

ELECTRIC UTILITY CAPACITY EXPANSION PLANNING
WITH THE OPTION OF INVESTING IN SOLAR ENERGY

by

Konstantin Staschus

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE
in
Industrial
Engineering and Operations Research

APPROVED:

Dr Hanif D. Sherali

Dr/Lester C. Frair

Dr Saifur Rahman

May, 1982
Blacksburg, Virginia

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor Dr Sherali for his guiding and encouragement. It was a great pleasure to work with him, and I particularly appreciated that he always had time when I needed his advice. I would also like to thank Dr Frair and Dr Rahman for their help.

CONTENTS

<u>Chapter</u>	<u>page</u>
I. INTRODUCTION AND BACKGROUND	1
II. LITERATURE REVIEW	5
Some Concepts Used in Electric Utility Planning	5
The Demand Encountered	5
Equipment Types	7
How to Meet the Demand	12
Solar Equipment	13
Pricing	15
Marginal Cost Pricing	15
Capacity Expansion Planning	18
Linear Programming Models	23
Nonlinear Programming Models	33
Static Models: Optimal Plant Mix	37
Traditional Stages of Capacity Expansion Planning	38
General Power Systems Optimization	42
Peak Load Pricing	43
Classical Peak Load Pricing Theory	44
The Regulated Utility	47
Stochastic Demand	47
Diverse Technology and a Neoclassical Approach	49
Applications	53
The Solar Energy - Electric Utility Interface	54
III. CAPACITY PLANNING WITH SOLAR ENERGY	59
Approach to the Problem	59
Problem without Solar Energy	60
Problem with Solar Energy	62
Convexity	68
Continuity and Differentiability	69
Continuity	69
Differentiability	70
Proof of Convexity of $h(x_0)$	72
Algorithm	76
Illustrative Example and Computational Results	79

IV.	CAPACITY EXPANSION PLANNING WITH SOLAR ENERGY . . .	83
	Formulation	83
	Solving the Expansion Problem without Solar Energy	86
	Solution Method with Solar Energy Via Decomposition	87
	Introduction and Outline	88
	Formulation and Decomposition Technique . . .	90
	The Optimal Dual Variables	97
	Dual Variables Associated with Equipment Types	99
	Dual Variables Associated with Demand Blocks	107
	Derivation of the Cut	110
	Algorithm	112
	General Vector Formulation of the Technique	114
	The Benders Decomposition Framework . .	114
	Convergence	117
	Expedient for Solving the Master Program .	119
	Illustrative Example and Computational Results	122
	Approach Using Lagrangian Dual Method	125
	Comparison with the Decomposition Method .	131
	Multiyear Expansion Planning Model	133
V.	MANY PERIODS OF DIFFERENT AVAILABILITY OF SOLAR ENERGY	141
	Introduction	141
	Formulation	144
	Convexity	147
	Algorithm	149
	Derivatives of $h(x_0)$	149
	Algorithm Statement	151
	Inclusion of Storage for Solar Energy	157
	Optimal Hydro Dispatching	159
	Optimal Storage Policy for Solar Installations	163
VI.	CONCLUSIONS	169

Appendix

	<u>page</u>
A. ALGORITHM FOR THE EXPANSION PROBLEM WITHOUT SOLAR ENERGY	171
Predetermination of Old Capacity Used at Optimality	171
Predetermination of Some Optimal New Capacities Steepest Descent Feasible Directions Algorithm	173
Algorithm	176
REFERENCES	179
VITA 189	

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1. Example of a Daily Load Curve	9
2. Example of a Yearly Load Curve	10
3. Construction of the Load Duration Curve	11
4. Block Representation of a Load Duration Curve . . .	30
5. Breakeven Analysis	31
6. Method of Z-Substitutes	32
7. Nonlinear Programming Approach	36
8. Example for Wenders Approach to Peak Load Pricing .	52
9. Approach for Solar Investment Analysis	67
10. Breakeven Analyses for Example Problem	82
11. Discretization of Expansion Planning Program	92
12. Network Interpretation of Transportation Problem .	108
13. Possible Cases of Positive and Degenerate Flows .	109
14. Decomposition Algorithm Logic	120
15. Varying Availability of Solar Energy	154
16. Derivative of Y with Respect to x_0	155
17. Ranges of x_0 With Different Overall Peak Loads . .	156
18. Dispatching of Hydro Units with Storage	162
19. Night Inverse Load Duration Curve with Storage . .	167
20. Optimal Plant Mix with Solar Energy and Storage .	168

Chapter I

INTRODUCTION AND BACKGROUND

The major goal of the electric utility industry is to provide reliable service at minimal cost given customers' varying demand. Uncertainties in forecasting contingencies such as future demand and forced outages of generation units greatly complicate this task. Mathematical programming has been used extensively in such areas as planning the optimal economic operation of power systems, scheduling hydro units, and finding the optimal capacity expansion path for generation capacity, transmission lines, etc. This thesis applies mathematical programming to generation capacity planning and expansion planning.

There are different kinds of power plants, e.g. hydroelectric, coal-fired, oil-fired, gas-fired, nuclear plants, etc. Until the 1960's, electric utilities expanded their capacity almost entirely with these traditional types of plants. The oil embargo of 1973 however, jolted people to an awareness that fossil fuels and even uranium supplies will eventually be depleted. For example, if current growth trends persist, oil and natural gas could be depleted as soon as 50 years from now.

Most politicians and scientists concerned with our energy future agree that alternatives to a fossil fuel economy must be found. Some stress the importance of nuclear energy, fission, fast breeder and eventually fusion. Others advocate a transition to renewable energy sources, primarily solar energy, for environmental and long-term economic reasons. The label solar energy actually includes a variety of energy sources such as wind, ocean thermal and biomass. This thesis will focus only on the direct use of solar energy.

It has been argued that the only feasible solution for short-term energy problems can be a large effort in conservation and solar energy. To accelerate the spreading of conservation and solar equipment, it has been suggested that electric utilities get involved in the marketing of solar equipment in some way. Measures like public utility financing or leasing out of solar installations, consulting or contracting for installation, and even direct involvement in the installation process have been discussed. Besides getting involved in solar home heating projects, utilities also have the option of investing in solar thermal stations for the direct production of electricity.

For both applications of solar technology, the same question arises for the utility: How much solar capacity would be economic to install? An important difference between solar energy and the more traditional energy sources for producing electricity complicates this question. While fossil fuels and uranium usually are available day and night, any season of the year, solar energy's availability is restricted to daytime hours, and varies with the seasons. Under cloudy conditions, no solar energy can be harnessed even during daytime; when cloudy periods fall in conjunction with the utility system's peak demand, solar energy's contribution towards the system's reliable capacity can approach zero. To increase solar energy's availability during night and cloudy periods, storage devices are often used. In most cases, the heat produced is stored in thermal mass, i.e. in rock beds or thick walls.

The purpose of this thesis is therefore to investigate a capacity plan and a capacity expansion plan for the electric utility, in which an optimal mix of equipment types - possibly including solar generation and storage capacity - must be selected and optimally dispatched in order to meet with an anticipated demand. Of necessity, the peculiarities of the availability and characteristics of solar energy must be explicitly accounted for in that analysis.

In the following chapter, an extensive literature review is given. Not only literature on capacity expansion planning will be covered, but also the economic literature on peak load pricing, and technical and economic literature on the electric utility - solar energy interface. Chapters 3 through 5 contain the results of this thesis. In Chapter 3, a plant mix or capacity planning algorithm is developed, taking into account non-availability of solar energy during night time, but not its restricted day time output due to cloudcover. Chapter 4 also only takes into account non-availability during night time; it examines the capacity expansion problem with solar energy. A decomposition approach is applied to solve that problem. Finally in Chapter 5, both non-availability during nights and intermittent availability during day time are considered in the context of a capacity planning algorithm. It is shown how the approaches used in Chapters 3 and 4 need be modified only slightly to accomodate that.

Chapter II
LITERATURE REVIEW

2.1 SOME CONCEPTS USED IN ELECTRIC UTILITY PLANNING

To facilitate a review of the literature, this section describes ways of representing the demand for electricity, equipment types, basic concepts of how to meet the demand and pricing issues. These concepts are described in basic books on electric utility planning, see for instance VARDI & Avi-Itzhak, 1981.

2.1.1 The Demand Encountered

Load Curve: In most utility planning models, demand is considered an exogenous variable (see ANDERSON, 1972). It varies widely between different times of the day and different seasons. A typical load curve for a day is given in Figure 1. The typical shape of a yearly load curve depends on whether the utility experiences summer or winter peaks. In southern states, where little heating during winter months, but much air-conditioning in summer months is required, the peak demand usually occurs in the summer. In northern states, with more heating and less air-conditioning require-

ments, winter peaks are prevalent. A typical load curve for a winter-peaking utility is given in Figure 2.

Load Duration Curve: If the chronological load curves are rearranged so that the peak load appears first, and the load decreases as duration increases, a load duration curve is obtained. If for some duration a , the load duration curve $f(\cdot)$ takes the value $f(a)=L$, for example, then this indicates that the load is greater than or equal to L for a duration of a . A load duration curve, together with the load curve it was derived from, is shown in Figure 3. It is usually assumed that load duration curves are monotone decreasing and differentiable, and that their inverses exist. (The presence of a base load, i.e. a load that is demanded during all the time, violates differentiability, but it is easily accommodated in most models.)

Often, areas under a load duration curve are classified as base, intermediate and peak loads. A base load is the amount demanded during the whole period, while peak loads are only demanded for short durations. An example of this classification is also shown in Figure 3.

Load Factor: In this context, an utility's load factor is defined as the ratio of average load during the period and

the peak load. Thus, the more peaked a load duration curve is, the smaller will be the load factor. Because a steep load duration curve makes necessary the installation of much equipment which is only used during the small fraction of time in which peak demands occur, it is desirable to have a relatively flat curve.

2.1.2 Equipment Types

Currently, the main methods of producing electricity are with hydroelectric plants (with or without storage reservoirs and with widely varying capacities), nuclear plants and fossil fuel fired plants (coal or oil-fired). Peaking units, mostly gas turbines or pumped-storage hydro plants, are also used. These plant types all have differing cost characteristics which make them suitable for supplying either base, intermediate or peak loads.

Gas Turbines have low capital but very high operating costs, mainly fuel costs, and are therefore used for peak loads only. Fossil fuel plants have high capital costs and lower operating costs; they are mostly used for intermediate loads. Nuclear plants have very low operating costs, but extremely high capital costs; they are cost-efficient only if run almost all the time, therefore they are used for base

loads only. Run-of-river hydro plants (without reservoirs) are used for base loads because of their almost zero operating costs. With reservoirs, the possible energy output of hydro plants is determined and limited by the water inflow, this is called an energy constraint. Then, each hydro plant's unique combination of capacity and energy available determines what section of the load curve it is operated in.

In utility planning methods, it is usually assumed that the operating costs of different plant types are ordered such that $g_1 < g_2 < \dots < g_I$, while for capital costs $c_1 > c_2 > \dots > c_I$ holds. This assumption is valid for the following reason. If both the capital and operating costs for a particular plant type are comparatively higher than for another type, the first plant type would never be economical to install.

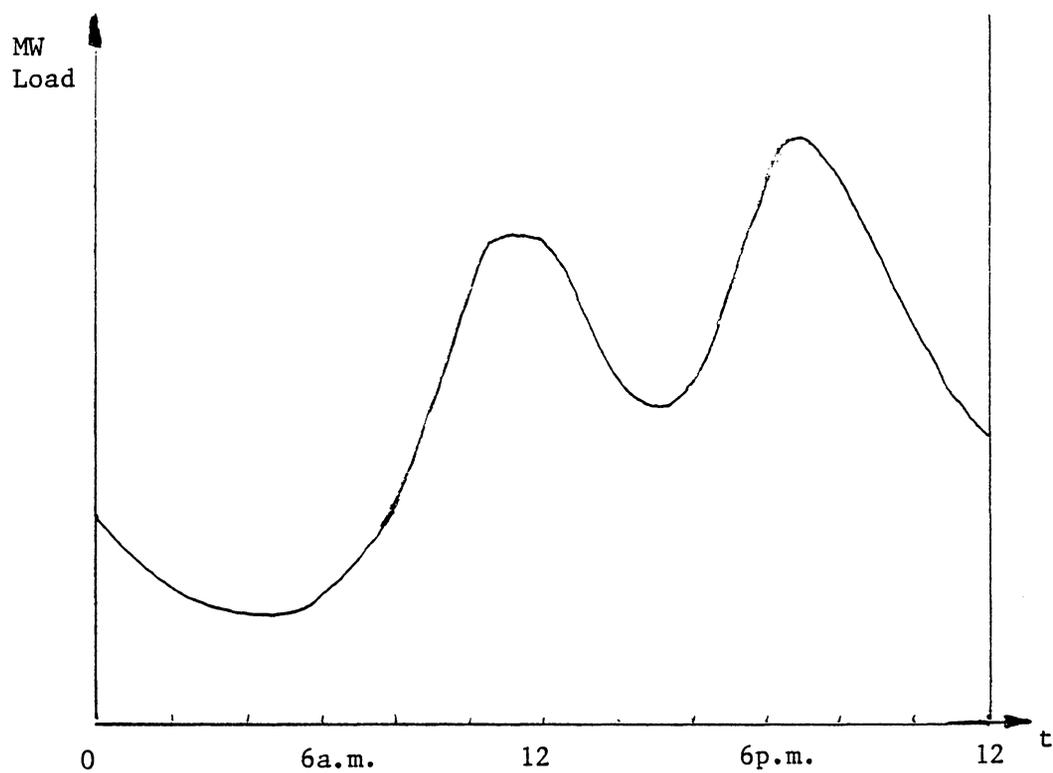


Figure 1: Example of a Daily Load Curve

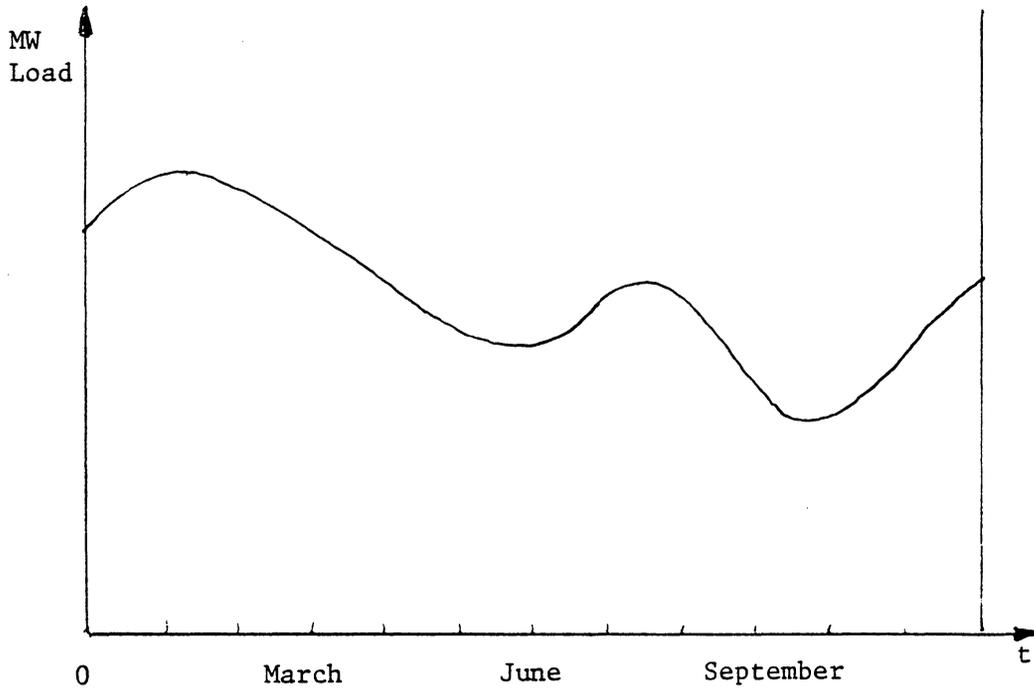


Figure 2: Example of a Yearly Load Curve

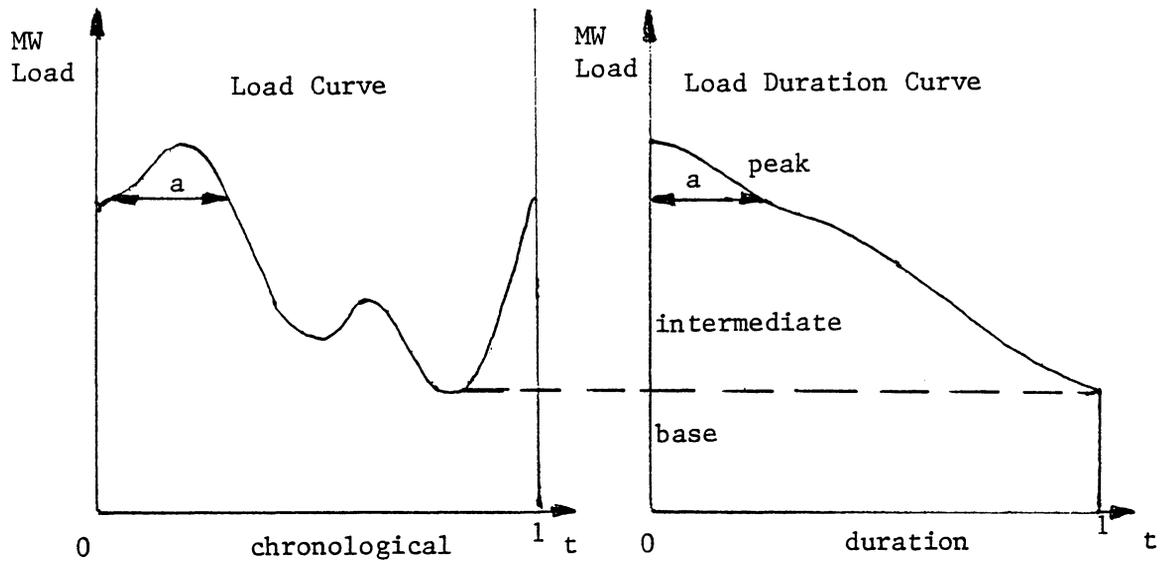


Figure 3: Construction of the Load Duration Curve

2.1.3 How to Meet the Demand

Once an equipment mix is selected, the problem is to meet the varying demand at minimal operating cost. This is achieved by merit-order dispatching. For any given load, the plant with the least operating cost is dispatched first, then the one with the next higher operating cost, and so on, until the demand is met. Thus, the units with the least operating costs will be operated during most of the time, while peaking units with high operating costs are really only used during the periods of peak demand.

In this context, a plant's capacity factor is defined as the fraction of time a plant is operated. This factor not only depends on the plant's position in the merit order, but also on its availability which can be restricted by maintenance requirements or forced outages.

It should be noted though that day to day on-line decisions may actually cause plants to be dispatched out of merit order, due to load fluctuations, shut down and start up costs, and other similar reasons.

For the whole utility, the loss-of-load probability is defined as the probability that demand will exceed total available capacity. It is usually given in terms of days during which load exceeds demand, in years; for example, a

well accepted standard is a loss-of-load probability of one day in ten years.

2.1.4 Solar Equipment

In this thesis, only the direct use of solar energy, either to produce electricity or to heat water or homes, is considered. Electricity from solar energy can be produced either with solar thermal power plants, called power towers, or with photovoltaic cells. Photovoltaic cells convert sunlight directly into electricity. They can be grouped into large or small arrays to produce energy in large or small quantities, respectively. While the operating cost of these cells is near zero, the relatively high capital cost so far restricts their use to extreme situations such as in space programs or in remote radar stations.

Pilot power towers, with capacities of up to 10 MW, are in operation or under construction in many parts of the world (see WEINGART, 1979, for a list). The plants consist of large arrays of mirrors positioned in a circle around a central receiver, tracking the sun's rays. These mirrors concentrate the sun's rays onto the central receiver, in which the concentrated energy generates steam. The steam is then sent through a conventional turbine and valve cycle,

the turbine drives a generator, which in turn produces electricity.

In most cases, part of the steam can be used to heat a large rock bed or a similar energy storage device. By drawing heat from storage, the installation can continue producing electricity during the night and during cloudy periods. Electricity can also be used to produce hydrogen by electrolysis as a means for longer term energy storage (see RALPH, 1972). The hydrogen can be easily stored, transported and used as a fuel for many purposes. However, such applications of solar energy are not economical yet.

For solar home heating, one distinguishes between passive and active solar systems (see, e.g., KREITH & West, 1980). Passive systems essentially use the building itself as a solar collector; large windows allow the sun's rays to enter the house, massive walls store the heat, vents are usually the only moving parts of the system. It is very economical to design new houses with passive solar features (see FELDMAN & Wirtshafter, 1980). Active systems mostly employ flat-plate collectors situated on the roof or in the garden. In the collector, which has a glass cover, a working fluid or gas is heated by the sun's energy, and circulated to a storage device. For storage, hot water, rock beds and other

methods are used. This method has already proven to be economical for hot water heating, and in parts of the country for home heating (see BEZDEK et al, 1979).

2.1.5 Pricing

The pricing problem is tied to other planning problems in electric utilities in three ways:

- Prices influence demand and thus both capacity expansion and operation cost.
- It will be shown in this section, and it has long been established in the economic literature, that at the welfare optimum, prices should reflect marginal costs.
- If linear or nonlinear programming is used to find the optimal capacity plan, the dual variables can be interpreted as marginal costs.

2.1.5.1 Marginal Cost Pricing

The rule that prices should equal marginal costs can be based both on formal mathematical analysis and on intuitive arguments. The mathematical derivation given below follows largely the one given by BERLIN et al, 1974, in their appendix. The intuitive arguments can be found in BOITEUX, 1964, as well as in BONBRIGHT, 1961.

In the economic literature, welfare (W) is defined as total benefit (TB) minus total cost to society ($SC=g(Q)$): $W = TB - SC$.

Let $P=f(Q)$ be the demand function which represents society's willingness to pay for varying amounts of commodity Q .

Then, total benefits are defined as the integral of the demand function: $TB = \int_0^{\hat{Q}} PdQ$.

Thus, $W = \int_0^{\hat{Q}} PdQ - g(Q)$.

Maximizing welfare yields the necessary condition:

$$\frac{dW}{dQ} = P - \frac{dSC}{dQ} = P - g'(Q) = 0, \text{ and therefore } P = g'(Q),$$

which is the result stated above that prices should be set at marginal costs.

BOITEUX (1964) presents arguments for pricing electricity at long-run marginal cost. Selling at cost ensures that society's economic optima are also seen as optimal by the individual or firm; marginal cost is the actual cost of expanding service to meet additional demand. It is also the cost of continuing service or the saving incurred when contracting service.

Pricing at short-run marginal cost would imply having constantly changing tariffs, depending on the instantaneous highest operating cost equipment. When more capacity is installed than is necessary to meet the peak demand, short-run marginal costs do not include any capital costs. But when the installed capacity is insufficient to meet peak demands, short-run marginal cost is essentially infinitely high, since it is impossible to serve any additional demand at peak times.

For these reasons, it is more reasonable to price electricity at long-run marginal costs, which do include capital cost components.

2.2 CAPACITY EXPANSION PLANNING

ANDERSON gives a good overview on modelling for capacity expansion planning for electric utilities. (For similar overviews, see BERRIE & Anderson, 1969, TURVEY ,1968, and BERRIE, 1967.) In his introduction, Anderson justifies the application of mathematical modelling to this problem. First, the enormous investments involved motivate a detailed treatment of the problem. Between 1945 and 1973, for example, electricity demand in the U.S. expanded at rates of about 7%, requiring investments of the order of hundreds of billions of dollars. (Current growth rates, however, are down to 1% to 2% per year.)

The complexity of the problem also requires the application of formal mathematical models. The problem of which plant types to operate under what load conditions, described in the previous section, is already quite complex. If the objective is to develop not only the static optimal plant mix, but a dynamic capacity expansion plan, then a plant's position in the merit order and its capacity factor in future years will be influenced by future investments. This as well as other complications make the capacity expansion problem so complex that mere guesswork, as opposed to a mathematical program, would most likely find only inferior, suboptimal solutions.

After this initial justification, Anderson proceeds to formulate in generic terms the mathematical programs most often used. The objective always is to minimize the sum of operating and capital costs, subject to such constraints as: All demand must be met with the capacity available at peak times as the upper capacity limit; if applicable, energy or capacity constraints for hydroelectric generation must be satisfied; and specified loss-of-load probabilities must not be exceeded.

In the most general case, the objective function takes the following mathematical form:

$$\min \sum_{v=1}^T \sum_{j=1}^J c_{jv} x_{jv} + \int_0^T \sum_{v=0}^t \sum_{j=1}^J g_{jv}(t) \cdot U_{jv}(t) dt, \text{ where}$$

$t=T$ = planning horizon

v = vintage (years of commissioning new plant)

$v=0$ = plants already installed

$j=1, \dots, J$ = plant types

c_{jv} = capital cost of equipment j for vintage v

x_{jv} = capacity of plant type j , vintage v

$g_{jv}(t)$ = operating costs of plant j, v (discounted)

$U_{jv}(t)$ = power output of plant j, v at instant t .

Thus, the first term in the objective function represents the system's total capital cost, while the second term represents total operating cost.

Often it is more convenient to use a discrete approximation of this function:

$$\min \sum_{v=1}^T \sum_{j=1}^J c_{jv} x_{jv} + \sum_{t=1}^T \sum_{v=0}^t \sum_{j=1}^J g_{jvt} U_{jvt} d_t ,$$

where d_t = width of time interval considered at time t .

The most important constraints may be stated mathematically in the following general terms:

$$\sum_{j=1}^J \sum_{v=0}^T U_{jvt} \geq Q_t, \quad t=1, \dots, T, \quad \text{where } Q_t \text{ is the demand at } t.$$

(Demand must be met at all times.)

$$0 \leq U_{jvt} \leq a_{jv} x_{jv},$$

for $j=1, \dots, J$; $v=0, \dots, T$; $t=1, \dots, T$;

where a_{jv} is the availability of plant j, v .

(No plant can be operated above its peak available capacity.)

The specified loss-of-load probability cannot be exceeded:

$$\Pr\left(\sum_{j=1}^J \sum_{v=0}^T a_{jv} x_{jv} - \hat{Q}_t \geq 0\right) \leq e, \quad t=1, \dots, T, \quad \text{where}$$

$$\hat{Q}_t = \text{yearly peak demand}$$

e = specified loss-of-load probability.

The reader may note that in many countries, system reliability is not secured by planning with this type of constraint, but by installing a certain margin of spare available capacity to meet demands above expected peak demand. Then, the reliability constraint takes the following form:

$$\sum_{j=1}^J \sum_{v=0}^T a_{jv} x_{jv} \geq \hat{Q}_t(1+m), \quad t=1, \dots, T,$$

where m denotes the margin of spare capacity.

It may be mentioned in passing here that more complex models include constraints dealing with peculiarities of mixed hydro-thermal systems and the issue of hydro scheduling and investment planning.

This general formulation of the capacity expansion problem can be modified in many ways. It can be extended so that issues like optimal replacement, optimal locations of plants, transmission optimization, nuclear fuel cycling or optimal storage policies for hydro plants are included. On the other hand, simplifying assumptions can be made which change the structure of the problem.

If one maintains a nonlinear load curve - as opposed to discretizing it -, then the durations of operation of plants

are nonlinear functions, thus resulting in a nonlinear objective function. Sometimes even nonlinear operating costs are assumed, which occur when a plant's efficiency changes with its loading. In this nonlinear case, the problem is essentially formulated as described above. For some time during the 1960's, most planners preferred nonlinear over linear models. Reasons for this and nonlinear formulations of the problem are given in the subsection on 'Nonlinear Programming Models'.

If the load duration curve is broken down into blocks of varying width d (as shown later in Figure 4), the problem can be formulated as a linear program. This approach is described in the following subsection.

If one is only interested in finding the optimal plant mix for the static case, all parameters dealing with vintage of new plant can be deleted from the problem. The resulting simpler problem is presented in the subsection 'Static Models - Optimal Plant Mix'.

Finally, the general traditional framework for applying any planning algorithms will be described, and limitations will be shown.

2.2.1 Linear Programming Models

MASSE & Gibrat, 1957, were the first ones to apply linear programming to investments in the electric power industry. In their 1957 Management Science article they justify in great length the use of this 'complicated' approach. But of course, compared to programs used today, Masse and Gibrat's is very simple. They do not consider any type of uncertainty, and they assume linearity of costs. Their program minimizes total cost subject to constraints for guaranteed power, peak power, annual energy production and limited funds; that is, instead of working with the load curve itself, they pick a few key values that represent its main characteristics (see also MASSE, 1962).

Most later linear programming applications represent the load duration curve with a histogram (see Figure 4), see, for instance, ANDERSON, 1972, BERRIE & Anderson, 1972, SHERALI et al, 1981a, BEGLARI & Laughton, 1974, ADAMS et al, 1972. The basic form of the linear programming models is:

$$\min \sum_{j=1}^J \sum_{v=1}^T c_{jv} x_{jv} + \sum_{j=1}^J \sum_{t=1}^T \sum_{v=0}^t \sum_{p=1}^P g_{jtv} U_{jtv} d_p ,$$

where d_p is the width of a block of the discretized load duration curve,

subject to linear constraints similar to the ones in the general nonlinear program. The two most important constraints, for meeting all demands and using no more than available capacity, take the following form:

$$\sum_{j=1}^J \sum_{v=0}^t U_{jtv} \geq Q_{tp}, \quad t=1, \dots, T; \quad p=1, \dots, P$$

$$U_{jtv} \leq a_{jvt} x_{jv}, \quad j=1, \dots, J; \quad v=0, \dots, t; \quad t=1, \dots, T; \\ p=1, \dots, P.$$

As is shown in SHERALI et al, 1981a, (the initial work in this area was done by TURVEY, 1968a) this basic LP-formulation yields the same optimal solution as the traditional, straightforward breakeven analysis. The breakeven analysis is demonstrated geometrically in Figure 5. Of course, it finds the optimal plant mix only on the basis of capital and operating costs, other constraints are not considered. Its rationale is that for an incremental load occurring for a duration a , it is cheapest to buy incremental capacity and to operate it for duration a such that the capacity type has the minimal $(c_j + g_j a)$ from among all plant types j .

Algebraically, e.g. for the case of three plant types, this is expressed as follows:

$$a_1 = \frac{c_1 - c_2}{g_2 - g_1}, \text{ and } a_2 = \frac{c_2 - c_3}{g_3 - g_2}$$

The breakeven analysis can be used as part of a complex, multiyear optimization, e.g. see PHILLIPS et al, 1969.

To be able to take into consideration the many uncertainties involved, e.g. in demand forecasting or with forced outages, one can use in the breakeven analysis a load duration curve that is adjusted for these stochastic effects, called the equivalent or combined load duration curve (for the derivation that works with convolving the involved probability distributions see VARDI et al, 1977a, or VARDI & Avi-Itzhak, 1981). BOOTH, 1972, uses this type of approach to find the expected value of energy produced by each plant, and then applies a dynamic program for optimization of both operation schedule and expansion path.

The method of Z-substitutes was developed since with the LP-formulation described above, one gets a very large number of constraints if one deals with many plant types and many load curve blocks (see, e.g., ADAMS et al, 1972, ANDERSON, 1972, BEGLARI & Laughton, 1974). This method uses the fact that moving from one block of the load duration curve to the

next one, as the demand level decreases, each plant is operated at most at the same level as in the previous block. Then, for one plant type j the capacity constraints for all load duration curve blocks p , $x \geq U_p$, can be replaced by just one constraint $x \geq \sum_{p=1}^P Z_p$, where Z_p is the reduction in power output level for the plant from block p to $p+1$ (if the indices for time t and vintage v are left aside for the moment). See Figure 6 for demonstration and derivation of this result.

An efficient and elegant approach to deal with large scale linear programs is to use decomposition techniques. COTE & Laughton, 1979, describe the successful application of the Benders Partitioning Method to power system planning. Their approach is described in detail below, since Benders Decomposition will be used extensively in this thesis, although in a manner quite different from the one used in COTE & Laughton.

They use a matrix representation of the capacity expansion problem to illustrate the approach. Let C and G be the vectors of capital and operating costs, associated with the vectors of capacity and operating decisions, X and Y , respectively. Let $AX \geq D$ represent demand constraints in matrix form, and let $GX + HU \geq B$ represent capacity const-

straints. Then, the expansion problem takes the following form:

Expansion Problem EXP

$$\begin{aligned} \min \quad & C^T X + F^T U \\ \text{s.t.} \quad & AX \geq D \\ & GX + HU \geq B \\ & X, U \geq 0 \end{aligned}$$

If X is fixed, a simple production problem results, which can be solved very efficiently by applying merit order loading. This suggests the representation of EXP as follows:

Decomposed EXP (DEXP)

$$\min_{X \in S(X)} \left\{ C^T X + \min \left[\begin{array}{l} F^T U \\ \text{s.t. } HU \geq (B - GX) \\ U \geq 0 \end{array} \right] \right\}$$

where $S(X) = \{X/AX \geq D, X \geq 0\}$.

Introducing dual variables Λ , the dual of the production problem can be written as follows:

Dual Production Problem DPP

$$\begin{aligned} \max \quad & \Lambda^T (B - GX) \\ \text{s.t.} \quad & \Lambda^T H \leq F \\ & \Lambda \geq 0 \end{aligned}$$

Note that the dual feasible region is independent of X , and the optimal solution will be at one of the extreme points of the feasible region. Therefore DPP can be represented as

$$\max_{\Lambda \in L} \Lambda^T (B - GX)$$

where $L = \{\Lambda_1, \Lambda_2, \dots, \Lambda_K\}$, and Λ_i represents an extreme point of DPP's feasible region.

After manipulations, the total problem can thus be written as follows:

Problem in Decomposed Form PDF

$$\min C^T X + y_0$$

$$\text{s.t. } AX \geq D$$

$$\Lambda^T (B - GX) \leq y_0 \quad \text{for all } \Lambda \in L$$

$$X \geq 0$$

Using these concepts, the following algorithm is quoted from COTE & Laughton, 1979:

Step 1: Solve a relaxation of PDF where the set L has been replaced by \bar{L} . This yields an investment plan \bar{X} together with a lower bound on the optimal value of PDF or equivalently of EXP. This relaxation will be called the master.

Step 2: Using the investment plan \bar{X} , the production problem is solved yielding a production schedule \bar{U} , dual variables $\bar{\Lambda}$ and an upper bound on the optimal value of EXP, given by $C^T \bar{X} + F^T \bar{U}$. $\bar{\Lambda}$ is added to the set L and control is returned to Step 1.

The algorithm is stopped when both lower and upper bound fall within a predefined value ϵ of each other. Because the set L is finite, the algorithm will converge to the optimal solution in a finite number of steps.

COTE & Laughton cite good computational results with the application of this method. They found it to be favorably comparable to other decomposition approaches and other methods of solving expansion planning problems.

The application of decomposition techniques to power system planning is a fairly new approach, and except for COTE & Laughton's paper, it has not been treated in the literature extensively.

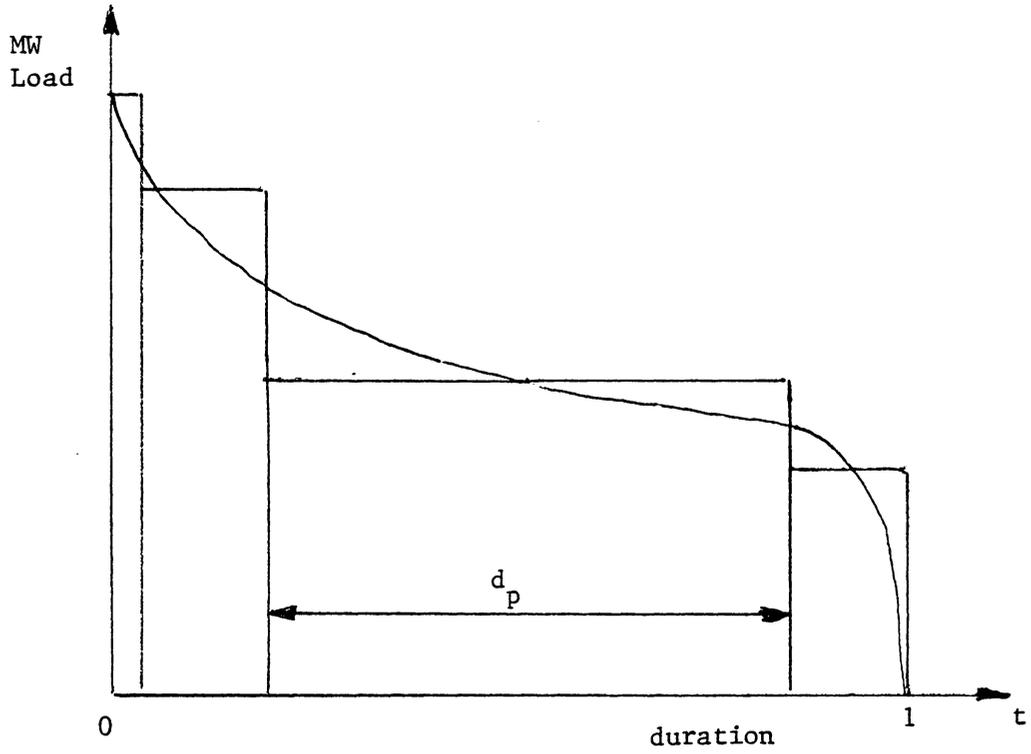


Figure 4: Block Representation of a Load Duration Curve

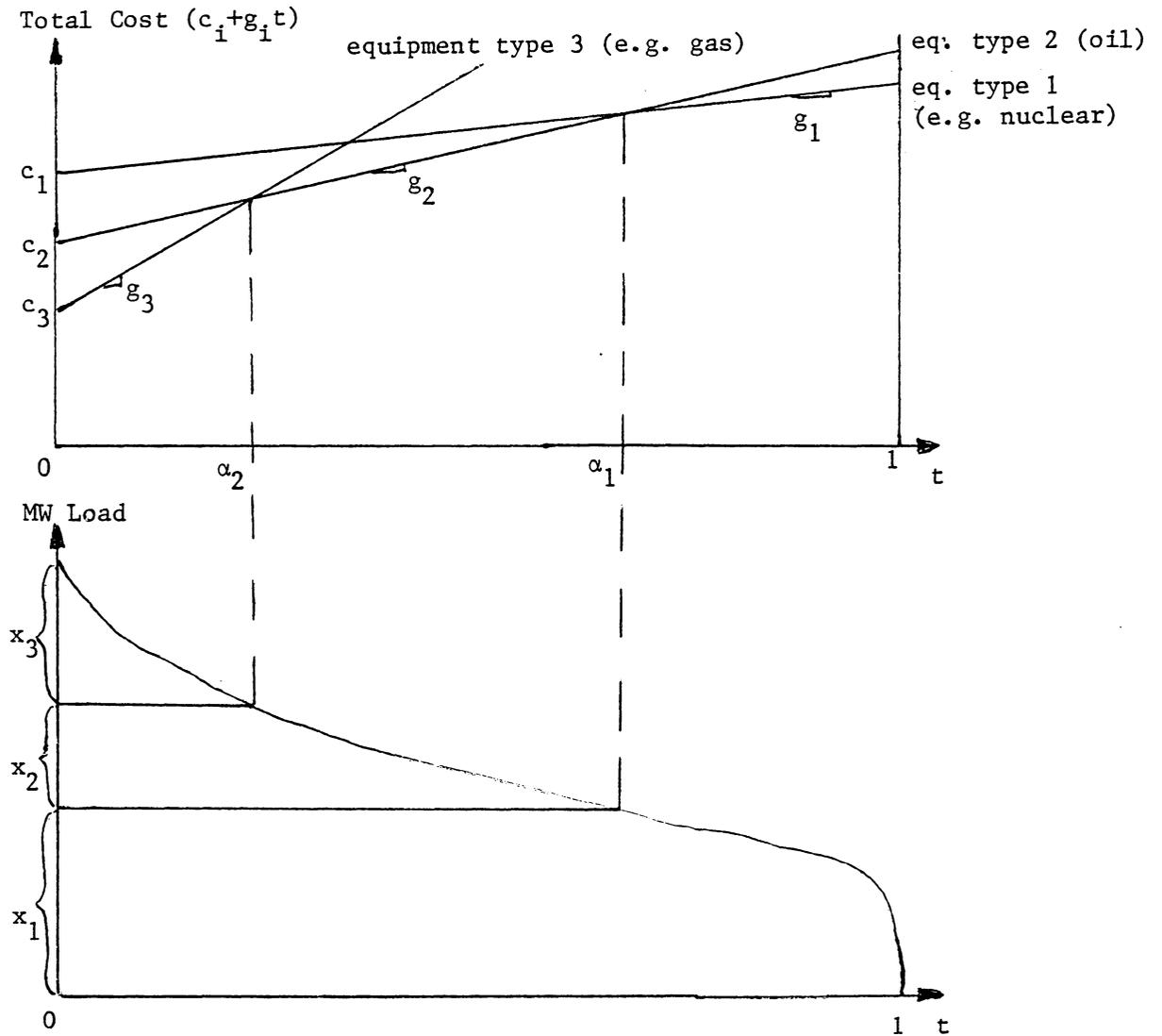


Figure 5: Breakeven Analysis

Derivation of method of z-substitutes:

From defn of LDC: $U_p \geq U_{p+1}$, for all p

Capacity constraints: $x \geq U_p$, for all p

Then, by defn of Z_p : $Z_p = U_p - U_{p+1} \geq 0$, for $p=1, \dots, P-1$, and

$$Z_P = U_P \geq 0$$

Summing these last two constraints yields $\sum_{p=1}^P Z_p = U_1 \geq 0$,

and since $x \geq U_p \geq U_{p+1} \geq \dots$,

the result $x \geq \sum_{p=1}^P Z_p$ follows.

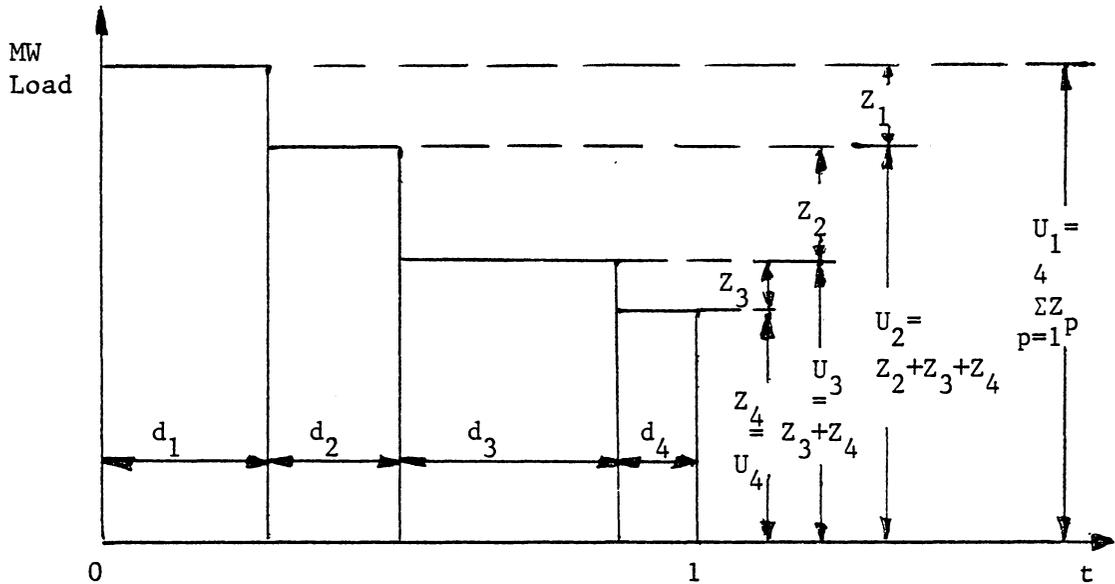


Figure 6: Method of Z-Substitutes

2.2.2 Nonlinear Programming Models

A major incentive for developing the method of Z-substitutes described in the previous subsection, was that computers in the early 1960's could not handle the vast number of capacity constraints in the original formulation. Before the Z-substitutes method came up, BESSIERE, 1971, and PHILLIPS et al, 1969, developed a nonlinear programming approach that overcame this constraint problem (also described in ANDERSON, 1972).

The main idea is to prearrange all installed and possible new plants in merit order before starting the algorithm. That is, for the operating costs g , the plants are indexed such that

$0 < g_1 < g_2 < \dots < g_W$ holds. Then, all operating variables and capacity constraints are satisfied implicitly and do not have to appear in the problem.

For the following derivation, let plant type j and vintage v be represented by the single index $w=1, \dots, W$. Let, as defined before, x_w be the capacity of w , and let $F(x)$ be the inverse load duration curve (the time index t is left aside for the moment). X is defined as

$$X_w = \sum_{u=1}^w x_u, \text{ as shown in Figure 7.}$$

Then, the cost of operating plant w in merit order is given by

$$\int_0^{X_w} g_w F(x) dx = g_w (G(X_w) - G(X_{w-1})), \text{ where } G(X_w) = \int_0^{X_w} F(x) dx.$$

The total operating cost is obtained by adding this over $w=1$ to W :

$$TOC = \sum_{w=1}^W (g_w - g_{w-1}) G(\sum_{u=1}^w x_w)$$

Using this total operating cost in the objective function, and reintroducing subscripts j, v and t , the capacity expansion problem takes the following form:

$$\begin{aligned} \min \quad & \sum_{j=1}^J \sum_{v=1}^T c_{jv} x_{jv} + \sum_{t=1}^T \sum_{j=1}^J \sum_{v=0}^t (g_{jvt} - g_{j,v+1,t}) G_t \left(\sum_{i=1}^j \sum_{u=0}^v x_{iu} \right) \\ \text{s.t.} \quad & \sum_{j=1}^J \sum_{v=0}^T x_{jv} \geq \hat{Q}_t, \text{ for } t=1, \dots, T, \\ & \text{where again } \hat{Q}_t \text{ is the peak demand in } t. \end{aligned}$$

Thus the total operating cost is given in closed form in terms of plant capacities and operating costs and the inverse load duration curve. Also, the capacity constraints are already implicitly satisfied. This is an enormous reduction in program size. However, the objective function is quite complex, nonlinear, nonseparable, but at least convex. So the program can be solved on computers without difficul-

ties; and this type of model has been used in France and Britain for quite some time, see, e.g., BESSIERE, 1971.

Later research though, after the Z-substitutes method had been developed, has preferred the linear programming approach, due to its relative simplicity and advantages with computer software.

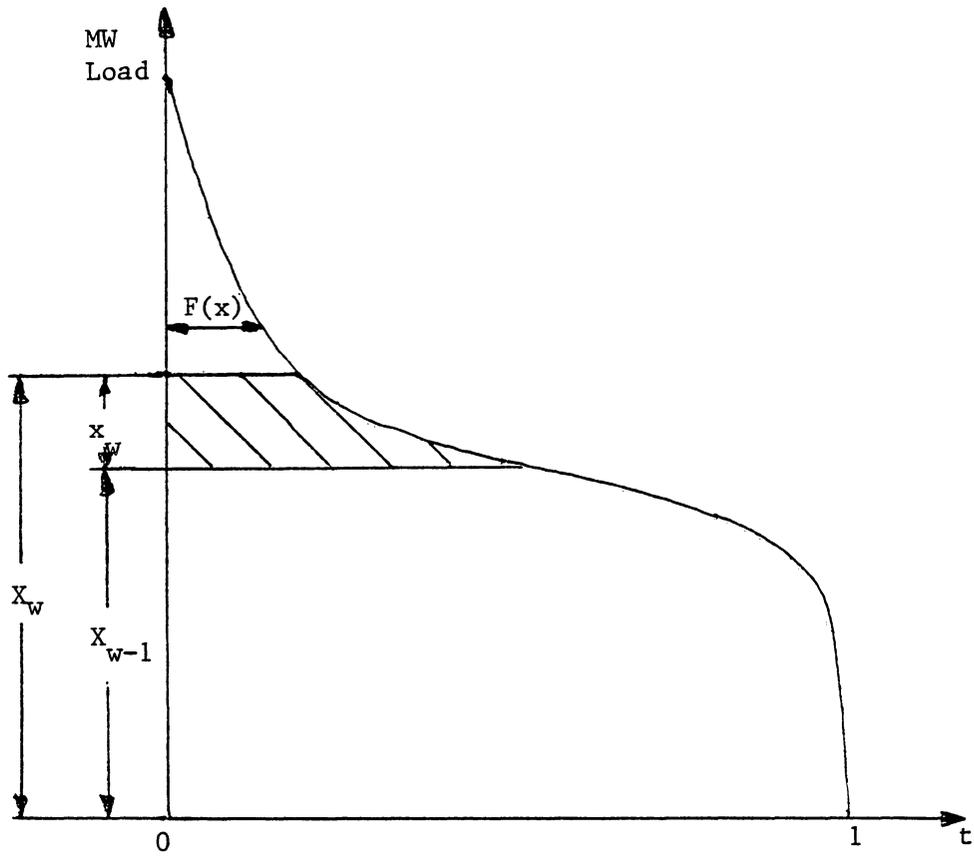


Figure 7: Nonlinear Programming Approach

2.2.3 Static Models: Optimal Plant Mix

The optimal plant mix problem can be regarded as a subproblem of the capacity expansion problem. It is modified so that already existent plants are not taken into account. Nor are investment decisions made for many future years; only the static optimal capacity mix is sought. Again, linear or nonlinear programming can be applied. The general - nonlinear - approach is described in detail in the Chapter 'Approach to the Problem', while the linear approach is shown below:

$$\begin{aligned} \min \quad & \sum_{j=1}^J c_j x_j + \sum_{j=1}^J \sum_{p=1}^P g_{jp} U_{jp} d_p \\ \text{s.t.} \quad & \sum_{j=1}^J U_{jp} \geq Q_p, \quad p=1, \dots, P \\ & U_{jp} \leq a_j x_j, \quad j=1, \dots, J; \quad p=1, \dots, P \end{aligned}$$

Of course, as with all other models, constraints for hydro generation, loss-of-load probabilities, and other 'local' constraints can easily be added.

If no other constraints are added, the simplest solution method for this program is via breakeven analysis, as described above.

2.2.4 Traditional Stages of Capacity Expansion Planning

As described in ANDERSON, 1972, the planning process consists of three stages: marginal analysis, simulation and application of global models.

- In the MARGINAL ANALYSIS stage, one starts with a more or less arbitrary reference solution and improves it by marginal substitutions of one feasible plant for another, maintaining the same capacity and energy output as in the reference solution. One difficulty with this planning method is that each plant examined has to be located on the system operating schedule so as to minimize costs. Also in complex systems, many marginal substitutions have to be made before the optimum is found. To tackle the first problem, one uses simulation programs. To tackle the second, one starts with a near-optimal so-called background plan as a reference solution. (For an example of a simple comparison between two investment alternatives, see TURVEY, 1963)

- SIMULATION programs calculate for each constellation of plants the minimal operating cost. The load curve is integrated, and the area multiplied by the respective operating costs. Plants are assumed to operate in merit order. Dynamic, integer and linear programming algorithms have been applied to this. Especially when applied to systems includ-

ing hydro units, the problem of finding the optimal operation schedules is very challenging, and much research has gone into this field (e.g., see EL HAWARY et al, 1979, GASSFORD & Karlin, 1958, BONAERT et al, 1971, ARVANITIDIS & Rosing, 1970, GAGNON et al, 1974, TYREN, 1969, MANNE, 1960, LOGENDRAU & Oudheusden, 1981). Similar problems are optimal load flow, minimal loss and economic dispatching. These problems are treated in more technical terms by electrical engineers, see e.g. HAPP, 1974, KIRCHMAYER, 1958, EL HAWARY et al, 1979. SASSON & Noulin (1969) give a unified approach for all three problems mentioned.

- GLOBAL MODELS try to find the optimal capacity expansion path with just one algorithm, unlike the time-consuming trial and error approach of the marginal analysis. On the other hand, to be able to keep a complex 30-year expansion plan from getting too huge for even today's efficient computers, one has to make some simplifying assumptions. To find the optimal expansion policy in more detail, one still has to apply marginal analysis and simulation programs on the results of the global models. For the global models, both linear and nonlinear programming approaches have been tried, some of which have been described in more detail above.

A few authors report how all three stages are applied on capacity expansion planning. BESSIERE, 1971, for instance, describes how a combination of marginal analysis and a nonlinear program is used for the Investment '85 model of Electricite de France, JENKIN, 1974, describes a similar approach for the British CEGB. Also, BERRIE, 1966, NITU et al, 1969, and GALLOWAY et al, 1966, apply this general type of approach in their computer programs.

One should be conscious of the following limitations of the approach described above, and of most models described in this literature review. First, demand is assumed to be exogenous; the anticipated loads obtained from more or less reliable forecasts is treated as input data. Some economic models examine America's or even the whole world's total energy system, of which the electricity system is a subsystem. In this more general setting demand can be treated as a variable. A few of these models are mentioned in the subsection on general power systems optimization.

Second, for the approach described in this subsection, the objective is taken to be cost-minimization. This assumption can be relaxed so as to maximize welfare. The resulting problem, in which demand is again a variable, can be treated with mathematical models of the type described in

'General Power Systems Optimization', or with simpler and much less complete economic models, some of which are described in the 'Peak Load Pricing' section.

Third, as ANDERSON, 1972, puts it,

the use of one or more investment models is (only) the first of several stages of the investment decision process. Engineering analysis of solutions follows and generally requires a revision of the solutions. The investment program finally selected must satisfy a number of engineering criteria regarding system stability, short-circuit performance, the control of watts, vars, and voltage, and the reserves and reliability of supply. The search for an investment program which satisfies engineering and economic criteria is an iterative, multi-disciplinary process.

Besides the authors mentioned above who generally follow the planning approach described here, other authors use more unusual approaches for finding the optimal expansion path. For example, LENCZ, 1969, combines a linear program with the method of operational games, PETERSEN, 1974, constructs a dynamic programming algorithm, NOONAN & Giglio, 1977, use a nonlinear mixed integer program and solve it with Bender's Decomposition, and LOUVEAUX, 1980, applies a multistage stochastic program with recourse.

COHON et al, 1980, develop a multiobjective linear program for power plant siting decisions; and DEES et al, 1980, and GARVER et al, 1976, examine the effects of load-growth uncertainties on long-range planning.

2.2.5 General Power Systems Optimization

In this subsection, a few more general models will be mentioned. In particular, the models described above do not include the price of electricity as a variable; they only treat capacity planning and operation scheduling. Since the price of electricity influences demand, changes in prices can have effects on these planning types. Therefore it makes sense to also examine more integrated energy models, which do treat prices and demand as variables.

Some authors deal with electric power systems as subsystems of a large overall energy system, for finding the optimal energy policy for the U.S. or even for the entire world (e.g. RATH-NAGEL & Voss, 1981, MANNE, 1974, ERLINKOTTER & Trippi, 1977, HOFFMANN & Wood, 1976). Others deal with power systems only (SASSON & Merrill, 1974, TURVEY, 1963).

URI, 1975, constructs an intertemporal-spatial model to find out an efficient allocation of electrical energy in the U.S. The objective is welfare maximization (as defined in the economic literature). The decision variables are the capacity expansion path, operation schedules and the price of electricity. The resulting nonlinear program is solved using the Kuhn-Tucker conditions.

ROWSE, too, develops a nonlinear program to solve for intertemporal prices, supply quantities and capacity additions. He also derives tradeoffs between prices, supply quantities, capacity, and environmental protection.

2.3 PEAK LOAD PRICING

When linear or nonlinear programming is applied to the capacity expansion problem, the resulting dual variables can be interpreted as marginal costs. In linear programs which represent the load duration curve in block form, the dual variables associated with demand constraints for the blocks of varying demand are the marginal costs of providing service during times of that demand (see, e.g., SHERALI et al, 1981a).

Economists using economic models which generally are much simpler than the capacity expansion models, have also derived marginal costs for periods of different demand for electricity. Their studies have led to the long-lasting discussion on peak load pricing.

2.3.1 Classical Peak Load Pricing Theory

The seminal paper on peak load pricing in the American literature is due to STEINER, 1957 (although some French economists, e.g. BOITEUX, 1960, originally 1949, had advocated marginal cost and peak load pricing earlier; DREZE, 1964, reviews these French contributions). Steiner uses the welfare maximization approach which is standard in the economic literature to examine optimal supply levels and prices for the case that periods of equal length - peak and off-peak periods - exist for the demand. In the main part of the paper, he explains his results for the case of only one peak and one off-peak period, employing graphical analysis of demand curves. The mathematical derivation for the many-period case, given in the Appendix, employs some unconventional maximization techniques; therefore the following mathematical derivation of the classical peak load pricing results follows the one shortly described in WENDERS, 1976.

Assume that the load duration curve is divided into three pricing periods, lasting d_1 , d_2 , d_3 fractions of the year, with respective demands of $Q_1 > Q_2 > Q_3$. Important assumptions are that the markets in the three periods are separable and independent and that the same capacity serves all periods.

Welfare is defined as the sum of producers' and consumers' surpluses, given by the integrals of demand curves minus costs. Maximizing this sum yields the optimal prices for the three periods.

$$\max W = \sum_{i=1}^3 d_i \int_0^{Q_i} P_i dQ_i - \sum_{i=1}^3 d_i g_i Q_i - CQ_1,$$

where W denotes welfare, the P_i prices and the g_i marginal operating costs for the different periods. C denotes per unit capital cost.

Finding the partial derivatives $\delta W / \delta Q_i$ and setting them equal to zero yields the following optimal prices:

$$P_1 = g_1 + C/d_1$$

$$P_2 = g_2$$

$$P_3 = g_3$$

This result shows that only peak users should bear any capital cost, while off-peak customers only pay operating costs. Since only peak users press against capacity, this result makes intuitive sense. It also reminds one of complementary slackness conditions obtained from mathematical programs (only for binding constraints, dual variables are nonzero; this can be read as: Only for periods pressing against capacity, marginal capital costs are nonzero).

A widespread and long-lasting discussion followed Steiner's initial article. The first authors to comment on Steiner, e.g. HIRSHLEIFER, 1958, agreed with Steiner's results in principle. WILLIAMSON, 1966, obtains slightly different results from Steiner's when introducing indivisibility of plants and varying lengths of periods. Later, several new aspects which brought the model closer to reality, were added to Steiner's approach, and different results were obtained.

Some of the related research which ensued includes the following (since these papers are not important for this thesis, they are given only with the author and a very short description of the contents): OFFICER, 1966 (optimality of pure competition), BUCHANAN, 1967, and CREW, 1969 (both discussing Officer's results), MOHRING, 1970 (case with increasing return to scale), PRESSMAN, 1970 (general mathematical formulation), PELTZMAN, 1971 (empirical look at pricing policies), NGUYEN, 1976 (possibility of storage), SORENSEN, 1976 (game-theoretic approach).

TURVEY, 1968b, relaxed Steiner's assumptions of constant marginal costs and of independency of demand curves in different periods. He suggests tariff experiments as the only way to find a tariff reasonable under realistic conditions.

Some other modifications of the classical model are treated in the following subsections.

2.3.2 The Regulated Utility

A few studies in the literature attempt to take into consideration the fact that electric utilities are regulated firms. The behaviour of such firms had first been examined by AVERCH & Johnson, 1962. Their main finding was that regulated firms often do not equate marginal rates of factor substitution to the ratio of factor costs, and thus do not operate at a welfare maximum.

BAILEY & White, 1974, BAILEY, 1972, BAUMOL & Klevorick, 1970, and ZAJAC, 1972, all discuss these results in the peak-load pricing context. The regulated case for which no negative net revenues are allowed is discussed by BAUMOL & Bradford, 1970. KLEVORICK, 1971, writes on the optimal fair rate of return to be set by the regulatory agency.

2.3.3 Stochastic Demand

Uncertainties in demand were first introduced by BROWN & Johnson, 1971. The welfare optimum they find has lower prices and higher supply quantities than in the riskless case; the net revenue is negative.

But VISSCHER , 1973, claims that these results of Brown and Johnson are not due to randomness, but to the assumption of an unrealistic rationing system.

CREW & Kleindorfer, 1978, show that without reliability constraints added, there are multiple optima to the welfare-maximization problem. Brown's & Johnson's results minimize the net revenue. After introducing the cost of rationing, the optimum features a high reliability level and maybe excessive investments.

Besides comparing monopoly pricing under welfare and profit maximization, MEYER, 1975, constructs a chance-constrained program, adding a loss-of-load probability constraint, and suggests introducing the measure of risk-efficiency. For every level of expected profits one would define the associated risk-efficient price-set as minimizing the variance of total profit.

Loss-of-load-probabilities as an instrument for pricing decisions are introduced by VARDI et al, 1977b. They find that by apportioning capital costs relative to curtailment probabilities, each period is actually charged according to the utilization it makes of the last kW of installed capacity.

A relatively simple method of incorporating uncertainties in load forecasts into a linear programming capacity planning algorithm is developed in SHERALI et al, 1981b. Again, the resulting dual variables represent marginal costs for periods of different demand.

2.3.4 Diverse Technology and a Neoclassical Approach

Results that differ from the traditional results in very important aspects are obtained when diverse technology is introduced to the peak load pricing problem. Some of this literature also employs models very similar to the ones often used in capacity planning. For instance, WENDERS' (1976) peak load pricing model is a special case of the model in SHERALI et al, 1981a.

Wenders considers a load duration curve broken into three blocks, and assumes that three equipment types are available. Employing breakeven analysis, he finds the optimal operation times for the three equipment types, and thus determines the pricing period served by each equipment type. An example he gives is shown in Figure 8. For this case, peak capacity only operates during the peak pricing period, intermediate capacity during both the intermediate and the peak period, and base capacity during all three periods.

Then, total operating costs are given by (where all symbols are as defined in previous sections)

$$OC = g_3 Q_3 + (d_1 + d_2)(Q_2 - Q_3)g_2 + d_1(Q_1 - Q_2)g_1 ,$$

while total capital costs are

$$CC = c_3 Q_3 + c_2(Q_2 - Q_3) + c_1(Q_1 - Q_2) .$$

The equation for welfare becomes

$$W = \sum_{i=1}^3 d_i \int_0^{Q_i} P_i dQ_i - CC - OC$$

Maximizing W with respect to output Q yields:

$$P_1 = g_1 + c_1/d_1$$

$$P_2 = [(d_1 + d_2)g_2 - d_1g_1 + c_2 - c_1] / d_2$$

$$P_3 = [g_3 - (d_1 + d_2)g_2 + c_3 - c_2] / d_3$$

Obviously in this result, off-peak periods do bear capital costs, too. This can intuitively be explained by the fact that off-peak users may not press against the total system capacity, but they do press against the capacity of base equipment.

Wenders then shows that an off-peak price will have no marginal capital cost component only in the following three cases: when only one kind of capacity is built; when supply and pricing periods coincide exactly; and when an off-peak pricing period is wholly contained within a supply period.

In all other cases, off-peak prices will contain a capital cost component.

Actually, WEINTRAUB, 1970, had been the first one to write on this so-called off-peak pricing. He, too, arrived at this result assuming different plant types being operated in merit order. CREW & Kleindorfer (1975 and 1976) extend these approaches and consider diverse technology, stochastic demand and rationing costs all in one model.

A neoclassical approach was first formulated by PANZAR, 1976. Among his propositions are the following: It is never optimal to operate at full capacity. Consumers in all periods should make contributions towards the cost of capital, provided the short-run returns to scale are decreasing. Under that condition it also holds that the larger the output in a period, the larger that period's contribution towards capital costs. Periods with larger optimal outputs have higher optimal prices. The difference between revenues and total costs at the optimum will equal the weighted sum of long-run scale elasticities in each period.

MARINO , 1978, extends Panzar's analysis and introduces bounds on input utilization. He obtains slightly different results; for instance he finds that optimal production at full capacity is possible.

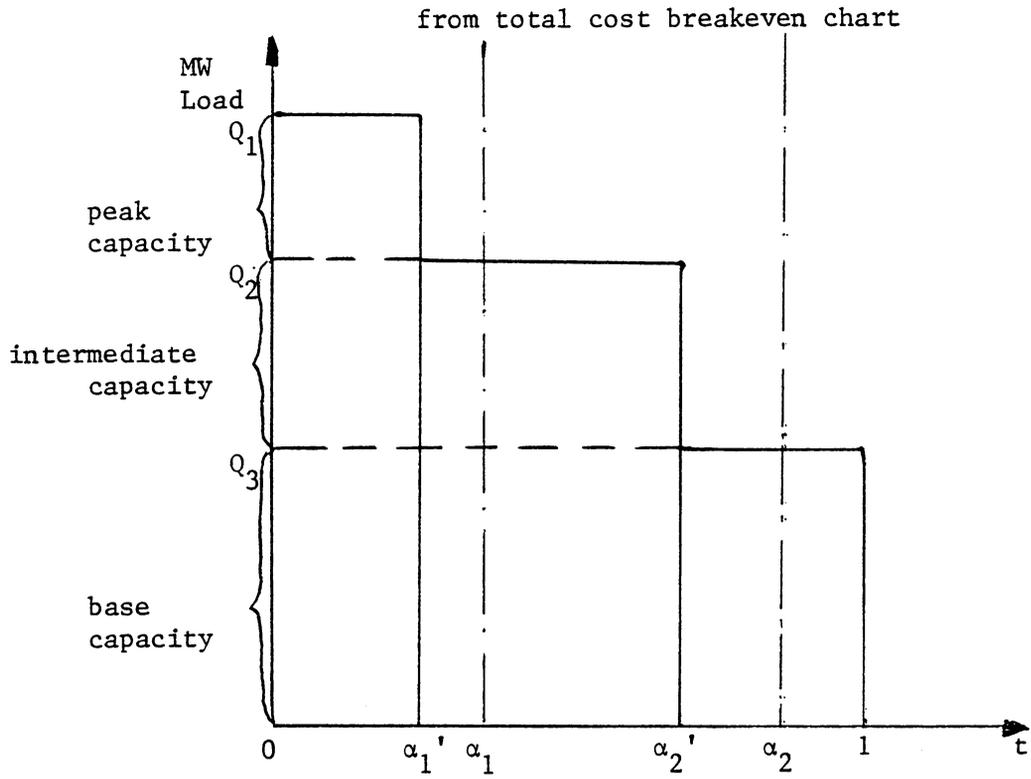


Figure 8: Example for Wenders Approach to Peak Load Pricing

2.3.5 Applications

The vast majority of all authors mentioned above agree on the fact that electricity should be priced at marginal cost. Differences in their opinions stem from differences in their modelling approaches. From the very simple welfare maximization model of Steiner's, with only two periods of equal length and only one type of plant, to the recent models of Crew's and Wender's which consider diverse technology and stochastic demand, the theory has come a long way. It has long found its way into public utility policy, beginning with the introduction of the 'tarif vert' in France in the 1950's, as described by NELSON, 1963, and others.

Not only from an economic welfare-maximization point of view, but also from the utility's load factor considerations peak load pricing makes sense. Electricity tariffs that deter demand at peak periods and encourage demand during off-peak hours can help to even out the load curve and thus can save capacity. In the U.S., marginal cost pricing is not as well established as in parts of Europe, but many utilities have at least some tariffs leaning towards it; and large-scale tariff-experiments are finally underway (see WENDERS & Taylor, 1976). Also SHEPHERD, 1966, and NELSON, 1963, give overviews of applications of marginal cost pricing.

2.4 THE SOLAR ENERGY - ELECTRIC UTILITY INTERFACE

One of the books advocating conservation and solar energy as solutions of the energy problem which received much public attention was 'Energy Futures', by STOBAUGH & Yergin, 1979. In its chapters on conservation and solar energy the respective authors, Yergin and Maidique, suggested that electric utilities were in an excellent position to help spread the use of these energy sources, for the following reasons:

- They have the potential for rapid market penetration, the customers trust them;
- they have access to low rates for borrowing money, and
- they can offer reliable service and maintenance.

Many different levels of utility involvement are being considered. FELDMAN & Wirtshafter, 1980, list and comment on them. Consumers, large and small producers, HVAC-contractors and utilities all express different interests and fears regarding this issue. A good overview on all the issues involved appears in BEZDEK & Cambel, 1981, and, in more detail, in 'The Role of Utility Companies in Solar Energy', 1978. Other discussion of this issue can be found in SMACKEY, 1978, ASHBURY & Mueller, 1971, ROSENBERG, 1977.

Investments in conservation are by far the most profitable energy investments today (see, e.g., FELDMAN & Wirt-

shafter, STOBAUGH & Yergin, KREITH & West, 1980). Solar hot water heating systems are economical in many areas of the country, even more if tax-credits are given. Solar space heating, especially in the passive form, is also economical in parts of the country (FELDMAN & Wirtshafter, BEZDEK et al, 1979). As prices for fossil fuels and electricity keep rising, these devices will become economically more competitive. Thus the utilities will be affected, anyway, as people increase the use of solar systems in their homes.

Many methods to determine if investments in solar energy are economical for a homeowner have been developed. One of the most-used is the f-CHART method as described and compared to other methods in KREITH & West, for example. Recently, more sophisticated, but also much more accurate simulation methods have been developed and successfully applied. FELDMAN & Wirtshafter for instance base their whole book on a simulation model.

The energy output of solar systems is simulated based on historical data on cloudcover and other determinants of solar energy availability at specific locations. At the same time the local utility's load, also based on historical data, and the household's ability to consume the solar energy are simulated. These three simulations interacting pro-

duce as output how much solar energy is produced in the house at all times, how much of it is consumed, how much backup energy from the utility is needed at what times, and how much this backup energy costs the utility and the homeowner.

Results of such simulation models have to serve as one input for models trying to determine how much utilities should get involved in solar energy, be it via marketing of solar homeheating equipment or with thermal solar power plants.

Pilot plants of large solar thermal power stations, up to 10MW capacity, are operating or are under construction in many parts of the world. For a list see WEINGART, 1979. Some authors try to show that in the long run solar energy alone, converted to electricity or used to produce hydrogen as a fuel, can be enough to satisfy the energy needs of the whole world. (See for example RAMAKUMAR et al, 1975, RALPH, 1972, WEINGART.)

Large-scale production of electricity from intermittent sources such as the sun's rays, the wind or the tides may cause technical and economical problems for power systems. Some authors fear that due to their intermittent structure they may not have any capacity value for the utility at all,

e.g. see BAE & Devine, 1978. More recently some authors showed how to quite accurately compute the capacity credit for production from intermittent sources, most often for wind machines and tidal power plants. Their results are that they do save the utilities conventional capacity, the amount being dependent on the probability distribution of the occurrence of wind or sunshine in relation to the utility's load curve, the structure and reliability of the company's other plants, and other factors (see SORENSEN, 1978, ANDREWS, 1976, PESCHON et al, 1978, KAHN, 1979, HASLETT & Diesendorf, 1981). These authors use mathematical modelling and simulation techniques.

Determining the capacity credit of solar or wind energy equipment does only a first step towards incorporating non-dispatchable technologies (NDTs) as decision variables into capacity expansion problems. Up to today, most researchers tried to determine the impact of NDTs on power systems by using simulation to find the 'negative load' due to the NDTs, i.e. the load that is saved due to NDTs at all times. However, this approach is not very appropriate to incorporate NDTs into expansion problems because for each penetration level of NDTs, a separate computer simulation has to be run; this entails a prohibitively large computing effort.

CARAMANIS et al, 1982, go a step further than examining the capacity credit of non-dispatchable technologies. They introduce such technologies as decision variables in capacity expansion models. In doing so, they have the same objective as this thesis. Their approach is, however, different from the one taken here. CARAMANIS et al employ a probabilistic approach to derive the load duration curve net of NDT generation. Although their method uses an approximating transformation of the involved joint probability functions, its accuracy has been tested and turned out very good. When a load duration curve net of NDT generation is generated for each capacity level of NDTs examined, the NDTs can be accommodated fairly easily in the context of a capacity expansion program. Thus, CARAMANIS' et al work is an approach parallel to the one used in this thesis.

Mathematical programming is also applied on the optimization of solar and wind energy systems themselves, i.e. on design issues. For instance, programs have been written to find the optimal tilt of solar collectors or the optimal wing design for wind machines. For examples, see SALIEVA, 1976, and DEVINE et al, 1978. DEVINE et al also give a list of what kind of optimization models have been applied on various alternative energy technologies.

Chapter III

CAPACITY PLANNING WITH SOLAR ENERGY

3.1 APPROACH TO THE PROBLEM

In this chapter, a capacity planning or plant mix problem will be examined. That is, it is assumed that no capacity whatsoever is installed yet. Features of the demand and the supply side of this problem, involving load duration curves, different equipment types and associated operating and capital costs, have already been discussed in the first section of the literature review. Throughout the following development, continuity of all load duration curves will be assumed. Let load duration curves be denoted by symbols $f(\cdot)$. It is also assumed that their inverses $f^{-1}(\cdot)$ exist in all cases. Also through all following chapters, let a superscript * denote the optimal value (with regard to some mathematical program) of the variable it is associated with.

Note also that all through this thesis, the planning problems examined are simplified cases of the real problem, because outages and reserve margins are neglected, and the operating costs are assumed linear.

3.1.1 Problem without Solar Energy

Analogous to the problem described in the literature review, the optimal plant mix problem in its basic form without the option of solar energy takes the following form:

$$\begin{aligned} \min H(x) &= \sum_{i=1}^I c_i x_i + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} F(z) dz \\ \text{subject to } &\sum_{i=1}^I x_i \geq P \\ &x \geq 0, \end{aligned}$$

where $H(x)$ = total cost function

c_i = annualized capital costs for equipment type i
with $c_1 > c_2 > \dots > c_I$

x_i = capacity of equipment type i

I = number of equipment types

g_i = operating costs for equipment type i
with $g_1 < g_2 < \dots < g_I$

$F(z)$ = inverse annual load duration curve

P = peak demand.

$Y_i = \sum_{j=1}^i x_j$, for $i=1, \dots, I$

As for instance shown in SOYSTER et al, 1981, this is a convex programming problem. Therefore, since a constraint qualification holds, the Kuhn Tucker conditions are both necessary and sufficient. Let

$$\nabla_i = \frac{\delta H}{\delta x_i} = c_i + \sum_{j=1}^{I-1} F(Y_j)(g_j - g_{j+1})$$

Then, the Kuhn-Tucker conditions assert that for some $\lambda \geq 0$,

$$\nabla_i = \lambda, \text{ if } x_i^* > 0$$

$$\nabla_i \geq \lambda, \text{ if } x_i^* = 0,$$

$$\text{where } \sum_{i=1}^I x_i^* = P, x_i^* \geq 0.$$

Thus, if β and γ , $1 \leq \beta \leq \gamma \leq I$ are two consecutive equipment types purchased (which implies $x_\beta > 0$ and $x_\gamma > 0$),

$\nabla_\beta = \nabla_\gamma$ implies:

$$c_\beta + \sum_{j=\beta}^{I-1} F(Y_j)(g_j - g_{j+1}) = c_\gamma + \sum_{j=\gamma}^{I-1} F(Y_j)(g_j - g_{j+1})$$

or $c_\beta + F(Y_\beta)g_\beta = c_\gamma + F(Y_\beta)g_\beta$, which implies

$$F(Y_\beta) = \frac{c_\beta - c_\gamma}{g_\gamma - g_\beta},$$

which is the formula used for calculation of operating times of different equipment in the breakeven analysis mentioned above. Thus, the capacities of the equipment types pur-

chased correspond precisely to breakeven point projections onto the load duration curve. The equipment types not purchased are always more expensive than some other equipment with respect to serving any given load.

3.1.2 Problem with Solar Energy

Now, consider the same problem for the case that solar energy equipment is available for installation. As described in the literature review, the main difference between solar and conventional equipment is that solar energy is not available at all times. Non-availability or partial availability during day time due to cloudcover, for instance, will be treated in Chapter 5. Storage devices that can smooth the availability of solar energy, will be treated in the same chapter. In this chapter, only the fact that output from solar equipment is restricted to the day time will be considered.

Due to the solar output being restricted to day time, the problem has to be broken up into two parts, namely, for day and night time. Capital costs are not affected. But operating costs and constraints have to be accounted for separately for the day time when solar energy is available, and for the night time when it is not available.

Thus, for the case that the option of investing in solar energy is available, the problem takes the following form:

$$\begin{aligned} \min \quad & \sum_{i=0}^I c_i x_i + \sum_{i=0}^I g_i \int_{X_{i-1}}^{X_i} \lambda F_d(z) dz + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} (1-\lambda) F_n(z) dz \\ \text{s.t.} \quad & \sum_{i=0}^I x_i \geq P_d \\ & \sum_{i=1}^I x_i \geq P_n \\ & x \geq 0 \end{aligned}$$

where x_0 = capacity of solar equipment

$$X_i = \sum_{j=0}^i x_j$$

$$Y_i = \sum_{j=1}^i x_j$$

c_0, g_0 = annualized capital and operating cost for solar equipment, with $g_0 < g_1$

λ = fraction of time that daylight is available, i.e. that the solar equipment can be used

P_d = peak day time demand

P_n = peak night time demand

P = overall peak demand

$F_d(z)$ = inverse annual load duration curve for day

$$= \begin{cases} f_d^{-1}(z) & \text{for } 0 \leq z \leq P_d \\ 0 & \text{for } P_d \leq z \leq P \end{cases}$$

$F_n(z)$ = inverse annual load duration curve for night

$$F_n(z) = \begin{cases} f_n^{-1}(z) & \text{for } 0 \leq z \leq P_n \\ 0 & \text{for } P_n \leq z \leq P \end{cases}$$

c_i, g_i, x_i are as defined before.

To facilitate the forthcoming analysis, the inverse curves are assumed smooth at the respective peak loads.

Then, their derivatives $F_d'(z)$ and $F_n'(z)$ can be assumed to exist for all $z, 0 \leq z \leq P$. The smoothness assumption can be justified since in case they are not smooth at the peak loads, as close a smooth approximation as desired can be found for any computational purposes. We emphasize that this assumption is merely a convenience in the development which avoids statements of differentiability and one-sided derivatives existing almost everywhere.

Also note that load duration curves and their inverses are generally assumed to be monotone decreasing.

By redefining the integrand and the limits of integration in the term representing day operating costs, the two operating cost terms for the day and for the night may be combined to give the following program.

$$\begin{aligned} \min \quad & \sum_{i=0}^I c_i x_i + g_0 \lambda \int_0^{x_0} F_d(z) dz + \sum_{i=1}^I g_{iY_{i-1}} \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda) F_n(z)] dz \\ \text{s.t.} \quad & \sum_{i=0}^I x_i \geq P_d \\ & \sum_{i=1}^I x_i \geq P_n \end{aligned}$$

$$x \geq 0,$$

which in turn can be written as

$$\begin{aligned} \min_{0 \leq x_0 \leq P_d} h(x_0) \quad , \text{ where} \\ h(x_0) = c_0 x_0 + \lambda g_0 \int_0^{x_0} F_d(z) dz \\ + \min \sum_{i=1}^I c_i x_i + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda) F_n(z)] dz \\ \text{s.t.} \quad \sum_{i=1}^I x_i \geq \max \{P_d - x_0, P_n\} \\ x \geq 0 \end{aligned} \quad (1)$$

If x_0 is held constant, the problem in Equation Set 1 is nothing else but a plant mix optimization problem involving only conventional equipment, which can be easily solved with breakeven analysis as described above. Thus, for each x_0 , one can find the minimal total cost in a straightforward way. See Figure 9 for a graphical description of the method. The task hence reduces to searching over x_0 for an optimal solution to the problem. In other words, we need to optimize the total cost curve $h(\cdot)$ with respect to the variable x_0 .

If this curve is convex, efficient search algorithms can be applied to find the optimal level of solar and conven-

tional capacity. Convexity will be proven in the following section; then, a search algorithm will be devised.

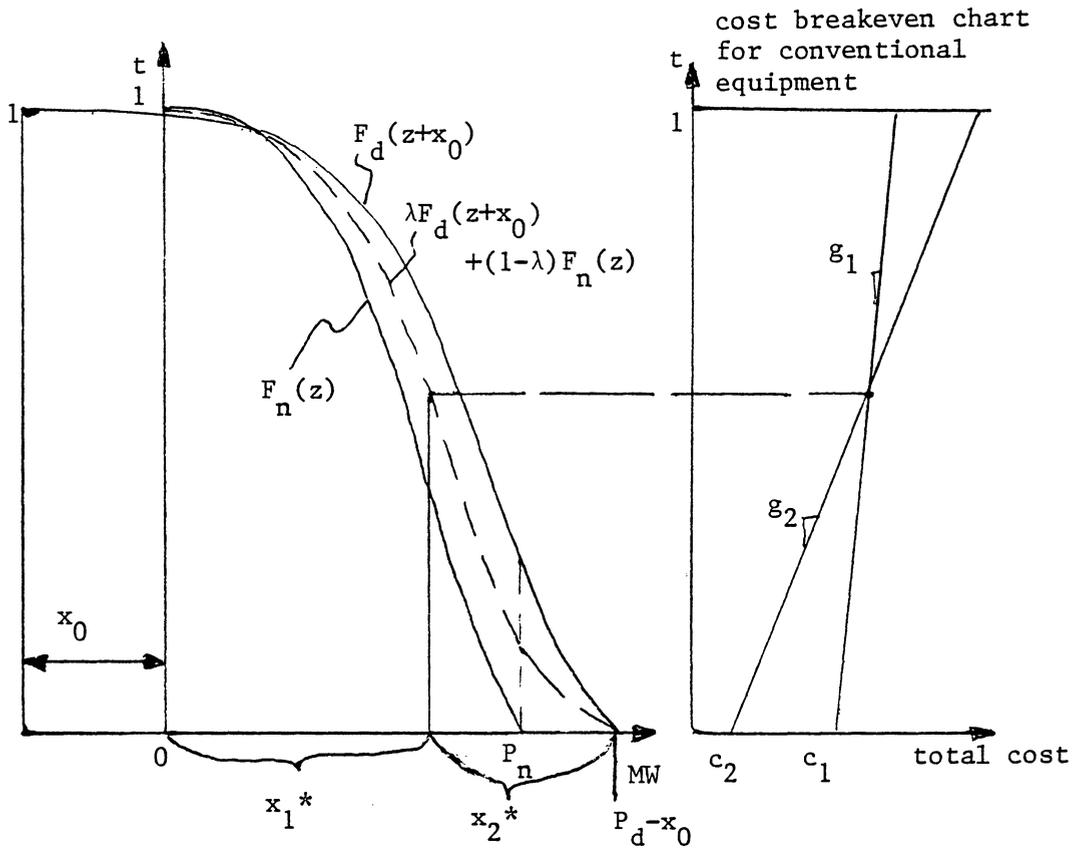


Figure 9: Approach for Solar Investment Analysis

3.2 CONVEXITY

In this section it will be shown that the first (right hand) derivative of the total cost function with respect to x_0 is an increasing function of x_0 . This in turn implies that the function is convex. For this analysis each variable will have to be treated as a function of x_0 , with the optimal capacities for conventional equipment types being determined via breakeven analysis for each level of solar investment x_0 .

Thus, for the remainder of this chapter, all conventional capacities x_i , $i=1, \dots, I$, shall denote optimal values as functions of x_0 as given by the program in Equation Set 1. In other words, in what follows, x_i is actually $x_i^*(x_0)$, $i=1, \dots, I$. Hence, the total cost function can be written as follows:

$$h(x_0) = c_0 x_0 + \lambda g_0 \int_0^{x_0} F_d(z) dz + \sum_{i=1}^I c_i x_i \\ + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda) F_n(z)] dz$$

Before proceeding with the convexity analysis, continuity and differentiability have to be investigated.

3.2.1 Continuity and Differentiability

Essentially, in order to investigate continuity and differentiability of $h(x_0)$, it is sufficient to study whether Y_i and X_i , for all i , are continuous and differentiable functions of x_0 . But $X_i = Y_i + x_0$, for $i=1, \dots, I$. Thus it is sufficient to consider whether Y_i is continuous and differentiable in x_0 .

3.2.1.1 Continuity

Let us establish continuity of $Y_i = Y_i^*(x_0)$. From breakeven analysis, the quantity Y_i (for a given x_0) is determined with respect to some fixed break point duration α_i according to

$$\alpha_i = \lambda F_d(Y_i + x_0) + (1-\lambda)F_n(Y_i)$$

Perturbing x_0 slightly by an amount Δx_0

(so that $P_d \geq x_0 + \Delta x_0 \geq 0$), the same has to hold for the duration, α_i , but for the resulting perturbed Y_i^+ . Hence,

$$\alpha_i = \lambda F_d(Y_i^+ + x_0 + \Delta x_0) + (1-\lambda)F_n(Y_i^+)$$

Combining both expressions into one equation yields

$$\lambda[F_d(Y_i^+ + x_0 + \Delta x_0) - F_d(Y_i + x_0)] + (1-\lambda)[F_n(Y_i^+) - F_n(Y_i)] = 0$$

Applying the Mean Value Theorem yields, for some convex combinations \bar{Y}_i between $(Y_i^+ + \Delta x_0)$ and Y_i , and \bar{Y}_i between Y_i^+ and Y_i :

$$\lambda[(Y_i^+ - Y_i + \Delta x_0)F_d'(\bar{Y}_i + x_0)] + (1-\lambda)(Y_i^+ - Y_i)F_n'(\bar{Y}_i)] = 0 \quad (2)$$

Rearranging terms, we obtain

$$(Y_i^+ - Y_i^-) [\lambda F'_d(\bar{Y}_i + x_0) + (1-\lambda) F'_n(\bar{Y}_i)] = -\lambda \Delta x_0 F'_d(\bar{Y}_i + x_0) \quad (3)$$

Now, as $\Delta x_0 \rightarrow 0$, since $F'_d(\cdot)$ is bounded, the right hand side of Equation 3 approaches zero, and hence so must the left hand side. First, suppose $i < I$. Then, the term in square brackets on the left hand side of Equation 3 is negative (load duration curves are monotone decreasing), even in the limit as $\Delta x_0 \rightarrow 0$. Therefore, $Y_i^+ \rightarrow Y_i^-$ as $\Delta x_0 \rightarrow 0$, which means that Y_i is continuous in x_0 . Further, since $x_0 + Y_I = P$, for all x_0 , it follows that Y_I is also continuous in x_0 .

Finally, note that $x_i = Y_i - Y_{i-1}$, for all $i=1, \dots, I$ (here, define $Y_0 = 0$). Thus, continuity of all x_i as functions of x_0 is also established.

3.2.1.2 Differentiability

By definition of the derivative

$$\lim_{\Delta x_0 \rightarrow 0} \frac{Y_i^+ - Y_i^-}{\Delta x_0} = \frac{\partial Y_i}{\partial x_0}$$

Noting that at $x_0=0$ and $x_0=P$ we restrict $x_0 \geq 0$ and $x_0 \leq 0$, respectively, the above definition represents one-sided derivatives in these two cases.

Again, from the development leading to Equation 2, we obtain (wherever the ratio below exists)

$$\lim_{\Delta x_0 \rightarrow 0} \frac{Y_i^+ - Y_i}{\Delta x_0} = \frac{-\lambda F'_d(\bar{Y}_i + x_0)}{\lambda F'_d(\bar{Y}_i + x_0) + (1-\lambda)F'_n(\bar{Y}_i)}$$

As $\Delta x_0 \rightarrow 0$, the difference between Y_i and Y_i^+ goes to zero, due to continuity of all Y_i with x_0 , thus forcing the mean values \bar{Y}_i and \bar{Y}_i to go to Y_i . Thus, wherever the ratio below exists,

$$\frac{\partial Y_i}{\partial x_0} = \frac{-\lambda F'_d(Y_i + x_0)}{\lambda F'_d(Y_i + x_0) + (1-\lambda)F'_n(Y_i)} = \frac{-\lambda F'_d(X_i)}{\lambda F'_d(X_i) + (1-\lambda)F'_n(Y_i)}$$

This derivation exists except where the denominator becomes zero. This happens when $x_0 = P_d - P_n$, and $Y_i = P_n$. In particular, this situation will occur for $x_0 = P_d - P_n$ and $i=I$. At that point, the solar equipment provides exactly the load by which the day peak exceeds the night peak. The conventional equipment has to meet equal loads during both day and night, so that

$$Y_I = \sum_{i=1}^I x_i = P_n. \quad \text{Thus,}$$

$$Y'_I(x_0) = \frac{-\lambda F'_d(X_I)}{\lambda F'_d(X_I) + (1-\lambda)F'_n(Y_I)} = \frac{-\lambda F'_d(X_I)}{\lambda F'_d(X_I) + 0} = -1, \quad P_d - x_0 > P_n$$

(continued)

$$Y'_I(x_0) = \begin{cases} 0 / (1-\lambda)E'_n(Y_I) = 0 & , P_n > P_d - x_0 \\ \text{undefined} & , x_0 = P_d - P_n \end{cases}$$

Thus except at $x_0 = P_d - P_n$ and $i=I$, all Y_i are differentiable functions of x_0 ; expressions for Y'_i have been developed. Together with the assumed differentiability of the inverse load duration curves this also establishes differentiability of the function $h(x_0)$, except at $x_0 = P_d - P_n$.

Therefore for establishing convexity properties, the function $h(x_0)$ will be broken up into two parts, for $x_0 > P_d - P_n$ and $x_0 < P_d - P_n$, respectively. Note that the only term causing non-differentiability at $x_0 = P_d - P_n$ is Y'_I . So the following proof of convexity will be identical for both sections of the function, as long as Y'_I is appropriately replaced by a variable δ such that

$$\delta = \begin{cases} -1 & \text{for the section } 0 \leq x_0 < P_d - P_n \\ 0 & \text{for the section } P_d - P_n < x_0 \leq P \end{cases} \quad (4)$$

3.2.2 Proof of Convexity of $h(x_0)$

For the sake of convenience, the function $h(x_0)$ is stated again:

$$h(x_0) = c_0 x_0 + \lambda g_0 \int_0^{x_0} F_d(z) dz + \sum_{i=1}^I c_i x_i + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda)E_n(z)] dz \quad (5)$$

THEOREM 1

The total cost function $h(x_0)$, stated in Equation 5, is a convex function.

Proof

Taking the derivative with respect to x_0 , for $x_0 \neq P_d - P_n$, yields:

$$\begin{aligned} h'(x_0) = & c_0 + \lambda g_0 F_d(x_0) + \sum_{i=1}^I c_i x_i'(x_0) + \sum_{i=1}^I g_i \lambda F_d(X_i) \\ & - \sum_{i=1}^I g_i \lambda F_d(X_{i-1}) + \sum_{i=1}^I g_i \lambda F_d(X_i) Y_i'(x_0) - \sum_{i=1}^I g_i \lambda F_d(X_{i-1}) Y_{i-1}'(x_0) \\ & + \sum_{i=1}^I g_i (1-\lambda) [F_n(Y_i) Y_i'(x_0) - F_n(Y_{i-1}) Y_{i-1}'(x_0)] \end{aligned}$$

This expression can be transformed in the following way.

From the definition of Y we get

$$x_i'(x_0) = Y_i'(x_0) - Y_{i-1}'(x_0).$$

Therefore,

$$\begin{aligned} \sum_{i=1}^I c_i x_i'(x_0) &= \sum_{i=1}^I c_i [Y_i'(x_0) - Y_{i-1}'(x_0)] \\ &= \sum_{i=1}^{I-1} Y_i'(x_0) (c_i - c_{i+1}) + c_I \delta \end{aligned} \quad (6)$$

where δ is defined in Equation 4.

From breakeven analysis one gets that for the duration,

α_i , representing a breakpoint between two equipment types, i and $i+1$,

$$\alpha_i = \frac{c_i - c_{i+1}}{g_{i+1} - g_i} = \lambda F_d(X_i) + (1-\lambda)F_n(Y_i)$$

This implies that

$$(c_i - c_{i+1}) = (g_i - g_{i+1})[-\lambda E_d(X_i) - (1-\lambda)E_n(Y_i)]$$

Substituting this into Equation 6 yields

$$\sum_{i=1}^I c_i x'_i(x_0) = \sum_{i=1}^{I-1} Y'_i(x_0)(g_i - g_{i+1})[-\lambda E_d(X_i) - (1-\lambda)E_n(Y_i)] + c_I \delta$$

Using this expression and rearranging the sums representing operating costs, the following expression for $h'(x_0)$ is obtained:

$$\begin{aligned} h'(x_0) = & c_0 + c_I \delta + \lambda F_d(x_0)(g_0 - g_1) \\ & + \sum_{i=1}^{I-1} Y'_i(x_0)(g_i - g_{i+1})[-\lambda E_d(X_i) - (1-\lambda)E_n(Y_i)] \\ & + \sum_{i=1}^{I-1} \lambda F_d(X_i) Y'_i(x_0)(g_i - g_{i+1}) + \sum_{i=1}^{I-1} (g_i - g_{i+1}) \lambda F_d(X_i) \\ & + \sum_{i=1}^{I-1} (g_i - g_{i+1}) Y'_i(x_0) (1-\lambda) E_n(Y_i) \end{aligned}$$

Cancelling terms appropriately yields

$$h'(x_0) = c_0 + c_I \delta + \sum_{i=0}^{I-1} (g_i - g_{i+1}) \lambda E_d(X_i), \quad x_0 \neq P_d - P_n \quad (7)$$

Note that for both sections of $h(x_0)$, $x_0 < P_d - P_n$ and $x_0 > P_d - P_n$, $c_I \delta$ is a constant (either $-c_I$ or 0). Thus, its derivative is 0 in both sections. Due to this fact, the second derivative of $h(x_0)$ exists and takes the same form for both sections:

$$h''(x_0) = \lambda F'_d(x_0)(g_0 - g_1) + \sum_{i=1}^{I-1} (g_i - g_{i+1}) \lambda F'_d(X_i) X'_i(x_0), \quad x_0 \neq P_d - P_n \quad (8)$$

Note that $(g_i - g_{i+1}) < 0$, for all i , due to the assumption that $g_0 < g_1 < g_2 < \dots < g_I$. Also, since the inverse load duration curves are assumed to be monotone decreasing, both $F'_d(z)$ and $F'_n(z)$ are nonpositive for all z .

From the definition of Y_i and X_i :

$$X'_i(x_0) = 1 + Y'_i(x_0) = 1 - \frac{\lambda F'_d(X_i)}{\lambda F'_d(X_i) + (1-\lambda) F'_n(Y_i)}$$

$$= \frac{(1-\lambda) F'_n(Y_i)}{\lambda F'_d(X_i) + (1-\lambda) F'_n(Y_i)}$$

Using the facts that all $(g_i - g_{i+1})$ are negative, that both $F'(\cdot)$ are nonpositive, and that all $X'_i(x_0)$ are positive, one obtains from Equation 8 that $h''(x_0) \geq 0$. Therefore $h(x_0)$ is convex, for $x_0 < P_d - P_n$ and for $x_0 > P_d - P_n$. To complete the proof, it must be shown that the right hand derivative is increasing in the neighborhood of $x_0 = P_d - P_n$ as well. This follows from Equations 7 and 8 by noting that $c_I \delta$ is (-1) for $x_0 < P_d - P_n$ and is 0 for $x_0 > P_d - P_n$. Hence, the proof is complete.

This important result can now be used in devising an efficient search algorithm that finds the optimal capacities for all equipment types.

3.3 ALGORITHM

For each value of x_0 , the optimal capacities of all conventional equipment types can be determined relatively easily with breakeven analysis. Thus, the total cost function can be treated as a function of one variable only, x_0 . Since it is a convex function, any local minimum is a global minimum. To find a minimum, the following search technique can be used. It can be viewed as a one-dimensional application of Newton's Method or the optimal gradient method (where the

optimal step length is found analyzing a Taylor's expansion of the objective function's gradient). (See, e.g., SIMMONS, 1975, for a description of these methods.)

In searching for the minimum, first it has to be determined whether the optimal x_0 is less than or greater than $(P_d - P_n)$, since at that point the derivative of $h(x_0)$ changes discontinuously. As can be seen from the expression for $h'(x_0)$ (Equation 7), there is a change of magnitude $+c_I$ in $h'(x_0)$ when going from $x_0 < P_d - P_n$ to $x_0 > P_d - P_n$. By examining the right hand derivatives at $x_0 = 0$ and $x_0 = (P_d - P_n)$ it can easily be determined what range the optimal x_0 lies in. These ideas are formalized below.

Capacity Planning Algorithm (CPA)

Step 1: If $h'(x_0=0) > 0$, then it is not economical to invest in solar energy at all.

If $h'(x_0=0) < 0$ and $h'(x_0=P_d - P_n - \epsilon) > 0$, then the optimal x_0 has to lie between 0 and $(P_d - P_n)$. Here, $h'(x_0=P_d - P_n +/\!-\epsilon)$ denotes the derivative slightly ($\epsilon > 0$) to the right/left of the non-differentiable point $P_d - P_n$.

If $h'(x_0=P_d - P_n - \epsilon) < 0$ and $h'(x_0=P_d - P_n + \epsilon) > 0$, then $x_0 = P_d - P_n$ is the optimal solution.

If $h'(x_0 = P_d - P_n + \epsilon) < 0$, then the optimal x_0 has to lie between $(P_d - P_n)$ and P_d . It will obviously never be optimal to install an x_0 that is greater than P_d .

Step 2: If it has been determined that the optimal x_0 lies either between 0 and $(P_d - P_n)$ or between $(P_d - P_n)$ and P_d , the following search technique is applied.

The first and second derivatives of $h(x_0)$ at the current x_0 are calculated. Then, approximating the function $h'(x_0)$ by a straight line with slope $h''(x_0)$, the root of $h'(x_0)$ is estimated. The root is then used as the current x_0 , and the procedure is repeated until $|h'(x_0)|$ is less than some specified accuracy $\epsilon > 0$.

This is a one-dimensional application of Newton's Method or the Optimal Gradient Method, for which convergence in a finite number of iterations is well-established, as long as one starts with a solution close enough to the optimum (see SIMMONS, 1975, Chapter 4).

To ensure convergence, therefore, and still maintaining algorithmic efficiency, the following procedure may be adopted. Perform m iterations of the algorithm described above. If the optimum is found within these m iterations, stop; otherwise do m iterations of the golden section search

(as also described in SIMMONS), starting with the smallest known interval of uncertainty. Repeat these two steps until the termination criterion is satisfied within Newton's algorithm.

3.4 ILLUSTRATIVE EXAMPLE AND COMPUTATIONAL RESULTS

To further illustrate the approach developed, Algorithm CPA is applied on the following example. To keep the example simple, a system with only three conventional equipment types, plus solar energy equipment, is assumed. The shape of the load duration curves is also chosen as simple as possible; triangular load curves are assumed.

Let the (annualized) capital and operating costs of all available equipment types be:

$$c_0 = 14, g_0 = 0; \quad c_1 = 10, g_1 = 10;$$

$$c_2 = 8, g_2 = 20; \quad c_3 = 6, g_3 = 40.$$

These cost values result in breakeven values of

$$\alpha_1 = (c_1 - c_2) / (g_2 - g_1) = 0.2, \text{ and}$$

$$\alpha_2 = (c_2 - c_3) / (g_3 - g_2) = 0.1.$$

The following functions are assumed as inverse load duration curves:

$$F_d(z+x_0) = 1 - (z+x_0)/20, \text{ and } F_n(z) = 1 - z/10.$$

For convenience, the expressions for $h'(x_0)$ and $h''(x_0)$ which will be needed during the algorithm, are restated below:

$$h'(x_0) = c_0 + c_I \delta + \sum_{i=0}^{I-1} (g_i - g_{i+1}) \lambda F_d'(X_i) \quad (7)$$

$$h''(x_0) = \lambda F_d''(x_0)(g_0 - g_1) + \sum_{i=1}^{I-1} (g_i - g_{i+1}) \lambda F_d''(X_i) X_i' \quad (8)$$

where

$$X_i' = [(1-\lambda)F_n'(Y_i)] / [\lambda F_d'(X_i) + (1-\lambda)F_n'(Y_i)] = 2/3.$$

This value for X' holds for all $X' < P_n$, due to the constant derivatives of the triangular load duration curves.

Then, Algorithm CPA proceeds as follows: (Note that for each iteration, the capacities of conventional equipment types are computed via breakeven analysis, as shown in Figure 10)

Step 1:

$$h'(x_0=0) = 14 - 6 - .5[10*1. + 10*0.4 + 20*0.2] = -1$$

Therefore the optimal x_0 will be positive.

$$h'(x_0=P=10) = 14-6 - .5[10*.5 + 10*.075 + 20*.0375] = 4.75$$

Therefore the optimal x_0 will be less than $P=10$.

Step 2:

Start with $x_0=0$, from above $h'(0) = -1$.

$$h''(0) = 10/(2*20) + [2/(2*3)][10/20 + 20/20] = 0.75$$

$$\text{Next } x_0 = h''/h' = 4/3$$

Second iteration:

$$h'(4/3) = 14-6 - .5[10*.933 + 10*.4 + 20*.2] = -2/3$$

Thus the optimal x_0 will lie between $4/3$ and 10 .

We quit the example here since the iterations shown are sufficient to demonstrate how Algorithm CPA works. As a matter of interest, the optimum for the example is

$$x_0=3.23; x_1=9.58; x_2=3.16; x_3=4.01.$$

This optimum is obtained within five iterations.

A Fortran computer program implementing Algorithm CPA was written and applied on several simple example problems. The example problems approximated day and night inverse load duration curves by fifth order polynomials. As in the example problem given above, three conventional equipment types are assumed to be available. For the example problems examined, only three to five iterations were needed to arrive at the optimum within a tolerance of 0.05; about 0.3 seconds of CPU time were used for each example.

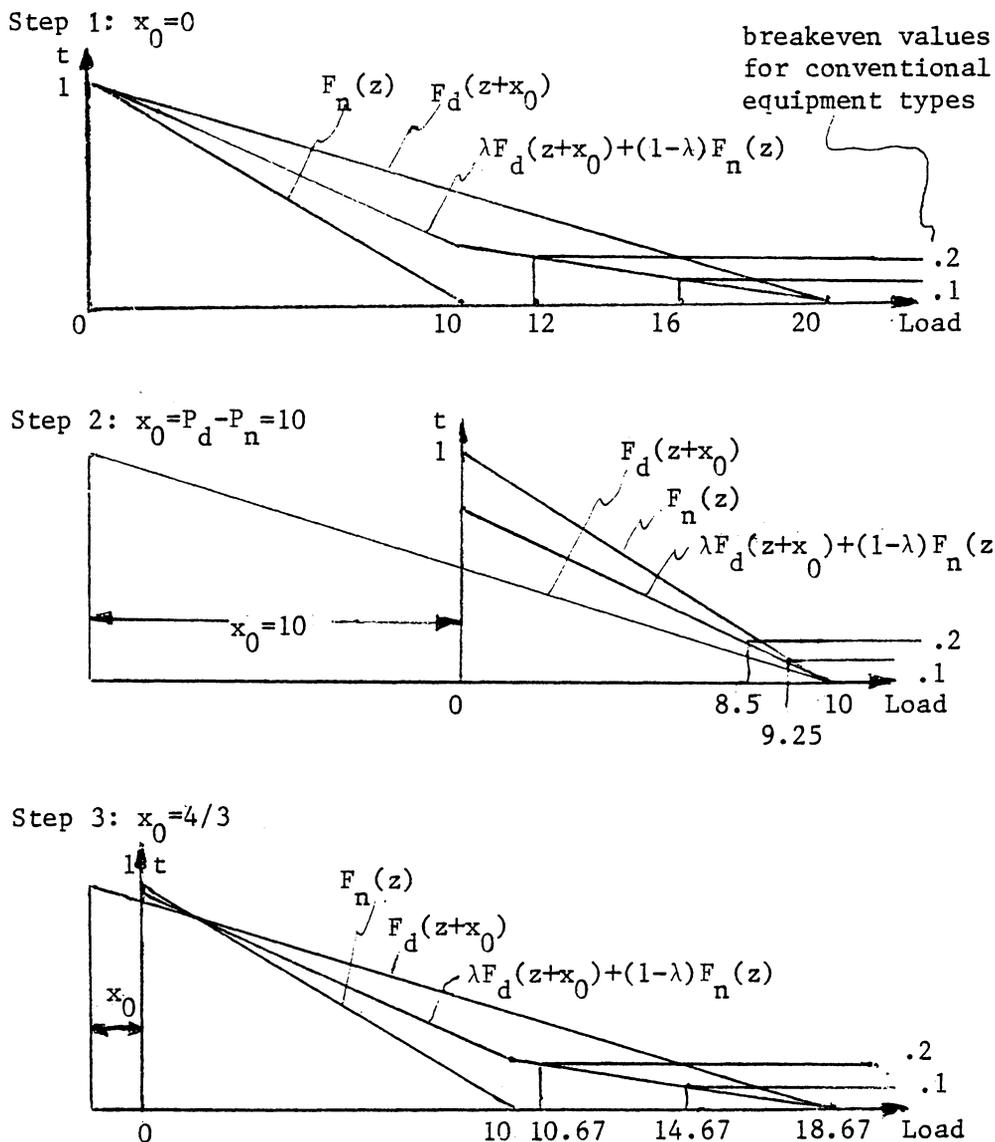


Figure 10: Breakeven Analyses for Example Problem

Chapter IV

CAPACITY EXPANSION PLANNING WITH SOLAR ENERGY

4.1 FORMULATION

The plant mix optimization described in the last chapter is unrealistic in one important aspect. It assumes that no capacity whatsoever is installed yet. Thus, solving this problem can give one an idea of what the ideal plant mix would look like. But with a bulk of the needed capacity already installed, the optimal expansion plan will almost always deviate from this ideal solution. However, this type of problem does form an important subproblem in solving more complicated problems as shown in section 4.7.

Expansion planning algorithms have to take into account that certain capacities of most equipment types, denoted by b_i , $i=0,1,\dots,I$, are part of the system already. (Again, the index 0 denotes solar equipment.) When these 'old' units are dispatched, respective operating costs of g_i , $i=0,1,\dots,I$, will be incurred. But no capital cost is assigned to these old units. Thus, instead of the plant mix decision variables x_i denoting total capacity of equipment i to be installed, the expansion problem has two types of decision variables.

The variables z_i , $i=1, \dots, I$, denote the amounts of old capacity of type i to be utilized. It can be more economical to install a new, efficient unit instead of keeping the old, inefficient one, when the new unit's capital cost is offset by the saving in fuel costs. Thus, z_i need not equal b_i , and is generally restricted to satisfy $0 \leq z_i \leq b_i$, $i=1, \dots, I$.

The variables $x_i \geq 0$, $i=0, 1, \dots, I$, denote the capacities of new equipment of type i to be installed. They have both capital costs c_i and operating costs g_i associated with it.

Then, the sums $z_i + x_i$, for all i , represent the capacity of type i available for dispatching. Similar to the definitions used in previous sections, let

$$Y_i = \sum_{j=1}^i (z_j + x_j), \quad i=1, \dots, I$$

$$X_i = \sum_{j=0}^i (z_j + x_j), \quad i=1, \dots, I$$

Note that the old solar capacity b_0 , if it exists, will always be kept due to its near-zero operating costs. Since $g_0 < g_i$, for $i=1, \dots, I$, it follows that $z_0 = b_0$. Else, $z_0 < b_0$ would imply that no conventional equipment at all would be economical to serve the day time load, which would render the problem uninteresting. Therefore no variable z_0 is introduced.

The program for finding the optimal one-year expansion plan takes the following general form.

Expansion Program EP

$$\begin{aligned} \min h = & c_0 x_0 + \lambda g_0 \int_0^{x_0 + b_0} F_d(z) dz + \sum_{i=1}^I c_i x_i \\ & + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z + x_0 + b_0) + (1-\lambda) F_n(z)] dz \\ \text{subject to: } & \sum_{i=1}^I (z_i + x_i) \geq P_d - x_0 - b_0 \\ & \sum_{i=1}^I (z_i + x_i) \geq P_n \\ & x \geq 0, \quad 0 \leq z \leq b. \end{aligned}$$

Note that for a fixed value of x_0 , a standard capacity expansion problem results, in which old plus new capacities of the conventional equipment types can be used at all times of the day. For this, the integrand in the objective function gives the equivalent inverse load duration curve. This problem can be solved by methods as described in the literature review. An efficient solution method, due to SHERALI (1982a), is referred to in the following section and detailed in Appendix A. This method will be used as part of the algorithm for solving the expansion problem with an option to invest in solar energy.

4.2 SOLVING THE EXPANSION PROBLEM WITHOUT SOLAR ENERGY

For a fixed x_0 , the problem with only the conventional capacities as decision variables takes the following form:

$$\min \sum_{i=1}^I c_i x_i + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda)F_n(z)] dz \quad (9)$$

$$\text{s.t. } \sum_{i=1}^I (z_i + x_i) \geq \max \{P_d - x_0 - b_0, P_n\} \quad (10)$$

$$x \geq 0, \quad 0 \leq z \leq b$$

When x_0 is fixed, this can be written as follows:

Problem Without Solar Energy (PWS)

$$\min \sum_{i=1}^I c_i x_i + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} F(z) dz$$

$$\text{s.t. } \sum_{i=1}^I (z_i + x_i) \geq P$$

$$x \geq 0, \quad 0 \leq z \leq b$$

where $F(z)$ denotes the aggregate inverse load duration curve (the integrand in Equation 9), and P denotes peak load (right hand side in Equation 10), which is modified appropriately by the solar capacity.

This problem can be solved with the algorithm due to SHERALI, 1982a, given in Appendix A. It consists of two phases. During a pre-optimization phase (I), optimal capac-

ities of old equipment (phase Ia), optimal capacities of a subset of new equipment (phase Ib) and a good quality starting solution (phase Ic) are determined. With part of the solution restricted to the optimal value in phase I, a problem of minimizing a convex, differentiable function, subject to a single generalized upper bounding constraint results. In phase II, a modified steepest descent feasible directions algorithm is used to solve this problem. Appendix A gives the details of this algorithm. For our purpose, we will henceforth assume that for a fixed $x_0 \in \{0, P_d\}$, the resulting capacity expansion problem is readily solvable.

4.3 SOLUTION METHOD WITH SOLAR ENERGY VIA DECOMPOSITION

The method presented here resembles the method known as Benders' Decomposition in important aspects (see BENDERS, 1962; for a description of decomposition techniques, see, e.g., LASDON, 1970). The approach to be presented can be regarded as an alternative to the Generalized Benders Decomposition technique due to GEOFFRION (1972).

4.3.1 Introduction and Outline

The nonlinear expansion planning program developed in the formulation section can be approximated by a linear program, as shown in Figure 11, by discretizing the load duration curves. The discretization results in several horizontal load segments, henceforth simply referred to as blocks. Let w_p represent the width of the load block p . To formulate the linear programming approximation, we need to introduce allocation variables y_{ip} , which denote the amount of load in block p served by equipment type i . Also, instead of just one constraint in the nonlinear case, requiring sufficient capacity to meet peak demand, two types of constraints have to be used. Demand constraints for all load blocks guarantee that enough capacity (of all types) is installed to meet each block's demand. Capacity constraints for all equipment types make sure that no more capacity is allocated via the y variables than is installed. The installed capacity is denoted by the sum of old and new capacity, $(x_i + b_i)$, for all i . Here it is not necessary to introduce z variables representing the old capacity actually utilized. If it is optimal to utilize less than b_i of some equipment i , the sum of the optimal allocation variables y_{ip} over all blocks p will be less than b_i . An appropriate capital cost c_i is assigned

to equipment i capacities whose total allocation $\sum_{p=1}^P y_{ip}$ exceeds the available supply b_i . This excess allocation is of course equal to x_i , the new capacity purchased.

Another major difference to the nonlinear formulation is that merit order loading is explicitly assumed in the nonlinear program. In the linear program to be given below, merit order loading will be a result of the optimization. That is, the optimal y variables will take such a form that equipments are loaded in merit order.

Now, without loss of generality, assume that the discretization is such that the widths w_p of all blocks p are equal to some standardized width Δ . This simplifying assumption will facilitate the forthcoming derivation in which we will obtain limiting forms of expressions as $\Delta \rightarrow 0$, i.e. as the discretization becomes very fine.

The use of the linear programming approximation and the simplifying assumption on the widths Δ are necessary ingredients of the decomposition technique to be developed in this chapter. The technique iterates between the nonlinear formulation, for which primal solutions are found, and a very finely discretized version with which the corresponding dual solutions are found. Then, the dual solutions are used to construct a Benders' Cut, and the next iteration is begun.

4.3.2 Formulation and Decomposition Technique

The discretized version of the expansion planning program takes the following form. (Compare with Figure 11)

Discretized Program DP

$$\begin{aligned}
 \min \quad & c_0 x_0 + \sum_{i=1}^I c_i x_i + \lambda \sum_{i=0}^I g_i \sum_{p=1}^{P_1} y_{ip}^d d_p^d + (1-\lambda) \sum_{i=1}^I g_i \sum_{p=1}^{P_2} y_{ip}^n d_p^n \\
 \text{s.t.} \quad & \sum_{i=0}^I y_{ip}^d \geq \Delta \quad p=1, \dots, P_1 \quad \text{--- } u_p^d \\
 & - \sum_{p=1}^{P_1} y_{ip}^d \geq -x_i - b_i \quad i=0, \dots, I \quad \text{--- } v_i^d \\
 & \sum_{i=1}^I y_{ip}^n \geq \Delta \quad p=1, \dots, P_2 \quad \text{--- } u_p^n \\
 & - \sum_{p=1}^{P_2} y_{ip}^n \geq -x_i - b_i \quad i=1, \dots, I \quad \text{--- } v_i^n
 \end{aligned}$$

$$x, y \geq 0,$$

where $x_i, g_i, c_i,$ are as defined in previous sections, and

P_1 = max index of load block containing day peak demand

P_2 = max index of load block containing night peak demand

y_{ip}^d = amount of equipment type i dispatched for
load block p during day

y_{ip}^n = amount of equipment type i dispatched for
load block p during night

d_p^d = length of load block p during day

d_p^n = length of load block p during night

Δ = width of all load blocks

$u_p^{d,n}$ = dual variables associated with demand constraints
 (as designated in the problem formulation DP)

$v_i^{d,n}$ = dual variables associated with capacity
 constraints (as designated in DP)

If one lets Δ , the width of the load blocks, approach zero, the number of blocks approaches infinity, and the linear programming solution approaches that of the nonlinear program. In the nonlinear program, merit order loading is implicitly assumed; the allocation or dispatching of the solar unit is taken care of by moving the day inverse load duration curve to the left by the amount x_0 . Thus, for infinitesimally small blocks of width Δ and a fixed solar capacity x_0 , an ordinary expansion problem involving only conventional equipment results. It can therefore be solved by the expansion planning algorithm presented in Appendix A.

This observation encourages the decomposition of problem DP into a master program, associated with determining solar capacity, and a subproblem, associated with determining the conventional capacities given a fixed choice of the solar capacity. Therefore, in the master program, the only two variables will be z , denoting the objective value, and x_0 . The master program then passes a value of x_0 , which is optimal for a current iteration, into the subproblem. In the subproblem, this value of x_0 is treated as a constant, so

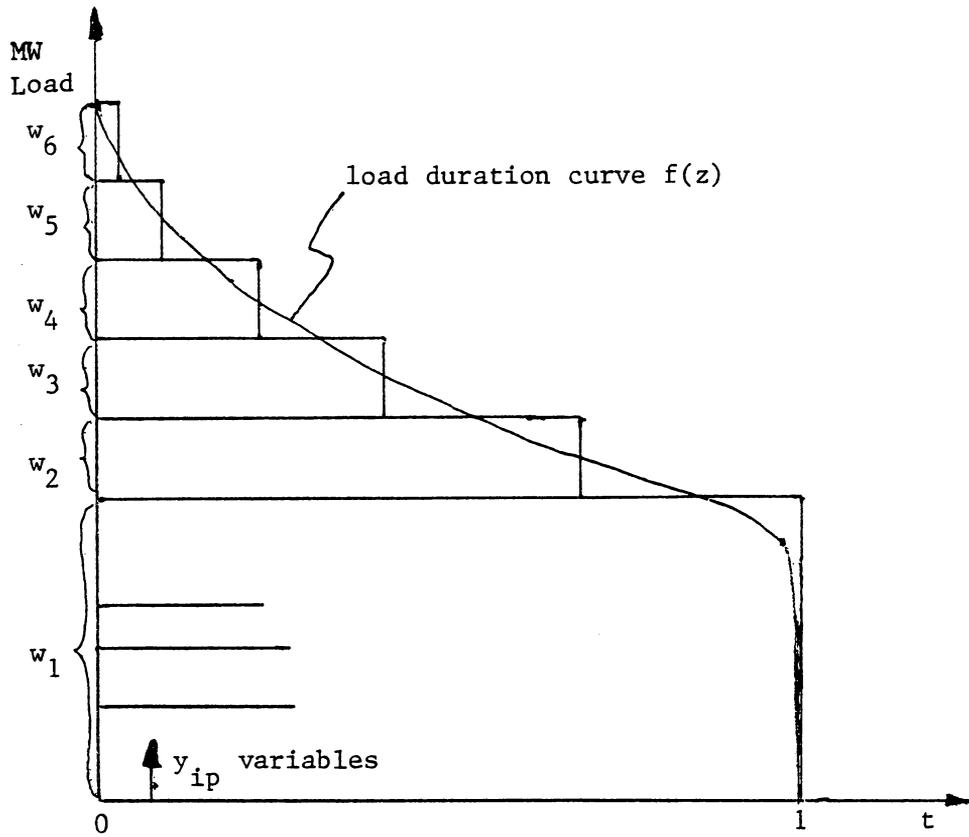


Figure 11: Discretization of Expansion Planning Program

that an ordinary expansion problem, not involving solar capacity as a decision variable, is obtained. The two programs, called MP and SP, respectively, are given below.

MP

min z

$$z \geq c_0 x_0 + h(x_0)$$

$$0 \leq x_0 \leq P_d - b_0$$

Here, $h(x_0)$ is the optimal objective value of the subprogram, for a fixed value of x_0 .

The subprogram consists of the rest of the expansion problem:

SP

$$h(x_0) = \min \sum_{i=1}^I c_i x_i + \lambda \sum_{i=0}^I g_i \sum_{p=1}^{P_1} y_{ip}^d d_p^d + (1-\lambda) \sum_{i=1}^I g_i \sum_{p=1}^{P_2} y_{ip}^n d_p^n$$

$$\text{s.t.} \quad \sum_{i=0}^I y_{ip}^d \geq \Delta \quad p=1, \dots, P_1 \quad \text{--- } u_p^d$$

$$- \sum_{p=1}^{P_1} y_{ip}^d + x_i \geq -b_i \quad i=1, \dots, I \quad \text{--- } v_i^d$$

$$- \sum_{p=1}^{P_1} y_{0p}^d \geq -x_0 - b_0 \quad \text{--- } v_0^d$$

$$\sum_{i=1}^I y_{ip}^n \geq \Delta \quad p=1, \dots, P_2 \quad \text{--- } u_p^n$$

$$- \sum_{p=1}^{P_2} y_{ip}^n + x_i \geq -b_i \quad i=1, \dots, I \quad \text{--- } v_i^n$$

$$x, y \geq 0$$

Note that the subproblem involves allocation variables, y , for both solar and conventional equipment, but capacity variables, x , only for conventional equipment.

Also note that the algorithm that will be applied to solve the nonlinear limiting version of the subproblem, involves a search technique, and thus only provides the primal optimal solution. But for applying a relaxation technique to solve the master program, a dual optimal solution of the subproblem is needed.

The dual of the subproblem takes the following form:

DSP

$$h(x_0) = \max \Delta \sum_{p=1}^{P_1} u_p^d + \Delta \sum_{p=1}^{P_2} u_p^n - \sum_{i=1}^I b_i (v_i^d + v_i^n) - (b_0 + x_0) v_0^d$$

$$\text{s.t. } -v_i^d + u_p^d \leq \lambda g_i d_p^d \quad i=0, \dots, I \quad ; \quad p=1, \dots, P_1$$

$$-v_i^n + u_p^n \leq (1-\lambda) g_i d_p^n \quad i=1, \dots, I \quad ; \quad p=1, \dots, P_2$$

$$v_i^d + v_i^n \leq c_i \quad i=1, \dots, I$$

$$u, v \geq 0 ,$$

where all variables are as defined before; the dual variables u and v are associated with demand and supply constraints, respectively, as indicated in SP.

Benders' decomposition, which will be applied here, involves the following logic. First, note that the feasible

region of DSP is independent of x_0 . Since it is a feasible, bounded linear program in non-negative variables, an optimum corresponding to any fixed x_0 will occur at one of the extreme points of the dual feasible region. Let the u and v values at these extreme points be denoted by u_{pq}^d , u_{pq}^n , v_{iq}^d , v_{iq}^n , with $q=1, \dots, Q$ representing the extreme points. Then, DSP can be written as

$$h(x_0) = \max \Delta \sum_{p=1}^P u_{pq}^d + \Delta \sum_{p=1}^P u_{pq}^n - \sum_{i=1}^I b_i (v_{iq}^d + v_{iq}^n) - (b_0 + x_0) v_{0q}^d \quad (11)$$

Using this expression for $h(x_0)$, the master program MP can be rewritten as follows.

MP1

min z

$$z \geq c_0 x_0 + \Delta \sum_{p=1}^P u_{pq}^d + \Delta \sum_{p=1}^P u_{pq}^n - \sum_{i=1}^I b_i (v_{iq}^d + v_{iq}^n) - (b_0 + x_0) v_{0q}^d$$

$q=1, \dots, Q$

$$0 \leq x_0 \leq P_d - b_0$$

Then, the decomposition algorithm works as follows.

Decomposition Algorithm DA

Start: Relax all Q constraints in MP1; find the optimal x_0 , z , denote them by \bar{x}_0 , \bar{z} .

Step 1: Solve subproblem SP with x_0 fixed at \bar{x}_0 ; obtain its objective function value $h(\bar{x}_0)$. Check if \bar{x}_0 and \bar{z} satisfy

$\bar{z} \geq c_0 \bar{x}_0 + h(\bar{x}_0)$, or, specifically,

$$z \geq c_0 x_0 + \Delta \sum_{p=1}^{P_1} \bar{u}_p^d + \Delta \sum_{p=1}^{P_2} \bar{u}_p^n - \sum_{i=1}^I b_i (\bar{v}_i^d + \bar{v}_i^n) - (b_0 + x_0) \bar{v}_0^d, \quad (12)$$

with the \bar{u}_p and \bar{v}_i denoting the optimal solution of DSP for \bar{x}_0 fixed. If they do, terminate with \bar{x}_0 (from MP1) and \bar{x}_i , $i=1, \dots, I$ (from SP) as the optimal solution. Otherwise, add a constraint of the form of Equation 12 to MP1, i.e. add

$$z \geq c_0 x_0 + \Delta \sum_{p=1}^{P_1} u_p^d + \Delta \sum_{p=1}^{P_2} u_p^n - \sum_{i=1}^I b_i (v_i^d + v_i^n) - (b_0 + x_0) v_0^d,$$

and proceed to Step 2.

Step 2: Solve MP1 and find a new \bar{x}_0 , \bar{z} . Return to Step 1.

Finite convergence of this algorithm is shown in BENDERS, 1962.

Recall that solving SP with the method described in Appendix A, only the primal optimal solution is obtained. However, for the Benders' cut as given in Equation 12, the optimal dual variables must be known. Thus it will be shown in the following subsection how the optimal dual variables can be obtained from the optimal primal variables. After

that, the algorithm will be restated, with explicit reference to how the needed dual variables can be computed. Using a general vector formulation of the problem, convergence of the algorithm will be proven.

4.3.3 The Optimal Dual Variables

For convenience, the Benders' cut as stated in Equation 12 is given again below.

$$z \geq c_0 x_0 + \Delta \sum_{p=1}^{P_1} u_p^d + \Delta \sum_{p=1}^{P_2} u_p^n - \sum_{i=1}^I b_i (v_i^d + v_i^n) - (b_0 + x_0) v_0^d \quad (12)$$

Note that the equal width of all load blocks, Δ , has been factored out in the cut expression. Thus, whereas all dual variables v , associated with equipment types $0, 1, \dots, I$, have to be calculated, for the dual variables u associated with the load blocks, only their sum over all blocks needs to be determined.

The primal subproblem is solved with the method described in Appendix A; this solution is optimal for an (infinitesimally) small value of Δ in Problem DP. Now, if all x variables are fixed at their optimal values, the subproblem (in the remaining allocation variables y) decomposes into two transportation problems, one for day and one for night time.

The primal and the dual of the transportation problem for day time is given below, with $\hat{v}_i^{d,n}$ and $\hat{u}_p^{d,n}$ denoting the dual variables associated with the supply and demand constraints, respectively. The corresponding problems for night time take the same form, except for the fact that the solar equipment type, $i=0$, does not appear at all. Other than that, only the superscripts d for day have to be replaced by n for night, in order to arrive at the night problems from the day problems. Therefore, in all following derivations only the day dual variables will be considered. The derivation for the night variables is analogous; only the final expressions will be given for them.

Primal Transportation Problem (PTP)

$$\begin{aligned} \min \quad & \sum_{i=0}^I \sum_{p=1}^{P_1} \lambda g_i^d d_p^d y_{ip}^d \\ \text{s.t.} \quad & - \sum_{p=1}^{P_1} y_{ip}^d \geq -x_i^* - b_i \quad i=0,1,\dots,I \\ & \sum_{i=0}^I y_{ip}^d = \Delta \quad p=1,\dots,P_1 \end{aligned}$$

$$y \geq 0$$

Dual Transportation Problem (DTP)

$$\begin{aligned} \max \quad & - \sum_{i=0}^I (x_i^* + b_i) \hat{v}_i^d + \Delta \sum_{p=1}^{P_1} \hat{u}_p^d \\ \text{s.t.} \quad & - \hat{v}_i^d + \hat{u}_p^d \leq \lambda g_i^d d_p^d \quad i=0,\dots,I ; p=1,\dots,P_1 \end{aligned}$$

$$v \geq 0$$

Note that the optimal solution to the overall dual subproblem (the dual to SP), u_p^d and v_i^d , solves the dual transportation problem DTP, too. Conversely, if \hat{u}_p^d , \hat{u}_p^n , and \hat{v}_i^d , \hat{v}_i^n , for all i and p , solve the day and night dual transportation problems, and in addition satisfy

$$\hat{v}_i^d + \hat{v}_i^n \leq c_i, \text{ for each } i=1, \dots, I, \quad (13a)$$

$$\text{with } \hat{v}_i^d + \hat{v}_i^n = c_i \text{ whenever } x_i^* > 0 \text{ for } i=1, \dots, I, \quad (13b)$$

then by duality, they also solve the dual subproblem DSP. In particular, if Equation 13 holds, then $u_p^d = \hat{u}_p^d$, and $v_i^d = \hat{v}_i^d$, for all i and p , and similarly, the analogous equivalences for the variables of the night problem, will hold. Below, an optimal dual solution to the above transportation problems is derived in closed form. Then it is demonstrated that these solutions satisfy Equation 13 for all i . In this manner, the dual solutions to be used for the Benders' cuts are derived.

4.3.3.1 Dual Variables Associated with Equipment Types

Consider the network interpretation of the subproblem, as shown in Figure 12 Equipment types $i=0, 1, \dots, I$ are interpret-

ed as supply nodes, with respective supplies of $(x_i^* + b_i)$. Blocks $p=1, \dots, P$, with P denoting the overall peak demand node, are interpreted as demand nodes, each having a demand Δ . The optimal allocation clearly loads the equipment types in merit order, and thus the positive flows will be distributed in the manner illustrated in Figure 12. Given an optimal basis tree graph, the dual variables are computed according to the formula

$$-\hat{v}_i^d + \hat{u}_p^d = \text{cost associated with basic arc connecting } i \text{ and } p$$

This formula is used in a recursive manner to compute the dual variables as follows. Let the final load segment to which equipment i gives positive flow be segment p . Suppose \hat{v}_{i+1}^d has been determined. Then, for the case when there is a basic arc $(i+1, p)$,

$$\hat{u}_p^d = \hat{v}_{i+1}^d + \lambda g_{i+1,p}^d, \text{ and } \hat{u}_p^d = \hat{v}_i^d + \lambda g_i^d. \quad (14)$$

$$\text{Thus } \hat{v}_i^d = \hat{v}_{i+1}^d + \lambda(g_{i+1} - g_i)d_p^d$$

Actually, there are two more possible cases for basic positive flow or degenerate arcs between nodes i , $i+1$, p and $p+1$ in an optimal basis. All three cases, together with the respective ways to compute \hat{v}_i^d from \hat{v}_{i+1}^d , are shown in Figure 13.

In all cases, d_p^d and d_{p+1}^d , respectively, denote the length of the load block associated with the first basic arc for equipment type $i+1$. Since a very fine discretization is assumed here, the lengths of two consecutive load blocks approach the same value, if their widths Δ approach zero (since the curve is assumed to be continuous). The value approached by both lengths is $F_d(X_i)$. If $F_d(X_i)$ is substituted for both d_{p+1}^d and d_p^d , all three expressions for \hat{v}_i^d become the same, namely,

$$\hat{v}_i^d = \hat{v}_{i+1}^d + \lambda(g_{i+1} - g_i)F_d(X_i). \quad (15)$$

The analogous expression for the night problem is

$$\hat{v}_i^n = \hat{v}_{i+1}^n + (1-\lambda)(g_{i+1} - g_i)F_n(Y_i). \quad (16)$$

By recursive substitution, one obtains

$$\hat{v}_i^d = \hat{v}_I^d + \lambda \sum_{j=i}^{I-1} (g_{j+1} - g_j)F_d(X_j) \quad i=0, \dots, I \quad (17)$$

The analogous expression for the night dual variables is

$$\hat{v}_i^n = \hat{v}_I^n + (1-\lambda) \sum_{j=i}^{I-1} (g_{j+1} - g_j)F_n(Y_j) \quad i=1, \dots, I \quad (18)$$

Due to the recursive structure of this computation, all variables v depend on the value assigned to $\hat{v}_I^{d,n}$. Thus, a value has to be assigned to $\hat{v}_I^{d,n}$ such that for all i , Equation 13 holds.

For a fixed x_0 , and for Δ approaching zero, the subproblem SP approaches a nonlinear program of the type of program PWS in the limit. The Kuhn Tucker conditions for this program will be used to show how Equation 13 holds for all i .

Note that this nonlinear programming subproblem is of the form

PWS1

$$\begin{aligned} \min \quad & \sum_{i=1}^k c_i x_i + \sum_{i=1}^k g_i \int_{Y_{i-1}}^{Y_i} F(z) dz \\ \text{s.t.} \quad & \sum_{i=1}^k x_i \geq P - B \end{aligned}$$

$$x_i \geq 0, \quad i=1, \dots, k$$

where $F(z) = \lambda F_d(z+x_0) + (1-\lambda)F_n(z)$ is the aggregate inverse load duration curve, $P = \max \{P_d - x_0, P_n\}$, and where the index k is defined as follows. As shown in Appendix A, the algorithm due to SHERALI (1982a) determines in phase Ia the capacities of old equipment used. If all existing capacities are used, then $k=I$ and $B = \sum_{i=1}^I b_i$. On the other hand, if it is determined that for some existing equipment $k \leq I$, $z_k^* < b_k$, then $x_j^* = 0$ for $j \geq k$ and $z_j^* = 0$ for $j > k$. Hence above, the index k corresponds to such an equipment type, and the variable x_k is in reality determining the value for z_k , so that

$c_k=0$, and $B = \sum_{i=1}^{k-1} b_i$. The above program will, by phase Ia of SHERALI, 1982a, automatically determine x_k^* as the value of $x_k^* < b_k$.

The Kuhn Tucker conditions for program PWS1 are:

$$\nabla_i - \pi - \pi_i = 0 \quad \text{for } i=1, \dots, k \quad (19)$$

$$\sum_{i=1}^k x_i = P - B \quad (20)$$

$$x_i \geq 0 ; \pi_i \geq 0 ; x_i \pi_i = 0, \quad i=1, \dots, k \quad (21)$$

where ∇_i = partial derivative of the objective function of PWS1 with respect to x_i

$$= \begin{cases} c_i + \sum_{j=i}^{k-1} (g_j - g_{j+1}) F\left[\sum_{m=1}^j (x_m + b_m)\right], & \text{for } i=1, \dots, k-1 \\ c_k & \text{for } i=k \end{cases} \quad (22)$$

π = dual variable for the single demand constraint

π_i = dual variables associated with $x_i \geq 0$, $i=1, \dots, k$

Now, for any two consecutive equipment types i and $i+1$, using Equation 19, $\nabla_i - \pi_i = \nabla_{i+1} - \pi_{i+1}$ gives,

noting $Y_j = \sum_{m=1}^j (x_m + b_m)$:

$$(c_i - \pi_i) + \sum_{j=i}^{k-1} (g_j - g_{j+1}) F(Y_j) = (c_{i+1} - \pi_{i+1}) + \sum_{j=i+1}^{k-1} (g_j - g_{j+1}) F(Y_j)$$

$$\text{or } (c_i - \pi_i) + (g_i - g_{i+1}) F(Y_i) = (c_{i+1} - \pi_{i+1})$$

$$\text{or } (c_i - \pi_i) + g_i F(Y_i) = (c_{i+1} - \pi_{i+1}) + g_{i+1} F(Y_i) \quad (23)$$

Observe that Equation 23 is a breakeven equation which asserts that the imputed capital costs for the equivalent breakeven analysis are $(c_i - \pi_i)$, $i=1, \dots, k$. That is, they are either c_i if $x_i^* > 0$ (whence $\pi_i = 0$), or they are less than or equal to c_i if $x_i^* = 0$ (whence $\pi_i \geq 0$). This forms the crux of the argument used below.

Thus, in reference to Figure 12, consider the assignment procedure for the dual variables, given in the following theorem. Note that since x_0 is treated as a constant in the subproblem, Equation 13 need not be shown to hold for v_0^d .

THEOREM 2

Let k be determined as for problem PWS1 above and let the Kuhn Tucker solution $(\underline{x}^*, \underline{\pi}, \underline{\pi})$ be determined through Equations 19, 20, 21. Suppose that \hat{v}_k^d and \hat{v}_k^n are assigned values such that

$$\begin{aligned} \hat{v}_k^d + \hat{v}_d^n &= \begin{array}{ll} 0 & \text{if } k < I \\ c_k & \text{if } k = I, x_k^* > 0 \\ \pi = c_k - \pi_k & \text{if } k = I, x_k^* = 0 \end{array} \end{aligned} \quad (24)$$

Further, suppose that \hat{v}_i^d and \hat{v}_i^n , $i=1, \dots, k-1$, are computed via Equations 15 and 16.

Then Equations 13a and 13b hold and therefore

$$v_i^d = \hat{v}_i^d \text{ and } v_i^n = \hat{v}_i^n, \text{ for all } i=1, \dots, I.$$

Proof (by induction)

First of all, observe that for $i > k$, $\hat{v}_i^d = \hat{v}_i^n = 0$ (due to non-binding supply constraints), and so Equations 13a and 13b hold. Now, suppose that we show that

$$\hat{v}_i^d + \hat{v}_i^n = c_i - \pi_i, \quad i=1, \dots, k \quad (25)$$

where π_i are the optimal dual variables given by Equations 19 and 20. Then the proof will be complete since $(\hat{v}_i^d + \hat{v}_i^n)$ will be c_i if $x_i^* > 0$ (as $\pi_i = 0$ in this case), and $(\hat{v}_i^d + \hat{v}_i^n)$ will be less than or equal to c_i otherwise (as $\pi_i \geq 0$ in this case). Hence, let us prove Equation 25 by induction.

Consider $i=k$. If $k < I$, then from Equation 22, $v_k = c_k = 0$. Further, from the \geq constraint in PWS1, $\pi \geq 0$. Equation 19 therefore asserts that

$$\pi + \pi_k = 0 \text{ or } \pi = \pi_k = 0.$$

Since Equation 24 sets $\hat{v}_k^d + \hat{v}_k^n = 0$ in this case, Equation 25 holds.

If $k=I$ and $x_k^* > 0$, then $\pi_k = 0$ and again, Equation 24 implies 25. Finally, if $k=I$ and $x_k^* = 0$, then Equations 21, 22 and 24 imply that

$$\hat{v}_k^d + \hat{v}_k^n = c_k - \pi_k \text{ which is precisely Equation 25. Thus Equation 25 holds for } i=k.$$

Assume that Equation 25 holds for some $2 \leq i+1 \leq k$, and consider Equation 25 for index i . Equations 15 and 16 imply:

$$\begin{aligned} \hat{V}_i^d + \hat{V}_i^n &= \hat{V}_{i+1}^d + \hat{V}_{i+1}^n + (g_{i+1} - g_i)[\lambda F_d(X_i) + (1-\lambda)F_n(Y_i)] \\ &= (c_{i+1} - \pi_{i+1}) + (g_{i+1} - g_i)[\lambda F_d(X_i) + (1-\lambda)F_n(Y_i)]. \\ &= (c_{i+1} - \pi_{i+1}) + (g_{i+1} - g_i)F(Y_i) \\ &= c_i - \pi_i \quad \text{by Equation 23.} \end{aligned}$$

Hence Equation 25 holds for i and the proof is complete.

Remark: In case π has to be determined to evaluate Equation 24, observe from Equations 19, 20 and 21 that $\pi = \nabla_i$ for any i such that $x_i^* > 0$, and one such i must exist by the algorithm in Appendix A when $k=I$.

The following strategy can be adopted for assigning values for \hat{V}_k^d or \hat{V}_k^n , respectively, satisfying Equation 24. If $k < I$, both $\hat{V}_k^d = \hat{V}_k^n = 0$. If $k=I$, then when equipment type k is not used to full capacity during either day or night, its supply constraint for that period is nonbinding at optimality, and the appropriate \hat{V}_k^d or $\hat{V}_k^n = 0$, necessarily. In that case, the respective other period has to get assigned the full cost c_k or π .

If both day and night make use of equipment k to full available capacity, there is no such necessary distribution of the capital cost of equipment k . In that case, a possi-

ble distribution would be the following. This distribution can be considered fair, since it allocates costs proportional to the usage of peaking unit k .

$$\begin{aligned}\hat{v}_k^d &= \{c_k \text{ or } \pi\} \cdot \{ \lambda F_d(X_{k-1}) / [\lambda F_d(X_{k-1}) + (1-\lambda) F_n(Y_{k-1})] \} \quad (26) \\ \hat{v}_k^n &= \{c_k \text{ or } \pi\} \cdot \{ (1-\lambda) F_n(Y_{k-1}) / [\lambda F_d(X_{k-1}) + (1-\lambda) F_n(Y_{k-1})] \}\end{aligned}$$

Since Theorem 2 implies that the v equal the \hat{v} variables, and the u equal the \hat{u} variables, from now on only u and v variables will be used.

4.3.3.2 Dual Variables Associated with Demand Blocks

For calculating the dual variables u , associated with load blocks p , the same formula as used for computing the variables v in Equation 14 can be used:

$$\begin{aligned}u_p^d - v_i^d &= \lambda g_{ip}^d, \text{ whenever } y_{ip}^d > 0 \\ \text{or } u_p^d &= v_i^d + \lambda g_{ip}^d, \text{ whenever } y_{ip}^d > 0.\end{aligned} \quad (27)$$

Now, define the sets S_0, S_1, \dots, S_k such that

$$S_i = \{p: y_{ip}^d > 0\},$$

so that $|S_i|$ equals the number of load blocks served by equipment type i . (Note that due to merit order loading, each set only contains consecutive load blocks.) This together with Equation 27 implies that

$$\Delta \sum_{p \in S_i} u_p^d = \Delta |S_i| v_i^d + \lambda g_i \sum_{p \in S_i} \Delta d_p^d$$

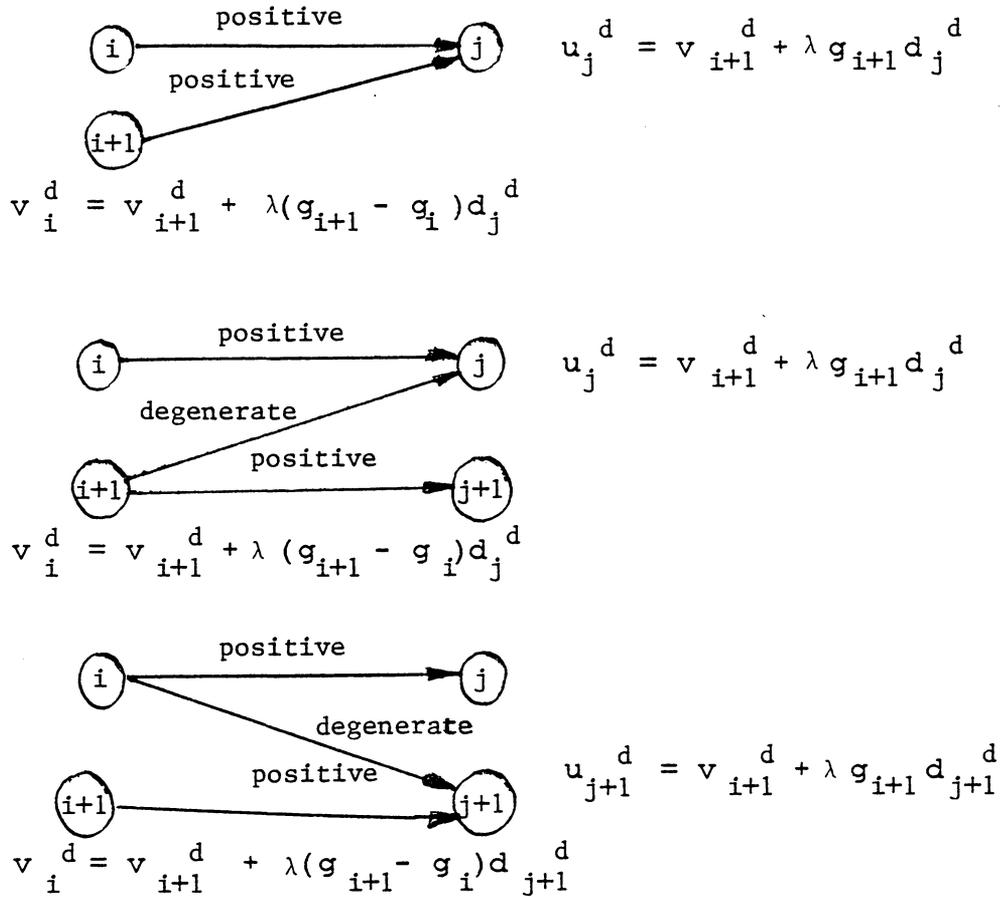


Figure 13: Possible Cases of Positive and Degenerate Flows

$$\Delta \sum_{p \in S_i} u_p^d \rightarrow (x_i + b_i) v_i^d + \lambda g_i \int_{x_{i-1}}^{x_i} F_d(z) dz ,$$

in the limit as the block width Δ approaches 0 and the linear program approaches the nonlinear program.

Therefore

$$\Delta \sum_{p=1}^P u_p^d = \Delta \sum_{i=0}^I \sum_{p \in S_i} u_p^d \rightarrow \sum_{i=0}^I [(x_i + b_i) v_i^d + \lambda g_i \int_{x_{i-1}}^{x_i} F_d(z) dz] \quad (28)$$

Again, the derivation for the night dual variables is analogous, and results in the following expression:

$$\Delta \sum_{p=1}^P u_p^n = \Delta \sum_{i=1}^I \sum_{p \in S_i} u_p^n = \sum_{i=1}^I [(x_i + b_i) v_i^n + (1-\lambda) g_i \int_{y_{i-1}}^{y_i} F_n(z) dz] \quad (29)$$

4.3.4 Derivation of the Cut

Recall the form of the Benders cut given in Equation 12:

$$z \geq c_0 x_0 + \Delta \sum_{p=1}^P (u_p^d + u_p^n) - \sum_{i=1}^I b_i (v_i^d + v_i^n) - (b_0 + x_0) v_0^d$$

The expressions involving u variables have been derived above in Equations 28 and 29. Thus, the cut becomes

$$z \geq c_0 x_0 + \sum_{i=1}^I x_i (v_i^d + v_i^n) + g_0 \lambda \int_0^{x_0 + b_0} F_d(z) dz \quad (30)$$

$$+ \sum_{i=1}^I g_i \int_{y_{i-1}}^{y_i} [\lambda F_d(z + x_0) + (1-\lambda) F_n(z)] dz + (\hat{x}_0 - x_0) v_0^d$$

In this term, x_0 only appears with the factors c_0 and v_0^d ; all other terms are constants once an x_0 is specified.

Therefore all terms that do not depend on x_0 can be taken together as β_q ; and the simple expression

$$z \geq c_0 x_0 + \beta_q - v_0^d (x_0 - \hat{x}_0) \quad (31)$$

results.

Consider the following term in Equation 30:

$$\sum_{i=1}^I x_i (v_i^d + v_i^n) .$$

In Theorem 2 it has been shown that $(v_i^d + v_i^n) = c_i$, whenever $x_i^* > 0$. Therefore, for the optimal solution of the subproblem,

$$\sum_{i=1}^I x_i (v_i^d + v_i^n) = \sum_{i=1}^I c_i x_i \quad (32)$$

has to hold. Therefore, β_q can be interpreted as the total cost of the expansion plan for the given value of \hat{x}_0 , excluding solar capital costs, but including solar operating cost. For this cut, consider the slope associated with the variable x_0 , $(c_0 - v_0^d)$. In Equation 17, a recursion formula for the v variables had been developed. Applied on v_0^d , it becomes

$$v_0^d = v_k^d + \lambda \sum_{j=0}^{k-1} (g_{j+1} - g_j) E_d(X_j) \quad (33).$$

Thus, the expression $(c_0 - v_0^d)$ takes the following form:

$$c_0 - v_0^d = c_0 - v_k^d + \sum_{j=0}^{k-1} (g_j - g_{j+1}) \lambda E_d(X_j) , \quad (34)$$

which is exactly of the same form as the one developed for $h'(x_0)$ in Chapter 3, for the plant mix program.

4.3.5 Algorithm

The algorithm to find the optimal expansion plan with the option of investing in solar energy may be implemented as follows.

Expansion Planning Algorithm EPA

Step 1: First, relax all constraints in the master program MP1. Its first solution will be $x_0=0$. Solve the non-linear subproblem for x_0 fixed at 0 with the method given in SHERALI, 1982a (Appendix A).

Check for Termination: If v_0^d resulting from subproblem SP, computed according to Equation 33, is less than or equal to c_0 , then stop, $x_0^* = 0$. (This is so since the next master program MP1 will give $x_0=0$ again.)

Step 2: Otherwise, $v_0^d > c_0$. Add a cut to the master program MP1 according to Equation 30 or 31. The next x_0 to be examined will be $x_0 = P_d - b_0$. Compute v_0^d (Equation 33). If $v_0^d > c_0$, terminate with $x_0^* = P_d - b_0$ (the master program would give again $x_0 = P_d - b_0$ in the next iteration). Otherwise, the optimal x_0 satisfies $0 < x_0^* < P_d - b_0$. Then, add a cut to the

. master program according to Equation 30 or 31, and proceed to step 3.

Step 3: Solve the master program MP1, including all cuts added so far, to obtain its solution \bar{x}_0, \bar{z} . Then, \bar{z} is a lower bound (LB) on the optimal solution. It is necessary to check if all relaxed constraints are satisfied; a violated constraint will otherwise be used to generate a cut.

Therefore, with \bar{x}_0 fixed, solve the subproblem SP with SHERALI's method. Its objective function value will be an upper bound (UB) on the optimum. Compute v_0^d (Equation 33) and β_q (Equations 30 and 31). Check if the following condition holds:

$$\bar{z} + \quad \geq c_0 \bar{x}_0 + \beta_q \quad (34)$$

for some specified accuracy ϵ .

If Equation 34 holds, then $LB + \epsilon \geq UB$, and the current solution \bar{x}_0 (from MP1) and $\bar{x}_i, i=1, \dots, k$ (from SP) is optimal (terminate). Otherwise, a cut is generated according to Equation 30 or 31, and step 3 is repeated.

In the following subsection, a general vector formulation of the problem and the algorithm will be given. It will be used to show that the total cost function is convex. This fact will make a graphical interpretation of the Algorithm

EPA possible; in addition, it will make it possible to simplify its step 3 considerably. Finally, convergence of EPA will be proven.

4.3.6 General Vector Formulation of the Technique

4.3.6.1 The Benders Decomposition Framework

The general form of the expansion problem, split up into master program and subproblem, is the following:

$$\min c_0 x_0 + h(x_0) \quad (\text{GMP})$$

$$0 \leq x_0 \leq P_d - b_0$$

$$\text{where } h(x_0) = \min \underline{c}x + \underline{d}y$$

$$\text{s.t. } \underline{A}x + \underline{D}y \geq \underline{b} - \underline{e}x_0 \quad (\text{GSP})$$

$$\underline{x}, \underline{y} \geq \underline{0}$$

Here, underscored letters indicate vectors or matrices, and \underline{e} is a vector of the form $\underline{e} = (0, \dots, 0, 1, 0, \dots, 0)$ with the 1 in the appropriate position (say, position r), where the capacity constraint for the solar equipment is located in the constraint matrix. \underline{A} and \underline{D} represent matrices of coefficients of \underline{x} and \underline{y} , respectively, in the constraints of SP; \underline{c} and \underline{d} represent the objective function coefficients of SP, and \underline{b} the right hand side vector (exclusive x_0).

Following Benders' decomposition approach, note that the subproblem can be written as follows (since the subproblem has as optimal solution for any $0 \leq x_0 \leq P_d$):

$$h(x_0) = \max_{q=1, \dots, Q} \underline{s}_q (\underline{b} - \underline{e}x_0) , \quad (35)$$

where $\underline{s}_{-1}, \dots, \underline{s}_{-Q}$ are extreme points of the dual feasible region

$$S = \{ \underline{s} : \underline{s}A \leq \underline{c} , \underline{s}D \leq \underline{d} , \underline{s} \geq \underline{0} \}.$$

Thus, s variables replace u and v variables used in previous formulations.

From the formulation in Equation 35 it follows that $h(x_0)$ is convex. In fact, it is convex for any discretization parameter and hence also as $\Delta \rightarrow 0$.

At any \bar{x}_0 , a support of $h(x_0)$ is given by $h(\bar{x}_0) + (x_0 - \bar{x}_0)h^+(x_0)$, where $h^+(x_0)$ is the right hand derivative of $h(\cdot)$ with respect to x_0 at $x_0 = \bar{x}_0$. But this derivative with respect to x_0 can be expressed as the optimal dual s_r^* , since optimal dual variables denote marginal rates of increase of objective values with respect to the right hand side. Also noting that the optimal dual solution has the same value as the optimal primal solution, the support can be written as

$$\underline{s}^* (\underline{b} - \underline{e}x_0) + (x_0 - \bar{x}_0)(-s_r^*) = \underline{s}^* (\underline{b} - \underline{e}x_0) , \quad (36)$$

where \underline{s}^* denotes the optimal dual solution to the subproblem. Note that in our general notation, s_r denotes the dual variable associated with the capacity constraint for solar equipment. Thus, it is analogous to the variable v_0^d used previously. To illustrate the equivalence of the general vector formulation of the Benders cut and its form given in Equation 31, v_0^d will be used from now on. Also, instead of the general \underline{s}^* and \bar{x}_0 above, we will use the particular values s_q and x_{0q} , respectively. Now, observe that as used in Equation 31,

$\beta_q = \underline{s}_q (b - \underline{e}x_{0q})$ denotes the total cost of the expansion plan (excluding solar capital costs) for a given x_{0q} . Then, the master program takes the following form:

min z

$$z \geq c_0 x_0 + \beta_q + v_0^d (x_{0q} - x_0) \quad \text{for all } q \text{ or } x_{0q}$$

$$0 \leq x_0 \leq P_d - b_0$$

The decomposition algorithm involves the following relaxation strategy. First, all constraints are relaxed, the master program is solved, and a solution (\bar{z}, \bar{x}_0) is obtained. Then, it is necessary to check if this solution satisfies $\bar{z} + \epsilon \geq c_0 \bar{x}_0 + h(\bar{x}_0)$.

If it does, then it is an optimal solution. Otherwise, the cut

$$z \geq c_0 x_0 + \beta_q + (\bar{x}_0 - x_0) v_0^d$$

is generated and added to the master program.

Note that this is exactly the form of the cut as given in Equation 31. It follows from Equation 36 that this is a support of the objective function of GMP; and since GMP is a generalized formulation of MP1, it is also a support of the objective function of MP1. This fact will be used below for a graphical interpretation of the EPA Algorithm. But first, a restatement of the algorithm in terms of supports of $h(\cdot)$ will be given, and finite convergence will be proven.

4.3.6.2 Convergence

Consider the following Algorithm GA, still stated in general vector notation, which is simply a restatement of Algorithm EPA.

General Algorithm GA

Start: Start with $\bar{x}_0 = 0$.

Check for Termination: If v_0^d resulting from the subproblem (which is an expansion problem not involving solar energy, and from which the dual v_0^d is computed via Equation 33) is less than or equal to c_0 , then stop; $x_0^* = 0$.

Otherwise, $v_0^d > c_0$, and the next \bar{x}_0 to be examined will be $\bar{x}_0 = P_d - b_0$.

Main Iteration Step: For each \bar{x}_0 examined, solve subproblem GSP.

If $\bar{z} + \varepsilon \geq c_0 \bar{x}_0 + h(\bar{x}_0)$, for a given $\varepsilon > 0$ tolerance, terminate. Otherwise, generate a cut

$$z \geq c_0 x_0 + \beta_q + (\bar{x}_0 - x_0) v_0^d$$

and add it to the master program GMP. Repeat.

THEOREM 3

The Algorithm GA stated above is finitely convergent.

Proof Assume it be not finitely convergent, i.e., an infinite sequence $\{z_\ell, x_{0\ell}\}$ is generated. This sequence determines an infinite sequence $\{v_{0\ell}^d\}$. Since all sequences are bounded, the Bolzano-Weierstrass Theorem implies that there exists a convergent subsequence indexed by L, say, such that $\{z_\ell\} \rightarrow \tilde{z}$, $\{x_{0\ell}\} \rightarrow \tilde{x}$, $\{v_{0\ell}^d\} \rightarrow \tilde{v}_0^d$; and by continuity of $h(\cdot)$,

$\{\beta_\ell = h(x_{0\ell})\} \rightarrow \tilde{\beta} = h(\tilde{x})$. Assume that $\ell \in L$ is renumbered consecutively.

Since by assumption the algorithm does not stop,

$$z_\ell + \varepsilon < c_0 x_{0\ell} + \beta_\ell, \quad \text{for all } \ell. \quad (37)$$

But since $(z_\ell, x_{0\ell})$ solves the master program which includes all previous cuts, one obtains

$$z_\ell \geq c_0 x_{0\ell} + \beta_{\ell-1} + (x_{0,\ell-1} - x_{0\ell}) v_{0,\ell-1}^d, \text{ for all } \ell \geq 2. \quad (38)$$

Taking limits for Equations 37 and 38, as $\ell \rightarrow \infty$, $\tilde{z} \geq c_0 \tilde{x}_0 + \tilde{\beta} \geq \tilde{z} + \epsilon$ results.

Since $\epsilon > 0$, this is a contradiction. Thus the algorithm is finitely convergent, and the proof is complete.

In concluding this subsection, we re-emphasize that the above convergence proof is basically for Algorithm EPA. Below, we will return to the original formulation of MP1 and SP, and use the convexity property shown in this subsection to simplify Algorithm EPA.

4.3.7 Expedient for Solving the Master Program

Consider the expression given for the cut in Equation 31:

$$z \geq c_0 x_0 + \beta_q - v_0^d (x_0 - \bar{x}_0)$$

Here, β_q is the total cost of the expansion plan for the value of x_0 obtained in the last iteration of the master program (excluding solar capital cost). Thus, the cut is a support for the convex total cost function $h(x_0)$, as shown in Figure 14. (Recall from Subsection 4.3.6 that $h(x_0)$ is convex.)

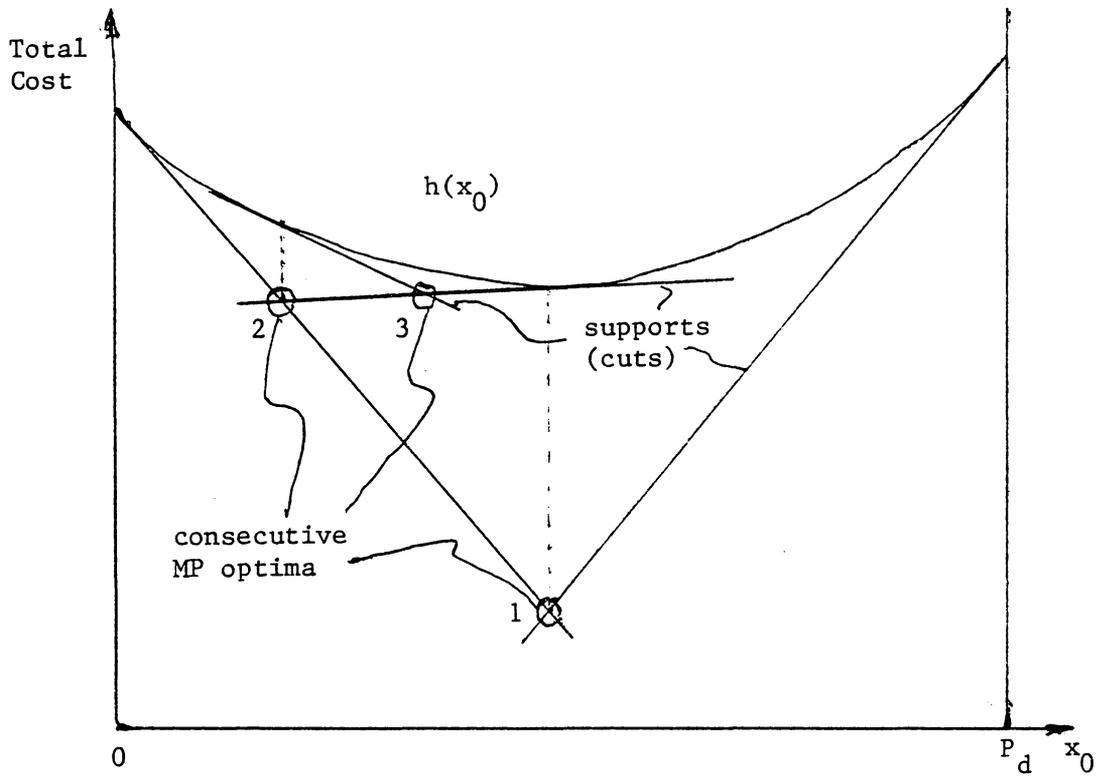


Figure 14: Decomposition Algorithm Logic

The master program incorporates supports added as constraints in previous iterations. At the minimum obtained in the master program, which is a linear program in the two variables z and x_0 , only two constraints will be binding. These binding constraints will correspond to the two cuts with the smallest positive and the least negative slope found so far. This is also depicted in Figure 14 .

Using this observation, step 3 of Algorithm EPA can be simplified. Instead of adding each cut to the master program and solving this ever growing linear program, only two constraints have to be kept in the master program, associated with the supports displaying the least positive and the least negative slope. In each iteration, the slope of the generated constraint is compared to the slopes of the two constraints of the last iteration. Out of these three constraints, the two with the smallest positive and the least negative slopes are kept.

In addition to that simplification, it is not necessary to solve the resulting linear program, either. The optimum of the master program will simply be at the intersection point of the two constraints with the most level slopes. This follows from convexity of $h(x_0)$.

The simplified step 3 of EPA is stated rigorously below.

Simplified Step 3 of Algorithm EPA

Let the two cuts present in the master program be denoted by cut 'pos' and cut 'neg' (for positive and negative slope). Find the intersection point \bar{x}_0 of the two cuts. For this \bar{x}_0 fixed, solve SP with SHERALI's method. Calculate v_0^d according to Equation 33. If $|c_0 - v_0^d| < \epsilon$, where ϵ is some specified accuracy, terminate; the current \bar{x}_0 and \bar{x}_i , $i=1, \dots, k$, as obtained from SP, are the optimal solution. Otherwise, if $(c_0 - v_0^d)$ is positive, disregard cut 'pos' (since due to convexity of $h(x_0)$ later cuts will have less steep slopes). Otherwise, disregard cut 'neg'. (Due to convexity of $h(x_0)$, a constraint having a less steep slope, will intersect with the constraint with opposite slope closer to the optimum.) Repeat step 3.

4.4 ILLUSTRATIVE EXAMPLE AND COMPUTATIONAL RESULTS

To further illustrate how Algorithm EPA works, it is applied on an example. The inverse load duration curves are approximated by fifth order polynomials, and the following cost coefficients and old capacities are assumed.

$$c_0=14, g_0=0, b_0=0$$

$$c_1=10, g_1=10, b_1=10$$

$$c_2=8, g_2=20, b_2=5$$

$$c_3=6, g_3=40, b_3=3$$

$$F_d(z) = (7/2 \cdot 20^5)z^5 - (33/8 \cdot 20^4)z^4 - (1/4 \cdot 20^3)z^3 - (1/8 \cdot 20^2)z^2 + 1$$

$$F_n(z) = (7/2 \cdot 10^5)z^5 - (33/8 \cdot 10^4)z^4 - (1/4 \cdot 10^3)z^3 - (1/8 \cdot 10^2)z^2 + 1$$

To find the capacities of conventional equipment types in each iteration, the algorithm given in Appendix A (due to SHERALI, 1982a) is used. Since the procedure of that algorithm is not a major subject of this thesis, the details of applying it are omitted, and only its results are given for each iteration. For convenience, the expressions for the Benders cut (Equation 30) and for v_0^d (Equation 33) are restated below:

$$z \geq c_0 x_0 + \sum_{i=1}^I x_i (v_i^d + v_i^n) + g_0 \lambda \int_0^{x_0 + b_0} F_d(z) dz \quad (30)$$

$$+ \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z + x_0) + (1 - \lambda) F_n(z)] dz + (\hat{x}_0 - x_0) v_0^d$$

$$v_0^d = v_k^d + \lambda \sum_{j=0}^{k-1} (g_{j+1} - g_j) F_d(X_j) \quad (33)$$

Algorithm EPA applied on the problem given above works as follows:

Step 1:

$$x_0 = 0 \implies z_1 = 10, x_1 = 1.92, z_2 = 5, x_2 = 0, z_3 = 3, x_3 = 0$$

$$v_0^d = 15. \implies \text{cut: } z \geq -x_0 + 130.2 \quad (\text{'neg'})$$

$$\text{Step 2: } x_0 = P_d = 19.96 \implies z_1 = 9.96, x_1 = z_2 = x_2 = z_3 = x_3 = 0$$

$$v_0^d = 0. \implies \text{cut: } z \geq 14x_0 + 32.7 \quad ('pos')$$

Step 3:

Intersection of the two cuts:

$$-x_0 + 130.2 = 14x_0 + 32.7 \implies x_{0 \text{ new}} = 6.5$$

$$x_0=6.5 \implies z_1=10, x_1=0, z_2=3.36, x_2=0, z_3=0, x_3=0$$

$$v_0^d = 5.74 \implies \text{cut: } z \geq 8.26x_0 + 104.6 \quad ('pos')$$

Therefore, cut 'pos' (from Step 2) is disregarded, a new 'pos' is constructed from Step 3 with $x_0=6.5$. The optimal x_0 will lie between 0 and 6.5.

Since this is sufficient to show how Algorithm EPA with modified Step 3 works, we quit the example here. The optimal solution, obtained in seven iterations, is given as a matter of interest:

$$x_0=1.97, z_1=10, x_1=0, z_2=5, x_2=0, z_3=2.93, x_3=0$$

This optimum satisfies the termination criterion, since the objective values of the master program (128.505) and the subproblem (128.506) lie within the specified tolerance of 0.05 of each other.

A Fortran computer program implementing Algorithm EPA was written and applied on the problem given above and several other example problems. The program uses the modified Step 3 of EPA; each subproblem is solved by applying the method

due to SHERALI, 1982a. It has to be noted, however, that the computer program is specifically designed to solve problems with only three conventional equipment types available; therefore computation times are faster than they would be with a more general program. For the example problems used, between six and eight iterations were needed to arrive at the optimum within the given tolerance; this required CPU times of about 0.7 seconds.

4.5 APPROACH USING LAGRANGIAN DUAL METHOD

In the foregoing sections, the reader may have observed that the solution technique developed resembles the tangential approximation method for solving Lagrangian duals. It turns out that the one year expansion problem can also be solved directly via duality, although not as efficiently as with the above proposed technique. Hence, this section is given only as a point of interest. Once the solution method is delineated, it will be compared to the decomposition method as given in Algorithm EPA.

We start with the same basic expansion program EP as given in Section 4.1. The Lagrangian dual of the problem may be defined as follows, by dualizing the constraints $z \leq b$.

LDP

$$\max \theta(u) \quad , \quad u \geq 0$$

LDSP

$$\text{where } \theta(u) = \min_{x_0} c_0 x_0 + \lambda g_0 \int_0^{x_0+b_0} F_d(z) dz + \sum_{i=1}^I c_i x_i \quad (35)$$

$$+ \sum_{i=1}^I u_i (z_i - b_i) + \sum_{i=1}^I g_i \int_0^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda)F_n(z)] dz$$

$$\text{s.t. } \sum_{i=1}^I (z_i + x_i) \geq P_d - x_0 - b_0$$

$$\sum_{i=1}^I (z_i + x_i) \geq P_n$$

$$x, z \geq 0$$

where LDSP stands for Lagrangian Dual Subproblem.

Again, Y_i is defined as $Y_i = \sum_{j=1}^i (z_j + x_j)$, for $i=1, \dots, I$

Note that for a fixed $\underline{u} \geq \underline{0}$, taking $\min \{c_i, u_i\}$ as respective capital costs for each $i=1, \dots, I$, and letting $t_i = (z_i + x_i)$, results in an equivalent capacity planning problem with the appropriate capital costs c_i or u_i . Thus, an optimum can fairly easily be obtained with the method described in Chapter 3. The rationale for using t_i is that the capacity planning algorithm shall treat z_i and x_i , for the same i , as different equipment types. Since they have the same operating cost g_i , it will be decided whether to make z_i or x_i positive by determining which one has less ca-

pital cost associated with it. This is achieved by using t_i and the appropriate capital cost.

$$\text{Also } Y_i = \sum_{j=1}^i (z_j + x_j) = \sum_{j=1}^i t_j, \text{ for } i=1, \dots, I,$$

will be guaranteed. Thus, all variables, integration bounds and constraints are analogous to the ones in Problem 1 as stated in Chapter 3, if t_i is used instead of x_i for all $i=1, \dots, I$ in Problem 1. Thus, Algorithm CPA as stated in Chapter 3 can solve problem LDSP for a fixed u . Consequently, consider the following decomposition approach.

The basic thrust of the algorithm resembles that of the foregoing section. The function $\theta(\cdot)$ is well known to be a concave (subdifferentiable) function of u . The master program approximates $\theta(\cdot)$ as the minimum of its tangential supports and thereby computes an upper bound (UB) on the problem. For the u vector which gives the optimum for the tangential approximation, the actual value of $\theta(u)$ is computed in the subproblem. This gives a lower bound (LB) for the optimum. If the upper and the latter lower bound are close enough, one terminates. Otherwise, one generates another support to $\theta(\cdot)$ at this u vector and repeats. These ideas are formalized below.

Let (x_0^j, x^j, z^j) be optimal in the master program for some generated u^j so far, $j=1, \dots, J$. Then, the tangential approximation is given by solving the

Master Program

max σ

$$\sigma \leq c_0 x_0 + \lambda g_0 \int_0^{x_0^j + b_0} F_d(z) dz + \sum_{i=1}^I c_i x_i^j + \sum_{i=1}^I u_i^j (z_i^j - b_i) \\ + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z + x_0^j) + (1-\lambda) F_n(z)] dz \quad , \quad j=1, \dots, J$$

$u \geq 0$

or, noting $\theta(u^j)$ from Equation 35, this may be compactly written as

LDMP

max σ

$$\sigma \leq \theta(u^j) + \sum_{i=1}^I (z_i^j - b_i)(u_i - u_i^j) \quad , \quad j=1, \dots, J \quad (36)$$

$u \geq 0$

Let (σ^{p+1}, u^{p+1}) solve this linear program. To check for feasibility to all relaxed constraints, one has to verify if z^{p+1} , or, approximately, $z^{p+1} - \epsilon$ for some tolerance $\epsilon > 0$, is less than or equal to the actual value $\theta(u^{p+1})$. Hence, one evaluates $\theta(u^{p+1})$ by solving LDSP.

If for the specified accuracy ϵ , $z^{p+1} \leq \theta(u^{p+1}) + \epsilon$, then

u^{p+1} is an ε -optimal dual solution. However, if

$z^{p+1} > \theta(u^{p+1}) + \varepsilon$, then a (p+1)st cut, of the form of Equation 36, is added to the master program and the procedure is repeated.

Once the optimal dual solution has been found, the corresponding optimal primal solution has to be determined. First of all, note that the problem is convex and a constraint qualification holds (see BAZARAA & Shetty, 1979, Chapter 6). Therefore no duality gap exists. To obtain the optimal primal solution, suppose that the final master program is of the form LDMP given above. Its dual is

$$\min \sum_{j=1}^J \lambda^j [\theta(u^j) - \sum_{i=1}^I (z_i^j - b_i) u_i^j] = \sum_{j=1}^J \lambda^j h^j \quad (37)$$

$$\begin{aligned} \text{s.t. } \sum_{j=1}^J \lambda^j &= 1 \\ \sum_{j=1}^J \lambda^j (b_i - z_i^j) &\geq 0, \text{ i.e. } \sum_{j=1}^J \lambda^j z_i^j \leq b_i \text{ for } i=1, \dots, I \end{aligned} \quad (38)$$

$$\lambda \geq 0$$

where h^j (see EP in 4.1) is the actual total capital and operating cost of the solution (x_0^j, x^j, z^j) .

Let λ^* be the optimal dual solution, and consider

$$(x_0^*, x^*, z^*) = \sum_{j=1}^J \lambda^{j*} (x_0^j, x^j, z^j).$$

Note that (x_0^*, x^*, z^*) is feasible to the original problem. This follows from the fact that for each j , (x_0^j, x^j, z^j) satisfies

$$x_0^j \geq 0, \quad x^j \geq 0, \quad z^j \geq 0 \quad (39),$$

and

$$x_0^j + \sum_{i=1}^I (z_i^j + x_i^j) \geq P_d - b_0, \quad \sum_{i=1}^I (z_i^j + x_i^j) \geq P_n; \quad (40)$$

hence, Equations 39 and 40 have to hold for (x_0^*, x^*, z^*) as well. Further from Equation 38, $z_i^j \leq b_i$ holds for all $i=1, \dots, I$. Thus (x_0^*, x^*, z^*) is feasible to the original problem. Now, by convexity of $h(\cdot)$,

$$h(x_0^*, x^*, z^*) = h\left[\sum_{j=1}^J \lambda^{j*} (x_0^j, x^j, z^j)\right] \leq \sum_{j=1}^J \lambda^{j*} h^j.$$

But by duality and assumed optimality,

$$\sum_{j=1}^J \lambda^{j*} h^j = z^{p+1} \leq \theta(u^{p+1}) + \epsilon.$$

Thus, $h(x_0^*, x^*, z^*) \leq \theta(u^{p+1}) + \epsilon$.

Since by Lagrangian duality, $\theta(u^{p+1})$ is a lower bound on the original problem optimum, the primal solution (x_0^*, x^*, z^*) is ϵ -optimal.

So it has been shown that the final master program gives the optimal primal solution as a convex combination with its dual variables.

4.5.1 Comparison with the Decomposition Method

The Lagrangian dual method developed above, involves the solution of a linear master program and a capacity planning subproblem. In the master program, an optimal value for the dual vector u^p is found, with the feasible region being bounded by constraints of the type of Eq.36. Then for this u , LDSP is solved with the method described in Chapter 3. If z obtained from the master program is less than the corresponding subproblem's objective value $\theta(u^p)$, then the primal optimal solution can be obtained from the values resulting in this final iteration as described above. Otherwise, a constraint of the type given in Equation 36 is added to the master program.

There is one striking difference between the decomposition and the Lagrangian dual approaches. In the decomposition algorithm, the solar capacity is treated separately from all other variables, i.e. in the master program. The resulting subproblem is a capacity expansion problem not involving solar energy and can be solved by search techniques such as SHERALI's.

The Lagrangian dual method decomposes the problem differently. In the master program, a set of shadow prices for

limited old capacities are optimized and transferred to the subproblem. This subproblem is a plant mix problem that involves solar capacity, but no old capacities, and can be solved with the search technique developed in Chapter 3.

Although both programs use a relatively easy linear program for their master programs, the decomposition method's LP is of a much more trivial nature, involving only two variables. In fact, the search for the optimum is conducted only over x_0 , while in the Lagrangian dual (LD) method a multivariate master program is used, which typically takes more solution time.

For the subproblems, one faces a tradeoff. The decomposition method excludes solar capacity from it, but a capacity expansion problem results. The Lagrangian dual method keeps the solar capacity in the subproblem, but obtains a simpler plant mix problem. Thus the subproblem in the decomposition approach is more difficult than in the LD method. However, it gives a more exact representation of the actual problem, since it considers $z \leq b$. Therefore one can expect less iterations between master program and subproblem in the decomposition approach.

Although the above arguments intuitively favor the decomposition approach, a formal comparison between the two al-

gorithms would have to be conducted empirically. However, due to the apparent disadvantages with the LD method, we merely discuss it as a point of interest; the decomposition approach is implemented in a computer program in the Appendix.

4.6 MULTIYEAR EXPANSION PLANNING MODEL

In multiyear expansion planning, one is confronted with load duration curves that change their peak loads and possibly their shapes over the years. One has to consider inflation and discounting on all costs. In the most general form, a multiyear expansion program involves different capital and operating costs for each year within the planning horizon and different load duration curves over the years in the planning horizon. The decision variables x_{it} needed for such a problem, represent the capacity of equipment type i , installed in year t , $t=1, \dots, T$, where T denotes the number of years in the planning horizon. Associated with the x_{it} are respective annualized capital costs c_{it} .

Another fact that complicates multiyear expansion planning is that a plant will change its position in the merit order over the years, as new, more efficient plant types get installed. This will have an effect on the operating cost.

Aside from inflation and discounting, however, the operating cost term g can be assumed to be constant over the years for our purposes. But more detailed models involving nonlinear operating costs even take into account higher operating costs during the first two to three years after installation.

A mathematical formulation of such a multiyear problem is given below. Note that construction lead times are neglected, but could be accommodated by assuming that $t=1$ is far enough in the future.

$$\begin{aligned} \min \sum_{i=0}^I \sum_{t=1}^T c_{it} D_t x_{it} &+ \lambda \sum_{i=0}^I \sum_{t=1}^T g_{it} K_t \int_{Y_{i-1,t}}^{Y_{it}} F_{dt}(z+x_0) dz \\ &+ (1-\lambda) \sum_{i=1}^I \sum_{t=1}^T g_{it} K_t \int_{Y_{i-1,t}}^{Y_{nt}} F_{nt}(z) dz \\ \text{s.t. } \sum_{i=0}^I (x_{it} + z_{it}) &\geq P_{dt} \quad t=1, \dots, T \\ \sum_{i=1}^I (x_{it} + z_{it}) &\geq P_{nt} \quad t=1, \dots, T \\ x &\geq 0, \quad 0 \leq z \leq b \end{aligned}$$

$$\text{where } Y_{it} = \sum_{j=1}^i (x_{jt} + z_{jt}) \quad , \text{ for } i=1, \dots, I$$

P_{d,n_t} = day or night peak demand in year t

$$K_t = \prod_{j=1}^t (1+m_j) / (1+r)^t$$

$$D_t = 1/(1+r)^t$$

m_t = inflation rate in year t

r = discount rate

b_i = old equipment of type i

λ = fraction of time that solar energy is available

and all other variables are as defined previously.

It may be noted that a program of this type is prohibitively large for any realistic applications. Further, the presence of solar capacities complicates the problem, and renders it hard to solve exactly. Some difficulties due to problem size may be ameliorated by making the following types of simplifying assumptions.

- the inflation rate is constant
- the shape of the load duration curves is constant
(i.e. the curve for year $t+1$ is the curve for year t , multiplied by some factor)
- peak loads grow at a constant rate.

However, no efficient decomposition technique - such as the one developed for the one-year expansion problem - is readily available, due to the complexity of multiyear decisions. This complexity has been sufficiently described in the literature review. With no such decomposition technique at

hand, the presence of the decision variable x_0 , representing solar capacity, prevents the straightforward application of expansion planning algorithms for conventional equipments as described in the literature review.

Therefore consider the following linear approximation of the nonlinear program given above. The linearization is, as before, based on the discretization of the load duration curves. Thus, $p=1, \dots, P$ denote the blocks into which load duration curve t is discretized, with corresponding widths w_{pt} and lengths d_{pt} , for each $t=1, \dots, T$. The energy supplied to load block p in year t by equipment i is denoted by Y_{itp} , with an associated per unit operating cost g_{it} .

$$\begin{aligned} \min & \sum_{t=1}^T \sum_{i=0}^I c_{it} D_t x_{it} + \lambda \sum_{t=1}^T K_t \sum_{i=0}^I \sum_{p=1}^P g_{it} Y_{itp}^d d_{pt}^d \\ & + (1-\lambda) \sum_{t=1}^T K_t \sum_{i=1}^I \sum_{p=1}^P g_{it} Y_{itp}^n d_{pt}^n \\ \text{s.t.} & \sum_{i=0}^I y_{itp}^d = w_{pt}^d \quad t=1, \dots, T ; p=1, \dots, P \\ & \sum_{i=1}^I y_{itp}^n = w_{pt}^n \quad t=1, \dots, T ; p=1, \dots, P \\ & \sum_{p=1}^P y_{itp}^d \leq \sum_{j=1}^t x_{ij} + b_i \quad t=1, \dots, T ; i=0, \dots, I \\ & \sum_{p=1}^P y_{itp}^n \leq \sum_{j=1}^t x_{ij} + b_i \quad t=1, \dots, T ; i=1, \dots, I \\ & x, y \geq 0 \end{aligned}$$

If the simplifying assumptions described above are made, a program about twice the size of ordinary capacity expansion programs results, since twice as many y variables and twice as many constraints are used as in a program not involving solar capacities. Although this is a very large program size for real world applications, too, the solution method - linear programming - presents no problem for this program. Today's computers and their linear programming codes are capable of solving such large scale programs.

However, program size can be reduced and solution time can be cut drastically, if one additional assumption is made. First, assume that forecasts for the load duration curves for all years within the planning horizon are available, a constant inflation rate and constant discount rates are given. The additional assumption is that the optimal expansion plan for the entire planning period is determined, but that all necessary investments are done in year one only. That is, the decision variables x_{it} , denoting purchase decisions for equipment i in year t , are replaced by variables x_i , denoting decisions for equipment i , to be purchased immediately. (In actual implementation, one could delay the purchase of fuel intensive equipment types which

will not be dispatched until the peak load increases sufficiently.) Using this assumption, the multiyear expansion program takes the following form.

$$\min \sum_{t=1}^T \sum_{i=0}^I c_i D_t x_i + \lambda \sum_{i=0}^I \sum_{t=1}^T \bar{g}_i \hat{K}_t \int_{Y_{i-1}}^{Y_i} F_{dt} (z+x_0) dz$$

$$+ (1-\lambda) \sum_{i=1}^I \sum_{t=1}^T \bar{g}_i \hat{K}_t \int_{Y_{i-1}}^{Y_i} F_{nt} (z) dz$$

$$\text{s.t. } \sum_{i=0}^I (x_i + z_i) \geq \max_{t=1, \dots, T} \{P_{dt}\}$$

$$\sum_{i=1}^I (x_i + z_i) \geq \max_{t=1, \dots, T} \{P_{nt}\}$$

$$x \geq 0, \quad 0 \leq z \leq b$$

$$\text{where } Y_i = \sum_{j=1}^i (x_j + z_j), \quad \text{for } i=1, \dots, I$$

$$P_{d,nt} = \text{day or night peak demand in year } t$$

$$\hat{K}_t = [(1+m)/(1+r)]^t / \sum_{j=1}^T [(1+m)/(1+r)]^j$$

$$\bar{g}_i = g_i \left\{ \sum_{j=1}^T [(1+m)/(1+r)]^j \right\}$$

$$m = \text{inflation rate}$$

$$r = \text{discount rate}$$

The factors \hat{K}_t can be interpreted as weighted fractions of time during which the respective load curves occur. Note that they are only dependent on t . Therefore, the T curves for day and night, respectively, can be equivalently treated as one aggregate curve, obtained by letting

$$F_d(\cdot) = \sum_{t=1}^T \hat{K}_t F_{dt}(\cdot), \quad \text{and} \quad F_n(\cdot) = \sum_{t=1}^T \hat{K}_t F_{nt}(\cdot).$$

Thus, the multiyear expansion problem reduces down to a one year expansion problem; the only difference being that the load curves first have to be aggregated. The resulting equivalent one year problem can be solved with the decomposition method described in this chapter.

For long planning horizons, however, this approach does not model reality correctly. The principal limitation is the assumption that all equipment has to be installed immediately. Since equipment that will be needed ten years from now, for example, has to be installed now, the method tends to select the cheapest equipment available for that future demand, so that capital costs do not have too big an impact on the present value function. If gradual installment would be allowed, however, (which is the case in reality), the true optimal solution might require purchase of equipment

with higher capital, but lower operating costs. In other words, the solution to the equivalent one-period problem tends to be less capital intensive and more fuel intensive than for the model which makes more frequent decisions in the planning period.

Chapter V

MANY PERIODS OF DIFFERENT AVAILABILITY OF SOLAR ENERGY

5.1 INTRODUCTION

Up to this point, only two different periods have been considered as far as availability of solar energy is concerned. Under many circumstances, this can be too simple an approach. Given below are a few examples of situations in which more than two periods of differing availability of solar energy may need to be modelled.

When power tower production of electricity is modelled, the plant's output is highly dependent on the intensity of sunshine hitting the area. Up to now, and probably for quite some time in the future, power towers have only been installed in arid areas with a very low probability of cloudcover. However, clouds do move in occasionally even over deserts, rendering the plant useless unless it employs storage devices. In many areas, the probability of cloudcover varies with time of day. For instance during summer in the southwestern deserts, mornings are almost always clear, while frequently rainstorms occur in the afternoon. Thus, one could define differing 'forced outage rates'

(FOR_r) for different times or periods ($r=1, \dots, R$) of day, due to different probabilities of cloudcover. Hence,

$a_r = (1 - FOR_r)$ defines availability of solar energy for the different periods $r=1, \dots, R$.

Another reason for varying intensity of solar energy is the varying angles of incidence of the sun's rays. For power towers this has only small impact since the mirrors are usually designed to track the sun. Nonetheless, the intensity of light varies as the distance through denser atmospheric layers, traversed by the rays, changes. Furthermore, around sunrise and dusk, the sun can be so low that some mirrors block the rays from hitting other mirrors, thus reducing the energy available for conversion into electricity. Hence again, one needs to account for varying availability of solar energy.

The other important area of application of solar energy is hot water and home heating. For this field, no exact capacity of the solar units can be defined. As mentioned in the literature review, the useful capacity for the utility has to be determined via simulation programs. This useful capacity depends on the weather pattern, the utility's load curve, and the household's ability to use the solar energy

produced. The manner in which weather patterns can influence solar output even in desert areas, has been described above. The varying nature of the load curve has been described in the literature review. Also the demand of a single household varies significantly with time. For instance during mornings a high demand can occur, when the family gets up, takes showers, and prepares breakfast. Afternoons will usually have lower demand; but probably in early evening, when dinner is cooked and heaters are turned on, the demand peaks once again.

When this fluctuating demand interacts with the varying output of the solar energy unit, quite different capacity savings for the utility can result, depending statistically on the time of day. Again, this situation can be modelled by incorporating more than two periods with differing availability of solar energy.

A third possible application of modelling varying availability of solar capacity lies in storage modelling. For power towers for instance, storage devices are often only sufficient to supply part of the day's power output, over only a portion of the night. In this case, at least three periods of different availability have to be accounted for:

day with full power output, evening with reduced power output produced from storage, and night with no power output. If relatively small storage devices are installed in solar home and hot water heating systems, a similar situation results for household modelling.

For all applications mentioned, an example of the derivation of the availability coefficients is given in Figure 15

5.2 FORMULATION

In this chapter, only the plant mix problem will be treated; i.e. it will be assumed that no old capacity is present. The capacity expansion problem can be treated similarly as the plant mix problem; it will not be detailed here.

Let $r=1, \dots, R$ denote different sub-periods of constant capacity availability of solar output, in the sense described in the previous section. Let $F_r(\cdot)$, $r=1, \dots, R$, denote the inverse load duration curves for these sub-periods, with respective peak loads P_r .

$$F_r(z) = \begin{cases} f_r^{-1}(z) & \text{for } 0 \leq z \leq P_r \\ 0 & \text{for } P_r \leq z \leq \hat{P} \end{cases}$$

where \hat{P} denotes the overall peak load

Further, let a_r denote availability of solar energy in per-

iod r . Finally, let θ_r denote the fraction of time that solar energy is available with factor a_r , that is, the fractional length of sub-period r .

Hence, $\sum_{r=1}^R \theta_r = 1$, $\theta_r \geq 0$, for $r=1, \dots, R$.

Observe that conceptually, the quantities θ_r play the role of probabilities. Consequently, the resulting program is very similar to the one analysed in Chapter 3, and also to the type of program analysed in SHERALI et al, 1981b.

Program for Many Periods (MPP)

$$\begin{aligned} \min h(x_0) = & c_0 x_0 + \sum_{r=1}^R g_0 \theta_r \int_0^{a_r x_0} F_r(z) dz + \sum_{i=1}^I c_i x_i \\ & + \sum_{i=1}^I g_i \sum_{r=1}^R \int_{Y_{i-1}}^{Y_i} \theta_r F_r(z + a_r x_0) dz \\ \text{s.t. } \sum_{i=1}^I x_i = & \max_{r=1, \dots, R} \{P_r - a_r x_0\} \\ x \geq & 0 \end{aligned}$$

where all variables are as defined before, in particular

$$Y_i = \sum_{j=1}^i x_j, \quad i=1, \dots, I$$

For a fixed x_0 , this program can be interpreted as a plant mix program involving only conventional equipment. The different load curves F_r associated with different prob-

abilities of occurrence θ_r can be treated via horizontal expectation as described in SHERALI et al, 1981b. That is, given x_0 , a single load curve

$$F(z, x_0) = \sum_{r=1}^R \theta_r F_r(z + a_r x_0)$$

is used as an aggregate equivalent load duration curve. In other words, as shown in Chapter 3, each inverse load duration curve is moved to the left by the amount $a_r x_0$, scaled, and then the resulting curves are summed up. Thus for a fixed x_0 , this program can be easily solved via breakeven analysis. Therefore the whole program can be viewed as a function of only x_0 , with the optimal levels of other capacities being determined for each value of x_0 (analogous to the approach in Chapter 3). If this optimal value function of x_0 , $h(x_0)$, could be proven to be convex, efficient algorithms of the type described in Chapter 3 could be applied to find the optimal solution. In the next section, convexity of $h(x_0)$ will be shown using arguments in general vector notation, similar to the formulation given in section 4.3.6.

5.3 CONVEXITY

To show convexity of $h(x_0)$, first program MPP is decomposed into a master program MPMP involving only x_0 , and a subproblem MPSP. Then, a discretized version of MPSP will be stated. Examination of its general formulation will lead to the proof of convexity for $h(x_0)$.

The master program MPMP takes the following form:

$$\min h(x_0) = c_0 x_0 + \theta(x_0) \quad (39)$$

$$0 \leq x_0 \leq \hat{P}$$

where $\theta(x_0)$ denotes the optimal solution of the resulting subproblem for a fixed x_0 , and \hat{P} denotes the overall peak load ($\max_{r=1, \dots, R} \{P_r\}$). Hence

MPSP

$$\begin{aligned} \theta(x_0) = \min & \sum_{r=1}^R g_r \theta_r \int_0^{a_r x_0} F_r(z) dz + \sum_{i=1}^I c_i x_i \\ & + \sum_{i=1}^I q_i \sum_{r=1}^R \int_{Y_{i-1}}^{Y_i} [\theta_r F_r(z + a_r x_0)] dz \\ \text{s.t. } & \sum_{i=1}^I x_i = \max_{r=1, \dots, R} \{P_r - a_r x_0\} \\ & x \geq 0 \end{aligned}$$

Consider the following discretized version of MPSP, labelled DMPSP.

$$\begin{aligned}
\theta(x_0, \Delta) &= \min \sum_{i=1}^I c_i x_i + \sum_{i=0}^I g_i \sum_{p=1}^P \sum_{r=1}^R y_{ip}^r d_p^r \theta_r \\
\text{s.t.} \quad &\sum_{i=0}^I y_{ip}^r \geq \Delta \quad p=1, \dots, P \quad ; \quad r=1, \dots, R \\
&-\sum_{p=1}^P y_{ip}^r + x_i \geq 0 \quad i=1, \dots, I \quad ; \quad r=1, \dots, R \\
&-\sum_{p=1}^P y_{0p}^r \geq -a_r x_0 \quad r=1, \dots, R \\
&x, y \geq 0
\end{aligned}$$

where x , c , g are as defined previously, and

d_p^r = length of block p in sub-period r

θ_r = fractional length of sub-period r

a_r = availability of solar capacity in sub-period r

Δ = standardized demand block width (assuming each P_r is some integral multiple of Δ)

y_{ip}^r = load in block p served by equipment type i , in sub-period r

P_r = maximal index of load block containing the peak demand in sub-period r

P = maximal index of load block containing the overall peak demand

Note that as $\Delta \rightarrow 0$, DMPSP will approach the original nonlinear subproblem, and $\theta(x_0, \Delta) \rightarrow \theta(x_0)$.

THEOREM 4

The function $h(x_0)$ as stated in Equation 39 is convex.

Proof

For a fixed Δ , the function $\theta(x_0, \Delta)$ is convex in x_0 since the optimal value function of a minimization linear program is convex in its right hand side. Consequently, since

$\theta(x_0) = \lim_{\Delta \rightarrow 0} \theta(x_0, \Delta)$, it follows that $\theta(x_0)$ is convex in x_0 . Noting Equation 39, this implies that $h(x_0)$ is convex in x_0 , and the proof is complete.

5.4 ALGORITHM

Since $h(x_0)$ is convex, the same type of algorithm as in Chapter 3 can be applied. Expressions for the first and second derivative of $h(x_0)$, whenever they exist, will be needed for the algorithm, and are derived in the following subsection.

5.4.1 Derivatives of $h(x_0)$

Let the capacities of all conventional equipments, x_i , $i=1, \dots, I$, be implicitly fixed at their optimal level for each value of x_0 . Then, the function $h(x_0)$ becomes

$$h(x_0) = c_0 x_0 + \sum_{r=1}^R g_r \int_0^{a_r x_0} \theta_r F_r(z) dz + \sum_{i=1}^I c_i x_i$$

$$+ \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} \left[\sum_{r=1}^R \theta_r F_r(z + a_r x_0) \right] dz$$

For the sake of simplicity, we will first obtain an expression for $h'(x_0)$ whenever it exists. Nondifferentiable points will subsequently be identified. Note that the first derivative of $h(x_0)$ involves the term $\sum_{i=1}^I c_i x_i'$. From breakeven analysis on the aggregate curve one obtains, (compare with Equation 6 and the subsequent development in Chapter 3)

$$\sum_{i=1}^I c_i x_i' = c_I Y_I' + \sum_{i=1}^{I-1} Y_i'(x_0) (g_i - g_{i+1}) \left[-\sum_{r=1}^R \theta_r F_r'(z + a_r x_0) \right].$$

As derived graphically in Figure 16, the expression for Y_i' is

$$\frac{\delta Y_i}{\delta x_0} = \frac{-\sum_{r=1}^R \theta_r a_r F_r'(z + a_r x_0)}{\sum_{r=1}^R \theta_r F_r'(z + a_r x_0)} > -1$$

Using these facts, one obtains for $h'(x_0)$

$$\begin{aligned} h'(x_0) &= c_0 + \sum_{r=1}^R (g_0 - g_1) a_r F_r'(a_r x_0) + c_I Y_I' \\ &\quad + \sum_{i=1}^{I-1} Y_i' (g_i - g_{i+1}) \left[-\sum_{r=1}^R \theta_r F_r'(Y_i + a_r x_0) \right] \\ &\quad + \sum_{i=1}^{I-1} (g_i - g_{i+1}) \sum_{r=1}^R \theta_r F_r'(Y_i + a_r x_0) (Y_i' + a_r) \\ h'(x_0) &= c_0 + c_I Y_I' + \sum_{i=0}^{I-1} (g_i - g_{i+1}) \sum_{r=1}^R a_r \theta_r F_r'(Y_i + a_r x_0) \quad (40) \end{aligned}$$

Note that $h'(x_0)$ exists whenever Y_I' is defined. However, Y_I' varies discontinuously. It takes the value $-a_r$, when for the current value of x_0 , $P_r - a_r x_0$ is the overall peak load. Thus differentiability only applies to pieces of the total curve.

Examining the terms in Equation 40 one finds that Y_I' is a constant and all other terms are differentiable at all points; therefore $h''(x_0)$ exists whenever $h'(x_0)$ exists. At such points, taking the second derivative yields

$$h''(x_0) = \sum_{i=0}^{I-1} (g_i - g_{i+1}) \sum_{r=1}^R a_r \theta_r F_r'(Y_i + a_r x_0) (Y_i' + a_r) \quad (41)$$

The expressions for $h'(x_0)$ given in Equation 40, and for $h''(x_0)$, given in Equation 41, are used in the algorithm described below.

5.4.2 Algorithm Statement

First, recall from above that Y_I' varies discontinuously, and that therefore $h(x_0)$ displays a kink where

$$P_r - a_r x_0 = P_s - a_s x_0, \text{ for some } r, s \in \{1, \dots, R\}.$$

Thus, the algorithm first has to determine for which r , $(P_r - a_r x_0)$ will be the overall peak load at the optimum. This is done by splitting the range of search over x_0 into different regions.

Many-Period Algorithm MPA

Step 1: Ranges of x_0 for which the different periods display the overall peak load, can be determined easily as shown in Figure 17. Then, $h'(x_0)$ is calculated according to Equation 40 at points just left and right of those x_0 -values where two periods have the same peak load, i.e. at points x_0 such that for load curve pairs $r, s \in \{1, \dots, R\}$,

$$P_r - a_r x_0 = P_s - a_s x_0. \quad (42)$$

Let these x_0 values be called mesh values. To the left and right of each mesh value, $h'(x_0)$ is calculated, using Equation 40 and the appropriate values for Y'_I corresponding to periods r and s . If the two values for $h'(x_0)$ have different signs, the x_0 value solving Equation 42 for the periods r and s currently examined is optimal. Otherwise, proceed to step 2.

Step 2: The period r has to be found for which $h'(x_0)$ is negative at the left mesh point and positive at the right mesh point of the range for which period r has the overall peak load. Then, the optimal x_0 lies between these two mesh values. Once such a range for the optimal x_0 has been determined, the optimum can be sought for using Newton's method as described in Chapter 3, with the appropriate steps being taken to ensure convergence.

Step 3: The first and second derivatives of $h(x_0)$ at the current x_0 are calculated according to Equations 40 and 41. Terminate, if $h'(x_0)$ is less than some specified accuracy $\epsilon > 0$. Otherwise, approximating the function $h'(x_0)$ by a straight line with slope $h''(x_0)$, the root of $h'(x_0)$ is estimated. With this root as the new current x_0 , step 3 is repeated.

To ensure convergence, the measures already described in the algorithm section of Chapter 3 have to be taken.

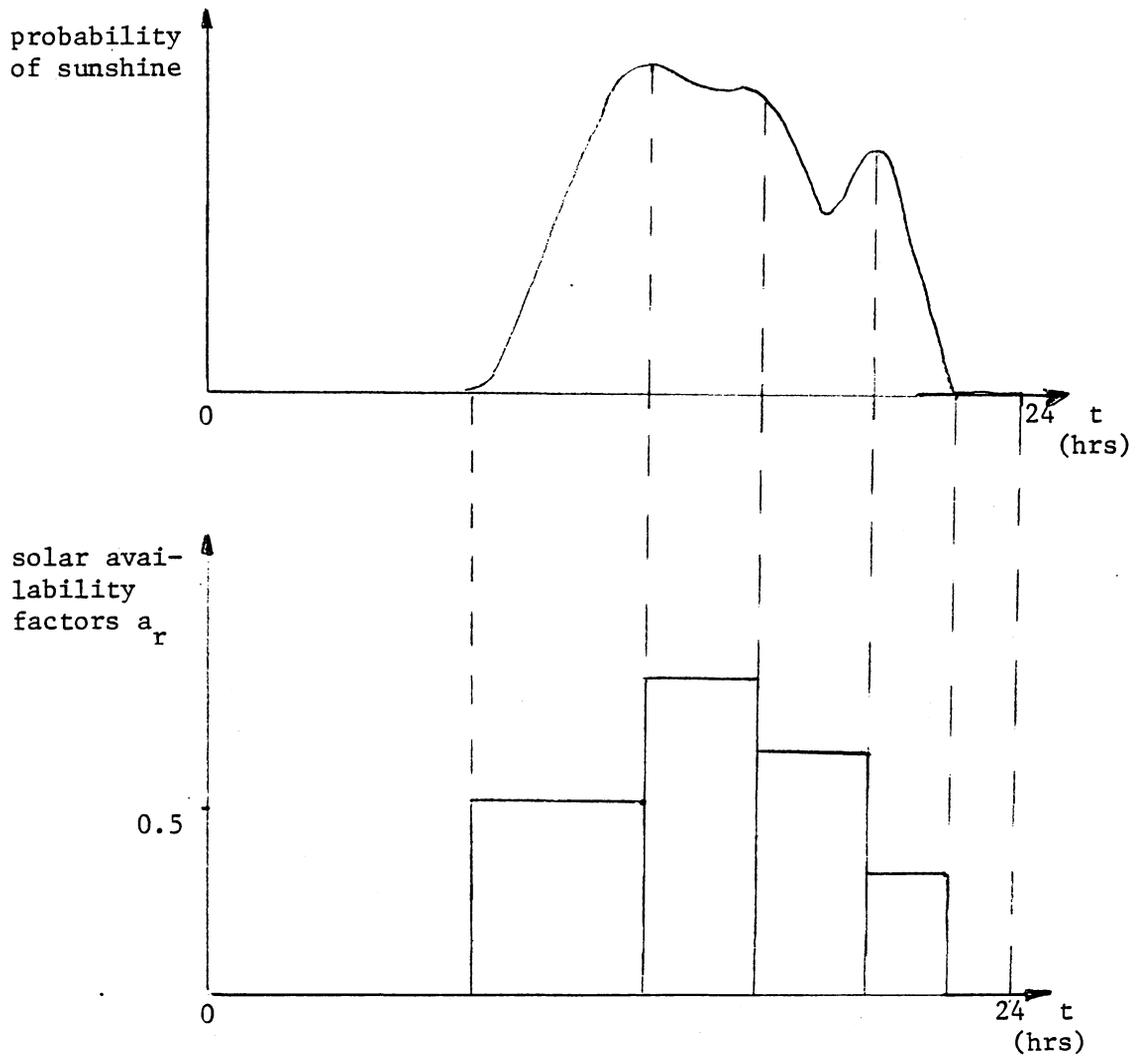


Figure 15: Varying Availability of Solar Energy

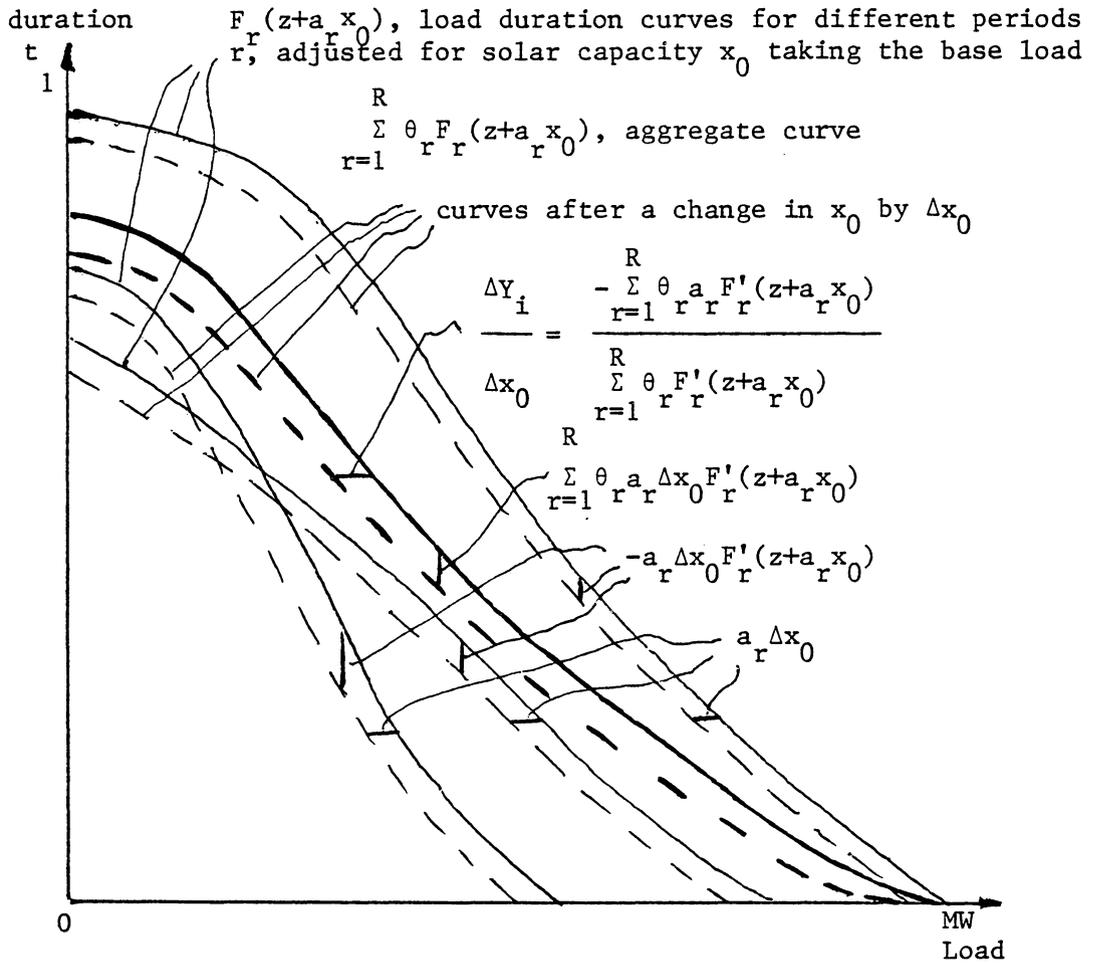


Figure 16: Derivative of Y with Respect to x_0

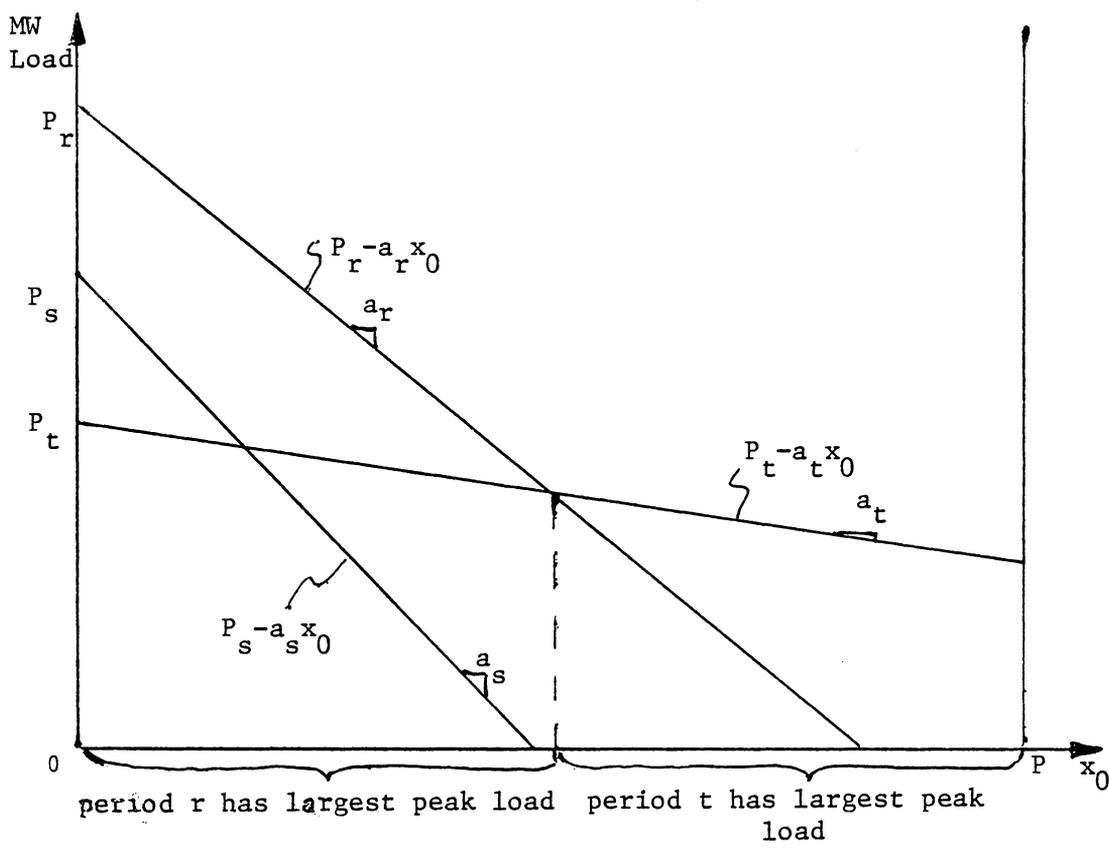


Figure 17: Ranges of x_0 With Different Overall Peak Loads

5.5 INCLUSION OF STORAGE FOR SOLAR ENERGY

As described above, storage can be modelled quite easily using different availability coefficients for different times of day. This approach models power tower systems realistically; current installations produce full output during day time, a constant reduced output for a few hours into the night, and are shut down during the rest of the night. For home heating systems, the approach is even more appropriate, since the interaction of the solar system, the utility and the household cause the effective capacity value of the solar system to vary with time.

However, if storage technology progresses such that it becomes possible for power towers to keep energy stored for longer parts of the night, the following approach to storage modelling could become more appropriate. Note, however, that the modelling approach to be described will be of no practical significance until penetration rates of solar energy into electric utility systems is very high. With the common structure of demand during night time - decreasing as the evening progresses, and not increasing before early morning -, stored solar energy would be optimally dispatched in the evening, anyway, unless the solar capacity is so large that no conventional capacity is needed to serve the

evening load. The following modelling approach is therefore given as a matter of interest only.

Assume that heat losses during one night are so small that it does not matter if the stored energy is used immediately after dusk or later in the night. Then the problem of when to dispatch the stored energy is essentially the same as the optimal dispatch problem for hydro plants with reservoir.

The optimal strategy is to dispatch the near zero operating cost stored energy at the time when it replaces most expensive conventional energy. Thus it has to be dispatched during times of peak demand, when it can replace fuel-intensive peaking units. Therefore, hydro plants with storage are dispatched such that the stored energy is exactly equal to the integral of the inverse load duration curve between the load at which hydro dispatching starts and this load plus the hydro plant's capacity. This way both the full capacity and all the stored energy are used. It can be easily shown that this standard hydro dispatching rule is indeed optimal.

5.5.1 Optimal Hydro Dispatching

In this section, we will demonstrate that at optimality, the hydro storage unit dispatches its stored energy at full capacity. The outline of the proof is the following. Consider two cases for an optimal solution, depending on whether the hydro plant replaces just one or two conventional equipment types, as shown in Figure 18. That is, case 1 occurs when the hydro storage is dispatched between two consecutive breakeven points; case 2 occurs when it is dispatched across some breakeven point. In either case, wasting energy by not using the full amount stored will not be optimal (trivial). Consider the case that the hydro energy only replaces one equipment, i.e. that its dispatching area is wholly contained in a conventional equipment's dispatching area. Then, dispatching the same amount of energy at less than full capacity (i.e. running it at higher loads) will save the same amount of conventional fuel, but more conventional capacity will be needed. Therefore it is optimal to dispatch hydro energy at full capacity.

Now consider the case that the hydro plant's dispatching position is exactly between two conventional plants (see Figure 18 for a visualization). The objective is to maximize savings, i.e.

$$\max f(y_1) = (x-y_1)c_1 + (y_2-x)c_2 + g_1 \int_{y_1}^x F(z)dz + g_2 \int_x^{y_2} F(z)dz \quad (43)$$

Assume that the hydro plant is not dispatched at full capacity. Then, the derivative of the objective with respect to y_1 represents the rate of change in objective value with respect to an increase in the capacity at which the hydro plant is used.

Expressing the stored energy, E_0 , through the area of the inverse load duration curve it serves,

$$E_0 = \int_{y_1}^{y_2} F(z)dz,$$

and taking derivatives yields

$$\delta E_0 / \delta y_1 = 0 = F(y_2)(\delta y_2 / \delta y_1) - F(y_1)$$

$$\text{or } \delta y_2 / \delta y_1 = F(y_1) / F(y_2) > 1$$

Using this and taking the derivative of the objective function of Equation 43, yields

$$\delta f / \delta y_1 = -c_1 + c_2 F(y_1) / F(y_2) - g_1 F(y_1) + g_2 F(y_1) \quad (44)$$

Since the mesh point between equipment one and two has been determined by breakeven analysis one can write

$(g_2 - g_1) = (c_1 - c_2)/F(x)$, where x denotes that mesh point. Thus Equation 44 becomes

$$\delta f / \delta y_1 = c_1 [-1 + F(y_1)/F(x)] + c_2 [F(y_1)/F(y_2) - F(y_1)/F(x)]$$

Since $F(y_1) > F(x) > F(y_2)$, this derivative is greater than zero. Therefore, in this case, too, it is optimal to dispatch a hydro plant at full capacity.

Note that this derivation assumes totally deterministic data. But in hydro scheduling, uncertainties about the river flow and about the actual load at any time are involved. Therefore for actual hydro generation scheduling, much more involved algorithms are used, mostly of the dynamic programming type.

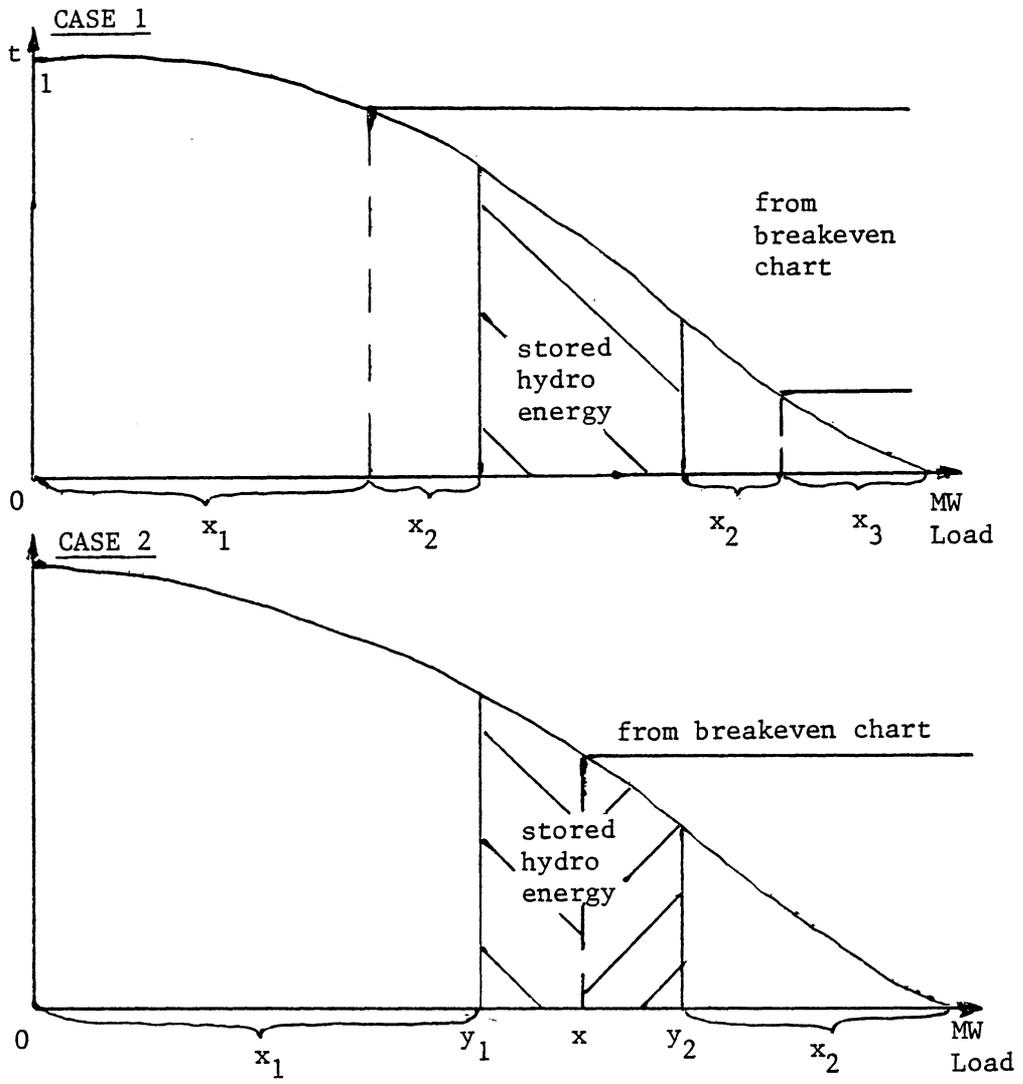


Figure 18: Dispatching of Hydro Units with Storage

5.5.2 Optimal Storage Policy for Solar Installations

For power tower modelling, under the condition that it does not matter if the stored energy is used immediately after dusk or later during the night, the solar storage dispatching problem is equivalent to the hydro dispatching problem. Thus the storage has to be dispatched such that the area cut out of the night load duration curve exactly equals the amount of stored energy. Here it is assumed that the energy drawn from storage is converted to electricity by the same turbine generator system as used during the day time. Due to heat losses from storage or other technological difficulties it may, however, be the case that the maximal power output is less than during day time. This can simply be modelled by introducing a factor $\eta < 1$ as a multiplier for the maximal available capacity at night. Thus, the solar capacity is denoted by x_0 at day time and by ηx_0 at night time.

Then, the plant mix problem with solar energy and storage takes the form given below. Note that this formulation assumes that for each x_0 , the optimal dispatching position of the storage is implicitly determined. The main differences between the programs with and without storage are the following. First, terms representing capital and operating

costs of the storage are included in the program involving storage. Also, the operating cost terms of conventional equipment are modified, due to the fact that the storage effectively cuts out part of the night load duration curve. This is represented by replacing the load curve function $F_n(z)$, used in all previous programs, by the curve $\hat{F}_n(z, x_0)$. This modified night load curve is constructed as shown in Figure 19, for any given x_0 .

$$\begin{aligned} \min & c_0 x_0 + E_0 (c_s + c_0 / \lambda) + g_0 \lambda \int_0^{x_0} F_d(z) dz + (1-\lambda) \eta g_s E_0 \\ & + \sum_{i=1}^I c_i x_i + \sum_{i=1}^I g_i \left\{ \int_{Y_{i-1}}^{Y_i} [\lambda F_d(z+x_0) + (1-\lambda) \hat{F}_n(z, x_0)] dz \right\} \\ \text{s.t. } & \sum_{i=1}^I x_i \geq \max \{ P_d - x_0, P_n - \eta x_0 \} \\ & x \geq 0 \end{aligned}$$

$$\text{and } E_0 = (1-\lambda) \int_D^{D+\eta x_0} F_n(z) dz = \text{stored energy}$$

D = load at which storage gets dispatched

c_0 = capital cost for solar equipment

E_0 / λ = solar capacity diverting energy to storage during day time

c_s, g_s = capital and operating cost for storage

$$\hat{F}_n(z, x_0) = \begin{cases} F_n(z) & 0 \leq z \leq D \\ F_n(z + \eta x_0) & D \leq z \leq P_n - \eta x_0 \\ 0 & \text{otherwise} \end{cases}$$

All other variables are as defined in previous chapters.

For fixed values of x_0 and E_0 , the night inverse load duration curve is modified by cutting a piece out at the appropriate place, and the day curve is moved to the left. An aggregate load duration curve can be obtained by summing the modified day and night inverse load duration curves, as done in Chapter 3. On this aggregate load duration curve, a breakeven analysis can be performed to determine the optimal conventional capacities for the fixed value of x_0 (as shown in Figure 20). Thus, the total cost function can be viewed as a function in only two variables, x_0 and E_0 , the solar and the storage capacity.

To find the minimum of this function, a two-dimensional search can be performed. Note however that due to the part of the load duration curve cut out due to the storage, neither the night nor the aggregate curves are continuous. Therefore, the optimal value function (as a function of x_0 and E_0) is hard to characterize. Due to this difficulty, no efficient search techniques of the optimal gradient type can be applied. The optimum can be found by systematically searching the part of the feasible region in which the optimum is believed to lie in, using any standard multi-dimen-

sional, non-derivative-based, nonlinear programming search technique. For descriptions of such techniques, see for example SIMMONS, 1975. Convergence of such a method to an optimal solution is not guaranteed for this problem; at best, one can hope for a good quality solution at termination.

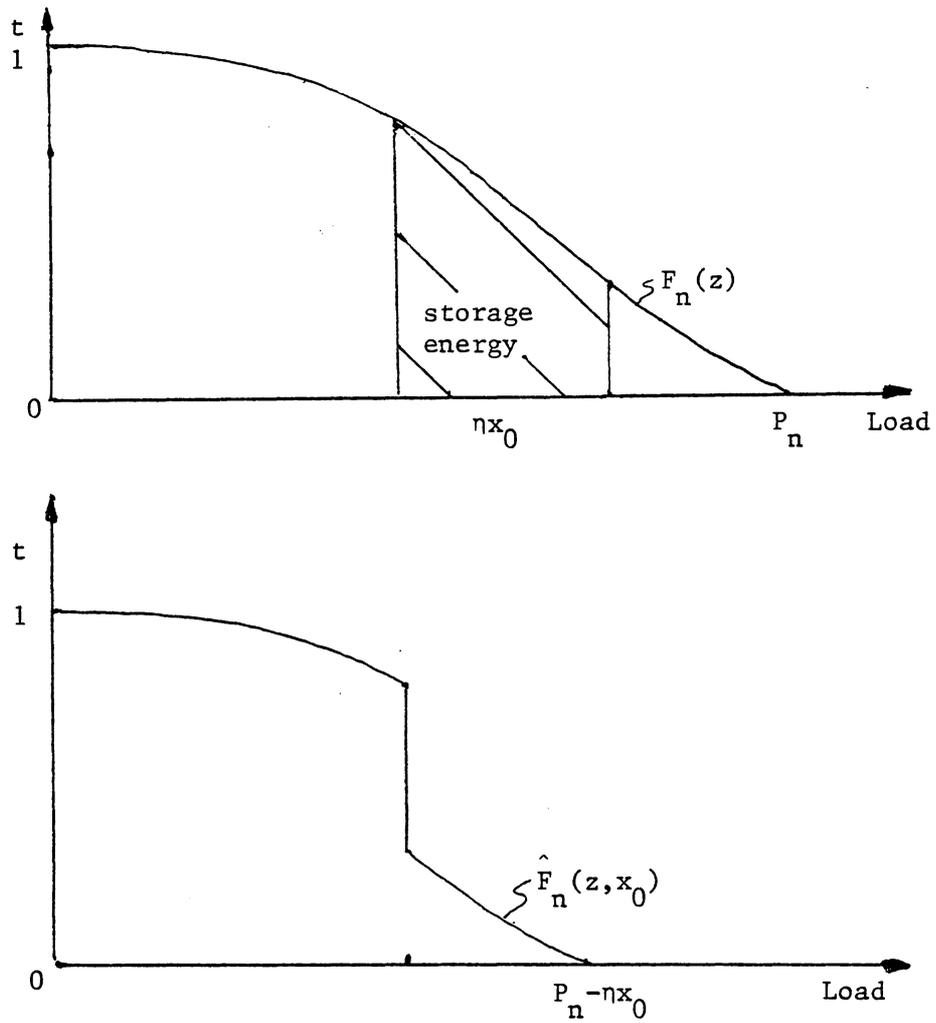


Figure 19: Night Inverse Load Duration Curve with Storage

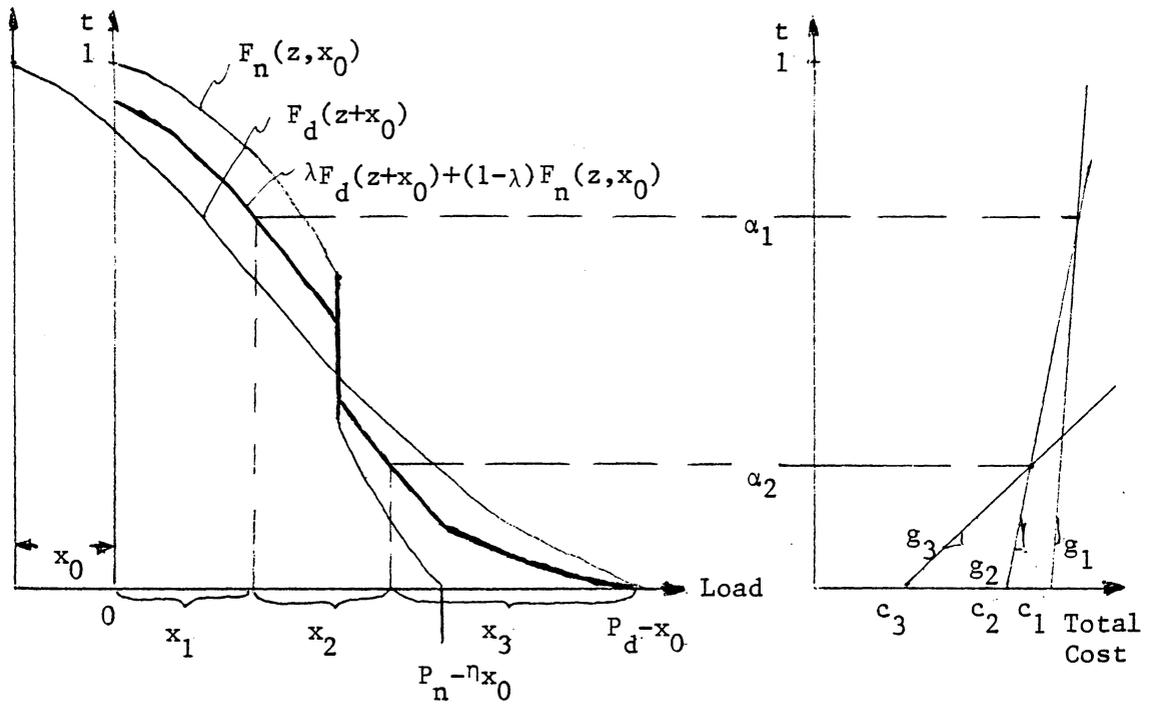


Figure 20: Optimal Plant Mix with Solar Energy and Storage

Chapter VI

CONCLUSIONS

In this thesis, methods have been developed to incorporate solar energy devices into electric utility capacity planning. Due to the non-availability of solar energy during night time, planning algorithms that are used today are not appropriate for treating solar energy installations.

Methods have been developed for finding the optimal plant mix and expansion plan. These models advance the state-of-the-art in the field of solar energy - electric utility interactions. Previous research has analyzed the impact of certain capacities of solar energy systems on an electric utility. For the most part, simulation models have been used, which by their nature, permit one to investigate only a limited number of investment levels for solar energy.

By contrast, this research analyzes how much solar energy capacity would be advantageous for an utility. For the plant mix and the one period expansion problem, convexity properties have been established that permit an implicit look at the whole feasible investment range. Efficient algorithms have been developed to find the optimal capacities of both conventional and solar equipment. These algorithms

have been implemented in Fortran computer codes; some computational results have been obtained.

A more general model that accounts for solar energy non-availability and variability, has also been developed. It models periods of different availability of solar capacity, by assuming a certain availability factor associated with different periods. This approach bears similarity with the derating method used in power system planning to account for forced outages. A more accurate and also more elegant approach to model forced outages is the equivalent load duration curve method. It is left for further research to modify the algorithm for different availability of solar capacity, so that it works with the equivalent load duration curve technique.

Appendix A

ALGORITHM FOR THE EXPANSION PROBLEM WITHOUT SOLAR ENERGY

In this appendix, the capacity expansion algorithm developed in SHERALI, 1982a, is described. This algorithm is used in Chapter 4 of this thesis to solve subproblems not involving solar capacity as a decision variable.

The algorithm involves two preliminary phases (Ia and Ib) in which the optimal capacities of all old equipment and some new equipment are determined. The final phase II then employs a steepest descent feasible directions algorithm. Phases Ia, Ib and II are described in more detail below.

A.1 PREDETERMINATION OF OLD CAPACITY USED AT OPTIMALITY

In this phase, the fact is used that if $z_k > 0$, then $z_1 = b_1, \dots, z_{k-1} = b_{k-1}$ must be true; and if $z_k < b_k$, then $z_{k+1} = \dots = z_I = 0$, and $x_k = \dots = x_I = 0$ must hold. This results from the assumption that old equipment has zero capital costs and its operating costs satisfy $g_1 < g_2 < \dots < g_I$. Thus, in order to find the optimal values of all z variables, it is enough to determine some z_k for which $0 < z_k < b_k$ holds, if there is any such equipment type k . Else, $z_i = b_i$ will hold for all i .

In the phase Ia algorithm to be cited below, the following additional notation is needed (for some index $1 \leq k \leq I$):

$f(\cdot)$ = load duration curve function

$f(0)$ = peak demand

$F(\cdot)$ = inverse load duration curve function

$$B_k = \sum_{j=1}^{k-1} z_j^*$$

$J(i)$ = index in a list of candidates for the purchase of new equipment which is less than i , and which gives the smallest breakeven point $\alpha \geq 0$ via the equation $g_i^\alpha = c_{J(i)} + \alpha g_{J(i)}$.

$$\gamma_i = \begin{cases} c_{J(i)} / (g_i - g_{J(i)}) & , \text{ if } J(i) \text{ exists} \\ 1 & , \text{ otherwise} \end{cases}$$

Then, the following algorithm is given in SHERALI, 1982a:

Initialization: Set $z_1^* = \min \{b_1, f(0)\}$, and $B_2 = z_1^*$, $k=1$.

Termination Step: If $B_{k+1} = f(0)$, stop; the problem is solved. If it is known that $z_k^* < b_k$, terminate; proceed to phase Ib. If $k = I$, terminate; it is known that $z_i^* = b_i$, $i=1, \dots, I$. Otherwise, increment k by one and proceed to the main step.

Main Step: If $b_k = 0$, set $z_k^* = 0$ and $B_{k+1} = B_k$ and return to the termination step. Otherwise, compute γ_k . If $\gamma_k \geq 1$, set $z_k^* = \min \{b_k, f(0) - B_k\}$. Let $B_{k+1} = B_k + z_k^*$ and return to the termination step.

On the other hand, if $\gamma_k < 1$, $b_k > 0$, check if $b_k > \min \{f(0) - B_k, f(0) - f(\gamma_k)\}$. If it is, then return to the termination step with the indication that $z_k^* < b_k$.

If $\gamma_k < 1$, $b_k > 0$ and $b_k \leq \min \{f(0) - B_k, f(0) - f(\gamma_k)\}$, then compute $\hat{z}_k, \hat{x}_1, \dots, \hat{x}_{k-1}$ and $\hat{c}_i, i=1, \dots, k-1$ as follows:

$$\hat{z}_k = b_k, \quad \hat{x}_1 = f(0) - \sum_{i=1}^k b_i, \quad \hat{x}_i = 0 \quad (i=2, \dots, k-1);$$

$$\hat{c}_k = 0, \quad \hat{c}_i = \hat{c}_{i+1} + (g_{i+1} - g_i) F\left[\sum_{j=1}^i (\hat{x}_j + b_j)\right], \quad i=k-1, k-2, \dots, 1.$$

Then, $z_k^* < b_k$ if any $\hat{c}_i > c_i$ for $i \in \{1, \dots, k-1\}$, and $z_k^* = b_k$ if $\hat{c}_i \leq c_i$ for $i=1, \dots, k-1$. If $z_k^* < b_k$, return to the termination step. If $z_k^* = b_k$, set $B_{k+1} = B_k + z_k^*$ and return to the termination step.

A.2 PREDETERMINATION OF SOME OPTIMAL NEW CAPACITIES

In this phase, some x-variables get fixed at their optimal values. This is achieved by using the breakeven chart for new equipments to find the operating ranges under which the total cost for respective equipment types is minimal. Then,

summing the optimal z_i^* determined in phase Ia, from equipment type 1 up and from equipment type I down, and taking into consideration that only new equipment types with index less than some k determined in phase Ia can be economical to install, upper and lower bounds for the optimal x_i^* are derived.

The algorithm for phase Ib is again cited from SHERALI, 1982a.

Let the range of duration values α for which equipment i has the cheapest total cost be $[\alpha_i^l, \alpha_i^h]$. Projecting this onto the load duration curve, a corresponding load interval $[L_i^l, L_i^h]$ with $L_i^l = f(\alpha_i^h)$ and $L_i^h = f(\alpha_i^l)$ is obtained. Let $L_i = L_i^h - L_i^l$.

For each new equipment type i considered for purchase, compute L_i , and set

$$L_i^+ = \text{length of } [L_i^l, L_i^h] \cap \left[\sum_{j=1}^{i-1} z_j^*, f(0) - \sum_{j=i+1}^I z_j^* \right].$$

Define L_{i1} and L_{i2} as follows:

$$L_{i1} = \begin{cases} \min \{ L_j^l - \sum z_m^* \} & \text{if such an } i \text{ exists} \\ j > i, L_j^+ > z_j^* & i < m < j \\ f(0) - \sum_{j=i+1}^k z_j^* & \text{otherwise} \end{cases}$$

$$L_{i2} = \begin{cases} \max_{j < i, L_j^+ > z_j^*} \{L_j^h + \sum_m z_m^*\} & \text{if such an } i \text{ exists} \\ \sum_{j=1}^{i-1} z_j^* & \text{otherwise} \end{cases}$$

If $L_i^+ \leq z_i^*$, set $x_i^* = 0$. For each considered equipment i for which $L_i^+ > z_i^*$, compute

$$L_i^- = \text{length of } [L_i^l, L_i^h] \quad [0, L_{i1}] \quad [L_{i2}, \infty].$$

If $L_i^- = L_i$, then $x_i^* = \max \{0, L_i^- - z_i^*\}$.

If $L_i^- - z_i^* > 0$, then $x_i^* \geq L_i^- - z_i^* > 0$. Record these lower bounds and variable indices. Check the following three conditions:

Let p and q be the first and last elements of the set J of candidate new equipments, ordered in increasing operating cost values.

If p and $p+1 \in J$, and if $x_{p+1}^* > 0$, then

$$x_p^* = \max \{0, L_p - \sum_{i=1}^p z_i^*\}.$$

If q and $q-1 \in J$, and if $x_{q-1}^* > 0$,

$$\text{then } x_q^* = \max \{0, L_q - \sum_{i=q}^I z_i^*\}.$$

If $\{r-1, r, r+1\} \in J$, and if $x_{r-1}^* > 0$, $x_{r+1}^* > 0$, then

$$x_r^* = \max \{0, L_r - z_r^*\}.$$

At the end of the phase, let

$$J^* = \{i \in J: x_i \text{ is not as yet fixed at some } x_i^*\}.$$

If $J^* = \emptyset$ or if the fixed variable capacities add up to $f(0)$, then the problem is solved. Otherwise, find an advanced starting solution and continue with phase II.

An advanced starting solution is found in the following way.

Pick the smallest index in J^* . With the other equipment capacities selected thus far being loaded in merit order from base upwards (including z_r^*) and from peak downwards, find the exposed part of strip L_r . Select x_r to fill this exposed part. Repeat with the other indices in J^* sequentially.

A.2.1 Steepest Descent Feasible Directions Algorithm

The algorithm described in SHERALI, 1982a, is a convergent specialization of Zoutendijk's steepest descent feasible directions approach (see SIMMONS, 1975, Chapter 8). Suppose some feasible solution $\bar{x} = (\bar{x}_1, \dots, \bar{x}_I)$ is given. Note that all z_i and some x_i are already fixed at their optimal values so that the only variables left in the problem are x_i ,

$i \in J^*$:

$$\min \sum_{i \in J^*} c_i x_i + \sum_{i=1}^I g_i \int_{Y_{i-1}}^{Y_i} F(z) dz$$

$$\text{s.t. } \sum_{i=1}^I (z_i + x_i) \geq P$$

$$z_i = z_i^*, \quad i=1, \dots, I \quad ; \quad x_i = x_i^*, \quad i \notin J^*$$

$$x_i \geq 0, \quad i \in J^*.$$

Then, the three steps described below are taken to arrive at an improved solution. The algorithm will converge finitely.

Step 1. Evaluate the gradient of the objective function at x .

$$\nabla_i = c_i + \sum_{j=i}^{I-1} (g_j - g_{j+1}) F(X_j), \quad i \in J^*.$$

Step 2. Find an improving, feasible direction (\bar{d}_i) if one exists. Otherwise, terminate with \bar{x} as the optimal solution. A steepest descent feasible direction is found via solving a trivial linear knapsack problem (knapsack problems are for instance described in BAZARAA & Jarvis, 1977).

Step 3. Take a suitable step length $\bar{\mu}$ along direction \bar{d} , so that $x_{i_{\text{new}}} = \bar{x} + \bar{\mu} \bar{d}_i$, for all $i \in J^*$. The step length advocated for rapid convergence of the method tends to obtain the optimality Kuhn Tucker conditions for the expansion problem. This step length is given by

$$\mu^* = \frac{\sum_{i \in J^*} (\nabla_i - D_i)(D_i' - \nabla_i')}{\sum_{i \in J^*} (D_i' - \nabla_i')^2}$$

where ∇_i' denotes the rate of change of ∇_i , and D_i and D_i' denote average values of ∇_i and ∇_i' , respectively. For the sake of convergence, if exact line searches need to be performed, the above step length μ^* is used as a starting estimate.

REFERENCES

Abbreviations for bibliography:

PSCC = Power Systems Computation Conference
PICA = Power Industry Computer Applications
Bell = Bell Journal of Economics
QJE = Quarterly Journal of Economics
AER = American Economic Review
JPE = Journal of Political Economy
JLE = Journal of Law and Economics
OR = Operations Research
MS = Management Science
PAS = IEEE Transactions on Power Apparatus and Systems

References

- ADAMS, N, Beglari, F, Laughton, M.A., Mitra, G.,
Mathematical Programming Systems in Electrical Power
Generation, Transmission and Distribution Planning,
Proceedings 4th PSCC, 1972, vol 1, 1.1/13
- ANDERSON, D., Models for Determining Least-Cost Investments
in Electricity Supply, Bell 3, 1972, 267-299
- ANDREWS, J.W., Energy-Storage Requirements Reduced in
Coupled Wind-Solar Generating Systems, Solar Energy 18,
1976, 73-74
- AVERCH, H., Johnson, L.L., Behaviour of the Firm under
Regulatory Constraint, AER 52, 1962, p 1052
- ARVANITIDIS, N.V., Rosing, J., Composite Representation of a
Multireservoir Hydroelectric Power System, PAS 89, 1970,
319-335
- ASHBURY, J.G., Mueller, R.O., Solar Energy and Electric
Utilities: Should They Be Interfaced?, Science 195,
1971, 445-450
- BAE, H.M., Devine, M.D., Optimisation Models for the
Economic Design of Wind Power Systems, Solar Energy 20,
1978, 469-481

- BAILEY, E.E., Peak Load Pricing under Regulatory Constraint, JPE 80, 1972, p 662
- BAILEY, E.E., White, L.J., Reversals in Peak and Off-Peak Prices, Bell 5, 1974, p 75
- BAUMOL, W.J., Bradford, D.F., Optimal Departures From Marginal Cost Pricing, AER 60, 1970, p 265
- BAUMOL, W.J., Klevorick, A.K., Input Choices and Rate of Return Regulation: an Overview of the Discussion, Bell 1, 1970, p 162
- BAZARAA, M.S., Jarvis, J.J., Linear Programming and Network Flows, John Wiley, New York, NY, 1977
- BAZARAA, M.S., Shetty, C.M., Nonlinear Programming, John Wiley, New York, NY, 1979
- BEGLARI, F., Laughton, M.A., The Combined Costs Method for Optimal Economic Planning of an Electrical Power System, PAS 94, 1975, 1935-1942
- BENDERS, J.F., Partitioning Procedures for Solving Mixed-Variables Programming Problems, Numerische Mathematik 4, 1962
- BERLIN, E., Cicchetti, C.J., Gillen, W.J., Perspective on Power, Ballinger Publishing Co, Cambridge, MA, 1974
- BEGLARI, F., Laughton, M.A., Model Building with Particular Reference to Power System Planning, in Energy Modelling, IPC Business Press, London, UK, 1974
- BERRIE, T.W., The Economics of System Planning in Bulk Electricity Supply, Electrical Review 181, Sept 1967
- BERRIE, T.W., Further Experience with Simulation Models in System Planning, Proceedings 2nd PSCC, 1966, part 5, 7.4
- BERRIE, T.W., Anderson, D., Power System Planning, Development Programs and Project Selection - a Discussion of Methods, Proceedings 4th PSCC, 1972, vol 1, 1.1/4
- BESSIERE, F., The Investment '85 Model of Electricite de France, MS 17, 1971, B192-B211

- BEZDEK, R.H., Cambel, A.B., The Solar Energy - Utility Interface, Energy 6, 1981, p 479
- BEZDEK, R.H., Hirshberg, A.S., Babcock, W., Economic Feasibility of Solar Water and Space Heating, Science 203, 1979, 1214-1220
- BOITEAUX, M., Marginal Cost Pricing, in Nelson, J.R. (ed), Marginal Cost Pricing in Practice, Prentice Hall, Englewood Cliffs, NJ, 1964
- BONAERT, A.P., Koivo, A.J., El Abiad, A.H., Optimal Scheduling of Hydro-Thermal Power Systems, Proceedings 7th PICA, 1971, 48-56
- BONAERT, A.P., Koivo, A.J., Computational Solutions for Integrated Hydro-Thermal Power Scheduling by Decomposition, Proceedings 8th PICA, 1973, 91-99
- BONBRIGHT, J., Principles of Public Utility Rates, Columbia University Press, New York, NY, 1961
- BOOTH, R.R., Power System Simulation Model Based on Probability Analysis, PAS 91, 1972, 62-69
- BROWN, G.Jr., Johnson, M.B., Public Utility Pricing and Output under Risk, AER 59, 1971, 119-128
- BUCHANAN, J., The Optimality of Pure Competition in the Capacity Problem: Comment, QJE 81, 1967, 703-704
- CARAMANIS, M.C., Tabors, R.D., Nochur, K.S., Schweppe, F.C., The Introduction of Non-Dispatchable Technologies as Decision Variables in Long-Term Generation Expansion Models, IEEE PES 1982 Winter Meeting, 1982
- COHON, J.L., Revelle, C.S., Current, J., Eagles, T., Eberhart, R., Church, R., Application of a Multiobjective Facility Location Model to Power Plant Siting in a Six-State Region of the U.S., Computers and OR 7, 1980, 107-123
- COTE, G., Laughton, M.A., Decomposition Techniques in Power System Planning: the Benders Partitioning Method, Electrical Power and Energy Systems 1, no. 1, April 1979

- CREW, M.A., The Optimality of Pure Competition in the Capacity Problem: Further Comment, QJE 83, 1969, 341-343
- CREW, M.A., Kleindorfer, P.R., On Off-Peak Pricing: an Alternative Technological Solution, Kyklos 28, 1975, 80-93
- CREW, M.A., Kleindorfer, P.R., Peak Load Pricing with a Diverse Technology, Bell 7, 1976, p 207
- CREW, M.A., Kleindorfer, P.R., Reliability and Public Utility Pricing, AER 68, 1978, 31-40
- DEES, D.L., Felak, R.P., Jordan, G.A., The Effect of Load Growth Uncertainty on Generation System Expansion Planning, Proceedings American Power Conference 40, 1978, 1222-1232
- DEVINE, M., Kumin, H., Aly, A., Operations Research Problems in the Economic Design and Operation of Solar Energy Systems, in Energy Modelling and Net Energy Analysis, Institute of Gas Technology, Chicago, IL, 1978
- DREZE, J.H., Some Postwar Contributions of French Economists to Theory and Public Policy, AER 54, 1964, 1-64
- EL HAWARY, M.E., Christensen, G., Optimal Economic Operation of Electric Power Systems, Academic Press, New York, NY, 1979
- ERLENKOTTER, D., Trippi, R., Optimal Investment Scheduling with Price-Sensitive Demand, MS 23, 1977, 1-11
- FELDMAN, S.L., Wirtshafter, R.M., On the Economics of Solar Energy, Lexington Books, Lexington, MA, 1980
- GAGNON, C.R., Hicks, R., Jacoby, S., Kowalik, J., A Nonlinear Programming Approach to a Very Large Hydroelectric System Optimization, Mathematical Programming 6, 1974, 28-41
- GALLOWAY, C.D., Garver, L.L., Kirchmayer, L.K., Wood, A.J., Generation-Transmission Expansion Planning, Proceedings 2nd PSCC, 1966, part 5, 7.6

- GARVER, L.L., Stoll, H.G., Szczepanski, R.S., Impact of Uncertainty on Long-Range Generation Planning, Proceedings American Power Conference 38, 1976, 1165-1174
- GASSFORD, J., Karlin, S., Optimal Policy for Hydroelectric Operations, in Arrow, K.J., Karlin, S., Scarf, H., Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, CA, 1958
- GEOFFRION, A.M., Generalized Benders Decomposition, Journal of Optimization Theory and Applications 10, 1972, 237-260
- HAPP, H.H., Optimal Power Dispatch, PAS 93, 1974, 820-829
- HASLETT, J., Diesendorf, M., The Capacity Credit of Wind Power: a Theoretical Analysis, Solar Energy 26, 1981, 391-401
- HIRSHLEIFER, J., Peak Loads and Efficient Pricing: Comment, QJE 72, 1958, p 451
- HOFFMANN, K.C., Wood, D.O., Energy System Modelling and Forecasting, Annual Review of Energy 1, 1976, p423
- JENKIN, F.P., Electricity Supply Models, in Energy Modelling, IPC Business Press, London, UK, 1974
- KAHN, E., The Reliability of Distributed Wind Generators, Electric Power Systems Research 2, 1979, 1-14
- KIRCHMAYER, L.K., Economic Operation of Power Systems, John Wiley, New York, NY, 1958
- KLEVORICK, A.K., The Optimal Fair Rate of Return, Bell 2, 1971, p 122
- KREITH, F., West, R.E., Economics of Solar Energy and Conservation Systems, CRC Press, Boca Raton, FL, 1980
- LASDON, L.S., Optimization Theory for Large Systems, Macmillan Publishers, New York, NY, 1970
- LENCZ, I., The Planning for the Power System Development with Mathematical Models from Two Parts, Proceedings 3rd PSCC, 1969, vol 2, P8

- LOGENDRAU, R., Oudheusden, D.L., Optimal Operation of a Bi-Fuel Thermal Power Plant by Mathematical Programming, European Journal of Operational Research 8, 1981, 226-234
- LOUVEAUX, F., A Solution Method for Multistage Stochastic Programs with Recourse with Application to an Energy Investment Problem, OR 28, 1980, 889-902
- MANNE, A.S., Electricity Investments under Uncertainty: Waiting for the Breeder, in Macrakis, M.(ed), Energy: Demand, Conservation and Institutional Problems, MIT Press, Cambridge, MA, 1974
- MANNE, A.S., Linear Programming and Sequential Decisions, MS 6, 1960, 259-267
- MANNE, A.S., A Model for Energy Technology Assessment, Bell 7, 1976, 379-406
- MARINO, A.M., Peak Load Pricing in a Neoclassical Model with Bounds on Variable Input Utilization, Bell 9, 1978, 249-259
- MASSE, P., Optimal Investment Decisions, Prentice Hall, Englewood Cliffs, NJ, 1962
- MASSE, P., Gibrat, R., Application of Linear Programming to Investments in the Electric Power Industry MS 3, 1957, 149-166
- MEYER, R.A., Monopoly Pricing and Capacity Choice under Uncertainty, AER 65, 1975, 326-337
- MOHRING, H., The Peak Load Problem with Increasing Returns and Pricing Constraints, AER 60, 1970, p 693
- NELSON, J.R., Practical Applications of Marginal Cost Pricing in the Public Utility Field, AER 53, 1963, 474-481
- NGUYEN, D.T., The Problem of Peak Loads and Inventories, Bell 7, 1976, 242-248
- NITU, V., et al, Models for the Study of Power System Development by Means of Digital Computers, Proceedings 3rd PSCC, 1969, vol 2, P9

- NOONAN, F., Giglio, R.J., Planning Electric Power Generation: a Nonlinear Mixed Integer Model Employing Bender's Decomposition, MS 23, 1977, 946-956
- OFFICER, L.H., The Optimality of Pure Competition in the Capacity Problem, QJE 80, 1966, 647-651
- PANZAR, J.C., A Neoclassical Approach to Peak Load Pricing, Bell 7, 1976, p 521
- PELTZMAN, S., Pricing in Public and Private Enterprises: Electric Utilities in the U.S., JLE 14, 1971, 109-147
- PESCHON, J., Lee, S.T.Y., Dechamps, C., Mathematical Models for Economic Evaluation of Nonconventional Electric Power Sources, in Energy Modelling and Net Energy Analysis, Institute of Gas Technology, Chicago, IL, 1978
- PETERSEN, E.R., A Dynamic Programming Model for the Expansion of Electric Power Systems, MS 20, 1974, 656-664
- PHILLIPS, D., Jenkin, F.P., Pritchard, J.A.T., Rybicki, K., A Mathematical Model for Determining Generation Plant Mix, Proceedings 3rd PSCC, 1969, vol 2, P2
- PRESSMAN, I., A Mathematical Formulation of the Peak Load Pricing Problem, Bell 1, 1970, 304-326
- RALPH, E.L., Large Scale Solar Electric Power Generation, Solar Energy 14, 1972, 11-20
- RAMAKUMAR, R., Allison, H.J., Hughes, W.L., Solar Energy Conversion and Storage Systems for the Future, PAS 94, 1975, 1926-1934
- RATH-NAGEL, S., Voss, A., Energy Models for Planning and Policy Assessment, European Journal of Operational Research 8, 1981, 99-114
- ROSENBERG, W.G., Conservation Investments by Gas Utilities as a Gas Supply Option, Public Utilities Fortnightly, Jan 1977, 13-21
- ROWSE, J., Intertemporal Pricing and Investment for Electric Power Supply, Bell 11, 1980, 143-165

- ROWSE, J., Toward Optimal Capacity Expansion for an Electric Utility, Canadian Journal of Economics 11, 1978, 447-469
- SALIEVA, R.B., Design Principles for Solar and Wind Power Installations, Applied Solar Energy 12, 1976, 41-45
- SANDIFORD, P.J., Bernholtz, B., Shelson, W., Three Applications of Operations Research in a Large Electric Utility, OR 4, 1956, 663-673
- SASSON, A.M., Merrill, H.M., Some Applications of Optimization Techniques to Power System Problems, Proceedings of IEEE 62, 1974, p 959
- SASSON, A.M., Noulin, P., Solutions for Load-Flow, Minimum-Loss and Economic Dispatching Problems, PAS 88, 1969, 399-409
- SCHERER, C.R., Estimating Peak and Off-Peak Marginal Costs for an Electric Power System, Bell 7, 1976, 575-601
- SHEPHERD, W.G., Marginal Cost Pricing in American Utilities, Southern Economic Journal 33, 1966, 58-70
- SHERALI, H.D., A Restriction and Steepest Descent Feasible Directions Approach to a Capacity Expansion Problem, working paper, Dept of IEOR, VPI&SU, 1982a
- SHERALI, H.D., Analysis of Network Structured Models for Electric Utility Capacity Planning and Marginal Cost Pricing Problems, working paper, Dept of IEOR, VPI&SU, 1982b
- SHERALI, H.D., Soyster, A.L., A Unifying Framework for Capacity Expansion Planning and Marginal Cost Pricing for Electric Utilities under Uncertainty, working paper, Dept of IEOR, VPI&SU, 1981
- SHERALI, H.D., Soyster, A.L., Murphy, F.H., Sen, S., Allocation of Capital Costs in Electric Utility Capacity Expansion Planning under Uncertainty, working paper, Dept of IEOR, VPI&SU, 1981b

- SHERALI, H.D., Soyster, A.L., Murphy, F.H., Sen, S., Linear Programming Based Analysis of Marginal Cost Pricing in Electric Utility Capacity Expansion, working paper, Dept of IEOR, VPI&SU, 1981a (also, European Journal of Operational Research, to appear)
- SIMMONS, D.M., Mathematical Programming for Operations Research, Prentice Hall, Englewood Cliffs, NJ, 1975
- SMACKEY, B., Should Electric Utilities Market Solar Energy?, Public Utilities Fortnightly, Sept 1978, 37-43
- SORENSEN, B., On the Fluctuating Power Generation of Large Wind Energy Converters, Solar Energy 20, 1978, 321-331
- SORENSEN, J.R., Tschirhart, J.T., Whinston, A.B., A Game Theoretic Approach to Peak Load Pricing, Bell 7, 1976, p 497
- SOYSTER, A.L., Murphy, F.H., Sen, S., A Convex Programming Approach for Determining Capacity Expansion Plans in the Electric Utility Industry, working paper, Dept of IEOR, VPI&SU, 1981
- STEINER, P.O., Peak Loads and Efficient Pricing, QJE 71, 1957, p 585
- STOBAUGH, R., Yergin, D. (eds), Energy Future, Random House, New York, NY, 1979
- The Role of Utility Companies in Solar Energy, Institute of Gas Technology, Chicago, IL, 1978
- TURVEY, R., Optimal Pricing and Investment in Electricity Supply, MIT Press, Cambridge, MA, 1968a
- TURVEY, R., On Investment Choices in Electricity Generation, Oxford Economic Papers 15, 1963, 279-286
- TURVEY, R., Peak Load Pricing, JPE 76, 1968b, p 101
- TYREN, L., Short-Range Optimization of a Hydro-Thermal System by a Gradient Method Combined with Linear Programming, Proceedings 3rd PSCC, 1969, vol 2, O.S.3
- URI, N.D., Towards an Efficient Allocation of Electrical Energy, Lexington Books, Lexington, MA, 1975

- VARDI, J., Zahavi, J., Avi-Itzhak, B., The Combined Load Duration Curve and its Derivation, PAS 96, 1977a, 978-983
- VARDI, J., Avi-Itzhak, B., Electric Energy Generation, MIT Press, Cambridge, MA, 1981
- VARDI, J., Zahavi, J., Avi-Itzhak, B., Variable Load Pricing in the Face of Loss-of Load Probability, Bell 8, 1977b, 270-288
- VISSCHER, M.L., Welfare-Maximizing Price and Output with Stochastic Demand: Comment, AER 63, 1973, 224-229
- WEINTRAUB, S., On Off-Peak Pricing: an Alternative Solution, Kyklos 23, 1970, 501-519
- WENDERS, J.T., Peak Load Pricing in the Electric Utility Industry, Bell 7, 1976, p232
- WENDERS, J.T., Taylor, L.D., Experiments in Seasonal Time-of-Day Pricing of Electricity to Residential Users, Bell 7, 1976, p 531
- WILLIAMSON, O.E., Peak Load Pricing and Optimal Capacity under Indivisibility Constraints, AER 56, 1966, 810-827
- ZAJAK, E.E., Note on Gold-Plating or Rate Base Padding, Bell 3, 1972, p 311

**The vita has been removed from
the scanned document**

ELECTRIC UTILITY CAPACITY EXPANSION PLANNING
WITH THE OPTION OF INVESTING IN SOLAR ENERGY

by

Konstantin Staschus

(ABSTRACT)

The problem of incorporating non-dispatchable energy sources such as solar energy into electric utility capacity expansion programs is as yet unsolved. This thesis develops methods to incorporate solar energy as a decision variable into capacity planning and capacity expansion planning algorithms. The model is based on variable or intermittent availability of solar energy. For capacity planning, certain convexity properties are established which lead to an efficient decomposition process using a Newton-type search method. For the capacity expansion planning problem, a modification of Benders' Decomposition is applied, which breaks up the problem into a master program containing the solar decision variable, and a subproblem which involves an expansion problem in conventional equipment types. The primal solution to the subproblem is found with existing algorithms, the dual solution can be obtained from the primal solution employing a network interpretation of the problem. This analysis leads to an efficient tangential approximation method. Illustrative examples, computational experience and further generalizations are also provided.