benefits of different proposed alternatives. We now proceed to examine the question of what LOEP level should be used for system expansions and later we examine the question of whether or not the addition of a particular plant to a given system plan is warranted. We will find that the latter approach is more flexible and yields some interesting conditions on system LOLP measures even though it is based on energy considerations.

6.1 Determination of LOEP* for Expansion Studies

In Chapters 4 and 5 we developed expressions for $C(LOEP)$ and $L(LOEP)$. Variations in the parameters in these expressions will have

an effect on LOEP* and we will examine the relevant sensitivities.

Each system will have its characteristic C(LOEP) and L(LOEP) functions and therefore our results should be understood as the outputs of a procedure for determining desirable reliability target levels rather than as numbers with an independent significance of their own which are applicable for all electric power systems.

In section 4.5 we remarked that the New York Power Pool had derived the following C(LOLP) function:

$$C(LOLP) = (34.5 - .75 \log (260LOLP)) \times 10^9$$

over $4 \times 10^{-3} > LOLP > 4 \times 10^{-5}$

We also remarked that if LOLP were linearly related to LOEP over this range (as we found to be the case in our Prosim runs) that it would be possible to express

$$\frac{d}{dLOEP} (C(LOEP)) = \frac{(-2.3)(.75)10^9}{LOEP}$$

We also know from Chapter 5, that for the same service area --- New York State--- the loss function could be expressed as:

$$L(LOEP) = (1200 \times 10^9) LOEP$$

To find LOEP* we must find the value of LOEP at which the total costs plus losses are at a minimum. This is given by:

$$\frac{d(C + L)}{dLOEP} = 0$$

$$\dot{L} = -\dot{C} \quad \text{at } LOEP = LOEP^*$$

$$\frac{(2.3)(.75) 10^9}{LOEP^*} = 1200 \times 10^9$$

$$LOEP^* = 1.45 \times 10^{-3}$$

Now, we showed in Chapter 3 that $LOEP < LOLP$ and we know that $LOLP$ levels today are set at 3.8×10^{-4} (one day in 10 years). This means that present reliability levels are too high if we believe the postulated loss model. We can get a better fix on how much lower they should be if we assume that $LOLP \approx 20 LOEP$, which means that at present, $LOEP \approx 1.9 \times 10^{-5}$ which is roughly one hundred times "more reliable" than the $LOEP^*$ we derived. If we believe our cost ($LOEP$) function, this means we are spending \$1.5 billion too much in present value dollars over the 20 year plan costs of \$34.5 billion.

Note that if we assume that

1) $C(LOEP)$ is logarithmic of the form $C = C_0 - C_1 \log(LOEP)$ and that 2) $L(LOEP)$ is linear, of the form $L(LOEP) = K_1(LOEP)$ then we can say that in general:

$$3) LOEP^* = \frac{2.3 C_1}{K_1}$$

This is a very useful result. It provides a handy rule of thumb.

If K_1 is doubled LOEP* is halved; and if C_1 is doubled, LOEP* also doubles. This is simply a reflection of the fact that the more expensive losses become, the higher the optimal level of reliability becomes. Similarly, if it becomes more expensive to provide extra generation than previously planned the lower the optimal level of reliability will become. Because of recent cost inflation in this sector, many utilities are considering doing just this.

The reason why this formulation for LOEP* is useful is that it now becomes simple to argue over the parameters C_1 and K_1 to see what effect they have on LOEP* without going through much computation.

Although there can be debate over large ranges of K_1 , C_1 can be estimated (as we discussed in Chapter 4) to be within a few percent of total thirty year plan expenditures.

6.2 The Costs and Benefits of Marginal Additions to Specific Expansion

Plans

6.2.1 Use of the Probabilistic Simulation Technique

We have derived a technique which would allow for evaluation of the effects on reliability of making changes to an arbitrary expansion plan. Note that this is a much more flexible and less difficult procedure than the one we outlined for expanding a system at a given LOEP level. We should not lose sight of the fact that this procedure also makes it

possible to test for optimality by ascertaining whether changes in the expansion programme yield benefits in excess of their costs.

We can use the procedure to test whether the plan is not optimal with respect to reliability questions. It does not tell us how to find the optimal expansion since it is a necessary, but not sufficient, condition for optimality, that changes in the plan not yield benefits in excess of their cost. We will describe the procedure we are advocating by detailing the results of a specific example. We will return to the abstract problem and starting from energy considerations derive an interesting condition on LOLP levels of optimal systems.

The specific example we chose to run involved testing the reduction of unserved energy throughout a thirty year plan brought about because of the purchase of an additional 100 mw. gas turbine in the first period of the plan. We also wanted to test the assumption that the addition would have an increasing effect on the reduction of unserved energy as we considered its addition to increasingly unreliable systems.

To these ends, we tested two hypothetical situations:

- 1) we added a 100 mw gas turbine in period 1 to the system 1 run in figure 43 of section 4.4.
- 2) we added the same turbine in period 1 to the system 6 run in that same figure.

System 1 is the more unreliable of the two and we expect that the addition will have greater impact in this system than in system 6. The

results on expected gigawatt hours unserved are displayed in fig. 5.7

An additional expenditure of \$10 million for the 100 mw gas turbine (at \$100/kw) brings about a reduction in present value of energy unserved (present value gwhrs) of:

- 1) 121 pvgwhrs in the unreliable system
- 2) 5.6 pvgwhrs in the reliable system

There are two ways to proceed from here. First, we can ask from an abstract viewpoint what value can be placed on these unserved gwhrs and then decide by comparing it to the additional \$10 million cost whether it was worthwhile. Remember that we showed in section 4.4 that the additional capital costs were the only relevant ones for gas turbines. There are some offsets due to possible fuel cost savings in the case of non peaking additions to plant. Second, we can turn the question around and ask what minimum price we must ascribe to the energy in order to make the addition worthwhile. If the maximum ascribable price is below this, we can say that the addition is probably a poor investment.

Using the first approach, we refer to the L(EU) formulation in section 4. and use the figure \$1170/mwhr. This means that:

- 1) system 1 benefits amount to $(1.17 \times 10^6)(121) \approx 142 \times 10^6$
- 2) system 2 benefits amount to $(1.17 \times 10^6)(5.6) \approx 6.6 \times 10^6$

Assuming that unserved energy is valued at \$1170/mwhr (nearly fourty times its sales price), our results state that the benefits of adding

a turbine to system 1 are worth \$142 million which are much greater than the turbine's cost, and the benefits of adding it to system 6 are only worth \$6.6 million which are below the turbine's cost.

Turning to the second approach outlined above, we ask how much is the minimum price we must be willing to pay before the investment is worthwhile. Let us assume that the price of power is \$30/mwhr and ask the above question in terms of how much of a multiple must this expected unserved energy command before it is advantageous to purchase the additional gas turbine. The general equation to be solved is:

$$\Delta C = -\Delta L$$

$$10^7 = (30)10^3 M \text{ (pvgwhrs of } \overline{\Delta EU})$$

For system 1 the equation becomes:

$$M \geq \frac{10^4}{(30)(121)} \approx 2.76$$

For system 6 the equation becomes:

$$M \geq \frac{10^4}{(30)(5.6)} \approx 59.5$$

The results say that in order to have the additional gas turbine pay off in system 1, we must be prepared to pay a multiple of 2.76 over its price of \$30/mwhr. If we assume our $L(\overline{EU})$ of section 4.5 to be correct, we see that \$1170/mwhr clearly implies the expenditure is worthwhile. In fact, we only must be willing to pay over \$83/mwhr to make

it a worthwhile investment.

Our results say that the system 2 multiple is 59.5. If we now assume that \$1170/mwhr is a correct cutoff figure, we see that the nearly \$1800 price implied by our analysis of system 6 indicates that the addition to the reliable system 6, is not worthwhile. Another thing these results say is that even under generous assumptions regarding the worth of the unserved energy, it does not seem to make sense to increase the reliability of system 6. In fact, one should make the deduction that it is possible to trim the reliability of system 6, i.e., reduce a few of the units in system 6, and obtain cost savings in excess of the additional unserved energy losses. The same style of argument can be made for the general problem and yields a few interesting marginal conditions.

6.2.2 Theoretical Analysis

1. Additional Costs

The additional cost of obtaining an extra generator consists mainly of the expenditures involved in purchasing and maintaining it. The total system fuel costs involved in producing the previously expected to be served energy have been reduced by virtue of the fact that the additional unit may displace previously operated more expensive capacity. If the additional unit is at the top of the loading order, system costs will remain unchanged since the added generator will not displace any other more expensive plants.

Roughly speaking, we can then say that:

$$\Delta C \approx CO, \text{ the capital cost for the addition plant.}$$

2. Reduction of System Losses

Additional plant makes it possible to reduce the amount of energy unserved in each of the years of the plan. Our goal now is to estimate how much energy would be affected in each year, appropriately value it by some multiple of its cost, discount future benefits to the present, and then sum them and compare the net total value of loss reduction to the additional cost involved.

We will make a brief aside to compare the effects on unserved energy of either 1. raising capacity on a committed future plant by 1 mw or 2. adding another mw of separate additional plant. The reduction of unserved energy is equal to the expected amount of additional generation.

$$1) \Delta \bar{E}_n = p_n \int_{IC}^{IC+1} F_{n-1}(x) = \Delta_1 \bar{EU}$$

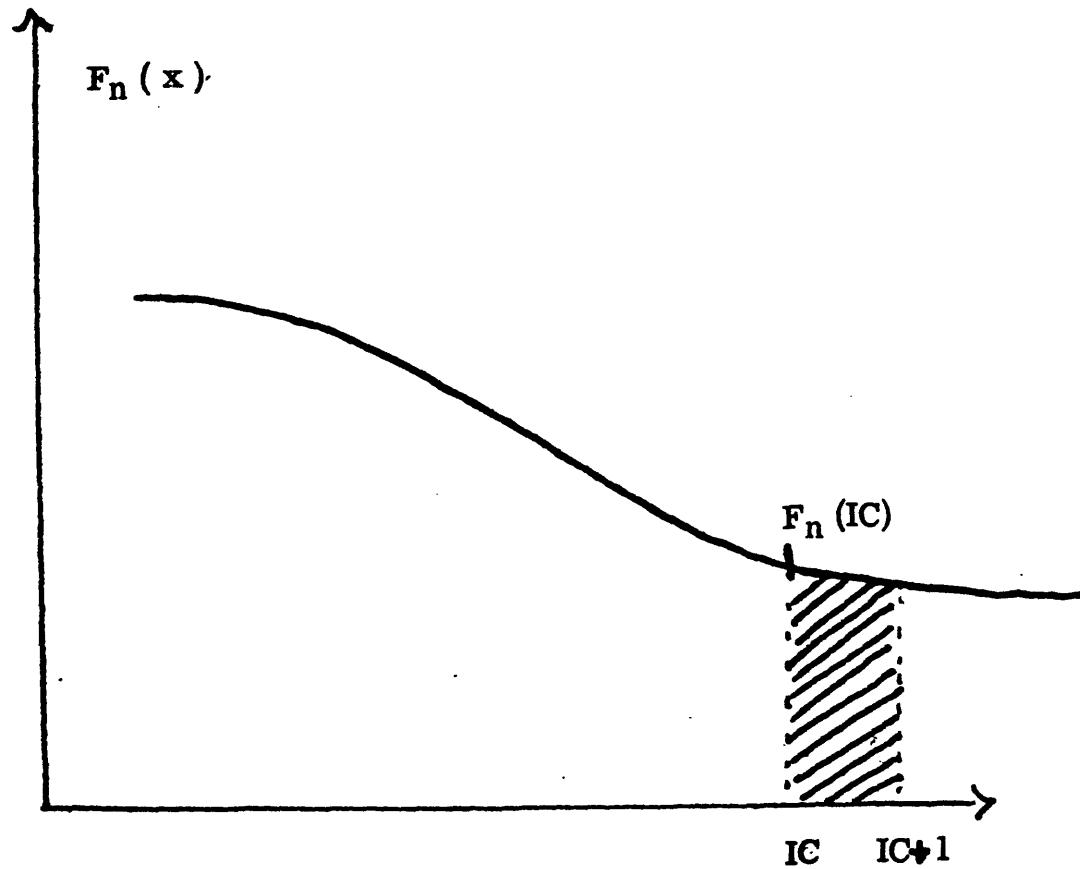
$$2) \bar{E}_{n+1} = p_{n+1} \int_{IC}^{IC+1} F_n(x) = \Delta_2 \bar{EU}$$

We can see that, at the margin, the reduction in system energy unserved is roughly

$$1) \Delta \bar{EU}_n \approx p_n F_{n-1}(IC)$$

$$2) \Delta \bar{EU}_{n+1} \approx p_{n+1} F_n(IC)$$

But we know that $F_n(x) > F_{n-1}(x)$ and so we see that we always get more of a change in unserved energy by adding a mw of separate capacity than adding another mw to a plant already committed, provided they are equally reliable capacity. Note also that we can interpret the system LOLP, $F_n(IC)$, as proportional to the marginal rate of reduction in system unserved energy.



$$\overline{E}_{n+1} = P_{n+1} \int_{IC}^{IC + C_{n+1}} F_n(x)$$

$$\frac{d}{dc_{n+1}} (\overline{E}_{n+1}) = F_n(IC + C_{n+1})$$

Figure 6.1 Marginal Changes in EU

The rate of change of unserved energy is $P_{n+1} F_n$ (IC) $\approx P_{n+1}$ LOLP.

This is true for every period in the plan that is being simulated, and therefore if LOLP (k) is the system LOLP in period k :

$$\Delta \bar{E}U(k) \approx LOLP(k) P_{n+1} T(k), \quad T(k) \text{ is the time represented in the LDC in period } k$$

and although the $F_n(x)$ for the periods k may change over time (it gets wider as the system grows and therefore $\Delta \bar{E}U(k)$ becomes less of a percentage of total system energy), the only thing that concerns us is the absolute magnitude of F_n (IC) over the periods (k) in the horizon.

In order to translate these reductions in unserved energy into reduced system losses we must evaluate them with a valuation constant L in terms of \$/mwhr of interruption.

$$\Delta L(k) = \Delta \bar{E}U_k \cdot T \cdot L \left[\frac{1}{1+d} \right]^k$$

The net change in present value losses, ΔL_{TOT} is given by

$$\Delta L_{TOT} = \sum_{k=1}^{30} L(k) \left[\frac{1}{1+d} \right]^k$$

but we note that $\Delta L(k)$ is independent of k in the case of a system expanded at a given LOLP target level and that

$$\sum_{K=1}^{30} \left[\frac{1}{1+d} \right]^k \approx 1/d \quad \text{for reasonable values of } d.$$

$$\Delta L_{TOT} \approx \Delta \bar{E}U_k \cdot T \cdot L \cdot 1/d$$

and at the margin, for an extra mw of plant,

$$\Delta L_{TOT} = LOLP P_{n+1} TL (1/d)$$

Let us compare this number to the ΔC number and see what it tells us about the value of LOLP when $\Delta C \approx \Delta B$, which is the point of optimal LOLP (LOLP*), for the given assumptions of T, L, d, and co.

$$\Delta L_{TOT} = LOLP (P_{n+1}) TL (1/d)$$

$$\Delta co = co$$

$$\text{Assume } d \approx 8\%$$

$$co \approx \$100/\text{kw}, \text{ or } \$100,000/\text{mw}$$

$$T = 260 \times 12 = 3120 \text{ hours/year}$$

$$L = \$30m/\text{mwhr} \quad m \text{ is a multiple of the sales cost for that unserved energy.}$$

$$P_{n+1} \approx .85$$

$$\text{In order that } \Delta L_{TOT} = \Delta co$$

$$\begin{aligned} LOLP^* &= \frac{co \cdot d}{P_{n+1} TV} \\ &= \frac{(100,000)(.08)}{(.85)(3120)(30)m} \end{aligned}$$

$$LOLP^* \approx .1/m$$

If the utility was a classical profit maximizer, it would decide to set the LOLP expansion level at $LOLP^* = .1$, or a LOLP criterion of 26 days per year or 260 times larger than the 1 day in 10 year criterion. Alternatively, the utility may value unserved mwhrs at rates a customer might use if he were not served; for example, let us check the implications

of L = \$1170. This would mean that

$$\text{LOLP}^* \approx .1/40 = 2.5 \times 10^{-3}$$

Since present LOLP values are 3.8×10^{-4} , this would mean that even at the above assumed loss valuation constant it would still be advantageous to accept target reliability criteria of approximately one day per year which are 10 times higher LOLP values than used at present. If we believe the New York Power Pool cost equation, these possible savings represent \$750 million savings in present value dollars over their twenty year plan without unreasonable increase in their losses suffered because of unserved energy.

Another way to look at this is to ask what implicit multiple would justify present LOLP levels used of 1/2600. From the equation it should be clear that $m = 260$ would do the job. This means that the customer must be willing to pay 260 times the normal price he pays for electric energy before the utility could find it advantageous to install one more megawatt of gas turbine capacity if the system had been expanded at the one day in ten year criterion.

The multiple is very large because we expect that the gas turbine will have to generate very small amount of energy in systems at levels of LOLP that have been used in the past. Our results are surprising and, if taken at face value, they suggest that reliability target levels

have been inordinately high in the past; even a multiple of 100 is too small to justify the marginal mw investment. We should note that this conclusion only gets worse the larger the size of the marginal addition becomes because additional megawatts past the first produce decreasing amounts of energy.

6.3 Analysis of the Results

We should qualify our results and make clear what are the implications of what we have shown. First and foremost, there is a great difference between the reliability level we try to operate at and the one we actually get. To be more specific, our treatment of uncertainty has not explicitly allowed for unexpected delays in completion dates. Our treatment has made no explicit allowance for load forecasting errors

or the occurrence of inordinately hot summers. Therefore, what our results really tell us is that if we could count on the assumptions we make to formulate our models, we could then aim to operate at much lower levels of reliability. Specifically, this means that if we had accurate (or conservative) representations of:

- 1) capacity on line (when it becomes available and what its forced outage rate will be)

and of

- 2) future expected system loads

Then we could say that presently targeted for reliability levels are much too high, maybe by an order of magnitude, i.e., LOLP could be 1 day/year or maybe higher (remember that because we recommend the use of 12 hour LDC's for calculations instead of DPLDC's, present techniques for calculation imply that we are operating at even higher levels of reliability than the one day in 10 year criterion. To repeat, if by using DPLDC's we acquire capacity to yield one day in ten year criteria, if we recalculate the reliabilities by using 12 hour LDC's we find that they are much more reliable yet).

We are left with a couple of unanswered questions, though. The methods we have presented deal with the capacity outage component of randomness. By setting conservative load forecasts and capacity availability schedules, as well as conservative forced outage rate assumptions, we can make the models accommodate these uncertainties as well.

Beyond this, we also have some protection because the actual point of customer disconnection comes for situations worse than when nominal load = capacity available as the LOLP calculation assumes.

Perhaps the way to deal with capacity delays is to calculate system reliabilities under best and worst conditions, see the sensitivity to these assumptions and plan for a reasonable level. Uncertainties in load forecasting can be treated in at least one of the following three ways. It would be possible to specify a probabilistic forecast model for each hourly load. We have to assume that we know the parameters that describe the data generating process. This is the approach used by Westinghouse and their parameters are historically determined. But this approach implicitly assumes that the parameters, and the process that generates the forecasts, can be known with more exactitude than the precision which can be obtained by assuming a forecast which is then inserted into the LOLP calculation. We do not have any reason to believe that one can forecast these parameters any better than one can forecast load directly, therefore this "treatment" does not solve the problem.

It would also be possible to specify for each hour in the period, a set of possible load levels each with different probabilities. It would then be easy to calculate for each level L_i , its associated $LOLP_i$ and set the LOLP for that hour, the expected fraction of the hour spent with $C < L$, equal to

$$LOLP = \sum_i P_i (LOLP_i)$$

The sum over the LOLP for all hours in the period would give the expected number of hours in the period spent with $C < L$.

Finally, it would also be possible to focus on two reasonable extremes for the system LDC and calculate the measures under both extremes. The same could be done with respect to the capacity model and some reasonable combination of these should be the input to the reliability planning models. At present it seems that the one day in 10 year criterion is used to plan over reliably knowing that because of delays, actual operating reliability will be much lower. Our point is that these two problems should be handled explicitly rather than buried together into the one day in ten year criterion. In this fashion it is possible to be explicit about the costliness of plant delays, something which is impossible to do with present techniques.

If we believe our assumptions though, the message of our calculation should be clear: in order to receive reliable service during a few peaking hours in the summer or winter, consumers pay a rate per kWhr slightly greater than two orders of magnitude for that energy than they pay on the average for their energy. However, because most energy consumption is off peak, the difference in average costs to a given consumer is likely to be only a few percent higher for the reliable versus the less reliable case.

This combination of circumstances leads to honest

disagreement between reasonable men. Although marginal costs for peak load service are very high, the total difference in yearly bills is not likely to be very large. Are consumers risk averse and therefore willing to pay the differential in order to avoid lower service reliability at peak consumption hours; or are they successful microeconomists who can be persuaded (through peak load surcharges, or through fervent load reduction requests or otherwise) to stop consuming the few rather expensive peakload kwhrs? A general answer to this question cannot be given, it depends on particular area needs and preferences.

In order to illustrate the approach we suggest should be taken with respect to treatment of errors in the load forecast and of delays in installation of equipment, we have performed two further PROSIM runs whose output is summarized in Figure 6.2. We should warn that the results are exaggerated because our example system is a small one and because we have used large perturbations. As usual, we remind the reader that the analysis is included so that our sensitivity analysis techniques be understood and not because of the specific results because each system will have its own peculiarities.

Expansion Strategy 1

<u>Period</u>	<u>Normal</u>		<u>2% higher load</u>		<u>delays</u>	
	<u>LOLP</u>	<u>LOEP</u>	<u>LOLP</u>	<u>LOEP</u>	<u>LOLP</u>	<u>LOEP</u>
1	3.8×10^{-4}	1.6×10^{-5}	7.4×10^{-4}	3.3×10^{-5}	1.1×10^{-3}	5.3×10^{-5}
2	6.6×10^{-4}	2.8×10^{-5}	1.3×10^{-3}	5.6×10^{-5}	2.2×10^{-3}	1×10^{-4}
3	7.1×10^{-4}	2.6×10^{-5}	1.5×10^{-3}	5.8×10^{-5}	2.1×10^{-3}	8.3×10^{-5}
4	3.9×10^{-4}	1.2×10^{-5}	1×10^{-3}	3.2×10^{-5}	1×10^{-3}	3.2×10^{-5}
5	3.9×10^{-5}	1×10^{-6}	3.7×10^{-4}	1×10^{-5}	3×10^{-4}	8×10^{-6}
6	9×10^{-4}	2.5×10^{-5}	2.6×10^{-3}	7.5×10^{-5}	1.8×10^{-3}	5×10^{-5}

Expansion Strategy 6

<u>Period</u>	<u>Normal</u>		<u>2% higher load</u>		<u>delays</u>	
	<u>LOLP</u>	<u>LOEP</u>	<u>LOLP</u>	<u>LOEP</u>	<u>LOLP</u>	<u>LOEP</u>
1	1.5×10^{-2}	9×10^{-4}	2.2×10^{-2}	1.4×10^{-3}	3×10^{-2}	2.1×10^{-3}
2	1.3×10^{-2}	8×10^{-4}	2.0×10^{-2}	1.2×10^{-3}	5×10^{-2}	2×10^{-3}
3	1×10^{-2}	5×10^{-4}	1.7×10^{-2}	9×10^{-4}	2.2×10^{-2}	1.2×10^{-3}
4	4×10^{-3}	1.6×10^{-4}	1.1×10^{-2}	4.5×10^{-4}	1.1×10^{-2}	4.7×10^{-4}
5	7×10^{-4}	2×10^{-5}	2.1×10^{-3}	6.7×10^{-5}	1.8×10^{-3}	5.8×10^{-5}
6	2×10^{-2}	9×10^{-4}	3.6×10^{-2}	1.7×10^{-3}	3×10^{-2}	1.4×10^{-3}

Figure 6.2 Sensitivity of Expansion Strategies 1 and 6 to

- a) actual loads being 2% higher than forecast.
- b) delays in installation of one base loaded unit in each period.

6.4 What is the Magnitude of the Cost Differentials Involved?

The answer to this question is of great import to consumers as well as to utility companies, especially in the light of their rapidly rising capital needs. Utility companies' demand for capital has been rising faster than ever before egged on by:

- 1) rapidly rising demand forecasts ranging from 7 to 9.5% per annum;
 - 2) the rapid capital cost inflation in this sector in the past five years over which generating equipment prices have tripled;
- and 3) the continual move toward nuclear generating facilities and hence, a more capital intensive form of generation.

Both sectors will view the possible cost saving from different viewpoints. The difference in total costs to consumers will be a small percent of total costs because there are so many other components of service costs, although in absolute terms the sum will be huge. For example, if we return to the New York Power Pool example, the 20 year present value cost difference between systems expanded at .01 days/year versus 1 day/year reliability is \$1.5 billion out of \$36 billion for total system costs; \$1.5 billion is a lot of money but is only 4.1% of total system costs. By extrapolation to the U.S. as a whole, this might represent \$30 billion present worth dollars over the next twenty years.

For the utility however, this argument takes on a different light. Profits are a fraction of total costs and thus a small savings in total costs is leveraged in its effect on gross profits and still further on net profits. A utility company would probably find it advantageous to investigate these savings if several of the disincentives to their doing so could be resolved or faced up to. One problem is that utility companies are exposed to a great deal of political pressure if loss of load incidents occur. A second problem is the fact that it is not clear that present forms of rate base regulation encourage any action on the reliability issue. The problem is complicated because of the fact that loss of load incidents are rather visible whereas cost savings are invisible; in other words, utility companies can tell the public they have saved money but this assurance does not have the same force as the loss of load incidents that occur.

A rough analysis will give the reader a better idea of the nature of the possible savings involved. Let us consider the alternatives of expanding a system at a 30% reserve margin versus expanding it at an 18% margin. These margins are today generally considered to be the respectable upper and lower limits of discussion. Let us assume that system investment costs are roughly proportional to their reserve margins so that if system 1 at 18% costs $I(k)$ in year k, system 2 will cost $(1.30/1.18) I(k)$ in year k which is simply $1.10 I(k)$. We will also assume that fuel costs stay roughly the same under either alternative.

The total cost differential between these two expansion strategies is then 10% of the capital expenditures associated with generation and related transmission investments. Depending on how large a percentage these are of total capital expenditures (distribution is a large component which is not differentially affected), this variation could have a reasonably strong impact on gross profits since the industry is so capital intensive and usually requires \$4 in capital assets to each \$1 in revenues (this is expected to grow to six to one by 1980. Source: Standard and Poor's Electric Utility Industry Survey July 13, 1972.) On a yearly basis, capital expenditures may be 9% of the extant capital base, and we are debating over a 10% swing in this 9%, or a .9% swing in total assets; this results in a 3.6% swing in total revenues and if we believe future forecasts it may represent a 5.4% swing in total revenues. These swings could have substantial effects on company profits.

6.5 Conclusions and Suggestions for Further Research

There are several conclusions and suggestions for further study that can be made after completing this work; we list them here for convenience although they have been suggested in the text:

- 1) For the purposes of evaluating electric power system generation supply reliability, it seems to make most sense to speak of a measure which is in some sense related to the expected fraction of energy sales curtailed --- such as LOEP--- rather than speaking

of a measure---such as LOLP--- which is related to the fraction of time spent in a shortage situation independent of how serious (in terms of load) they may be when they occur.

2) Many load models may be used in conjunction with the calculation of the LOEP measures: among them load duration curves of the daily peak, twelve hour peak usage and hourly variety.

Of the three, the DPLDC would yield the least reliable measures and the hourly LDC would yield the most reliable set of measures.

Since most production, and the greatest probabilities of load curtailment, are encountered for the twelve hours of peak daily use, we recommend that this type of load duration curve be used for planning purposes. It is the percentage of energy lost during these hours that determines the percentage of output lost.

3) The measures that are presently in use are well suited for treating the component of randomness associated with availability of generation capacity. The measures are not well suited to dealing with the vagaries of facility completion; it is not clear that we would even know how to describe these events in a probabilistic model.

In the text we discussed the various problems that the measures had with explicit treatment of load forecasting uncertainties. The gist of our comments is that the uncertainties must be reflected in the LDC which is used for computations; because once this LDC is developed, LOLP and LOEP computations in essence assume

that it is a "deterministic" LDC and proceed to calculate: 1) the expected fraction of time in the period represented by the LDC that the system will have $C \leq L$ and 2) the expected fraction of energy sales in the period that the system will not be able to make.

Because of the above reasons, several runs should be made to ensure that the assumptions regarding unit completion dates and regarding load forecasts are set at reasonable levels.

4) If proper attention has been given to the previous factors, then it seems that target levels for system operation reliability, in terms of LOLP, should be lower than the present one day in 10 year criterion which has been computed on the basis of daily peak load data. As we discussed, if the computation were based on a more realistic LDC model, the computations of actual LOLP levels would show that we have had reliability in excess of that level.

It would also seem that guideline mixes would change as a function of desired reliability: the greater the target level, the greater the peaking component.

5) Rapid inflation in the capital costs of generation, and a slowing down in the economies of scale and in the efficiencies of operation of newer equipment, are starting to force a rethinking of the reliability question in utility company management. Greater effort should be developed toward ensuring that institutional problems, such as rate of return regulation, should not present a barrier to such

economizing by utility companies.

6) One clear conclusion of our analysis is that peak load users contribute inordinately to system expense. It is precisely this user who should be paying the 260 multiple on his peak load energy consumption. This suggests that aggressive, and yet effective, peak load reduction techniques be devised. This is a topic involving elasticities of power consumption and consequently outside the present scope of this work. These results suggest that much creative use might be made of the interruptible power contract perhaps limiting the maximum number of mwhrs a customer might be exposed to lose under utility company orders.

7) One possible extension of this work involves the use of this loss framework for determining optimal system maintenance strategies. If one could construct a loss function which varied over different times of the year, i. e., L (in \$/kwhr interrupted) were more expensive in the summer because of the obvious impossibility of doing office work and its associated discomfiture, and lower in the fall, it would then be possible to design maintenance strategies which minimized yearly losses. One possible result of this approach would be to change the present efforts at equalizing LOLP levels in the summer and winter when they would be useful, and higher in the fall and spring.

8) In our work, we have assumed that LOL events started at $C = L$. This is not quite accurate as we have discussed, and load reduction schemes are brought into effect at smaller levels of C while actual customer disconnections do not occur until C is a few percent above the nominally demanded load which has been reduced through voltage reductions, etc. Now, this suggests that loss of load penalties should be assessed from the beginning of the violation of the spinning reserve criteria and not from the point $L = C$ as is implicitly done now. However, the first few mwhrs of load reduction do not cause anyone significant pain and therefore what we would ideally need is a function $V(x)$ which would start at 0 for the first few mwhrs of disconnection, and rise till it reached the value L at the point where actual customers were disconnected. If this were done, system reliabilities could be yet smaller than the ones we derived. See figure 6.3.

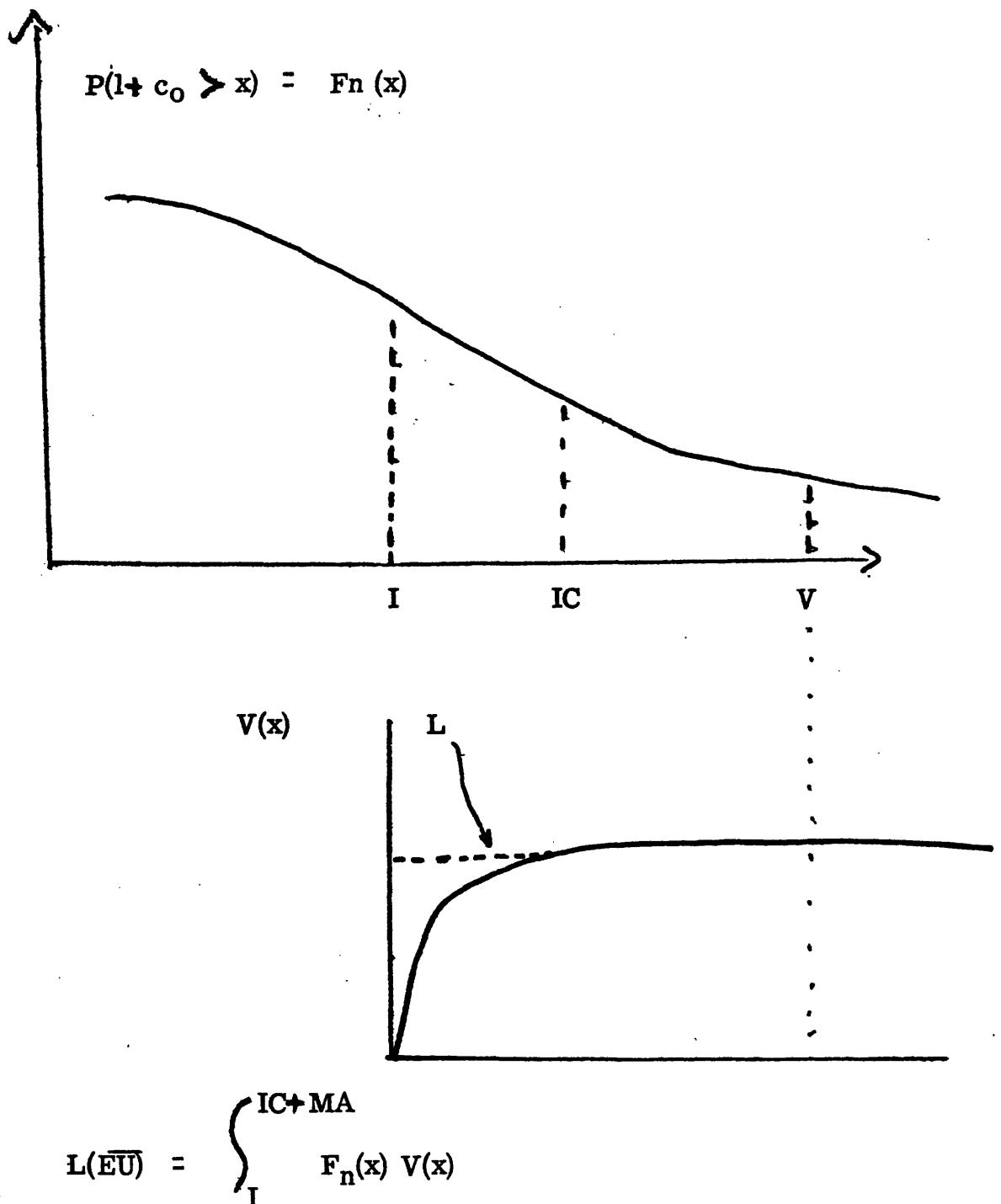


Figure 6.3 A possible formulation for $L(\overline{EU})$

9) We have already discussed in the text what has been done by owners of computer installations to ensure high service quality; these people have acquired "ups" systems at additional cost to themselves. In their case, it is clear that it would not be economically feasible for the utility to provide such high quality service. The same is not true of all other users who may be hurt by reduction of reliability. There may be some users who might be inordinately hurt by additional blackouts and these cases should be isolated and something should be done to either protect them against the worst possible losses or to ensure that they will acquire backup generation for a minimum level of critical facilities (such as elevators.)

But if these customers who could suffer great losses are accommodated for, through mandatory backup supplies or through special arrangements in the loss of load management strategy pursued by the utility, then our linear loss assumption would tend to overestimate the amount of damages caused by a loss of load and this would only buttress our conclusions regarding system reliability levels. The ~~onus~~ is on the companies to develop reasonable loss of load management schemes that will cause minimal disruption. This effort could be associated with the work that several utilities are doing in trying to find a 25% of system load that could be tripped automatically in order to prevent runaway system breakdowns.

10) Our results have been based on the analysis of generation supplies to a given pool; they have not explicitly included interpool arrangements. Our results are not affected by this; if on the basis of interpool agreements it is possible to provide for our desired reliability levels, then this only says that intrapool reliability can be made much smaller yet. However, we should point out that interpool models for reliability evaluation are still not well understood (especially on the basis of how they handle correlated loads) and this is one of the reasons why we centered our analysis on the pool level. We should also emphasize that there is a great need that further investigation be made regarding the assumption of statistical independence among system generators.

11). Although we have focused on the reliability of generation supply and concluded that it is presently too high, it would seem that more work could be done to ensure that transmission and distribution systems deliver that reliability to the customer. Over the past several years, as verified by Federal Power Commission Reports under order No. 331 , it has been these two sectors that have caused most of the loss of load incidents.

12) Prosim calculations have an interesting advantage over typical LOLP calculations. The output of a PROSIM calculation includes the entire system-margin cumulative distribution function. This makes it possible to see the sensitivity of the reliability measure

to errors in the forecast and to changes of definition of the point at which serious load shedding begins. Of course, PROSIM also computes the expected energies produced by the system's peaking units.

We should also mention an important computational result we found. In our runs we found that if we used a computation step size that was much larger than the size of the largest common factor of the nameplate capacities available to the system, the computing errors introduced by this problem could easily swamp the small numbers, such as system LOEP and LOLP, that we were trying to compute. One way to protect against this would be to go toward a step size of 1 mw, however, this would needlessly expand computation time. A reasonable compromise would be to segregate the units into those very small ones and the rest. Then one could round off the capacities in the latter group to facilitate obtaining a large "largest common factor" which would be used as the step size for further computation. After these were finalized, it would be possible to return to the smaller units and use a smaller step size for them. To summarize, it would seem more logical to obtain the computationally correct answers to a rounded off incorrect problem than to obtain wrong results to the accurate problem.

BIBLIOGRAPHY

AIEE Subcommittee on Application of Probability Methods,
"Application of Probability Methods to Generating Capacity
Problems," AIEE Transactions, Part III (Power Apparatus and
Systems) Vol. 79, 1960, pp. 1165-82.

Anderson, D., "Models for Determining Least Cost Investments
in Electricity Supply," Bell Journal of Economics and Management Science, Spring, 1972.

Baleriaux, H.; Jamouille, E.; and Gueretechin, Fr. Linard de.
"Establishment of a Mathematical Model Simulating Operation
of Thermal Electricity-Generating Units Combined With Pumped
Storage Plants." Translated by A. Marty. Extract from Revue E (Edition S.R.R.E.), Vol. 5, No. 7 (1967), 1-24.

Berrie, T. W., "The Economics of System Planning in Bulk Electricity Supply," Turvey, Ralph W., Public Enterprise.

Billington, Roy, Power System Reliability Evaluation, Gordon and Breach, Science Publishers, N.Y., 1970.

Boone, C. C., "The Financial Impact of Outages," paper delivered at 39th Annual Executive Conference, October, 1968.

Booth, R. R., "The ABC of Probabilistic Simulation," Feb., 1971.

Day, J. T.; Shortley, P. B. and Skooglund, J. W., "Expected Value of Generation Deficit: A Supplemental Measure of Power System Reliability," IEEE P.A.S. Transactions, Sept.-Oct. 1972, pp. 2213-23.

Deaton, Paul F., "Utility System Integration and Optimization Models for Nuclear Power Management," M.I.T. Nuclear Engineering Department Doctoral Dissertation, May, 1973.

Edison Electric Institute, Electric Generating Equipment Forced Outage and Availability Data.

"Employment and Earnings," States and Areas 1939-71. Bulletin 1370-9, U.S. Department of Labor, Bureau of Labor Statistics.

Farrar, D. L. and Woodruff, F., "A Model for the Determination of Optimal Electric Generating System Expansion Patterns," M.I.T. Energy Laboratory Publication #10, February, 1973.

Federal Power Commission, Docket No. R-405, Order No. 445,
"Statement of Policy on Power Supply Interruptions."

Federal Power Commission 1970 National Power Survey, Vol. 1.

Federal Power Commission, "Quarterly Reports on Electric Power Interruptions."

Garver, L. L., "Effective Load Carrying Capability of Generating Units," IEEE Transactions, PAS-85, No. 8, August 1966, pp. 910-19.

Hall, J. D.; Ringlee, R. J. and Wood, A. J., "Frequency and Duration Methods for Power System Reliability Calculations: Part I - Generation System Model," IEEE Transactions, PAS-87, No. 9, September 1968, pp. 1787-96.

Mabuce, Estel M., Wilks, Richard L. and Boxerman, Stuart B., "Generation Reserve Requirements - Sensitivity to Variations in System Parameters," IEEE Summer Power Meeting Paper No. C72481-0.

National Electric Reliability Council, "Impact of a 12 month delay of new nuclear and fossil-fired steam generating units on the adequacy of electric power in the United States," February 1972.

Nelson, J. R., Marginal Cost Pricing in Practice, Prentice Hall, Englewood Cliffs, N.J., 1964.

New England Planning Committee, "Interconnected New England Generation Study Generation Task Force Report No. 4," May 1971.

New York City Environmental Protection Administration, Emergency Control Board, "Regulations Pertaining to the Air Pollution Warning System," June 9, 1971.

Northeast Power Coordinating Council, "First Report," September 1968.

Patton, A. D. and Holditch, D. W., "Reliability of Generation Supply," IEEE Transactions, PAS-87, No. 9, September 1968, pp. 1797-1803.

Ringlee, R. J. and Wood, A. J., "Frequency and Duration Methods for Power System Reliability Calculations - Part II - Demand Model and Capacity Reserve Model," IEEE Transactions, PAS-38, No. 4, April 1969, pp. 375-88.

Schofield, G. H. and Marsh, Wilbur D., "The Effect of Reliability on Utility Financial Results," Public Utilities Fortnightly September 24, 1970, Vol. 86, No. 7.

Shipley, R. Bruce, Patton, A. D. and Denison, J. S., "Power Reliability Cost vs. Worth," IEEE Winter Power Meeting, Paper No. T72 143-1.

State of New York Public Service Commission. Case 25937 "Proceeding on Motion of the Commission as to the Plans and Procedures of Electric Corporations for Load Shedding in Time of Emergency."

Subcommittee on Communications and Power of the Committee on Interstate and Foreign Conference. U.S. House of Representatives, "Electric Power Reliability, 1969-1970," Serial Nos. 91-94, 91-95.

Turvey, Ralph W., Optimal Pricing and Investment in Electricity Supply, The M.I.T. Press, Cambridge, 1968.

Vennard, E., The Electric Power Business, McGraw-Hill, N.Y., 1970.