

NAT'L INST OF STANDARDS & TECH R.I.C.



A11102616027

Witzgall, Christoph/The parametric cost-
QC100 .U5753 NO.1225 1986 V198 C.1 NBS-P

NBS Technical Note 1225

The Parametric Cost-Revenue Model (“PAREC”) for Electronic Message Service Systems

Christoph Witzgall, Patsy B. Saunders, Ralph E. Schofer

National Bureau of Standards

QC — BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
100 NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS N
U5753 BS NBS NBS NBS NBS NBS NBS NBS NBS NBS N
NO.1225 BS NBS NBS NBS NBS NBS NBS NBS NBS NBS NBS
1986
C.2



The National Bureau of Standards¹ was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, the Institute for Computer Sciences and Technology, and the Institute for Materials Science and Engineering.

The National Measurement Laboratory

Provides the national system of physical and chemical measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; provides advisory and research services to other Government agencies; conducts physical and chemical research; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

- Basic Standards²
- Radiation Research
- Chemical Physics
- Analytical Chemistry

The National Engineering Laboratory

Provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

- Applied Mathematics
- Electronics and Electrical Engineering²
- Manufacturing Engineering
- Building Technology
- Fire Research
- Chemical Engineering²

The Institute for Computer Sciences and Technology

Conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

- Programming Science and Technology
- Computer Systems Engineering

The Institute for Materials Science and Engineering

Conducts research and provides measurements, data, standards, reference materials, quantitative understanding and other technical information fundamental to the processing, structure, properties and performance of materials; addresses the scientific basis for new advanced materials technologies; plans research around cross-country scientific themes such as nondestructive evaluation and phase diagram development; oversees Bureau-wide technical programs in nuclear reactor radiation research and nondestructive evaluation; and broadly disseminates generic technical information resulting from its programs. The Institute consists of the following Divisions:

- Ceramics
- Fracture and Deformation³
- Polymers
- Metallurgy
- Reactor Radiation

¹Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted; mailing address Gaithersburg, MD 20899

²Some divisions within the center are located at Boulder, CO 80303.

³Located at Boulder, CO, with some elements at Gaithersburg, MD.

NBS
GORD
105752
10.1435
1/16
62

NBS Technical Note 1225

The Parametric Cost-Revenue Model ("PAREC") for Electronic Message Service Systems

Christoph Witzgall
Patsy B. Saunders

with an Appendix by:

Christoph Witzgall
Ralph E. Schofer

Center for Applied Mathematics
National Engineering Laboratory
National Bureau of Standards
Gaithersburg, MD 20899

Revised September 1985
Issued November 1986



U.S. Department of Commerce
Malcolm Baldrige, Secretary
National Bureau of Standards
Ernest Ambler, Director

National Bureau of Standards
Technical Note 1225
Natl. Bur. Stand. (U.S.)
Tech. Note 1225
90 pages (Nov. 1986)
CODEN: NBTNAE

U.S. Government Printing Office
Washington: 1986

For Sale by the Superintendent
of Documents,
U.S. Government Printing Office
Washington, DC 20402

ABSTRACT

This report describes a methodology for analyzing costs and benefits of a national electronic message service system and of similar systems involving alternative configurations of unspecified size. At the core of this methodology is a mathematical model, called PAREC, which provides information on sizes and configurations which optimize profit or return on investment. It is based on a technique for optimal selection of subconfigurations due to J.M.W. Rhys. A parametric minimum cost network flow algorithm was developed for the purpose of solving the resulting optimization problems.

Key words: combinatorial optimization; communication systems; cost-benefit; electronic mail; linear-programming; location theory; network optimization; parametric optimization; return on investment; RHYS technique; system configuration

ACKNOWLEDGEMENTS

We are grateful to Mr. Emile S. Sherrard of the U.S. Postal Service for his active participation in developing the concepts of the PAREC model and for his continued interest and encouragement. Mr. Lester Gardner played a key role in shaping the computer implementation and in transferring the program to the USPS computer facilities. Dr. John Walter and Mr. Charles Siegel contributed their thoughts in many discussions. All were at the time with the sponsoring organization. Mr. Ralph E. Schofer of NBS contributed as a consultant to the authors through much of the project. During the prolonged and often interrupted period of readying this material for final publication, several colleagues at NBS offered editorial suggestions. However, we particularly appreciate the repeated and exceptionally thorough reading by Dr. Alan J. Goldman of Johns Hopkins University. It was also under his guidance and supervision in his former capacity as Chief of the Operations Research Division that this work was carried out.

TABLE OF CONTENTS

	Page
ABSTRACT	
INTRODUCTION.....	1
1. The Purpose of PAREC.....	3
1.1 Basic System Concept.....	4
1.2 Network Selection.....	6
1.3 Data Requirements.....	7
1.4 Seed Costs, Build-up Costs, and Overhead.....	9
1.5 Rules of Thumb for Seed Costs.....	15
1.6 Formulation of the PAREC Model.....	16
1.7 Parametrization.....	19
1.8 Demand Elasticity.....	22
1.9 Return on Investment.....	23
1.10 The Impact of System Overhead.....	23
1.11 Extensions of PAREC.....	26
2. Interpreting the Results of PAREC.....	27
2.1 Computer Inputs and Interpretive Data.....	27
2.2 The VF-diagram.....	28
2.3 The HF-diagram and the VFF-diagram.....	37
2.4 Investment Units.....	37
2.5 Using Seed Costs.....	39
3. How to Run PAREC.....	40
3.1 Overview.....	40
3.2 General Program Comments.....	41
3.3 Input Requirements.....	45
3.4 Output.....	49
3.5 Error Messages.....	57
4. Description of PAREC Program.....	58
4.1 Data Structure.....	58
4.2 Subroutines.....	63
APPENDIX A: The Network for the Dual Problem.....	69
APPENDIX B: The Use of Customized Stationery.....	75
REFERENCES.....	84

INTRODUCTION

In terms of its specific motivating application, this report deals with the question of how to place an optimal number of Electronic Message Service (EMS) facilities ("how many" and "where"), balancing the costs of such installations against the revenues they would realize.* The report is based on earlier work sponsored by the U.S. Postal Service.

The question arises during the planning phase of an EMS-system (EMSS) when such optimized networks would provide both estimates for the size of a mature system and guidelines for the system's build-up over time. A related question arises in performing the economic analyses of proposed EMSSs: which of the many possible networks should be selected as candidates for a break-even analysis?

In Spring 1974, a number of potentially applicable mathematical methods for site selection were reviewed by one of the authors (WITZGALL [1]). Two of these were recommended for further development and implementation:

- 1) Maximization of incremental patronage volume (or related measures of benefit) based on Origin Destination Information System (ODIS) [2] data by successively adding those facilities which each at its turn offers the maximum incremental benefit (SHERRARD [4]). In this fashion, networks are found which are approximately optimal for a given number of stations. (Methods of this kind are frequently called "greedy" algorithms.)
- 2) Adoption of a general approach introduced by RHYS [5] for cost-benefit optimization of configurations. Based again on ODIS data, it will simultaneously optimize the size of the network and the location of its installations. It can be applied to configurations which contain several different kinds of installations, e.g. terminals hooked up to stations.

The first approach has been further advanced and implemented as the MAXBEN model by GARDNER [6]. The second approach has been pursued by the authors and has led to the development of the PAREC (= "Parametric Revenue Cost") model, which is the subject of this report. One advantage of the PAREC approach is that it produces an actually optimal sequence of "nested" networks, or more generally, configurations. (The MAXBEN model produces sequences that are only approximately optimal.) Since the optimal PAREC solutions are solutions to linear programming problems, they are accompanied by dual-variable values

* Disclaimer: Any references to companies in the private sector do not imply recommendation or endorsement by the National Bureau of Standards of their reports or products. Any opinions expressed in this report are those of the authors and do not represent a position taken by the National Bureau of Standards.

which may be used for sensitivity analyses. While PAREC may be computationally more expensive than some of the simple greedy schemes, it may compare favorably to the more complex versions of MAXBEN. The version of the PAREC model described in this report had been implemented and a code delivered to sponsor in 1976. A first version of this report had been issued in 1977 as a final report at the conclusion of an interagency project agreement with the USPS.

It was the uncertainty that clouds demand forecasts that motivated the authors to develop a parametric variant of the RHYS-method. When applied to ODIS data, this parametric approach yields among other results a "break-even configuration" in the sense that no other network can break even if this one fails to do so. The same network also maximizes return on investment, provided there can be any (positive) return at all. An adaptation of the RHYS-method to volume-determined sizing of facilities is accomplished by introducing the concepts of "seed-cost" and "build-up cost."

The PAREC technique to be advanced in this report thus consists of two parts: (i) a particular method of collecting and organizing data, especially cost items, to be carried out by the analyst in preparation for the application of the PAREC computer model; (ii) the actual application of the PAREC computer model, and the subsequent interpretation of the computational results.

Subsequently, it was realized that the techniques advanced in this project exemplify a valuable generic tool for the cost-benefit analysis of configurative systems: they permit break-even and profitability determination in case the size and the layout of the configuration itself are not known at the outset, and can be used to determine optimal network sizes along with optimal facility locations. Although the discussion is carried out in the particular context of an EMSS, applicability to a broad class of technologies is evident. For these reasons, the decision was made by NBS to make this material more generally available, albeit belatedly, in the form of an NBS Technical Note.

1. The Purpose of PAREC

PAREC investigates the economic performance of an EMS technology without a priori deciding on a particular configuration or "network", that is, on the size, location and number of EMSS facilities. The choices open to PAREC in selecting a suitable network are, however, restricted to all possible subnetworks of a user-specified "universal" network or configuration. PAREC also requires knowledge of fixed and variable costs for facilities, covering both operations and investment. In addition, the rate-structures under which customers will be charged for using the EMSS need to be stipulated. For instance, the user may want to include costs and benefits which reflect the impact of an EMSS on traditional USPS services. For demand information, PAREC relies on ODIS data. The idea is that ODIS data are indicators of the relative strength of the EMS market between pairs of ODIS areas, and that the actual demands for an EMSS can be estimated by multiplying these relative market indicators by a general "demand level parameter".

Each "run" of PAREC covers all possible values of the demand level parameter. Each run can therefore be considered as providing a whole "family" of model evaluations. This will illustrate the sensitivity of the size of the optimal configuration to changes in the general demand level. In particular, PAREC specifies brackets within which the demand level may change without changing the optimal configuration. The optimal configurations for specified demand levels can thus be found by "table look-up."

A remarkable theoretical result is that the optimal configurations are "nested", that is, in the sequence of optimal configurations resulting for increasing demand levels, each configuration is contained in the subsequent one. In other words, optimization at a given demand level does not compromise the optimality of subsequent expansions.

The smallest of these nested networks are called "break-even" networks, because it can be shown that if they are not economical, then there exists no economical EMS system. PAREC can therefore be used to determine the economic feasibility of a particular EMS technology.

A further useful theoretical result of this study is the observation that the economic feasibility of a particular EMSS as well as the size of the optimal configuration depends essentially on the values of the ratio

$$h = \frac{c}{L \cdot r} ,$$

where c is the annual facility cost, L the demand level, and r the direct revenue per item. This statement will be made more precise later on in this report.

1.1 Basic System Concept

In this section, we describe the general structure of an EMSS as presumed by the present version of PAREC.

The purpose of an EMSS is the electronic transmission and reception of messages or other information, henceforth called EMS-items. These enter the EMSS--either directly or by way of the postal collection system--at an EMS-terminal, where their information is transformed into electronic signals. The signals are then sent by some local communication system (e.g., telephone) to an EMS-station, where they are sorted and prepared for long-distance transmission to another station. From this station they are forwarded--again by local communications--to their destination terminals, where they are printed and sealed, or otherwise converted into deliverable form. The long distance communication is generally seen as utilizing a communications satellite, owned or leased; in any case, it is such that any station in the configuration can communicate directly with every other station. The configuration can be visualized as a network, whose links form paths which connect terminals with each other.

Each EMS-terminal serves a zone, within which the EMS-items transmitted via this terminal originate and the items received at this terminal are to be delivered. Zones should not overlap. Each terminal hooks up to a particular station, but a station may serve several zones, which are then said to form a region. The zones themselves may consist of one or more ODIS-areas. Thus areas combine into zones, and zones into regions, which also will not overlap (see Fig. 1). More often than not, a single area will be an entire zone. Similarly a single zone may be an entire region. Overlapping of zones is ruled out, because a major assumption of the PAREC model is that each zone is served either fully or not at all. Service within a zone is provided by its corresponding terminal, through which all service to customers in the zone is assumed to be channeled. Zones without terminal are considered without service.

Different zones will generate different levels of demand for EMS. It is assumed that sufficient equipment has been installed in any zone and region to satisfy the demand generated in and for this zone and region. Thus different terminals and stations will in general have different sizes. A study by PHILCO [8] introduced the concept of modules for measuring the size of an installation: a module is a minimal collection of all equipment necessary to do the job, and any installation is described as consisting of one or more such identical modules. In our case, there would be 1-module, 2-module, 3-module, and so on, terminals, stations, or satellites. The number of modules required for each installation is determined by the fixed capacity of each single module and by the annual volume anticipated at the installation. PAREC will retain the term "module" for "smallest operational equipment" or "base line" unit. However, it will consider fractional sizes in excess of 1.0, like "2.3-module station" or "1.7-module terminal." This practice will be justified in Section 1.4 below.

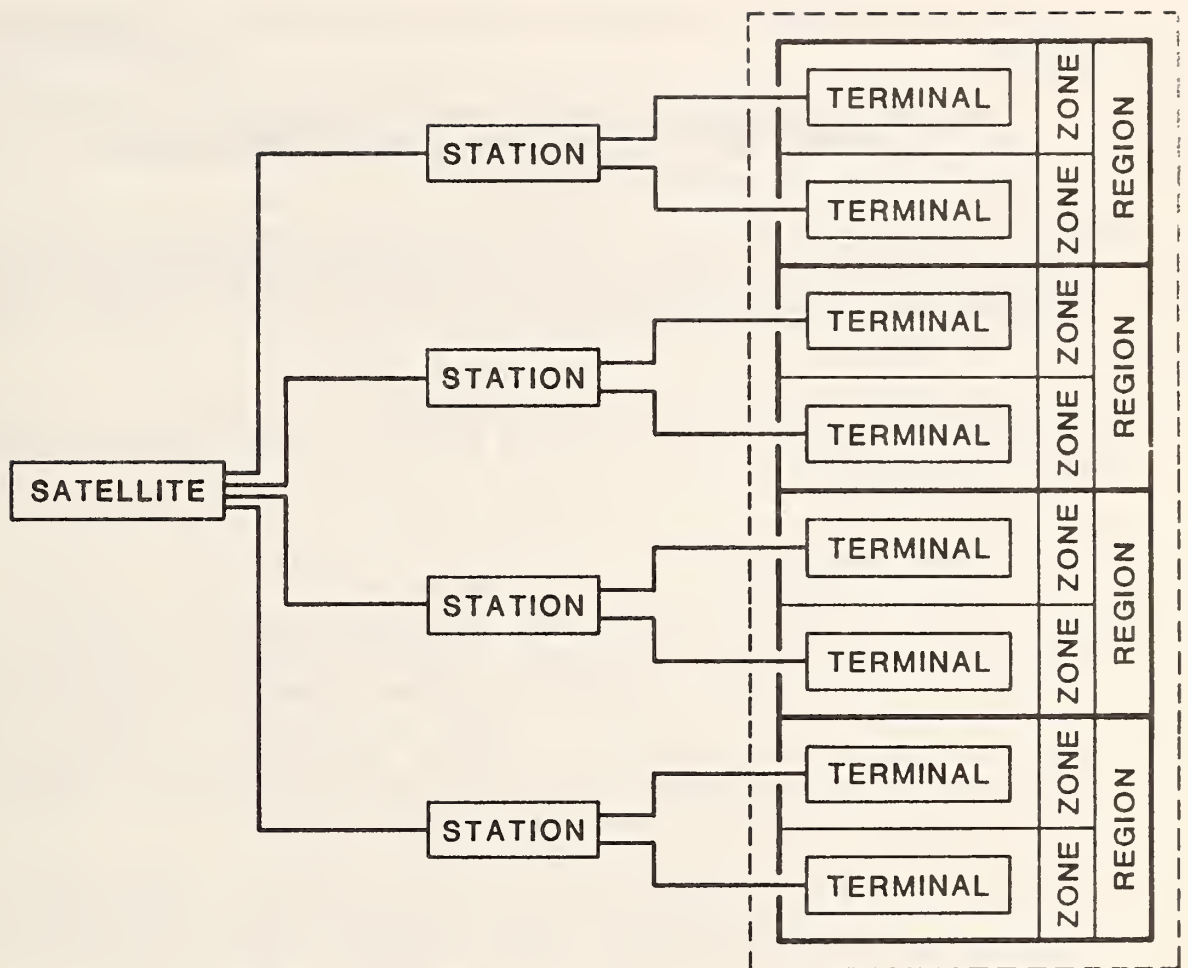


Figure 1: Schematic representation of an EMSS.
Dashed lines --- indicate area served

Note that any mature EMSS is bound to provide various distinct services, to which different rates apply, and which put different loads on various kinds of equipment. In other words, there will be several categories of EMS-items.

The following assumption simplifies both data analysis and mathematical theory:

Homogeneity assumption: The EMS demand is homogeneous, that is, the proportions of the various categories of EMS-items are the same on all potential links.

This assumption permits us to consider an average EMS-item whose cost and revenue profile is a proper mix of the profiles of the various categories. As a consequence, volumes and capacities can be measured in terms of average EMS-items.

The homogeneity assumption has been adopted for this study. In Section 1.11, a possible extension of PAREC will be discussed in which the EMS-demand need not be homogeneous.

1.2 Network Selection

The purpose of PAREC is to determine the optimal number of terminals and stations as well as their locations. These facilities are considered located, if the zones and regions they serve have been determined. The questions of precisely where to place terminals within their zones, or stations within their regions, are not addressed. In fact, terminals or stations need not be physically within the boundaries of their respective zones or regions, and may even be split into parts located in different places.

The main criterion of optimality is to maximize profit. PAREC accomplishes this by selecting from potential zones those which are actually to be served, i.e. will be assigned a terminal. The potential zones must be specified at the outset, along with a grouping of zones into nonoverlapping potential regions wherever desired. A potential region requires a single station in order to operate terminals in any one or more of its zones. No terminal can function without being connected to a station. Placing stations in all regions, and terminals in all zones, would result in the universal configuration, of which all other configurations are subsystems, and which has been mentioned at the beginning of this chapter.

The analyst will have to gather, digest, and aggregate a large amount of diverse information in order to arrive at the input data for PAREC. Demand estimates will be particularly critical and hard to come by. This presents a special challenge to the modeler during the early stage of the planning: to derive useful results which do not rely on precise knowledge of the general level of demand. Our answer has been to develop a "parametric" approach, which is able to identify special preferred network configurations, and which will be described in Section 1.7.

The three following sections will deal with data sources, forms and kinds of information needed, and simplifying assumptions. They will provide guidelines to the analyst on how to prepare information for input into PAREC.

1.3 Data Requirements

The main elements of a cost-benefit analysis are cost and revenue. Corresponding data for an EMSS will be discussed in detail.

To start with, the potential zones and regions will have to be specified. In general, each pair of potential zones will have a specific demand (anticipated EMS-item flow) between them, a specific rate schedule, and specific "direct" per item costs. All these data are necessary for determining the operating profit for each pair of potential zones. This profit will be realized whenever both zones are included in the EMS-network configuration.

A direct or variable cost is any cost which increases proportionally to the quantity of EMS-items processed, e.g. the cost of paper on which the messages are printed for delivery (ignoring volume discounts). Costs which are incurred by the system as it stands regardless of actual volume processed are called indirect or fixed costs. Examples are station administration costs, building costs, costs of purchasing or leasing machinery, and so on. To the extent that labor requirements are dictated by the amount of machinery present rather than actual volume processed (= "operating labor"), labor costs are also indirect costs, especially when contracted for on an annual basis. The same is true for maintenance agreements.

The PAREC model requires that every indirect cost be expressed as an annual indirect cost. Most indirect costs will naturally arise in this form. Exceptions are purchase costs for equipment or facilities. Life time information for such equipment or facilities can be utilized to develop suitable amortization and depreciation schedules which permit the conversion of purchase costs to repeated annual costs. The amount of equipment needed depends, of course, on capacities and on anticipated demand.

Purchase of real estate, site-finding, negotiations, legal requirements, promotions, and similar costs associated with the initiation and start-up of an installation appear to occur only once. Transforming such costs into regular and equal annual costs would thus depend on fixing a time "horizon" after which these costs would have to be fully amortized. We submit, however, that most of the above costs are, in reality, recurring. Indeed, installations will outgrow their localities and are likely to move for this and other reasons. Initiating an installation could thus be regarded as an investment which depreciates and which will have to be repeated after a certain "life time," however long that may be. This interpretation permits the transformation of initiation costs into equivalent annual expenses based on a life time estimate by the analyst.

The annual indirect costs at each installation depend on its size. That size depends on the anticipated volume to be processed, which in turn depends on which other zones and regions are served by the system. When hypothetically assessing the effects of adding or deleting installations--and this is the

essence of mathematical optimization--it is therefore not possible to fix a total annual indirect cost for an individual installation in a way analogous to calculating the operating profit for a potential link. It is, however, possible to determine the annual indirect costs for a module of a potential terminal, station, or satellite.

Since there are locational variations in labor, construction, and land costs, the annual indirect costs of modules may be different in different potential zones. Indications are, however, that such variations while possibly substantial do not influence the results of PAREC in a major way. The present version of PAREC is therefore based on the simplifying

Equal costs assumption: The annual indirect costs of each kind of module do not depend on the location of the module.

It would not be difficult however, to incorporate local cost variations into the model, should the need for a more refined analysis arise.

It is not enough to determine the various cost-elements on a per module basis, because of the "economies of scale" shown by some cost-elements like administration, buildings, operational labor, and others. If the number of modules increases within a given installation, these costs increase less than proportionally. To be very precise, one would have to know each cost element not only for the first module, but also for the second module, third module, and so on. It is clear that such a study would be much too detailed for the system planning situation in which PAREC is to be used. There are options that should be left open at the configuration planning level--like leasing versus buying or enlarging of buildings--which would all show different degrees of economies of scale. In the following sections, we will propose therefore some simplified procedures for estimating the relationship between the size of an installation and its costs.

To sum up, in order to determine the optimal network configuration, or to establish the economic viability of a particular proposed network, the following information will have to be provided:

- (1) Potential zones and regions, which are the candidates for having terminals and stations installed;
- (2) a service mix, i.e., the proportions of service categories (different types of EMS-items);
- (3) an estimate for the relative demand between each two potential zones (the actual demand will depend in addition on the general demand level);
- (4) rates and direct costs per average EMS-item (for each link, if rates and/or direct costs vary);
- (5) capacities of equipment;
- (6) life-times of purchased equipment and facilities;
- (7) cost per module of purchased equipment and facilities;
- (8) annual indirect costs of leases, labor and maintenance, administration, and so on, for each size and kind of installation.

This information has to be processed in order to serve as input to PAREC. Instructions for this processing are provided in the following sections.

1.4 Seed Costs, Build-up Costs and Overhead

In this section, we propose a way of handling varying sizes of installations and their influence upon the selection of EMS configurations.

Consider those annual indirect costs at installations that are most closely proportional to their size, namely the annual indirect costs for leasing or amortizing and depreciating machinery. If installations are viewed as consisting of replications of identical modules, then the annual equipment costs will be integer multiples of annual equipment costs per module. In what follows, we will reassess this point of view and introduce the concepts of "seed costs" and "build-up" costs.

As a hypothetical but illustrative example, consider a single terminal module of capacity 3 and cost 2. The curve which plots required terminal costs against demand volumes then takes the form of a simple staircase (Fig. 2). This cost-volume relationship is plainly unrealistic.

First, there will almost certainly be different types of machinery for identical tasks, varying as to their costs and capacities. As a simple example, we consider two terminal modules, one at cost 1.5 and capacity 2, and the other one as before at cost 2 and capacity 3. A combination of smaller and larger modules will then provide an optimal mix for various demand levels. This has a smoothing effect on the staircase curve, as shown in Fig. 3. Note that the bigger module still determines the "slope" of the new staircase curve. Note also that the "rises" between steps repeat in periods of 3 after step 6. In reality, this smoothing effect will be much more pronounced because there will be more equipment types and also kinds of machinery for which a choice of types will be available.

Second, the capacity limitations are not sharply defined. There is a "grey area" between recommended full load and actual overload. This means that the rises of the steps are not vertical and that the cost-volume relationship is more realistically expressed by a curve of the form (Fig. 4). The real cost-volume curve for the equipment at an installation therefore lends itself quite well to an approximation by a straight line with positive vertical intercept (= "cost at zero"). The analyst is thus faced with the selection of a suitable approximation. To guide him in this endeavor, we will develop some "rules of thumb" in Section 1.5.

In Fig. 4, the intercept is selected in such a fashion that everywhere, except at the origin, the cost-curve is slightly below the straight line. This represents therefore an increased estimate of the annual costs. This increase is not expected to distort the model's representation of reality. Besides, it might be justified in terms of prudent acquisition of excess equipment to cover peak loads and anticipated expansion. Also, errors in forecasting the actual demand and resistance against frequent equipment changes will result in equipment mixes which are not optimal.

The PAREC model requires that all annual indirect costs are approximated in this way as linear (i.e. straight line) functions of the size of the installation. We call the cost at zero the seed cost. The slope of the line is a cost per item, and we call it the build-up cost. Thus every annual

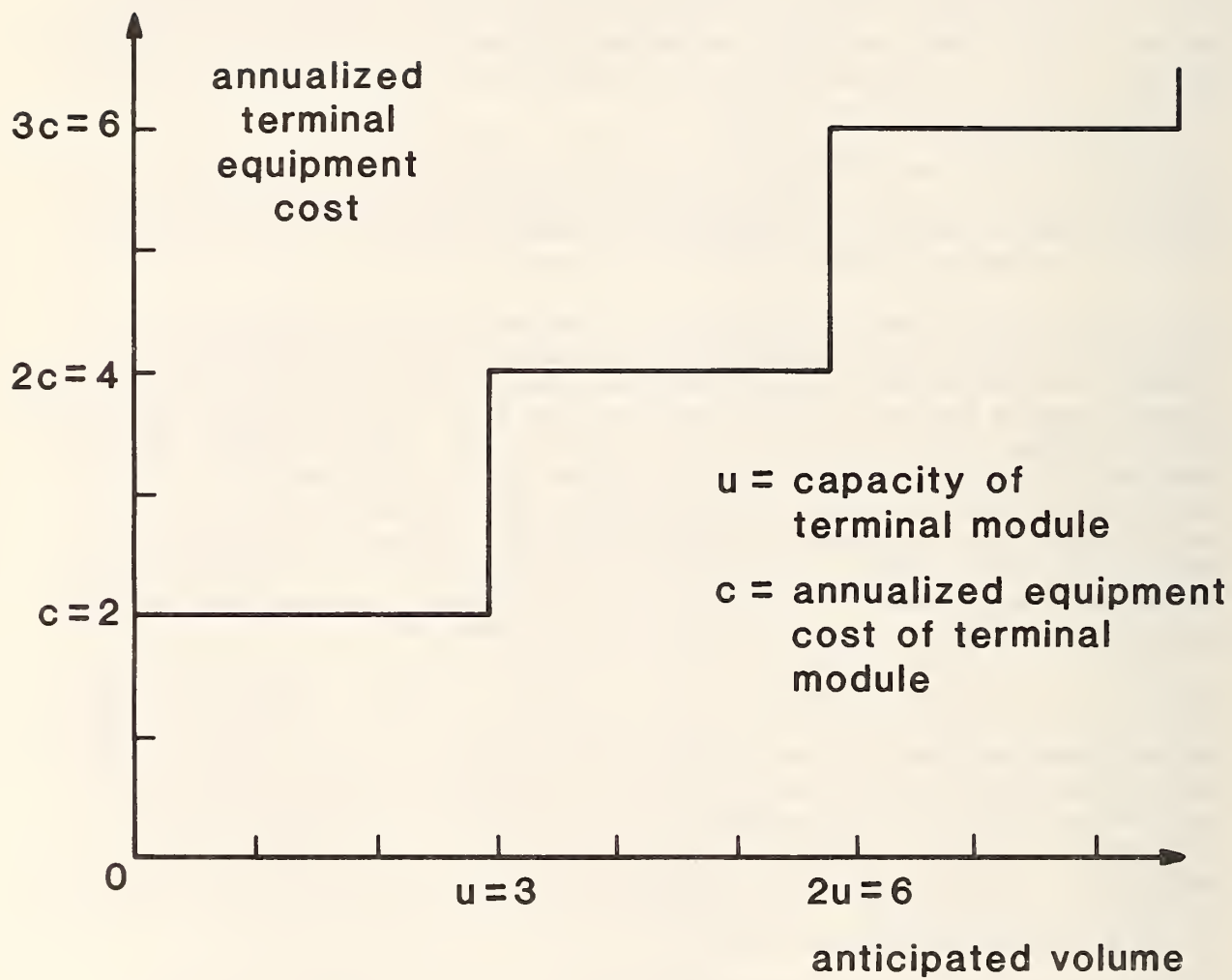


Figure 2: Schematic increase of equipment cost with anticipated volume

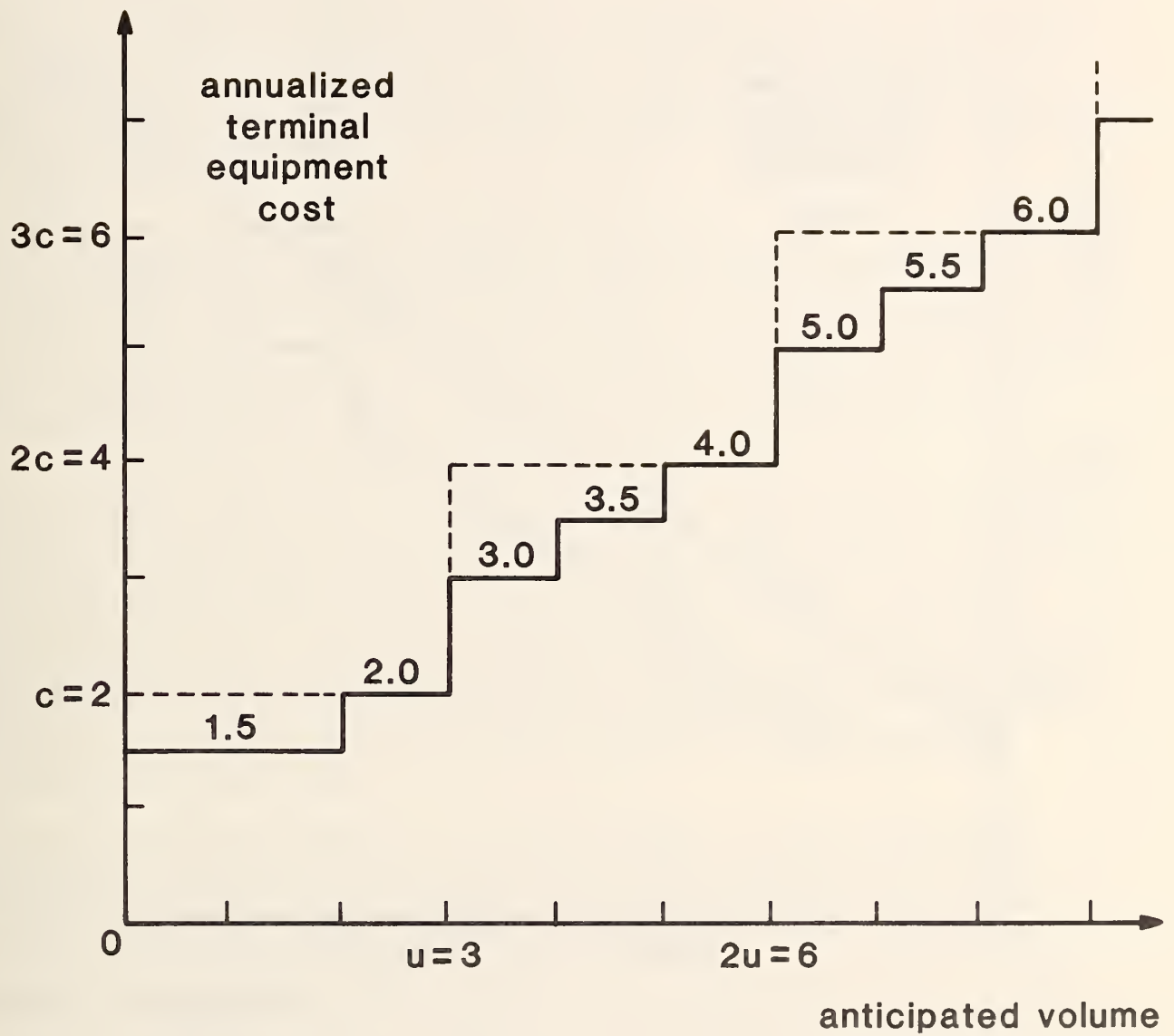


Figure 3: Smoothing effect caused by availability of modules of different sizes

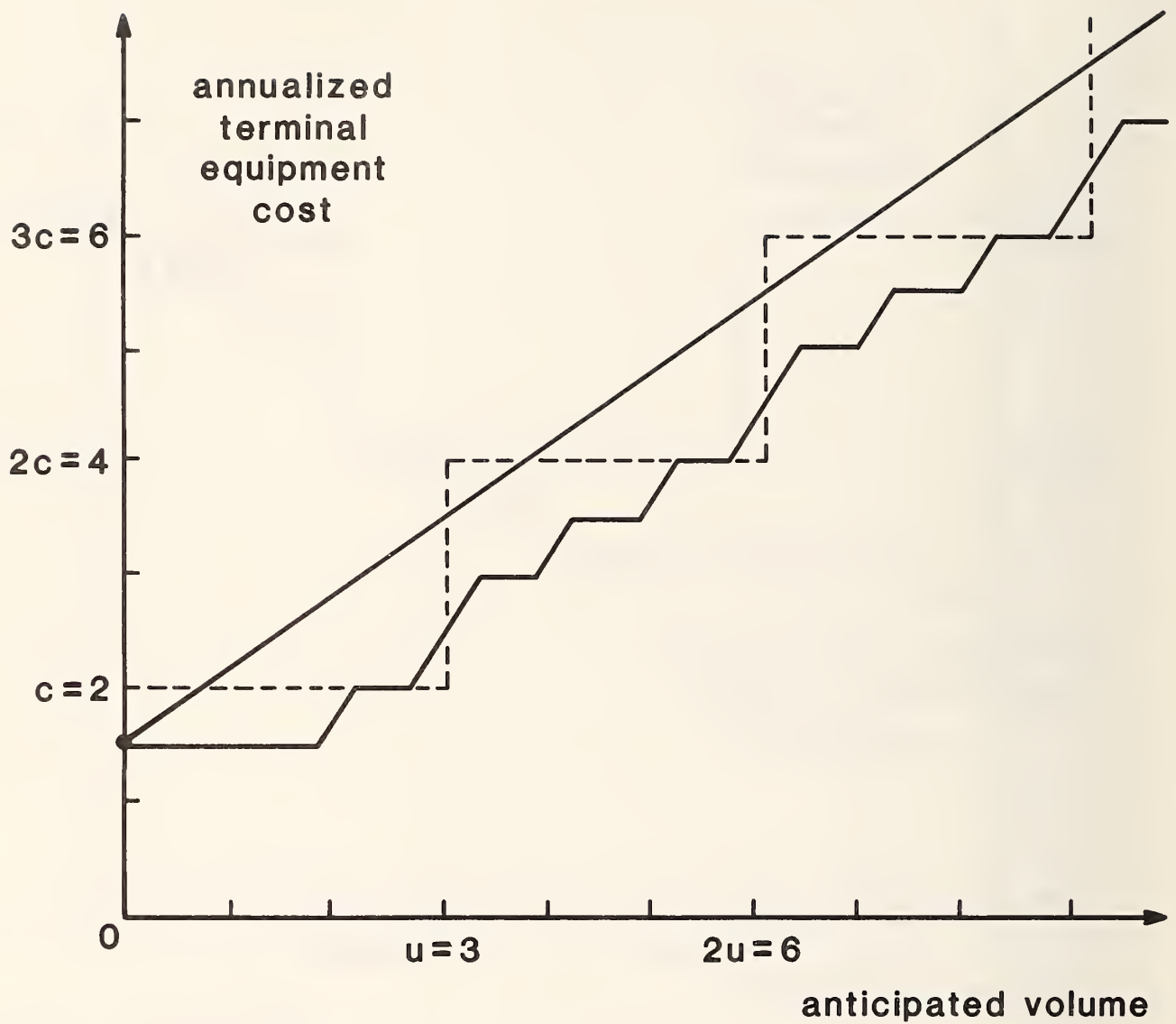


Figure 4: Taking into account the fact that capacity limitations are not sharply defined, the cost-volume curve is approximated by a straight line.

equipment cost is split into seed cost and build-up cost times volume processed. We have described this split-up, based on a straight line approximation, for annual equipment cost.

A similar procedure is to be followed for those cost elements that, unlike equipment costs, show economies of scale. Qualitatively, the cost-volume curve will be of the form shown in Fig. 5. It is clearly not possible to approximate equally well for small and for large volumes. If more accuracy is needed, then an extension of the present mathematical method will be required (see Section 1.11).

Some further remarks about the build-up costs are in order. They constitute per item costs which can be attached to each average EMS-item that passes the installation. Any such item, provided it does not stay within its region of origin, passes two terminals and two stations, each time being counted as part of the volume processed. Thus terminal and station build-up costs will have to be assessed twice. (If one were to assess "transmission build-up costs" and "reception build-up costs" separately, the result would be the same as that from assessing the mean of these two build-up costs twice.)

Intraregional traffic--if considered at all--may require different treatment. Note, however, that some of the transmission tasks and some of the reception tasks are both executed even for intra-regional traffic, albeit at the same station. Thus one may argue to some extent for a double assessment of station build-up costs for EMS-items which never leave their region. Charging satellite build-up costs to intra-regional messages is even more questionable. (For these and other differences between intra- and extra-regional traffic, the present PAREC program gives the analyst the opportunity of specifying a blanket profit adjustment factor for intra-regional traffic.) In what follows, we assume equal build-up costs for all EMS-items, mainly for simplicity of exposition.

Turning to system overhead: it includes expenses for the long-distance communication system, headquarters, software development, promotion, etc. Equipment costs stem from the communication system. They may be channels leased on an annual basis; they may be communication satellites owned or leased. The equipment size will be determined primarily by the anticipated volume to be processed. Direct communication costs may be present. Annual indirect equipment costs will have to be approximated in the manner described earlier by straight lines with respect to anticipated volume to be processed, and will thus be split into a general seed-cost and an additional build-up cost per item.

The extent of administrative, developmental and promotional costs at headquarters will exhibit definite economies-of-scale with respect to the size of the entire configuration. If the analyst considers it necessary, he may generate still another seed cost and a build-up cost to be distributed over those units in which the size of the configuration is measured. For instance, if the size is measured by number of stations, the station seed cost will have to be incremented accordingly.

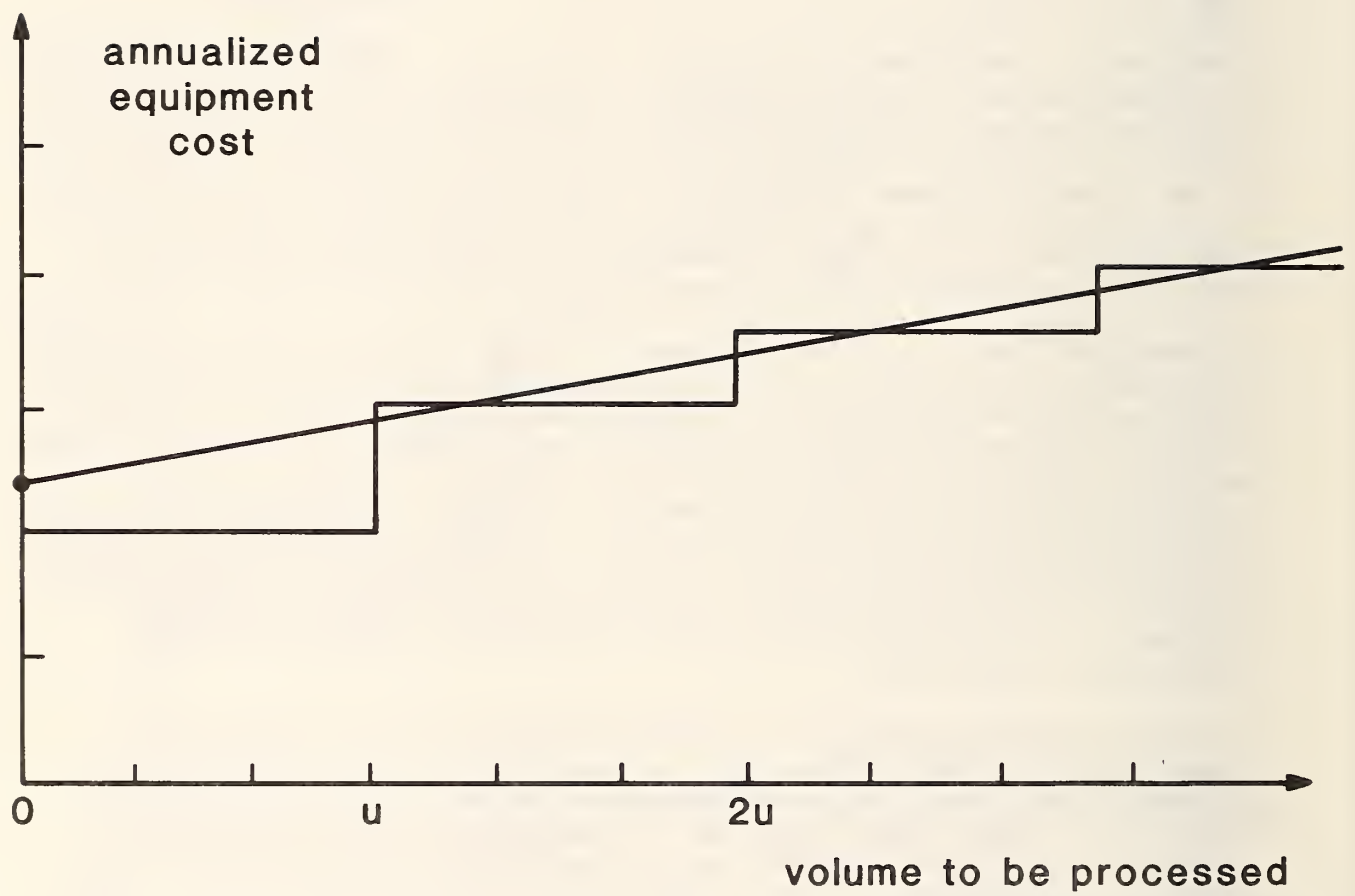


Figure 5: Cost-Volume curve in presence of economies of scale.

1.5 Rules of Thumb for Seed Costs

The determination of the cost-volume curves for the annual indirect cost-elements of installations, involving the selection of build-up and seed costs, is the task of the analyst. It is part of preparing the input for PAREC. Without intending to preempt this role of the analyst, we suggest three simple rules of thumb for estimating total seed costs and build-up costs for an installation. These rules are intended for the case in which cost-information is available mainly in the form of total annual indirect costs per module, with little information as to the cost elements taken into account or the economies-of-scale anticipated. In other words, we are addressing the case in which the cost functions is a step function of the form indicated in Fig. 2. For many of the early planning applications of PAREC, only the "relative" sizes (i.e. mutual ratios) of seed costs will matter, and in these and similar cases, the very rough estimates recommended below are justified.

For each module, terminal or station, estimate that portion of its annual indirect cost which is due to equipment, namely the

total annual indirect equipment cost c_e per module.

The remainder of the indirect costs will be mainly associated with facilities and their administration. We therefore call this remainder the

total annual indirect facility cost c_f per module.

The equipment cost increases essentially in proportion to size if the latter is increased, whereas the facility cost is subject to economies-of-scale. In addition, one needs to know the capacity of the module:

maximum annual volume u per module.

In the case of the equipment costs, we approximate the cost-volume curve by a straight line precisely as described in Fig. 4 of the previous section. This suggests the first Rule of thumb: For terminals and stations,

equipment seed cost = $.75 \cdot c_e = .75 \cdot (\text{cost-at-zero})$ [\$ / yr]

equipment build-up = c_e / u [\$ / # items]

In the case of facility costs, we assume that the second module requires only half of the cost of the first one, the third requires only one third, the fourth requires one fourth, and so on. This economies of scale schedule is the one illustrated in Fig. 5, and the straight line approximation shown there is the one that corresponds to the following second Rule of thumb: For terminals and stations,

facility seed cost = $1.2 \cdot c_f = 1.2 \cdot (\text{cost-at-zero})$ [\$ / yr]

facility build-up = $.25 \cdot c_f / u$ [\$ / # items]

Again this is a suggestion from which the analyst may well want to deviate.

With regard to system overhead, we recommend the following procedure: distinguish (as before) between annual equipment cost per module c_e and annual facility cost per module c_f . The former refer to communications cost, cost, the latter to administrative, development and promotional costs. We then recommend considering build-up costs only for equipment.

Rule of thumb: for system overhead,

$$\text{overhead seed cost} = .75 \cdot c_e + c_f \text{ [$/yr]}$$

$$\text{overhead build-up} = c_e/u \text{ [$/\# items]}$$

Here u denotes again the maximum annual module capacity.

Seed costs and build-up costs represent a way of estimating individual installation costs in response to estimated volume of service. To achieve this, the actual estimated volume processed at, say, a terminal, is multiplied by the terminal build-up costs. The resulting annual cost is added to the annual seed cost to give the annual fixed cost representative of terminal size. More precisely, for given annual volume requirement, we have in the case of a terminal:

$$\begin{aligned} \text{annualized fixed cost of terminal [$/yr]} = & \\ & \text{terminal equipment seed cost [$/yr]} \\ & + \text{terminal facility seed cost [$/yr]} \\ & + \text{annual volume [\# items/yr]} * \text{terminal build-up cost [$/\# items]} \\ & + \text{annual volume [\# items/yr]} * \text{facility build-up cost [$/\# items]} \end{aligned}$$

1.6 Formulation of the PAREC Model

The profit realized by a particular EMS-configuration can be expressed in terms of demands, rates, direct costs, seed costs and build-up costs. Let

$$P = \text{annual profit,}$$

and for potential zones i and j where $i < j$, let

$$p_{ij} = \text{operating profit realized between the zones}$$

$$v_{ij} = \text{annual volume between zones (if both served)}$$

$$r_{ij} = \text{per-item revenue (= rate)}$$

$$\gamma_{ij} = \text{total per-item direct costs}$$

and

$$\beta = \frac{\text{total per item build-up costs}}{(\text{same for each EMS-item in the system})}$$

$$c = \frac{\text{annual (seed) costs per terminal}}{(\text{same for all potential zones})}$$

$$d = \frac{\text{annual (seed) costs per station}}{(\text{same for all potential regions})}$$

$$e = \text{annual (seed) costs for system overhead.}$$

To allow for the fact that not all potential zones are necessarily reached by the system--and in order to provide variables for the optimization--we introduce the following 0,1-indicators:

$$y_{ij} = \begin{cases} 1 & \text{if zones } i \text{ and } j \text{ are serviced} \\ 0 & \text{else} \end{cases}$$

$$x_i = \begin{cases} 1 & \text{if zone } i \text{ has terminal} \\ 0 & \text{else} \end{cases}$$

$$z_h = \begin{cases} 1 & \text{if region } h \text{ has station} \\ 0 & \text{else} \end{cases}$$

$$w = \begin{cases} 1 & \text{if there are any stations in the entire system} \\ 0 & \text{else} \end{cases}$$

In what follows, we will use the symbol

$$h(i)$$

to denote the region h which contains zone i . The indicators then must satisfy

$$y_{ij} = x_i \cdot x_j = \min \{x_i, x_j\}$$

$$(1.6.1) \quad z_h = \max \{x_i : h(i) = h\}$$

$$w = \max \{z_h : \text{all } h\}$$

These three conditions, respectively, reflect the fact that the volume v_{ij} will be captured precisely if there are terminals in both zones i and j , the fact that a region has a station precisely if at least one of its zones has a terminal, and finally the fact that overhead will not be present, unless there is a network.

The profit of the configuration singled out by the indicators is then given by the formula

$$(1.6.2) \quad P = \sum_{i,j} p_{ij} \cdot y_{ij} - c \cdot \sum_i x_i - d \cdot \sum_h z_h - e \cdot w$$

where

$$(1.6.3) \quad p_{ij} = (\rho_{ij} - \gamma_{ij} - \beta) \cdot v_{ij}$$

expresses the annual profits in terms of volumes, rates, direct and build-up costs.

The PAREC optimization problem then reads:

$$(1.6.4) \quad \begin{array}{l} \text{Maximize} \quad \sum_{i,j} p_{ij} \cdot y_{ij} - c \cdot \sum_i x_i - d \cdot \sum_h z_h - e \cdot w \\ \text{subject to} \quad y_{ij} \leq x_i, y_{ij} \leq x_j, x_i \leq z_h(i), z_h \leq w \\ \text{where} \quad x_i = \begin{cases} 0 \\ 1 \end{cases} \end{array}$$

It is assumed that

$$(1.6.5) \quad p_{ij} \geq 0, c \geq 0, d \geq 0, e \geq 0.$$

It can be shown that if all operating profits p_{ij} and the seed costs are actually positive, then any optimal solution $\{y_{ij}, x_i, z_h, w\}$ will be a set of 0,1-indicators satisfying conditions (1.6.1). If $d = 0$ or $p_{ij} = 0$ for some pairs of potential zones, then the values z_h and y_{ij} , respectively, do not matter. They can therefore be redefined in terms of the values of x_i so as to satisfy (1.6.1). This also ensures that the new values of y_{ij} and z_h are either 0 or 1. If $e = 0$, one may redefine w analogously or drop the variable w altogether. In any case, a configuration is identified and this configuration maximizes profit.

For all non-empty configurations, $w = 1$. The term w therefore acts very much like a common constant added to the function which is to be maximized. As such constants don't affect optimizations, one wonders whether the term ew is necessary. We will discuss this question in Section 1.10.

Problems of the form (1.6.4) have been studied by RHYS [5], who has shown that they are "duals" of network flow problems. Solving these network flow problems--and particularly efficient solution methods are available for this task (see FORD and FULKERSON [9])--will then solve the original problems as well. This solution technique has been incorporated into PAREC. The reader should keep in mind, that the networks for which the flow problems are formulated and solved are not the configuration networks of EMSS: they are "abstract" networks which arise from the latter by "dualization" (see for instance GASS [10]). Details of the actual dualization are presented in Appendix A to this report.

1.7 Parametrization

It is apparent that an actual PAREC optimization will require a considerable data-collection and analysis effort. Cost and operational data will depend on the details of a particular system definition. A market analysis would in general be needed to define expected demands as well as viable services.

The demand forecasts are particularly critical as well as difficult. Estimates have been attempted (e.g. GENERAL DYNAMICS [10]) by identifying that percentage of conventional mail to be considered "replaceable" by an EMS. The judgement of replaceability, however, had to be made without the benefit of a precise definition of the service to be performed.

In addition--and this is the most crucial aspect--the information which would indicate which percentage of the replaceable mail actually would be replaced, may be controversial and subject to frequent revisions

On the positive side, the case has been made (e.g. SHERRARD [4]) that the distribution of present or past mail flows as observed, for instance, by ODIS [2], will provide reasonably good indications as to the relative distribution of EMS demands. In other words, given for each pair of potential zones i and j a

reference volume \bar{v}_{ij} ,

the actual volume can be estimated as a multiple of the reference volume,

$$(1.7.1) \quad v_{ij} = L \cdot \bar{v}_{ij} ,$$

using a single general parameter, the

$$(1.7.2) \quad \text{demand level } L > 0 .$$

Defining the

$$(1.7.3) \quad \text{reference (operating) profit } \bar{p}_{ij}$$

by

$$\bar{p}_{ij} = (\rho_{ij} - \gamma_{ij} - \beta) \cdot \bar{v}_{ij} ,$$

the profit formula (1.6.2) becomes

$$(1.7.4) \quad P = L \cdot \sum_{i,j} \bar{p}_{ij} \cdot y_{ij} - c \cdot \sum_i x_i - d \cdot \sum_h z_h - e \cdot w .$$

If the demand level $L > 0$ is sufficiently small, then no network configuration will turn a profit. In other words, the "empty" network will be optimal. As L increases, it will reach a "break-point" just beyond which some non-empty network configuration will be optimal. The former network will be called the

break-even configuration ,

because if this network does not break even, for the actual demand level, then no other non-empty network will. As L increases beyond the first break-point, it will eventually reach a second break-point, just beyond which a second network configuration will be optimal. This process continues until finally all potential zones are included.

The result will be an increasing sequence of break-points

$$(1.7.5) \quad 0 < L_1 < L_2 < \dots ,$$

together with a sequence of optimal network configurations, each optimal for values of L between two particular successive break-points. The full network will be optimal for all values beyond the last break-point. It can be shown--a proof would require too much mathematical background material to be presented here--that the sequence of configurations is "nested" as L increases; that is, each network configuration is contained in the next one. The break-even configuration, in particular, will be contained in all subsequent optimal networks corresponding to large values of L .

The reference profits can all be multiplied by a common positive factor, and this does not influence the optimization except for rescaling the break-points (1.7.5). The break-even network, in particular, remains unaffected. In other words, only the relative values of the reference profits--and therefore only the relative values of the reference volumes--matter as far as the parametric analysis is concerned. This is, of course, the main reason for considering a parametric approach in the first place. It should be noted, furthermore, that this relativity extends to the seed costs c , d , and e . Indeed, replacing them by proportional values

$$c \rightarrow Mc, d \rightarrow Md, e \rightarrow Me,$$

where $M > 0$, leads to the profit

$$P = L \cdot \sum_{i,j} \bar{p}_{ij} \cdot y_{ij} - Mc \cdot \sum_i x_i - Md \cdot \sum_h z_h - Me \cdot w.$$

Dividing the function which is to be maximized by a positive number does not affect the optimization. In particular, dividing the above expression for the profit by M yields the original parametric problem, except that L is replaced by L/M , again causing the break-points (1.7.5) to be rescaled.

An even further reduction in data requirements results from the following

Flatness assumption: The per item rates charged, as well as the direct costs incurred, are independent of the origin and the destination of the EMS-item.

The locational independence of the (small) direct costs is expected to hold for most systems. The flat rate on a national level is in tune with present rate policies and has many obvious advantages.

Under the above flatness assumption, we have

$$\rho = \rho_{ij}, \quad \gamma = \gamma_{ij}$$

for the per item revenues and costs respectively. Since the build-up costs are already independent of location, one has a locationally independent

$$\text{per-item operating profit} \quad \pi = \rho - \gamma - \beta.$$

The reference profits \bar{p}_{ij} are proportional to the reference volumes \bar{v}_{ij} --and we have just seen that such a rescaling does not affect the optimization. The parametric problem can therefore be formulated in terms of a rescaled parameter

$$(1.7.6) \quad K = \pi L$$

as maximizing

$$(1.7.7) \quad P = K \cdot \sum_{i,j} \bar{v}_{ij} \cdot y_{ij} - c \cdot \sum_i x_i - d \cdot \sum_h y_h - e \cdot w.$$

Thus under the flatness assumption, the rates and the direct costs do not change the relative values of the reference profits, and the latter can be replaced by reference volumes. Consequently,

the break-even configuration and the nested sequence of additional optimal network configurations depend exclusively on the relative values of the reference volumes and the relative values of the three seed costs c , d , and e .

1.8 Demand Elasticity

In what follows, we assume flat rates and flat direct costs, although similar analyses can be conducted for more general cases.

The general demand level L will be in fact rate-dependent. Specifically, one may assume some demand function, say, of the form

$$(1.8.1) \quad L(\rho) = L_0 \cdot \exp(-\rho/a) .$$

This assumption corresponds to that of a linear demand elasticity of $-\rho/a$.

The profit level $K = \pi L$ (1.7.6) will then also be a function of ρ ,

$$K(\rho) = (\rho - \gamma - \beta) \cdot L_0 \cdot \exp(-\rho/a) .$$

Observe that

$$K'(\rho) = \frac{K(\rho)}{\rho - \gamma - \beta} - \frac{K(\rho)}{a} .$$

It follows that $K(\rho)$ assumes a maximum of

$$(1.8.2) \quad K_{\max} = a \cdot L_0 \cdot \exp(-1 - (\gamma + \beta)/a)$$

for

$$\rho_{\max} = a + \gamma + \beta .$$

This is an optimal rate selection with respect to the above demand assumption. PAREC can be used to identify the optimal configuration for this parameter value.

Demand functions of the form

$$L(\rho) = L_0 \cdot (\rho/a)^{-\epsilon} = L_0 \cdot \exp(-\epsilon \cdot \log(\rho/a))$$

are more commonly used. They are characterized by the constant elasticity $-\epsilon$. The profit function then takes the form

$$K(\rho) = (\rho - \gamma - \beta) \cdot L_0 \cdot \exp(-\epsilon \cdot \log(\rho/a))$$

It follows that the optimal rate is of the form

$$\rho_{\max} = (\gamma + \beta)/(1 - 1/\epsilon)$$

Note that this optimal rate exists only if $\epsilon > 1$. Otherwise, the profit increases steadily with the rate ρ .

We feel the first of the two demand functions discussed here is more realistic in the context of demand for mail delivery.

1.9 Return on Investment

Several different definitions of return on investment (= ROI) have been used by analysts in different contexts. The reason for this lies in varying interpretations of the term "investment." Does it refer only to one-time outlays or does it include recurring annual expenses? Does one restrict oneself to purchases, and exclude leasing? (The difference may be irrelevant for planning, but crucial for tax purposes.) Is "investment" restricted to capital goods or does it include the acquisition of services, "know-how," "good will," etc.?

For the purposes of assessing the performance of an EMSS, we advance the following concept of an initial investment consisting of the total seed costs for installations and overhead. In terms of the indicators in Section 1.6, the annual equivalent of the initial investment is therefore given by

$$c \cdot \sum_i x_i + d \cdot \sum_h z_h + e \cdot w .$$

The return on annualized investment thus takes the form:

$$ROAI = \frac{\sum_{i,j} p_{ij} \cdot y_{ij} - (c \cdot \sum_i x_i + d \cdot \sum_h z_h + e \cdot w)}{(c \cdot \sum_i x_i + d \cdot \sum_h z_h + e \cdot w)}$$

We state without proof that

The break-even configuration maximizes the return on annualized investment (provided there is profit at all).

If the promotional effect of the number of terminals on the general demand level L is taken into account, then a slightly larger configuration might be required to maximize the return on initial investment.

1.10 The Impact of System Overhead

In problem formulation (1.6.4) the "objective function," namely the profit to be optimized, takes the form

$$P = \sum_{i,j} p_{ij} \cdot y_{ij} - c \cdot \sum_i x_i - d \cdot \sum_h z_h - e \cdot w .$$

If the optimal solution is a non-empty configuration, then $w = 1$. The term ew then acts most of the time as an additive constant to the objective function, suggesting the question whether the term ew could be ignored.

Clearly, the term ew cannot be ignored altogether, since the system overhead significantly affects the break-even pattern and the return on investment. It will turn out, however, that one can solve without the term ew and then adjust for it later.

Specifically, consider the reduced problem (compare (1.6.4)):

$$\begin{aligned}
 (1.10.1) \quad & \text{Maximize} \quad \sum_{i,j} p_{ij} \cdot y_{ij} - c \cdot \sum_i x_i - d \cdot \sum_h z_h \\
 & \text{subject to} \quad y_{ij} \leq x_i, \quad y_{ij} \leq x_j, \quad x_i \leq z_h(i) \\
 & \text{and} \quad x_i = \begin{cases} 0 \\ 1 \end{cases},
 \end{aligned}$$

where the term $e \cdot w$ and the variable w have been dropped.

Now let

$$\hat{y}_{ij}, \hat{x}_i, \hat{z}_h$$

be an optimal solution to (1.10.1). The connection between the two problems (1.6.4) and (1.10.1) is then expressed by the following observation:

If

$$(1.10.2) \quad \sum_{i,j} p_{ij} \cdot \hat{y}_{ij} - c \cdot \sum_i \hat{x}_i - d \cdot \sum_h \hat{z}_h > e,$$

i.e., if the profit after deducting station costs and terminal costs covers the seed costs for system overhead, then

$$\hat{y}_{ij}, \hat{x}_i, \hat{z}_h, \hat{w} = 1$$

is an optimal solution to (1.6.4). Roughly speaking, both (1.6.4) and (1.10.1) are optimized by the same configuration. If, however, (1.10.2) is not satisfied, i.e. if there will be no profit after including the station seed cost e , then the empty configuration will be best for (1.6.4).

When the above result is applied to the parametric analysis, it follows that the sequence of optimal networks generated in Section 1.7 for $e > 0$ is a subsequence of the sequence of optimal networks for $e = 0$. More precisely, the break-even network for $e > 0$ is one of the optimal networks in the sequence for $e = 0$, and all subsequent networks are common to both sequences. Corresponding break-points (1.7.5) will be equal, except for the first one in the case $e > 0$.

If

$$\begin{matrix} (1) & (1) & (1) \\ y_{ij} & , & x_i & , & z_h \end{matrix}$$

is the break-even solution for $e = 0$, and if one puts for some $e_0 > 0$,

$$\bar{L} = \frac{c \cdot \sum_i^{(1)} x_i + d \cdot \sum_h^{(1)} z_h + e_0}{\sum_{i,j}^{(1)} \bar{p}_{ij} \cdot y_{ij}},$$

where \bar{p}_{ij} are the reference profits of section 1.7 (compare (1.7.4)), then $e = e_0$ and $e = 0$ will have the same break-even configuration if and only if

$$(1.10.3) \quad L_1 < \bar{L} \leq L_2,$$

where L_1 and L_2 are the first two break-points in the case $e = 0$.

If (1.10.3) is satisfied--and the two break-even configurations therefore coincide--then all subsequent networks will coincide also. In the case $e = e_0$, the first break-point will be \bar{L} rather than L_1 . All subsequent break-points are identical.

Note that $e_0 > 0$ implies $\bar{L} > L_1$. Hence (1.10.3) can be violated only by having $\bar{L} > L_2$. In this case, let

$$\begin{matrix} (2) & (2) & (2) \\ y_{ij} & , & x_i & , & z_h \end{matrix}$$

correspond to the second optimal configuration in the sequence for $e = 0$. If the value $e = e_0$ violates (1.10.3), redefine

$$\bar{L} = \frac{c \cdot \sum_i^{(2)} x_i + d \cdot \sum_h^{(2)} z_h + e_0}{\sum_{i,j} \bar{p}_{ij} \cdot y_{ij}}.$$

Clearly, $L_2 < \bar{L}$. If $\bar{L} \leq L_3$, where L_3 is the third break-point in the $e = 0$ sequence, then the break-even configuration for $e = e_0$ is the second optimal configuration for $e = 0$, and so on.

In this fashion, one can determine the break-even configuration for any $e > 0$ from the sequence of network configurations which are optimal when $e = 0$. As noted earlier, this also applies to the problem of maximizing return on investment.

1.11 Extensions of PAREC

The methodology for cost-revenue analysis of configurations, developed by J. M. W. RHYS [5] and extended in the present document, is quite general and can be applied to systems other than EMSS. In particular, it is flexible enough to accommodate different kinds of EMSS configurations, i.e. configurations involving other installations besides stations and terminals.

For instance, an EMSS may provide a special service capable of reproducing legal documents, engineering drawings and the like. This would require the installation of expensive special facsimile equipment at terminals. The demand for the service, on the other hand, may not support having such special equipment at every terminal, and one would like to know which terminals to select.

This type of analysis will lead to a different kind of network, in which there are now two links between some parts and which involves seed costs of a third kind of installation, namely special facsimile equipment, and estimated annual operating profits for the special services in addition to the estimated annual operating profits p_{ij} (see Section 1.6) for the main services. PAREC's input processing and those PAREC portions which generate the "dual" network would have to be generalized to handle this more involved problem, but its overall structure and--most importantly--its solution algorithms would not have to be changed.

Other extensions require iterative applications of PAREC. For instance, instead of introducing seed costs for installations, one may want to specify actual numbers of modules. In this case, a run with seed costs may be executed in order to arrive at estimates for volumes to be processed. The number of modules is then selected for each installation on the basis of the estimate, and the run is repeated with the corresponding revised installation costs. PAREC would have to be changed only slightly to allow different annual indirect costs at each installation.

The RHYS approach is indeed remarkable for its breadth and adaptability to a wide range of problem formulations.

2. Interpreting the Results of PAREC

When discussing the information required by an analyst planning to apply PAREC, we have to distinguish between computer inputs which are needed for the actual computer run, and interpretive data, which help provide the economic interpretation for the results of PAREC runs. This way of organizing the model's application has the disadvantage of requiring an additional, albeit very moderate, effort of hand calculation using PAREC-produced diagrams. On the other hand, there is a considerable advantage in that a single run of PAREC can serve to analyse a variety of economic parameters simultaneously.

In what follows, we will briefly describe both the computer inputs to PAREC and the interpretive data. We will then show how the results of a PAREC run and the interpretive data can be combined to provide answers to various questions concerning the economic performance of an EMS technology.

2.1 Computer Inputs and Interpretive Data

The first decision faced by the analyst who plans to use PAREC is to define the network of potential zones and regions (see Section 1.2). Each potential zone will consist of one or more ODIS areas as specified by the user. A potential zone will be served by a single terminal if served at all. The analyst then groups the potential zones into potential regions, each to be served by a single station. This description of the full potential network is a computer input. See section 3.3 for details.

ODIS data are also computer inputs. They determine the typical mail volumes between pairs of potential zones. These volumes will serve as relative demand indicators or reference volumes. Reference volumes may be given on a daily or annual basis, differing by a factor of 300.

Finally two nonnegative integer numbers C and D are specified. The ratio of the two numbers should reflect the ratio c/d of annual (seed) costs per terminal ($= c$) to the annual (seed) costs per station ($= d$):

$$\frac{C}{D} = \frac{c}{d}$$

The actual sizes of C and D , as selected by the analyst for input into PAREC, determine the "coarseness" with which intermediate networks are analysed by PAREC. If C and D are large, run-times are expected to be shortened, possibly at the expense of skipping an intermediate configuration (see Section 3.2).

The computer input to PAREC thus represents network relationships, the relative annual seed costs of facilities, and relative demands between potential zones. This input is sufficient to produce the sequence of nested preferred configurations.

Note that the computer inputs do not include the general demand level, the costs of facilities or the rates per item. All these data are necessary for the economic evaluation. However, they are not needed for the computer run of PAREC, and therefore constitute interpretive data. A complete list of interpretive data is as follows (compare Section 1.6):

- L = general demand level
- c = annual (seed) costs per terminal
- d = annual (seed) costs per station
- e = annual (seed) costs per system overhead
- r = per item direct net revenue
- p = rate minus per item direct cost

Note that the interpretive inputs represent the totals of many individual costs, some of them annualized. Sections 1.3 through 1.5 describe the process by which the analyst arrives at these totals. This process has not been computerized so far, since it is a non-iterative process and since some of it relies on the judgement of the analyst.

2.2 The VF-diagram

For the sake of simplicity of exposition, we assume in this section that each particular region consists of just one potential zone. This amounts to assuming that each station and its corresponding terminal are considered as one unit (although not necessarily placed in the same location). We assume further that all these facilities are of the same size, and we denote the annual (seed) cost for each combined facility by c . (Previously, we used this letter to denote the annual (seed) costs just for terminals.) Later on we will discuss more general situations.

Under our simplifying assumption the size of the network can be measured by the number n where

$$n = \text{number of facilities.}$$

Whenever two facilities i and j are included in the network configuration, then the demand for EMS between the corresponding potential zones will be "captured." This captured demand will be the ODIS derived reference volume times the general demand level parameter L . Since L is a common multiplier, it will be advantageous to measure the captured demand in terms of reference volumes rather than actual volumes.

Given the number n of facilities, the question arises how to optimally deploy these facilities, i.e. what potential zones to schedule so that the captured reference volume (and therefore the captured actual volume) will be maximized. Given a solution to this problem for each n (there exist approximate solution methods; see Gardner [6] and Witzgall [1]), it is possible to plot the maximum captured volume $V(n)$ versus the number of facilities. We call the corresponding diagram a VF-diagram (Fig. 6).

The profit for each configuration with $n > 0$ is given by

$$(2.2.1) \quad P(n) = L \cdot V(n) \cdot r - n \cdot c - e.$$

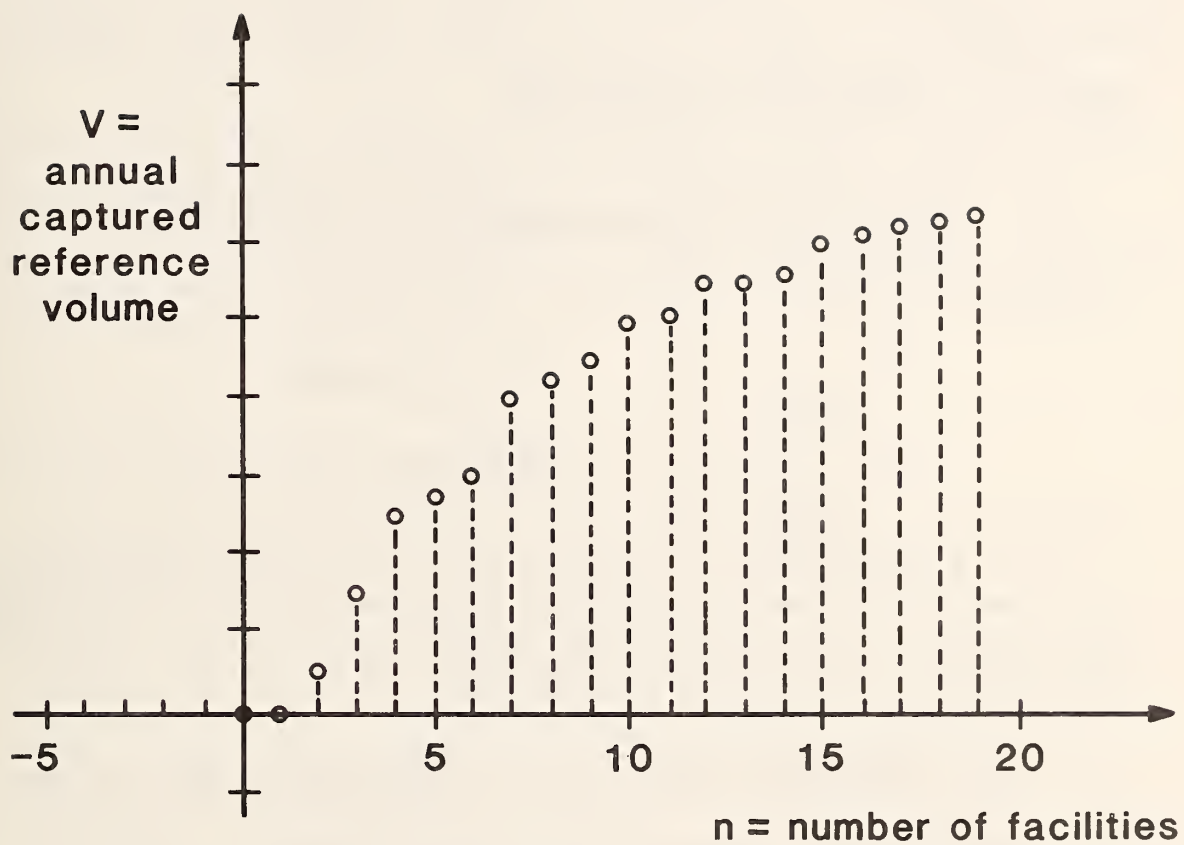


Figure 6: VF-diagram for single (combined terminal and station) facilities.

Thus

$$\frac{P(n)}{L \cdot r} = V(n) - nh - g$$

where

$$(2.2.2) \quad h = \frac{c}{L \cdot r}, \quad g = \frac{e}{L \cdot r}.$$

We now ask

Question 1: Given the interpretive data

$$L, c, e, r,$$

what is the most profitable configuration size n ?

To answer this question, we note that for $k > 0$:

$$\begin{aligned} \frac{P(n_0 + k) - P(n_0)}{L \cdot r \cdot k} &= \frac{V(n_0 + k) - V(n_0)}{k} - h \\ \frac{P(n_0) - P(n_0 - k)}{L \cdot r \cdot k} &= \frac{V(n_0) - V(n_0 - k)}{k} - h. \end{aligned}$$

If n_0 is a profit-optimal configuration size, then the first of the above expressions must be nonpositive and the second expression must be nonnegative for all values $k > 0$ for which the expressions are defined. The above conditions for a profit-optimal configuration can be combined into the following necessary condition for optimality:

$$(2.2.3) \quad \frac{V(n_0 + k) - V(n_0)}{k} \leq h \leq \frac{V(n_0) - V(n_0 - k)}{k}, \quad k > 0.$$

This condition is also sufficient for profit-optimality if $P(n_0) > 0$, i.e., if

$$(2.2.4) \quad \frac{V(n_0)}{h} - n_0 \geq \frac{g}{h} = \frac{e}{c}.$$

In other words, (2.2.3) characterizes the best configuration without considering system overhead. If the latter is covered, then n_0 is indeed profit-optimal. If system overhead is too large to be covered, then $n_0 = 0$ is profit-optimal, i.e. it is best not to build a system at all.

Condition (2.2.3) can be interpreted geometrically. The quotients

$$\frac{V(n_0 + k) - V(n_0)}{k}, \quad \frac{V(n_0) - V(n_0 - k)}{k}$$

represent slopes of "chords" through two points of the VF-diagram, namely,

$$(n_0, V(n_0)), (n_0 + k, V(n_0 + k))$$

and

$$(n_0 - k, V(n_0 - k)), (n_0, V(n_0)),$$

respectively. Thus (2.2.3) ensures that all chords to the left of the point $(n_0, V(n_0))$ have a slope not less than h , whereas all slopes to the right have a slope not exceeding it. Another way of stating the same fact is to say that the point $(n_0, V(n_0))$ is the first point touched if a straight line of slope h is lowered from above. We call this line the h-line (Fig. 7).

Condition (2.2.4), which completes the optimality criterion, also has a simple geometric interpretation. Indeed, the left part of the inequality (2.2.4) denotes the (positive) distance between zero and the point at which the h-line intersects the horizontal coordinate axis. The ratio e/c expresses the annual system overhead in multiples of annual (fixed) facility costs. We call the point which lies e/c units to the left of zero on the horizontal coordinate axis the overhead point.

We are now able to answer Question 1.

Answer to Question 1 based on VF-diagram:

A configuration represented by a point $(n_0, V(n_0))$ in the VF-diagram is most profitable only if $(n_0, V(n_0))$ is a contact point of the h-line with

$$h = \frac{c}{L \cdot r}.$$

To make entirely sure that a configuration which passes the above test is most profitable, its profitability as such has to be established: if this particular configuration is profitable, it is also the most profitable. To decide whether the configuration is profitable, find the place at which the h-line intersects the horizontal coordinate axis: if to the left of the overhead point, the configuration is profitable; otherwise, it is not. The distance between the two axis points indicates the amount of profit (or loss) in terms of multiples of facility costs (Fig. 7).

Note that the optimality of a configuration depends first of all on the ratio h (2.2.2), and next on the ratio e/c . These are the main interpretive data. The smaller h , the better the economic outlook. For a method of determining the rate so as to optimize h see Section 1.8. Many points

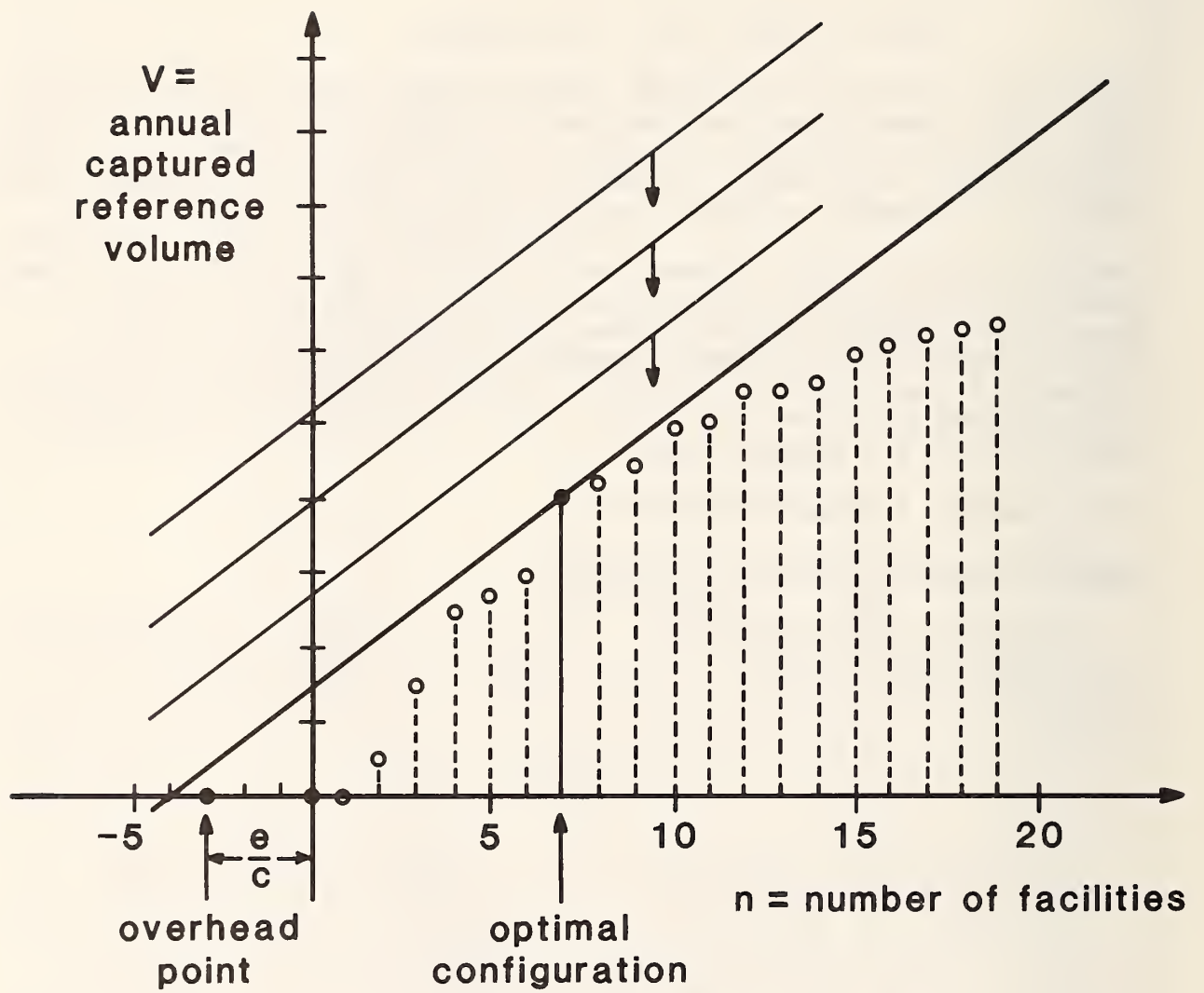


Figure 7: h-line and overhead point in VF-diagram

$(n, V(n))$ in the VF-diagram will not correspond to profit-optimal configurations since they can never be first touched by lowering a line of any slope. Those that can be touched, thus satisfying the necessary condition (2.2.3) for some slope h , characterize what we call preferred configurations. They are the upper vertices of the "convex hull" of the points $(n, V(n))$ (Fig. 8). The PAREC algorithm determines the sequence of preferred configurations. It turns out that the preferred configurations are nested.

The preferred configurations are potentially profit-optimal for some slope h if they lie on the h -line, i.e., if they satisfy (2.2.3) but not necessarily (2.2.4). If only preferred configurations are considered in the FV-diagram, then the following question can be answered:

Question 2: What are the "break-points" $h_0 > h_1 > h_2 > \dots > h_m > 0$ such that the same configuration is potentially profit-optimal for $h_j \leq h \leq h_{j-1}$ but not for $h > h_{j-1}$ or $h < h_j$?

Answer to Question 2 based on VF-diagram:

Draw the chord through the point $(0,0)$ and the closest preferred point $(n_0, V(n_0))$ to the right. Then h_0 is the slope of this chord. The empty configuration is profit-optimal if $h \geq h_0$. Next draw the chord through the point $(n_0, V(n_0))$ and the next preferred point to the right. h_1 is the slope of this chord. The "break-even configuration" represented by $(n_0, V(n_0))$ is potentially profit-optimal for $h_1 \leq h \leq h_0$, and so on. In other words, the boundary slopes of the convex hull are the desired break-points h_j .

The annual system overhead e is a critical quantity. It contains central administrative, promotional, and developmental costs. But most of all, it is determined by the costs of a central communications system, be it an owned or leased communications satellite, or the fixed portion of yearly payments on a communications contract. It is therefore natural to ask

Question 3: Given annual facility costs c and annual system overhead costs e , what is the biggest value h for

$$h = \frac{c}{L \cdot r}$$

which will break even and what configuration will achieve this?

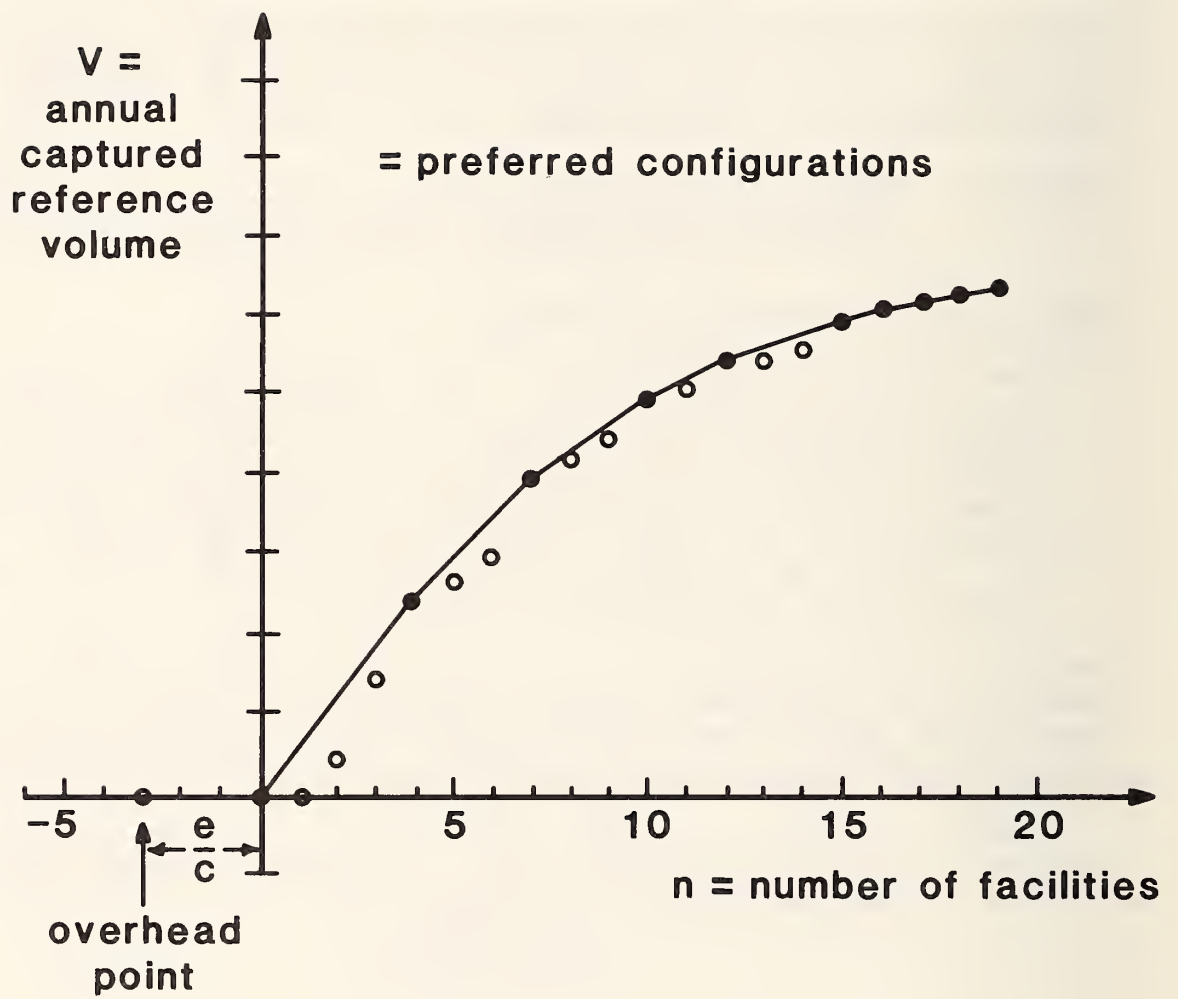


Figure 8: Preferred configurations and connecting chords

From the previous developments, the answer is clear:

Answer to Question 3 based on VF-diagram:

Pivot a straight line at the overhead point and rotate it downwards from above until it touches a configuration point $(n_0, V(n_0))$. The slope \bar{h} of the line through $(n_0, V(n_0))$ is the break-even value for h . The corresponding configuration is that represented by $(n_0, V(n_0))$. As shown in Section (1.7), this configuration also minimizes the annualized return on investment (Fig. 9).

Question 4: Given a value for

$$h = \frac{c}{L \cdot r}$$

and the ratio

$$\frac{e}{c}$$

of system overhead versus single facility cost, what are the smallest and the largest configuration sizes for which the EMSS breaks even?

Answer to Question 4 based on VF-diagram:

Draw a line of slope h through the overhead point and find its left and right intersection points with the curve of the diagram. The sizes n_1 n_2 at which these intersections occur, closely approximate the minimum and maximum break-even sizes.

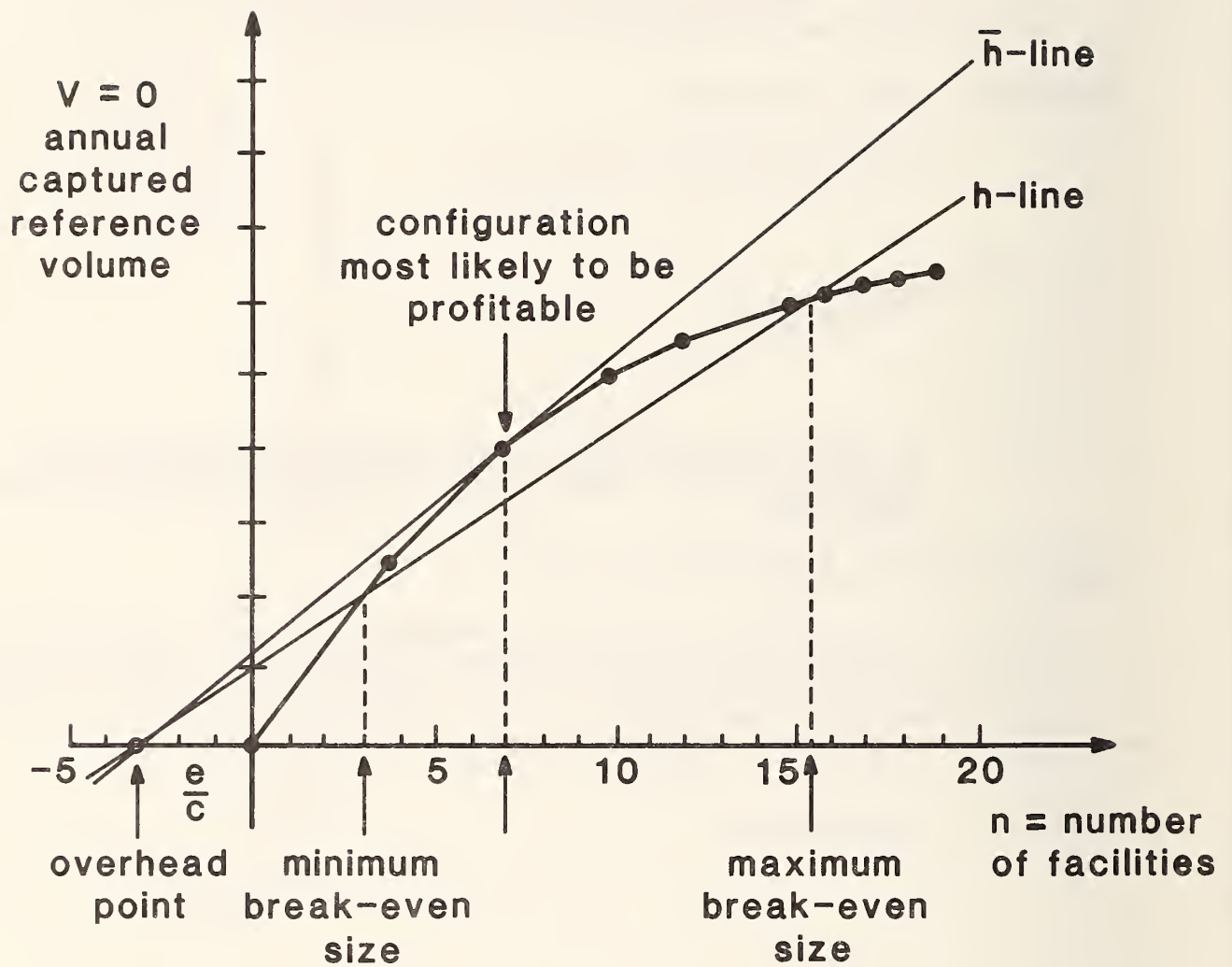


Figure 9: Answers to Questions 3 and 4

Since reference volumes and number of facilities are rather arbitrary measures of actual volume and cost, respectively, it is sometimes useful to present the VF-diagram on a percentage basis. In this case, the values h and e/c have to be rescaled.

2.3 The HF-diagram and the VFF-diagram

Two other diagrams are sometimes useful because the VF-curve tends to be rather close to linear and this makes the determination of the h -line inaccurate. The HF-diagram plots the break-points

$$h_0 > h_1 > \dots > 0$$

of question 2 in section 2.2 versus the configuration size (Fig. 10). A horizontal line drawn at height h will then intersect the HF-diagram at the profit-optimal configuration size.

The VFF-diagram plots $V(n)/n$ rather than $V(n)$. Its peak indicates the break-even configuration. Minimum and maximum break-even sizes, disregarding overhead costs, are indicated by the intersections of the horizontal line at height h with the curve of the VFF-diagram.

Since both diagrams have the same coordinate units, we can draw them together in the same coordinate system (Fig. 10).

2.4 Investment Units

If two different kinds of facilities, terminals and stations, are considered for deployment and some regions consist of several zones, then we cannot measure system size by the number of facilities. For this reason, we introduce an arbitrary investment unit relative to which we express facility sizes. Stating, for instance, that a terminal amounts to 3 investment units and a station to 5 investment units means that the annual fixed terminal costs and the annual fixed station costs are in the ratio 3 to 5. Similarly, system overhead e is expressed in terms of investment units.

The VF-diagram then lists best configurations consisting of a prescribed number of investment units. Heuristic procedures for solving the deployment problem are very complicated and expensive. Again, however, there are preferred configurations which can be found efficiently using the PAREC algorithm. The annual reference volumes captured by these configurations can be plotted against the number of investment units in an VF-diagram.

A related generalization concerns the case of different size terminals and different size stations specified for different zones and regions, respectively. For instance, one might want to specify that a terminal in Manhattan should be twice the size of a terminal in Atlanta. Such size specifications can be expressed in investment units. PAREC methodology is capable of handling this case. However, the present computer version of PAREC still assumes that terminals are of the same size regardless of location. The same is true for stations. Adjusting the PAREC code to handle different size

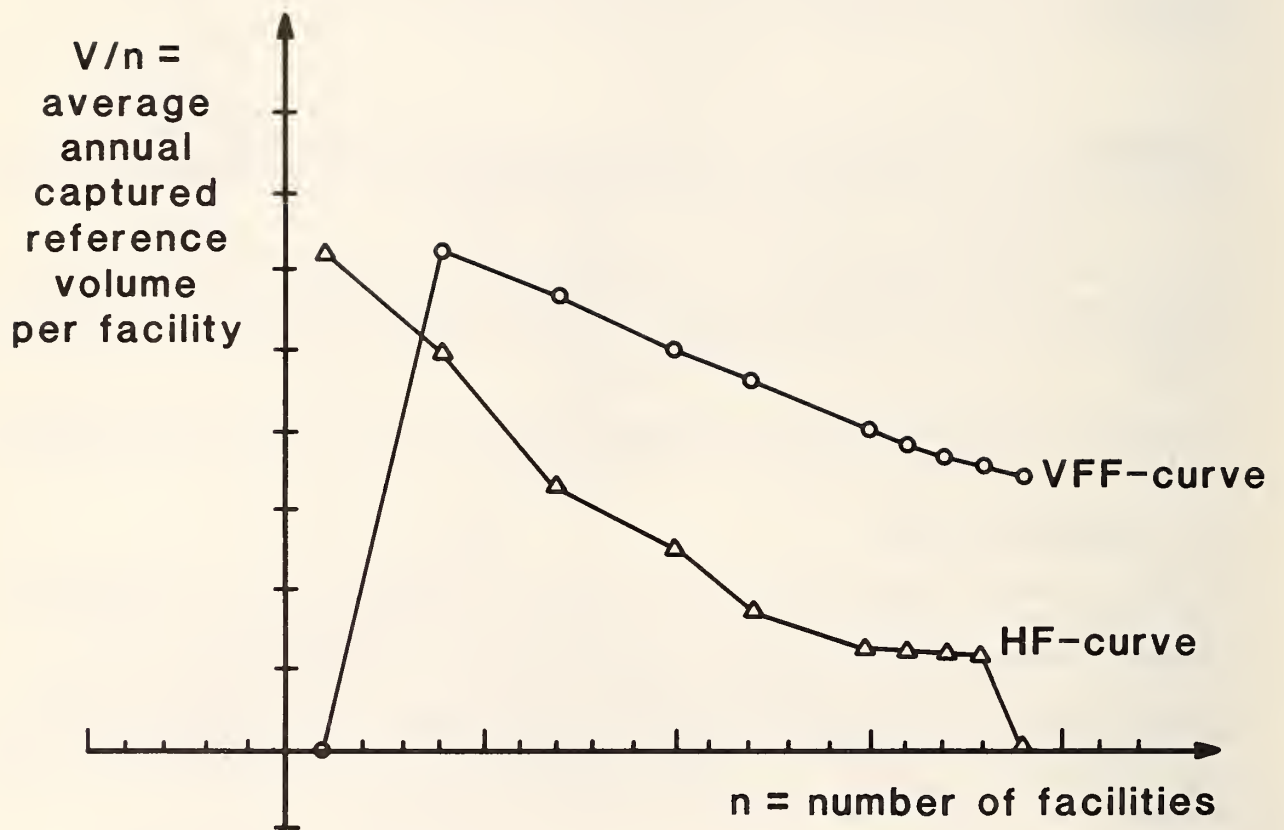


Figure 10: VFF and HF-diagrams

specifications has been deferred since no immediate need for it has been apparent. Note that different size specifications for facilities pose a problem different from that of having the model determine the required size of each facility based on demand. That problem is handled by the introduction of seed costs and build-up costs, as will be briefly described in the next section.

The use of the VF-diagram in the case of investment units is still essentially the same as outlined in Section 2.2. The only difference is that h is now defined as the ratio

$$h = \frac{f}{L \cdot r} ,$$

where

f = annual (seed) cost per investment unit.

When marking the overhead point, the ratio e/c is to be replaced by the number of investment units measuring system overhead.

2.5 Using Seed Costs

Seed costs can be used to estimate the sizes of individual terminals and stations--as well as system overhead size--in response to demand incurred at these installations.

The technique is experimental. It is not part of the PAREC model as such, but rather an option of the interpretation which is carried out separately. How to determine seed costs is described in Section 1.4.

Suppose seed costs and the concurrent build-up costs have been determined. Then a particular configuration, say, an optimal one, is selected for further examination. This includes determining reference volumes of incoming and outgoing EMS items at each individual zone and each individual station. These volumes are multiplied by the demand level parameter L to arrive at the actual demand forecasts. These are then multiplied by the build-up costs, arriving at an annual build-up cost for each individual installation. Analogously, total volume is multiplied by overhead build-up costs. These costs are then interpreted as annual fixed costs at these installations.

The seed-cost technique of estimating installation size is intended as an approximation and should probably be employed on an iterative basis. More development is needed for this purpose.

3. How to Run PAREC

The problem to be solved and the solution method employed have been described in Section 1 of this document. Section 3 is directed towards a hypothetical analyst/user of the PAREC program who will be concerned with input, output, and having a sufficient general knowledge of the program to understand the implications of the input parameters and the nature of the output reports. Details of the program will be discussed in Section 4. Although operational versions of the PAREC program may no longer be available at this time, the more detailed descriptions of the mode of use and the structure of the program will supplement the more theoretical material in the previous sections. It will also provide guidelines for implementation as well as documentation should the need for an analogous model develop.

3.1 Overview

The scenario upon which the model operates is described by a set of nonoverlapping potential zones. The potential zones are grouped into regions, such that the terminals which one might install in the zones of such a region will connect to a single station for that region. The problem then is to select the actual zones from among the potential zones so as to optimize profit. The annual fixed (seed) costs c (per terminal) and d (per station) are expressed in investment units (see Section 2.4), the size of which is chosen by the user. To allow for uncertainties in the demand estimates, a parametric approach to this optimization problem has been developed in Section 1.7, and PAREC implements this parametric approach.

The parametric approach is based on reference volumes, which express relative magnitudes for each pair of potential zones. The real annual volumes derive from these reference volumes by multiplication with a general demand level. The parametric approach, as described in Section 1.7, considers this demand level as a parameter to be increased from zero on up. In this process an increasing sequence of nested networks is created, each optimal for some demand level bracket. The first such non-empty network is of special interest as it represents the break-even network and maximizes return on annualized investment.

The PAREC program now differs in its parametric approach from the procedure outlined in Section 1.7. Instead of multiplying the reference volumes by a parameter, it multiplies the total station and terminal costs by another parameter, essentially the reciprocal of the first. Varying this other parameter from zero on up will produce the same networks as previously, but in reverse sequence. Thus one starts with the network that includes all potential zones, and successively deletes batches of zones until nothing is left. The last non-empty network then represents the break-even configuration and maximizes return on investment disregarding system overhead.

Two features are available which were not mentioned in Part 1. First, it is possible to force in some of the potential zones. Those zones will be included in any network produced by PAREC without regard to their profitability. Second, a locality factor can be specified, which multiplies those reference volumes which correspond to intra-regional demand. The reason for this feature is to permit an easy evaluation of different assumptions about the extent and the costs of intra-regional communications.

The PAREC program is set up for solving one or more cases in a single run. This saves computer time in that the expensive operations of processing the tape of reference volumes (see below) and setting up the dual network (see Section 1.6 and Appendix A) are carried out only once in each run. The various cases may differ in their terminal and station seed costs as well as in their selection of forced-in zones.

The typical application of PAREC will involve the use of the ODIS data base. From this data base, current annual mail volumes--first class, air mail or mixtures thereof--can be extracted for pairs of ODIS-areas, henceforth called simply "areas". The idea is to use these annual volumes as reference volumes for PAREC. A magnetic tape with this information, the ODIS-tape, has to be provided for PAREC. (Recall that only the relative magnitude of the reference volumes matters in most contexts, and that therefore daily ODIS-volumes could be used instead of annual ODIS-volumes.)

The areas for which such volume data are specified on the ODIS-tape constitute the raw material from which the potential zones are defined by selection and aggregation. First those areas which are to be included in the analysis are selected. This selection feature has been included to avoid the costly task of regenerating the ODIS-tape whenever some areas are disregarded. In most cases, these selected areas will already be the potential zones from which the optimal selection is to be made. The analyst, however, has the option of defining aggregate zones consisting of more than one area. The reference volumes connected with such zones are aggregated by PAREC from the reference volumes of their respective areas. All selected areas which are not included in any aggregate zone are considered stand-alone zones. Henceforth the single word "zones" will refer to both aggregate and stand-alone zones. Zones may be grouped into aggregate regions - as mentioned before - in order to permit their sharing of common stations.

3.2 General Program Comments

The program PAREC was written and debugged at NBS on a UNIVAC 1108 under the EXEC 8 operating system and a FORTRAN V compiler. It has also been made operational at the USPS on an IBM 360/65 with the FORTRAN IV, G compiler. The program consists of a main routine, eleven major subroutines, and two utility subroutines used in association with reading and validating input. All subroutines are explained in detail in Section 2.7.

Storage requirements for PAREC depend largely on the numbers of area codes, selected areas, aggregate zones, zones, and regions. On the UNIVAC the code and minor local variables occupy about 6K words of storage. The following expression is used to calculate the storage needed for variables associated with the potential EMS-network and the dual network:

$$2*JODI+10*SARA+6*TNZO+27*INZO+4*INZO(INZO-1)+21*KNRG+ZMAX+6$$

where

JODI = total number of areas,
 SARA = number of selected areas,
 TNZO = number of aggregate zones,
 INZO = total number of zones,
 KNRG = number of aggregated (i.e., multi-zone) regions
 ZMAX = work array used as a stack to keep a record of out-of-kilter arcs; suggested size is about 250 words.

For a 100 area scenario, where all areas are selected and there are no aggregate regions, i.e. JODI = SARA = INZO = 100 and TNZO = KNRG = 0, the variables require about 44K words for storage. (Since most of the variables are integer, many with known upper bounds, the storage requirements could be reduced on the IBM 360/65 by taking advantage of the 1/2 word and 1/4 word storage capabilities.)

The basic logical structure of PAREC is illustrated in the flow diagram of Fig. 11. The first part of the program is devoted entirely to input and storage. As discussed previously, the main thrust of the model is the solution of a parametric network flow problem. This is accomplished in the program by starting with very low (zero) seed costs, solving a parametric flow problem, increasing the costs, modifying the flows, and solving another parametric flow problem. As this process continues and the costs increase, the size of the EMS-network decreases, starting from a full network of all potential zones and regions and decreasing to a minimal network which contains, at most, only those zones which are forced to be a part of the network regardless of cost.

During this process, when the EMS-network changes by dropping one or more zones, a "step" is said to occur, some output is printed, and certain data are stored to describe the optimal EMS-network at each such step. Note that a step does not necessarily occur each time that a network flow problem is solved: due to so-called degeneracy effects in linear programming, several parameter changes may be needed before the optimal solution changes, that is, one or more zones are dropped. The user supplies, therefore, a maximum on the number of iterations, i.e., the number of passes through the network flow routine to be made without dropping at least one zone. If after this maximum number of iterations there is no step, the program terminates. The iteration count is reset to zero after each step.

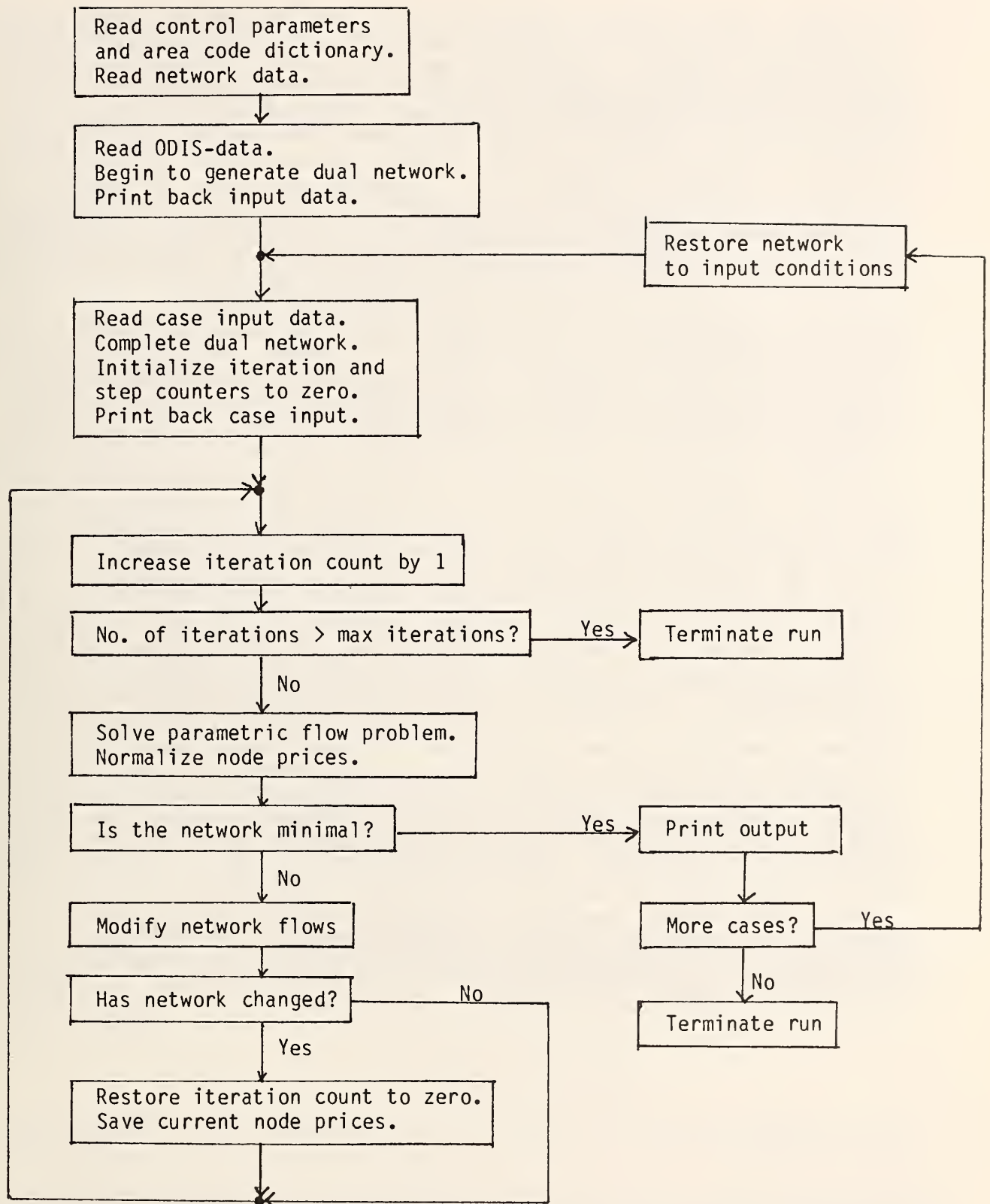


Figure 11: Flow diagram of PAREC

Eventually in the solution process the costs will increase to values such that the optimal EMS-network is minimal. When this occurs, a detailed description of the minimal network is printed as well as a summary of network statistics (size, volume, etc.) for each step. The entire parametric solution procedure is then said to be complete for one case. A case is defined as the specification of values of the seed costs and a list of any zones or regions which are forced in to the solution. PAREC allows the user to process more than one case during a computer run. If another case is to be processed, then before reading the input data for the next case, the potential EMS-network is restored to the original input conditions, i.e. volumes are recalculated and the restrictions which force zones and regions into all solutions are removed. When all cases have been processed the program terminates.

Run time is determined not only by the size of the potential EMS-network, but also by the number of investment units, in which the terminal and station costs, c and d , are expressed. If the investment units are selected to be large, thus leading to "small" cost parameters C and D expressing c and d , respectively, then each step will require many iterations, i.e. many small increases in costs will occur before the total cost increase is sufficient to require the exclusion of one or more zones and regions from an optimal network solution. Apparently because there is so little change between iterations, the time per iteration is quite small, but many iterations are needed to accomplish a step. However, when the costs are sufficiently small, the user can be fairly certain of obtaining all steps, i.e. each time there is a change in the set of optimal zones and regions, the data associated with the change will be available to the user. On the other hand, if the cost parameters are "large", there will be few iterations per step and, in fact, the process may even skip over steps which would appear if smaller costs were used. In an extreme case, it is possible for the seed costs to be so large that the program produces only two optimal network configurations, the zero cost maximal network (all zones included) and the large cost minimal network (including only forced in zones). When cost parameters are large, there are relatively large changes in the dual network between iterations, which increases the time per iteration. Obviously, somewhere between the "small" seed costs and the "large" seed costs there exists a range of seed cost values which are most appropriate. As the user gains experience in applying PAREC, he will become adept at selecting such values. On a network of 256 areas, 100 zones, and 22 regions, the total run times (on the UNIVAC 1108, EXEC 8 Operating System at the National Bureau of Standards) varied between approximately 24 minutes and 29.5 minutes for values of the seed costs C fixed at 100 and D ranging from 0 to 200.

3.3 Input Requirements

This section will describe the meanings and formats of data required as input to PAREC. It should be noted when reading the data description that certain groups of data are optional, i.e. under certain run scenarios they need not be supplied.

Control Record. The first input record gives values to the four variables LIMIT, JODI, STOP, and FLOC in the format (3I10, F10.2). As explained in section 4.1, LIMIT is the maximum total amount of storage to be used for all variables relating to both the potential EMS-network and the dual network. LIMIT should not exceed the dimension for the array X as given in the routine MAIN. The present dimension of X is 47500. Further explanation of this is given in section 4.1. JODI is the total number of areas to be considered for selection during the current run. STOP is the maximum number of iterations allowed. FLOC is a locality factor used when volume data are aggregated for regions. It indicates what fraction of the volume between zones in the same region will be considered. When FLOC is equal to 1.0, all intra-regional volume is included.

Area Code Dictionary. This group of input records supplies a three character code for every area which might be encountered during a particular run. The areas which are selected as stand-alone zones or as part of aggregate zones of the potential EMS-network must be identified by one of the codes in this dictionary. There are 20 codes per record, in the format 20(A3,1X) starting in column 1, for a total of $|(JODI-1)/20|+1$ records of this type. (The symbol $||x||$ denotes the largest integer less than or equal to x .) This dictionary must include all areas which are encountered in the data for reference volumes by origin/destination pairs which will be discussed later in this section.

Network Definition Data. Three groups of records may be used to define respectively the areas, zones, and regions for each run. Because the three groups are similar in format, certain remarks applying to all three groups will be made first before giving specific details pertinent to each individual group.

Each group begins with a set-up record which has an asterisk in column 1 and a keyword starting in column 2. The three keywords are "AREAS", "ZONES" and "REGIONS". Following the set-up record for each group are the definition data for that group. The following format is used for the definition data in all three groups:

(1X, A3, 1X, 4A4, 12(1X,A3))

The information being interpreted by this format is described separately for each group.

Area Selection Data. This group of records indicates which of the JODI areas in the area dictionary are to be selected as part of the network for this run. There is one area selection record for each selected area. The maximum number of these records is JODI and the actual number is counted in PAREC and stored in the variable SARA (selected areas). Each record contains two items: the three character code for area and up to 16 characters of a name for the area.

The code must match one of the codes in the area code dictionary and duplications are not allowed, i.e. a code cannot appear on more than one area selection card. There are no restrictions on the name other than the 16 character maximum. The code is in alphanumeric format in columns 2-4; the name is in alphanumeric format, left-justified in columns 6-21.

Zone Definition Data. This group of records defines the aggregation of areas into zones. These data are optional, i.e. if there are no aggregate zones then there should be no data of this type. There is one zone definition record for each aggregate zone. The actual number of these cards is counted and stored by PAREC in the variable TNZO (total number of aggregate zones). Each aggregate zone has associated with it a three character code which may not duplicate any of the area codes in the area code dictionary. Each zone also is given a name of up to 16 characters. There are no other restrictions on these names. An aggregate zone can include up to a maximum of twelve of the selected areas. The areas to be included in each zone are given by their three character area codes. Zones cannot overlap, so that an area code can appear on only one zone definition record. Within the list of area codes for a particular zone, blanks are ignored, but a zone definition card with no area codes is not allowed. It is not necessary to provide a zone definition record for "stand-alone" zones, i.e. zones which consist of only a single area. In summary, each zone definition record contains a zone code, three alphanumeric characters in columns 2-4, a zone name of up to 16 characters in columns 6-21, and a list of up to 12 area codes in columns 23-25, 27-29, 31-33, ... , 67-69, i.e. with the format 12(1X,A3) starting in column 22.

Region Definition Data. This group of records, which starts off with the set-up record "*REGIONS", defines the consolidation of aggregate zones and stand-alone zones, i.e. areas which are not included in any aggregate zone, into regions. There is one region definition record for each region and if there is no consolidation of zones into regions, there should be no data of this type. The number of these records is counted and stored by PAREC in the variable KNRG. For each region the definition record contains a 16 character region name and up to a maximum of 12 area and/or zone codes to be included in the region. (Note that region codes are not used.) When an area code is specified, it must be one of the selected areas and it must be a stand-alone area which is not a part of any aggregate zone. When a zone code is specified, it automatically forces all areas within that zone to be included in the region. Neither area codes nor zone codes can appear on more than one region definition record. Within the list of area and zone codes for each region, blanks are ignored except that all blanks is not allowed. Thus for each region there is a 16 character name in columns 6-21, and a list of up to twelve area or zone codes given by format 12(1X,A3) starting in column 22, i.e. the codes are in columns 23-25, 27-29, 31-33, . . . , 67-69.

Case Data. The purpose of these data is to define particular cases to be executed by this run. A case is defined by the values of the installation costs C and D and a list of zones or areas which are to be "forced in". If a zone is forced in, then there must be a terminal in that zone, regardless of cost and benefit. If an area is forced in, then there must be a terminal in that area, or, if the area is part of a zone, there must be a terminal in that zone. There is no limit to the number of cases per run.

The first record in this group is the set-up record, "*CASE". The set-up record is then followed by a cost record containing the values of C and D in the format 2I10. This record is followed by the forced in data. Each forced in record has the letters FOR in columns 2-4 and a list of up to 12 codes for areas or zones which are to be forced in. There can be more than one forced in record per case. If there are no forced-in areas or zones, there are no forced in records.

If there is more than one case per run, additional case records follow, starting with another set-up record, "CASE, followed by the cost record giving new values for C and D and the forced in records, if any. After the last case, the record "*END" indicates the end of the run. It should be noted that forcing is not cumulative, i.e. at the end of each case, all zones and areas are freed from being forced so that every zone and area to be forced in must be explicitly listed on the forced in records for each case.

Volume Data. For each origin/destination (O/D) pair of selected areas, PAREC requires as input a reference volume for that pair. As explained in section 2.1, this reference volume can be either the amount of mail going from the origin to the destination or else a profit for the origin and destination depending on the purpose of the model as expressed by the way the data are prepared by the analyst. These data consist of one record for each O/D pair. Each record contains the area code for the origin, the area code for the destination, and the reference volume, in the format (1X, A3, 1X, A3, F11.0). The order of the records is important. It is assumed that the data may include information for all O/D pairs in the area code dictionary and not just of the selected area O/D pairs. However, the O/D pairs must be sequentially ordered in accordance with their order in the area code dictionary. In other words, origins must occur in the same order as they occur in the area code dictionary, and the data for each origin must be grouped together so that for each origin, the destinations appear in the same order as in the area code dictionary.

Sample Input. Fig. 12 is a listing of all input (except volume data) for an example run. From the first record it can be seen that the dimension of the storage array X is +7500, there will be data for 32 areas on the volume input, no more than 100 iterations should occur, and all intra-regional volume is to be included. The network definition data describes a scenario with 32 areas, 30 of which are combined to form 11 zones. The two remaining areas are then stand-alone zones, giving a total of 13 zones. Six of the zones are combined to form 3 regions. The (seed) costs C and D are 1000 and 200, respectively, indicating that stations are twice as expensive as terminals. There are no areas or zones forced into the solution. The output associated with this example will be discussed in the next section.

	47500		32		1000		1.00													
21	70	71	73	100	200	206	210	212	463	464	600	604	606	750	752	760	761	800	802	
803	900	902	903	940	941	962	980	981	983	984	987									

*AREAS

21 BOSTON, MA
70 NEWARK, NJ
71 NEWARK, NJ
73 JERSEY CITY, NJ
100 NEW YORK, NY
200 WASHINGTON, DC
206 PRINCE GEORGES
210 BALTIMORE, MD
212 BALTIMORE, MD
463 GARY, IN
464 GARY, IN
600 N. SUBURBAN, IL
604 S. SUBURBAN, IL
606 CHICAGO, IL
750 RICHARDSON, TX
752 DALLAS, TX
760 FORT WORTH, TX
761 FORT WORTH, TX
800 DENVER, CO
802 DENVER, CO
803 BOULDER, CO
900 LOS ANGELES, CA
902 INGLEWOOD, CA
903 INGLEWOOD, CA
940 SAN FRANCISCO, CA
941 SAN FRANCISCO, CA
962 SAN FRANCISCO, CA
980 SEATTLE, WA
981 SEATTLE, WA
983 TACOMA, WA
984 TACOMA, WA
987 SEATTLE, WA

*ZONES

70 NEWARK, NJ	70	71	73
200 WASHINGTON, DC	200	206	
210 BALTIMORE, MD	210	212	
463 GARY, IN	463	464	
606 CHICAGO, IL	600	604	606
760 DALLAS-FT. WORTH	750	752	760 761
800 DENVER, CO	800	802	803
900 LOS ANGELES, CA	900	902	903
940 SAN FRANCISCO, CA	940	941	962
980 SEATTLE, WA	980	981	987
983 TACOMA, WA	983	984	

*REGIONS

NATIONS CAPITAL	200	210
CHICAGO AREA	463	606
SEATTLE-TACOMA	980	983

*CASE

100	200
-----	-----

*END

Figure 12: Sample input data

3.4 Output

Output consists of seven different types of information. Fig. 13 is an example of the output for the example described in the previous section with 32 areas, 13 zones, and 4 regions and volume data retrieved from a USPS-supplied tape. The reader should refer to the example output while reading the following discussion.

The first page of output consists of run header information. This output gives the title ("PARAMETRIC COST REVENUE MODEL") the version and release date, space for user supplied comments if any, and statistics on the numbers of areas, zones, regions, fraction of local volume, and total available volume.

The second type of output, the zone input description, may require several pages. Each zone is identified by an ordinal number, followed by the three character zone code or, in the case of stand-alone zones, the area code, and by the zone or area name. The percentage of the total volume associated with each zone is labelled "Volume Available". For aggregate zones all of the areas in the zone are listed by code and by name.

The region input description is quite similar to the zone description. Each region is identified by an ordinal number and a name. The percentage of total volume is presented for each region along with a list of zones included in the region. This output may also require several pages.

A fourth type of output is the case header description. This output displays a case number and the values of the cost parameters C and D. Although it is not shown in the example, this output also includes a list of zones, if any, which are forced into all solution networks. The forced zones are identified by an ordinal number, the area or zone code, the area or zone name, and the percentage of total volume. For forced aggregate zones, there is a list of the included areas. This output can require more than one page.

The fifth type of output, the deletion output, is the first output that is associated with the actual iterative solution procedure. This output occurs each time there is a step. Included in the output are the step number, number of iterations, h-value, and a list of all zones which are deleted at that step. The output from the first step starts a new page. All subsequent output of this type continues on the same page or subsequent pages if needed.

The sixth and seventh output types occur only after the iterative process is complete and the minimal network is obtained. The sixth type of output is the final zone output. Each zone in the final EMS-network is identified by its original ordinal number, the zone code, and the zone name. The percentage of total volume is given along with the name of the region, if any, to which the zone belongs. Zones in this list are ordered by decreasing volume.

The final output type gives the case history. This output gives the total number of steps, i.e. the number of networks found excluding the entire network associated with zero costs. Then a table is printed giving, for each step, the percentage to total volume of captured volume, the actual captured volume, the number of terminals, the number of stations, and the percent of fixed cost.

PARAMETRIC COST REVENUE MODEL

USPS-NBS VERSION 02, RELEASED 05/01/1977

SCENARIO :

LINK DATA:

REMARKS :

AREAS = 32, ZONES = 13, REGIONS = 3, L-FACTOR = 1.00
TOTAL AVAILABLE VOLUME = 2961921

Figure 13: Output from Sample Problem

- - - - -
C A S E 1
- - - - -

TERMINAL COST PARAMETER: C = 100

STATION COST PARAMETER: D = 200

FORCED ZONES AND AREAS :

Figure 13 (Continued)

 INPUT 0

- - - - -
 Z O N E S
 - - - - -

ZIP	NAME	VOLUME AVAILABLE	ZIP	NAME
---	---	-----	---	-----
1 21	BOSTON, MA	5.50		
2 70	NEWARK, NJ	6.49	AREAS	
			73	JERSEY CITY, NJ
			71	NEWARK, NJ
			73	NEWARK, NJ
3 100	NEW YORK, NY	17.96		
4 200	WASHINGTON, DC	11.85	AREAS	
			206	PRINCE GEORGES
			208	WASHINGTON, DC
5 210	BALTIMORE, MD	7.41	AREAS	
			212	BALTIMORE, MD
			210	BALTIMORE, MD
6 463	GARY, IN	2.33	AREAS	
			464	GARY, IN
			463	GARY, IN
7 606	CHICAGO, IL	15.12	AREAS	
			606	CHICAGO, IL
			604	S. SUBURBAN, IL
			600	N. SUBURBAN, IL
8 760	DALLAS-FT.WORTH	4.29	AREAS	
			761	FORT WORTH, TX
			760	FORT WORTH, TX
			752	DALLAS, TX
			750	RICHARDSON, TX
9 800	DENVER, CO	3.87	AREAS	
			803	BOULDER, CO
			802	DENVER, CO
			800	DENVER, CO
10 900	LOS ANGELES, CA	9.46	AREAS	
			903	INGLEWOOD, CA
			902	INGLEWOOD, CA
			900	LOS ANGELES, CA
11 940	SAN FRANCISCO, CA	7.89	AREAS	
			962	SAN FRANCISCO, CA
			941	SAN FRANCISCO, CA
			940	SAN FRANCISCO, CA
12 980	SEATTLE, WA	5.06	AREAS	
			987	SEATTLE, WA
			981	SEATTLE, WA
			980	SEATTLE, WA
13 983	TACOMA, WA	2.78	AREAS	
			984	TACOMA, WA
			983	TACOMA, WA

Figure 13 (Continued)

INPUT 0

- - - - -
R E G I O N S
- - - - -

	NAME	VOLUME AVAILABLE	ZIP	NAME
---	-----	-----	---	-----
1	NATIONS CAPITAL	19.26	ZONES	
			210	BALTIMORE, MD
			200	WASHINGTON, DC
2	CHICAGO AREA	17.45	ZONES	
			606	CHICAGO, IL
			463	GARY, IN
3	SEATTLE-TACOMA	7.83	ZONES	
			987	TACOMA, WA
			986	SEATTLE, WA

Figure 13 (Continued)

*****	- - - - -	
STEP 1	D E L E T E	ITERATIONS = 83
*****	- - - - -	H= 764
9 800 DENVER, CO		
*****	- - - - -	
STEP 2	D E L E T E	ITERATIONS = 4
*****	- - - - -	H= 773
8 760 DALLAS-FT.WORTH		
*****	- - - - -	
STEP 3	D E L E T E	ITERATIONS = 5
*****	- - - - -	H= 810
12 980 SEATTLE, WA		
13 983 TACOMA, WA		
*****	- - - - -	
STEP 4	D E L E T E	ITERATIONS = 12
*****	- - - - -	H= 947
1 21 BOSTON, MA		
2 70 NEWARK, NJ		
3 100 NEW YORK, NY		
4 200 WASHINGTON, DC		
5 210 BALTIMORE, MD		
6 463 GARY, IN		
7 606 CHICAGO, IL		
10 900 LOS ANGELES, CA		
11 940 SAN FRANCISCO, CA		

Figure 13 (Continued)

CASE 1

- - - - -
Z O N E S
- - - - -

	ZIP	NAME	VOLUME REALIZED		ZIP	NAME
---	---	-----	----		---	-----
3	100	NEW YORK ,NY	15.78			
7	606	CHICAGO, IL	12.60	REGION		
						CHICAGO AREA
4	200	WASHINGTON, DC	10.49	REGION		
						NATIONS CAPITAL
10	900	LOS ANGELES, CA	7.92			
5	210	BALTIMORE, MD	7.04	REGION		
						NATIONS CAPITAL
11	940	SAN FRANCISCO, CA	6.31			
2	70	NEWARK, NJ	6.88			
1	21	BOSTON, MA	5.05			
6	463	GARY, IN	2.24	REGION		
						CHICAGO AREA

Figure 13 (Continued)

CASE 1

- - - - -
H I S T O R Y
- - - - -

NUMBER OF NETWORKS FOUND = 4

STEPS	PERCENT VOLUME	CAPTURED VOLUME	NUMBER TERMINALS	NUMBER STATIONS	PERCENT FIXED COST
-----	-----	-----	-----	-----	-----
0	100.00	2961921	13	10	100.00
1	92.26	2732811	12	9	90.91
2	84.44	2501124	11	8	81.82
3	73.51	2177404	9	7	69.70
4	.00	0	0	0	.00

END OF RUN

Figure 13 (Continued)

3.5 Error Messages

When any type of error condition is detected in PAREC, a diagnostic message is printed and the program terminates. Many of the diagnostics are associated with improperly prepared input data. These diagnostics arise for errors such as set-up cards (those with "*" in column 1) which are either wrong or out of order, area codes which are not included in the area code dictionary or which duplicate an already selected area, duplicate zone codes, areas in more than one zone, zones in more than one region, etc. The particular messages associated with each type of input error are phrased so as to be self-explanatory. In each case, the last card read is printed after the message.

There are two diagnostic messages related to errors in reading the ODIS data. The first occurs when the reading gets out-of-step, i.e. if the origin/destination pairs on the volume tape do not include all possible area pairs in the area code dictionary or the pairs on the tape are not lexicographically ordered according to their order in the area code dictionary. The second message occurs if any volume is either negative or larger than a program parameter BIG which is 999999999 (set in subroutine GENNET to be described in Section 4.2). A negative value, of course, does not make sense. A volume larger than BIG could mean either that the volume is incorrect or that BIG is too small and should be increased.

Two diagnostic prints are associated with difficulties in solving a parametric flow problem. The first message, "INCOMPATIBLE BOUNDS" occurs when a lower bound in the dual network exceeds the corresponding upper bound. The occurrence of this error in PAREC usually means that the parameter BIG is too small. The user should increase BIG by modifying subroutine GENNET (see Section 4.2). A second message associated with the parametric flow model is "NO FEASIBLE FLOW COULD BE FOUND." If this message occurs while running PAREC, it indicates that some type of data or storage problem has occurred and gone undetected through all the other error checks of the program. Another message occurs in subroutine MODFLO and reports an "OUT-OF-KILTER STACK OVERFLOW." This indicates a need to increase the dimension of the areas X in the main program (see Section 4.1).

There are two groups of additional error messages which we will not discuss in detail in this document. The first group consists of a complete set of input checks, and the corresponding messages are self explanatory. The second group of error messages such as "BAD PARAMETER", "SOS PRICES" arise from internal consistency checks, and the sponsor was instructed to contact the authors should one of these messages occur during a model run.

4. Description of PAREC Program

This section is intended for a programmer/user of the PAREC computer code who is concerned about the details of the program which are of importance for maintenance and/or modification. The section includes a description of the data structures and subroutines of the code. Input, output, and error messages were discussed in Section 3.

4.1 Data Structure

The need for a somewhat special data structure in PAREC is apparent when one considers the size of the potential EMS-networks to be handled and the need for a user to specify varying numbers of the four entities: areas, aggregate zones, zones, and regions. There are many arrays associated with each of these entities. If each array were to be stored separately, changes in the numbers of any entities could necessitate recompiling the program with revised DIMENSION statements. Alternatively one could just set aside huge amounts of storage for all of the arrays, but for realistically sized networks, this simply would require entirely too much storage. What is needed is some form of dynamic storage allocation.

In PAREC a type of dynamic storage is accomplished by use of a single array, X . The various attributes for each entity are stored or stacked "end-to-end" by subroutine INZONE in the array X . Pointers are used to locate the first variable of each type. Then when INZONE calls EXECUT, the calling sequence specifies addresses or locations within the X array while the subroutine statement uses the names of individual arrays which need only be dimensioned by 1 in all subsequent subroutines.

An additional advantage of this type of storage is that there is no reason to require the user to specify how many of each entity there will be. In fact, INZONE determines these numbers by counting the amount of input data. The number of selected areas is stored as SARA, the number of aggregate zones is stored in TNZO, the total number of zones is stored in INZO, and the number of aggregated (i.e., multi-zone) regions is stored in KNRG. (Recall that the total number of entries in the area code dictionary, an input to MAIN, is JODI.) The total storage needed within the X array, as shown on the left-hand side in the following expression, must be less than the input parameter LIMIT.

$$2*JODI+10*SARA+6*TNZO+27*TNZO+4*INZO*(INZO-1)+21*KNRG+6<LIMIT.$$

LIMIT must be no less than the dimension of X .

The order in which the arrays are stored within X is given in Table 1. A brief description of each array will follow.

XZIPJ(J): three character code for area J; the Jth entry in the area code dictionary.

SELEJ(J): running selection number for area J. If SELEJ(J) = P, then area J is the Pth area to be selected; if 0, area J is not selected.

JZIPS(S): position of selected area S in the area code dictionary.

XNAMS(4*S-3)-XNAMS(4*S): the name input in INZONE for selected areas. There are 4 computer words per area with 4 characters per word.

IZONS(S): number of the zone containing the selected area S.

SNXTS(S): chain pointer to next selected area in the same zone as area S. The last link is negative number of aggregate zone.

TVOLS(S): total volume realized with the Sth network configuration.

NTRMS(S): total number of terminals with the Sth network configuration.

NSTNS(S): total number of stations with the Sth network configuration.

XZIPT(T): user supplied 3-character code for aggregate zone T.

SNAMT(4*T-3)-XNAMT(4*T): user supplied name for aggregate zone T. There are 3 computer words per zone with 4 characters per word.

START(T): pointer to the start of the chain (stored in SNXTS) of selected areas contained in aggregate zone T.

SAIZI(I): selected area which stands alone in zone I, if positive. If negative, its absolute value is the number of the aggregate zone to which it corresponds.

KREGI(I): number of the region associated with zone I.

INXTI(I): chain pointer to next zone in same region as zone I. If negative, its absolute value is number of region to which zone I belongs.

VOLMI(I): total volume realized at zone I. It is the average of all volume received at I and all volume originating at I.

IORDI(I): the position which zone I occupies in a list of zones ordered by volume realized.

XNAMK(4*K-3)-XNAMK(4*K): user-specified name for region K. There are 4 computer words per region with 4 characters per word.

IENTK(K): pointer to the start of the chain (stored in INXTI) of zones contained in region K.

COUNT(K): the number of zones contained in region K.

VOLMK(K): the total volume realized in region K. This is the average of volume originating at K and volume destined for K.

NORGA(A):* the origin of arc A of the dual network.

NDSTA(A): the destination of arc A of the dual network.

COSTA(A): the cost associated with arc A of the dual network.
BDUPA(A): the upper bound on flow through arc A of the dual network.
FLOWA(A): the flow through or across arc A of the dual network.
PARMA(A): the parameter for modifying the flow through arc A of the dual network.
AFORA(A): the next arc after arc A in the forward star of the node which is the origin of arc A .
ABACA(A): the next arc after arc A in the backward star of the node which is the destination of arc A .
PRICN(N):* price associated with node N of the dual network.
SPRCN(N): stored price for node N of the dual network.
ALABN(N): a label associated with node N of the dual network.
AFORN(C): the first arc in the forward star of node N.
ABACN(N): the first arc in the backward star of node N.
NSTKN(N): the node in the Nth position on the stack.
AOKSZ(Z): the arc in the Zth position on the list of out-of-kilter arcs.

*Variables indexed by A, N, or Z refer to arcs, nodes, and out-of-kilter arcs of the network of the dual problem as described in Appendix A.

TABLE 1. ARRAYS STORED WITHIN ARRAY X

	<u>ARRAY NAME</u>	<u>POINTER</u>	<u>SIZE</u>
1.	XZIPJ	LXZIPJ	JODI
2.	SELEJ	LSELEJ	JODI
3.	JZIPS	LJZIPS	SARA
4.	XNAMS	LXNAMS	4*SARA
5.	IZONS	LIZONS	SARA
6.	SNXTS	LSNXTS	SARA
7.	TVOLS	LTVOLS	SARA
8.	NTRMS	LNTRMS	SARA
9.	NSTNS	LNSTNS	SARA
10.	XZIPT	LXZIPT	TNZO
11.	XNAMT	LXNAMT	4*TNZO
12.	START	LSTART	TNZO
13.	SAIZI	LSAIZI	INZO
14.	KREGI	LKREGI	INZO
15.	INXTI	LINXTI	INZO
16.	VOLMI	LVOLMI	INZO
17.	IORDI	LIORDI	INZO
18.	ZNAMK	LXNAMK	4*KNRG
19.	IENTK	LIENTI	KNRG
20.	COUNK	LCOUNK	KNRG
21.	VOLMK	LCOUNK	KNRG
22.	NORGA	LNORGA	ARCS*
23.	NDSTA	LNDSTA	ARCS
24.	COSTA	LCOSTA	ARCS

TABLE 1. (continued)

	<u>ARRAY NAME</u>	<u>POINTER</u>	<u>SIZE</u>
25.	BDUPA	LBDUPA	ARCS
26.	FLOWA	LFLOWA	ARCS
27.	PARMA	LPARMA	ARCS
28.	AFORA	LAFORA	ARCS
29.	ABACA	LABACA	ARCS
30.	PRICN	LPRICN	NODS*
31.	SPRCN	LSPRCN	NODS
32.	ALABN	LALABN	NODS
33.	AFORN	LAFORN	NODS
34.	ABACN	LABACN	NODS
35.	NSTKN	LNSTKN	NODS
36.	AOKSZ	LAOKSZ	ZMAX**

*NODS and ARCS are parameters of the dualized problem with the following relationships to input data:

$$\text{NODS} = \text{INZO} + \text{KNRG} + 1$$

$$\text{ARCS} = (\text{INZO}) * (\text{INZO} - 1) / 2 + 2 * \text{INZO} + \text{KNRG}$$

**ZMAS is set equal to the size of all remaining storage in the X array. It must be strictly positive and although its value could well be large (depending upon the dimension of X), it need never exceed ARCS.

4.2 Subroutines

This section will provide a relatively brief description of the purpose for each subroutine. The diagram in Fig. 14 illustrates the relationships between calling routines. Subroutine entry prints are shown in Fig. 15.

PAREC consists of a main program, MAIN, a subroutine for reading in the area/zone/region data, INZONE, and a main driver subroutine, EXECUT, which controls all of the looping and logic discussed in Section 3.3 by calling, in the appropriate order, the other seven major subroutines. The two utility subroutines, REAREC and LOOKUP, are used to read the ODIS-data tape and to verify card input respectively. Each of these program components will be described in the paragraphs which follow.

MAIN. The primary purpose of MAIN is to read in the control card and the area code dictionary. This code also sets up the main storage array X and calls subroutine INZONE. The dimension of X must be no less than the value of the input parameter LIMIT.

Subroutine INZONE. This subroutine is responsible for reading in all of the area/zone/region data cards which describe the potential EMS-network for the current run. The input to this routine was described in section 3.3. Much of the code in INZONE is devoted to determining and setting up the amount of storage needed for all the arrays associated with the areas, zones, and regions of the potential EMS-network as well as the node and arc data of the dual network. The details of this storage procedure were discussed in Section 4.1.

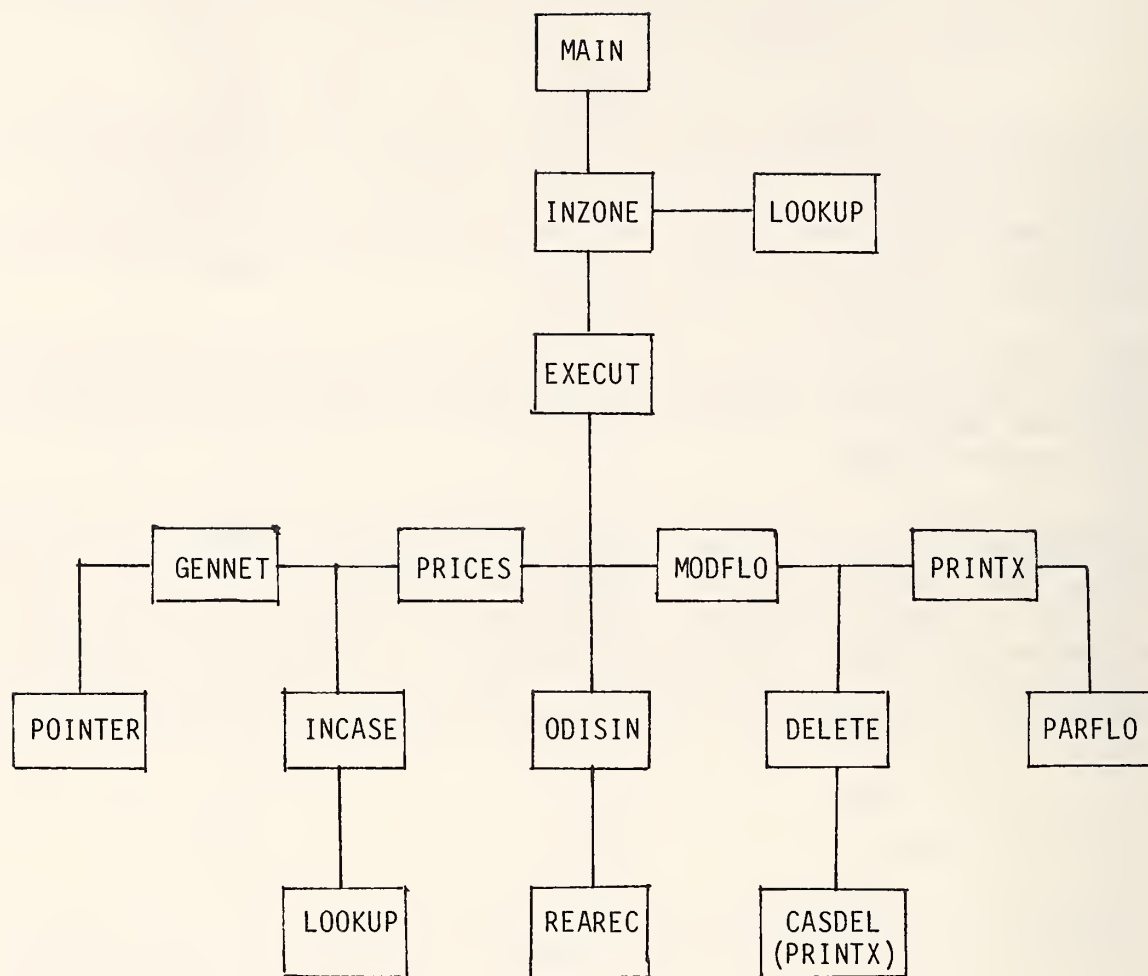


Figure 14: Calling relationship among subroutines

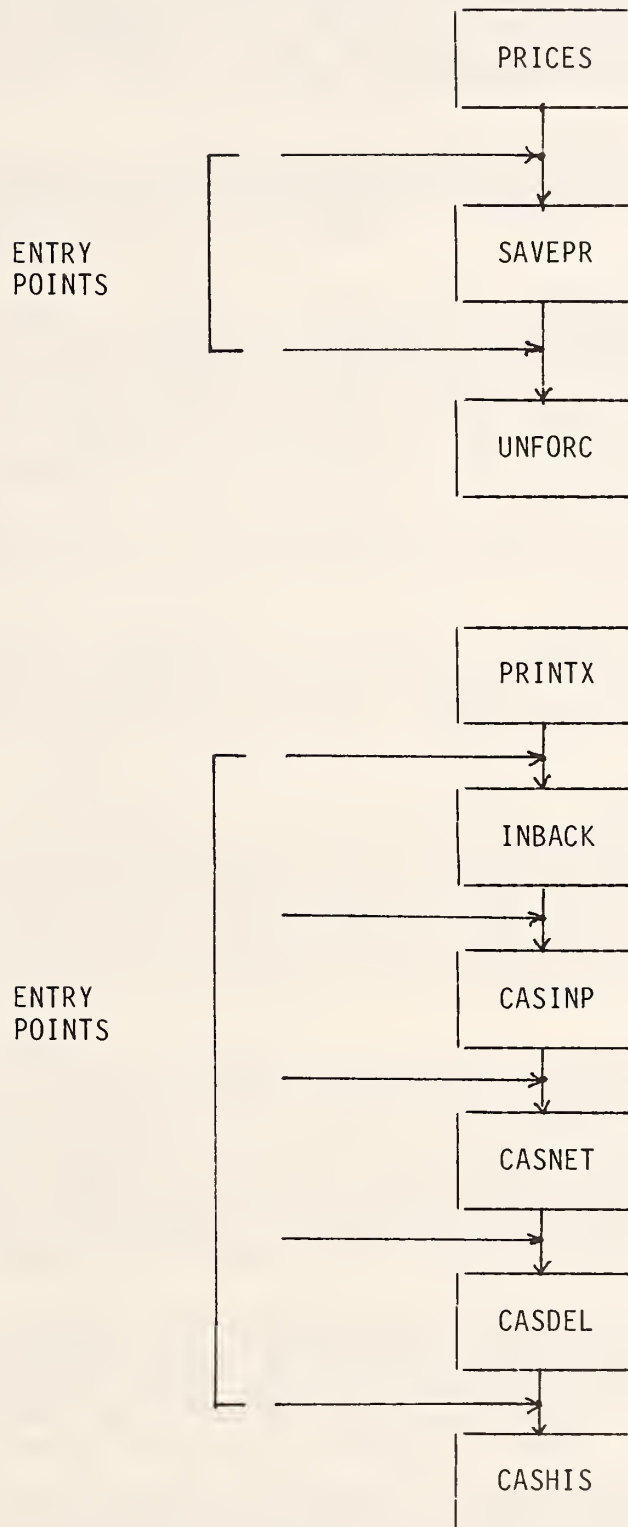


Figure 15: Subroutine entry points

Subroutine ODISIN. The primary purpose of ODISIN is to direct and control the logic needed to read from the ODIS-magnetic tape the reference potential volume data for the selected pairs of origin/destination areas. Total volumes are accumulated by zone and by region. ODISIN calls a short utility routine, REAREC, to perform the actual tape read.

Subroutine GENNET. The purpose of GENNET is to set up the dual network. Values are given to all or parts of all of the following arrays: NORGA, NDSTA, COSTA, and BDUPA. For further discussion of the dual network see Appendix A.

Subroutine INCASE. This subroutine reads the input which specifies the particular cases which are to be included in the run. Recall that the case data consist of the (seed) costs, C and D, and a list of areas and aggregate zones, if any, which are to be forced in as part of all EMS network configurations generated during the run. INCASE calls the utility subroutine LOOKUP to verify the list of forced areas and zones. Values are assigned to all or parts of all the following arrays: FLOWA, PARMA, COSTA, PRICN, and SPRCN. The arrays are associated with the dual network described in Appendix A.

Subroutine POINTER. The purpose of POINTER is to set up for each node in the dual network the lists of all arcs originating at that node and all arcs terminating at that node. The lists are referred to as forward and backward stars. Values are given to the following arrays: AFORN, ABACN, AFORA, and ABACA.

Subroutine PARLO is the code which actually solves a parametric network flow problem. By operating on the dual network, this subroutine provides a solution which can be translated directly into the set of zones and regions which comprise the optical EMS-network for particular values of the seed costs c and d. Further details on the problem solved by this algorithm are available in Appendix A.

Subroutine PRICES (Entry Points: SAVEPR and UNFORC). Subroutine PRICES with its two entry points, SAVEPR and UNFORC, is set up as if it were three separate subroutines, i.e. there are no overlapping executable statements. Thus three separate purposes are served. PRICES checks to see if the dual network is empty, i.e. if the current EMS network configuration consists of at most those zones and areas which are forced to be a part of the configuration. The logical variable EMPTY is set equal to TRUE when this condition obtains.

The entry SAVEPR is called with the single argument STEP whenever a step occurs, i.e. whenever the EMS-network configuration changes. This code saves the current dual network node prices, PRICN, in the array SPRICN, recalculates the total potential volume by zone, VOLMI, and computes and stores the total volume, total number of terminals, and total number of stations for the step.

After one complete case is finished, if another case is to be processed during the same computer run, the entry UNFORC is called. This code resets the values of COSTA so that no areas and zones are forced into the EMS-network. It also recalculates the full potential volume for each zone, VOLMI.

Subroutine DELETE. The purpose of subroutine DELETE is to detect the occurrence of changes (deletions) in the EMS-network. When deletions occur, TEP is increased by one and the appropriate node and arcs are eliminated from the dual network by revising the arrays associated with the forward and backward stars: AFORN, ABACN, AFORA, and ABACA. Subroutine CASDEL is called to print the deletec zone.

Subroutine MODFLO. This subroutine utilizes the parameters PARMA to modify the flows, FLOWA, in such a way that the dual network becomes infeasible and another parametric network flow solution is needed.

Subroutine PRINTX (Entry Points: INBACK, CASINP, CASNET, CASDEL, and CASHIS). Subroutine PRINTX and its five entries is set up as if there were six subroutines, each with a separate purpose. Together this group of codes is responsible for all of the non-diagnostic output from PAREC. A local variable LMAX controls the amount of output on each computer page. The current value of 46 does not include heading prints.

Since the output has already been described in Section 3.4, this discussion will simply indicate which entries are responsible for each section of output. The call to PRINTX which occurs immediately at the start of EXECUT prints all of the heading, comments, and data on the first page of output except for the total available volume. Entry INBACK prints the total available volume on the first output page as well as all of the data as supplied on input to describe the zones and regions of the entire potential EMS-network. Entry CASINP prints the output which describes the scenario for a particular case. This includes the seed costs as well as a list of any forced zones or areas. Entry CASDEL is called by subroutine DELETE. A call occurs once for each zone that is deleted, with the end result that at each step, CASDEL prints the list of zones which are to be deleted from the EMS-network. Entry CASNET prints the detailed data describing the smallest EMS-network generated for each case. The array IORDI is used to order the zones by volume realized. Entry CASHIS prints the history for the case, i.e. the data which are stored for each distinct EMS-network, that is, for each step.

Subroutine REAREC. This is a utility subroutine called by ODISIN to read from magnetic tape the potential volume data by origin/destination pair.

Subroutine LOOKUP. This is a utility subroutine called by both INZONE and INCASE to verify the three-digit codes which identify areas and zones. The parameters of the subroutine statement are (*, NWDS, DISTN, WORD, NPOS). DICTN is an array to be scanned in search of the value of WORD.

The length of the array is NWDS, and if found, NPOS is the position of WORD within the array. The symbol * is used to stand in the place of a statement number within the calling subroutine to which the return will transfer if WORD is not found in DICTN.

Subroutine EXECUT. Much of the logic and looping as explained in Section 3.2 is the responsibility of EXECUT. The flow chart of EXECUT, Fig. 16, illustrates the relationship among the various subroutines and entry points which EXECUT accesses in controlling the execution of PAREC. This flow chart, together with the discussion in Section 3.2, should provide a complete understanding of EXECUT as it ties together the subroutines of PAREC.

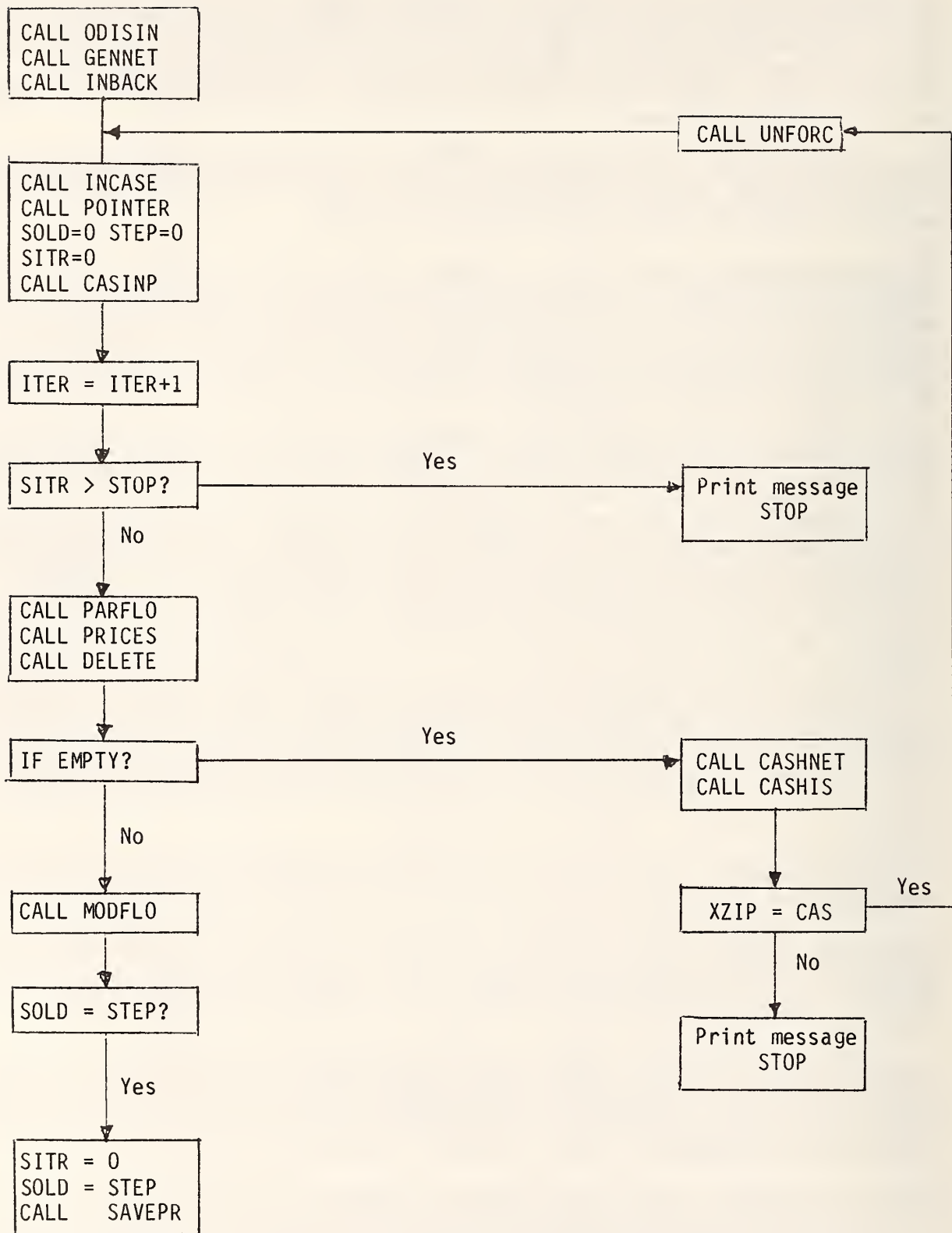


Figure 16: Flow chart for subroutine EXECUT

APPENDIX A: The Network for the Dual Problem

This appendix describes the dualization procedure employed by the PAREC program in order to generate a "dual" network, i.e. an abstract network to which a network flow optimization algorithm is then applied (Section 1.6).

Let I (INZO in the program) be the total number of zones and let K (KNRG in the program) be the total number of aggregated (i.e., multi-zone) regions. Also for all zones, i.e. for $i = 1, 2, \dots, I$, let $\sigma(i) = i$ if zone i is a stand-alone zone and let $\sigma(i) = k+1$ if zone i is a part of aggregated region k . Furthermore for $i = 1, 2, \dots, I$ let $\gamma(i) = 1$ if zone i is "forced in" (Section 3.1) and let $\gamma(i) = 0$ if zone i is not forced in.

With the above notation and the overhead term deleted (Sections 1.6, 1.10), the problem can be stated as

$$\text{Maximize} \quad \sum_{i=1}^{I-1} \sum_{j=i+1}^I p_{ij} \cdot y_{ij} - \sum_{\substack{i=1 \\ \sigma(i)=i}}^I (c+d) \cdot x_i - \sum_{\substack{i=1 \\ \sigma(i)>I}}^I c \cdot x_i - \sum_{k=1}^K d \cdot z_k$$

Subject to (note that by definition $\sigma(i) > i$ if and only if $\sigma(i) > I$, and that $k = \sigma(i) - I > 0$ is the region to which zone i belongs provided it is not a stand-alone zone)

$$(A.1) \quad y_{ij} - x_i \leq 0, \quad 1 \leq i < j \leq I,$$

$$(A.2) \quad y_{ij} - x_j \leq 0, \quad 1 \leq i < j \leq I,$$

$$(A.3) \quad -x_i \leq 0, \quad 1 \leq i \leq I, \quad \gamma(i) = 0,$$

$$(A.4) \quad -x_i \leq -1, \quad 1 \leq i \leq I, \quad \gamma(i) = 1,$$

$$(A.5) \quad x_i \leq 1, \quad 1 \leq i \leq I, \quad \sigma(i) = i,$$

$$(A.6) \quad x_i - z_k \leq 0, \quad 1 \leq i \leq I, \quad k = \sigma(i) - I > 0,$$

$$(A.7) \quad z_k \leq 1, \quad 1 \leq k \leq K,$$

where

p_{ij} = revenue associated with having both zones i and j included in the network

c = terminal (seed) cost

d = station (seed) cost.

In order to form the dual of the above linear program, we introduce a multiplier corresponding to each constraint:

$$u_{ij} \geq 0 \quad \text{for } 1 \leq i < j \leq I, \quad (y_{ij} - x_i \leq 0),$$

$$w_{ij} \geq 0 \quad \text{for } 1 \leq i < j \leq I, \quad (y_{ij} - x_j \leq 0),$$

$$r_i \geq 0 \quad \text{for } 1 \leq i \leq I, \quad \gamma(i) = 0, \quad (-x_i \leq 0),$$

$$r_i \geq 0 \quad \text{for } 1 \leq i \leq I, \quad \sigma(i) = 1, \quad (-x_i \leq -1),$$

$$s_i \geq 0 \quad \text{for } 1 \leq i \leq I, \quad \sigma(i) = i, \quad (x_i \leq 1),$$

$$s_i \geq 0 \quad \text{for } 1 \leq i \leq I, \quad k = \sigma(i) - I > 0, \quad (x_i - z_k \leq 0),$$

$$t_k \geq 0 \quad \text{for } 1 \leq k \leq K, \quad (z_k \leq 1).$$

The dual problem is then

$$\text{Minimize} \quad \sum_{\substack{i=1 \\ \sigma(i)=i}}^I s_i + \sum_{k=1}^K t_k - \sum_{\substack{i=1 \\ \gamma(i)=1}}^I r_i,$$

subject to

$$(A.8) \quad u_{ij} + w_{ij} = p_{ij}, \quad 1 \leq i < j \leq I,$$

$$(A.9) \quad s_i - r_i - \sum_{j>i} u_{ij} - \sum_{j<i} w_{ji} = - \begin{cases} c & \text{if } \sigma(i) > I \\ c + d & \text{if } \sigma(i) = i \end{cases}, \quad 1 \leq i \leq I,$$

$$(A.10) \quad - \sum_{\substack{i=1 \\ \sigma(i)-I=k>0}}^I s_i + t_k = -d, \quad 1 \leq k \leq K,$$

$$u_{ij} \geq 0, \quad w_{ij} \geq 0, \quad s_i \geq 0, \quad r_i \geq 0, \quad t_k \geq 0.$$

From (A.8) we have $w_{ij} = p_{ij} - u_{ij}$ and

$$\sum_{j < i} w_{ji} = \sum_{j < i} p_{ji} - \sum_{j < i} u_{ji} \geq 0 .$$

Substituting this into (A.9) and adding an additional superfluous constraint yields the following problem:

$$\text{Minimize} \quad \sum_{\substack{i=1 \\ \sigma(i)=i}}^I s_i + \sum_{k=1}^K t_k - \sum_{\substack{i=1 \\ \gamma(i)=1}}^I r_i ,$$

subject to

$$\sum_{j > i} u_{ij} - \sum_{j < i} u_{ji} - s_i + r_i = \begin{cases} -c & \text{if } \sigma(i) > I \\ -c + d & \text{if } \sigma(i) = i \end{cases} - \sum_{j < i} p_{ji}$$

for $i = 1, 2, \dots, I$,

$$\sum_{\substack{i=1 \\ \sigma(i)=k+I}}^I s_i - t_k = d , \quad k = 1, 2, \dots, K, \quad \text{and}$$

$$\sum_{\substack{i=1 \\ \sigma(i)=i}}^I s_i + \sum_{k=1}^K t_k - \sum_{i=1}^I r_i = - \sum (\text{right hand sides of all other constraints})$$

$$= -c \cdot I - d \cdot \bar{K} + \sum_{i=1}^{I-1} \sum_{j=i+1}^I p_{ij}$$

$$0 \leq u_{ij} \leq p_{ij} , \quad r_i \geq 0 , \quad s_i \geq 0 , \quad t_k \geq 0 ,$$

where \bar{K} is the total number of regions, both aggregated and single zone.

Note that the last equation is indeed superfluous. After it has been added, the sum of all constraint equations is zero since each variable appears exactly twice with opposite signs.

The network associated with this formulation is constructed as follows: the nodes are numbered $1, 2, \dots, I+K+1$, i.e. there is a node corresponding to each constraint. Each arc has as its origin the equation (node) in which the variable (arc) appears with a positive sign and as its destination the equation (node) in which it appears with a negative sign. The nodes also may be thought of as corresponding to the zones and aggregated regions of the potential EMS-configuration with the additional "dummy" node. The $I \cdot (I-1)/2$ arcs associated with the variables u_{ij} , $1 \leq i < j \leq I$, form a complete graph--complete in the undirected sense where there is an arc for each pair of different nodes--on the set of nodes associated with zones. However, each arc is directed from a lower to a higher zone number. Variables s_i , $1 \leq i \leq I$, correspond to arcs leading to the zone-node i from either the dummy node (for stand-alone zones) or the appropriate region-node. Variables r_i , $1 \leq i \leq I$, correspond to arcs leading from zone-node i to the dummy node. Finally, variables t_k , $1 \leq k \leq K$, belong to arcs from the dummy node to the nodes $I + k$ representing aggregated regions. Thus there is a total of

$$I \cdot (I-1)/2 + 2 \cdot I + K$$

arcs.

The right-hand sides in the dual problem represent the source strength at each node in the network as defined by the Kirchhoff balance equation

$$\text{Flow out} - \text{Flow in} = \text{source strength}.$$

Whereas the flow in the network is typically considered variable, the source strength is considered constant or, more precisely, flow independent, --as are the upper and lower bounds. A node is a proper source if its source strength is positive. It is a proper sink if its source strength is negative. Network flow algorithms generally start with a flow that is balanced in the Kirchhoff sense for given source strengths, but may violate the bounds of some arcs. For the above network an initial flow is given by

$$u_{ij} = p_{ij}, \quad 1 \leq i < j \leq I$$

$$r_i = 0, \quad 1 \leq i \leq I$$

$$s_i = - \begin{cases} c & \text{if } \sigma(i) > I \\ c + d & \text{if } \sigma(i) = i \end{cases} + A_i, \quad 1 \leq i \leq I,$$

$$t_k = -d - c \cdot I_k + \sum_{\substack{i=1 \\ \sigma(i)=k+I}}^I A_i, \quad 1 \leq k \leq K,$$

where

$$A_i = \sum_{j>i} p_{ij}, \quad 1 \leq i \leq I,$$

and I_k denotes the number of potential zones in aggregated region k .

For the parametric analysis we work with the reference profits \bar{p}_{ij} (1.7.3) and the demand level L (1.7.2):

$$p_{ij} = L \cdot \bar{p}_{ij} .$$

Dividing the objective function of the original maximization problem by L thus yields

$$\sum_{i=1}^{I-1} \sum_{j=i+1}^I \bar{p}_{ij} \cdot y_{ij} - \frac{1}{L} \left(\sum_{\sigma(i)=i}^I (c+d) \cdot x_i - \sum_{\sigma(i)>I}^I c \cdot x_i - \sum_{k=1}^K d \cdot z_k \right) .$$

In the network for the dual problem this change is reflected in an altered initial flow which arises from the one given above by the replacement:

$$p_{ij} \rightarrow \bar{p}_{ij} , \quad c \rightarrow \frac{c}{L} , \quad d \rightarrow \frac{d}{L} .$$

This flow is split into a determinate portion, which is independent of L , and an indeterminate portion, which contains the factor $1/L$. Consider the minimum-cost flow problem with the determinate flow as the initial flow. (The choice of initial flow determines the source strength of each node and, thereby implicitly the problem to be solved.) The node prices

$$x_i = 1 , \quad 1 \leq i \leq I ,$$

$$z_k = 1 , \quad 1 \leq k \leq K ,$$

$$q = 0 \quad \text{for the dummy node,}$$

satisfy the kilter conditions (FORD and FULKERSON [8]) for optimality together with the determinate flow. The above prices thus indicate that the universal configuration is optimal, which is to be expected since the demand level L is infinitely large in this case.

To start the parametric procedure we add a multiple of the indeterminate portion to the initial flow. This multiple exceeds by a given $\epsilon > 0$ the largest multiple that can be added without violating the kilter conditions. The modified initial flow defines a new problem which is then optimized. This process of modifying and optimizing is repeated, if necessary, until there is a change in the optimal prices x_i . This means that at least one node has dropped out. After recording the resulting new configuration, the process of modifying, solving and recording resumes by successively increasing $1/L$ until finally $x_i = 0$ for all zones i that are not forced in.

Table A.1 shows the major variables associated with each arc of the dual network. For every arc there is an origin node, a destination node, a cost, an initial determinate flow, lower and upper bounds on the flow, and a parameter correction quantity. These correspond, respectively, to the PAREC arrays NORGA, NDSTA, COSTA, FLOWS, BDUPA, and PARMA. The initial values of the arrays are shown in the table.

ARC	NORGA	NDSTA	COSTA	BDLOA	BDUPA	FLOWA	PARMA
For all zone pairs (i,j) with $i < j$	i	j	0	0	\bar{p}_{ij}	\bar{p}_{ij}	0
For all zones i	$\sigma(i)-I$ if $\sigma(i) > I$ dummy else	i	0 if $\sigma(i) > i$ 1 else	0	∞	$\bar{A}_i = \sum_{j>i} \bar{p}_{ij}$	-c if $\sigma(i) > i$ -(c+d) if $\sigma(i) = i$
For all zones i	i	dummy	-1 if $\gamma(i)=1$	0	∞	0	0
For all aggregated regions k	dummy	k	1	0	∞	$\sum_i \bar{A}_i$ $\sigma(i)=k+I$	-d-c•I _k

Table A.1 Initial Values of Network Arrays

APPENDIX B:

The Use of Customized Stationery
for Electronic Message Service Systems

by

Christoph Witzgall
Ralph E. Schofer

1. INTRODUCTION

This appendix presents a brief discussion of two modes of operation for an Electronic Message Service System (EMSS) in which customized preprinted stationery is used at the output terminal. Some of the preprinted material may be "stuffers," that is, enclosures which are identical for large groups of addresses. Examples include advertisements and return envelopes. But preprinted material may also serve as "message stock," that is, special stationery or "forms" on which EMS-messages are to be written. Such materials include letterhead stationery, checks, invoices, and billing stock.

Of particular interest appears to be the first of the aforementioned modes, which we call the "active loading" mode. It is made possible by the computerized environment in which messages may be collected in computers memory prior to being printed out. In the active loading mode, the preprinted message stocks and their stuffers are loaded into the output facility in more or less arbitrary sequence. The computer identifies the message stock and recalls from memory all those messages which required this particular kind of preprinted message stock. These messages are then printed out while and as long as the desired kind of stationary is in the hopper. This mode contrasts with a "passive loading" mode in which the sequence of the messages is essentially predetermined and the computer signals in each case for the appropriate message stock to be retrieved from inventory and loaded into the hopper.

2. BACKGROUND

This Appendix repeats the material of an informal working paper presented to the U.S. Postal Service in December 1976. It was prompted by work by Arthur D. Little, Inc. (ADL), on assessing the market for various EMSS options. The EMSS options examined by ADL would accomplish output by facsimile reproduction of mail items.

After the first draft of this working paper had been completed on November 24, 1976, our attention was drawn to a manuscript by Mr. J.W. McGinn [12], which presents similar ideas of mail generation using customer-specified preprinted stationery.

Also, a letter of June 25, 1976, from Mr. Harvey L. Pastan of ADL to Mr. McGinn on "the ... ability to handle billing forms on a non-facsimile basis" was made available to us. The point of departure of that letter was the availability in the future of an "electronic printing" technology, which would print forms and billing information "in synchronism." The images of forms and/or stuffers for mass mailings then need to be transmitted only once to "... be followed by the continuous transmission of billing information required for each mail piece such as name, address, zip code, and billing data..." This corresponds essentially to Option (D) examined by ADL and described later in this Appendix.

3. PURPOSE OF STUDY

We feel that digitally generated mass mailings are particularly suited for EMS and should form the backbone of its market. The capability for capturing this market would then constitute the main test for the viability of any EMSS.

We are concerned that "on-line" printing of such mail items may not be competitive from the point of view of efficiency with "off-line" printing and subsequent filling-in of digital billing data. The latter permits the customer to use his own forms and paper, and technology for it is already available.

Through this study we hope, therefore, to stimulate and/or intensify the discussion of techniques and operating procedures for paper handling and inventory management in an EMS environment which utilizes customized, "off-line" preprinted stationery.

4. FINDINGS BY ARTHUR D. LITTLE, INC.

Recently, Arthur D. Little, Inc., produced a category by category assessment of the portion of the market, presently served by First Class (FC), which would be served by an EMSS (see Arthur D. Little, Inc. [11]).

After ruling out that portion of the market which is captive to FC mail service because of the need to transport "artifacts," ADL identified "personalness" and "image" as requirements which are foremost in further limiting user acceptance of EMS.

The requirement of "personalness" pertains to the bulk of Household to Household mail, which for social and emotional reasons is clearly the domain of traditional FC mail, regardless of the quality of EMS.

The requirement of "image" pertains to the selection of stationery or paper stock--sizes, colors, letterheads, logos, embossing, graphics, watermarks, etc.--for highly repetitive and mostly commercial or business mailings. ADL found that demand for EMS hinges on the degree to which EMS would permit the projection of a desired image: EMS is perceived as reducing the capability for projecting "image," the extent of this reduction depending on the technological sophistication and reproduction quality opted for by the EMS implementation. Other substantial service improvements, e.g., reduced rates and preparation costs, or increased speed of delivery, must therefore be offered to offset this loss of image.

One such service improvement is to provide the option of digital input for computer-generated mass mailings--bills, invoices, statements of accounts, checks, etc. The advantage for the originator of such mass mailings lies in the fact that it obviates the physical creation of mail pieces at the point of origin, an activity now performed by the mailers or by service companies under contract to mailers.

In view of these facts, Arthur D. Little compared four EMS options with respect to their impact on the market.

- Option (A): a straight black-and-white high-quality facsimile service.
- Option (B): same as (A), but including a digital input option.
- Option (C): a color facsimile service of restricted quality, with digital input option.
- Option (D): a color printing service ("on-line") of photographic quality, with digital input option.

The estimated market share of Option (D) was of considerable size, about six times as great as that estimated for Option (A).

It is important to note that, in all the above estimates, ADL assumes a capability of handling "stuffers" as well as "multiple pages." It is our understanding that, in the currently prevailing USPS scenario, every EMS-item, including stuffers, is printed from scratch in each instance, but that the scanning and transmission of repetitive information is avoided.

5. DETACHMENTS

There is a practical reason in addition to "image" which might cause originators of mass mailings to request their own stationery. Parts of bills are frequently designed to be detached and returned along with the payment. If these detachments are to be processed automatically upon receipt, then they may well be required to meet precise specifications as to size and quality of paper. In any case, the returns are likely to be expected to be of uniform quality and appearance. A mass mailer might therefore object to a mode of operation which involves, say, digital mail generation on customized stationery for local addresses, and the use of facsimile mail generation for out-of-town ones.

6. ADDITIONAL OPTIONS

In this working paper, we will consider two additional options. The first one (E) assumes complete flexibility in selecting message stock and stuffers for digitally created mass mailings.

- Option (E): same as (B) but ("off-line") preprinted customer-selected stationery and enclosures.

We feel that this option warrants discussion because

- it permits total image projection;
- it precludes the need for "on-line" printing of repetitive material, thereby increasing output efficiency;
- it does not infringe upon legitimate interests of the Printing Industry;
- it lowers the requirements for "on-line" printing capabilities and capacities, with substantial concomitant savings in capital investment and maintenance;
- it is based on technology that is already available to a large extent.

On the debit side are

- increased requirements for paper-handling capabilities and storage
- slightly increased manual and managerial operations
- need for bulk shipments and distribution of paper materials.

The main benefit of this option, as well as those discussed by ADL, is that sorting can be performed internally in the computer. The mail pieces are then generated in sorted sequence, say bundled by mail carrier.

The same holds for the second additional option. In this case, the use of preprinted material is restricted to "pure stuffers," i.e., enclosures without individual messages, such as advertisements and return envelopes. All EMS-generated messages are printed out on standard EMS message stock. Logos would have to be printed on an addressee-invariant cover letter, which would be one of the pure stuffers. Such "Logo stuffers" may represent a satisfactory solution to the image projection problem. Of course, this would have to be confirmed by market research.

Option (F): same as (B) but utilizes preprinted pure stuffers.

This allows for more flexibility in operating procedure and may require a lower level of paper-handling capability than Option (E). In particular, mass mailings need not be printed all together in one batch: the sorting process may therefore extend over several different mass mailings.

7. ACTIVE VERSUS PASSIVE LOADING

Consider the mailing of credit card charges by a major oil company. Typically, each item in this mailing consists of five subitems:

- (i) outer envelope
- (ii) monthly statement
- (iii) purchase receipts
- (iv) return envelope
- (v) advertising enclosures.

The monthly statement will contain digital information as to amounts due and customer identification. Subitems (iv) and (v) are pure stuffers. Facsimile copies of the purchase receipts (iii)--black on white is probably quite adequate--are generated either separately or are reproduced in reduced sizes on summary pages--each page containing a number of invoice facsimiles or excerpts. Finally, the outer envelopes will have to be addressed.

If a mailing of monthly statements is to be accomplished via EMS under Options (E) or (F), then special message stocks would have to be provided on the spot for subitems (i) and (ii). Furthermore, preprinted pure stuffers must be on hand. In other words, special stocks must be in hoppers prior to the generation of the mail pieces in question. Operating procedures will have to center around the availability of special stock at the right time. There are two basic philosophies for achieving this:

"Active Loading": ("message stock seeks message"). The printing of accumulated messages requiring identical message stock and stuffers is triggered by the presence on the hopper of the required special stock.

"Passive Loading": ("message seeks message stock"). All conceivably needed special stocks are present on the hopper. Printing is triggered by events outside the printing environment.

In the active loading mode, all EMS-items pertaining to the same mass mailing are collected in a computer-internal memory file--one such file for each mass mailing. Hoppers are periodically loaded with the special stocks required for a particular mass mailing, e.g., the hypothetical credit-card statement. All individual statements which arrived since the last printing have been collected on a file. They are now printed in one continuous procedure until either the memory file or the hopper is empty. This mode of operation is essentially the one proposed and examined by McGinn.

In the passive loading mode, each EMS-message is dealt with as it arrives. The mail pieces to be generated are not grouped by special stock required. The passive loading mode is therefore more flexible, and permits "on-the-spot sorting to carrier" of larger populations of mail pieces than those consisting of a single mass mailing.

On the other hand, all special stock must be in the hopper all the time. This might be more readily accomplished, if all special stock consisted of pure stuffers (Option (F)). One can also cut down on the number of special stocks in the hopper by restricting their use to particular time periods. The mailing of the credit card statements may, for instance, be restricted to Mondays. Then the corresponding special stock would have to be in the hopper on Mondays only.

We summarize: Option (E) is oriented towards active loading only, whereas the advantage of Option (F) would be to permit passive loading. All four EMS-options discussed by ADL permit passive loading, or rather render the distinction moot.

8. SAFEGUARDS

There is clearly a danger of printing an EMS-message on the wrong stock, of misplacing stuffers, etc. There is no way of making any system entirely fool-proof. Complete automation appeared to be an answer, but more recent management philosophy tends towards utilizing the diligence of intelligent and responsible personnel, and employing highly sophisticated technology in a subordinate role intended to assist people in their job.

For the sake of discussion, we describe some concepts of safeguards. We assume manual loading of the hoppers. The stock, however, should be prepackaged and supplied with identification tags which can be read or scanned by computer at each particular hopper being loaded. The computer then will receive messages which identify the material that is presently being loaded as well as the hopper location.

Further safeguards may be desired to protect against errors in tag-preparation and prepackaging. For instance, after reading the identification tag of some particular stock, visual images and/or verbal descriptions of the stock the computer expects to be associated with this tag could be displayed on a TV-screen. The operator then can check whether the stock he put in the hopper does indeed fit the description. Locking devices might ensure that no further stock can be added or removed.

So far, we have thought in terms of several hoppers which are in use simultaneously. Of course, one may just have a single hopper, and have all subitems "hang together," i.e., input one large sheet of stock which then would be printed, cut, folded, stuffed and sealed. This single input sheet approach would to some extent reduce the possibility of mix-ups.

9. INVENTORY

The problems of keeping an inventory of special stocks are considerable, but not insurmountable. Two different modes of storage present themselves.

"Discriminate Storage": There is a separate bin for each special stock.

"Indiscriminate Storage": Packages of special stock are not stored according to their kind.

When operating in the discriminate storage mode, the operator typically knows what special stock is required to be loaded, looks up the bin in which this stock is stored, goes and gets it. When operating in the indiscriminate storage mode, the operator picks up some designated package of special stock without necessarily knowing what it contains. The computer identifies the stock from its identification tag, and proceeds to print those messages requiring this stock that have accumulated in its memory.

Discriminate storage management is necessary for passive loading. It also is assumed by McGinn for his local mail generation scheme. The problem with discriminate storage is that it requires a potentially prohibitive number of bins. Indiscriminate storage holds therefore a definite advantage if there are many special stocks to be stored. It also avoids the operational tasks of reassigning and consolidating bins, and is probably more amenable to the implementation of computer-controlled loading strategies.

Non-local mass mailings might be handled as follows: packages of special stock are sent by Parcel Post to prospective terminals. There they are entered into a queue. Upon their being loaded, the digital information, having arrived earlier and been stored in the computer memory, is now printed. For frequent mass mailings of large size, it may pay always to keep a supply of special stock at hand.

10. SORTING VERSUS SEQUENCING

As McGinn shows, the pay-off for generating mail locally from digital input lies in the operational advantages gained by letting the computer "sort by carrier" in its memory the mail items to be generated. They will then be generated in this sequence. This reduces the cost of operations in that the generated mail can now be handed directly to the carriers without having to go through an expensive sorting process.

For local mail there are no other important operational advantages to be gained. Since most mass mailings are actually local, e.g., most utility bills, sorted printing has to be considered the key to any pay-off on a national level.

For this reason, some more detailed examination of sorting in the context of mail generation is in order. First a clarification of terms: "sorting" by, say, carrier means physically depositing mail pieces into carrier-assigned specific bins or pouches. When we talk about "sorting" in the computer memory, as we did at the beginning of this section, we are obviously referring to something quite different: the mail pieces are printed out in a particular sequence, namely, ordered by carrier. Thus all mail pieces for carrier 1 are

printed first, to be followed by those for carrier 2, and so on. While this does not constitute actual "sorting" as defined above, it certainly facilitates it drastically. We will use the term "sequencing" for this process of having the computer arrange the generation of the mail pieces in a particularly advantageous sequence. It is our understanding that McGinn's concept of mail generation involves "sequencing" rather than "sorting."

For active loading patterns, where items of the same mass mailing are generated consecutively, sequencing takes place within each printing cycle of a particular mail item. For instance, all those credit-card statements of our hypothetical brand that are printed in one batch are sequenced within this batch. It is clear, on the other hand, that sequencing becomes rapidly less advantageous as the above batch size decreases. It follows, that in an operational environment which includes many mass mailings of moderate size or nonlocal origin, the advantages gained from sequencing alone are not sufficient to justify the operation.

Another consideration along this line is that local mass mailers may conceivably be asked to generate their mail in a particular sequence and to deliver it to the Post Office in this ordered state. The same operational advantages would then accrue as if the Postal Service itself had sequenced the mail.

For these reasons, it may be necessary to provide a capability for actual sorting. The computer has the address information necessary to guide each generated mail piece to its proper bin. Items from different mass mailings may thus find their way into the proper bins, and the advantages gained from sorting are not restricted to very large mass mailings.

Some savings may be achieved by combining the process of printing with the marking of each mail piece by magnetic ink so that these mail pieces can then be entered directly into standard sorting machinery, circumventing the need for address recognition by Postal Service personnel.

11. CONCLUSIONS

In this paper, we have examined and to some extent advocated the concept of mail generation proposed by McGinn for local mass mailings into prospective EMS-systems, so as to accommodate the key market of digitally created mass mailings. The technology for this concept is essentially available, and it appears to offer a potential for operational savings by the Postal Service in spite of the admittedly thorny problems of operation, inventory, and distribution caused by the use of customized stationery and preprinted stuffers.

The reasons for using the latter are, on the one hand, to avoid the costs inherent in printing repetitive information "on-line" rather than "off-line" and, on the other hand, to be fully able to meet customer requirements for maintenance and projection of "image."

The examination carried out in this paper has been on the conceptual level. The paper is not a final report on a detailed investigation, but rather a prospectus of ideas and concerns meriting further discussion.

REFERENCES

- [1] Witzgall, C.: Mathematical Methods of Site Selection for Electronic Message Systems (EMS), NBSIR 75-737, National Bureau of Standards, June 1975.
- [2] ODIS Data Base: Origin-Destination Information System. Data Collection Procedures, Methods Handbook M-61, United States Postal Services, December 1972.
- [3] Ruckman, R.: Electronic Mail Handling Network Conceptualization, United States Postal Service, June 1972.
- [4] Sherrard, E.: Two Networks for Electronic Transmission of Inter-office Mail, United States Postal Service, September 1973.
- [5] Rhys, J. M. W.: A Selection Problem of Shared Fixed Costs and Network Flows, Management Science, 17, 200-207, 1970.
- [6] Gardner, L.: The Maximum Benefit Model for EMSS Station Selection. AMSD - Working Paper, United States Postal Service Research Laboratories, Rockville, MD, January 1976.
- [7] PHILCO: Conversion Subsystems for the Electronic Mail Handling Program. Task 4 Report: Baseline EMH Conversion Subsystem Data Base. May 1973.
- [8] Ford, L. R., and Fulkerson, D. R.: Flows in Networks, Princeton University Press, Princeton, NJ, 1972.
- [9] GENERAL DYNAMICS: Study of Electronic Handling of Mail. Collection and Distribution. Report R69-046-2, March 1970.
- [10] Gass, S. I.: Linear Programming, 3rd edition, McGraw-Hill, New York, NY, 1969.
- [11] Arthur D. Little, Inc.: The First Class Mail Stream as a Candidate for EMS, Final Report to the United States Postal Service, February 1977.
- [12] McGinn, J. W.: Concept for a local digital mail system; informal document, U.S. Postal Service, 1976.

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET <i>(See instructions)</i>	1. PUBLICATION OR REPORT NO. NBS/TN-1225	2. Performing Organ. Report No.	3. Publication Date November 1986
4. TITLE AND SUBTITLE The Parametric Cost-Revenue Model ("PAREC") For Electronic Message Service Systems			
5. AUTHOR(S) Christoph J. Witzgall, Patsy L. B. Saunders, Ralph E. Schofer			
6. PERFORMING ORGANIZATION <i>(If joint or other than NBS, see instructions)</i> NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE GAITHERSBURG, MD 20899		7. Contract/Grant No. N.A.	8. Type of Report & Period Covered Revised Sept. 1985
9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS <i>(Street, City, State, ZIP)</i> Same as #6 above.			
10. SUPPLEMENTARY NOTES <input type="checkbox"/> Document describes a computer program; SF-185, FIPS Software Summary, is attached.			
11. ABSTRACT <i>(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)</i> This report describes a methodology for analyzing costs and benefits of a national electronic message service system and of similar systems involving alternative configurations of unspecified size. At the core of this methodology is a mathematical model, called PAREC, which provides information on sizes and configurations which optimize profit or return on investment. It is based on a technique for optimal selection of subconfigurations due to J.M.W. Rhys. A parametric minimum cost network flow algorithm was developed for the purpose of solving the resulting optimization problems.			
12. KEY WORDS <i>(Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)</i> combinatorial optimization; communication systems; cost-benefit; electronic mail; linear-programming; location theory; network optimization; parametric optimization; return on investment; RHYS technique; system configuration			
13. AVAILABILITY <input checked="" type="checkbox"/> Unlimited <input type="checkbox"/> For Official Distribution. Do Not Release to NTIS <input checked="" type="checkbox"/> Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. <input type="checkbox"/> Order From National Technical Information Service (NTIS), Springfield, VA. 22161		14. NO. OF PRINTED PAGES 90 15. Price	

NBS *Technical Publications*

Periodical

Journal of Research—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. Issued six times a year.

Nonperiodicals

Monographs—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The Journal of Physical and Chemical Reference Data (JPCRD) is published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements are available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—Disseminates technical information developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

Consumer Information Series—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the above NBS publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.

Order the following NBS publications—FIPS and NBSIR's—from the National Technical Information Service, Springfield, VA 22161.

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NBS Interagency Reports (NBSIR)—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Service, Springfield, VA 22161, in paper copy or microfiche form.

U.S. Department of Commerce
National Bureau of Standards
Gaithersburg, MD 20899

Official Business
Penalty for Private Use \$300