

City Research Online

City St George's, University of London

Citation: Solomon, S. I. (1974). Planning of natural regime hydrological networks. (Unpublished Doctoral thesis, The City University)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/38690/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

PLANNING OF NATURAL REGIME HYDROLOGICAL NETWORKS

by

S.I. SOLOMON

(Civil-Hydraulic Engineer, University of Bucharest)

A Thesis

Presented for the Degree of Doctor of Philosophy

Department of Civil Engineering

The City University, LONDON, England.

1 9 7 4

THE CITY UNIVERSITY LIBRARY,
BT. JOHN STREET, LONDON, E.C.1.

TABLE OF CONTENTS

	Page
ABSTRACT	i
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	iv
LIST OF TABLES	vi
SUMMARY	S-1
CHAPTER 1 - INTRODUCTION	1-1
1.1 Hydrological Networks: Definition and Classification	1-3
1.1.1 Definition	1-3
1.1.2 Classification	1-6
1.2 Purpose and Scope	1-14
1.3 Present Status and Trends in Planning of Natural Regime Networks	1-16
1.3.1 Review of International Literature	1-16
1.3.2 Review of International Practice	1-30
CHAPTER 2 - PLANNING FRAMEWORK	2-1
2.1 Criteria, Objectives and Decision Variables	2-1
2.1.1 Criteria	2-2
2.1.2 Objectives	2-4
2.1.3 Decision Variables	2-7
2.2. Optimization	2-8
2.2.1 Optimization Technique for the Planning of the Natural Regime Section	2-8
2.2.2 Transitional Section	2-11
2.3 Economic and Other Considerations	2-13
2.3.1 Priority Development	2-13
2.3.2 Conservation	2-15

CHAPTER 3	DATA AVAILABILITY AND DATA HANDLING	3-1
3.1	Data Availability	3-3
3.2	Data Handling	3-5
CHAPTER 4	ERRORS IN HYDROLOGICAL INFORMATION AS RELATED TO NETWORK PLANNING	4-1
4.1	Errors of Hydrological Information Related to Planning Decision Variables	4-4
4.2	Estimation of Error Variance Parameters	4-10
4.2.1	Error Due to Information Transfer in Space	4-11
4.2.2	Error Due to a Limited Sampling Period	4-40
4.2.3	Error of Measurement	4-48
CHAPTER 5	COSTS	5-1
5.1	Canadian Costs	5-2
5.1.1	Installation Costs	5-3
5.1.2	Operation and Maintenance	5-6
5.1.3	Data Processing	5-7
5.1.4	Data and Information Interpolation	5-8
5.1.5	Other Costs	5-11
5.2	Costs in Other Countries	5-11
CHAPTER 6	OPTIMIZATION	6-1
6.1	Planning of the Natural Regime Network in Areas not Designated for Development	6-3
6.1.1	Error Reduction	6-3
6.1.2	Additional Costs	6-4
6.2	Planning of the Natural Regime Network in Areas Designated for Development within the Planning Horizon	6-9
6.2.1	General Considerations	6-10
6.2.2.	Dynamic Programming Applied to Hydrological Network Design	6-11

CHAPTER 7 STATION DISTRIBUTION 7-1

7.1 Geophysical Considerations 7-1

7.1.1 Model-homogenous Regions 7-2
7.1.2 Physiographical-hydrological Regionalization 7-9

7.2 Economic Considerations 7-11

7.2.1 Costs of Installation and Operation 7-13
7.2.2 Economic Development Considerations 7-14

7.3 Environmental and Aesthetic Considerations 7-14

CHAPTER 8 PRACTICAL APPLICATIONS 8-1

8.1 Application to Natural Regime Section Assuming no Economic Development 8-2

8.1.1 Error Variance Parameters 8-2
8.1.2 Costs 8-5
8.1.3 Results 8-7
8.1.4 Sensitivity Analysis 8-10

8.2 Application to Natural Regime Section Assuming Significant Economic Development (Transitional Section) 8-17

8.2.1 Basic Assumptions 8-17
8.2.2 Steps Involved in the Hypothetical Application 8-13
8.2.3 Application to Canadian Network 8-26
8.2.4 Optimum Path for Reaching a Certain Level of Accuracy 8-34
8.2.5 Sensitivity Analysis 8-37
8.2.6 Sensitivity Analysis Assuming Cyclic Economic Development 8-40
8.2.7 Recommendations with Respect to the Application of the Dynamic Programming Technique to Network Planning 8-43

CHAPTER 9 CONCLUSIONS AND RECOMMENDATIONS 9-1

9.1 Conclusions and Recommendations for Network Planning 9-1

9.2 Recommendations for further studies 9-7

REFERENCES R-1
APPENDIX A A-1
APPENDIX B B-1
SUPPLEMENTARY REFERENCES SR-1

ABSTRACT

This thesis is intended as a theoretical discussion and a practical guide for the planning of the hydrological networks in areas where the hydrological regime has not been disturbed by man. Although the scope of this thesis is limited to the natural regime section of the hydrological network, the writer suggests that the methodology developed may have applications to other sections of the hydrological network and to related networks, particularly the meteorological one.

The main elements which are at the basis of the planning technology developed are data banks, and modelling and optimization procedures, the combined use of which is made possible by the computer. The method of planning suggested is based on the idea that the development of the hydrological network should proceed through a reasonable allocation of effort and funds for data acquisition and data interpretation. Methods of information interpolation based on the analysis of the available data in conjunction with data from other related networks and information on the terrain characteristics are emphasized as a complement or an alternative to direct measurement.

While taking his lesson from past and present, the writer recognizes that developing technology will soon change some,if not most, of the current hydrometric

practices. Hence, the writer also considers the planning process as an effort to reconcile the future to the present.

ACKNOWLEDGEMENTS

The writer wishes to express his sincere appreciation and thanks to Professor P.O. Wolf for his overall supervision, guidance and kind encouragement during the course of this research. He is also heavily indebted to [redacted] and others with whom he had many fruitful discussions on the subject of this thesis.

Discussions with many officials of the Canadian Department of the Environment, Inland Waters Directorate, have also been helpful in completing the research. The writer wishes also to express his thanks to the Department of the Environment for providing the required data and entrusting him with assignments which made possible the application of the research results to a series of significant practical problems.

* * * * * * * * * * * * * * *

LIST OF FIGURES

		Page
Fig. 1.1	Suggested Classification of Hydrologic Networks	1-8
Fig. 3.1	Basic Layout of the Physiographical Land use - Land cover and Hydro-meteorological Data Bank	3-6
Fig. 3.2	Square Grid System Superimposed on Basin Topography for Automatic Determination of Drainage Basin Area Along a River	3-9
Fig. 3.3.	Square Grid Estimation of Distribution of Mean Annual Temperature in Southern Ontario.	3-12
Fig. 4.1	Relationship Between Standard Error of Estimate of Average Runoff and Hydro-metric Network Density for Various Data Transfer Methods	4-27
Fig. 4.2	Relationship Between Relative Standard Error of Long-term Average Flow Coefficient of Variation and Coefficient of Skew of Annual Flow Versus Length of Sampling Period	4-44
Fig. 5.1	Relationship Between Gross Income Per Capita and Density of Hydrological Stations	5-13
Fig. 6.1	Stages and States for Optimization by Dynamic Programming of Hydrologi-cal Network	6-13
Fig. 6.2	Stages and States for Application of Dynamic Programming to Network Planning	6-14
Fig. 8.1	States and Stages used in the Dynamic Programming Application to Network Planning	8-22
Fig. 8.2.	Computational Steps in the Dynamic Programming Application to Network Planning	8-25
Fig. 8.3	Optimum Conditions for Transition from a State of a Given Stage to another State of the Next Stage	8-29

Fig. 8.4	Dynamic Programming Application to Network Planning - Cost of Expanding and Operating the Network in Millions of Dollars	8-30
Fig. 8.5	Dynamic Programming Application to Network Planning - Losses Corresponding to various states (Error Levels) in Millions of Dollars	8-32
Fig. 8.6	Dynamic Programming Application to Network Planning - Losses and Operation costs corresponding to various Transition Paths in Millions of Dollars	8-33
Figs. A-1 to A24	Southern Ontario - 5 x 5 km Grid - Terrain Characteristics	A-6 to A12
Figs. A-25 to A27	Southern Ontario - 5 x 5 km Grid - Climatic Characteristics	A13
Fig. B-1	Square Grid and Isoline Distribution of Runoff - Southern Ontario	B5

LIST OF TABLES

		Page
Table 2-1	Accuracy, Cost and Time Considerations for Establishing the Objectives of the Natural - regime network	2-6
Table 4-1	Values of E_s^* for the Square Grid - Multiple Regression Technique for Hydrological Information Interpolation	4-26
Table 4-2	Comparison Between Error of Estimate of Mean Annual Runoff Obtained from an Isohyetal Map (calibration sample) and Square Grid Technique (Validation Sample)	4-32
Table 4-3	Expected Values of E_s^* obtained from Parametric Model - Square Grid Technique	4-39
Table 4-4	Expected Values of $E_{t,50}^*$ obtained from Data at Six Gauging Stations in Canada	4-47
Table 8-1	Value of the Error Variance Parameter E_s^*	8-4
Table 8-2	Optimum Values of Planning Variables for Various Objectives (Interest Rate 7%).	
	a. Parametric Model - Square Grid Technique	8-11
	b. Correlation - Square Grid Technique	8-12
Table 8-3	Variation of Optimum Number of Supplementary Stations Corresponding to Largest Error Reduction per dollar Spent within a Five-Year Period (Western & Northern Canada Conditions)	8-14
Table 8-4	Summary Data on Effect of Station Discontinuation	8-16
Table 8-5	Results of Optimization by Dynamic Programming of Simulated Canadian Natural Regime Network for the 1971-1985 Period	8-35

LIST OF TABLES

		Page
Table 8-6	Sensitivity Analysis: Variation of Optimum Number of Stations with Variation of Basic Economic Assumptions	8-41
Table 8-7	Sensitivity Analysis: Effect of Economic Cycles-Costs of Operating and Expanding Network	8-44
Table B-1	Correlation Between Runoff and Terrain Characteristics in Southern Ontario	B-2
Table B-2	Comparison Between Measured (actual) and Estimated (Predicted) Runoff for the Calibration Sample (Runoff Estimates from Correlation-Square Grid Technique)	B-3
Table B-3	Comparison between Measured (Actual) and Estimated (Predicted) Runoff for the Validation Sample (Runoff Estimated from Correlation - Square Grid Technique)	B-4

SUMMARY

This thesis contains a theoretical discussion and a practical approach to the planning of natural-regime hydrological networks. It is based on international theoretical developments and practical experience, particularly on work in this area carried out during the last few years (1967-1973) in Canada and on research done by the writer for the purpose of this thesis.

The approach to planning here suggested is based on the concept that a hydrological network does not consist only of measurement stations and the time series of hydrological data obtained at these stations, but also of a system of processing these data to obtain hydrological information as well as synthesised time series of hydrological data and/or hydrological information at any point of the area covered by the network. Hydrological information is defined in this thesis as quantitative definitions with statistical qualifications of the hydrological characteristics of river basins. The necessity of coordinating the planning of related networks is implicit in the suggested approach.

The planning methodology suggested consists of two optimization phases.

The first optimization phase, which is applied to the whole area covered by the natural-regime net-

work, involves the determination of the main network characteristics which would optimize the network according to an internal objective such as least cost for achieving a certain level of accuracy within a given time interval, or highest accuracy for a given cost (budget). The main characteristics considered in this phase are the number and location of data-acquisition points, the duration of measurement at various points, the frequency of flow measurement, and the technique of interpolation of information. Other characteristics, such as the level of sophistication of instrumentation, could be included using the same procedure. The optimization method used in this phase is the technique of multivariate minimization by the steepest descent. It requires the definition of a series of relationships between the error of estimate of the hydrological characteristic considered and the error sources, such as limited sampling in space and time, limited number of flow measurements for the definition of the stage-discharge relationships, etc. The crucial problems for this first phase are the selection of the information-transfer technique, and the determination of the number of data-acquisition points and of their location so as to avoid redundancy on the one hand and gaps in coverage on the other. The solution of these problems is greatly facilitated by using a bank of hydrometeorological and physiograph-

ical data. This data bank makes it possible both to develop models and transfer information by using data obtained from other networks and sources, and to define hydrological regions which can be used for selecting non-redundant locations of the stations. The combined data bank and techniques developed on its basis also appear to be adequate tools for the gradual acceptance in the network of new methods of data acquisition, particularly those based on remote sensing. The collection of environmental data is also facilitated by the data bank and its associated techniques of analysis.

The first phase is sufficient for the planning of the natural-regime hydrological network in areas where economic development is not envisaged within the planning time horizon. For those areas where such development is foreseen, a second optimization phase is necessary. In the framework of this phase one considers the combined economic conditions of the planned development and of the network. It requires as prerequisites a forecast of the economic development and of relationships between error of estimate of the hydrological characteristics and losses incurred in the developments envisaged. On the basis of this information and of the optimization technique used in phase one it is possible to obtain, by means of dynamic-programming techniques, the variation of the network characteristics during the

planning period which would produce the largest benefit from the forecast development after allowing for network costs.

The first phase is illustrated by using real examples based on the Canadian experience. Because of lack of the prerequisite information required, the second phase has to be discussed on the basis of hypothetical conditions.

The thesis has inherent limitations due to the restricted scope and some gaps in the supporting information in a few areas. Sensitivity analysis is used in areas where such gaps exist, and the results of such analysis indicate the areas where further research would be needed. The thesis is intended to be both theoretically rigorous and practical and may be used as a planning tool, particularly to assess the optimum magnitude and geographic distribution of the main planning variables in areas where the development of the natural-regime network is in the incipient stages, and to eliminate redundancy in areas where the network is overdesigned. Additional research is recommended for further developing some sections of the study. The results of such research combined with those obtained from developed and developing networks and improvements in data-acquisition techniques may bring about the improvement of the method suggested in this study, and perhaps its replacement by a better one.

1. INTRODUCTION

A large number of international and national organizations active in the field of hydrology and many hydrologists have shown during the last decade or two an increased preoccupation with the problems of designing and planning hydrological networks. Although it may appear somewhat illogical that the countries which have recently started to show the most active interest in the problem of network planning have already developed to a large extent their hydrological networks, the situation can be explained at least partially by the following factors: lack of hydrological knowledge during the initial stages of the network development, the pressure on the limited funds and staff to provide data to satisfy immediate needs, and the lack of clearly defined objectives for the network. To these factors one might add the lack of a clear definition of the hydrological network, and of its components and functions.

In all branches of knowledge, and hydrology is no exception, the philosophy of the respective activity was developed much later than the accumulation of knowledge. However, the fact that almost all important hydrological networks were developed in an empirical manner does not mean that those countries which are now in the initial stages of development of their hydrological networks must follow the same

path. Moreover, in many countries with a developed hydrological network there are areas which are badly underdesigned or where there is no hydrological network at all. These areas as well as those which initially or ultimately contain redundant stations will benefit from the development of a scientific approach to hydrological network planning. In addition, those countries and areas which have an adequate hydrological network could still benefit from a design and planning analysis by streamlining their operation and co-ordinating it with the operation and development of related networks. In addition, these benefits may also extend to preparing for the assimilation of the new, more efficient technology of data acquisition and information interpolation. Thus, the increasing interest in the planning and design technology of hydrological networks is perfectly justified.

The ready availability for processing of hydrological data, the development of the statistical analysis and modelling technology as applied to hydrology, the creation of data banks, and the development of system analysis and optimization techniques all of which was made possible by the advent of the digital computers, combined to create the conditions for a significant effort for developing a methodology of planning and designing of hydrometeorological networks. The result of this effort is

reflected in the technical literature and in the national hydrometeorological programmes reviewed partly in Section 1.3. In three of the larger countries of the world, the U.S.A., the U.S.S.R., and Canada, where the problems of network design are most pressing owing to the existence side by side of areas with overdesigned and underdesigned networks, the results of this effort have produced conditions which permit a more rational approach to hydrological network planning. This study is based on the results obtained in the above-mentioned countries, particularly Canada, and on research carried out by the writer for the purpose of this thesis. Before discussing the purpose and scope of this thesis, it will be necessary to clarify the definitions of the hydrological network and its subdivisions and of the concepts of hydrometric data and hydrological information.

1.1 Hydrological Networks: Definition and Classification.

1.1.1 Definition.

The definition of the hydrological network accepted in this thesis is identical to that accepted in Report No. 12 of the World Meteorological Organization (Rodda et al., 1969): "A hydrological network is a programme for systematically acquiring the

requisite information, processing and disseminating it in a like manner". The emphasis is ours, and the interpretation given in this study to the above definition is different from that given in the report mentioned. According to the authors of the report the interpretation of this definition in its broadest sense would be that the hydrological network includes a time series of observations at individual stations and alternative forms of data collection, and also the means for storing, handling, and publishing the data (p. 4 in Rodda et al, 1969). In the interpretation given to the above definition in this thesis, the network includes also the systematic processing of data to obtain information at any point of the area covered by the network. We define in this study data as numerical expressions of measurements on some quantifiable characteristics of phenomena, and information as quantitative definitions with statistical qualifications of these characteristics. A series of tables with daily flow values at a station represent data. The statement that average flow at a certain point of interest is 100 cfs with an error of ± 10% at the 99% confidence level represents information.

This definition and interpretation of a hydrological network may seem controversial, particularly to some agencies responsible for the operation of such networks. However, there are three compelling

reasons for which this interpretation should be accepted. The first relates to the fact that without data interpretation for obtaining generalized information the network is incapable of being improved. We cannot determine where data are scarce and where they are redundant, where the frequency or technique of measurement should be changed, and where and how the hydrological network could benefit from data collected by other related networks and vice versa. The second reason is related to efficiency. Leaving the interpretation of data and generalization of information to the users, as is done in some countries, is extremely inefficient since the same data are processed repeatedly by various people who use different approaches according to purpose, technical ability, budgetary limitations, etc., and who seldom obtain the same answer to an identical question. The third reason is related to safety. It is unsafe to use, for example, inconsistent estimations of flood peaks and volumes. In a cascade of dams, the safety of a dam downstream is equal to or lower than the safety of any of the dams upstream. If by underestimation of the flood peak and volume at a dam upstream this dam is unsafe, the safety of all dams downstream is thus reduced independently of the accuracy of the information used for the design of the downstream dams (unless the dams downstream are

all designed for the event of the failure of the underdesigned dam upstream, which is very costly).

The writer does not suggest, of course, that the organization operating the hydrological network should become a kind of unchallenged "one-party hydrological government" establishing by decree what are the hydrological characteristics of the nation's rivers. The studies on which the collection of information is based must be published and submitted for discussion. Various organizations with hydrological research ability should be called to co-operate with the organization operating the network in the process of data interpretation and of obtaining and generalizing hydrological information. But all this activity must be controlled and co-ordinated by the organization operating the network in order that operation and development may be consistent and optimized.

The writer has considered the problems of the planning of natural-regime hydrological networks in the framework of the interpretation of its definition as given above.

1.1.2 Classification

The literature is abundant in hydrological network classifications based on a variety of viewpoints (see for example Langbein, 1965, Carter and

Benson, 1969, Rodda et al., 1969, Solomon, 1971). We retain in this study the classification suggested by the writer (Solomon, 1971) developed in co-operation with W. B. Langbein. This classification is considered by the writer the most suitable for the purpose of hydrological network planning for considerations which are discussed further in this section. The suggested classification is presented in Fig. 1.1.

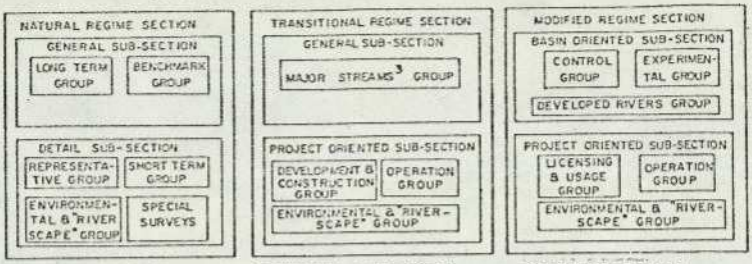
The suggested classification divides the hydrological network into three sections (natural-regime, transitional, and modified-regime), according to the level of interference by man with the hydrological regime. Each section is subdivided into two subsections reflecting variations in space coverage by various types of stations or network-related activities (surveys). Each subsection consists of one to four groups of stations, the groups being constituted on the basis of the use made of the data collected.

The definition of each section, subsection, and group is given in Fig. 1.1.

Further subdivision of the hydrological network according to specific purpose (e.g. flood warning in contrast to hydroelectric plant operation), characteristic measured (level, flow, sediment discharge, etc.), and other criteria is conceivable

FIGURE 1.1

SUGGESTED CLASSIFICATION OF HYDROLOGIC NETWORKS



NATURAL REGIME SECTION
The section of the hydrologic network which covers river basins whose hydrologic regime has not been affected by man.[1]

GENERAL SUB-SECTION
That portion of the natural regime section which collects hydrologic data from rivers of various sizes by means of stations distributed over the whole area covered by the natural regime section, distributed according to the general variation of the physiographic characteristics.

LONG TERM GROUP
Those stations which are operated for a long period of time to provide stable hydrologic characteristics of the natural flow and basic data for extending short time series of data obtained at other stations.

BENCHMARK GROUP
Those stations which are to be maintained indefinately in basins unaffected by man to assess long time trends in the hydrologic logic phenomena.

DETAIL SUB-SECTION
That portion of the natural regime section which collects hydrologic data by means of dense station coverage distributed according to the detailed variation of the physiographic characteristics, or by areal surveys.

REPRESENTATIVE (AND RESEARCH GROUP)
Those stations which collect date in small subregions to assist in understanding the hydrologic processes in their relation to the physiographic and climatic factors.

SHORT TERM GROUP
Those stations operated only for the length of time required to establish the correlation with the flows of long-term and/or benchmark stations, or to calibrate models using climatic data as inputs.

ENVIRONMENTAL AND "RIVER-SCAPE" GROUP
Those stations operated (and occasionally surveys carried out) with the special purpose of obtaining data on the aquatic environment, and on the particularities of the surrounding landscape in relation to the presence of the river.

SPECIAL SURVEYS
Surveys done by means of special stations (e.g. particular time, minimum flows within a period).[2]

1. Basins in which the changes in land use-land cover have reached a steady state can be included, from the time when the steady state was reached, in the natural regime section.
2. Special surveys may extend occasionally over all three sections of the network.

TRANSITIONAL REGIME SECTION
The section of the hydrologic network which covers river basins whose hydrologic regime is affected in some areas or where development is imminent.

GENERAL SUB-SECTION
That portion of the transitional regime section which collects hydrologic data from all major rivers within the area covered by the transitional section by means of stations distributed according to the general variation of the physiographic characteristics and development potential, with the purpose to provide direct information on major streams, and methods of extrapolation of information obtained at stations located on smaller streams.

MAJOR STREAMS GROUP
Same definition as above.

PROJECT ORIENTED SUB-SECTION
That portion of the transitional section which collects data at specific locations where projects are proposed, under construction, or in operation.

DEVELOPMENT AND CONSTRUCTION GROUP
Those stations established at ungauged project sites to check estimates made on the basis of data transfer, to collect data difficult to obtain by data transfer (e.g. ice conditions) and for various purposes including forecasting during the construction.

OPERATION GROUP
Those stations established for obtaining data required for day to day decisions on water management.

ENVIRONMENTAL AND "RIVER-SCAPING" GROUP
Those stations operated (or occasionally surveys carried out) for the special purpose of obtaining data on the changes in the aquatic environment due to changes in the river regime or land environment in connection with the construction or operation of projects, change in land use, etc.and on changes in the landscape and aesthetic quality of the area due to the changes in the hydrologic regime.

3. Major streams are considered in the frame of this classification those streams which could be used during the foreseeable future in projects of international, national or metropolitan significance, or whose floods would produce damages of significance.

MODIFIED REGIME SECTION
The section of the hydrologic network which covers river basins whose hydrologic regime is modified due to regulation, diversion, or other artificial control, and/or by significant, unsteady activity in the river basin.

BASIN ORIENTED NETWORK
That portion of the modified regime section which collects data from river basins affected by man's activity which can not be accounted for because of its diffuse nature.

CONTROL GROUP
The group of stations whose purpose is to monitor the altered hydrologic characteristics due to man's activity of a diffuse nature and which can not be consequently controlled on an individual project basis.

EXPERIMENTAL (AND RESEARCH) GROUP
Those stations which are operated with the purpose of assesing experimentally the effect of man's activity on the hydrologic regime of the rivers.

DEVELOPED RIVERS GROUP
Those stations which are operated with the purpose of obtaining the data required for adjusting the flow data as much as possible to the natural flows.

PROJECT ORIENTED NETWORK
That portion of the modified regime section which collects data at specific locations where projects have been developed.

LICENSING AND USAGE GROUP
Those stations which are operated with the purpose of measuring the amounts of water used and returned changes of water quality, water available to be shared between several users according to interjurisdictional or international agreements, etc.

OPERATION GROUP
Those stations which are operated for obtaining data required for day to day decisions on water management.

ENVIRONMENTAL AND "RIVER-SCAPE" GROUP
Those stations operated (or occasionally surveys carried out for the special purpose of obtaining data on the changes in the aquatic environment due to changes in the river regime or land environment in connection with the operation of projects, changes in land use etc., and on changes in the landscape and aesthetic quality of the area due to the changes in the hydrologic regime.

but is not of any significant interest from the viewpoint of this thesis.

It is important to discuss at this point some basic differences between the two major sections of the hydrological network, the natural-regime section and the modified-regime section, which are relevant to the planning of this network. The first difference relates to the possibility of obtaining information from the data collected. Whereas data obtained from the natural-regime network stations can be used as such for obtaining and generalizing information (this information being improved by the availability of physiographical data and data from related networks but not directly dependent on them), in the case of the modified-regime network, data from the gauging stations alone are not sufficient for this purpose. In the latter case one has to have, in addition to the data from the network, data on the operation of the various projects in the river basin, on changes in land cover and land use, and data from other related networks. This should not be read as an indication that the writer does not consider inter-network co-ordination important for the natural-regime network. On the contrary, such co-ordination is important both for improving the accuracy of the information and for estimating the changes of the hydrological regime which may be expected

when development takes place. What is pointed out here is that the natural-regime network can function independently, whereas the modified-regime network cannot. Because of this the modified-regime network cannot be planned independently.

The second major difference between the two sections refers to the purpose for which the collected data are used. In general, data from the hydrological network are used for the following purposes:

a) Resource inventory
b) Geophysical research
c) Decision-making (selection between development alternatives, or between corrective alternatives)
d) Planning (analysis of alternatives)
e) Design
f) Construction
g) Operation (including flood-warning)
h) Commercial (e.g. real estate activities)

The natural-regime network basically provides data for the users concerned with purposes a and b, and to the extent to which developments extend into the area covered initially by this network, for the users concerned with purposes c to e. The emphasis is usually on information, occasionally on data, and rarely on real time data.

The modified-regime network provides in principle data for users concerned with all of the above purposes, but the emphasis is usually on purposes c to h. Most users need data in addition to information and frequently real time data.

This latter difference in uses generates two additional differences between the two major sections of the hydrological network. One deals with the type of data collected (required), the second with the economic objectives to be used in the planning of the network.

Whereas the users of the natural-regime network most frequently require data and information on flows and occasionally on sediment discharge and water quality, the users of the modified-regime network usually require in addition data on levels, river cross sections, roughness, and ice characteristics, and have heavier requirements for water quality and sediment data. This adds a further complication to the problem of obtaining generalized information in the framework of the modified-regime network since most of the above-listed hydrological characteristics, in addition to having unsteady statistical characteristics due to man's interference, have a discontinuous variation in space, and are therefore very difficult, if not impossible to interpolate. The system of data acquisition used currently is

practically incapable of producing information of this type much beyond the gauging stations. The developing remote-sensing systems of data acquisition might free at some future date the modified-regime network of its dependency upon stations, but this does not change the substance of the problem since the obtaining of information will still depend largely on direct measurement. The point which is made here is that in the case of the modified-regime network the requirements of the user dictate the location of the stations; the instrumentation; frequency and types of measurements; and method of storing, transmitting and disseminating the data. Very little is left for the design and planning of the modified-regime network.

With regard to the economic objectives to be used in planning there is a basic and obvious difference between the two major sections of the hydrological network. In the case of the modified-regime network, which is always closely connected with economic development, it is easy to set an economic optimization objective (even in the case of a corrective project such as a pollution control scheme). This objective is usually formulated as the requirement that the marginal cost of increasing the amount of data or the accuracy of the information should be equal to or less than the

marginal benefit obtained from the additional data or the increased information accuracy. It is much more difficult to set up an economic optimization objective for the natural-regime network. In this case the information may never result in any additional, economically accountable benefit. The natural-regime network is related to the economy only loosely, since it provides an inventory of a resource, and estimating benefits from the network is possible only when development materializes. At this stage the network enters into the transitional section and its economic objective approaches that of the modified-regime network. (As will be shown in Chapters 2 and 6 optimization of this transitional section uses a combination of objectives borrowed from both the natural-regime and modified-regime sections). Under these circumstances the only possible avenue left is to consider the natural-regime network as an independent economic unit which has the function of producing data and information. The economic optimizing objective can be then formulated in terms of the general policy of the country with respect to water resources, as the requirement to obtain a certain level of information accuracy at a minimum cost within a certain period of time, or of obtaining the maximum accuracy possible for a given cost

within a given time, or other similiar permutations as discussed in Sections 2 and 6.

1.2 Purpose and Scope.

The purpose of this thesis is to provide a theoretical analysis and practical guide applicable to various conditions for the planning of the natural-regime section of the hydrological network. The design techniques developed in this thesis consider the optimization of the network with respect to the following decision variables: number and location of stations; duration of station record; number of flow measurements; and method of data transfer. However, other decision variables concerning the network, such as the level of sophistication of the instrumentation, could be included using the same design principles. The purpose of this study includes also the development of the methodology for the planning of the hydrological network at the stage when a portion of the natural-regime section is transformed into a portion of the transitional section in the framework of the planning of economic development in the basins involved.

The scope of this thesis is restricted to the analysis of the planning of the natural-regime network as a whole, at a national scale, and covers

only in a general manner the problems of detail planning of the subsections and groups of this section as defined in Fig. 1.1. Furthermore, the scope of this thesis is restricted to one type of data and information, namely flow. However, the methodology developed may be applied with slight modifications to other types of data networks such as the sediment, water temperature, and water quality networks. The scope of the analysis is generally restricted to the currently used methodology of data acquisition, but consideration is given to the possible changes in this technology and to the implications for the planning of the natural-regime hydrological network.

The scope of this thesis was restricted to the planning of the natural-regime network and the methodology of the planning of the transitional network at the stage of development planning for the following reasons: the study must be kept to a manageable size; the modified-regime network has relatively few decision variables amenable to planning and the economic aspects of the planning of this section of the network have already been studied extensively, as witnessed by the list of papers given in the "Supplementary References"; there are pressing needs for techniques of planning the natural-regime network in many countries, particularly in the developing ones.

1.3 Present Status and Trends in Planning of Natural Regime Networks

The significant differences which exist in the design and planning of the natural-regime, transitional, and modified-regime networks are not always clearly recognized by all organizations responsible for such planning and design or by the hydrologists who published papers on this subject. We shall confine our review of the current practices and trends in this field only to those countries and authors that recognize at least to a certain extent this difference. However, we offer in the "Supplementary References" a list of papers in which there are some theoretical or practical considerations which may be of use in the planning of natural-regime networks, although this section of the hydrological network is not clearly identified by the respective authors.

1.3.1 Review of International Literature

During the last fifteen years or so the World Meteorological Organization has been very active in organizing international co-operation in the field of hydrometeorological network planning. In connection with this the WMO has organized meetings and symposia, and has sponsored the research and publication of papers and reports on the subject.

In 1958 the WMO published a basic paper by Kohler on the design of hydrological networks which was followed shortly afterwards by a Seminar on planning and operation of hydrological networks (WMO, 1960). In 1965 the WMO published the comprehensive "Guide to Hydrometeorological Practice", and organized in co-operation with the International Association of Scientific Hydrology a Symposium on Design of Hydrometeorological Networks (WMO, IASH, 1965). Subsequently the WMO commissioned a study on hydrological network design (Rodda et al., 1969). The Commission for Hydrology of the WMO has appointed a "Rapporteur on Network Design" in the person of W. B. Langbein to prepare inter alia a Casebook on Hydrological Network Design Practice including contributions from various authors active in this field (WMO, 1972).

The basic principles of hydrological network planning which have been postulated in various forms in the above-mentioned references are summarized by Rodda et al. (see section Streamflow Networks, pp 18-19). According to these principles the hydrological network should consist of a basic and an auxiliary network. We discuss here a paragraph on the basic network from Rodda et al. since it illustrates some of the problems raised by the published literature. According to Rodda et al. the

basic network should be planned as a tool to collect full observational data for the whole territory. It should enable derivation of values for the main characteristics of the hydrological regime by means of interpolation between base points." (Although not clearly stated here, it is implied that the auxilary network is also used for this purpose). "The accuracy of the interpolated characteristics is defined by the accuracy of the degree of variability in time and space, and also by the particular demands connected with water use." (The report seems to amalgamate here the obtainable accuracy and the desirable or required accuracy). "This indicates the main principle of distribution of the basic network - its optimum density depending on the physiographic variability influencing the hydrological regime." The problems of desirable accuracy in connection with water use are not considered in the above stated main principle of distribution of the basic network, and this led the writer to the conclusion that Rodda et al. imply that the basic network belongs to the natural-regime section of the hydrological network. Rodda et al, recommend that "the duration of observations in the basic network should be unlimited". The recommended density varies from one station per 2,500 km^2 to one station per 12,500 km^2 on medium

size rivers (5,000 to 50,000 km^2), and sampling methods are recommended for small basins having an area of less than 5,000 km^2. Auxiliary stations, according to Rodda et al. are aimed toward explaining, in conjunction with the basic network, the territorial variability of the hydrological characteristics of the various water bodies, and include also temporary posts installed during short term investigations (special survey stations in our classification), and stations installed for the design of large hydraulic projects (transitional stations in our classification). No specific recommendations are made regarding the density of the auxiliary network, but it is implied (see Table 4 in Rodda et al.) that a density varying from one station per 500 to 1,000 km^2 in mountainous areas to one station per 1,500 to 2,000 km^2 in temperate regions represents the desirable range. These densities are slightly higher than those recommended earlier by WMO (1965). From the viewpoint of the planning in time of the development of the network Rodda et al. presents somewhat contradictory ideas (see p.7 and p. 18 in Rodda et al.) However, a free interpretation of these ideas could be that the basic network should be developed gradually in time, whereas the auxiliary stations should initially be installed in a large number and then either discontinued, or moved to other locations.

Of more significance for the present study are a series of basic ideas related to the planning of the natural-regime network contained in various reports and papers published under the sponsorship of the WMO but which, to our knowledge, have not yet been systematically applied. These basic ideas can be summarized as follows:

- The natural-regime network should be used for inventory purposes and should be established well ahead of development (Langbein, 1965).

- A minimum level of data collection is required in a basic hydrological network regardless of prospective economic development and should provide information needed to estimate or synthesize hydrological characteristics of any locality in the national domain on the basis of correlation with terrain factors (Solomon, 1965; Wallis and Anderson, 1965; Kohler, 1965).

- In connection with the above, hydrological networks should not be considered to include only stations and the corresponding series of data, but also short term surveys (and, in our interpretation, other data) which could help in evaluating the hydrological, geological, and climatological framework of an area to a sufficient degree to enable the estimation of water quantity and quality anywhere in the region (Langbein, 1965; Solomon, 1965; Cavadias and Solomon, 1972).

- Coordination between networks may provide

means for generalization of data and information and thus contribute to increased efficiency of the networks (Langbein, 1965; Kohler, 1965).

- Models could be used to transfer data and information from one network to another (Linsley and Crawford, 1965), and model parameters should be correlated with physiographical characteristics of the basins to enable generalization of data and information (Linsley and Crawford, 1965).

- There are not yet clearly defined objectives of accuracy for the natural regime network (Rainbird, 1965) and density recommendations such as those given for example in WMO, 1965 and Rodda et al., 1969, represent an expedient adopted in the absence of any better solution (Linsley, 1965).

- Recent developments in instrumentation: remote controlled recording and teletransmitting ground instruments, and remote sensing from the ground or aerial platforms are likely to produce significant changes in the technology of data acquisition, and may change in the future the basic conditions for the planning of the natural-regime hydrological network (Langbein, 1965; Linsley, 1965; Rodda et al., 1969; Solomon, 1970).

Another international organization which has been active in the area of hydrological network design, particularly during its sponsorship of the International Hydrological Decade, is UNESCO.

This organization has appointed an "Ad Hoc Working Group on Network Planning and Design" which presented its findings in UNESCO 1966, and later a "Panel of Experts on Network Planning and Design", whose conclusions and recommendations are reported in UNESCO 1967. The basic philosophy of these reports is generally similar to that adopted in the WMO documents reviewed above.

An interesting manifestation of international concern and co-operation in the field of hydrological networks was the CENTO Seminar on Evaluation of Water Resources with Scarce Data which took place in Teheran (CENTO, 1969). Although the Seminar was concerned mainly with problems related to the transitional or modified regime networks, some of the discussions were relevant to our problems. Thus Nace (1969), after discussing the vast literature on the subject of hydrological network design, concludes: "A theoretically based system probably could be devised for design of networks that would adequately[1] define the hydrology of the natural areas. For practical purposes, however, the natural occurrence of water provides only the initial basis for water development. Thereafter the data network must be primarily for operation and management purposes and theoretical

---

[1]. As in most statements of this type the definition of the word "adequate" is not given.

design principles become irrelevant." And later, in the concluding remarks of the Seminar, Nace points out that "... where data are scarce, the new programs should not be merely to add to the bulk of the data or to increase the density of the networks. The goal should be to increase the effectiveness of data that are available or will be obtained."

The review of national literature has to start with the U.S.A. which has produced by far the richest literature related in various ways to hydrological network planning. However, papers dealing specifically with the planning of hydrological networks in general, and with natural-regime networks in particular, are not very numerous. We shall confine our discussion to the papers dealing specifically with hydrological network planning. The other papers which may be of some interest to our problem are listed in "Supplementary References".

The first significant paper on this subject is probably that published by Langbein (1954), in which he advanced the idea of dividing the natural-regime network into basic stations operated for a long period, and secondary (roving) stations operated for short periods at various locations. The next significant publication was the paper published by Benson (1962), in which he developed a correlation between a hydrological characteristic and various river basin characteristics which permits the estimation of the hydro-

logical characteristic at any point of the area for which the correlation was developed. Later Fiering (1965) suggested an optimization technique for the planning of the hydrologic network operation based on the concept of information content introduced by Fisher (1949) who defined it as the reciprocal of the variance of a particular parameter. Fiering assumes a budgetary constraint and tries to determine how many stations should be discontinued, and how many years should the remaining stations be further operated to maximize the information content. Fiering's suggestion did not find any practical applications because his analysis was based on unrealistic operational and statistical assumptions which are not confirmed by the actual data. In the same year Wallis and Anderson (1965) published a paper which indicated the possibility of correlating sediment data with terrain characteristics. In their paper, Wallis and Anderson suggested that certain rotations of the coordinate axes could lead to the development of statistical relations presenting the possibility of obtaining a better interpretation of the physical meaning of the relationships. Matalas (1967) further developed Fiering's idea of optimization of a gauging system. He eliminated the budgetary constraint and introduced the idea of optimization based on the division of the network into regions. According to the technique developed by Matalas the optimum

operation is reached when the number of stations in each region is such that the ratio of marginal information obtained to the corresponding unit cost is the same in each region. This idea has merit, but the technique suggested was incomplete since the accuracy of interpolated data is not a unique function of the number of stations operated, and the manner in which the regionalization can be obtained was not clearly defined. In addition, the possibility of other objectives, such as the obtaining of a certain accuracy within a given time period, was not considered. In the same year Benson and Matalas (1967) published a study on the possibility of synthesizing time series of stream flow using statistical relationships between hydrological and terrain characteristics. This paper was followed by two others on the same general subject by Benson and Carter (1970) and by Thomas and Benson (1970). These latter papers are at the basis of the network planning programme currently under way in the U.S.A., and which will be discussed briefly in Section 1.3.2. Theoretical developments on the errors affecting hydrological data and on the use of models in data synthesis which may be of some significance in the planning of the transitional section of the hydrological network have been presented by Kisiel and his co-workers in a series of papers. A paper by Kisiel (1971) summarizes most of this work. Other papers by Kisiel and his co-workers

are listed in the "Supplementary References". An interesting review on possible applications of systems analysis to hydrological network planning is contained in a contribution by Dawdy, Moss and Matalas (1972) to the Casebook on Hydrological Network Design Practices (WMO, 1972).

In the USSR, similarly to the situation noted above in the U.S.A., the literature related indirectly to hydrological network design is fairly extensive (see "Supplementary References"), but that dealing specifically with the planning of hydrological network design is quite limited. In addition to the paper by Uryvaev (1965) which presents a simplified empirical approach to the problem, the only papers published in the Soviet literature and known to us which are of some significance for the planning of natural-regime hydrological networks are three papers published by Kuzin (1965), Karasev (1968) and Alexeev (1971). Kuzin presents a methodology for hydrological regionalization based on reasonable principles which is, nevertheless (see Section 7.1.2) very difficult to apply in practice. Karasev suggests a method of assessing the number and location of the basic stations of the natural-regime network based on one hand, on Kuzin's regionalization, and on the other on a minimum and maximum size of the incremental river basin corresponding to a basic gauge. The minimum basin

size is based on the idea that the incremental flow measured by a gauging station should be at least equal to the measurement error with a certain level of confidence at the station upstream, and the maximum incremental area is defined from the condition that the coefficient of correlation between two successive stations on a river or between two gauging stations on two adjacent rivers should be such that the error of interpolation should be lower than the error of measurement. An arbitrary number of auxiliary stations (15%) and a certain number of stations on major rivers is added by Karasev to those determined as shown above to form the natural-regime hydrological network. Karasev applied his technique to the whole of the USSR and the results are discussed in Section 1.3.2 which includes also a critique of the approach.[1] Alexeev (1971) tried to develop a formula for establishing the admissible distance between gauges for which one obtains an admissible error of data transfer between three gauging stations located around the point where the data transfer is made. This type of technique is no longer of any practical use since interpolation techniques using all stations in the region are currently applied.

---

[1]. It is not clear from Karasev's paper if and how he accounts for the superimposition of the error of interpolation over other errors.

In Canada the literature on hydrological network planning reflects mainly the efforts of the Water Survey of Canada, currently a branch of Environment Canada, to put the planning of the hydrological network on a rational basis. Thus, on the basis of large scale hydrological studies carried out in the Province of Newfoundland the writer and his co-workers (Solomon et al., 1968) presented the advantages of information interpolation by joint estimation of precipitation, temperature, and runoff, by means of a square-grid hydrometeorological, land-use land-cover, and physiographical data bank. The possible future changes in hydrological network operation conditions due to the development of new techniques of data acquisition, and their implications for the planning of the hydrological network were discussed later by the writer (Solomon, 1970). Cuthbert (1971) published a study on the application of the square-grid hydrometeorological and physiographical data bank for interpolation of hydrological information in the Maritime Provinces using the techniques developed by Solomon in his Newfoundland studies.

Pentland and Cuthbert (1971), using the results of studies for the planning of the hydrological network commissioned by the Water Survey of Canada described in Section 1.3.2 (particularly Shawinigan 1970; Ingeldow, 1970) published a paper on the development

of synthetic hydrological information based on the square-grid data bank. In the same year the writer and his co-workers published a paper (Solomon et al., 1971) on the use of a parametric model in conjunction with the square-grid data bank in solving several problems related to hydrological network planning. The use of the square-grid data bank for hydrological regionalization and other applications are discussed in other papers by the writer (Solomon 1972 a, b; Cavadias and Solomon, 1972). Further discussions on the above-mentioned papers are included in various sections of this thesis. A significant amount of documentation on the network planning activity in Canada is included in the reports prepared by consultants for this purpose, discussed briefly in Section 1.3.2. Other papers published by Canadian hydrologists which are indirectly related to the planning of the natural-regime hydrological networks are listed in the "Supplementary References".

According to the information available to the writer there is some activity regarding the planning of natural-regime networks going on in a series of other countries. This is reflected for example in the literature published by hydrologists from France (Roche, 1967; Dubreuil, 1969), the UK (Water Resources Board, 1969), Romania (Goldstein, 1968), Hungary (Szesztay, 1967), Australia (Chapman, 1972), New

Zealand (Toebes, 1965), and the Netherlands (van der Made, 1972). All these papers represent valuable contributions to the information on specific hydrological conditions prevailing in the respective countries, but the network planning methodology described generally presents little that is different from what was suggested in the WMO and other papers analysed earlier in this section.

1.3.2 Review of International Practice.

We confine this review to the U.S.A., the U.S.S.R., and Canada. This limitation is introduced because the methodologies developed in these three countries for planning the natural-regime network represent three distinct viewpoints, whereas the methodologies developed in other countries can be reduced in most cases to various combinations of these three.

a) The U.S.A. The U.S. Geological Survey, which is responsible for the national hydrological network, currently is carrying out an evaluation programme of the existing network, which is used as a basis for planning the expansion or contraction of the network. This programme is based on a study mentioned in Section 1.3.1. (Thomas and Benson, 1970), and on a series of studies carried out in several states. An example of such studies is the Report on the Evaluation of the Stream-Gauging Programme in Kansas (Jordan and Hedman, 1970) which is summarized by Langbein in the Casebook on Hydrometeorological Design Practices (WMO,

1972). The programme is based on the concept that stream-flow information could be needed at any point of the area covered by the network and that the network must be designed to accommodate this need. The hydrological data needed are classified in four categories out of which three (data required for the planning and design of water resources projects, long term streamflow data to define trends, and environmental data) can be considered to correspond to the data obtained from the natural-regime and transitional section of the network according to our classification (Section 1.1.2). Each stream is classified as principal or minor according to a drainage basin size criterion. The planning objectives for the natural regime network were established on the basis of engineering judgment and include a level of accuracy of interpolated (transferred) information equivalent to the accuracy obtained from direct observations for a period of 10 years in the case of minor streams; and for a period of 25 years in the case of principal streams. Interpolation of information (data transfer) is obtained by means of multiple correlations with physiographical characteristics, the above being considered to include also climatical characteristics such as average precipitation and snowfall in the corresponding basin, various precipitation intensity indices, an evaporation index not clearly defined in the papers published, etc. If the multiple correlat-

ions succeed in estimating the various hydrological characteristics considered with the accuracy required, the network is considered satisfactory and the discontinuation of some of the stations is considered. In the contrary case the installation of additional stations is recommended. The location of new stations on minor streams is recommended to be selected in such manner that a good areal distribution is obtained, and the stations are located in basins having different physical characteristics. The location of new stations on principal streams is recommended to be selected in such manner that each new station measures the flow of a basin approximately twice as large as that of the station immediately upstream. If only a few of the hydrological characteristics cannot be interpolated with the required accuracy, then partial record stations (e.g. for the measurement of low flows) are recommended.

The main points which are debatable in this programme are: the arbitrary selection of the objectives, the reliance on the standard error of estimate of the data included in the correlation, which may be completely different from that of predicted (interpolated or extrapolated) data;[1] the failure to recog-

---

[1]. Some indication of the errors one may expect when new information is estimated is included in Table 11 of Thomas and Benson, 1970, but this is not considered in the planning process.

nize in the planning analysis that the interpolation error is superimposed over other errors, particularly the error due to the length of the sampling period; the lack of an error-density analysis to be used as a basis for estimating the number of additional stations required; and the use of data obtained from the meteorological network in a subjective manner (use of isoline maps, of various correction factors obtained from particular papers, etc.). On the whole, the programme may have merit under the specific conditions prevailing in the U.S., but the writer considers that its adoptation in other countries would require considerable reconsideration, equivalent practically to the development of a completely new approach.

It is interesting to note that in 1970 there were in Kansas about 120 stations of which 51 can be considered to belong to the natural-regime network (Jordan and Hedman, 1970). The evaluation study concludes that some 15 new stations would be necessary in order to reach the desired accuracy (for both minor and principal streams). This leads to an approximate network density of one station in 1,400 sq. mi.

b) The USSR. The writer does not have direct information on the current network planning activity in the USSR and hypothesizes that the network evaluation and development programme included in Karasev's paper

(1968) represents more or less an official viewpoint. On the basis of the network planning methodology described in Section 1.3.1, Karasev made a network evaluation by regions and sub-regions and concluded that in the European portion of the USSR, 10 out of 18 sub-regions already have an "optimum" number of stations or more, and in the Asiatic portion of the USSR only 8 out of 29. The extreme Arctic area and the deserts were excluded from the analysis. In the Caucasus and Central Asia a simplified analysis was applied since the complex hydrological and physiographical characteristics of these areas make the application of the methodology suggested by Karasev very difficult. The study recommends an increase in the total number of natural regime stations from the present (1968) number of 4,847 to 6,492, which corresponds roughly to a density of one station per 1,100 sq. mi.

As in the case of the U.S. methodology, it is debatable that the USSR methodology could be applied elsewhere for the following reasons: it uses a subjectively defined statistical-physiographical optimization criterion; it fails to recognize the effect of super-imposition of errors on network planning; it uses a system of regionalization which includes a large amount of subjective judgment and probably an excessive amount of manual work (Kuzin,

1965); it includes qualitative indications which are used for quantitative assessments without proper definition (e.g. the minimum size of the basin which has a hydrological regime which reflects "regional" and not "local" conditions); there is no concern for coordination with other networks; the possible advantages of various methods of interpolation other than isolines and simple correlations between data at neighbouring stations are not considered; and, as recognized by Karasev himself, the analysis suggested cannot be applied to areas having complex hydrological and physiographical conditions. Nevertheless, the writer knows of one instance in which Karasev's methodology was applied for the planning of the natural-regime network outside the USSR, namely in a study of the Hydrological network of the Province of Quebec, Canada (CEQEAU, 1971). In this study Karasev's methodology was preferred over other methodologies considered. An analysis of the Quebec study indicates however, that the unavailability of a systematic, consolidated presentation of the methodology developed in Canada, which is the object of the present thesis, explains to a large extent the preference given to Karasev's method in that study.

c) Canada. The Water Survey of Canada of the Department of the Environment, initiated in 1968 a series of studies for the planning of the hydrological net-

work which were carried out in co-operation with consulting engineering firms and universities (Shawinigan, 1968, 1970; Ingeldow, 1970; CEQEAU, 1971; Solomon, 1972c). Some of these area studies tested methodologies developed elsewhere, as already mentioned (CEQEAU, 1971), but they were used mainly as a framework for developing an original methodology for planning the natural-regime network. The latter is based on the use of an interfacing data bank, which makes possible the combined use of physiographical, land-use land-cover, and hydrometeorological information (Shawinigan, 1968, 1970; Ingeldow, 1970; Solomon, 1972c). Optimization techniques which permit a rational allocation of funds between the various activities which may improve the accuracy of the hydrometeorological information were applied within the framework of these studies (Shawinigan, 1970; Solomon, 1972c). The use of models for information transfer was investigated in two of the studies (Shawinigan, 1970; CEQEAU, 1971) and the significance of this technique for natural-regime network planning was analysed by Solomon (1972c). Methods of hydrological regionalization based on the use of the interfacing data bank, which can be applied in an objective manner and used for obtaining indications on the rational distribution of new gaugings were also developed in these

studies (Shawinigan, 1968, 1970). Currently the Water Survey of Canada is carrying out a project for combining the various regional square-grid data banks developed within the framework of the various studies mentioned into a national one to be used in practical problems of information transfer and hydrological forecasting. Furthermore, the Water Survey of Canada is considering the extension of the natural-regime network from the present 1,200 stations to the 2,000 stations recommended in the various studies. The recommended number of stations of the natural-regime network would correspond roughly to one station per 1,500 sq. mi.

It should be noted however, that the methodology of the network planning has evolved gradually in the studies carried out in Canada, which were not based on a systematic approach available before the studies were initiated. This is reflected by some variation in the conclusions presented in the various reports of the successive studies. A systematization and consolidation of the theoretical and practical work done in the area of hydrological network planning in Canada is, therefore, highly desirable.

2. PLANNING FRAMEWORK

The development of a rational planning methodology of the natural-regime network requires the definition of the planning criteria, objectives, and decision variables. Furthermore, it requires the selection of the appropriate mathematical techniques for determining the values of the decision variables, i.e. the optimization techniques. These problems are discussed in this section. In addition this section contains supplementary considerations on social-environmental factors which cannot be directly included in the mathematical planning analysis but may influence the selection of the decision variables.

2.1 Criteria, Objectives, and Decision Variables.

We define in the framework of this study as criteria those hydrological characteristics which are used to indicate the adequacy or otherwise of the hydrological network. We define as objectives the levels of accuracy of estimating the criteria which are required in order to consider the hydrological network as adequate. Accuracy is expressed in percent and is defined in the present thesis as the difference between 100% and the relative standard error of estimate in percent. Decision variables are those numerical and/or operational character-

istics of the network which are determined in the planning process in such a way that they make it possible to achieve the planning objectives.

2.1.1 Criteria.

Current international practice and literature indicates a wide variation in the selection of the criteria. Thus in the U.S.A. (Thomas and Benson, 1970), a number of 71 criteria (called indices of streamflow characteristics) were used. These criteria ranged all the way from the long-term average flow to the 10 year peak flood flow to the lag-one serial correlation coefficient. One could obviously increase ad infinitum the number of criteria, using for example floods with various probabilities, or the serial correlation coefficients with various lags. As indicated by Karasev (1968), in the USSR the main criterion is the long-term average flow. In Canadian practice the criteria used have been the long-term average flow, the coefficient of variation of the annual flow, monthly average flows (Shawinigan, 1970; Solomon, 1972c), and occasionally maximum and minimum flows with various probabilities (Ingeldow, 1970; CEQEAU, 1971).

In fact it would be impossible to select criteria which should be considered valid for all situations. The criteria may well vary from one country to another, or even within a country from one region to another.

In regions in which the flows vary smoothly and can be readily regulated, as for example in some areas of Canada and the USSR, the most important from the viewpoint of the natural-regime network may appear to be the long-term average flow and the coefficient of variation of the annual flow which determine the amount of water available and the volume necessary for full regulation. In very arid areas both the above-mentioned criteria may loose their significance. In such areas the criteria may become the number of days with measurable river flow, or the proportion of surface supply water which is infiltrating to the ground water during a rainy season, etc. A truly flexible method should be capable of using for the planning of the network any criterion selected in accordance with the particular problems of the area considered. The methodology used in this thesis offers this flexibility. Although the practical applications presented are limited to a few criteria considered significant for Canada, the methodology is developed to a level which permits the use of any criterion, provided that the required relationships for the determination of the errors of estimate and of the related costs can be established well enough empirically. Empirical relationships were preferred over theoretical ones since most hydrological data do not satisfy the basic

statistical assumptions on which are based the currently used theoretical derivations of relations between error of estimate and the decision variables. It should be noted that with the change in the selected criteria the decision variables may also change; for example if the criterion is suspended sediment, the decision variables would have to include the number of stations with sediment measurements; if the criterion is maximum annual level, the decision variables would have to include the number of stations with partial record of maximum annual levels.

2.1.2 Objectives.

As has been indicated in Section 1.3, the international literature and practice is extremely vague with respect to the objectives of the natural-regime network. The WMO publications give indications of desirable station densities in lieu of objectives (WMO, 1960, 1965). In the U.S. the level of accuracy which is considered as the desirable objective is equated to that obtainable from a given number of years of direct observations, and this number is determined on the basis of experience and engineering judgment (Thomas and Benson, 1970; Jordan and Hedman, 1970). Obviously experience and judgment come to a large extent after the network has been in operation for a certain period of time, as is the case in the U.S., and it is doubtful that this experience could

be successfully transferred to other countries. In the USSR Karasev (1968) sets out the objective of keeping the error of interpolation within the limits of the order of magnitude of the error of measurement, which assumes that the stations have been already operated for a very long period of time to eliminate the error due to the limited period of sampling. Obviously, networks in the development stage do not fulfil this condition. In Canada two possible ways of establishing the objectives were considered: one is based on the maximization of the incremental information per unit cost, and the other on the idea that the definition of the objective should be a matter of administrative decision based on the knowledge of the minimum cost to achieve this objective and on the Government's priorities based on its economic and socio-environmental policy (Shawinigan Engineering 1970, Solomon, 1972C)

We consider in this thesis two complementary methods of determining the network objectives. The first refers to the natural-regime network proper, the second to the development phase of a natural-regime network into a transitional one in relation to economic development planned in the area.

The basic objective of the planning of the natural-regime network is considered in this thesis to be the achievement of the maximum possible accuracy

of the interpolated information for a given budget. In addition, according to the existence of various constraints regarding the desirable accuracy, such as time limits, the objective of the natural-regime network can be formulated in various manners as indicated in Table 2.1.

TABLE 2.1

Accuracy, cost, and time considerations for establishing the objectives of the natural-regime network.

No.	Accuracy	Cost	Time	$\frac{A_f - A_i}{C}$
1	$A_f \geqslant A_o$	Results	Results	Maximized
2	$A_f \geqslant A_o$	Results	$T \leqslant T_o$	Maximized
3	Maximized	$C = C_o$	$T = T_o$	Results
4	$A_f = A_o$	Minimized	$T = T_o$	Results

In the above table A_f and A_i represent the final and initial accuracy.

In the case of the transitional network, actually of that portion of the natural-regime network which is transformed into a transitional network in connection with planned economic development, the objective accepted in this study is formulated as follows: to maximize the incremental net benefit from the projects included in the planned development due to increased accuracy of the hydrological information, taking into account the incremental hydrological network costs involved in the increase of the accuracy.

The mathematical techniques required for the

application of these objectives are discussed in principle in Section 2.2 and in more detail in Chapter 6.

2.1.3 Decision Variables.

The definition of the objective of network planning in conjunction with an analysis of its structure and operation makes it possible to determine the basic decision variables. These are selected from the various network characteristics, taking into account the possible impact on accuracy and costs (the basic elements of the objective formulation). A more detailed analysis of these two elements is given in Chapters 4 and 5. On the basis of these considerations the following list of decision variables has been compiled:

- number, location, instrumentation, and date of installation of new gauging station;
- number, location, and date of discontinuation of existing stations;
- the method of information interpolation;
- frequency of observations and measurements;
- changes in instrumentation at existing stations; and
- the time horizon of the planning operation.

As will be shown in Chapters 4 and 6, for practical considerations, only a few of the decision variables are determined in the framework of the

planning analysis proper. Others could be determined more readily in the framework of subsequent detailed design and operation optimization computations.

2.2. Optimization

From the discussion of the objectives and decision variables in Sub-sections 2.1.2. and 2.1.3, it follows that the problem of the planning of the natural-regime network is relatively simple, whereas that of the transitional network is more complicated both from the viewpoint of the mathematical techniques and of the background information required.

2.2.1 Optimization Technique for the Planning of the Natural-Regime Section.

In the case of the planning of the natural-regime section the problem is reduced to the minimization or maximization of a function F of the decision variables relating cost to accuracy

$$F \text{ (Decision variables)} \rightarrow \begin{matrix} \text{min.} \\ \text{or} \\ \text{max.} \end{matrix}$$

subject to a series of constraints.

The two major problems related to the solution of the problem are the determination of the shape of the function F and of the constraints, and the finding of a suitable optimization technique. The development of relationships between the decision variables

on one hand and accuracy and cost on the other, to be used for the definition of the function F is discussed in Chapters 4 and 5. It can be seen from these Chapters that, allowing for some simplifications, it is possible to write the function F explicitly in terms of the numerical decision variables for any of the objectives 1 to 4 given in Table 2.1, as well as of the constraint functions, i.e. the relations between the constraints and the decision variable

$$C \text{ (Decision variables)} = C_o$$

and

$$A \text{ (Decision variables)} = A_o$$

which are required in the case of the objectives 3 and 4 in Table 2.1.

Analytical solutions based on the computations of first and second order derivatives with respect to the various decision variables were tried first and found complicated and expensive. The solution used finally was the Cauchy gradient or steepest descent (ascent) method which employs Rosenbrock's algorithm (Wilde and Beightler, 1967). Several computer subroutines for the application of the method are available (e.g. Subroutine RNK 7 at McGill University and Subroutine CLIMB at the University of Waterloo). These subroutines will minimize or maximize a function of several variables subject to a series of constraints, indicating also the cases when there is no minimum or maximum within the searched

domain. The application of the technique is straightforward in the case of objectives 1 and 2, Table 2.1. In the case of objectives 3 and 4 shown in the same table, it is required to use Lagrange multipliers in order to obtain the objective function. The principles of the application of Lagrange multipliers to the solution of such problems can be found in any standard book on optimization techniques (e.g. Wilde and Beightler, 1967). A particularly clear presentation of the theoretical basis and principles involved in the application of the technique can be found in Cooper, L. and D. Steinberg, Introduction to Methods of Optimization, W.B. Saunders, 1970, pp. 119-124.

Since in the application of the steepest descent technique there is occasionally a danger of finding a local instead of a general minimum, the application of the technique was checked on several occasions by means of the mapping of the functions in terms of a reduced number of decision variables (3). The results were found to be correct in all cases.

Because of the simplicity of application, the technique can be readily used in sensitivity analyses and regional studies which require a large number of changes in the constants of the objective and constraint functions. The results of such sensitivity analyses are discussed in Section 8.1.

2.2.2 Transitional Section

Tangible economic benefits can be obtained from the natural-regime network only to the extent to which this is included in the transitional network, i.e. to the extent to which economic development is planned and occurs in areas covered initially by the natural-regime network. These economic benefits represent actually the economic justification for the expenses incurred with the natural-regime network.

Several methods were investigated in relation to the optimization of the network in the phase when the natural-regime section is transformed into the transitional section. These investigations led to the conclusion that the problem could be solved best by means of dynamic programming (Solomon, 1972c, d). The following information is required for the application of the technique (in addition to the information required for the planning of the natural-regime section):

a) A forecast for the area concerned of the planned economic development primarily in fields requiring hydrological information;

b) Relationships between errors in estimation of hydrological information and losses due to such errors in various types of projects.

If this information is available, the optimization of the hydrological network in the area

designated for development can proceed according to the following steps:

1) Select stages (points in time) preferably to coincide with points in time when significant new developments are forecast to occur; equal time intervals between stages are preferable but not essential;

2) Select range of error of estimate which may be expected to characterize the natural-regime network at the beginning of each stage in the area; divide these ranges into several (4 to 8) intervals. The average of each of these intervals is considered in the following as a state;

3) Compute for each time interval between two stages the minimum cost of increasing the accuracy corresponding to each state at the beginning of one interval to the accuracy corresponding to that state at the end of the interval (objective 4 in Table 2.1);

4) Compute estimated economic losses due to errors in the hydrologic data for each stage and for each transition from one state at the beginning of the stage to each of the possible states at the beginning of the next stage;

5) Add cost of network development to losses due to errors corresponding to the transition from one state to the next and use the currently available dynamic programming techniques to determine the path corresponding to the minimum total costs (cost of network plus cost of losses).

The optimum path determined in this manner would indicate for each stage the decision variables and consequently will give the basic information for the optimization of the network development plan. Further elaboration of the technique is given in Section 6.2, and the results of a practical application are discussed in Sub-Section 8.1.2.

2.3 Economic and Other Considerations

Although the primary purpose of the natural-regime hydrological network is the obtaining of hydrological information in areas not affected significantly by man (or in areas where man's activity has a steady pattern), its planners have to recognize that sooner or later some of these areas will be developed. In connection with this, the planning of the natural-regime hydrological network has to consider two additional problems; the priority development of the network in connection with development potential and the conservation of certain areas of hydrological and environmental interest.

2.3.1 Priority Development

The planning objectives of the natural-regime network as defined in Sub-Section 2.1.2. do not make any distinction between regions with higher and lower development potential, except for areas already designated for development. Nevertheless, it is

advisable to consider in the network planning the possibility of giving priority to areas in which development is likely to occur sooner than in others. There are two reasons for this: the first is economic, the provision of a better hydrological background to development plans when these will be considered; the second is hydrological, to provide a denser network in areas in which the duration of operation of the natural network will be shorter and thus compensate by areal sampling for the deficiency in time sampling.

In order to proceed with priority planning in areas with higher development potential it would be advisable to separate these areas and apply different constraints in the planning objectives. If for example, objectives 1, 2, or 4 Table 2.1 are considered, one could use higher values than in the other areas for A_o in the definition of the objective. If objective 3 is used, then the budget constraint C_o could be divided between the regions in accordance to their areas and development potential. Obviously this differentiated distribution of constraints or funds has to be based on judgment and rigourous rules cannot be offered. However, the following suggestion can be used in a first approximation:

If objectives 1, 2, or 4 are used, select for the various regions values of A_o proportional to the

reciprocal of the square root of the number of years for which each region is expected to remain undeveloped.

If objective 3 is used, distribute C_o among the various regions proportional to the product between their area and the reciprocal of the square root of the number of years for which the corresponding region is expected to remain undeveloped.

2.3.2 Conservation.

As development proceeds, certain areas are allocated for perpetual conservation as natural parks, wild-life reserves, etc. One criterion for the selection of certain areas for conservation could be the uniqueness of their hydrological features, their "riverscape" (e.g. particular water falls, particular areas of interchange between surface and ground waters, lakes and their supply areas, etc.) in combination with particular ecological conditions (unique vegetation-faunal eco-systems, vanishing species, etc.). The delineation of such areas has to be done in the early stages of network planning, using maps, aerial surveys, and any other pertinent sources of information. After the delineation of areas of potential interest from this viewpoint, it is advisable to complete the information to the largest extent possible with ground surveys.

It is advisable to plan the hydrological network

in such areas in cooperation with the organizations responsible for conservation in order to arrange that the conservation areas include whole hydrological basins to make possible the establishment of bench-mark gauging stations in the conservation areas. These should be complemented with other hydrological programmes which may be of interest in such areas to the organizations involved in water resources and/or conservation problems. This problem is discussed further in Section 7.3.

3. DATA AVAILABILITY AND DATA HANDLING

From the interpretation of the definition of a natural-regime hydrological network accepted in this thesis, according to which the hydrological network should be designed and operated to provide hydrological information at any point of the area covered by the network, it follows naturally that the data with which the natural-regime network operates are not limited to the classical hydrometric data (levels, velocities, flows, sediment concentrations). The sphere of the data which are pertinent for the hydrological network extends much further to include data required for the development of models to interpolate hydrological information. Regardless of the type of model used* all take advantage to a certain

---

* We consider in the framework of this thesis three basic types of models: Statistical, in which relationships between hydrological phenomena and causative factors are determined basically through a trial and error process, with physical analysis being used only as a means of making the process more efficient; deterministic-empirical, in which relationships are determined on the basis of physical-mathematical considerations or in which empirical (statistical) relationships determined elsewhere are considered valid in the study area; and parametrical, in which deterministic, empirical and statistical techniques are combined.

extent of the fact that hydrological phenomena are an arc of the hydrological cycle, and that the areal variability of this and all the other arcs of the hydrological cycle depends to a large extent on the physiography of the area. Consequently the data available to the hydrological network for fulfilling its tasks can be grouped into three main categories;

1) Hydrometric data;
2) Hydrological cycle data;
3) Physiographical data*;

In most countries, including Canada from which the examples illustrating this thesis are taken, data included in the two latter groups are collected normally by agencies other than those running the hydrological network. The above have to solve in this case, the problem of co-ordination between networks and surveys for obtaining the desired data in the appropriate amount, distribution, and form. This together with the problem of overall efficiency are not simple matters by any means. These aspects of the controversial problem of overall design of networks and surveys extend far beyond the scope of this thesis, although it is conceivable that the

---

* In this thesis the physiographical data are considered to include topographical, geomorphological, pedological and land cover data, and surficial and bedrock geological data, as well as climatic data. whereas terrain data include all of the above, except climatic data.

methodology developed in this thesis might be expanded later to the analysis of the more general problem. For the purpose of this thesis the assumption is made that the hydrometric network initially uses the hydrological cycle and physiographical data as provided by the other agencies, and - on the basis of its internal optimization - recommends changes to the other agencies. The latter could then optimize their operation on the basis of these requirements and of those from other users, and by an iterative process some form of overall optimum design could be finally reached.

3.1 Data Availability.

On the basis of the foregoing discussion, data availability for the planning of the natural-regime section of the hydrological network could be defined by means of an inventory of data collected by various agencies which would indicate:

a) The number of hydrometric gauging stations, average density and density per statistical-hydrological and physiographical-hydrological region as defined in Chapter 7 (initially a preliminary, tentative regionalization could be used), instrumentation, period of record, types of observations, frequency of measurement for various intervals of the phenomena measured, the statistics of the error of measurement,

and the regional distribution of the characteristics listed above. This should be supplemented by a description of the format of data storage, computer readability, and soft and hardware indications on possibilities of direct computer access to the data.

b) Similar information on networks for the arcs of the hydrological cycle other than the hydrometric one (meteorological, soil moisture, groundwater, evaporation, snow, and glacier networks and surveys).

c) Information on the physiographical (topographical, geomorphological, pedological and land cover, and geological) surveys, including average and regional density of observations, statistics of measurement and interpolation errors of various map representations, the storage characteristics, availability of digitized maps.

d) Delineation of research basins and information as in a), b), and c) above for each research basin, and summaries of pertinent research reports.

Unfortunately, this information is not yet available in the format required for efficient data handling. Development of an interfacing data bank for water resource planning, now in a preliminary stage of development in Canada by the writer for the Department of the Environment, is intended to fill this important gap. For the purpose of this thesis we used the available hydrometeorological, physiographical, and land cover data bank described by

Solomon and Qureshi (1972).

3.2 Data Handling

The elaboration of a methodology of planning the development of natural-regime networks which could have applicability in areas other than the one for which it was initially designed required the development of an objective system of handling the data, and of applying models for information transfer. Several techniques of achieving this could be imagined. One such technique is that developed in Canade for the Water Survey described by Solomon and Qureshi (1972). A series of studies for improving it are currently under way, but even at the present level of development it serves in a satisfactory manner the needs of natural-regime network planning. Therefore, this technique was used in the present thesis for the illustrative examples of the methodology developed.

This technique consists of the development of an interfacing data bank providing a systematic framework for storing and processing hydrological, physiographical and land-use land-cover data and for the objective application of various methods of information interpolation. It consists basically of the following components (Fig. 3-1):

a) The space-time reference system;

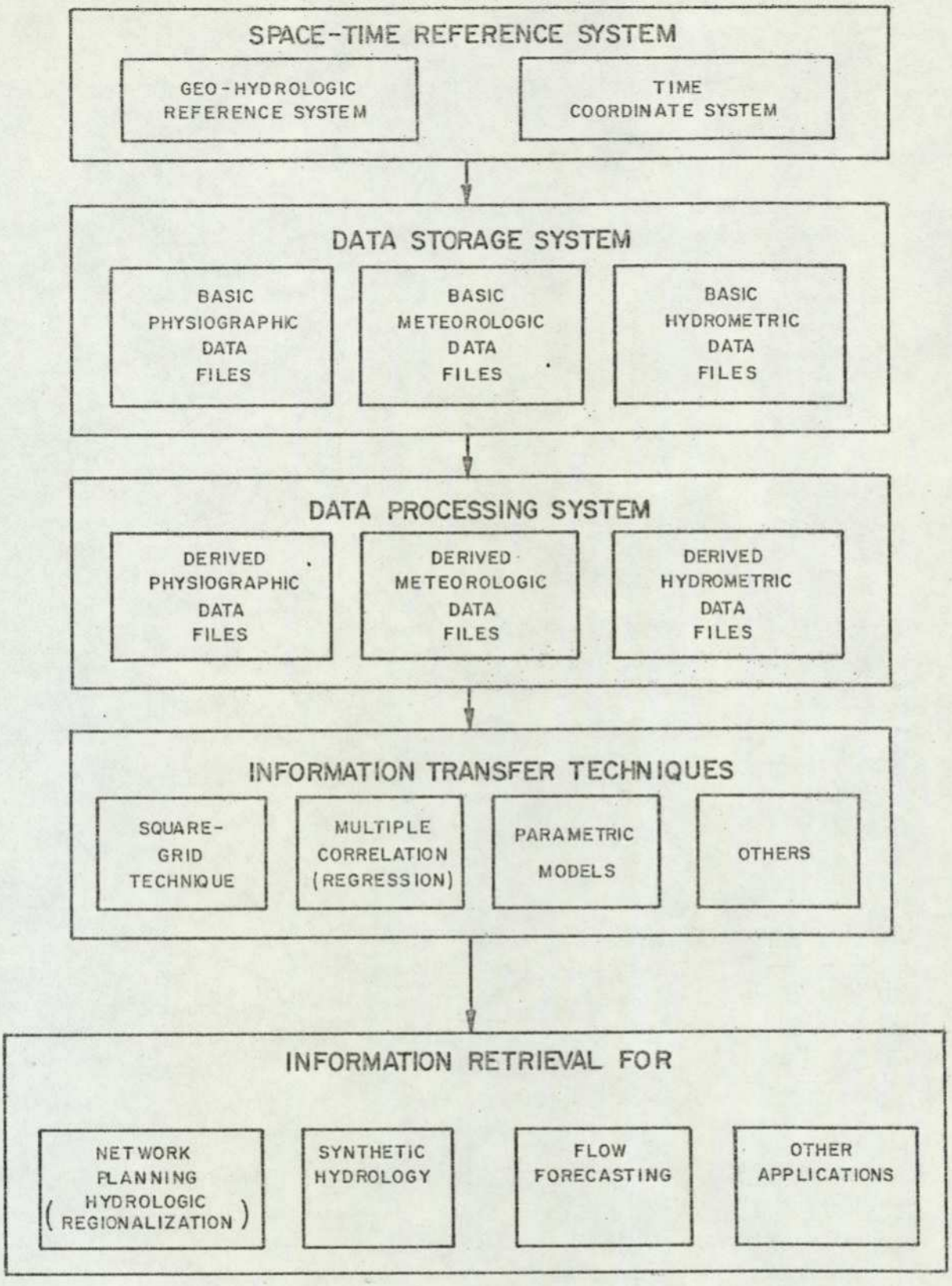


Figure 3-1 BASIC LAYOUT OF THE PHYSIOGRAPHICAL LAND-USE LAND-COVER AND HYDROMETEOROLOGICAL DATA BANK

b) Data storage including data screening;

c) Data processing;

d) Information transfer techniques;

e) Information retrieval.

The space-time reference system consists in turn of a geo-hydrological reference system and a time co-ordinate system.

The geo-hydrological reference system contains two elements: the first is a square grid system which consists of a matrix of squares covering the area investigated and corresponding to the UTM* reference system. The grid interval as used at present is uniform (10 x 10 km in Northern and Western Canada, 5 x 5 km in Southern Ontario), but a technique based on digitized maps and making it possible to use a system with variable grid intervals was developed recently. (Solomon, 1974)

The index number of the row and column of the square which is uniquely related to the UTM identification system provides the required information on the location of each square. A sub-system of coordinates within each square can be added and may provide the possibility of locating further details regarding the gauging stations, the main river channels, etc.

The second element of the geo-hydrological reference system consists of a technique of identifying the runoff path starting with the divide

---

* Universal Transverse Mercator

between basins and ending at the sea. This enables the determination within each square containing a divide between basins of how many distinct basins are supplied by the runoff from the square considered, and the corresponding areas and direction of outflow with respect to the adjacent squares. For those squares which do not contain divides, the "inflowing areas or squares" and the square in which the outflow occurs are indicated (Fig. 3-2). This permits the determination at any point of the area of the drainage basin above it, the composition of squares and square sub-divisions constituting it, and the flow path of any surface runoff. A system of automatically delineating the flow directions, flow paths and watershed boundaries based on digitized maps was developed recently by the writer and his co-workers (Solomon, 1973, 1974).

The time coordinate system consists of an origin and reference time intervals, both of which can be selected arbitrarily according to the requirements of the problem at hand. Current research in connection with the time reference system is aimed toward introducing a variable time interval in accordance with the rate of variation of the time series investigated, in line with techniques of data acquisition used in communication theory.

Component b) data storage, consists of hydrometric and meteorological data, and of digitized

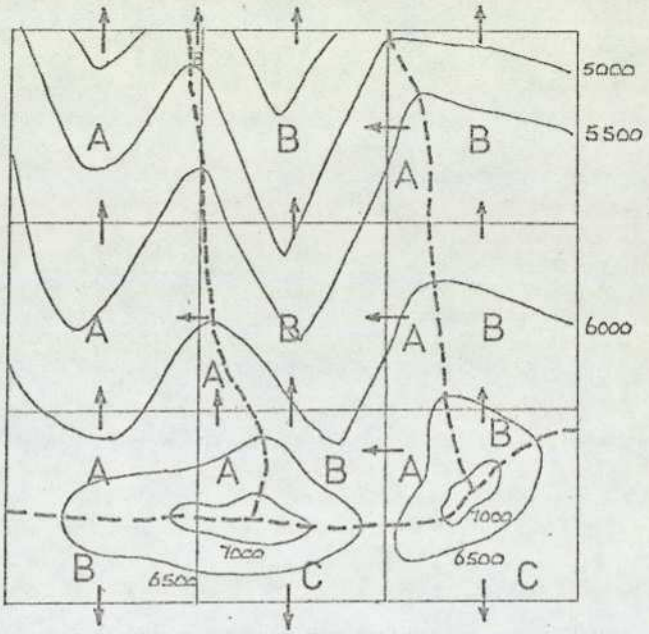


Fig. 3-2 SQUARE GRID SYSTEM SUPERIMPOSED ON BASIN TOPOGRAPHY FOR AUTOMATIC DETERMINATION OF DRAINAGE BASIN AREA ALONG A RIVER

Note: 1) Capital letters indicate the subareas included in each square draining in different basins

2) Arrows indicate main direction of runoff as determined by the computer program

3) For details on the computer program, see Solomon (1974)

topographical, physiographical, land-use, and land-cover maps.

The hydrometric data in the active storage (discs) include at present only monthly flow and sediment discharge time series recorded at the gauging stations. The passive storage (tapes) also contain the corresponding daily values.

The meteorological data stored in the active storage are limited currently to the monthly precipitation and temperature time series recorded at the gauging stations. The passive storage contains the corresponding daily values, and also radiation records for Southern Ontario. Other meteorological data, or data regarding other arcs of the hydrological cycle have not been stored so far since they are too scarce to be of any use in the interpolation of hydrological information.

The physiographical data stored at present consist basically of the following records:

Elevation in each grid knot;*

Percentage of square covered by forest, lake marsh, barren land, agricultural land, urbanized land*

---

* From topographical maps 1:250,000, 1953, edition 2, produced by the Army Survey Establishment, R.C.E. Copies available at the Map Distribution Office, Department of Mines & Technical Surveys, Ottawa.

An index of soil permeability (in Southern Ontario only);*

The data stored are screened using various statistical techniques (particularly the coefficient of variation and limits, graphical and analytical correlations, etc.).

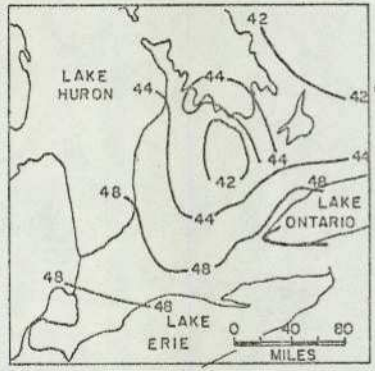
The obtaining of physiographical data from maps was initially a manual operation. At present this operation is carried out by means of digitizers. Further research aimed toward obtaining digital maps from the original survey data and programmes simulating the operations through which the mapping is carried out or possible improvements of these operations are comtemplated. Computer digital mapping of meteorological data (precipitation and temperature) is currently carried out successfully using the data bank as shown by Solomon and Qureshi (1972), as illustrated in Fig. 3.3.

Component c) data processing, contains three groups of operations:

The first group includes the computation of "derived" physiographical characteristics such as slopes, barrier heights, shield factors, etc. These "derived" physiographical characteristics are described in Appendix A. They can be computed readily

---

* From the map "Physiography of Southern Ontario" by Chapman and Putnam.



MEAN ANNUAL TEMPERATURE
DISTRIBUTION BY DEPT. OF TRANSPORT

GEORGIAN BAY
LAKE SIMCOE
TORONTO
LAKE ONTARIO
NIAGARA FALLS
LAKE ST. CLAIR
LAKE ERIE

LONG TERM AVERAGE
TEMPERATURE (DEG. F. /10)

0	=	300	=	309
1	=	310	=	319
2	=	320	=	329
3	=	330	=	339
4	=	340	=	349
5	=	350	=	359
6	=	360	=	369
7	=	370	=	379
8	=	380	=	389
9	=	390	=	399
A	=	400	=	409
B	=	410	=	419
C	=	420	=	429
D	=	430	=	439
E	=	440	=	449
F	=	450	=	459
G	=	460	=	469
H	=	470	=	479
I	=	480	=	489
J	=	490	=	499
K	=	500	=	509
L	=	510	=	519
M	=	520	=	529
N	=	530	=	539
P	=	540	=	549
Q	=	550	=	559
R	=	560	=	569
S	=	570	=	579
T	=	580	=	589
U	=	590	=	599
V	=	600	=	609
W	=	610	=	619
X	=	620	=	629
Y	=	630	=	639
Z	=	640	=	649

Fig. 3-3
SQUARE GRID ESTIMATION OF DISTRIBUTION OF MEAN ANNUAL TEMPERATURE IN SOUTHERN ONTARIO

for each square from the topographical data stored, and their computation is fully automated.

The second group of processing operations includes the computation of the statistics of the distributions of the physiographical characteristics of various river basins. It is to be noted that these computations can be made automatically for any river basin at any point of the main stem or of the tributaries.

The third group of processing operations is concerned with the statistical analysis of the hydrometric and meteorological time series. As a result of such operations one can obtain long term, annual, monthly and daily means and other statistics of the variables of interest (flow, precipitation, temperature); cross and serial correlation coefficients for various lags, statistics of the random components of the time series, etc.

Component d), techniques for interpolation of information based on the data bank are discussed in Chapter 4. Other applications of the data bank are discussed in Chapter 7.

4. ERRORS IN HYDROLOGICAL INFORMATION AS RELATED TO NETWORK PLANNING.

A complete analysis of the problem of errors in hydrological information is an extremely complicated task, and as indicated by the very long list of supplementary references on this subject, even the review of published literature in this field would require a whole volume. In view of this, the writer decided to confine this chapter to the analysis of errors in hydrological information as related to network planning, i.e. to that part of the analysis which is required for the development of the techniques of hydrological network planning, in particular, of its natural regime section.

After carrying out this analysis, the writer arrived at two major conclusions related to errors in hydrological information. These are advanced here because they help in the understanding of the subsequent sections. The first conclusion is that it is necessary to divide the errors in hydrological information into two broad classes : the first one includes errors which relate to planning decision variables (e.g. number of stations or density of gauging stations used for information interpolation, duration of sampling, frequency of measurement and type of equipment) and a second which relates to design and operation decision variables (e.g. actual location of gauging stations to obtain maximum stability of stage-discharge relationship, timing of flow measurement, and staff and equipment allocation to obtain

as wide a range of measurement as possible, number and duration of velocity measurements, etc.). Although consideration must be given in the network planning process to both types of errors, one cannot include in the planning process a complete error analysis which relates to the second category, both because of lack of data and of computational complexity. The influence of the second type of errors has to be assumed, in this case to be more or less constant and estimated from available data, experience, and suggested changes in equipment, personnel, station location, etc. When design is completed and further experience from operation of the revised network is obtained, a revision of the estimated errors related to design and operation variables can be introduced and the planning conclusions adjusted to the newly developed estimates. This may require further adjustments in design and operation. However, it is not expected that significant revisions would result from such iteration. It is more likely that technological advances in fields related to network planning, such as changes in methods of information interpolation, or in the general technology of geophysical surveying are likely to produce revisions of the decision variables involved in planning of much larger significance than those brought about by the revision of the initial estimates of the effect on errors of factors related to design and operation. This should be carefully considered in areas where new technology is introduced on an experimental basis.

The second conclusion is that analytical expressions of relations between errors and decision variables

in planning represent only indications of a qualitative nature and that these could be used only as guides to the structure of empirical relationships which should be used in the planning of the hydrological network. This is related to the fact that analytical derivation of such relationships is based on a set of assumptions about the statistical characteristics of hydrological data which do not correspond to their actual characteristics. As was shown by Wallis and Matalas (1972), even when one develops a complex analytical solution, it is impossible to solve such relationships analytically, and the best one can do is to simulate large numbers of data if actual observations are not available in order to derive the required practical relationships. Since simulated data have built-in characteristics that do not correspond entirely to those of actual data, even such an approach may be doubtful. The writer prefers to base the network planning techniques obtained to the extent possible on recorded hydrological data. Since such data may not be available in some areas for which it is intended to plan a hydrological network, this may produce great difficulties in practical applications. However, the transfer of relations between errors and decision variables is acceptable in such applications and the problem is reduced therefore to establishing such relationships in areas with similar

physiography. This is possible in almost all areas of the world. Where this is not possible, the establishment of pilot study areas in conjunction with theoretical research could probably provide the basis for at least preliminary planning, to be adjusted after a number of years of operation, as the initially presumed relationships are corrected on the basis of the accumulated data.

4.1 Errors of Hydrological Information Related to Planning Decision Variables

As will be shown in Section 4.2.1, there are several techniques for interpolation of hydrologic information which are based on very different hydrological and mathematical considerations. Although almost all of these techniques (except the isoline method) can be expressed in the form of some algorithms, it is practically impossible to subject them to a rigorous statistical analysis. However, since most of these techniques rely on least squares optimization procedures, they are similar to the multiple regression technique for which Matalas and Gilroy (1968) have derived the following expression for the estimated error variance V'_{tt} of a prediction at a site not used in the multiple regression analysis:

$$\hat{V}'_{tt} = (1+ \frac{1}{N_s} + \phi''_{tt})\sigma^2(\delta) + \{\rho' + \frac{1-\rho'}{N_s} + (1-\rho')\phi''_{tt}\}\sigma^2(\varepsilon). \qquad (4.1)$$

where N_s is the number of data used to develop the multiple regression between the hydrological characteristic and the independent variables considered;

ϕ''_{tt} is the (t,t) element of the matrix ϕ'';

$\phi'' = \phi'(\phi')^T$; and

$\phi' = X' A^{-1} X^T$, an $N' \times N_s$ matrix,

in which

N' is the number of ungauged sites for which estimates are made;

$X = |(x_{jk} - \bar{x}_k)|$ is an $N_s \times m$ matrix,

m being the number of independent variables (physiographical characteristics) used in the multiple regression;

x_{jk} is the k-th independent variable for the j-th gauged basin;

$$\bar{x}_k = \frac{\sum_{1}^{N} x_{jk}}{N_s} ; .$$

$X' = |(x'_{jk} - \bar{x}_k)|$ is an $N' \times m$ matrix;

x'_{jk} is the k-th independent variable in the j-th ungauged basin;

$A = X^T X$

A^{-1} = the inverse of A

$\sigma^2(\delta)$ is the average error variance due to space sampling error in a given regression,

$\sigma^2(\varepsilon)$ is the average error variance associated with the time sampling error at the stations used in the regression;

ρ' is the average correlation between the series of data z_{ij} used to define the given hydrological characteristic y_j (a mean, a variance) where z_{ij} is the i^{th} observation at station j (the cross-correlation coefficients).

Equation 4.1 is based on a series of assumptions in accordance with classical regression analysis namely:

(1) For any set $X_m = | x_1 \ldots\ldots, x_m |$ the variance of y is homoscedastic i.e. the variance is the same,

(2) For a given set X_m the associated y's are normally and independently distributed with mean, $\bar{E}$

$$\bar{E}\{ y|X_m \} = \alpha + \sum_{k=1}^{m} \{\beta_k (x_k - \bar{x}_k) \},$$

where

α is the multiple regression constant and β_k are the multiple regression coefficients, and variance $V\{ \delta + \varepsilon | X_m \} = \sigma^2(\delta) + \sigma^2(\varepsilon)$.

(3) The elements of X_m are fixed (error-free variables).

As indicated in the introductory discussion to this section, none of the above assumptions holds for the hydrological variables which form the object of our analysis. However, Eq. (4.1) gives some insight in the causes of increased error variance at an ungauged site, which are related particularly to the differences between the physiographical characteristics (independent variables) of the ungauged basin considered and the average characteristics of the sample used to establish the correlation. This indication is important and will be considered in the delineation of model-homogenous hydrological regions (Section 7.1) The further implication of Eq. 4.1 is that the total error variance is smaller when the cross correlation is smaller, an indication first that dense networks are inefficient, and that sampling be should be done in such manner that cross correlation avoided as much as possible. This shows that from a planning viewpoint it is interesting to obtain a decrease of the cross correlation by stratified sampling, that is by installing gauges in basins having distinct hydrological regimes. This can be arrived at by means of delineation of physiographically homogenous regions (Section 7.2).

If the error variance due to interpolation is determined by split sampling, the expected value of

$$(1 + \frac{1}{N_s} + \bar{\varphi}''_{tt}) \, \sigma^2(\delta)$$

can be estimated by $\rho^{*2}(\delta)$, the error variance determined from the validation portion of the sample.

Numerous computations of the value of ϕ''_{tt} made on the basis of sets of physiographical characteristics obtained within "model-homogeneous regions" as defined in Sub-section 7.1.1 show that ϕ''_{tt} is proportional to $\frac{1}{N_s}$ and has values which oscillate around an average* of $\frac{1.5}{N_s}$. Consequently:

$$(1 - \rho') \left(\frac{1}{N_s} + \phi''_{tt}\right) \simeq \frac{2.5\ (1-\rho')}{N_s}$$

Furthermore, in the variance component $\sigma^2(\varepsilon)$ we shall consider, in addition to the error variance due to time-sampling error $\{\delta^2(t)\}$, also the error of measurement $\{\sigma^2(m)\}$. Thus, for our purposes, Eq. 4.1 will be used as:**

$$\hat{V}_{tt} = \sigma^{*2}(\delta) + \left\{\rho' + \frac{2.5(1-\rho')}{N_s}\right\} \{\sigma^2(t) + \sigma^2(m)\}$$

Starting from the generally accepted assumption that the error variance of estimate of an expected value is proportional to the reciprocal of the number of independent measurements, we introduce the following notations:

$$N_s \sigma^{*2}(\delta) = E^{*2}_s$$

$$N_t \sigma^2(t) = E^{*2}_t$$

$$N_m \sigma^2(m) = E^{*2}_m$$

in which

---

* Hardison and Gilroy (personal communication) indicate that, on the basis of the assumption that the physiographical characteristics are normally distributed, the expected value of ϕ''_{tt} is $\frac{1.33}{N_s}$. Because of the skewness of the actual distributions of the physiographical characteristics, the coefficient of proportionality is closer to 1.5.

** $\sigma(t)$ and $\sigma(m)$ are assumed to be uncorrelated.

E_s^* is an error variance parameter characterizing the error due to information transfer;

E_t^* is an error variance parameter characterizing the error due to limited sampling in time;

N_t is the average number of independent* observations in time carried out at the hydrological stations;

E_m^* is an error variance parameter characterizing the error due to the limited number of flow measurements;

N_m is the average number of independent** flow measurements carried out at the hydrological stations.

With the above notations, Eq. 4.2 can be written as:

$$E^2 = \frac{E_s^{*2}}{N_s} + \{\frac{E_t^{*2}}{N_t} + \frac{E_m^{*2}}{N_m}\} \{\rho' + \frac{2.5(1-\rho')}{N_s}\} \qquad (4.3)$$

(in which for the sake of simplicity we replace the symbol V_{tt} by E). The values of all the error variance parameters depend on the hydrological characteristic considered, the network configuration, and the hydrology of the area. In addition, E_s^* depends on the method of information transfer considered, and E_m^* on the operation policy of the network and on the stability of the measurement sections. In view of the specific character of the values of the error variance parameters it was suggested by the writer that, for the purpose of hydrological network planning, these parameters should be determined empirically by split sampling techniques.

It should be noted that, in the discussion which follows, the error variance parameters will be considered in terms of "relative error variance parameters" as the absolute error

---

* See Sub-section 4.2.2 for the meaning of independent observations in this case and on the related problem of auto-correlation.

** See Sub-section 4.2.3 for further discussion of the problems related to the error of measurement.

obtained from the empirical calculations is always divided by the actual value of the variable considered. For the sake of brevity, however, we shall generally use the term, "error variance parameter". To facilitate the interpretation of the relative error variance parameters, the following quite realistic example is given. Let us assume:

$$E_s^* = 3, \quad E_t^* = 0.3, \quad E_m^* = 1, \quad N_s = 100,$$

$$N_t = 20, \quad N_m = 200, \quad \text{and} \quad \rho' = 0.5.$$

Then Eq. (4.3) gives the following expected relative error:

$$E = \sqrt{\frac{9}{100} + (0.5 + \frac{1.25}{100})(\frac{0.09}{20} + \frac{1}{200})}$$

$$E = 0.308 \quad \text{or} \quad 30.8\% \ .$$

4.2 Estimation of Error Variance Parameters

As was mentioned in the previous section, the method suggested for estimating the error variance parameters E_s^*, E_t^*, and E_m^* is based on the application of a split sampling technique and consists of the following:

The sample of data available is divided into two parts: A calibration sample, which is used to derive the given model for information transfer; a set of validation samples, which is used to estimate the expected value of the error variance, and from this

the variance parameter. The data on dependent and independent variables used for validation as well as for calibration are assumed to be error-free, the only error being related to the model. Of course, this assumption is very far from being true. However, any consideration of errors in the data would make the problem intractable from a practical viewpoint. As will be shown in Sections 4.2.2 and 4.2.3, some allowance is made in the estimation of the error variance paramaters to account to some extent for such simplifying assumptions.

4.2.1 Error Due to Information Transfer in Space.

Values of E^*_s for various techniques of interpolating hydrologic information were estimated using results obtained mainly from the writer's work (Solomon, 1972c), although occasionally results of work carried out by others were also used. The work carried out by the writer has indicated that there are significant differences between the values of E^*_s obtained by using various techniques of information interpolation. In fact, these results indicate that differences in the technique of information interpolation represent in some cases the most significant factor to be considered in hydrological network planning.

It should be noted that in the analysis of E^*_s

one has to consider N_s corresponding to a certain reference area. In the following N_s is considered to be included in a reference area of 100,000 square miles. When an area A different from 100,000 square miles is considered, E_s has to be multiplied by

$$\Pi = \sqrt{\frac{A}{100,000}}$$

to obtain the equivalent variance parameter for the given area.

The various methods of information interpolation discussed further are presented in order of increasing sophistication. It should not be considered that the more sophisticated methods are necessarily preferred by the writer to the simpler ones, or that the most sophisticated method is presented as the end of a line beyond which further progress is not possible. On the contrary, the writer considers that each method may have its usefulness under certain circumstances, and can be improved upon or replaced by a better one. In the planning process the choice of one or another of the methods is determined by economic considerations which will be discused in Chapters 5 and 6, and by data availability. The fact that in most cases the more sophisticated methods lead to greater economy of the network operation as a whole than the simpler ones give an indication on the economic value of hydrological research.

a) Interpolation by maps.

One of the simplest and widely used methods of

interpolation of hydrological information is the use of isoline maps. The usual procedure is to trace the isolines on the basis of the information available at the gauging stations, but using as supporting information maps of related physiographic (topographic, meteorologic) information. An interesting example of the application of the method is that given by Coulson (1967). In his paper Coulson presents a map of the isolines of long-term average runoff in Southern Ontario using data from 75 hydrological gauging stations (N_s = 200 stations per 100,000 square miles) and an isohyetal map based on about 80 meteorological stations. The corresponding variance parameter is $E_s^* = 2.5$

A more sophisticated approach is used by Nordenson (1968) who applies an iterative approach for the co-ordinated preparation of precipitation, evaporation, and runoff maps. In fact, the iterative approach is applied only in humid areas where a map of precipitation is prepared first on the basis of precipitation records and topography, then a map of actual evaporation is estimated using basically the assumption that actual evaporation = precipitation - runoff, plus data on topography and temperature. Where the map of actual evaporation thus derived shows inconsistencies, the map of precipitation is adjusted up to the point at which these inconsistencies are eliminated. A map of runoff is derived subsequently as the difference between precipitation and actual evaporation values. The runoff map

is then compared with the recorded runoff. The observed discrepancies are then corrected by again adjusting the precipitation and actual evaporation maps. In arid areas, where according to Nordenson actual evaporation and runoff depend basically on precipitation and soil characteristics, and runoff in addition on the areal distribution of the precipitation (storm characteristics), Nordenson suggests a procedure based on the estimation of a relationship between storm rainfall and soil characteristics on one hand and direct runoff on the other, of the type developed earlier by the U.S. Soil Conservation Service (1971). Unfortunately, Nordenson does not give any indications on the statistics of the error of estimate and therefore it is not possible to estimate the variance parameter E^*_s. The technique suggested by Nordenson for humid areas has the merit of recognizing that isohyetal maps can be used for improving runoff maps and vice versa. His techniques involve a great deal of subjective decisions, and consequently require an experienced hydrologist to obtain good results.

The Soviet literature abounds in examples of applications of isolines as a technique for interpolating hydrological characteristics (see for example Uryvaev, 1965). The Soviet literature available to the writer does not contain data on the statistics of the error of information interpolation of the estimates made by means of isolines. Most Soviet hydrologists recognise that the isoline technique is

applicable only for interpolation of hydrological acteristics of large basins (of the order of several thousands of square miles), and that the estimates interpolated for smaller basins have to be corrected using coefficients which take into account the differences between the physiographical characteristics of the respective basin and those of the region in which the basin is located. This leads naturally to the application of more rigorous techniques such as correlations between hydrological and physiographical characteristics* which is discussed next.

b) Correlation with physiographical characteristics.

Correlations between physiographical and hydrological characteristics have been used for many years, particularly by means of graphical techniques, but their analytical application has been expanded widely during the last decade owing to the availability of computers. A new dimension has been added to this technique as applied in hydrology by the introduction of computerized data banks (see Chapter 3), which makes possible the rapid inclusion computation of physiographical (particularly of terrain) characteristics.

The technique assumes basically that there is

---

* By physiographical characteristics we shall understand in this thesis, location, topographic, land-use and land cover, pedological, geological, hydromorphological, and meteorological characteristics. We shall designate as terrain characteristics all the above except the meteorological ones.

either a linear or log-linear relationship between the hydrological characteristic considered (the"dependent" variable) and the physiographical characteristics (the "independent" variables). The computational method used is usually the stepwise multiple regression technique in which the "independent" variables are entered in the relationship in decreasing order of the amount of the total variance of the dependent variable explained by the respective "independent" variable. If the ratio between the explained variance and the residual (unexplained) one is smaller than a given pre-set value for any of the "independent" variables not included yet in the relationship, the computation is stopped.

Correlations between hydrological and physiographical characteristics have been used widely in the U.S.A. and in Canada both for information interpolation and in the frame of hydrological network planning. However, there are significant differences in the practical application of the technique in the two countries as outlined further.

The application of the multiple regression technique in the U.S.A. has been described in detail by Thomas and Benson (1970). The hydrological characteristics used as dependent variables in the U.S. studies include low flows with various probabilities and durations, peak flows with various probabilities, flood volumes with various probabilities and durations, long-

term means of the annual and monthly flows, the standard deviations and coefficients of skew of the annual and monthly flows, and the serial correlations of the annual and monthly flows with a one-year lag. It is worth mentioning that in the U.S. studies the hydrological characteristics were expressed in terms of total basin discharge (cfs), that the dependent variables displayed a large variance in connection with the variance of the drainage areas considered,* and that the variance of the specific runoff (flow per square mile) represented only a very small fraction of the total variance. However, the variance of the specific runoff is of larger significance in information interpolation than the variance of the total basin discharge. Note also that the hydrological characteristics which are expressed in probabilistic terms are affected by a supplementary error in addition to the error of measurement and error of limited period of sampling, namely an error related to the discrepancy between the actual probability curve of the given variable and the type of probability curve used to estimate it.

The physiographical characteristics used in the U.S. studies can be grouped into four categories:

---

* In the Eastern region, one of the regions used in the study by Thomas and Benson to illustrate the application of their method, the drainage basins used varied from 11,560 sq. mi. to 5.9 sq. mi.

hydromorphological, comprising drainage area, main channel length, basin width, channel characteristics (main channel entrenchment, channel width, median size of stream-bed material, meander ratio defined as the ratio of the main channel length to the length of the valley), bifurcation ratio as defined by Horton (1932), mean flow distance defined as the average distance that a particle of water must travel to leave the drainage area as streamflow; topographical, comprising surface storage defined in terms of percentage of total drainage area covered by lakes, ponds and swamps, a soil index related to the potential maximum infiltration as determined by the U.S. Soil Conservation Service (19 71) and the area of alluvial soils; and meteorological, containing mean long-term precipitation, mean long-term snow pack, storm precipitation indexes, an evaporation index, and temperature indexes. The technique was illustrated through applications in four regions, but there is a significant variation in the way in which the technique is actually applied in these four regions, both from the viewpoint of the "independent" variables considered and of the methods used for estimating the "independent" variables. Thus, the evaporation and temperature indexes and the alluvial area are used as "independent" variables only in the Central region where the storage index is not used. Long-term annual precipitation was determined for two regions (Southern and Western) from isohyetal maps,

for one region (Central) from Thiessen polygons, and for one region (Eastern) partly from an isohyetal map and partly from a "precipitation anomaly map" derived for this region. The results of the study are summarized in Tables 6-9 of Thomas and Benson (1970) which gives the regression coefficients, standard errors of estimate (for the calibration sample) and the coefficients of correlation for each step of each correlation. The data required for the computation of the variance parameter E^*_s based on a validation sample are not presented in the paper, but the writer estimated from additional data obtained from the U.S.G.S. long-term flow data that for two of the four regions considered (the Central and Western) the variance parameter E^*_s is higher than that obtained for the isohyetal technique. This might be related to the particular conditions of the two regions. However, it is quite possible that the high values of the error variance parameters obtained are related to the way in which the technique was applied, and which could have had an adverse effect on the results. We refer particularly to the possible effect of multicollinearity between the independent variables (i.e. the existence of significant multiple correlation between one "independent variable" and several others , for example between drainage area on one hand and drainage elevation, precipitation, and snow on the other). Multicollinearity is probably responsible for the fact

that about two-thirds of the 182 correlations presented in the paper include only three independent variables or less. The drainage area is included in all correlations, the average precipitation in 128 correlations and snow in 55 correlations. Since drainage area is included in most cases simply as a proportionality factor and as such is an irrelevant variable, and precipitation and snow estimates are frequently affected by significant errors, the practical usefulness of the correlation developed for the interpolation of hydrological information may be in some doubt. The techniques used for the estimation of some of the "independent variables" also casts some doubt on the results of the studies. Thus, the mean basin elevation in the Central and Southern Regions was estimated as the mean of the elevations at the 10% and 85% points along the main stem channel. It is not surprising that the elevation determined in this way was not included in any of the correlations determined for the Central Region, and only in one of the 45 correlations determined for the Southern Region.

The technique described above was used widely in the U.S.A for assessing the adequacy of the existing hydrological networks, and for planning the extension or contraction of the network, in the manner described by Jordan and Hadman (1970). The assessment was made by developing multiple correlations as described above and

by comparing the standard error of estimate of these correlations with the standard error of estimate due to the sampling error resulting from the limitation of the record duration. The sampling error corresponding to a 10-year record was considered as the accuracy target for minor basins, and that corresponding to a 25-year record as the accuracy target for major basins. The network is considered adequate and its contraction might be considered if the standard error of estimate obtained from the multiple correlations is lower than the target error, and inadequate and requiring extension in the contrary case. The writer considers the basic philosophy and the practical application of the suggested approach to network planning as doubtful for the following reasons:

i) the target accuracy cannot be considered to be uniform independently of the regional conditions and priorities; in addition, it is obviously not simple to define in a general manner the difference between a minor and major river basin;

(ii) without the use of split sampling techniques it is not possible to provide realistic estimates of the errors of information interpolation since the assumptions made about the

statistical characteristics of the variables in the development of the correlations do not correspond to their actual characteristics, and therefore the standard error of estimate of the calibration sample will be different, in the general case, from the standard error of estimate of an independent sample;

(iii) the target accuracy test suggested does not take into account that the error of information interpolation is just one component of the total error of estimate; without considering the other components it is not possible to design and/or operate the hydrological network efficiently;

(iv) there is so large an amount of subjective decision making involved in the selection of the "independent variables" and of the techniques that it requires a very experienced hydrologist to estimate these variables;

(v) although it is obvious that the accuracy of the interpolated information is dependent on the accuracy of information obtained from other networks, the methodology suggested does not include any indications on the co-ordinated planning of hydrological, meteorological and other related networks.

A somewhat different approach to the problem of information interpolation by multiple regression technique was adopted in Canada. This approach was essentially developed by the writer. It uses as "independent variables" terrain characteristics estimated as described in Appendix A. The independent variables are screened for multicollinearity and only those which do not display a multiple-correlation coefficient significant at the 1% level are retained. The multiple correlations are developed using as independent variables only terrain characteristics because meteorological characteristics are affected by significant errors and since they are closely correlated to the terrain characteristics, cannot be used as "independent variables". The meteorological information is nevertheless used for the interpolation of hydrologic characteristics but in the frame of an iterative technique in which hydrological and meteorological information successively play the role of independent variables, as will be shown further.

In actual applications, because of time and budget limitations, the suite of dependent variables was limited to the long-term mean annual and mean monthly runoff, their standard deviations (coefficient of variation) and coefficients of skew. Calibration samples of various sizes were used in each case to obtain indications in the variation of the standard error of inter-

polation with the specific density (number in reference area) of the gauging stations. The parameter E_s^*, was established in each case from validation samples.

The application of the technique for interpolation of hydrological information requires the application of the model homogeneous hydrological regionalization method described in Section 7.1. After application of this method and computation of the relative error variance obtained in interpolation of the various independent variables for various sample sizes in Southern Ontario (which is completely included in one model homogeneous region) and Western and Northern Canada (where five model homogeneous regions were delineated), it was possible to estimate the values of E_s^* for various hydrological characteristics. These are summarized in Table 4-1.

The best graphical relationship which could be obtained between error and number of stations (station density) in the case of the multiple regression technique applied to the mean annual runoff is shown in Figure 4-1, from which, for the sake of clarity, the data points were omitted.*

c) Square grid simultaneous interpolation of average runoff and precipitation.

It was recognized in the early stages of the application of the square-grid data bank that it would be possible to improve the accuracy of interpolation

---

* Figure 4-1 including data points and similar figures for other hydrological characteristics can be obtained on request.

of the long-term mean annual runoff by using the meteorological data as an additional dependent variable, rather than considering them as part of the "independent" variables, because of the multicollinearity between the meteorological and terrain variables. Two different methods of using the meteorological data in this manner have been developed by the writer and applied so far*. The first method refers to the simultaneous interpolation of the areal distribution of long-term mean annual runoff and precipitation, and is described in this paragraph. The second method is concerned with the synthesis of time series of monthly runoff (flow) data and is discussed under (d) below.

The method described here for the definition of the distribution of mean annual runoff using a data bank of terrain characteristics and the available meteorological data was presented initially by Solomon et al (1968) and was extensively applied by the writer, and also by others (Cuthbert, 1970, Ingledow, 1970, Obdekoff, 1970, Foyster, 1973) in various hydrological studies.

The basic assumption required for the application of the method is that in the area investigated the losses to groundwater from precipitation or runoff

---

* A third method which is a further refinement of the method described in this subsection is being currently developed by G. Cavadias.

TABLE 4-1

VALUES OF E_s^* for the Square Grid - Multiple Regression Technique of Hydrological Information Interpolation

	Long Term Mean Annual Runoff	Coefficient of Variation of Runoff
Annual flow	1.5	1.3
January	3.2	1.6
February	2.5	1.9
March	2.6	1.7
April	1.3	1.3
May	1.3	1.2
June	3.3	1.7
July	3.5	1.3
August	3.7	1.8
September	3.3	1.6
October	3.3	1.6
November	3.2	1.6
December	3.3	1.6

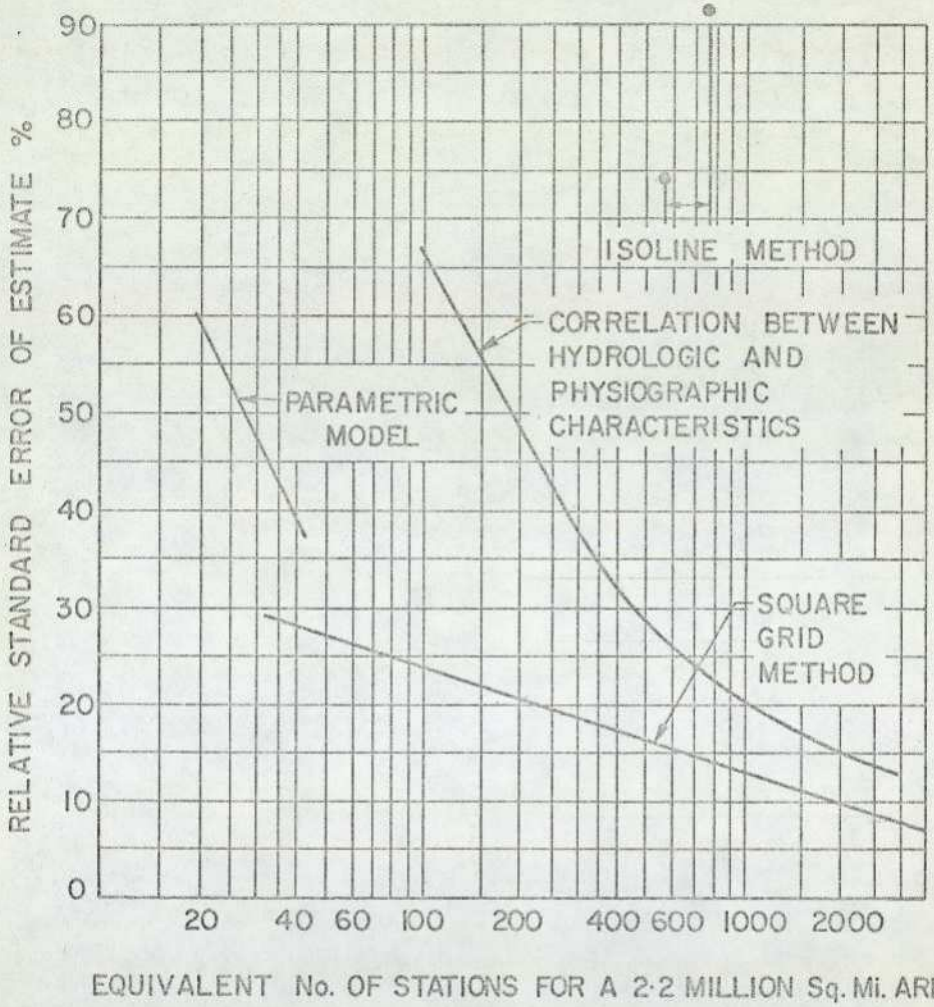


Figure 4-1 – RELATIONSHIP BETWEEN STANDARD ERROR OF ESTIMATE OF AVERAGE RUNOFF AND HYDROMETRIC NETWORK DENSITY FOR VARIOUS DATA TRANSFER METHODS.

(Data points not shown to preserve the clarity of the graph; figures including data points available on request).

are either negligible compared to the surface runoff or can be estimated from data or techniques independent of the one described here. The method consists basically of an iterative procedure in which estimates of the distribution of precipitation made initially without considering runoff data are successively improved by considering the latter data. In the first step of the technique one computes two correlations; one between precipitation as dependent variables and those terrain characteristics which are considered relevant (location, topographic, and topographic location, See Appendix A) as independent variables, the other between temperature as dependent variable and the same independent variables. On the basis of these two correlations one estimates in each square the long-term mean annual precipitation and temperature, and, from them, the values of the long-term mean annual evaporation by means of Turc's formula, (Turc, 1954)* or of Langbein's graph (Langbein, 1948) for estimating mean annual actual evaporation from precipitation and temperature data. Using the results of the above computations, one estimates in a square i the long-term mean annual runoff as:

---

* A correction of the temperature value to be used in Turc's formula has to be introduced when this is applied in climatic conditions with large temperature variations, such as those encountered in Canada. The method for computing this correction is described in Solomon, 1972b.

$$R_i = P_i - E_i - Gw_i \qquad (4.4)$$

in which P_i, E_i, and Gw_i, are respectively the long-term mean annual values of precipitation, actual evaporation, and losses to the groundwater in square i. A calibration sample of N basins is then selected, and for each basin j of this sample one computes the estimated runoff as:

$$R_{ej} = \frac{\Sigma R_i \; a_i}{\Sigma \; a_i} \qquad i = 1,2,\ldots b_j \qquad (4.5)$$

where a_i is the area of the portion of square i included in the basin j, and b_j is the number of squares included in basin j. The factor k_j is computed then for each basin using the formula:

$$k_j = \frac{R_{rj}}{R_{ej}}$$

in which R_{rj} is the long-term mean annual runoff recorded in basin j.

If $k \neq 1$, it is assumed that all the error is in the precipitation estimate and the precipitation is corrected in each square i of the basin j using the formula:

$$P_i^{(1)} = k_j^{(i)} (P_i - E_i) + E_i + Gw_i \qquad (4.6)$$

Where the index in brackets represents the number of the

iteration. The computation is then repeated starting from the correlation between precipitation and physiographic characteristics, but using this time, in addition to the precipitation records at the precipitation gauging stations, the estimated values of the precipitation $P_i^{(1)}$ in each of the squares of each of the b_j calibration basins. The iterations are repeated as long as necessary to reduce the expression

$$E = \frac{\sum_{1}^{N} \{k_j^{(m)} - 1\}^2}{N}$$

below a pre-established limit. After reaching this limit, a correlation is computed between the unexplained errors in the estimation of the runoff as the dependent variable and the terrain characteristics which were not used in the correlation between precipitation and temperature on one hand and terrain characteristics on the other, namely land-cover and land use, pedological, and geological characteristics. A final correlation between runoff and terrain characteristics is then obtained by combining the correlation between precipitation - actual evaporation - ground-water and related terrain characteristics with the correlation of unexplained error of runoff and the remaining terrain characteristics mentioned above. An example of the results obtained by applying the technique in Southern Ontario is given in Appendix B.

It is obvious that the coefficient of correlation and the standard error of estimate obtained in this case does not give any indication on the error of interpolation which will be obtained, since one can repeat the iteration process as long as desired to reduce the standard error of estimate as much as desired. For this technique particularly, only a check by means of an independent validation sample can indicate to what extent the runoff-terrain correlations derived as described above can be used for the purpose of interpolation of runoff values. In the case of the Southern Ontario application a 13 basin validation sample was used. The runoff for these 13 basins was computed using the correlation obtained between mean annual runoff and terrain characteristics shown in Appendix B. The results of the computation are given in Table 4.2. In the same table we show, for comparison purposes, the relative errors made in the estimation of the runoff from Coulson's isohyetal map (Coulson, 1967) for nine of the river basins which belong both to our validation sample and to the calibration sample used by Coulson. As can be seen from Table 4-2, the standard error of estimate for these nine basins is about 50% higher in the case of the isohyetal technique than in the case of our technique, although these stations belong to the calibration sample in the case of the isohyetal technique, and to the valid-

TABLE 4-2

COMPARISON BETWEEN ERROR OF ESTIMATE OF MEAN ANNUAL RUNOFF OBTAINED FROM AN ISOHYETAL MAP (CALIBRATION SAMPLE)* AND SQUARE GRID TECHNIQUE (VALIDATION SAMPLE).

River	Station	Drainage Area	Error(in %) = (Estimated - Recorded / Estimated X100) Isohyetal Map	Square Grid Technique
Saugeen	Walkerton	1,080	+1.0	-9.4
Maitland	Donnybrook	812	+6.2	-10.1
Grand	Galt	1,360	-	+8.2
Speed	Guelph	199	-13.6	-2.7
Grand	Brantford	2,010	-	+3.7
North Thames	Fanshawe Dam	560	-	+2.5
North Thames	London		-	-5.4
Credit	Erindale	262	-18.7	-1.3
Etobicoke	Summerville	41	+2.4	+0.7
Humber	Weston	204	-14.7	+10.5
Little Don	Lansing	41	-19.6	-6.1
Thames	Ealing	478	+6.5	+5.5
Big Otter	Vienna	256	+9.4	-1.9

* From Coulson, 1967, Table 4.

ation sample in the case of our technique.

The application of the technique to the various provinces and territories of Canada showed similar results, and a remarkable good relation was obtained between the relative standard error of estimate corresponding to the validation samples and the specific number of hydrological gauging stations. This relationship is shown in Fig. 4-1. The value of the error variance parameter for interpolation of mean runoff accepted by the writer for this method is $E_s^* = 0.75$. A comparison of this value and the graphical relationship given in Fig. 4.1, obtained on the basis of data from Northern and Western Canada and Southern Ontario, shows that the value of E_s^* is overestimated in this case for the lower network densities. We prefer to have an overestimation of this parameter in areas of low densities, since in practical applications this would provide a safety factor against unexpected difficulties in applying the technique in such areas.

d) Models

A large number of papers related to the application of hydrological models to the synthesis of time series of flow data on the basis of recorded or synthesised time series of meteorological data, or on the basis of purely statistical considerations has been published during the last few years. Some of these papers include also discussions on possible applications of such hydrological models to hydrological network design.

However, a review of all these papers would require the extension of this paragraph beyond reasonable limits, and we shall confine our discussion on this subject to one recent Ph.D Thesis by Barton (1971) which dwells specifically with the synthesis of monthly runoff records for ungauged basins and includes some considerations regarding hydrological network design. Other papers related to the subject and having some relevance to the problems debated in this subsection are listed in "Supplementary References".

Barton uses a combined deterministic-statistical model and correlates the model parameters with the pedological-geological and topographical characteristics of the basin. This enables him to interpolate the model parameters for ungauged basins. Then, by using synthetically developed precipitation data, he synthesizes time series of flow data for ungauged rivers. Barton's procedure, although more advanced than similar ones developed earlier by others, raises some problems from the viewpoint of its application to network design such as: the use of a subjective technique for optimizing the model parameters; the lack of possibility of synthesizing time series of flow data at several river basins which would display a cross-correlation pattern similar to the recorded one; and lack of clear indications on the relationship between specific number of hydrological and meteorological gauging stations and the standard error of estimate of the interpolated information. Because of these problems, it is questionable that the results of Barton's study could be used directly in hydrological network design.

The writer suggests that most of these problems encountered in the application of models for information interpolation and in related hydrological network planning can be solved much more easily if the methods are developed in conjunction with the square-grid data bank described in Chapter 3. Preliminary results of such applications were discussed by Solomon and Qureshi(1971) Since the publication of this paper the writer has carried out a series of supplementary studies on this subject for the purpose of this thesis which have contributed further to the analysis of the possibilities offered by the application of hydrological models to hydrological network design.

The study reported by Solomon and Qureshi was based on 34 river basin models developed in Western and Northern Canada in an area assumed to be extremely scarcely gauged from the hydrological viewpoint (15 hydrological stations per million square miles) and scarcely gauged from the meteorological viewpoint (430 meteorological stations per million square miles). The studies carried out by the writer for the purpose of this thesis are based on the development of 33 river basin models in an area assumed to have a moderately dense hydrological and meteorological network (130 hydrological stations and 300 meteorological stations per 100,000 square miles).

For both studies the deterministic-statistical model used was almost identical and was described by Solomon

and Qureshi. The use of this model does not imply in any way that it is considered "better" by the writer than any of the other models available. The experience of the writer and of others tends to indicate that almost all reasonable models give comparable results which depend mainly on the accuracy of the input and output data used for calibration, and that the capacity of the model to be used for interpolation of information decreases with the number of true (endogenous) model parameters. Because of this, the number of model parameters was limited in the applications discussed to 8. The model was designed to suit Canadian conditions, in the sense that it requires as input only precipitation and temperature data, the only meteorological data which can be reasonably interpolated at any point of the Canadian territory. Furthermore, to facilitate the use of the model in practical applications of hydrological network planning, it was designed in such manner as to make possible its application by any hydrologist without requiring subjective decisions of any sort.

As indicated above, the model is designed to accept as inputs mean monthly precipitation and temperature time series of data and to produce corresponding mean monthly runoff time series of data. The input data are estimated using the precipitation and temperature records at the meteorological stations and monthly correlations between these data and the terrain characteristics which are relevant from a meteorological viewpoint (location,

topographic, and combined topographic-location characteristics) in the squares corresponding to the meteorological stations. For each basin j and month i, the estimated precipitation P_{eij} is computed as the weighted average of the precipitation estimated from the correlation equation in each of the squares of the basin, corrected for local bias using the formula:

$$P_{cij} = P_{eij} + \frac{\sum_{1}^{m} \{P_{rkji} - P_{ekji}\}}{m}$$

where P_{cij} is the corrected precipitation for basin j and month i;

P_{rkji} is the recorded precipitation at station k, month i, station k being one of the m stations included in the rectangle circumscribed to the basin j; and P_{ekji} is the precipitation estimated from the correlation between precipitation and terrain characteristics in the month i, at station k.

For interpolation of hydrological information the models developed are transferred to ungauged sites by correlating the parameters of the models used in the calibration sample with the terrain characteristics. The correlations between the monthly precipitation and temperature data and the terrain characteristics makes it then possible to synthesize time series of precipitation and temperature in the ungauged basins and then, time series of runoff data by means of the transferred models. Because of the inherent correlation between the

meteorological data estimated at the ungauged sites, the synthesized flows at various stations in the areas display also cross correlations similar to those observed in nature. The meteorological input used may be based on recorded or synthetic data, but is introduced always in the same way, i.e. on the basis of correlations between the meteorological data and the terrain characteristics. The meteorological data synthesis is carried out by synthesizing time series of regression coefficients of correlations between precipitation and temperature as dependent variables and physiographic characteristics as independent ones.

In both the Northern and Western Canada and in the Ontario studies an error - network density analysis was carried out by successively reducing the size of the calibration sample. In all cases the error variance parameter was computed using validation samples. The error analysis was carried out for the long-term mean annual and mean monthly runoff and for their coefficients of variation. In addition, for the Southern Ontario study we have computed the error of estimate of serial and cross correlation of the monthly flows. The values of E_s^* obtained for the two regions are different and are shown in Table 4.3.

The reason for the differences obtained is related probably to the very different ratios between number of meteorological and hydrological stations used in the two regions. Whereas this ratio (considering the whole

TABLE 4.3

EXPECTED VALUES OF E_s^* OBTAINED FROM PARAMETRIC MODEL - SQUARE GRID TECHNIQUE

	Long-Term Mean Annual Runoff		Coefficient of Variation of Runoff	
	Northern & Western Canada	Southern Ontario	Northern & Western Canada	Southern Ontario
Annual Runoff	0.53	1.4	0.82	1.2
January "	1.60	3.1	1.72	1.5
February "	1.61	2.4	1.83	1.7
March "	1.08	2.6	1.29	1.7
April "	1.08	1.2	1.33	1.2
May "	0.43	2.1	0.69	1.3
June "	1.08	3.1	1.40	1.6
July "	0.54	3.3	0.75	1.4
August "	1.08	3.5	1.40	1.7
September "	1.51	3.1	1.94	1.5
October "	1.51	3.2	1.72	1.6
November "	1.72	3.2	2.04	1.6
December "	1.29	3.1	1.50	1.7

set of hydrologic stations) was 13 for Western and Northern Canada, it was only 2.3 for Southern Ontario. Because of the lower ratio in Southern Ontario, the amount of information introduced through the model in the hydrological time series is relatively lower in that area than the one in Western and Northern Canada. It is likely that a more rational technique for estimating E_s^* would have to consider in the case of models a value of N_s which would consist of the number of gauging stations augmented by an amount representing a function of the number of metorological stations in the area. Other techniques of interpolation of hydrological information would probably benefit also by considering this "equivalent hydrological value of the meteorological stations". Research in this direction in the framework of co-ordinated meteorological network planning is recommended. However, for the purpose of the practical application carried out in this thesis (Section 8), the values of E_s^* considered were those obtained for Western and Northern Canada since it was felt that they are closer to the average Canadian conditions, in areas of natural hydrological regime.

4.2.2. Error Due to a Limited Sampling Period.

A special study was carried out for the purpose of this thesis to develop a relationship for the estimation of errors due to a limited sampling period which would be suitable for planning purposes. The study had

as its initial point results obtained by others regarding the variance of the error of estimate of statistics of hydrological characteristics obtained from time series of auto-correlated data and time series of flow data extended by correlation with other time series of flow data obtained at long-term gauging stations (Wallis and Matalas, 1972). Because of the unavailability of an analytical solution for estimating the variance of the error of estimate of hydrological characteristics obtained from auto-correlated data augmented by multiple correlation, Wallis and Matalas used a Monte Carlo technique for obtaining such estimates.

As the technique used by Wallis and Matalas is based on several assumptions on the statistical characteristics of the time series of data which do not correspond entirely to the actual properties of the data, and as their work is not providing a complete analytical solution, the writer decided to use an empirical approach for the derivation of E_t^*, based on the following considerations:

It is practically impossible to determine the time sampling error because one does not have an "error-free" infinite length value to compare with in the general case of time series of hydrological data. Furthermore, the "effective number of years of record" is reduced and the error variance accordingly increased because of the effect of serial correlation (Yevjevich, 1972). In most cases in Canada, data for a period of record

longer than 50 years are not available. If one uses a 50-year record to obtain the "error-free" data, then one underestimates to a certain extent the error due to limited time sampling. Furthermore, the use of the coefficient of serial correlation to decrease the number of years of record is questionable, since this coefficient varies and frequently decreases as the period of record increases, and values obtained from available records would over-estimate the error values if the results are applied to longer periods. However, if one uses the actual number of years of record as the "effective number of years of record", this leads to another underestimation of the error variance.

In summary, if one elects to estimate the error variance parameter E_t^* assuming that the longest period of record, e.g. a 50 - year period, produces the error-free data (from the viewpoint of time sampling error) and uses the actual number of years of record as the "effective number of years of record", an underestimation of the actual value of E_t^* would be obtained. Let k denote the ratio between the square of the actual error variance parameter E_t^{*2} and the value estimated on the basis of the above mentioned assumptions $E_{t.50}^{*2}$. The value of k is larger than one and experience shows that it has most frequent values between 1.50 and 2.0. It was shown that in Equation 4.3 giving the total error variance in terms of its components, the error variance para-

meter E_t^{*2} is multiplied by a factor $\{\rho' + \frac{2.5\ (1-\rho')}{N_s}\}$. Since for the usual values of ρ' and N ($\rho' = 0.5 - 0.7$ N = 40 - 50), the value of this factor varies between 0.525 and 0.72, we can write:

$$E_t^{*2} \simeq (1.5 \div 2.0)\ E_{t,50}^{*2}$$

$$\{\rho' + \frac{2.5(1-\rho')}{N_s}\}\ E_t^{*2} \simeq (0.525 \div 0.72)\ E_t^{*2}$$

$$\{\rho' + \frac{2.5(1-\rho')}{N_s}\}\ E_t^{*2} \simeq E_{t,50}^{*2} \qquad (4.7)$$

On the basis of the above considerations, the writer decided that, for planning purposes, it would be acceptable to replace in Eq. 4.3 the term $\{\rho' + \frac{2.5(1-\rho')}{N}\}E_t^{*2}$ by $E_{t,50}^{*2}$, computed using the above stated assumptions. Consequently, the computation of $E_{t,50}^{*}$ was carried out as follows:

Six stations located in various regions of Canada and having a period of record of at least 50 years were selected and the hydrologic characteristics (long-term mean, long-term means of monthly flows, etc.) estimated, considering the most recent 50-year period. The assumption was made that a 50-year period of record represents the "true" values of the hydrologic statistics, and the statistics obtained from ten five-year

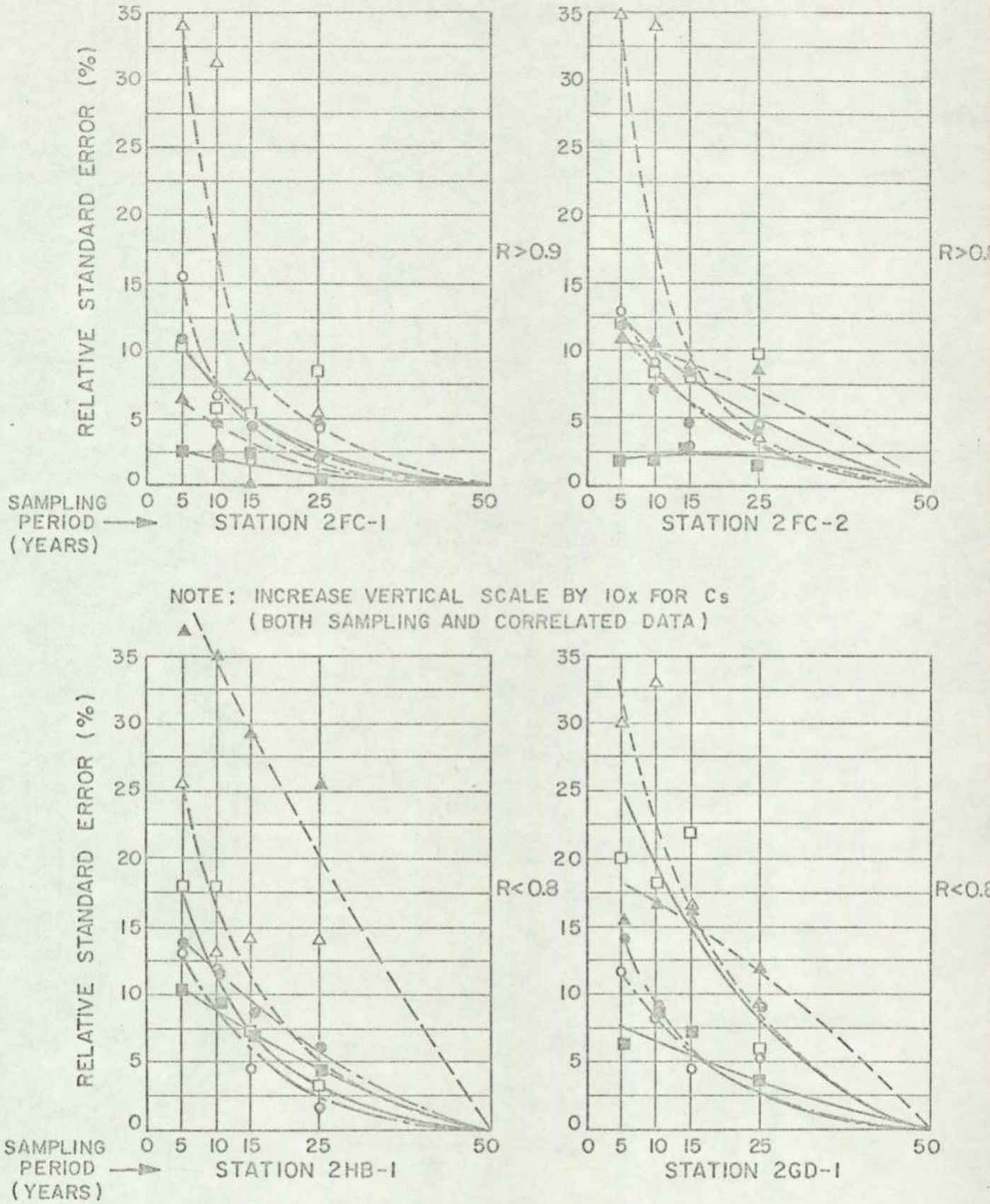


Figure 4-2

RELATIONSHIP BETWEEN RELATIVE STANDARD ERROR OF LONGTERM AVERAGE FLOW COEFFICIENT OF VARIATION AND COEFFICIENT OF SKEW OF ANNUAL FLOW VERSUS LENGTH OF SAMPLING PERIOD.

samples, five ten-year samples, etc., were compared to the "true" statistics and the corresponding relative standard errors. The expected values of the relative standard errors were then plotted against the length of the period of record as can be seen from the four series of graphs on Figure 4.2. There is a relatively smooth continuous reduction of the relative standard error, except in some cases, for the 25-year period for which the number of samples (two) makes the standard error computation inconclusive. A period of record of five years results in relative standard errors of estimate of the mean of 12-15%, while a period of 15 years reduces it generally to below 5%. The error in the determination of C_v is about two to three times larger than that of the mean flow, while the error in the determination of the coefficient of skew is about ten times that of the mean flow.

In order to obtain indications of the improvement (or otherwise) of the estimation of the flow statistics which may result from synthesizing a longer period of record by correlation with a neighbouring station, all the short period samples considered were completed to 50 years of monthly data by using a correlation with the closest station having a 50-year period of record. The flow statistics were then derived for each expanded sample, and the relative standard error computed. The results are also shown on Figure 4-2. It should be

noted that monthly log correlations were used according to the technique described by Langbein (1960). The correlations for synthesizing data at stations $2FC_1$ and $2FC_2$ were good (coefficients of correlation higher than 0.9), while the correlations for synthesizing data at stations $2HB_1$ and $2GD_1$ were poor (coefficient of correlation varying between 0.6 and 0.8. The improvements obtained in the estimation of the mean were most of the time insignificant and occasionally an error increase was observed at stations where the coefficients were poor, as expected also from the simulated analysis carried out by Wallis and Matalas (1972). It is noteworthy that for a period of 15 years of record at all stations investigated, after synthesizing the data by correlation, the errors of the means became larger than those obtained from the use of the actual record data only. However, significant improvements were generally obtained for the other statistics (the coefficients of variation and the coefficients of skew) with one significant exception regarding the coefficient of variation at one station, the one with the lowest coefficients of correlation where the error increased drastically.

Although these investigations are too limited to allow definite conclusions to be drawn, they seem to indicate that a period of record of about 15 years would produce the best return for short-term stations

TABLE 4.4

EXPECTED VALUES OF $E^*_{t,50}$ OBTAINED FROM DATA AT SIX GAUGING STATIONS IN CANADA

	Long-term mean	Coefficient of Variation
Annual Runoff	0.10	0.40
Spring "	0.36	0.45
Summer "	0.46	0.42
Fall "	0.33	0.60
Winter "	0.33	0.54
January "	0.31	0.65
February "	0.33	0.54
March "	0.31	0.60
April "	0.45	0.75
May "	0.48	0.27
June "	0.51	0.48
July "	0.48	0.40
August "	0.33	0.40
Sept. "	0.33	0.60
October "	0.40	0.96
November "	0.40	0.66
December "	0.35	0.60

in this area, and that information synthesis by correlation should concentrate on synthesis of other statistics than the average. It is unlikely that correlation coefficients less than 0.8 can be accepted for extending data for any of the statistics. In any case, further investigations, both theoretical and empirical, are warranted on this problem.

The expected values of $E^*_{t,50}$ obtained from data at 6 stations are given in Table 4.4.

4.2.3. Error of measurement

The error of measurement is assumed to consist of two parts, one related to the error of flow measurement (E_{mf}), the other (E_{mN_m}) related to the number of flow measurements available for obtaining the stage discharge relationship. Other sources of error (such as errors in stage records) can be ignored in the planning process since they are related mainly to operational problems, and since many sources of error of such type are currently eliminated by introducing automated and more reliable equipment. Assuming independence between the two sources of errors, one can write

$$E_m^2 = E_{mf}^2 + E_{mN_m}^2$$

The first component of the error of measurement was studied in three significant studies on errors of flow measurement that have been carried out during the last

few years by Dickinson (1967), Burkham and Dawdy (1969) and Goldstein (1968).

Dickinson's study is most comprehensive and his conclusions are significant mainly from the viewpoint of design and operation of the network. Most relevant from the viewpoint of network planning are the following conclusions:

"The magnitude of measurement error in a single discharge measurement and the degree of minor random shifting of a rating curve are reflected in the co-efficient of variation (relative standard error) from regression for the sample data. The average coefficients for the upper portion of the curve are +9.37 and -8.40 percent; and the average coefficients +17.3 and -10.9 percent on the lower portion.

Summer mean daily discharge values can be expected to be estimated within +2.40 and -2.33 percent error bounds, at the 80 percent confidence level; and within +3.82 and -3.65 percent, at the 95 percent level. Mean daily estimates in the winter months can be said to be within + 8.60 and -8.34 percent error bounds, at the 80 percent confidence level; and +13.7 and -13.1 percent, at a 95 percent level.

Future summer discharge measurements* may be

---

* These are equivalent to validation data in our terminology.

expected to lie within +12.9 and -11.2 percent of the rating curve values, 80 percent of the time, and within +21.5 and -16.9, 95 percent of the time. Single discharge measurements made in the winter months can be expected to fall within +47.6 and -39.6 percent of the curve, with 80 percent confidence; or within +76.9 and -59.6 percent, within 95 percent confidence.

If a single rating curve is employed to estimate all mean daily discharges during the summer, the percentage error bounds on the total summer flow, on each monthly flow, and on each mean daily value are equivalent

During the winter months, when daily flows are estimated independently, the percentage error bounds on a monthly flow estimate are a function of the percentage error bounds on a summer mean daily flow estimate, and the relative magnitude of the monthly estimate and the divisive discharge value.

The percentage error bounds on an annual discharge estimate are negligibly affected by the errors in winter flow estimates. They are approximately equal to the product of the ratio of summer to a total annual flow and the percentage error bounds on a mean daily discharge estimate in the summer months."

From a planning viewpoint, the important conclusion is that the error of an average does not depend in the case of summer flows*on the number of data averaged.

---

*Dickson understands by "summer" the ice free period of the year.

Dickinson's conclusion on the errors during the winter months is slightly confusing and difficult to apply practically because of the introduction of the concept of "divisive discharge value" (which is defined by Dickinson as the discharge arising from velocities less than one foot per second and/or depths less than one foot). In fact, most methods of determining daily flows in winter in Canada (see for example, Rosenberg and Pentland,1966) do not imply an independent estimation of each flow, and consequently in this case also the errors are strongly correlated and independent of the number of days used for the computation of the given statistics.

An element which has not been investigated by Dickinson, but was considered by the writer because of its significance for planning, was the error due to differences between the true rating curve and the rating curve obtained from a limited number of flow measurements. Although for particular cases this difference will depend on a large number of factors, such as the range of the measurements, the shape of the cross-section, the hydraulic conditions at the cross-section, etc., on the average the most important factor is probably the number of flow measurements. At any rate, the latter factor is the one which is most important from the viewpoint of planning the network, whereas the others are design and operation factors.

In view of this, a special investigation was carried

out to estimate the relationship between the error of estimate of a daily flow and the number of flow measurements used for the estimation of the stage-discharge relationship. The investigation was based on flow values determined at stations with well defined stage-discharge relationships which were compared with flow values obtained by using stage-discharge relationships determined on the basis of various samples of flow measurements drawn from the "population" of flow measurements available. This made it possible to determine the error of flow estimate as a function of the number of flow measurements.

The stage-discharge curves were computed using the same equation as that used by Dickinson,

$$Q = b_1 (H - H_o)^{b_2}$$

where Q is the discharge, H, the stage, H_o the stage at zero discharge, b_1 and b_2 parameters; b_1 and b_2 were determined by trial and error, using an optimization technique (the H_o selected was that for which the coefficinent of correlation of the regression of log Q versus log (H-Ho) was a maximum).

The investigation was carried out at 5 stations where the conditions were appropriate. It was found that an equation relating the standard error of measurement E_{mN_m} and N_m, having the form:

$$E^2_{mN_m} = \frac{E^{*2}_{m} N_m}{N_m}$$

can be obtained in each case.

The values obtained for $E^{*}_{mN_m}$ were:

River and station	$E^{*}_{mN_m}$
Saugeen R. near Port Elgin	0.27
Otter Creek at Tilsonburg	0.92
Young Creek near Victoria	0.74
Thames R. at Byron	0.20
Gull R. at Norland	0.07

Given the large variation of the value of $E^{*}_{mN_m}$ obtained at the various stations, it is recommended that further research should be carried out on this subject. However, for the practical applications presented in this thesis, an average value $E^{*}_{mN_m} = 0.44$ was selected. A further simplification of the analysis was introduced based on the above quoted conclusions of Dickinson's work from which it would appear that the error of flow measurement is approximately equal to the error due to limited number of flow measurements for the usual number of flow measurements at a station (200 to 300). On this basis it was decided that in further analysis the term $\{\rho' + \frac{2.5\ (1-\rho')}{N_s}\} E^{*2}_{m'}$ which for most frequent values of ρ' and N_s can be estimated to be approximately equal to (0.525 to 0.72) $E^{*2}_{m'}$, be replaced by $\frac{E^{*2}_{mN_m}}{N_m}$, since

$$E^{*}_{m} = \sqrt{E^{2}_{mN_m} + E^{2}_{mf}} \approx 1.41\ E^{*}_{mN_m}$$

and (0.525 to 0.72) $E_m^2 \simeq E_{mN_m}^2 = \frac{E_{mN_m}^{*2}}{N_m}$

In summary, the analysis carried out in the last three sub-sections indicates that it is acceptable to estimate the total error variance at an ungauged station from the following relationship:

$$E^2 = \frac{E_s^{*2}}{N_s} + \frac{E_{t,50}^{*2}}{N_t} + \frac{E_{mN_m}^{*2}}{N_m} \qquad (4.8)$$

with the variance parameters estimated as shown in sub-sections 4.2.1, 4.2.2. and 4.2.3.

It should be emphasized that this formula represents an estimation of the expected value of the error variance acceptable for application for planning purposes and not for individual estimates of the error at a given site. In this latter case, one has to consider a formula of the type suggested by Matalas and Gilroy (1968), which takes into account the deviation of the characteristics of the given basins from the average characteristics of the sample used to calibrate the model.

5. COSTS

Planning of the hydrological network is carried out either in the framework of a budgetary constraint if the purpose of the network is for inventory and research, or for maximizing returns in the case when the data from the network are used for the design and operation of projects. In both cases it is essential to have data on the costs of installing and operating the network and on the processing of data, and for the development and application of models for interpolation of information.

In discussing the problem of costs one has to keep in mind that these costs can vary rapidly in all countries of the world for two basic reasons: cost escalation and changes in technology. We suggest that in network planning analyses cost escalation can be neglected as long as there are no reasons to believe that the cost escalation for the given activity has a different rate from the average, since the average cost escalation represents basically only an adjustment of numerical values to a variable (but common) unit of measurement. Changes in technology, however, may significantly affect the costs by continuously reducing the ratio between labour and investment costs. Specific labour costs have the tendency to rise much faster than specific investment costs. In areas of fast technological progress, the ratio labour costs/investment costs may in fact

decrease significantly. One can, therefore, expect that in areas where significant technological progress is currently being made, unit costs will decrease as the technological advances are implemented. Since such progress is being made in the area of the hydrological network due to the application of computers and data bank techniques, remote sensing, modelling, etc., one can reasonably expect that specific costs in this area will be at least stable within the next decade, corresponding in fact to a relative decrease in these costs.

We shall discuss in this Chapter basically the Canadian costs on which the writer has some information (Section 5.1). A brief discussion on problems of costs of hydrological networks in other countries is given in Section 5.2. This discussion is of a very general nature only since, due to the more or less confidential nature of this type of costs we did not have any factual data available for a quantitative analysis.

5.1 Canadian Costs.

There are several types of hydrological networks in Canada: Federal, Provincial, networks of the hydro-electric companies, etc. Specific costs of various components of these networks vary widely in relation to the size of the network, the natural conditions

access, and occasionally to the level of technology, etc. In addition, there are problems related to the accounting system used, some of the costs being reported inclusive of overhead expenditures, some others excluding such costs. The costs as estimated in this study represent costs incurred by the Federal network exclusive of overhead resulting from the activity of the central offices in Ottawa.

5.1.1 Installation Costs.

Hydrometric stations in Canada usually consist of a simple structure which supports a staff gauge, in the case of a non-recording station*, or a concrete well connected with the river which protects the level-recorder in the case of recording stations. More recently, with the introduction of pressure gauges, a new type of structure which is slightly less expensive than the classical recorder well, has been put in use.

Because of the scarce population in most areas and of the relatively low cost and higher accuracy obtained from recording devices, the trend in Canada

---

* Occasionally, where possible, the staff gauge used to be mounted on an existing structure, often a bridge pier or abutment. Since in many cases the bridge interferes with the natural conditions of the river, this practice has been discouraged in recent years.

is to install recording gauges at all sites independently of the size of the river basin. In order to obtain a stable stage-discharge relationship, where possible, the gauging stations are installed immediately upstream of natural control sections which, in Canada, are the head sections of rapid reaches. When no control is available, an attempt is usually made to solve the problem of instability of the stage-discharge relationship, which is usually related to sedimentation-erosion, ice, and vegetation, by means of very frequent discharge measurements. The construction of artificial controls, which is practised widely in some areas of the world, is not accepted practice in Canada, where such controls are built only for small experimental and research basins. However, as is indicated by simple calculations, for small rivers the construction of artificial controls might be a more economical solution than frequent measurements. Even if one assumes that the cost of a flow measurement for a small river is as low as $200 (in actual fact with the cost of processing, etc., it is much higher) and that one has to carry out about 25 flow measurements per year over and above those which would be carried out in the case of a stable control to obtain a comparable accuracy, one comes to the conclusion that costs of artificial controls even as high as 20-30 thousand dollars per structure would

be acceptable if the maintenance costs do not exceed 2% per year. However, such control structures could be considered only for river basins with drainage areas of up to 200 square miles since for larger basins the costs of the structures would exceed the economical limit mentioned. It is true that such structures may produce some disturbance of the natural sediment and goundwater regime, but with proper design such disturbances could be reduced significantly.

Whenever possible the flow measurements are carried out without special installations, by wading, walking over the ice or, at higher levels, from the nearest bridge. When this is not possible, a cable-way or a similar installation has to be included in the gauging station and this considerably increases the costs.

Installation costs are different in southern and northern regions because of differences in climatic and foundation conditions, and in the difficulty of access, transporting materials, etc. On the average, one can consider in preliminary calculations the following installation costs for stations provided with recorders (Analog recorders are cheaper than digitizing recorders, but the trend is for the latter to become cheaper).

	Northern Regions (North of the 55° parallel)	Southern Regions (South of the 55° parallel)
	(dollars)	(dollars)
Without cable-way or other installation for crossing the river	2,500	1,500
With cable-way or with special installation for crossing the river	4,000	2,500

5.1.2 Operation and Maintenance.

Costs of operation and maintenance, including the costs of equipment for flow measurement, vary also with the climatic and accessibility conditions, the stability of the station, the size of the river, etc. The primary problem in operation and maintenance is related to the necessity of carrying out flow measurements more frequently during peaks which occur, particularly in spring, simultaneously over large areas. Currently the main transportation mode for the flow measurement teams is the airplane, particularly in northern areas, large costs being

involved in the waiting time during the measurement. A drastic change in flow measurement technology, involving the use of remote sensing techniques which would make it possible to carry out the operation mainly from the air will probably significantly increase the efficiency of operation of the network. Currently the average cost of operation and maintenance of a station is of the order of $1,500 per year for a station located in the southern portion of Canada and of $4,000 per year for stations located in the northern area.

One may expect that with the increase in the density of the network which occurs steadily in relation to economic development, the specific operation and maintenance costs will also decrease, or at least will not increase.

5.1.3 Data Processing.

Data processing is currently done partly manually and partly by computer. Basically, the computation of flow measurements and the tracing of stage discharge relationships is done manually. In some districts the computation of daily flows on the basis of stage recorder log-sheets and stage discharge relationships is done manually, whereas in others some of these operations are done by computer. Computation of monthly and yearly means and preparation for printing is done by computer. In fact, the "Water

Resources Papers" which publish daily flow data, are copies of computer printouts of the data which are stored on magnetic tape as a result of the last stage of the processing of the data. Because part of the operations are carried out in the field and in the district offices and part in the central offices of the Water Survey of Canada, it is very difficult to make an accurate estimate of the costs involved in processing the data for one hydrometric station. A rough estimate of these costs is $500 per station per year. It is reasonable to expect a decrease of these costs during the next few years as a result of increasing automation of data processing, e.g. the introduction of digital recorders and digital chart readers.

5.1.4 Data and Information Interpolation.

Data and information interpolation is not done currently on a routine basis by the Canadian agencies operating hydrological networks. However, the Water Survey of Canada which operates the Federal network, is working toward integrating the data bank described in Section 3 in its system of operation with the purpose of using it for data and information interpolation. Experience accumulated so far in the various offices which have used the data bank for information interpolation indicates a large variation of specific costs.

For a small number of stations (less than 10) the cost of information interpolation by means of a parametric model based on monthly time intervals is approximately equal to the cost of interpolation by means of multiple regression techniques, and represents about 2,000 dollars per station for a period of 10 to 20 years.* The specific costs may decrease to about $600/station if the number of stations increases to one hundred or more because the parametric model does not present any supplementary complications when applied to a larger number of stations, and there is a significant element of economy of scale when the number of river basins for which it is applied increases. The costs and complications generated by programs for simultaneous data interpolation by multiple regression - Markov model techniques at a larger number of stations (over 10) appear to be very significant, and we did not have data available for estimating the related costs.

The costs involved in the use of lumped parametric models with daily intervals are about 3 times higher than those for a model with monthly intervals. The costs involved in applying the multiple regression – Markov-model technique with daily intervals are about

---

* Cost estimates refer to the synthesis of time series of data taking into account autocorrelation and cross correlation.

10 times higher than those involved in the application of the technique using monthly averaged data. Distributed models with small time and space intervals are still in the development stage and it is not yet possible to estimate their cost. Currently, these costs appear to be many times higher than the costs of simple, lumped models.*

All the above-quoted cost figures are based on the assumption that the programmes required are available and that all input data are stored in a comprehensive hydrometeorological land-use land-cover and physiographical data bank and can be retrieved readily on the basis of available programmes.

The available experience indicates that the cost of setting up such a data bank varies from 1 dollar per square mile for areas of the order of a few tens of thousands of square miles, to 15-20 cents per square mile for areas of two million square miles, or more. It is likely that the further development of remote sensing techniques and of digitizing equipment and the availability of user-oriented programmes for comprehensive, interfacing data banks will significantly reduce the costs of setting up such data banks and of using parametric models before the end of the current decade.

---

* The writer developed and applied a distributed model in 1974 (Solomon, 1974) for a basin of 110 sq. mi. The cost of running such a model is about $10,000 per year.

5.1.5 Other Costs.

These are mainly administrative (overhead) costs or costs of research and development. It is extremely hard to realistically separate such costs from the total budget of an organization operating a hydrological network. Although these costs probably vary slightly with increases and decreases in the number of gauging stations, it is more likely that their variation is related mainly with the extent and nature of the research and development activity. This in turn depends on the budgetary constraints, level of economic activity, and priority given by the administration to water resources problems. A very rough estimate of these costs for Canada is that they are of the order of 3 million dollars per year.

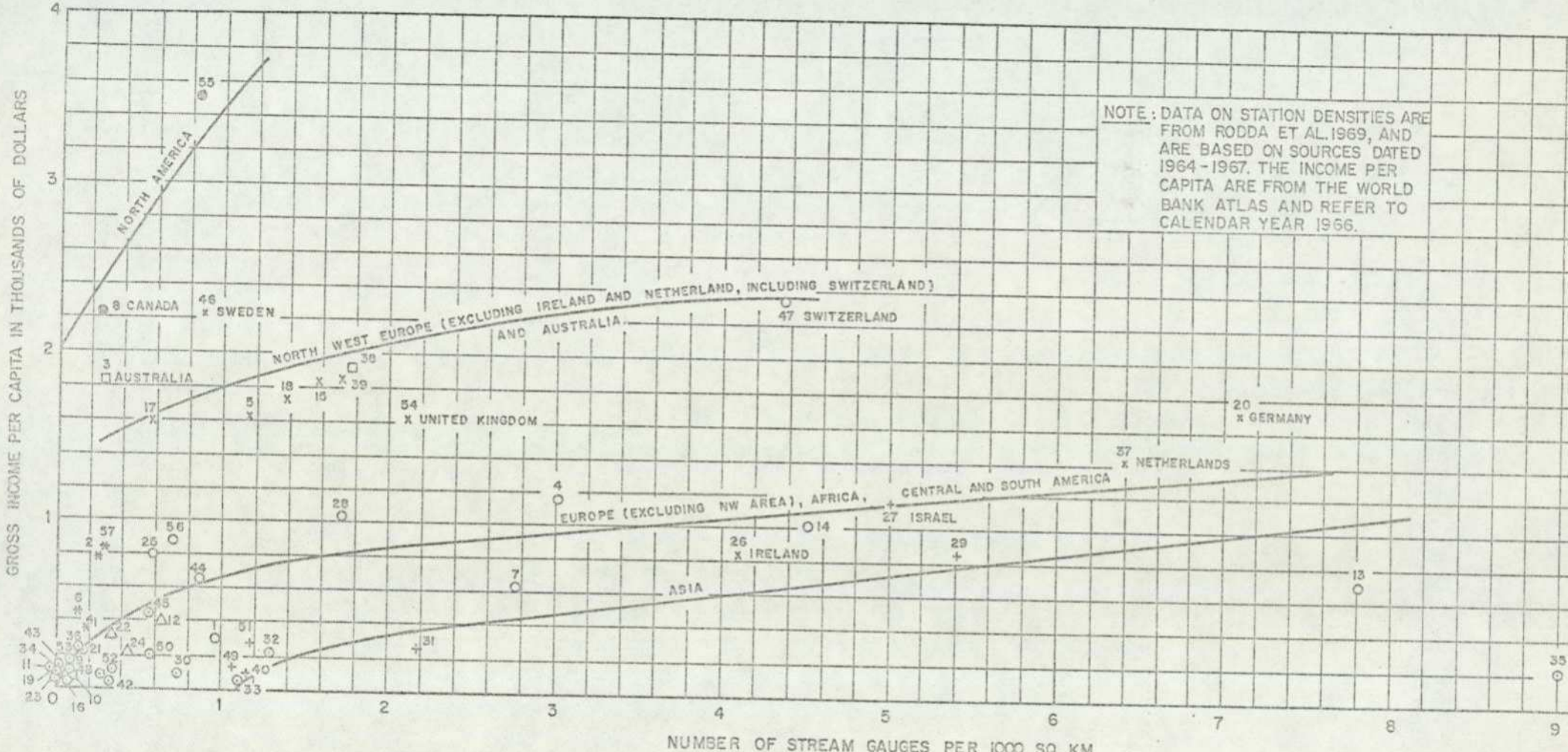
If the figures given above are of a correct order of magnitude, it appears that the annual costs of the Canadian Federal hydrological network are of the order of 10 million dollars per year. Consequently, efforts to improve the efficiency of this operation even by á few percent will be worthwhile in economic terms.

5.2 Costs in Other Countries.

Data on the number of hydrometric stations in various countries of the world can be obtained from

Rodda et al., 1969. A rough estimate of the number of hydrometric stations for other countries could be obtained from Fig. 5.1, taken from Shawinigan, 1970. This figure gives the relationship between gross income per capita and hydrological network density for various areas of the world.

As already mentioned, information on costs of hydrological networks were not available for the purpose of this thesis. Judging from the general economic and geographical-hydrological conditions, it can be inferred that specific costs in the U.S. are lower than in Canada, and in the U.K. and Western Europe much lower (probably 60-70% of the Canadian costs). However, due to the larger number of stations in the U.S. (where there are about 20,000 hydrometric stations), the annual operation and maintenance costs are of the order of one-hundred millions of dollars. Using the data on the number of hydrometric stations obtained from the above mentioned sources, it may be inferred that in the U.K. and Western Europe the annual costs of operating and maintaining the networks are also of the order of millions of dollars, or tens of millions of dollars in the larger countries such as the U.K. or France. In the U.S.S.R. and Eastern Europe the specific costs are probably lower than in Western Europe because of lower costs of labour. Nevertheless, the



1 ALBANIA
2 ARGENTINA
3 AUSTRALIA
4 AUSTRIA
5 BELGIUM
6 BRAZIL
7 BULGARIA
8 CANADA
9 CENTRAL AFRICAN REP.
10 CHAD
11 CONGO (D.R.)
12 COSTA RICA
13 CYPRUS
14 CZECHOSLOVAKIA
15 DENMARK
16 ETHIOPIA
17 FINLAND
18 FRANCE
19 GAMBIA
20 GERMANY (F.R.)
21 GHANA
22 GUATEMALA
23 GUINEA
24 HONDURAS
25 HUNGARY
26 IRELAND
27 ISRAEL
28 ITALY
29 JAPAN
30 KENYA
31 KOREA
32 LIBERIA
33 MALAWI
34 MAURITANIA
35 MAURITUS
36 MOROCCO
37 NETHERLANDS
38 NEW ZELAND
39 NORWAY
40 PAKISTAN
41 PERU
42 RWANDA
43 SENEGAL
44 SPAIN
45 SOUTH AFRICA
46 SWEDEN
47 SWITZERLAND
48 TANZANIA
49 THAILAND
50 TUNISIA
51 TURKEY
52 UGANDA
53 U.A.R.
54 UNITED KINGDOM
55 U.S.A.
56 U.S.S.R
57 VENEZUELA
58 YUGOSLAVIA
59 ZAMBIA

● NORTH AMERICA
✱ SOUTH AMERICA
△ CENTRAL AMERICA
× NORTH – WEST EUROPE
○ EUROPE (LESS N.W. EUROPE)
+ ASIA
⊙ AFRICA
□ AUSTRALASIA

Figure. 5-1 – RELATIONSHIP BETWEEN GROSS INCOME PER CAPITA AND DENSITY OF HYDROLOGICAL STATIONS from Shawinigan, 1970

total costs in these countries are probably of the same order of magnitude as in Western Europe. Specific costs in developing countries may still be lower than in Eastern Europe and the U.S.S.R. because of much lower labour costs. However, in many of these countries the existing networks are completely inadequate, and it is probably required to spend significant amounts of money to bring them to a satisfactory if not optimum level. As an example which should be mentioned, the Central American Countries have spent during the last 4 years about 7 million dollars on the expansion of their network (UN, 1970). All this seems to indicate that the world's total costs of operating, maintaining, and expanding hydrological networks represents many hundreds of millions, if not billions, of dollars and emphasizes again the economic significance of improving their efficiency.

6. OPTIMIZATION

The error analysis carried out in Chapter 4 indicates that there are basically two types of sources of error: sources related to network planning decision variables such as number of stations, duration of sampling, number of flow measurements, technique of information interpolation; and errors related to the detailed design and operation of the network, such as selection of station sites, distribution in time of the flow measurements, etc. The analysis carried out in Chapter 4 indicates that an attempt to include simultaneously both types of sources of error in an optimization analysis would be a practically impossible task because of the large number of relationships, unknowns, variables, and constraints. To make the problem manageable, a step-wise optimization is suggested in which the decision variables are divided into planning variables and design and operation variables. The planning variables are optimized first using assumed values for the design and operational variables and for their effect on the combined error. The results obtained are then used with the relationships between the errors and the design and operation variables to optimize the latter. If the results obtained are significantly different from the initially assumed

values, an iteration of the first stage is done, and the operation is continued until convergence is obtained.

As was shown in Chapter 4, some of the relationships between error and decision variables cannot be determined analytically and require the determination of empirical coefficients. Since these coefficients may be in turn affected by errors, it is necessary to carry out a sensitivity analysis to determine the sensitivity of the decision variables to the errors in these coefficients, so that research is directed toward the improvement of the definition of the more significant coefficients. As results from such research become available, adjustments to the values of the decision variables will be made as indicated by the consideration of these new results.

We shall discuss in this section the optimization of the planning variables, namely the number of stations, the duration of the sampling, number of flow measurements, and the technique of interpolation of information. As we have indicated in Chapter 2, we consider separately the optimization of the natural regime network in areas not designated for economic development during the planning interval (Section 6.1) and in areas designated for economic development (Section 6.2). Some of the problems of design of the network (particularly the problem of distribution in space of the stations) are discussed in Chapter 7.

6.1 Planning of the Natural-regime Network in Areas not Designated for Development

A tenative solution to this problem was presented by the writer in earlier papers and reports (Solomon, 1971, 1972c, d). The following treatment of the problem represents an improved version of the earlier solution.

From the discussion in Chapter 4, it was concluded that the relative error at an ungauged site is given by the following relationship (Eq. 4.7)

$$E^2 = \frac{E_s^{*2}}{N_s} + \frac{E_{t,50}^{*2}}{N_t} + \frac{E_{mN_m}^{*2}}{N_m}$$

The explanation of the symbols used in the above equation was given in Chapter 4.

6.1.1 Error Reduction

Assuming that the network consisting at present of N_{si} stations has been already operated for an average number of N_{ti} years, and that an average number N_{mi} of flow measurements was already carried out, the problem is to investigate the planning strategy for a supplementary number N_{ts} of years. During this period a percentage p of stations will be discontinued, but the total number of stations will be increased up to N_{sf}, and the total number of flow measurements per station up to N_{mf} by making n_{ms} flow measurements per year at each station during the period N_{ts}.

The error reduction can be estimated from the expression:

$$\Delta E = \sqrt{\frac{E^{*2}_{si}}{N_{si}} + \frac{E^{*2}_{t,50i}}{N_{ti}} + \frac{E^{*2}_{mN_m i}}{N_{mi}}} - \sqrt{\frac{E^{*2}_{sf}}{N_{sf}} + \frac{E^{*2}_{t,50f}}{N_{tf}} + \frac{E^{*2}_{mN_m f}}{N_{mf}}} \qquad (6.1)$$

where

$$N_{tf} = \frac{N_{si}\, N_{ti} + (N_{sf} - \frac{pN_{si}}{100}) N_{ts}}{N_{sf}} \qquad (6.2)$$

and

$$N_{mf} = \frac{N_{si}\, N_{mi} + (N_{sf} - \frac{pN_{si}}{100}) N_{ts} n_{ms}}{N_{sf}}. \qquad (6.3)$$

The values of E^{*}_{si}, $E^{*}_{t,50i}$ and $E^{*}_{mN_m i}$ corresponding to the current conditions can be readily estimated using calibration-validation techniques. The values of E^{*}_{sf}, $E^{*}_{t,50f}$ and $E_{mN_m f}$ for various possible changes in data processing and interpolation techniques have to be estimated together with the costs of such changes from research work carried out on methods of interpolation of hydrological information in areas with dense networks and at stations with frequent measurements.

6.1.2 Additional Costs

The additional cost ΔC for operating the network during the N_{ts} years will consist of four components:

$$\Delta C = C_0 + \Delta C_1 + \Delta C_2 + \Delta C_3 \qquad (6.4)$$

in which:

C_0 are annual costs which consist of a fixed portion (administration overhead) and a variable cost (which depends on the extent and nature of studies for information interpolation, stage-discharge extrapolation, etc.);

ΔC_1 is the cost of installing new stations;

$$\Delta C_1 = (N_{sf} - N_{si}) C_1 \tag{6.5}$$

where C_1 is the average cost of installing new stations (a simplification of the study requires the assumption that all new stations are installed at the beginning of the planning period[1]);

ΔC_2 is the cost of operating the network (excluding flow measurements and instrumentation).

$$\Delta C_2 = (N_{sf} - \frac{pN_{si}}{100}) N_{ts} \; C_2 (PWN_{ts}) \tag{6.6}$$

where C_2 is the cost of operating one station one year and PWN_{ts} is the average current worth of one dollar over a period of N_{ts} years.

$$\Delta C_3 = \Delta C_{03} + (N_{sf} - \frac{pN_{si}}{100}) N_{ts} (PWN_{ts}) C_3 \tag{6.7}$$

where ΔC_{03} being the cost of new instrumentation, and C_3 the cost of one flow measurement.

In the above equations N_{sf}, N_{ts}, n_{ms}, p, and C_0 represent the basic planning decision variables for

---

[1] As shown in Section 6.2 this assumption is not necessary if the dynamic programming technique is applied.

planning the network; E^*_{si}, $E^*_{t,50i}$, and $E^*_{mN_m i}$ represent the initial network characteristics in the area considered; E^*_{sf}, $E^*_{t,50f}$ and $E^*_{mN_m f}$ represent together with C_0 the technical-economic improvement potential in the planning and operation of the network and the availability of data from other networks; and PWN_{ts}, C_0, C_1, and C_2, reflect the prevailing technical-financial conditions.

6.1.3 Objectives and Optimization

The various objectives of the planning of the natural regime network can be expressed by the following relationships:

$$\frac{\Delta C}{\Delta E} \rightarrow \min \qquad \text{(a)}$$

if we wish to obtain a minimum cost per unit of error reduction without any further constraints;

$$\frac{\Delta C}{\Delta E} \rightarrow \min, \text{ with } N_{ts} \leq N'_{ts} \qquad \text{(b)}$$

if we wish to obtain a minimum cost per unit of error reduction within a given period of operation of the network N'_{ts};

$$\frac{\Delta C}{\Delta E} \rightarrow \min, \text{ with } \Delta E \geq \Delta E_i \qquad \text{(c)}$$

if we wish to obtain a minimum cost per unit of error reduction but for an error reduction larger than a given value ΔE_i;

$$\Delta E \rightarrow \max, \text{ with } N_{ts} \leq N'_{ts} \qquad (d)$$

if we wish to reduce the error as much as possible within a given number of years N'_{ts};

$$\Delta E \rightarrow \max, \text{ with } \Delta C \equiv \Delta C_o \qquad (e)$$

if we wish to reduce the error as much as possible within a given total budget;

$$\Delta E \rightarrow \max, \text{ with } \Delta C = \Delta C_j \text{ and } N_{ts} \leq N'_{ts} \qquad (f)$$

for same conditions as above, but within a time limit N'_{ts};

$$\Delta C \rightarrow \min, \text{ with } \Delta E = \Delta E_j \qquad (g)$$

if we wish to obtain a certain error reduction ΔE_j at a minimum cost;

$$\Delta C \rightarrow \min, \text{ with } \Delta E = \Delta E_k \text{ and } N_{ts} \leq N'_{ts} \qquad (h)$$

if we wish to obtain a certain error reduction at a minimum cost and within a certain time interval.

Other constraints could be introduced according to particular circumstances, e.g. an upper limit for n_{ms} the number of flow measurements per year per station in connection with manpower limitations, etc.

The application of the technique is straightforward in cases (a), (b), (c) and (d) since it involves the optimization of a function which has some

of its variables subject to constraints. In the cases (e), (f), (g) and (h), for which, in addition to obtaining an optimum value of the function we have to also satisfy an equation, e.g. $\Delta C = \Delta C_o$ and which are in fact most interesting from a practical viewpoint, the solutions require the introduction of Lagrange's multipliers. Thus, for example, in case (e) the function to be optimized is written as

$$e = \Delta E + \lambda(\Delta C - \Delta C_o) \qquad (6.8)$$

in which λ is a Lagrange multiplier.* With this the problem is reduced to the previous case. This optimization can be done by means of a steepest descent (ascent) technique of multi-variate optimization for which there are well-documented computer programmes (e.g. CLIMB at the University of Waterloo). The technique is capable of circumventing difficulties related to the fact that some of the parameters of the functions to be optimized are not constant and cannot be expressed analytically. In this case, these parameters can be inputted in a tabular form.

The application of the technique requires initial values of the parameters to be optimized. If the initial values are not appropriately selected, it is possible that a local instead of an overall

---

*See Wilde and Beightler (1967) for details on this technique. A very clear explanation of the principles of the technique can be found in Copper, L. and D. Steinberg (1970) Introduction to Methods of Optimization, W.B. Saunders, pp.119-124 .

maximum is obtained. This may represent a significant drawback to the method. To circumvent this, it is necessary to obtain an initial "mapping" of the objective function by computing its values for various selected combinations of the values of the variables within their limits of variation. Although the number of such combinations increases rapidly with the number of variables and intervals considered, the computation is not excessively expensive since the computer time involved is not significant (of the order of a few seconds for several thousands of values on the IBM 360-75). This aspect of the problem is further discussed in Chapter 8 in connection with the application of the method described in this section to an actual study.

One of the major advantages of the method consists of the fact that it enables one to estimate the error reduction caused by changes in techniques of information interpolation or other improvements in the planning or operation of the hydrological network over the usual methods of increasing the accuracy by installing new stations, measuring more frequently and for a longer period of time, etc. The practical relevance and replacement value of research thus become obvious.

6.2 Planning of the Natural-Regime Network in Areas Designated for Development within the Planning Horizon.

6.2.1 General Considerations

In such areas one could produce an argument for making a certain set of decisions with respect to the natural-regime network, on the basis of the economic value of the hydrological information which is used for the design of the development projects. The literature contains several techniques of approaching such a problem either from the viewpoint of individual gauges and projects (Dawdy et al., 1970; Moss, 1970; Shawinigan Engineering, 1970; Davis and Dvoranchik, 1970) or from the viewpoint of the network for the design of a certain type of structure (Wilson, 1972).

In this section, we suggest a method for planning the natural-regime network in developing areas by means of dynamic programming. The method is based on the combined application of the method of planning the network in non-developing areas (Section 6.1), and on estimates of the monetary value in design and operation of projects due to incremental accuracy of hydrological information as estimated by means of any one or a combination of methods included in the references mentioned above. In addition, the application of the method requires a forecast of the economic development in the area considered, particularly of water resources-related development. The suggested method is discussed in the next sub-section.

6.2.2 Dynamic Programming Applied to Hydrological Network Design

The application of dynamic programming to the planning of the natural-regime hydrological network in areas designated for development can be carried out only if sufficiently detailed forecasts on the planned development are available. This is usually the case for periods of up to ten years since for projects of significance, feasibility studies are usually carried out with a lead time of 5-10 years.

If a period of 10 years is considered,[1] for the purpose of computation this period is divided into two 5-year intervals. The limits of the intervals constitute the stages. A series of values for the error of estimate of the hydrometric characteristics chosen as the planning criterion are then selected for each stage and these constitute the states (Fig.6.1). The transition from one state to another within a time interval is always along a path which minimizes the expenses, using the steepest descent technique described in the previous section. It is obvious that,

---

[1] The planning interval cannot realistically exceed 10 years and stages of 5 years each appear reasonable to use. The application of the technique with more than 3 stages is complicated and short cuts have to be used in this case, as explained in the application presented in Chapter 8. However, the use of more than 3 stages might be necessary because of the lag between the installation of new stations and the moment when the benefits from these stations (in terms of increased accuracy) are obtained.

as indicated in Fig.6.1, one can consider only transitions from a larger to a smaller error. From the results of the optimization computations for the transition from one state to another within each interval, one can estimate the related costs and their present worth (Fig.6.2a).

One can then consider all the developments which are planned to occur within each interval and by using one of the techniques available (e.g., Dawdy et al., 1970; Moss, 1970; Shawinigan Engineering, 1970) one can estimate a relationship between error of estimate of the hydrological characteristics and losses (from investment and operation) due to these errors for each of the planned developments. If one assumes that the error of estimate of the hydrological characteristics affecting a development is approximately equal to the average error at the beginning and end of the corresponding interval, it is possible then to estimate the losses due to such errors for each transition from one state to another within each interval. (Fig.6.2b) The present worth of these losses is then added to the corresponding present worth of the extension and operation of the hydrological network and the total costs (actual costs plus losses) are thus estimated for the transition from one state to another for each interval (Fig. 6.2c). It is easy then to compute the total costs for each possible path which could be followed from

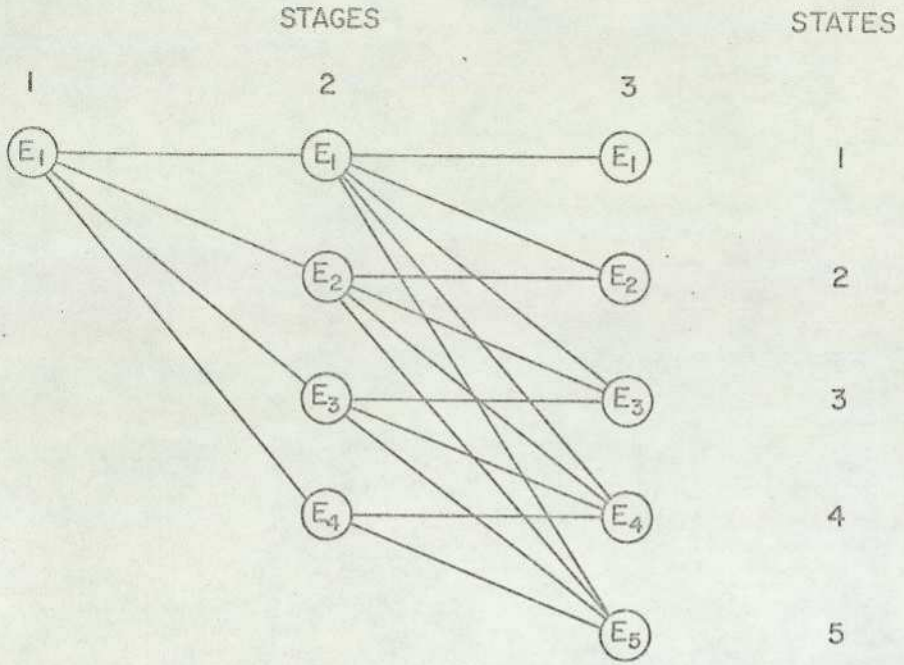
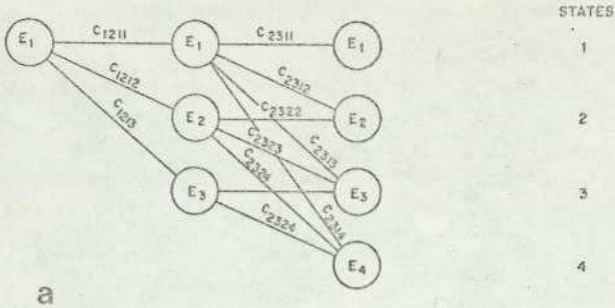
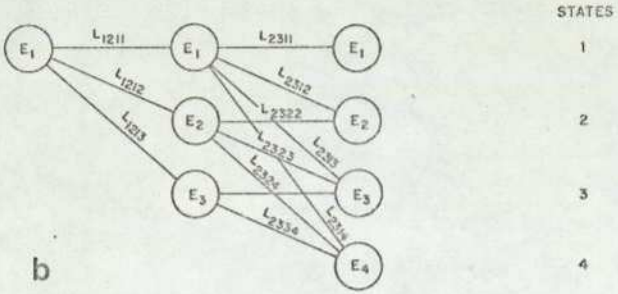
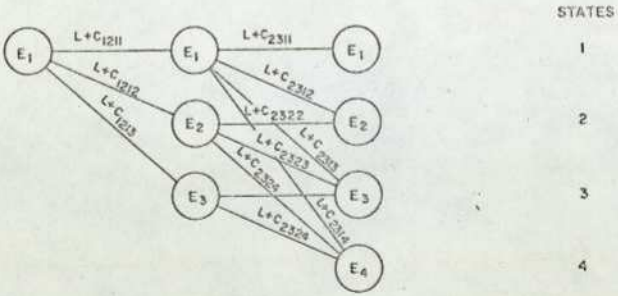


Fig. 6·1 STAGES AND STATES FOR OPTIMIZATION BY DYNAMIC PROGRAMMING OF HYDROLOGICAL NETWORK







NOTE: C = COST OF NETWORK OPERATION AND INVESTMENT
L = LOSSES DUE TO UNCERTAINTY (ERROR) IN HYDROLOGIC DATA
FIRST TWO INDEXES REFER TO TRANSITION FROM ONE STAGE TO ANOTHER,
LAST TWO INDEXES REFER TO TRANSITION FROM ONE STATE TO ANOTHER.

Figure 6.2 – STAGES AND STATES FOR APPLICATION OF DYNAMIC PROGRAMMING TO NETWORK PLANNING.

Stage 1 to Stage 3 and choose that path which would correspond to a minimum combined cost and losses (Fig.6-2c).

In some cases, difficulties may be encountered in the application of the technique because of the probabilistic character of some, if not all, economic forecasts. These difficulties could be circumvented by using expected values instead of "deterministic" values with respect to the losses.

The main practical problem in the application of the technique consists in the determination of the relationship between error of estimate of the hydrometric characteristics and losses incurred in investment and operation of various types of projects. In spite of the work done so far in this area, the problem requires further attention, particularly the development of such relationships for a series of typical projects, and methods of adjusting the results to various conditions, thus making possible an inexpensive estimate of the error-loss relationship for a large variety of potential projects. As long as such relationships are not available, the application of the technique described in the present section will be limited to hypothetical cases of the type described in Chapter 8. Nevertheless, as will be shown in Chapter 8, the application of the technique in the

framework of sensitivity analysis can be a useful tool for planning purposes even if well defined error-loss relationships are not available to the planner.

7. STATION DISTRIBUTION.

Once the number of stations which correspond to the optimum conditions according to the selected criterion has been determined, the next step consists in selecting the location of the stations in such a manner that the actual operation of the network satisfies as closely as possible the assumptions made in the optimization computations. In a first stage the location of the station is selected in a regional manner, i.e. decisions are made on the number of stations within a certain region and on the approximate size of the basins to be gauged. This first stage is discussed in this Chapter. In a second stage the selection of the appropriate locations of gauging stations has to be made from the viewpoint of the hydraulic and hydrological conditions (stability of the section, the presence of a hydraulic control, ice conditions, vegetation, etc.), accessibility, etc. This second stage is beyond the scope of this thesis.

The regional distribution of the gauging stations requires the examination of geophysical, economic, and environmental-aesthetical considerations, as discussed in Sections 7-1, 7-2, and 7-3 respectively.

7.1 Geophysical Considerations.

From a geophysical viewpoint, the distribution of the optimal number of stations has two basic aspects.

The first is related to ensuring within a region which is "model-homogeneous", i.e. a region in which one can apply basically the same model for interpolation of hydrological characteristics (or synthesis of hydrological time series), the minimum number of stations which will make it possible to interpolate (synthesize time series of flow)with the accuracy determined in the frame of the optimization process (Section 7.1.1). The second aspect is related to the necessity of avoiding the installation of redundant stations within an area which is physiographically and hydrologically homogeneous, i.e. in which the various river basins have almost identical input, physiographical characteristics, and outputs. The information gained by installing more than one or two stations within such an area is small and, other conditions being equal, it is much more advisable to distribute the stations more or less uniformly between such areas (Section 7.1.2).

7.1.1 Model-homogeneous Regions

The concept of the model-homogeneous region arises from the intention of delineating regions in which the term ϕ_t'' in Eq.(4.1) is minimized, and thus the error of information interpolation is also reduced.

The delineation of a model-homogeneous region can be done as suggested by Solomon (1972a) in the following manner:

a) The hydrological characteristic for which the regionalization is made is chosen and a model of information transfer is selected for the given characteristic on the basis of available data on obtainable accuracy and costs as shown in Section 6.1. The existing hydrometric stations are equally divided into a calibration sample and a validation sample, and the parameters of the model determined from the calibration sample.

b) If the model parameters correlate directly with the physiographical characteristics of the corresponding basins, as for example in the case of multiple regression techniques, one can proceed directly to the estimation by means of the model of the hydrological characteristic at the validation and calibration stations, and estimate the errors. If the model parameters do not correlate directly to the physiographical characteristics, as for example in the case of parametric models, one first correlates model parameters with physiographical characteristics, synthesizes the former at the validation stations and then proceeds to estimate the errors at the calibration and validation stations.

c) One then separates the whole area into two "super-regions":

(i) The "gauged" super-region which includes all basins which actually have gauges or whose physiographic characteristics fall within the

range of the corresponding physiographic characteristics of the gauged river basins.

(ii) The "ungauged" super region containing all other basins.

A first approximation of the number of new gauging stations (n_s') which should be allocated to each "super-region" (or region) can be then obtained by using the formula

$$n_s' = 5 \sum_{1}^{k} \frac{C_{vpi} - C_{vsi}}{C_{vpi}} - n_s \qquad (7.1)$$

where k is the number of significant and potentially significant physiographic characteristics. Significant physiographic characteristics are those included in the correlations used for direct and indirect modeling. Potentially significant are those physiographic characteristics which have not been selected as significant in the model, but for which the coefficient of variation for the gauged basins C_{vsi} is less than 0.5 of the population coefficient of variation C_{vpi}, and n_s is the number of existing stations.

d) The "gauged" super-region is then examined for possible further subdivision. This is done by examining the geographical distribution of the validation errors. If the geographical distribution of the errors as established by visual inspection or preferably by statistical tests (e.g. run tests) can be considered random, then the region is considered homogeneous (the whole "gauged super-region" is considered to belong

to one model-homogeneous region). If the geographical distribution of the errors indicate a geographical trend for grouping in various ranges, then this result is construed as an indication that the selected model does not apply to the whole "gauged" super-region, and that further second order subdivision is indicated. A second order subdivision defined at this stage includes all basins which have their physiographical characteristics within the range of the physiographical characteristics of the basins whose errors were included within the chosen error range used as a criterion for subdividing the "gauged super-region".

e) After defining the second order subdivision the procedure is iterated in the sense that the model is applied separately to each second order subdivision and step (d) applied again for possible third order subdivision, etc.

Figure 7-1 shows the result of the application of this technique by the writer to Western and Northern Canada (Solomon, 1972a) using the mean annual runoff as the planning criterion and a multiple regression as the model applied. The significant effect of topography in relation to the surrounding oceans on the delineation of model-homogeneous regions is obvious in this case. It is interesting to note that the application of the same technique to the Island of Newfoundland for the same criterion and model as above lead to the delineation of only two regions, and of one

single region for the whole of Southern Ontario (for the area shown in Figure 7-2).

It should be noted that the procedure requires that each subdivision includes a large number of stations to make the correlations between model parameters and physiographical characteristics meaningful.* It is also required that a large portion (at least 50%) of the stations are reserved for validation, to avoid the influence on the regionalization process of the fact that one obtains a better "curve fitting" effect as the size of the subdivision and the corresponding number of stations (data) decreases. If the regionalization process indeed reflects the regional difference in the hydrological conditions not covered by the physiographical conditions considered and by the model of information interpolation chosen, a continuous reduction and randomization of the errors of estimate will be observed for the validdation sample. In the contrary case, one usually obtains an increase in the errors of the validation sample, although randomization is not ruled out.

The need for many regions is generally an indication of an unsatisfactory statistical model and/or unsatisfactory suite of physiographical characteristics. Possible improvements in the physiographical

---

* If the number of gauges in a gauged region is not sufficiently large (at least 3-4 times the number of significant physiographical characteristics in the model) that region is attached to the "ungauged super-region" and treated accordingly.

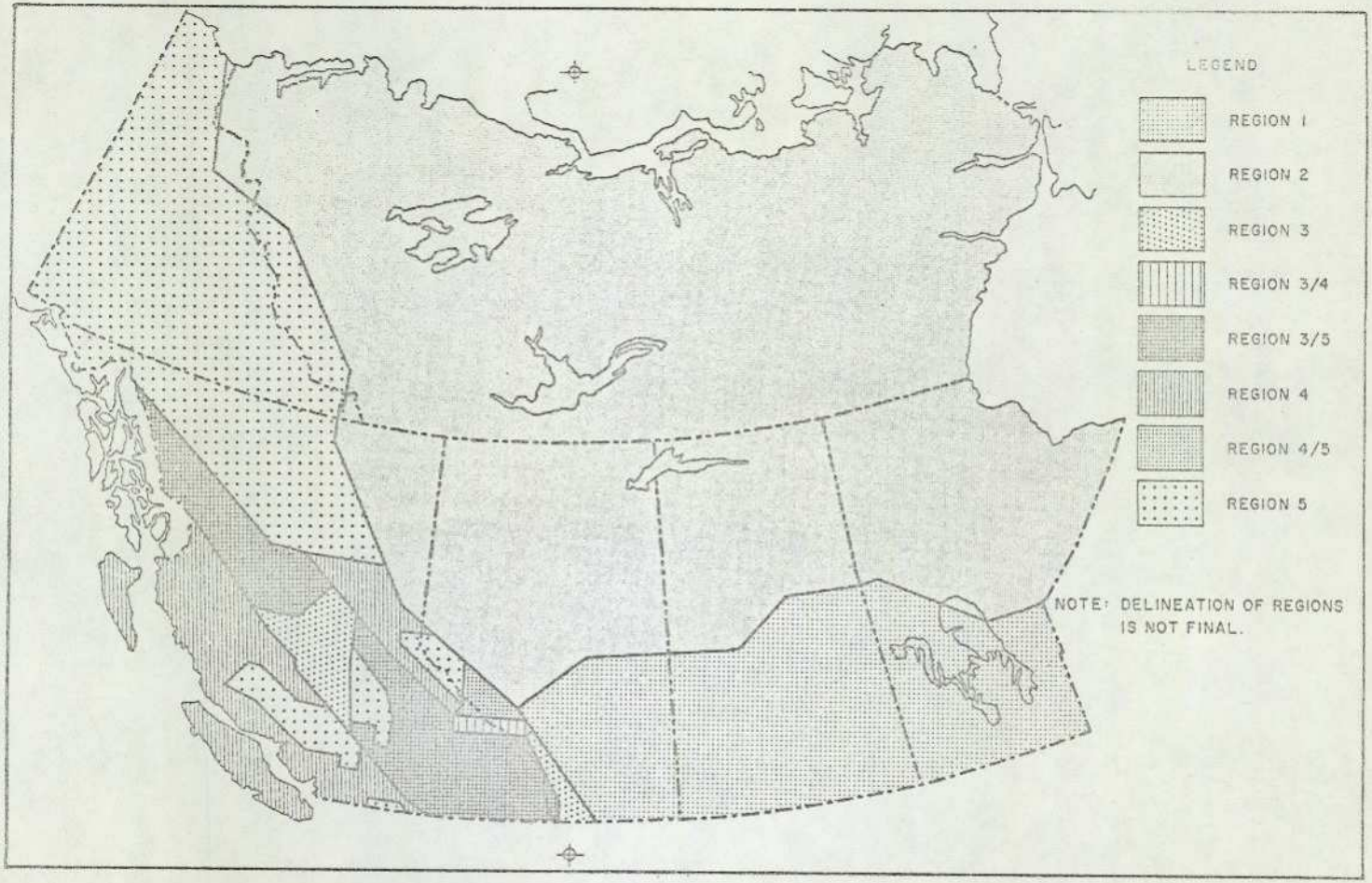


Figure 7-1 - MODEL- HOMOGENEOUS HYDROLOGICAL REGIONS IN WESTERN AND NORTHERN CANADA

characteristic suite and in the model should be investigated.

Differences in region delineation for various hydrological characteristics should be expected because of the different physiographical characteristics which determine them (e.g. in the case of maximum and minimum flow). However, large differences between region boundaries of the same area when they are delineated for different hydrological characteristics are probably in most cases also an indication of the weakness of the model and suite of physiographical characteristics considered.

Once the subdivision (regions) has been determined and the number of new stations estimated for each in accordance with Eq. (7.1), one can check if the total number of stations corresponds to that determined in accordance with the optimization technique applied to the whole area (Chapter 6). If the number of stations determined is different from that obtained from the optimization computations, one can change the number of stations allocated in a proportional manner, so as to obtain the optimum number of stations. If a significant increase (more than 20%) is necessary, this may indicate that the model used in the optimization computations is not applicable and an iteration of these computations may be necessary.

The practical application of this regionalisation procedure is facilitated very much by the use of the square grid interfacing bank of physiographical land-cover land-use, and hydrometeorological data described in Chapter 3. This data bank makes it possible to apply the procedure in a completely computerised manner.

7.1.2 Physiographical-hydrological Regionalisation

Physiographical-hydrological regionalisation is carried out with the purpose of achieving the equivalent of stratified sampling. Physiographical-hydrological regions are regions within which the significant physiographical factors as defined for example from correlations between hydrological and physiographical characteristics or between model parameters and physiographical characteristics vary within narrow limits. One can expect that within regions defined in this manner the hydrological characteristics which have been used in the selection of the significant physiographical characteristics vary also within narrow limits.

The technique for delineation of physiographical regions can be applied using the following steps:

a) Define correlations between various hydrological characteristics (selected as criteria for the regionalisation or for network planning) or model parameters and physiographical characteristics and select all significant physiographical characteristics;

b) Analyze the coefficients of variation of each non-significant physiographical characteristic of the sample used and for those physiographical characteristics for which the sample C_v is less than 0.2 compute also the C_v of the population for the given physiographical characteristic. If the population C_v is larger than 0.4 include the corresponding physiographical characteristic in the list of potentially significant physiographical characteristics;

c) Select ranges (shades) for each significant and potentially significant physiographical characteristics (at least three shades for each characteristic),

e.g. Shade 1: $X > \bar{X} + \frac{1}{2} Sd_x$

Shade 2: $\bar{X} + \frac{1}{2} Sd_x > X > \bar{X} - \frac{1}{2} Sd_x$

and Shade 3: $X < \bar{X} - \frac{1}{2} Sd_x$

where $\bar{X}$ is the average value of the physiographical characteristics and Sd_x its standard deviation;

d) Define the $N_r = m^n$ physiographical regions where m is the number of shades and n the number of characteristics included in the regionalization, and delineate the area of each region by overlaying the corresponding maps of the physiographical characteristics subdivided into the selected ranges.

The last operation is best achieved by means of computers using data bank techniques as described in Chapter 3. An example of such physiographical regionalization carried out in Southern Ontario is illustrated in Figure 7.2.

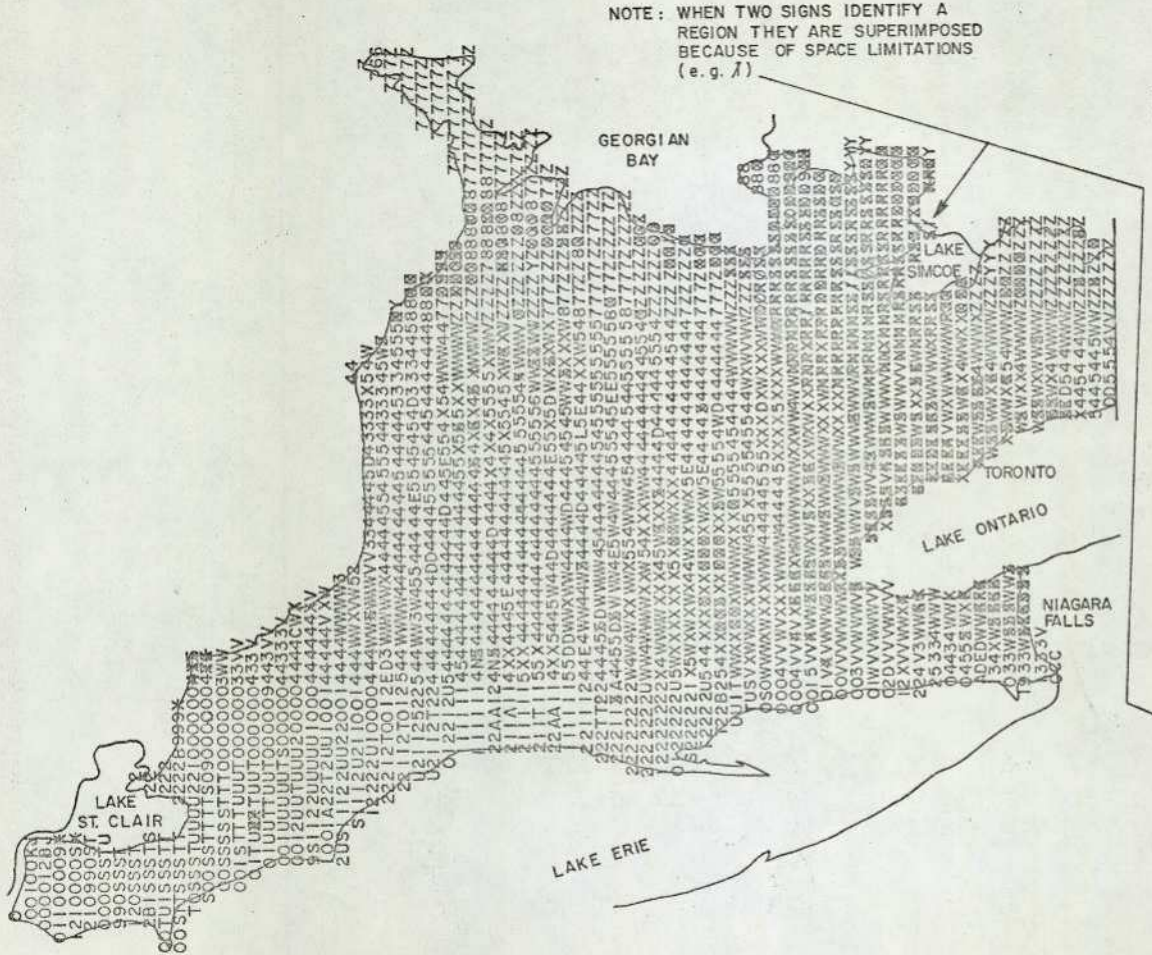
Once the physiographical regions have been defined, one should test the assumption that basins located in the same physiographical-hydrological region have a similar hydrological regime by comparing the hydrological characteristics (in dimensionless or comparable terms such as runoff) for all basins located within the same physiographical region. Following this validation test, the new stations allocated are then distributed to obtain as uniform a distribution between the physiographical-hydrological regions as possible.

7.2 Economic Considerations

Two distinct sets of economic considerations have to be taken into account in connection with the distribution of the new stations within the optimized design:

a) considerations related to the costs of installing and operating the stations;

b) considerations related to economic planning in the area covered by the network.



LEGEND

POSITION OF L's
SYMBOL USED ON MAP

LBSW	1	1	1	1	1	1	1	1	1
LABL	0	1	2	3	4	5	6	7	8
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	1	1	1	1	1	1	1	1	1
LABL	9	A	B	C	D	E	F	G	H
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	1	1	1	1	1	1	1	1	1
LABL	I	J	K	L	M	N	P	Q	R
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	2	2	2	2	2	2	2	2	2
LABL	S	T	U	V	W	X	Y	Z	0*
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	2	2	2	2	2	2	2	2	2
LABL	1*	2*	3*	4*	5*	6*	7*	8*	9*
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	2	2	2	2	2	2	2	2	2
LABL	A*	B*	C*	D*	E*	F*	G*	H*	I*
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	3	3	3	3	3	3	3	3	3
LABL	J*	K*	L*	M*	N*	P*	Q*	R*	S*
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	3	3	3	3	3	3	3	3	3
LABL	T*	U*	V*	W*	X*	Y*	Z*	0/	1/
LLAT	1	1	2	2	2	3	3	3	3
LPIN	1	2	3	1	2	3	1	2	3
LBSW	3	3	3	3	3	3	3	3	3
LABL	2/	3/	4/	5/	6/	8/	B/	C/	D/
LLAT	1	1	1	2	2	2	3	3	3
LPIN	1	2	3	1	2	3	1	2	3

KEY FOR L's

BSW =	0	<50	>50	FT x 10
LBSW	1	2	3	
ABL =	0	<30	>30	(UNIT = SQ. KM/4)
LABL	1	2	3	
LAT	<30	>30 & <60	>60	
LLAT	1	2	3	
PIN	<200	>200 & <500	>500	
LPIN	1	2	3	

Figure 7-2 : PHYSIOGRAPHICAL - HOMOGENEOUS HYDROLOGICAL REGIONS IN SOUTHERN ONTARIO

7.2.1 Costs of Installation and Operation

The average costs of installing and operating a station have been discussed in Chapter 5. As was pointed out there, these costs vary with many factors, of which the most important are the size of the river and accessibility. Since costs increase significantly with the size of the river basin, and since from the viewpoint of accuracy of information interpolation techniques it is more advantageous to have data from basins of small sizes (a few hundred square miles) for which the average physiographical characteristics do not display excessive variation, one should prefer, other considerations being equal, the installation of gauging stations on small rivers for the extension of the natural regime network. Furthermore, it is also obvious that from the viewpoint of the installation and operation costs, other conditions being equal, it is preferable to install stations in more accessible areas. However, in practice, there is a tendency to over-emphasize the requirement of accessibility, resulting in the installation of stations too close to each other and failing to meet the requirements of a regionally balanced distribution as indicated by the statistical-hydrological and physiographical-hydrological regions (Section 7.1). In all cases considerations of balance in station distribution should take precedence over those of accessibility, the latter criterion being

applied only for choosing between two locations of equal merit from the viewpoint of regionally balanced and physiographically stratified sampling.

7.2.2 Economic Development

This problem is analysed in detail in Section 8.1.2. We note here only that in areas designated for development the emphasis may shift from the small rivers to the larger, but sampling considerations must not be neglected.

7.3 Environmental and Aesthetic Considerations

A significant group of factors which has to be considered in the selection of sites for hydrometric stations is that related to environmental and aesthetic considerations. As stated by Langbein (1972b,) "many decisions on river and reservoir development are beginning to be influenced as much by environmental factors as by hydrological factors. There appears, therefore, to be justification for river data to include information on such matters in order that they may be considered in water resources planning in its initial stages."

The problem has two aspects in relation to hydrological networks: a major one related to the relationship between the environment as a whole and the water resources as a part of this whole, and a minor one related to the integration of the structure and equipment of the hydrological station and of the

activity related to it into the corresponding environmental setting. The latter problem is related to actual station site selection and design and is outside the scope of this thesis. It has to be considered alongside all other problems of the practical selection of a gauging site after the regional site was selected, such as the problem of site accessibility, section stability etc. The only reason for which this problem was mentioned here is because it is often overlooked in practice. The first aspect is relevant to the planning of the hydrological network and deserves further discussion.

The relationship between environment and water resources has regional as well as local implications. The regional implications are related to the complex role played by water in molding the environment, e.g. its role in the determination of the vegetation and animal life of the region, erosional and sedimentation patterns etc. The planning of a hydrological network which should be useful for environmental assessment requires, as indicated by Langbein (1972b), "an inquiry among landscape specialists, ecologists, and conservationists on the facts which are significant" from an environmental-hydrological viewpoint. Once these facts have been selected, it might become feasible to delineate "hydrological-environmental" regions or "physiographical-environmental" regions in a similar manner as for the

delineation of the "physiographic-hydrological" regions (Section 7.1.2). Data banks and environmental-hydrological models could also play a significant role in such regionalization. Once the regions have been delineated it would then be necessary to organize the network so that each region is gauged so as to satisfy hydrological requirements.

The local implications are related to the aesthetical value of the river valley landscape - the "riverscape" to use a word coined by Leopold (1969). The local problem could be treated in a manner similar to the regional one, with the difference that in this case the elements which have to be considered are those which in the opinion of various people have significance from an aesthetical viewpoint. Because of this, regionalization would have a subjective character. This is further complicated by the difficulty of rating these factors and attaching figures to their qualitative variation. Because of this, as pointed out by Howe (1971), "techniques for environmental description are just in their infancy". Nevertheless, one can ascertain that there are a series of physical, chemical, biological, and human factors which can be measured and which certainly have a significant influence on the aesthetic perception of riverscapes by human beings. These factors were listed by Langbein (1972b). If one could regionalize the river valleys according

to the space variation of these factors and, on this basis, set up a network to measure those factors which vary in time, the results of such measurements could be used with subjective ratings set up according to the nature of uses and user preferences to produce information needed in decision-making in this area. A technique of producing such information by means of factor analysis with application to the recreational use of reservoir shorelines has been developed by Jaakson (1970), and could also be of some use in the regionalization process. However, such regionalization currently requires expensive and time-consuming ground surveys. It is,therefore, suggested that research be carried out on the possibility of "environmental-hydrological regionalization" by means of interfacing data banks and remote sensing, particularly aerial photography.

8. PRACTICAL APPLICATIONS

The methodology described in the previous sections was applied by the writer for planning the hydrological network in various areas of Canada. Preliminary results of this application were described by Solomon (1972c). Some results obtained in the area of hydrological regionalization have already been presented in Chapter 7. We shall, therefore, in this section discuss only the results of the application of the methodology to the estimation of optimum planning variables as described in Chapter 6. Since there are uncertainties related to many input data, a large portion of these practical applications is devoted to the analysis of the sensitivity of the results to these uncertainties.

The methodology described in Section 6.1 was applied to the natural regime section of the hydrological network in an area of Canada covering about 2.2 million sq. mi. assuming in a first approximation that no significant development will take place in the natural regime section. The results of this application are described in Section 8.1. In a second approximation, the hypothesis of significant development in these areas was made and the corresponding planning variables determined using the dynamic programming technique described in Section 6.2. The results of this application are discussed in Section 8.2.

8.1 Application to Natural Regime Section Assuming no Economic Development

This part of the practical application assumes that the objective of the analysis is to estimate the optimum planning variables for Western and Northern Canada (i.e. the Prairie Provinces, British Columbia, Northwest Territories and the Yukon). Because of the incomplete level of the supporting studies required for the application of the optimization technique and the statistical character of the relationships between error of estimate of hydrological characteristics and network parameters, the results of the analysis are not exact. In order to investigate the order of magnitude of the effects of the approximation involved, a sensitivity analysis was carried out to examine the effect of variation in the basic relationship between the error of information interpolation and number of stations on the optimum network strategy, considering various planning objectives. Furthermore, the sensitivity analysis has been extended to examine the effect of changing the hydrological characteristic which is considered as the planning criterion on the optimization results.

8.1.1 Error Variance Parameters

The error variance parameters to be used in Eq.4.7 were determined on the basis of the results presented in Chapter 4. For the error variance parameter E_s^* two methods of information interpolation were considered:

(1) The parametric model

(2) Correlations between physiographical and hydrological characteristics.

It should be noted that the analysis was carried out in both cases on the assumption that the above two methods are applied in conjunction with the physiographical, land-use land-cover and hydrometeorological data bank described in Chapter 3. This makes both methods more efficient than the ones in current use. That the correlation method is more efficient when used in combination with the above mentioned data bank was demonstrated in the Network Planning Study for Western and Northern Canada (Shawinigan Engineering, 1970). A similar increase in efficiency has not yet been demonstrated for the parametric model technique. However, it is unlikely that this method could be even attempted without the use of the square grid technique in areas such as the Canadian North where meteorological data are very scarce. In the writer's opinion, only such a technique can provide a practical and objective procedure of estimating the required input data for setting up a parametric model. Therefore, in order to emphasize that the two methods are used in conjunction with the physiographical, land-use land-cover and hydrometeorological data bank which uses a square grid system for storing the data, the designations employed for them subsequently in this chapter are parametric

model - square grid technique and correlation-square grid technique.

To provide a basis for sensitivity analysis the graphical relationship between error of information interpolation E_s and number of stations N_s was used to produce three sets of values for the error variance parameter E_s^*: a high value corresponding to the upper envelope of the data points, an average value and a low value corresponding to the lower envelope. The values of E_s^* were adjusted to take into account that the area considered is 2,200,000 square miles (Section 4.2.1, third paragraph). The values thus obtained are summarized in Table 8.1.

TABLE 8.1

Values of Error Variance Parameter E_s^*

Hydrological Characteristic	Error Level	Method of Information transfer	
		Parametric Model sq. grid	Correlation sq. grid
Average annual runoff	high	3.70	9.00
	average	2.50	7.30
	low	1.30	5.00
Coeff. of variation of annual runoff	high	5.00	7.75
	average	3.75	6.00
	low	2.50	4.00
Mean May runoff	high	2.75	8.00
	average	2.00	6.00
	low	1.25	4.00

The error variance parameters $E^*_{t,50}$ were obtained from the results shown in Table 4.4. The error variance parameter $E^*_{m_i m}$ was assumed to be equal to 0.44 (see sub-section 4.2.3).

8.1.2 Costs

The following assumptions were made in consultation with the Water Survey of Canada, and on the basis of the discussion included in Chapter 5:

C_0 = annual costs of network operation including overhead, data processing and studies, but excluding costs of new stations, station operation and flow measurement

C_0 = 800,000 dollars/year if the data transfer method used is the correlation-square grid method, and

C_0 = 1,200,000 dollars/year if the data transfer method used is the parametric model - square grid technique.

The basic estimates are that the administration and other fixed costs at each of the four districts are of the order of $180,000/year, and that the costs of studies for the correlation - square grid technique would be of the order of $80,000/year and for the parametric model -

square grid technique of the order of $480,000/year[1]. In both cases it is assumed that these costs include costs for obtaining detailed grid data in areas where the variation of the physiographical characteristics and the size of the river basins would require such detail, and that the analysis of flow data is based on daily or at least weekly time intervals.

C_1 (cost of installing a new station)

C_1 = 3,000 dollars/station

C_2 (cost of operating one station for one year excluding administrative and fixed costs)

C_2 = 700 dollars/year per station

C_3 (cost of carrying out one measurement)

C_3 = 100 dollars/measurement

ΔC_{03} (cost of new instrumentation)

ΔC_{03} = 500,000 dollars (only once at the beginning of the study period)

Note that ΔC_{03} is of no consequence in the optimization process, this being a constant.

---

[1] This amount is overestimated on purpose. It is unlikely that such studies cost more than $250,000 to $300,000/year for the period of time investigated. The overestimation is introduced to emphasize the advantages of the model - square grid technique even if it is assumed that the cost of the related studies is very high.

The above figures are of necessity average ones and do not reflect the large variation of costs which characterize the study area because of the variation of station density and of the geographic and economic conditions. Because of the lack of detailed data it is not possible to apply the planning analysis separately to various regional subdivisions of the area, e.g. a northern and a southern region. However, if such details become available, it would be perfectly feasible to carry out such regional analyses in a similar way as for the whole area. This would provide information for regional planning which to a certain extent may affect the network design. In addition, it would also be useful to compare the results of such regional analyses with those obtained for the area as a whole, to obtain an estimation of the approximations involved in the results obtained because of the use of averaged values.

8.1.3 Results

Using the data assumed above the methodology described in Section 6.1 was applied to determine the optimum values of the planning variables for various criteria as shown in that section, for the two basic methods of data transfer, viz. the correlation - square grid and the parametric model- square grid techniques. The following objectives were examined:

Minimum cost per unit of error reduction i.e. objective a in sub-section 6.1.2, and 1 in Table 8.2.

Maximum error reduction within a given time interval (8 years) i.e. objective d in sub-section 6.1.2 and 2 in Table 8.2.

Minimum cost per unit of error reduction for a given error reduction (3%), i.e. objective c in sub-section 6.1.2 and 3 in Table 8.2.

Minimum cost per unit of error reduction within a given period of operation (5 years) i.e. objective b in sub-section 6.1.2 and 4 in Table 8.2.

Three different hydrological characteristics were considered as planning criteria:

- the mean annual flow (criterion A in Table 8.2)
- the coefficient of variation of the mean annual flow (criterion B in Table 8.2)
- the mean flow of the month of May (criterion C in Table 8.2)

As the currently available optimization routine (CLIMB) requires significant computer time when the number of variables to be optimized is large and as the influence of the measurement error on long term averages is not large, it was decided to consider that the number of flow measurements per year should be considered as fixed ($n_{ms} = 10$). For the same reason the percentage of gauging stations

to be discontinued was considered only for three given values (p = 0, p = 25% and p = 50%). In the first stage of the analysis the value of E_s^* in the relationship between error of information transfer and number of stations was considered to correspond to the average values shown in Table 8.1.

The results summarized in Table 8.2 (in which the optimum number of stations was rounded up to the nearest 50) indicate the following:

The error reduction with increased number of gauging stations is smaller for the model - square grid technique than for the correlation-square grid technique, but this occurs because the application of the technique reduces drastically the initial error. Reducing the error becomes more difficult as the starting error decreases.

The objective of least cost per error reduction is useful only for the very short range strategy, because it indicates the optimum only for one year of operation, disregarding the optimization of a longer period, since the maximum error per unit cost objective is always obtained at the beginning of the period considered.

The objective of largest error reduction within a time interval leads to the upper limit of number of stations in most cases (except for May - model technique) indicating also that this criterion is not a very good one for planning. The fact that for May and the model technique the criterion does not lead to the maximum

number of stations seems to indicate that the model is more efficient for the interpolation of monthly values that have a large variability.

Most useful for planning purposes appear to be the criteria of least cost for a certain error reduction, and least cost per error reduction for a given time period. By combining results obtained from the application of these two criteria, and considering the budgetary constraints one can find the most adequate combination of number of stations to be operated for a certain time, when a certain annual budget is available and a certain level of error reduction is desired. However, since we do not yet have very clearly defined relationships between errors and the planning variables, actual applications of the methodology described would benefit most from sensitivity analyses which would indicate which elements of the decision process would have the largest influence on the decision variables. This is discussed in the next sub-section.

8.1.4 Sensitivity Analysis

A sensitivity analysis was carried out to ascertain primarily the influence of possible variations in the relationship between error of information interpolation and number of stations. Three different values of the corresponding error variance parameters for each of the two methods of data transfer were used as described in sub-section 8.1.1 (Table 8.1).

TABLE 8.2

Optimum Values of Planning Variables for Various Objectives
(Interest rate 7% p.a.)

A. Parametric Model-Square Grid Technique

Objective*	Criterion*	Error reduction**	N_{s_f}**	N_{t_s}**	p%**	Final error (Ei - ΔE)**	Cost (Present Worth)[1]
1	A	1.6	650	1	25	13.2	3.3
	B	1.3	550	1	0	22.4	2.9
	C	1.4	400	2	0	18.1	1.4
2	A	6.5	2000[2]	8	0	8.3	35.2
	B	7.6	2000	8	0	16.1	35.2
	C	3.9	500	8	0	15.6	15.5
3	A	3.0	800	3	0	11.8	8.9
	B	3.0	650	3	0	20.7	7.7
	C	3.0	400	6	0	16.5	10.9
4	A	3.8	900	5	25	11.0	13.6
	B	4.0	750	5	25	19.7	12.1
	C	2.8	400	5	0	16.7	9.3

[1] millions of dollars

[2] the maximum limit of N_s was set at 2000.

* See section 2.1 for definition and p.8.8 for identification of objectives and criteria.

** See sub-section 6.1.1 for definitions.
(E_i = Error at the beginning of the study period.)

TABLE 8.2 (cont'd)

Optimum Values of Planning Variables for Various Objectives
(Interest rate 7% p.a.)

B. Correlation - Square Grid Technique

Objective*	Criterion*	Error reduction (ΔE)**	N_{s_f}**	N_{t_s}**	p%**	Final error (Ei - ΔE)**	Cost (Present Worth)[1]
1	A	10.5	900	1	0	25.3	4.2
	B	4.7	700	1	0	28.6	4.7
	C	3.7	650	1	0	30.6	3.0
2	A	19.0	2000[2]	8	0	16.8	32.0
	B	14.1	2000	8	0	19.2	32.0
	C	13.4	2000	8	0	20.9	32.0
3	A	3.0	750	1	0	32.8	3.6
	B	3.0	700	1	0	30.3	3.3
	C	3.0	650	1	0	31.3	3.1
4	A	12.7	1000	5	0	23.3	11.9
	B	8.7	900	5	0	24.6	12.3
	C	7.6	800	5	0	26.7	11.6

[1] millions of dollars

[2] the maximum limit of N_1 was set at 2000.

* See section 2.1 for definition and p.8.8 for identification of objectives and criteria.

** See sub-section 6.1.1 for definitions.
(E_i = Error at the beginning of the study period.)

In addition, the influence of the following factors was investigated in the sensitivity analysis:

- level of sophistication of data transfer technique (high: corresponding to models, low: corresponding to multiple correlations);
- and the level of variability of the hydrological characteristic considered in the planning analysis: low (long-term average flow), medium (standard deviation of mean annual flow) and high (average monthly flow in May).

The results of the sensitivity analysis are summarized in Table 8.3.

As can be seen from this table, the number of supplementary stations required increases rapidly with the decrease of the sophistication of the technique of data transfer, but the increase is much slower if the relationships adopted reflect a lack of confidence in the possible success and further improvement in the techniques of information transfer. Table 8.3 indicates also a rapid increase of the supplementary number of stations as the variability of the hydrological characteristics used in the planning process decreases.

Although some of these results are to be expected from the simple analysis of the survey process, in some instances the influence on the rational decision-making process of the adoption of a given information transfer technique or the use of one or another hydrological criterion is somewhat surprising.

TABLE 8.3

VARIATION OF OPTIMUM NUMBER OF SUPPLEMENTARY STATIONS CORRESPONDING TO LARGEST ERROR REDUCTION PER DOLLAR SPENT WITHIN A FIVE YEAR PERIOD (WESTERN + NORTHERN CANADA CONDITIONS)*

Level of sophistication of information transfer technique	Level of confidence in success of application of information transfer technique	Level of variability of hydrological characteristic considered		
		High (May aver.)	Medium (S.dev. of annual flow)	Low (Long term aver. flow)
High (models)	high (lower envelope)	0	100	150
	medium (best fit)	0	350	500
	low (upper envelope)	50	500	600
Low (correlation)	high (lower envelope)	250	350	600
	medium (best fit)	400	500	600
	low (upper envelope)	500	600	600

* The computations were carried out assuming $n_{ms} = 10$, $p\% = 0$, and $N_{si} = 400$.

(For definition of symbols see Sub-section 6.1.1).

To investigate the effect of the variation of p, the percentage of discontinued stations, a further sensitivity analysis was carried out, in which significantly different values of p (p = 0, p = 25% and p = 50%) were considered.

Although the discontinuation of 50% of the stations is not a realistic proposition, this was nevertheless investigated to obtain an extreme case which would point out clearly the potential benefits or otherwise which may be derived from such discontinuation. In order to simplify the computer programme, the hypothesis was made that all stations designated for discontinuation would be removed from operation in the very first year of the planning period. This is matched by the hypothesis that all new stations are installed in the same year. Any other hypothesis would have involved various alternatives to be considered, increasing significantly the computational complexity. As shown in Table 8.4 this analysis has indicated that discontinuation of stations does not produce anything more than insignificant differences in the results for both methods of information interpolation.

These results indicate that the advantages of station discontinuation are minimal at least in the short range. This is probably due to the low cost of station operation compared to the other costs. The analysis of the cost and accuracy formula indicates that for a long range analysis the significance of station discontinuation becomes larger.

TABLE 8.4

Summary Data on Effect of Station Discontinuation[1]

A) Model - square grid

Objective	Minimum cost per unit of error reduction			Minimum cost per unit of error reduction for a 5 year period		
Percentage of stations discontinued, p	0	25	50	0	25	50
Final error, E_f%	13.1	13.2	13.3	10.7	10.9	11.1
Cost per unit of error reduction	2.2	2.2	2.3	3.7	3.6	3.6
Total No. of stations for optimum cost per unit of error reduction, N_{sf}	650	650	650	900	900	900
No. of stations in operation at end of period, $N_{sf} - pN_{si}/100$	650	488	325	900	675	450

B) Correlation - square grid

Percentage of stations discontinued, p	0	25	50	0	25	50
Final error, E_f%	26.5	26.5	26.6	23.2	23.2	23.3
Cost per unit of error reduction	0.4	0.4	0.4	1.1	1.0	1.0
No. of stations for optimum cost per unit of error reduction, N_{sf}	900	800	800	1000	1000	1000
No. of stations in operation at end of period $N_{sf} - p\,N_{si}/100$	900	600	400	1000	750	500

[1] Total number of stations $N_{si} = 400$

A sensitivity analysis to check the influence of the rate of interest on the planning variables indicates that this rate has an extremely small influence on the planning variables.

The conclusion of the above analyses is that the factor which has the largest influence on the planning variables, particularly the number of stations, is the technique of information interpolation used, followed in importance by the hydrological characteristic used as the planning criterion. Further research is, therefore, required to establish more reliable relationships between error of information interpolation and number of stations for various techniques of information interpolation and various criteria.

8.2 Application to Natural Regime Section Assuming Significant Economic Development (Transitional Section)

8.2.1 Basic Assumptions

In this hypothetical application the very extreme assumption was made that all forecast development in the water resources field will take place in river basins which may be classified as belonging to the natural regime section.[1]

[1] In fact during the period investigated a large part of this development will take place in Northern Regions which are basically included in the natural regime section. It appears now that the forecast of Bruce and Maasland was conservative and that development particularly in hydro-electric plants and resource exploitation will be faster than forecast. Thus, the assumption made might not be in fact far from the actual figures.

In the hypothetical application made which simulates very roughly the Canadian conditions, the year 1985 was considered as the planning horizon and the year 1971 as the starting point. The interval was divided into three five-year periods. The forecast of investment in water resources related development used in this application was obtained from Bruce and Maasland (1968, p.48). The estimates given by Bruce and Maasland were multiplied by a factor of 1.26 to take into account the difference in cost estimates in 1968 and 1971 due to inflation. Note that the estimates given by Bruce and Maasland do not include costs related to water pollution control, except for costs of sewers. We did not include an estimate for such costs since it is presumed that water pollution control projects will be designed only in areas with non-stationary hydrological conditions. Since this is not always the case, this compensates to a certain extent for the hypothesis that the remaining development takes place entirely in areas with stationary (natural) hydrological regimes. The effect of possible errors in the forecast of economic development is analysed in sub-section 8.2.5.

Based on the studies carried out by Shawinigan Engineering (1970) and by Wilson (1972) a rough relationship between error in the standard deviation of yearly flows E_s(%) and annual economic losses due to erroneous design L was accepted in the form

$$L = aE_s^2 I \qquad (8.1)$$

where a is a constant depending on the type of project and I is the investment. Obviously, not all projects are affected by an error in standard deviation of the annual flow, but an error in this hydrological characteristic corresponds to a network configuration which is usually associated with errors in all other hydrological characteristics. Since the errors in various hydrological characteristics are, in most cases, strongly correlated and since any type of project affected by the hydrology of the area will be affected by errors in at least one hydrological characteristic it may be assumed that the losses affecting any type of project are directly related to errors in the estimation of the standard deviation of the annual flow.

From the Shawinigan study mentioned above, the value of a was estimated to be 0.00016, i.e. an error of 10% in the standard deviation would result in an annual loss of 1.6% of the investment. The effect of uncertainty in estimating the value of a is discussed in sub-section 8.2.5.

8.2.2. Steps Involved in the Hypothetical Application

The dynamic programming technique was applied basically as described in Section 6.2., but with some slight modifications which are explained further. The steps required for the application of the technique are: (see also flowchart, p.8.21).

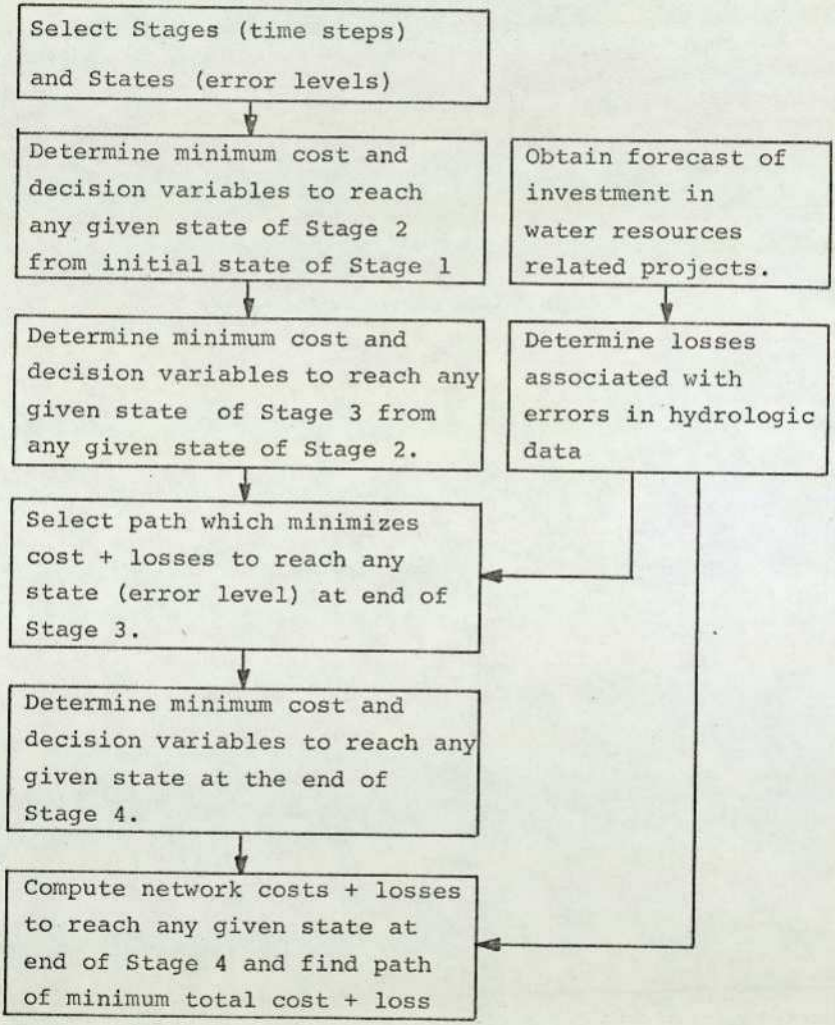
Step 1: Select stages (time steps) and states (error levels). The beginning of each five-year interval (1971, 1976, 1981) and the end of the last period (1985) were considered as stages (4 stages) and various levels of error attainable in each stage, as states (Fig.8-1).

Step 2: Determine minimum cost and decision variables to reach a given state (error level) at end of first time interval (stage 2). If one accepts that the average level of the error in the period between two stages is the average of the corresponding error levels of the two states considered, it is possible to estimate from Eq.8.1 the losses for the transition from one state to a given stage to another state of the following stage due to the error in hydrological information for any possible transition path.* It is also possible to compute for each path followed from the initial stage (1) to the various states of stage 2 the minimum cost required for obtaining the corresponding accuracy level by means of the method presented in Section 6.1, considering the objective of minimum cost for a certain error reduction and a given time interval. By adding the minimum cost for reaching the

---

* Although the analysis is shown as being carried out in a lumped manner, i.e. by assuming average conditions for the whole area, there is no reason why it could not be applied in a distributed manner, i.e. separately for various regions having various hydrological and/or economic development characteristics, provided that the data required are available.

SCHEMATIC FLOW CHART OF APPLICATION
OF DYNAMIC PROGRAMMING TECHNIQUE
TO HYDROMETRIC NETWORK PLANNING



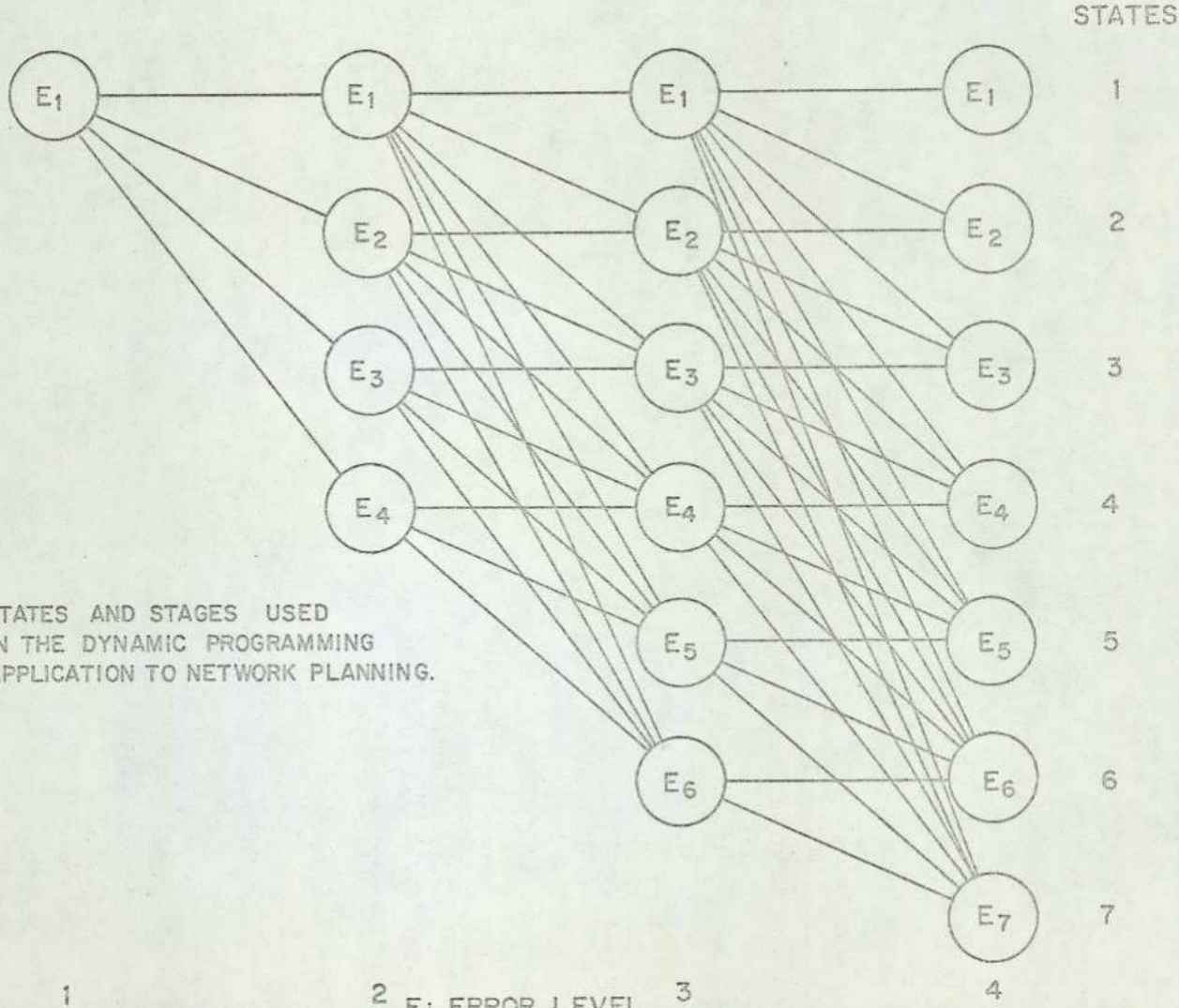


Figure 8-1 – STATES AND STAGES USED IN THE DYNAMIC PROGRAMMING APPLICATION TO NETWORK PLANNING.

corresponding error level to the associated losses one determines the minimum overall cost of getting from the initial state to any state (error level) of Stage 2, Fig.8.2a.

Step 3: Determine minimum cost and decision variables to reach any of the states (error levels) at end of second interval (Stage 3) starting from any state of Stage 2. Once the decision variables (N_s = number of stations, p = percent of existing stations to be discontinued, n_{ms} = number of flow measurements per station per year) have been determined for Stage 1 - Stage 2, it is possible to determine the same minimum costs and the corresponding decision variables for getting from any of the states of stage 2 to any of the states of stage 3, using the same technique as for Step 2, Fig.8.2b.

Step 4: Select path which minimizes network costs and project losses to reach any state (error level) at end of second interval (Stage 3) and note corresponding decision variables. For the transition from Stage 3 to Stage 4 the computation could become very complicated since one could arrive to each state of Stage 3 along several paths, which means that the transition from states of Stage 3 to states of Stage 4 must consider various initial values of the pertinent variables N_{si}, N_{mi}, p . In order to avoid this complication, it was decided to use a somewhat unorthodox procedure, in the sense that the computation for the transition from Stage 3 to Stage 4 was computed only

for those initial conditions which correspond to the economically optimal path of arriving at the states of Stage 3. This decision was taken because the economic forecast for the last period of five years is so uncertain, and technological changes could be so significant at that time, that a decision which would optimize the whole 1970-1985 period, but not the 1970-1980 period taken separately, would not be acceptable.

Step 5: Determine minimum costs and losses to reach any state (error level) at end of third time interval (Stage 4). Consequently, before computing the transition from Stage 3 to Stage 4, the optimum (minimum cost and losses) obtained paths for Stage 1 - Stage 3 were determined as illustrated in Fig.8.2c, by combining minimum costs of expanding and operating the network with the corresponding project losses due to errors in data. Using the decision variables obtained for the optimum (minimum) paths for reaching any state of Stage 3 it is then possible to compute the minimum costs of expanding and operating the network for all possible transition paths from Stage 3 to Stage 4, and the corresponding losses (Fig.8.2d). On this basis, it is then possible to derive the optimum path for the whole period as illustrated in Fig.8.2e.

Because of the simplicity of the dynamic programming computations these can be carried out by

STAGE 1 \ STATE	1	2	3
1	α	β	γ

α, β, γ = incremental costs + losses, stage 1 to stage 2

(a)

STAGE 2 \ STATE	1	2	3	4
1	A	B	C	D
2	–	E	F	G
3	–	–	I	J

STAGE 3

A, B, C, ··· = incremental costs + losses, stage 2 to stage 3

(b)

STAGE 2 \ STATE	1	2	3	4
1	$A+\alpha$	$B+\alpha$	$C+\alpha$	$D+\alpha$
2	–	$E+\beta$	$F+\beta$	$G+\beta$
3	–	–	$I+\gamma$	$J+\gamma$

STAGE 3

$B+\alpha$ --- = cumulated cost stages 1 to 3

(c) $\boxed{E+\beta}$ ---, minimum cumulated costs for given state, stage 3

STAGE 3 \ STATE	1	2	3	4	5
1	a	b	c	d	e
2	–	f	g	h	i
3	–	–	j	k	ℓ
4	–	–	–	m	n

STAGE 4

a, b, c, ···· = incremental cost + losses, stage 3 to stage 4

(d)

STAGE 3 \ STATE	1	2	3	4	5
1	$A+\alpha+a$	$A+\alpha+b$	$A+\alpha+c$	$A+\alpha+d$	$A+\alpha+e$
2	–	$E+\beta+f$	$E+\beta+g$	$E+\beta+h$	$E+\beta+i$
3	–	–	$C+\alpha+j$	$C+\alpha+k$	$C+\alpha+\ell$
4	–	–	–	$J+\gamma+m$	$J+\gamma+n$

STAGE 4

$A+\alpha+b$ --- = cumulated cost state 1 to state 4, with optimum path state 1 to state 3

(e)

$\boxed{C+\alpha+k}$ optimum solution

Figure 8·2 – COMPUTATIONAL STEPS IN THE DYNAMIC PROGRAMMING APPLICATION TO NETWORK PLANNING.

hand. The determination of the decision variables corresponding to a minimum cost, to get from one state (error level) to another within a given time interval, has been carried out for this study using a steepest descent technique based on Rosenbrock's algorithm for which there are several standard programs available. In the case of the present study the "CLIMB" program available at the University of Waterloo has been used.

A complete illustration of the application of the technique to Canadian conditions based on the assumptions described at the beginning of this sub-section is presented in the next sub-section.

8.2.3 Application to Canadian Network

The computations which simulated roughly the Canadian conditions were made considering an initial number of natural regime stations $N_s = 2000$, and an interest rate of 6%. The area with stationary hydrological regime was estimated to be 2,000,000 sq. mi., and the corresponding density of the existing network was thus estimated at one station per 1,000 sq. mi.

The error function was determined from the results obtained in Chapter 4 to be

$$E^2 = \frac{540000}{N_s} + \frac{1600}{N_t} + \frac{1936}{N_m} \text{ (per cent)} \qquad (8.2)$$

where N_s is the number of stations in the area considered, N_t is the average duration of operation in years and N_m is the average number of flow measurements per station. For the assumed initial conditions

---

* The conservative assumption of a value E_s^* corresponding to its higher possible value was accepted in these computations.

(N_s = 2000, N_t = 10 and N_m = 100) this corresponds to an error of 21.2%. The method of interpolation assumed is correlation between hydrological and physiographical characteristics. Switching from this method to a parametric model method would result in significantly higher benefits and this should be further investigated.

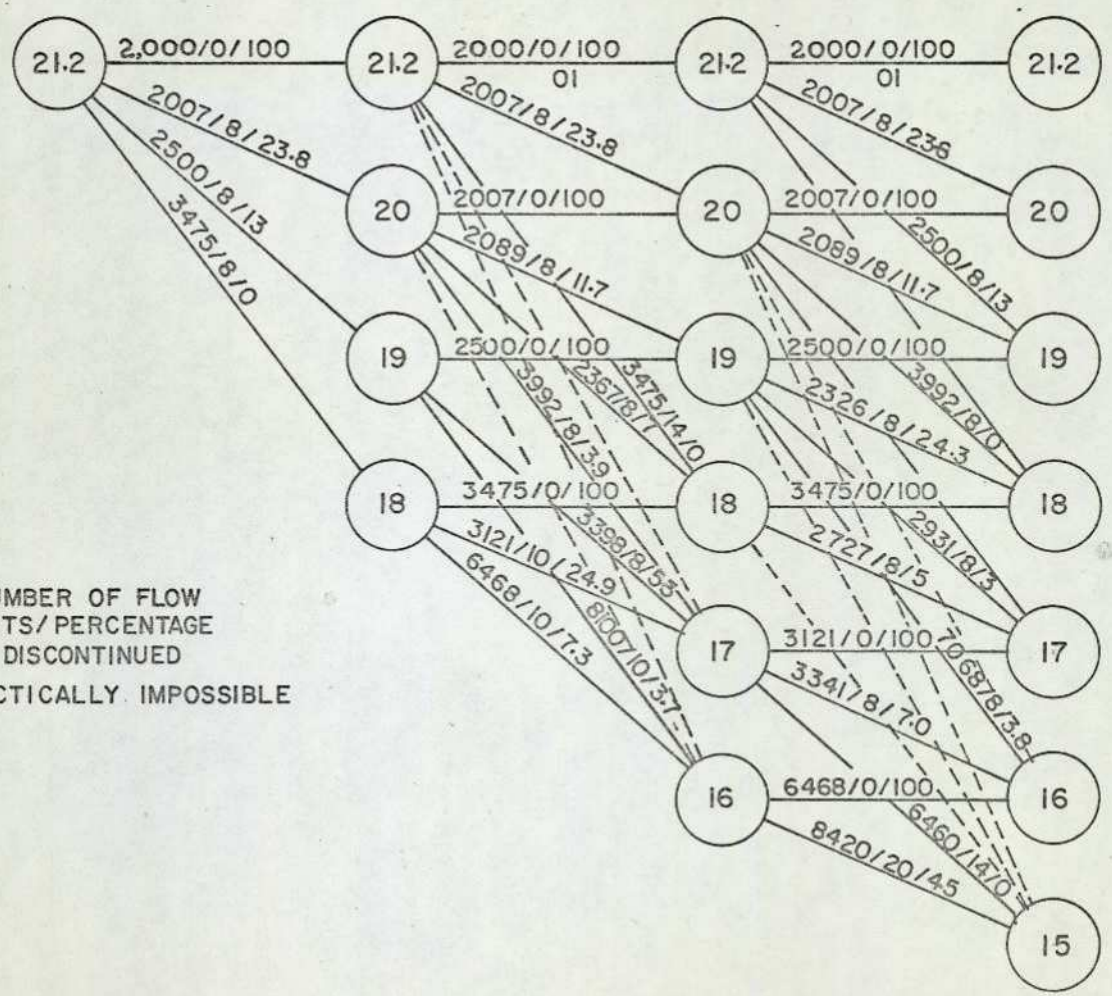
The network investment and operation costs were estimated assuming that the cost of installing a new station is $3,000, that of operating a station (exclusive of measurement costs) is $700/year, and that the cost of a flow measurement is $100. With these assumptions the cost of operating a station, assuming an average of 8 flow measurements per year is $1,500/year. New stations were assumed to be installed at the beginning of each period. All costs were "present worthed." Other network operation costs (central administration overhead, studies, etc.) were estimated at $2,000,000/year.

On this basis and using the techniques of optimization for achieving minimum cost in reducing the error of estimate by a given amount described in Section 6.2, the required number of supplementary stations, the required number of discontinued stations, and required number of flow measurements per year per station were estimated for the transition from the assumed initial state of Stage 1 to any state of Stage 2 and from any stage of Stage 2 to any state of Stage 3

as shown in Fig.8.3. The dotted lines in Fig.8.3 indicate impossible transitions, i.e. transitions which require an increase in accuracy which is not possible within the corresponding time interval.

On the basis of the results indicated in Fig.8.3, the corresponding costs of the network were computed and "present worthed" using the unit costs indicated above. The results of these computations are indicated in Fig.8.4. In this figure we have indicated by dotted lines and eliminated from consideration, in addition to the impossible transitions, the transitions implying a complete interruption of the network operation for a period of 5 years or more. Next, on the basis of the errors corresponding to the various states and using Eq.8.1 with a = 0.0016, we have computed the losses due to errors in the hydrological estimates affecting the investment and operation of projects related to water resources for the periods 1971-1975 and 1976-1980. The estimated investment considered for the first period was 11.2 billion and for the second, 13.8 billion. The results of these computations are summarized in Fig.8.5.

Finally, the cost of expanding and operating the network was added to the value of the losses due to errors in hydrological estimates affecting projects related to water resources, for each transition from the initial state, Stage 1, to the various states of Stage 2, and further, from the states of Stage 2 to



LEGEND :

STATIONS/NUMBER OF FLOW MEASUREMENTS/ PERCENTAGE OF STATIONS DISCONTINUED

---- PRACTICALLY IMPOSSIBLE

Figure 8·3 — OPTIMUM CONDITIONS FOR TRANSITION FROM A STATE OF A GIVEN STAGE TO ANOTHER STATE OF THE NEXT STAGE.

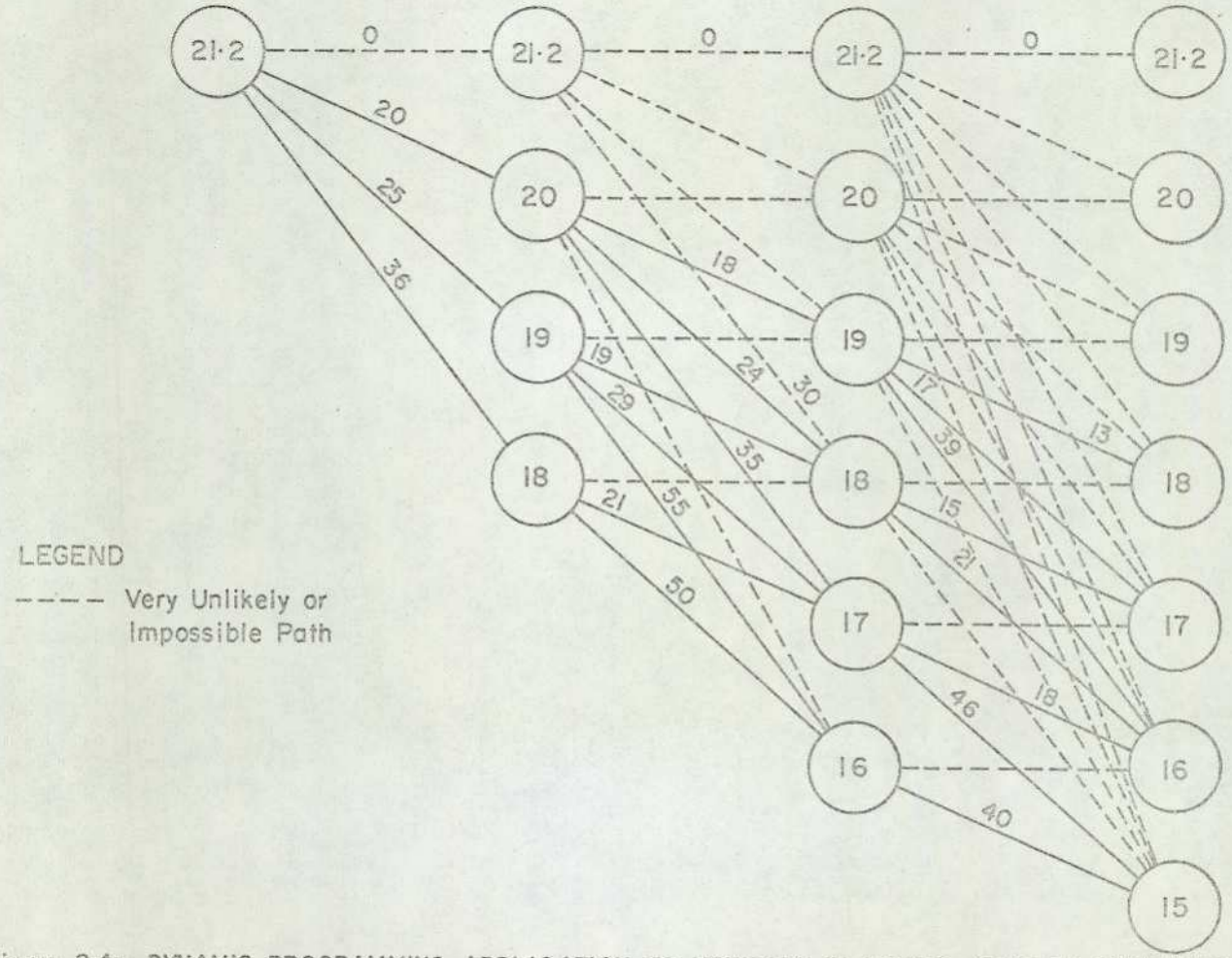


Figure 8-4 – DYNAMIC PROGRAMMING APPLICATION TO NETWORK PLANNING – COST OF EXPANDING AND OPERATING THE NETWORK IN MILLION OF DOLLARS.

the states of Stage 3. The results of these computations are summarized in Fig.8.6. On the basis of these results the optimum minimum cost paths for arriving at any state of Stage 3 were determined. These paths are indicated with thick lines in Fig.8.6.

On the basis of the results thus obtained, it was possible to determine the number of stations and the other network parameters which would characterize the network in any state (error level) at Stage 3, if an optimum path is selected to reach this state. With this information it was then possible to compute the optimum (least cost) network characteristics for transitions from any state of Stage 3 to any state of Stage 4 (Fig.8.3) and the corresponding present worthed costs (Fig.8.4).

The losses corresponding to transitions from any state of Stage 3 to any state of Stage 4 were computed in a manner similar to that described for losses from Stage 2 to Stage 3, assuming an investment of 16.4 billion dollars. The present worth of the estimates of these losses is shown in Fig.8.5.

The cost of expanding and operating the network was added to the value of the losses due to errors in hydrological estimates for each transition from any of the states of Stage 3 to any of the states of Stage 4 as shown in Fig.8.6. On this basis, it was possible to choose the optimum (minimum cost and losses) for a complete path from Stage 1 to Stage 4. This corresponds to a decrease of the error of the standard

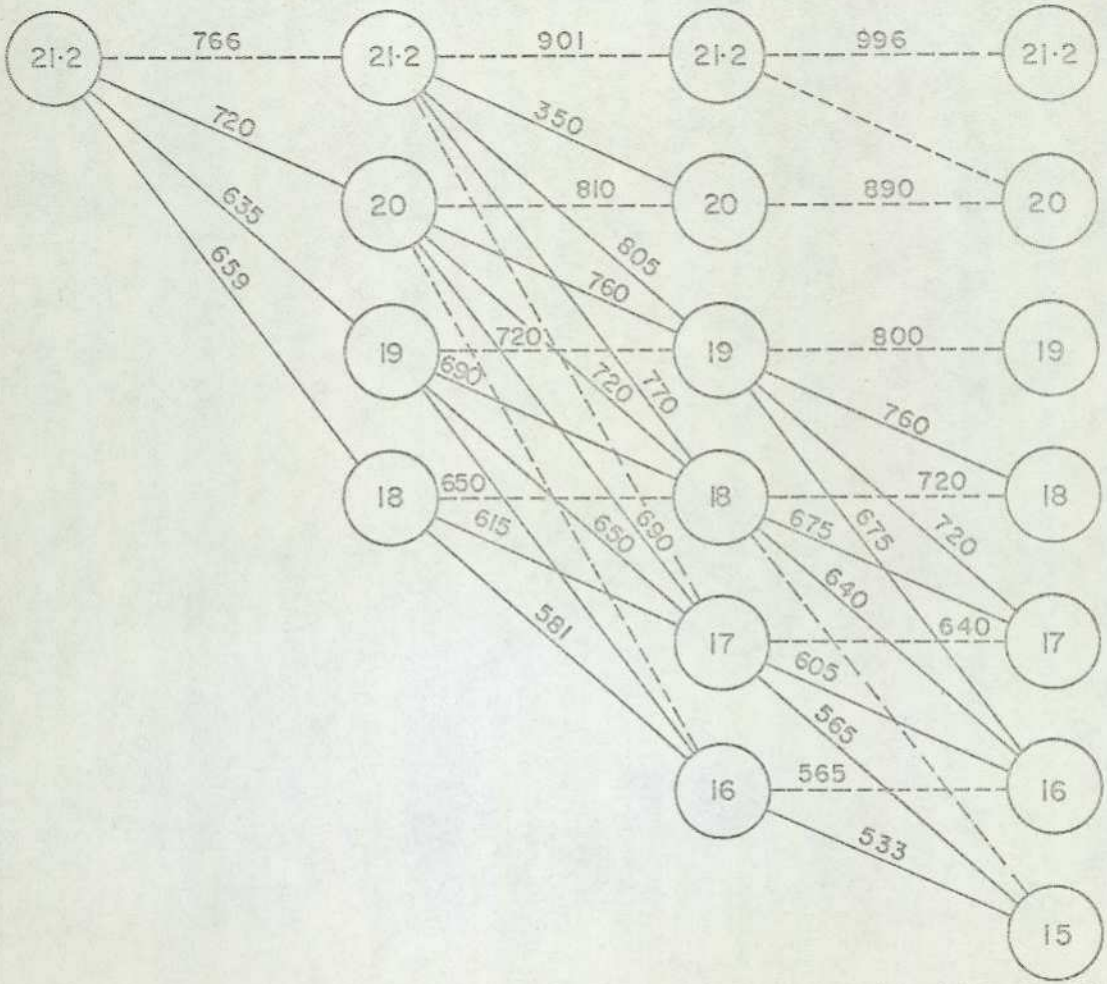


Figure 8·5 – DYNAMIC PROGRAMMING APPLICATION TO NETWORK PLANNING - LOSSES CORRESPONDING TO VARIOUS STATES (ERROR LEVELS) IN MILLION OF DOLLARS.

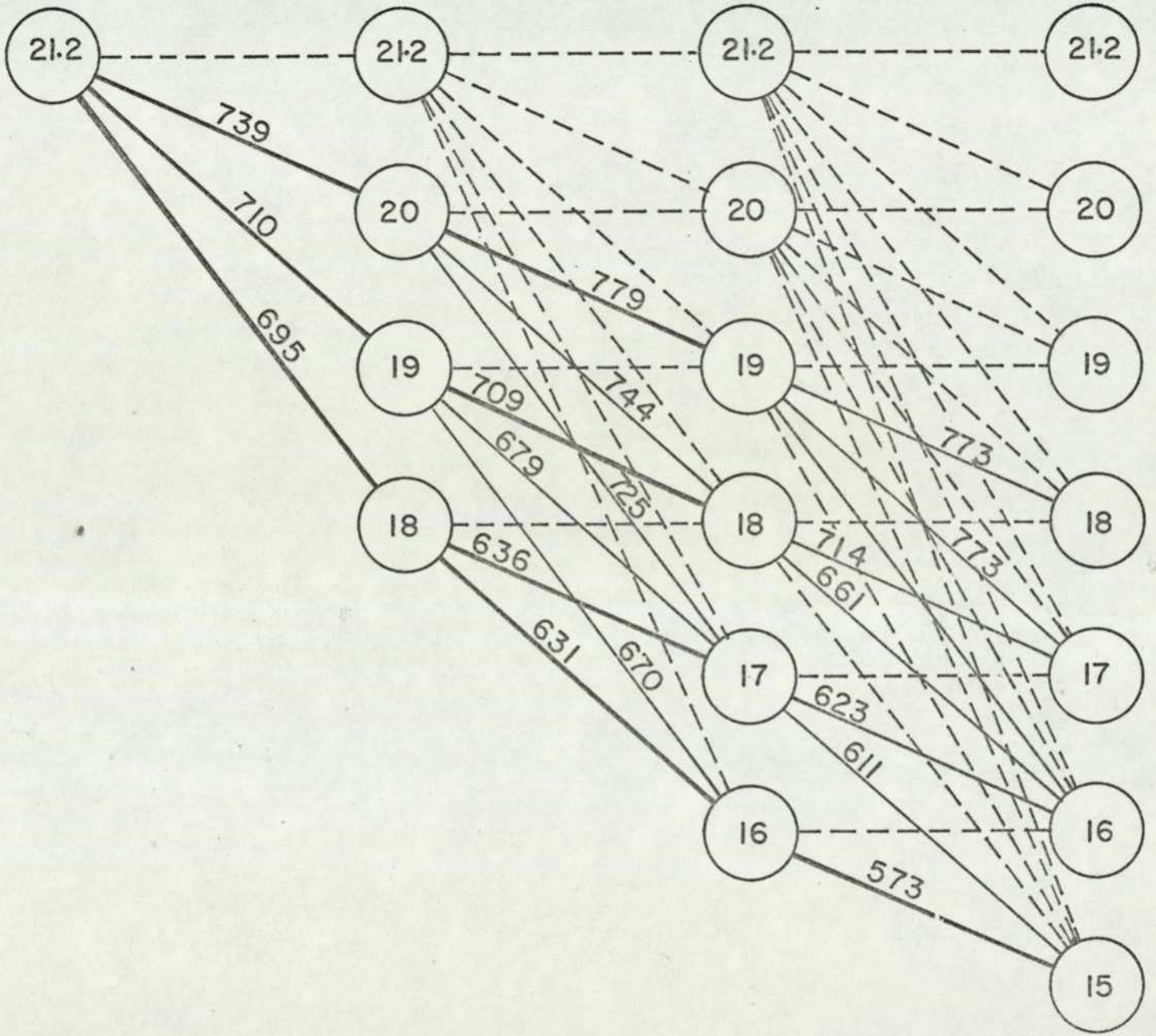


Figure 8·6 — DYNAMIC PROGRAMMING APPLICATION TO NETWORK PLANNING – LOSSES AND OPERATION COSTS CORRESPONDING TO VARIOUS TRANSITION PATHS IN MILLIONS OF DOLLARS.

deviation of the annual flow from 21.2% in Stage 1 to 18% in Stage 2, to 16% in Stage 3 and 15% in Stage 4. A complete discussion of the implications of these results is given in Table 8.5, which indicates the economics of the network operation and the main decision variables considering one, two or three five-year periods. The table indicates benefits/cost ratios from the operation of the network. This ratio increases significantly as the period considered is extended, since most benefits from investment in the network will be reaped in the future. The results indicate that the obtaining of the optimum returns from the network would require an expansion of the network by about 1500 stations for the period 1971-1975, a further addition of about 3000 stations for the period 1976-1980, while discontinuing about 250 stations, and finally for the period 1981-1985 a displacement of 3000 stations from old to new sites. Since the results indicate the necessity of a significant effort for the network expansion during the first two periods, we have examined further in sub-section 8.2.5 the effect on these conclusions of various departures from the basic assumptions, to ascertain their sensitivity to the errors in the basic assumptions.

8.2.4 Optimum Path for Reaching a Certain Level of Accuracy

A by-product of the previous analysis is the possibility of determining by dynamic programming

TABLE 8.5

Results of Optimization by Dynamic Programming of Simulated Canadian Natural Regime Network for the 1971 - 1985 Period

(All values in millions of dollars, "present worth" to 1971, except net benefit per dollar which is dimensionless and decision variables which have the stated dimensions, N_1 = 2000 stations, interest rate: 6% p.a.)

	Results for: the 1971-75 period only	the 1971-80 period only	the 1971-80 period only
Losses due to errors in standard deviation of mean annual flow assuming that hydrological network ceased operation in 1971	766	1667	2663
Losses due to errors in standard deviation of mean annual flow, assuming hydrometric network operating at optimum level	659	1240	1773
Cost of operating the network at optimum level	36	86	126
Net benefits	71	341	754
Net benefits per dollar spent for hydrological network operation (%)	197	396	599
Decision variables: Number of supplementary stations required at end of period	1475	4463	6420
No. of stations to be discontinued in this period	0	253	2911
No. of flow measurement per station/year	8	10	18
Error of standard deviation of annual flow at end of period (%)	18	16	15

techniques the optimum (least cost) path for reaching stepwise a certain level of accuracy during a certain period of time or,conversely, the optimum level of accuracy which can be obtained for a given amount of money over a given period of time in non-development areas. For such areas, one has to consider only the cost of network operation and expansion in relation to the error level and arrive at the cost of reaching a certain accuracy in a given period of time. The solution of this type of problem is illustrated clearly by the following dynamic program examples, based on the data obtained from Fig.8.4.

The example assumes that one is interested in reaching an error level of 17% during the period 1971-1980. As indicated in Fig.8.4 one can reach this level along the following paths and at the following costs:

Path (Errors %)			Costs (millions of dollars)		
1971	1975	1980	1971-75	1975-80	Total
21.2	20	17	20	35	55
21.2	19	17	25	29	54
21.2	18	17	36	21	57

It is obvious that in this case the path corresponding to a gradual decrease of the error from 21.2 to 19 and 17% would be optimum, although the difference from a 21.2 - 20 - 17% path is not large. However, since the optimum path obtained is also closer to the

path which would be obtained in the case when economic development is taking place in the area, the difference between the two solutions becomes more significant from a practical point of view.

Thus, the dynamic programming technique appears to be also applicable to areas of the stationary regime network in which economic development is not taking place within the planning horizon, but for which certain accuracy objectives or budgetary allocations have been established.

8.2.5 Sensitivity Analysis

Since many of the basic assumptions may be subject to significant errors, a sensitivity analysis was carried out to estimate the influence of these errors on the results, and make possible the selection of the significant assumptions which should be further investigated to consolidate the preliminary results obtained above. The sensitivity analysis examined the effect on the basic recommendations on network expansion of the following factors:

a) variation in investment in projects related to water resources;

b) variation in cost of expansion and operation of the network;

c) variation in interest rates;

d) variation in the relationship between losses due to errors in the estimate of hydrological characteristics affecting the design and operation of projects related to water resources and the size of such errors.

In addition, the sensitivity analysis examined the effect on optimum network expansion of cyclicity in economic development.

Because of the large number of cases to be analysed and of the corresponding enormous number of computations, the sensitivity analysis was carried out using an approach different than that used previously, namely:

- only three stages were considered
- the states were characterized by number of stations rather than error level
- the number of flow measurements per year were assumed constant (10) since it did not appear to be a significant decision variable and it showed little variation in the analysis described in sub-section 8.2.3.
- no allowance was made for station discontinuation.

With these simplifications the computations were carried out otherwise as described in sub-section 8.2.3 for a number of 81 cases representing combinations of:

- 3 interest rates (4%, 6%, 8%);
- 3 levels of costs of network expansion and operation (as assumed in sub-section 8.2.3, 80% of costs assumed in sub-section 8.2.3, 120% of costs assumed in sub-section 8.2.3);
- 3 levels of investment (as assumed in sub-section 8.2.3, 70% of investment assumed in sub-section 8.2.3, 130% assumed in sub-section 8.2.3);
- 3 levels of relationships between errors in estimation of hydrological characteristics and investment and operation losses in projects related to water resources (50%, 100% and 150% of losses assumed in sub-section 8.2.3.)

The results of the sensitivity analysis are summarized in Table 8.6, which gives the optimum number of stations for the transition from Stage 1 to Stage 3 for each combination of assumptions regarding the interest rate, cost of expansion and operation of network, level of investment, and relationship between error and losses. The results summarized in Table 8.6 indicate the following:

a) The optimum total number of stations for the accepted assumptions oscillates around 5500 for the total period.

b) A reduction in network expansion and operation costs of about 20% does not affect significantly the number of stations, whereas an increase of about 20% will reduce the optimum number by about 500 stations (10%).

c) A reduction in the loss function of 50% will reduce the optimum number by about one thousand (20%) and an increase of the loss function of 50% will increase the number of stations by 500 (10%).

d) Changes in interest rate affect very little the optimum number of stations.

e) A decrease of 30% of the economic development forecast will reduce by about 500 (10%) the optimum number of stations and an increase of 30% would have the opposite effect.

f) The detailed computations indicate that even under most conservative assumptions (highest interest rate, lowest investment level, highest network costs, lowest relationship between error and losses) the benefit investment ratio is higher than 100%, i.e. each dollar invested in the network produces one dollar return during the first five years of operation.

8.2.6 Sensitivity Analysis Assuming Cyclic Economic Development

Since Canada in common with many other countries presents a cyclic economic development, a second sensitivity analysis was done to examine the effects of

SENSITIVITY ANALYSIS: VARIATION OF OPTIMUM NUMBER OF STATIONS WITH VARIATION OF BASIC ECONOMIC ASSUMPTIONS

INTEREST RATE	LOSSES	COSTS OF OPERATING AND EXPANDING NETWORK								
		Lower by 20% than costs in section 3			As in Section 3			Higher by 20% than costs in Section 3.		
		INVESTMENTS			INVESTMENTS			INVESTMENTS		
		Lower by 30%	As in Section 3	Higher by 30%	Lower by 30%	As in Section 3	Higher by 30%	Lower by 30 %	As in Section 3	Higher by 30%
4%	Higher by 50% than in Sec.3	6000	6000	6000	5500	6000	6000	5000	6000	6000
	As in Sec.3	5000	6000	6000	5000	5500	6000	4500	5000	5500
	Lower by 50% than in Sec.3	4500	5000	5000	4000	4500	5000	4000	4000	4500
6%	Higher by 50% than in Sec.3	6000	6000	6000	5500	6000	6000	5000	5500	6000
	As in Sec.3	5000	5500	6000	5000	5500	6000	4500	5000	5500
	Lower by 50% than in Sec.3	4000	4500	5000	4000	4500	5000	3500	4000	4500
8%	Higher by 50% than in Sec.3	5500	6000	6000	5500	6000	6000	5000	5500	6000
	As in Sec.3	5000	5500	6000	5000	5500	5500	4500	5000	5500
	Lower by 50% than in Sec.3	4000	4500	5000	4000	4500	4500	3500	4000	4500

TABLE 8-6

cyclicity on the optimum number of stations. In this analysis the ten year economic development forecast was first considered to be 30% lower than that given by Bruce and Maasland for the first five-year period, and 30% higher for the second five-year period. In a second set of computations the inverse situation was considered: for the first period of five years the forecast was increased by 30% relatively to the reference values, and in the following period the corresponding figures were reduced by 30%. All other variables considered in the preceding analysis were allowed to vary in exactly the same way as described in sub-section 8.2.5. The results of this sensitivity analysis are summarized in Table 8.7, and indicate the following:

a) The optimum number of stations oscillates in this case also around 5500.

b) The variation in the loss function of ±50% has the same effects as previously indicated (sub-section 8.2.5).

c) Interest rates have little effect on the optimum number of stations.

d) In the case of a cyclic development of the economy with a recession followed by a boom it is more advantageous to have a larger number of stations installed in the first period (about 500 stations) than in the case of a

boom followed by a recession. As is known, the reverse policy is usually being applied.

It is interesting to note that the computations described above indicate that if the planning horizon is only 10 years, the results of the computations show invariably that the number of stations should be maintained approximately constant during the second five-year period, whereas an increase is indicated for this period if a 15-year period is considered. This difference is related to the fact that an increased accuracy is obtained from new stations only after a certain period of operation (about 5 years) and consequently, by truncating the study period to 10 years only, the effects of the expansion of the network during the second five-year period are not accounted for. Because of this it is advisable that a plan for the expansion of the network for a period of n years should use a dynamic programming analysis covering n+5, or, preferably n+10 years.

8.2.7 Recommendations with Respect to the Application of the Dynamic Programming Technique to Network Planning

In the light of the results obtained from the sensitivity analysis it appears that one could make the following recommendations:

1. To expand vigorously the natural regime hydrological network in areas where economic development is planned, since such expansion is justified even under most conservative assumptions regarding the economic and technical conditions.

SENSITIVITY ANALYSIS: EFFECT OF ECONOMIC CYCLES
COSTS OF OPERATING AND EXPANDING NETWORK

INTEREST RATE	LOSSES		Lower by 20% than costs in section 3			As in Section 3			Higher by 20		
			INVESTMENTS			INVESTMENTS			INVESTMENTS		
			First period 30% less. Second period 30% more.	As in Section 3	First period 30% more. Second period 30% less.	First period 30% less. Second period 30% more.	As in Section 3	First period 30% more. Second period 30% less	First period 30% less. Second period 30% more.	As in Section 3	First period 30% more. Second period 30% less
4	Higher by 50%		6000	6000	6000	6000	6000	5500	6000	6000	5500
	As in Sec.3		6000	6000	5500	6000	5500	5000	5500	5000	5000
	Lower by 50%		5000	5000	4500	4500	4500	4000	4500	4000	4000
6	Higher by 50%		6000	6000	6000	6000	6000	5500	6000	5500	5500
	As in Sec.3		6000	5500	5500	5500	5500	5000	5500	5000	5000
	Lower by 50%		5000	4500	4500	4500	4500	4000	4500	4000	4000
8	Higher by 50%		6000	6000	6000	6000	6000	5500	6000	5500	5500
	As in Sec.3		6000	5500	5500	5500	5500	5000	5500	5000	5000
	Lower by 50%		5000	4500	4500	4500	4500	4000	4500	4000	4000

TABLE 8-7

2. To carry out studies to define as precisely as possible the relationship between error in estimating hydrological variables and the related losses in investment and operation for various types of projects since it is the most important factor in the selection of optimum decision variables.

3. To refine further the application of the dynamic program technique to network planning in such a manner as to eliminate the constraints and simplifications encountered in the current application. This refers particularly to the constraints imposed in defining the states which in the current applications are limited in number and have to take given values and to the simplifications included in the sensitivity analysis.

9. CONCLUSIONS AND RECOMMENDATIONS

The main conclusion of this thesis is that hydrological networks can be planned according to given basic criteria. The planning process here envisaged covers the number and regional distribution of gauging stations, the method of interpolation of information, the duration of observation and the frequency of measurements. However, additional research is needed before it will be possible to obtain the full return from the planning principles developed here. In view of this, we give in Section 9.1 the conclusions and recommendations for current network planning and in Section 9.2. recommendations for further research.

9.1 Conclusions and Recommendations for Network Planning

The hydrological network should be planned in such a way as to incorporate the possibility of synthesizing hydrological data and information at any point (gauged or ungauged) in the most efficient manner, i.e. at least cost per unit of supplementary accuracy achieved.

The hydrological network should be considered as consisting of three sections. Section 1, the natural regime section, is the basic one which provides data for broad national planning to hedge against the development of unanticipated needs, and to provide background information for the design of more intensive data collection programmes. Sections 2 and 3 correspond to the levels of information required for water resources and water-resource management respectively. It is

clear that Sections 2 and 3 develop on the basis of Section 1.

For Section 1 the basic principles of network planning are:

1) A network operated on a fixed budget should be planned to minimize the error of estimation of the hydrological variables involved;

2) A network operated on a criterion of a given minimum acceptable accuracy should be planned for a minimum cost.

In the case of Sections 2 and 3, the costs to be minimized include network costs and also the expected opportunity loss, i.e. the difference between net returns from a project designed and operated on the basis of the available level of accuracy of the data and on the basis of a higher level of accuracy which can be obtained by additional data collection.

Errors in the values of hydrological data have three basic sources: errors of sampling in space, errors of sampling in time, and errors of measurement. Errors of sampling in space and in time can be reduced by adequate planning of the network. Errors of measurement are related mainly to the operation of the network.

Adequate planning of sampling in space requires hydrological regionalization to avoid the over-sampling of basins having similar characteristics, and under-sampling of basins having distinctive characteristics.

Hydrological regionalization can be achieved by interfacing a physiographic, land-use, land-cover data bank with a hydrometeorological data bank.

Adequate planning of sampling in space requires also the use of efficient techniques of data transfer. A parametric model based on the square grid technique, although its potential has not yet been fully explored, appears to be one of the most efficient data transfer techniques. The application of the technique indicates the usefulness of the interfacing of data banks and the necessity of co-ordinated planning of inter-related networks.

The efficiency of sampling in time is reduced by the autocorrelation of the time series of recorded flows. Efficient sampling requires that, other conditions being comparable, the river sections selected for gauging sites should be located on rivers having low autocorrelation, i.e. rivers with small areas and small natural storage. However, this consideration should not detract from the necessity of obtaining an adequate space sampling in which rivers of various autocorrelation characteristics should be included. Adequate technology for removing the autocorrelation effect from the time series of flow data could also improve the efficiency of time sampling. This emphasizes again the potential usefulness of good hydrological models.

If station locations are selected according to the indications obtained from the hydrological regionalization analysis and in such manner as to reduce as much as possible the effect of auto-correlation, the space sampling error is primarily a function of the number of stations and of the method of data transfer used, and the time sampling error is primarily a function of the length of record and cross correlation (which in turn is a function of station density). However, it is extremely difficult to obtain realistic analytical relationships between error of sampling in space and time on one hand, and station density and length of record on the other. Because of this, it is necessary to use in network planning calculations empirical relationships between the above mentioned variables developed on the basis of data obtained from densely gauged areas and from long-term flow records.

Planning of the natural regime section of the hydrological network can be done by means of a variety of techniques of which the steepest descent techniques are currently the best. However, these techniques are still not very efficient and it is desirable to develop more efficient techniques for this purpose.

The optimum number of stations obtained in planning the natural regime section of the hydrological network is very sensitive to the type and reliability of the data transfer technique considered in the planning

analysis. The number of supplementary stations which correspond to the optimum conditions decreases sharply with an increase in the sophistication and reliability of the technique of data transfer. It appears, therefore, that returns from hydrological research carried out to increase the reliability of such techniques are important.

Planning of the transitional regime section of the hydrological network may be achieved by means of dynamic programming. The application of the technique requires the availability of economic forecasts, and of relationships between error of estimate of hydrological characteristics and losses incurred by various water resources or related projects because of such errors. Such relationships are available so far only in a very general and theoretical form, and work to obtain realistic relationships of this type should be encouraged. Because of the lack of precision of the relationships between errors and related losses, planning for the transitional regime section has to be carried out in the framework of a sensitivity analysis.

The results of sensitivity analyses for the planning of the transitional regime section indicate that the main planning variable -- the supplementary number of stations -- is sensitive only to a limited extent to the variations in the error-loss function and to the economic forecast. This seems to indicate

that the method could be used for a preliminary planning analysis even without precise relationships between errors and losses. However, the results are sensitive to the length of the planning period considered: the number of stations decreases as the period of time over which the analysis is extended becomes shorter. The analysis indicates that the planning period should be extended by 5 to 10 years beyond the time horizon for which planning is being attempted, e.g. 15 years, to obtain stable indications for the first five years.

The sensitivity analysis indicate that in the case of a cyclic economic development it would be preferable to increase the number of stations during the recession period of the cycle and vice-versa, because of the lag between the time when new stations are installed and the time when the benefits from such new stations are reaped. The practice of discontinuing stations during economic recession periods is very inefficient.

The annual costs to the international community of all hydrological networks is of the order of several hundred millions of dollars. Implementation of the planning techniques described in this thesis, particularly in conjunction with the parametric model - square grid technique of data transfer and of other techniques based on interfacing data banks, can improve significantly the efficiency of

the hydrological betworks of various countries and produce savings of many millions of dollars in the network and large additional returns from water resources and related projects.

Since high benefit/cost ratios are obtained from the hydrological network even under most conservative assumptions in areas in which economic development is planned, vigorous development of the network taking into account both the physiography and the requirements of the planning and design process is recommended.

9.2. Recommendations for Further Studies.

In each Chapter a number of detailed suggestions for further research have been made and need not be repeated here. The following recommendations, however, are of an altogether higher order of importance.

This thesis indicates that it is possible to increase significantly the efficiency of hydrological networks by interfacing hydrometeorological, physiographical, land-cover, and land-use data banks, and by applying suitable models for information interpolation. This increased efficiency results in significant savings if the network is planned along the lines of the methodology developed in this thesis. In connection with the application of this methodology it is necessary to pursue further research in three main directions.

The first direction leads to research to ascertain if it would be possible to further increase the efficiency

of the hydrological network by improving the models for information interpolation. Improvements could probably by obtained by reducing the parametric and/or statistical components of the models, by decreasing the time intervals considered in the models; and by using areally distributed models. However, in order to obtain new models of practical significance, it would be necessary to obtain models which could use data from interfacing data banks without excessively increasing the requirements with respect to these data banks and for which the monetary and time expenses would compare favourably with those involved in the existing models.

The second direction of research to be pursued in relation to the major finding of this thesis relates to the planning of other geophysical networks and surveys. Research should be carried out to investigate the possibility of using the methodology developed in this thesis for planning co-ordinated networks and surveys. Attention should be paid in the framework of such research to the interfacing of ground networks with remote sensing surveys which, in the opinion of this writer, hold great promise in relation to co-ordination and increased linking of geophysical surveys.

The third direction of research which should be followed up in relation to the findings of this thesis is on the possible benefits of international co-operation in planning and setting up geophysical and socio-economic

surveys and data banks. Research should include the comparison of benefits of large-scale (continental or world-wide) systems with those of national systems, in relations to economy of scale, increased possibilities of information interpolation, and an earlier introduction of advanced technology in developing countries.

REFERENCES

Alexeev, G.A. (1971). Optimal Spatial Interpolation and Its Errors in Case of Various Distances between Observational Points of Watersheds Centres, Paper presented at the IASH International Symposium on Mathematical Models in Hydrology, Warsaw, 1971.

Barton, B.M.J. (1971) Synthetic Generation of Monthly Runoff Records for Ungauged British Catchments, Ph.D. Thesis, Department of Civil Engineering, University of Newcastle.

Benson, M.A. (1962) Factors Influencing the Occurrence of Floods in the Southwest, U.S. Geological Survey Water Supply Paper 1580-8, Washington, D.C.

Benson, M.A. and R.W. Carter (1970). Areal Extension of Stream-flow Information by Multiple Regression Techniques, Proceedings of the AWRA Symposium on Water Balance in North America, Banff, 1969, Published by AWRA, Illinois, Urbana.

Benson, M.A. and N.C. Matalas (1967). Synthetic Hydrology Based on Regional Statistical Parameters, Water Resources Research, Vol. 3, No. 4.

Bruce, J.P. and D.E.L. Maasland (1968). Water Resources Research in Canada, Special Study No. 5, Science Secretariat, Queen's Printer, Ottawa.

Burkham, D. and D. Dawdy (1969). Error Analysis of Stream-flow Data for Alluvial Streams, U.S. Geological Survey Professional Paper 655 C, Washington, D.C.

Carter, R.W. and M. A. Benson (1969). Concepts for the Design of Stream-flow Data Programs, U.S.G.S. Open File Report.

Cavadias, G.S. and S.I. Solomon (1972). Future Developments in Hydrologic Networks, Paper V1-2-1 in Casebook on Hydrological Network Design Practice, Published by WMO, W.B. Langbein, Editor, Geneva.

CENTO, (1969). Seminar on Evaluation of Water Resources with Scarce Data, Teheran, 1969, Imprimerie Catolique, Beirut.

CEQEAU (1971). Rationalisation du réseau hydrométrique du Québec, Report for Environment Canada, University of Quebec, Québec.

Chapman, T.G. (1972). The Australian Representative Basins Programme, Section III - 3.1 in Casebook on Hydrological Network Design Practice, Published by WMO, W.B. Langbein, Editor, Geneva.

Chapman L.J. and D.F. Putnam (1960). The Physiography of Southern Ontario, University of Toronto, Toronto.

Coulson,A. (1967). Estimating Runoff in Southern Ontario , Technical Bulletin No. 7, Department of Energy, Mines, and Resources, Ottawa.

Cuthbert, D.R. (1970). The Application of the Square Grid Technique for the Determination of Regional Hydrologic Relationships for the Province of New Brunswick, Canada Department of Fisheries and Forestry, Ottawa.

Davis,D.R. and W.M. Dvoranchik, (1970). Evaluation of the Worth of Additional Data, University of Arizona, Department of Systems Engineering, Tucson, Arizona.

Dawdy, D.R., H.E. Kubik, and E.R. Close (1970). Worth of Stream-flow Data for Project Design - A pilot study, Water Resources Research, Vol. 5, No. 4.

Dawdy, D.R., M.E. Moss and N.C. Matalas (1972). Application of Systems Analysis to Network Design, paper IV - 4.1, in Casebook on Hydrological Network Design Practice, Published by WMO, W.B. Langbein, Editor, Geneva.

Dickinson, W.T. (1967). Accuracy of Discharge Measurements, Hydrology Paper No.20, Colorado State University, Fort Collins, Colorado.

Dubreuil, P. (1969). Nature et répartition des stations hydrometriques des réseaux en France au début de 1968, Cahiers de l' ORSTOM Série Hydrologie, Vol.6, No. 2.

Fiering, M. B. (1965). An Optimization Scheme for Gaging, Water Resources Research, Vol. 1, No. 4.

Fisher, R.A. (1949). The Design of Experiments, Oliver and Boyd, London, England.

Foyster, A.M. (1973). Application of the Grid Square Technique to Mapping of Evapotranspiration, Journal of Hydrology, Vol. 19, pp 205-226.

Goldstein, M. (1968). Errors in Current Meter Measurements (in Romanian) Hidrotehnica, No. 11.

Gupta, S.K. (1974). A Distributed Digital Model for Estimation of Flows and Sediment Load from Large Ungauged Watersheds, Ph.D Thesis, University of Waterloo, Waterloo, Ontario.

Horton, R. E. (1932). Drainage Basin Characteristics, Transactions of the American Geophysical Union, Vol. 13, pp 350-361.

Howe, C. W. (1971). Benefit - Cost Analysis for Water System Planning, American Geophysical Union - Water Resources Monograph 2, Washington, D.C.

Ingledow T. and Associates Limited (1970). Hydrometric Network Plan for the Provinces of Newfoundland, New Brunswick, Nova Scotia and Prince Edward Island, Report for Canada Department of Energy, Mines and Resources.

Jaakson, R. (1970). A Method to Analyse the Effects of Fluctuating Reservoir Water Levels on Shoreline Use, Water Resources Research, Vol. 6, No. 2.

Jordan, R.P. and E.R. Hedman (1970). Evaluation of Stream Gaging Program in Kansas U.S.G.S. Open File Report.

Karasev, I.F. (1968). Principles for Distribution and Prospects for Development of a Hydrologic Network, Soviet Hydrology No. 6, Translated from Trudy GG 1, Vol. 164.

Kisiel, C.C. (1971). Efficiency of Parameter and State Estimation Methods in relation to Models of Lumped and Distributed Systems. United States - Japan Bi-Lateral Seminar on Systems Approach to Hydrology, Honolulu.

Kohler M.A. (1958). Design of Hydrological Networks, WMO Technical Note No. 25.

Kohler M.A. (1965). Co-ordinated Network Planning, WMO/IASH Symposium on the Design of Hydrological Networks, Quebec, 1965, Vol. I, IASH Publication No. 60, Gentbrugge.

Kuzin, P.S. (1965). Methodology and Practical Techniques of Hydrological Regionalization (in Russian, O metodike i prakticheskikh priemakh gidhologicheskogo raionirovaniya), Trudy GGI, Vol. 126.

Langbein, W.B. (1948). Annual Runoff in the United States, Circular J2, U.S. Geological Survey, Washington, D.C.

Langbein, W.B. (1954). Stream Gauging Networks,IASH Publication No. 38.

Langbein, W.B. (1960). Hydrologic Data Networks and Methods of Extrapolating or Extending Available Hydrologic data, Hydrologic Networks and Methods, WMO Flood Control Series,No. 15.

Langbein, W.B. (1965), National Networks of Hydrological Data, Proceedings of the IASH Symposium on Design of Hydrological Networks, Quebec, 1965, Vol. II, IASH Publication No. 68, Gentbrugge.

Langbein, W.B. (1972 a),Evaluation of Stream Gauging Program in Kansas, a Case Example Abstracted from a 1970 eport by P.R. Jordan and E.R. Hedman, in Casebook on Hydrological Network Design Practice, Published by WMO, W.B. Langbein, Editor, Geneva.

Langbein, W.B. (1972b), Riverscope and River Environmental Surveys, Section II-61 in Casebook on Hydrological Network Design Practice, published by WMO, W.B. Langbein, Editor, Geneva.

Leopold, L.B. (1969). Quantitative Comparison of some Aeshetic Factors Among Rivers, U.S. Geological Survey Circular 620. Washington, D.C.

Linsley, R.K. (1965). Symposium on Hydrometeorological Networks, a Summary, WMO/IASH Symposium Design of Hydrological Networks, Quebec, 1965 Vol. II, IASH Publication No. 68, Gentbrugge.

Linsley, R.K. and N.H. Crawford (1965). Co-ordination of Streamflow and Precipitation Networks, WMO/IASH Symposium on the Design of Hydrological Networks, Quebec, 1965; IASH Publication No. 67, Gentbrugge.

Matalas,N.C. (1967). Optimum Gauging Station Location, I.B.M. Scientific Symposium on Water and Air Resources, Yorktown Heights, N.Y.

Matalas,N.C. and E.J. Gilroy (1968). Some Comments on Regionalization in Hydrologic studies, Water Resources Research, Vol. 4, No. 6.

Moss M.E. (1970). Optimum Operating Procedure for a River Gaging Station Established to Provide Data for Design of a Water Supply Project, Water Resources Research Vol. 5, No. 6.

Nace, R.L. (1969). Hydrological and Related Data Programs in the United States. (See also Conclusion Remarks by same author), CENTO Seminar on Evaluation of Water Resources with Scarce Data, Teheran, 1969, Imprimerie Catolique, Beirut.

Nordensen, T. J. (1968). Preparation of Co-ordinated Precipitation, Runoff and Evaporation Maps, WMO/IHD Project No. 6, Geneva.

Obedkoff, W. (1970) A Study of the Grid Square Method for Estimating Mean Annual Runoff, M.A. Sc. Thesis, University of British Columbia.

Pentland, R.L., and D.R. Cuthbert (1971). Operational Hydrology for Streams by the Square Grid Technique, Water Resources Research, Vol. 7, No. 2.

Rainbird, A.F. (1965). Discussion on the paper by Linsley, R.K. and N.H. Crawford, Co-ordination of Stream-flow and Precipitation Networks, WMO/IASH Symposium on the Design of Hydrological Networks, Vol. II, p.644, IASH Publication No. 68. Gentbrugge.

Rodda, J.C., W.B. Langbein, A.G. Kovzel, D.R. Dawdy and K. Szesztay (1969) Hydrological Network Design, Needs, Problems and Approaches, Reports on WMO/IHD Projects, Report No. 12, World Meteorological Organization, Geneva.

Roche, M. (1967). Composition et rationalisation du réseau hydrométéorologique, Cahiers de l'ORSTOM, Vol. 4. No. 3

Rosenberg, H.B. and R.L. Pentland, (1966) Accuracy of Streamflow Records, Proceedings of the 1966 Eastern Snow Conference.

Shawinigan Engineering Company Limited (1970). Hydrometric Network Planning Study for Western and Northern Canada, Report for Canada Department of Energy, Mines and Resources.

Shawinigan Engineering Company Limited (1969). Planning of a Hydrometric Network for the Province of Ontario, Report for Canada Department of Energy Mines and Resources.

Solomon, S.I. (1965). Discussion of the paper by A. Bleasdale, Development and Design of Rain Gauge Networks with Special Reference to the U.K., Symposium on the Design of Hydrological Networks, Quebec 1965, Vol. I, pp 156-157 ,IASH Publication No. 67, Gentbrugge.

Solomon, S.I. (1970). Hydrometric Surveys, Proceedings of the AWRA Symposium of Water Balance in North America, Banff, 1969, Published by AWRA, Urbana, Illinois.

Solomon, S.I. (1971). Budget Allocation Planning for Hydrometric Networks, Paper presented at the Fifty Second Annual Meeting of the AGU, Washington.

Solomon, S.I. (1972a). Multi-regionalization and Network Strategy, Paper III - 3.3 in Casebook on Hydrological Network Design Practice, Published by WMO, W. B. Langbein, Editor, Geneva.

Solomon, S.I. (1972b). Joint Mapping, Paper III-21 in Casebook on Hydrological Network Design Practice, Published by WMO, W.B. Langbein, Editor, Geneva.

Solomon, S.I. (1972c). Natural Regime Hydrological Planning - A methodological discussion, International Symposium on Uncertainties in Hydrologic and Water Resources Systems, 1972, Vol. II, University of Arizona, Tucson, Arizona.

Solomon, S.I. (1972d). General Report: Data Network Design - Part II, International Symposium on Hydrologic and Water Resources Systems, 1972, Vol. III, University of Arizona, Tucson, Arizona.

Solomon, S,I. and Associates, (1973). WATERSTAT , A System for Interfacing Data Banks for Water Resources Planning and Management, Report for Canada Department of Environment.

Solomon, S.I. and Associates (1974). A Hydrologic Model for Environmental Impact Assessment in the Rouge, Little Rouge, Petticoat and Duffins Watersheds, North Pickering Project, Ontario Ministry of Housing.

Solomon, S.I., J.P. Denouvilliez, E. J. Chart, J.A. Woolley and C. Cadou (1968). The Use of the Square Grid System for Computer Estimation of Precipitation, Temperature and Runoff, Water Resources Research. Vol. 4, No. 5.

Solomon, S.I. and A.S. Qureshi (1972). Hydrological Data Banks - Present Status and Future Potential, EIC Engineering Journal Jan. Feb.

Solomon, S.I., A.S. Qureshi, and U. Korngold (1971). Use of a Parametric Model as a Tool for Hydrometric Network Planning, IASH Symposium on Mathematical Models in Hydrology, Warsaw, 1971.

Szesztay, K. (1967). On Concepts and Methods of Hydrological Data Acquisition, Draft Report on WMO/IHD Projects.

Thomas D.M. and M.A. Benson, (1970). Generalization of Streamflow Characteristics from Drainage Basin Characteristics, U.S. Geological Survey Water Supply Paper 1975, Washington, D.C.

Toebes, C. (1965). The Planning of Representative and Experimental Networks in New Zealand, IASH Symposium on Representative and Experimental Areas, Budapest, IASH Publication No. 66, Gentbrugge.

Turc, L. (1954). Le bilan d'eau des sols, Relations entre les précipitations, l'évaporation et l'écoulement, Annales Agronomiques, Vol.5, No. 4.

U.K. Water Resources Board, (1969). Proceedings of the U.K. Water Resources Board Symposium on Data Retrieval and Processing.

U .N. (1970). Projecto Hydrometeorologico Centro Americano, Pamphlet issued by U.N W.M.O. and the Governments of Costa Rica, El Salvador, Guatamala, Honduras, Nicaragua and Panama.

UNESCO, (1966). Ad Hoc Working Group on Network Planning and Design, Final Report of Co-ordinating Council for IHD, Second Session Paris, Doc. NS/204, Annexe IV, Paris.

UNESCO (1967). Panel of Experts on Network Planning and Design, Final Report of First Session, Doc. SC/IHD/11, 5 annexes, Paris.

Uryvaev, V.A. (1965). Basic Principles governing the Design of Hydrological Networks, WMO/IASH Symposium on the Design of Hydrological Networks, Quebec, 1965, IASH Publication No. 67, Gentbrugge.

U.S. Soil Conservation Service (1971). Engineering Handbook, Section 4 - Hydrology, U.S. Department of Agriculture, Washington, D.C.

Van der Made, J.W. (1972). Streamflow and Water Levels, Section I-5.1 in Casebook on Hydrological Network Design Practice, Published by WMO, W. B. Langbein, Editor, Geneva.

Wallis,J.R. and Anderson H.W. (1965). An Application of Multivariate Analysis to Sediment Network Design, WMO/IASH Symposium on Design of Hydrological Networks, Quebec, 1965, Vol. I, IASH Publication No. 67, Gentbrugge.

Wallis,J.R. and N.C. Matalas (1972). Information Transfer via Regression in Markovian Worlds, International Symposium on Uncertainties in Hydrologic and Water Resources Systems, 1972, Volume I, University of Arizona, Tucson, Arizona.

W.M.O. (1960). Hydrological Networks and Methods, WMO Flood Control Series, No. 15.

W.M.O. (1965). Guide to Hydrometeorological Practices, 1st Edition, Secretariat of the World Meteorological Organization, Geneva.

W.M.O. (1972). Casebook on Hydrological Network Design Practice, W.B. Langbein, Editor, Report No. 324, Secretariat of the World Meteorological Organization, Geneva.

W.M.O./IASH (1965). Symposium on the Design of Hydrological Networks, Quebec, 1965, IASH Publications 67 and 68, Gentbrugge.

Wilde, D.J. and C.S. Beightler (1967). Foundations of Optimization, Prentice Hall, Englewood Cliffs, N.J.

Wilson,K.C. (1972). Benefit - Accuracy Relationship for Small Structure Design Floods, Water Resources Research, Vol. 8, No. 3.

Yevjevich, V. (1972), Probability and Statistics in Hydrology, Fort Collins, Colorado, Water Resources Publication.

APPENDIX A

Physiographical Characteristics by Square Grid Technique

Many of the applications presented in this thesis are based on examples obtained from the analysis of the hydrology of Southern Ontario in conjunction with the physiographical characteristics of this area. We give in this Appendix the maps with the square grid values of the physiographical characteristics (terrain and climatic characteristics) of this area and the corresponding definitions of these characteristics. A general description of the technique used to determine the values of the physiographical characteristics is given in Solomon et al, (1968). A very detailed description of the technique can be found in Gupta (1974).

The physiographical characteristics were computed for a 5 x 5 km square grid, each square represented one quarter of the 10 x 10 km square of the Universal Transverse Mercator Projection grid which is printed on the 250,000 topographical maps produced by the Canada Department of Energy, Mines and Resources.

The terrain characteristics include:

1. Average elevation of each square (ELE), estimated as the average of the elevations of its four corners and stored as tens of feet. The square grid map of the elevations of the region is shown in Fig. A.1.

2. Local slope (SLO), estimated as slope of a plane defined by the elevations of the two southern corners and the left hand northern corner; stored as the tangent of the angle of inclination x 10^5 (Fig. A.2)

3. Azimuth of the slope (AZI) defined as the angle in degrees x 10 between the north direction and the horizontal projection of the line of steepest descent of the local slope plan, (Fig. A.3)

4. Barrier in the south direction (BHS), defined as the difference between the average elevation of the square considered and the highest elevation encountered in the south direction from this point until the ocean is reached; actual values used were determined for each square as the weighted average of the barrier height for the square considered and of the two pairs of adjacent squares on each side in a direction perpendicular to that for which the value of the barrier height is sought (stored as tens of feet, Fig. A.4.)

5,6,7, 8, 9, & 10. Barrier heights in the southwest, west, north-west, north, north-east and south-east (BSW), BHW, BNW, BHN, BNE. BSE) are defined similarly to the barrier height in the south direction. (Figs. A.5, A.6, A.7, A.8, A.9 and A.10).

11. Distance to the Great Lakes in south direction (DLS) defined as the distance from the Great Lakes in the

south to the given square; actual values were computed as weighted averages of the distances to the sea of the given square and of two pairs of adjacent squares on each side in a direction perpendicular to that for which the value of the distance is sought (stored in km. Fig. A.11)

12,13, 14,15, 16, & 17 Distance to Great Lakes in south-west, west, north-west, north, north-east and south-east directions (DSW, DLW, DNW, DLN, DNE, and DSE), defined in a similar manner as distance to the sea in the north direction. (Figs. A.12, A.13, A.14, A.15, A16 and A17).

18 Shortest distance to the Great Lakes defined as the shortest distance to one of the Great Lakes independently of the direction (Figure A.18)

19 Longitude index (LON) defined as the co-ordinate of the square in the west-east direction starting from the most western column of the grid squares (stored as the number of squares).

20 Latitude index (LAT) defined as a co-ordinate of the square in the south-north direction starting from the most southern line of squares (stored as the number of squares).

21 Area of of small lakes (ALA) defined as the area of small lakes (excluding the Great Lakes) included in a given square (stored as square kilometers x4; Fig. A.19.)

22 Area of bogs and swamps (ABS) defined previously as the area covered by swamps, marshes, and bogs in the (stored as square kilometers x4; Fig. A.20).

23 Area of forest defined as the area covered by forest in the grid square (stored as square kilometers x4; Fig. A.21)

24 Built-up land area index (ABL) defined as the areas covered by urban and large rural centers in the square (as shown on the 1:250,000 map) (stored as square kilometers x4; Fig. A.22)

25 Number of small dams in a square (NSD) defined as the number of small retention dams (with a retention volume of less than 100 acre ft.) stored as a number (Fig. A.23)

26 Permeability index (PIN) defined as a weighted average of soil indexes; the soil indexes are based on the ranking of the soils in order of increasing permeability as defined by Chapman and Putnam (1960). (Fig. A.24).

Note that in the case of other areas of Canada other terrain characteristics (such as shield effect, regional slope, area of glaciers, etc.,)were added. The definition of these additional factors is given by Solomon (1972a).

The climatic characteristics used in the study were long-term mean precipitation, temperature, and actual evaporation. (Figures A.25, A.26 and A.27).

Temperature values were obtained directly from a correlation between temperature data at meteorological stations and terrain characteristics. Precipitation and actual evaporation data were obtained through an iterative process in which these characteristics were considered to be interdependent with the runoff (Section 4.2.1c)

SOUTHERN ONTARIO - 5 x 5 km GRID - TERRAIN CHARACTERISTICS

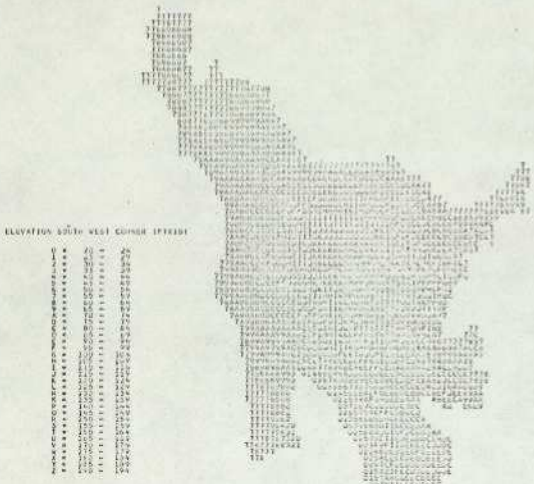


Fig. A-1 ELEVATION AT SOUTH-WEST CORNERS OF THE SQUARES



Fig A-2 SLOPE OF THE SQUARES

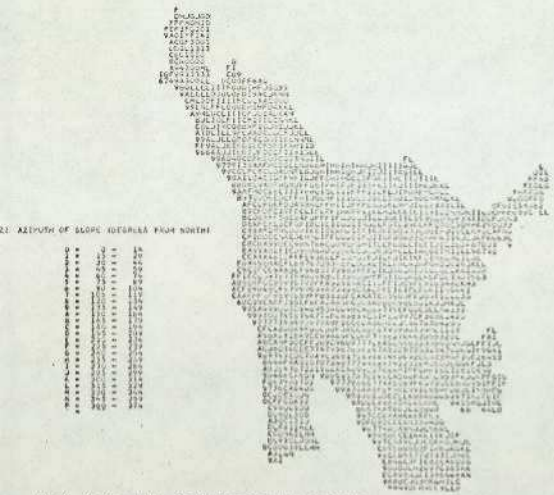


Fig. A-3 AZIMUTH OF SLOPE OF THE SQUARES

SOUTHERN ONTARIO – 5 x 5 km GRID – TERRAIN CHARACTERISTICS

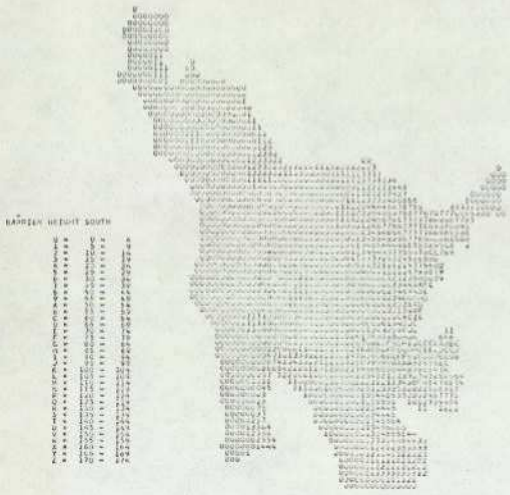


Fig. A-4 BARRIER HEIGHT – SOUTH (FT/10)



Fig. A-5 BARRIER HEIGHT – SOUTH WEST (FT/10)

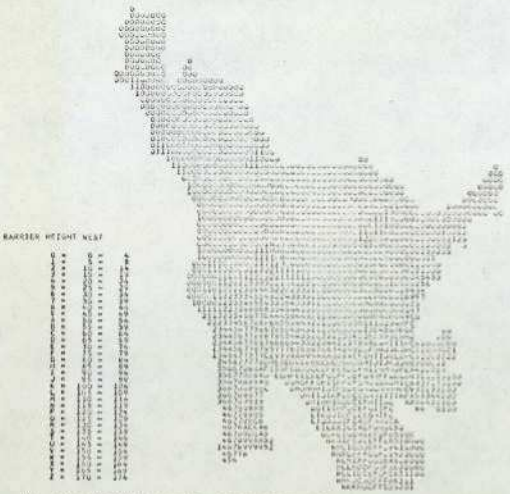


Fig. A-6 BARRIER HEIGHT – WEST (FT/10)



Fig. A-7 BARRIER HEIGHT – NORTH WEST (FT/10)

SOUTHERN ONTARIO – 5 x 5 km GRID – TERRAIN CHARACTERISTICS

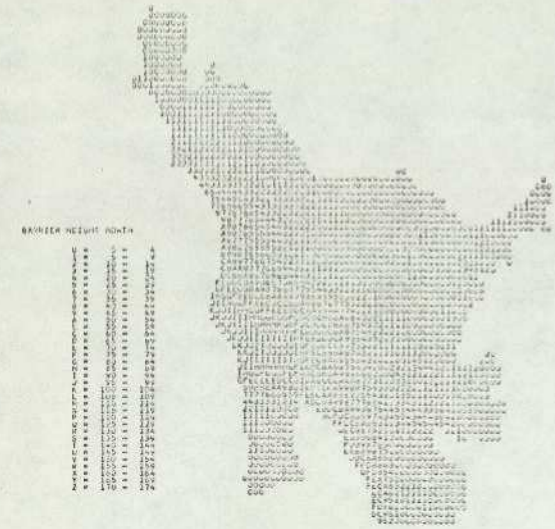


Fig. A-8 BARRIER HEIGHT – NORTH (FT/10)



Fig. A-9 BARRIER HEIGHT – EAST (FT/10)

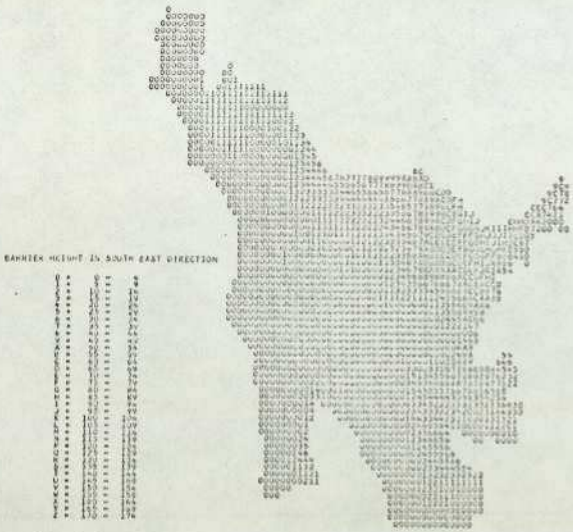


Fig. A-10 BARRIER HEIGHT – SOUTH EAST (FT/10)

SOUTHERN ONTARIO – 5 x 5 km GRID – TERRAIN CHARACTERISTICS



Fig. A-11 DISTANCE TO GREAT LAKES – SOUTH



Fig. A-12 DISTANCE TO THE GREAT LAKES – SOUTH WEST

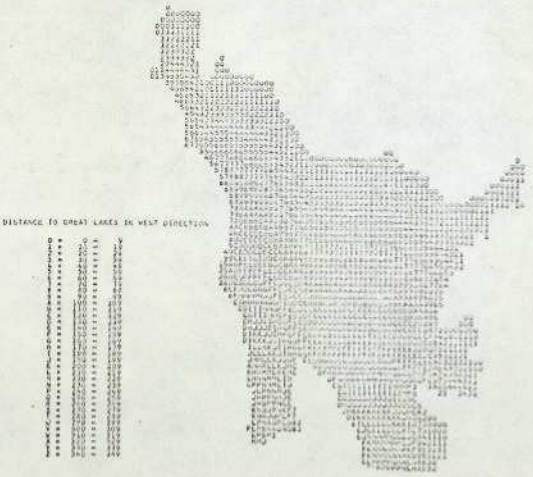


Fig. A-13 DISTANCE TO THE GREAT LAKES – WEST

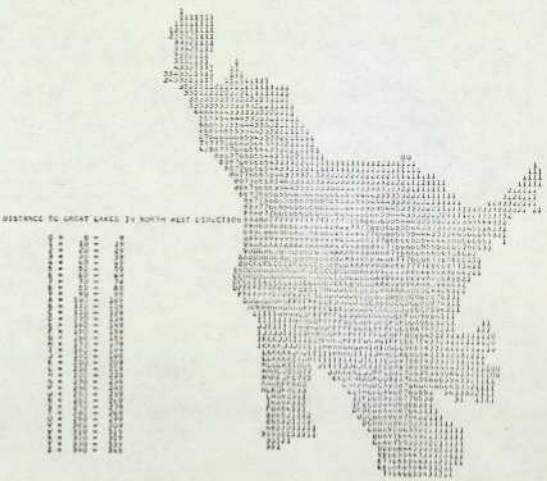


Fig. A-14 DISTANCE TO THE GREAT LAKES – NORTH WEST

SOUTHERN ONTARIO – 5 x 5 km GRID – TERRAIN CHARACTERISTICS



Fig. A-15 DISTANCE TO THE GREAT LAKES – NORTH



Fig. A-16 DISTANCE TO THE GREAT LAKES – EAST

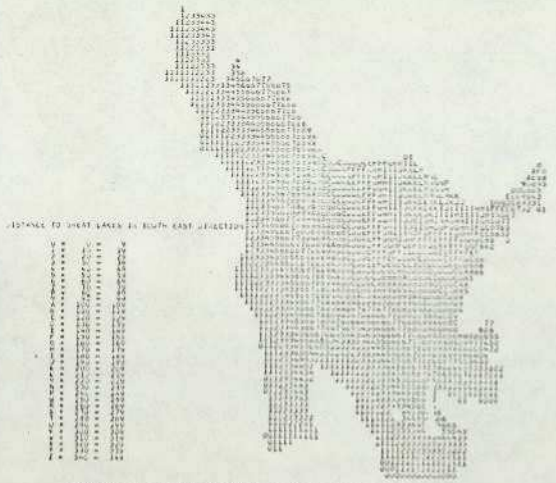


Fig. A-17 DISTANCE TO THE GREAT LAKES – SOUTH EAST

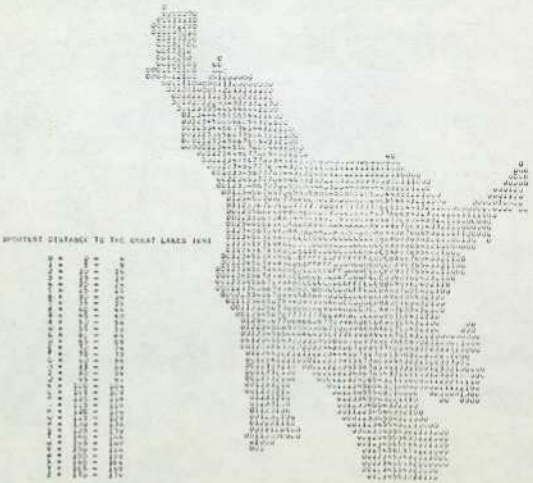


Fig. A-18 SHORTEST DISTANCE TO THE LAKES

SOUTHERN ONTARIO – 5 x 5 km GRID – TERRAIN CHARACTERISTICS

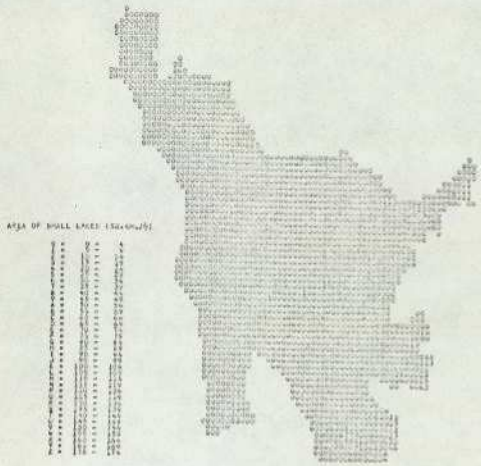


Fig. A-19 AREA OF SMALL LAKES IN THE SQUARES

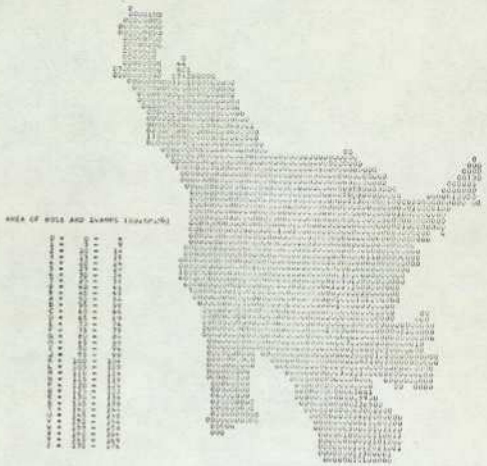


Fig. A-20 AREA OF BOGS AND SWAMPS IN THE SQUARES

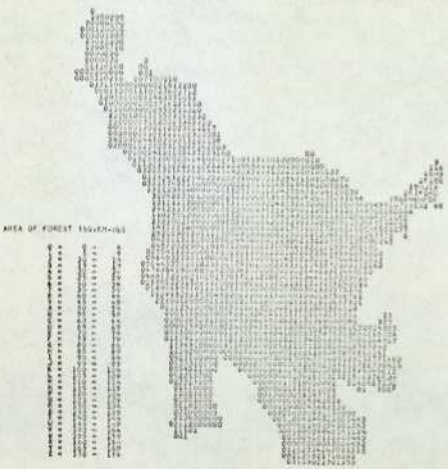


Fig. A-21 AREA OF FOREST IN THE SQUARES

SOUTHERN ONTARIO - 5 x 5 km GRID - TERRAIN CHARACTERISTICS

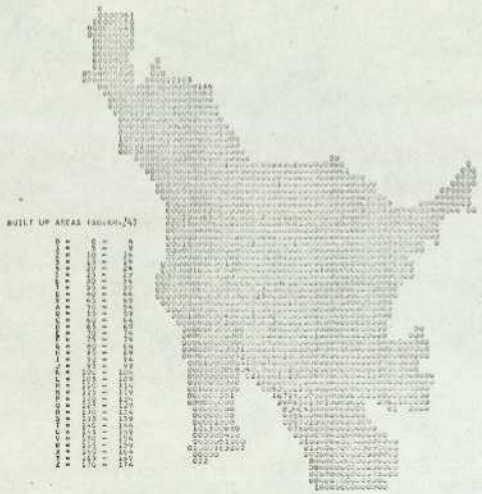


Fig. A-22 BUILT-UP AREAS IN THE SQUARES



Fig. A-23 NUMBER OF SMALL DAMS IN THE SQUARES

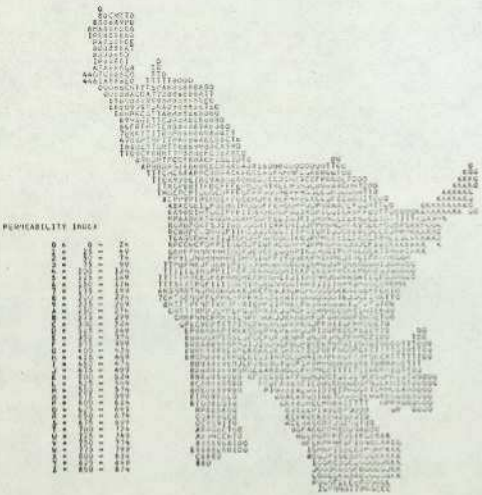


Fig. A-24 PERMEABILITY INDEX IN THE SQUARES

SOUTHERN ONTARIO – 5 x 5 km GRID – CLIMATIC CHARACTERISTICS

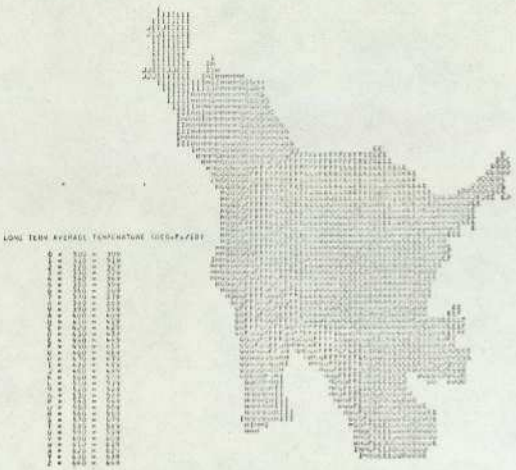


Fig. A-25



Fig. A-26

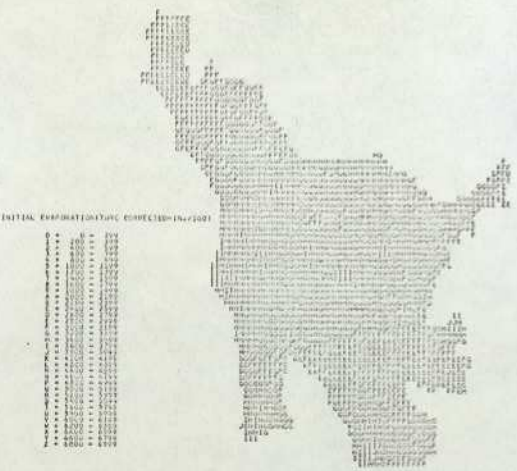


Fig. A-27

APPENDIX B

Sqare Grid Definition of Runoff in Southern Ontario.

To illustrate further the technique for interpolation of runoff data by the square grid technique (Sub-Section 4.1.2.c), we give in this Appendix some details on the results obtained in applying the technique to Southern Ontario.

Table B-1 shows the results of the regression equation obtained after the final analysis. The table gives the regression constant, the regression coefficients and their standard error, the standard error of the dependent variable (in inches) and the correlation coefficient. Table B-2 presents a comparison between the measured (actual) and estimated (predicted) runoff data and their ratio for the 33 river basins used as the calibration sample. Table B-3 gives the same data for the 13 river basins used as the validation sample. Figure B-1a presents the square grid distribution of the runoff obtained from the regression equation defined in Table B. 1 and Figure B-1b, the runoff distribution as determined using the isoline (isorhets) Technique (Coulson, 1967).

STANDARD ERROR OF DEP VARIABLE = 0.3164E 00

CONSTANT 0.1119755E 02

VARIABLE	COEFF	STANDARD ERROR
X - 1	0.8983486E-01	0.1332281E-02
X - 2	-0.1516847E-02	0.2453935E-03
X - 4	0.8151654E-04	0.7471361E-05
X - 5	-0.1895995E-02	0.7162004E-04
X - 6	0.2217357E-02	0.4837206E-04
X - 7	0.1061139E-02	0.2902898E-04
X - 8	0.1670277E-01	0.2864455E-03
X - 10	-0.3575703E 00	0.8671900E-02
X - 11	-0.2132229E 00	0.5520000E-02
X - 12	-0.5658507E 00	0.1062256E-01
X - 14	0.5440282E-01	0.3623136E-02
X - 15	-0.1752475E 00	0.6227537E-02
X - 16	-0.2235561E 00	0.5279044E-02
X - 17	0.1370331E 00	0.5812302E-02
X - 18	0.5332461E-01	0.5913403E-02
X - 19	-0.4549317E-09	0.2940341E-10
X - 20	0.3642538E-01	0.5868040E-02

COEFFICIENT OF CORRELATION =0.9061

X(1) = LON
X(2) = AFO PERCENT
X(3) = ABL PERCENT
X(4) = SLO (FT/FT*100,000)
X(5) = DBW (FT*10**1/3)*(KM)
X(6) = DBS (FT*10**1/3)*(KM)
X(7) = DBE (FT*10**1/3)*(KM)
X(8) = SDL (KM)**1/3
X(9) = DNW (KM)**1/3
X(10) = DLW (KM)**1/3
X(11) = DSW (KM)**1/3
X(12) = DLS (KM)**1/3
X(13) = DLE (KM)**1/3
X(14) = BHN (FTX10)**1/3
X(15) = BHW (FTX10)**1/3
X(16) = BSW (FTX10)**1/3
X(17) = BHS (FTX10)**1/3
X(18) = BHE (FTX10)**1/3
X(19) = PIN **3
X(20) = NDS
X(21) = RUN (INCH/100)**1/3

TABLE B-I : CORRELATION BETWEEN RUNOFF AND TERRAIN CHARACTERISTICS IN SOUTHERN ONTARIO

STATION NO.	--RUNOFF (INCHES)-- ACTUAL	PREDICTED	(ACT./PRED.)	STATION	
74	17.60	17.27	1.018	02FA001	SAUBLE RIVER AT SAUBLE FALLS
75	17.51	16.39	1.068	02FB007	SYDENHAM RIVER NEAR OWEN SOUND
76	16.84	16.74	1.005	02FC001	SAUGEEN R.NEAR PORT ELGIN(BELOW 2FC2)
77	14.87	15.73	0.944	02FC011	CARRICK CREEK NEAR CARLSRUHE
78	15.69	15.50	1.011	02FE005	MAITLAND RIVER ABOVE WINGHAM
79	12.88	12.69	1.014	02FF002	AUSABLE RIVER NEAR SPRINGBANK
80	12.43	13.03	0.953	02FF003	PARKHILL CREEK NEAR PARKHILL
81	11.68	12.08	0.966	02GA010	NITH R.NEAR CANNING(BELOW 02GA018)
82	13.83	14.88	0.929	02GA017	CONESTOGO RIVER AT DRAYTON
83	12.77	11.96	1.067	02GA029	ERAMOSA RIVER ABOVE GUELPH
84	11.95	11.64	1.025	02GB008	HORNER CRK NEAR MOUNT VERNON(BELOW 2GB
85	12.19	12.07	1.009	02GB009	KENNY CREEK NEAR BURFORD
86	9.98	11.34	0.879	02GB010	MCKENZIE CREEK NEAR CALEDONIA
87	12.06	12.55	0.960	02GC010	BIG OTTER CREEK AT TILLSONBURG
88	14.05	12.72	1.104	02GD004	MIDDLE THAMES RIVER AT THAMESFORD
89	15.47	13.07	1.183	02GD009	TROUT CREEK NEAR ST.MARY'S
90	13.33	12.91	1.032	02GD010	FISH CREEK NEAR PROSPECT HILL
91	12.49	12.29	1.015	02GD012	THAMES RIVER AT WOODSTOCK
92	14.57	13.78	1.056	02GD018	AVON RIVER BELOW STRATFORD
93	11.08	11.51	0.961	02GE003	THAMES R.AT THAMESVILLE(BELOW 02GE002)
94	11.78	11.80	0.997	02GG002	SYDENHAM RIVER AT ALVISTON
95	8.98	10.84	0.828	02GG004	BEAR CREEK ABOVE WILKESPORT
96	11.80	10.69	1.103	02HA006	TWENTY MILE CREEK AT BALLS FALLS
97	10.93	10.91	1.001	02HA007	WELLAND RIVER AT MERRITTS CHURCH
98	9.66	11.71	0.824	02HB001	CREDIT RIVER NEAR CATARACT
99	9.72	9.72	0.999	02HB004	EAST OAKVILLE CREEK NEAR OMAGH
100	12.03	10.85	1.107	02HB008	WEST CREDIT RIVER AT NORVAL
101	10.45	9.30	1.123	02HC005	DON RIVER AT YORK MILLS
102	11.91	10.96	1.086	02HC006	DUFFIN CREEK AT PICKERING
103	10.67	10.39	1.026	02HC013	HIGHLAND CREEK NEAR WEST HILL
104	9.83	11.54	0.851	02HC018	LYNDE CREEK AT WHITBY
105	7.21	9.61	0.750	02HC022	ROUGE RIVER NEAR MARKHAM
106	8.67	8.69	0.997	02HC023	COLD CREEK NEAR BOLTON

TABLE B-2 : COMPARISON BETWEEN MEASURED (ACTUAL) AND ESTIMATED (PREDICTED) RUNOFF FOR THE CALIBRATION SAMPLE (RUNOFF ESTIMATES FROM CORRELATION - SQUARE GRID TECHNIQUE)

STATION NO.	--RUNOFF (INCHES)-- ACTUAL	PREDICTED	(ACT./PRED.)	STATION	
3	16.61	15.06	1.102	02FC002	SAUGEEN RIVER NEAR WALKERTON
4	15.99	14.37	1.112	02FE004	MAITLAND RIVER NEAR DONNYBROOK
7	11.77	12.74	0.923	02GA003	GRAND RIVER AT GALT
8	11.26	11.96	0.941	02GA015	SPEED RIVER BELOW GUELPH
11	12.22	12.68	0.963	02GB001	GRAND RIVER AT BRANDFORD
15	12.92	13.25	0.974	02GD003	NORTH THAMES RIVER BELOW FANSHAWEDA
17	14.00	13.24	1.057	02GD007	NORTH THAMES RIVER AT LONDON
24	10.64	10.50	1.012	02HB002	CREDIT RIVER AT ERINDALE
25	8.84	8.90	0.993	02HC002	ETOBICOKE CREEK NEAR SUMMERVILLE
26	8.21	9.07	0.904	02HC003	HUMBER RIVER AT WESTON
27	10.06	9.41	1.068	02HC004	LITTLE DON RIVER NEAR LANSING
29	12.47	13.16	0.947	02GD001	THAMES RIVER NEAR EALING
30	13.16	12.91	1.018	02GC004	BIG OTTER CREEK NEAR VIENNA

TABLE B-3 : COMPARISON BETWEEN MEASURED (ACTUAL) AND ESTIMATED (PREDICTED) RUNOFF FOR THE VALIDATION SAMPLE (RUNOFF ESTIMATES FROM CORRELATION - SQUARE GRID TECHNIQUE)

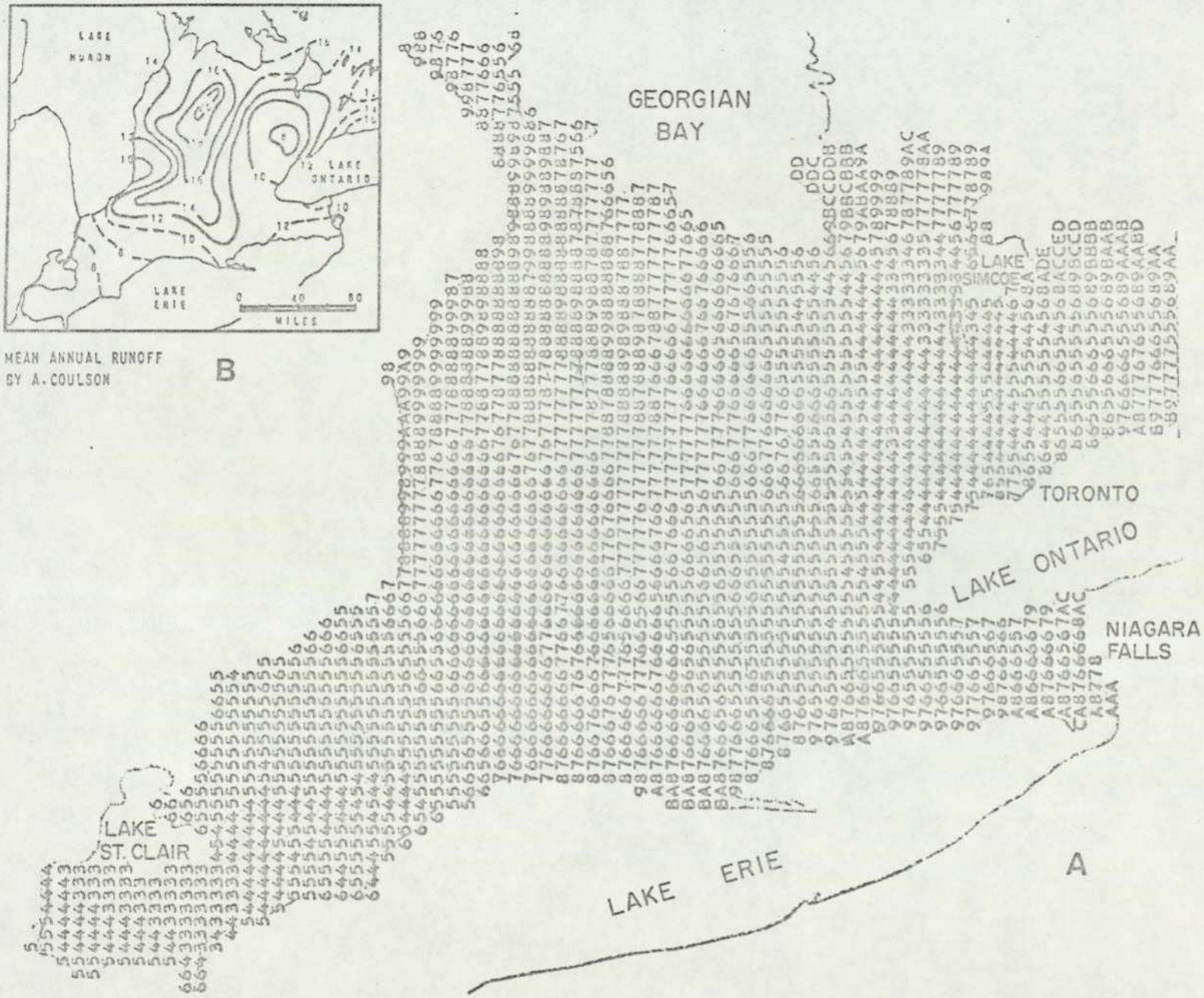


FIGURE B-1 : SQUARE GRID AND ISOLINE DISTRIBUTION OF RUNOFF IN SOUTHERN ONTARIO

SUPPLEMENTARY
REFERENCES

Allison, S.V. (1967). Cost, Precision and Value Relationships of Data Collection and Design Activities in Water Development Planning, Hydraulic Laboratory, University of California, Berkeley, Water Resources Center Control No. 120.

Andreyanov, V.G. (1969). Study of Influence of Physical-Geographical Factors on the River Runoff Characteristics (In Russian: Ob issledovanii vliyaniya fiziko-geograficheskikh faktorov na kharakteristiki rechnogo stoka). Meteorologiya i Gidrologiya, No. 8.

Beard, L.R. (1965). Use of Interrelated Records to Simulate Streamflow, ASCE Water Resources Engineering Conference, Mobile, Alabama.

Beard, L.R. (1972). Uncertainties in Reservoir and Lake Management, General Report. Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resource Systems, Vol. III, University of Arizona, Tucson, Arizona.

Beard, L.R. (1973). Status of Water Resource Systems Analysis, PASCE, Journal of the Hydraulics Division Vol. 99 No.HY4, April.

Belore, H.S. (1974). Estimating Snow Accumulation by Interfacing Hydrological and Meteorological Models with a Physiographic Data Bank, M.Sc. Thesis submitted at the University of Waterloo.

Belousov, S.L., L.S. Gandin, and A.S. Mashkovich (1968). Processing of Operative Meteorological Information by Means of Electronic Computers (In Russian: Obrabotka operativnoi meteorologicheskoi informatsii s pomoshchyu elektronnykh vychislitel'nykh mashin). Gidrometeoizahet, Leningrad.

Bleasdale, A. (1965). Development and Design of Rain Gauge Networks with Special Reference to the UK, WMO/IASH Symposium on the Design of Hydrological Networks, Quebec, IASH Publication No. 67, Gentbrugge.

Bruce, J.P. (1965). Effects of New Instrumentation on Snow Cover and Precipitation Networks, WMO/IASH Symposium on the Design of Hydrological Networks, Quebec, IASH Publication No. 67, Gentbrugge.

Charbonneau, R., J.P. Fortin and G. Girard (1971). Précision et sensibilité des modèles paramétriques, Runoff From Snow and Ice. NRC Subcommittee on Hydrology, Symposium No. 8, Quebec City.

Chow, V.T. (1965). Sequential Generation of Rainfall and Runoff Data, ASCE Water Resources Conference, Mobile, Alabama.

Cornell, C.A. (1972). First Order Analysis of Model and Parameter Uncertainty, Proceedings of International Symposium on Uncertainties in Hydrological and Water Resource Systems, Vol. III, University of Arizona, Tucson, Arizona.

Crow, F.R. and S.D. Hottman (1973). Network Density of Temperature Profile Stations and its Influence on the Accuracy of Lake Evaporation Calculations, Water Resources Research Vol. 9, No. 4.

Daniel, J.F. (1970). Satellite (ERTS-A) Network of Ground Data Sensors: A User-Oriented Experiment, Preprint No. 29, National Symposium on Data and Instrumentation for Water Quality Management, Madison, Wisconsin University.

Davis, D.A. and A. Coulson (1967). Hydrological zones in the Headwaters of the Saskatchewan River, Technical Bulletin No. 6, Inland Waters Branch, Department of Energy, Mines and Resources, Ottawa.

Davis, D.R. (1971). Decision Making Under Uncertainty in Systems Hydrology, Technical Report No. 2, Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona.

Dawdy, D. (1972). Worth of Data for Inference and Decision Making, General Report, Proceedings of the International Report on Uncertainties in Hydrologic and Water Resource Systems, Vol. III, University of Arizona, Tucson, Arizona.

Dawdy, D.R. and J.M. Bergmann (1969). Effect of Rainfall Variability on Streamflow Simulation, Water Resources Research, Vol. 5, No. 5.

Dawdy, D.R. and W.B. Langbein (1960). Mapping Mean Areal Precipitation, IASH Bulletin 19.

De Coursey Donn G, Objective Regionalization of Peak Flow Rates, Proceedings of the Second International Symposium in Hydrology, Fort Collins Colorado, Water Resources Publications.

Dimaksyan, A.M. and V.P. Petrov (1967). Determining Maximum Reliability and Efficiency of Hydrometeorological Systems and Instruments, Trudy GGI No. 149. Translated in Soviet Hydrology, No. 6.

Diskin, M.H., N. Buras, and S. Zamir (1973). Application of a simple Hydrologic Model for Rainfall-Runoff Relations of the Dalton Watershed, Water Resources Research Vol. 9 No. 4.

Dragoun, F.J. (1969). Effects of Cultivation and Grass on Surface Runoff, Paper presented at the 50th Annual Meeting of AGU, Washington D.C.

Duckstein, L. and C.C. Kisiel (1970). Efficiency of Hydrologic Data Collection Systems, Role of Type I and II Errors, Paper presented at the Sixth Water Resources Conference, Las Vegas, Nevada.

Eagleson, P.S. (1967). Optimum Density of Rainfall Networks, Water Resources Research, Vol. 3, No. 4.

Eagleson, P.S. and W.J. Shack (1967). Some Criteria for Measurement of Rainfall and Runoff, CSIRO Symposium on the Collection of Field Data, Interscience Publishers, New York.

Edson, C.G. (1951). Parameters for Relating Unit Hydrographs to Watershed Characteristics, Transactions AGU, Vol. 32, No. 4.

Fabret, A., Th. Thoré, B. Bayer and A. Castex (1971). Modèles physiques schématiques pour l'étude de la formation des crues à partir des précipitations et de la propagation de l'onde de crue, Houille Blanche, No. 3.

Fiering, M.B. (1963). Use of Correlation to Improve Estimates of the Mean and Variance, U.S. Geological Survey Professional Paper 434 - C.

Fiering, M.B. (1969). Economic Measure of Effect of Extreme Data , PASCE Journal of Sanitary Engineering Division, Vol. 95, No. SA4.

Gavrilov, A.M. (1968). The Problem of Hydrological Stations Distribution in the Area of Large Hydroelectrical Developments (In Russian: K voprosu razmeshcheniya giorologicheskoi seti v raionakh krupnykh gidrouzlov)Trudy GGI, Vol. 164.

Gavrilov, A.M. and L.V. Liss (1969). Example of Water Budget Computations at a Hydroelectric Plant (In Russian: Primer vodnobalansovykh raschetov na GES) Trudy GGI, Vol. 172 .

Gruffydd Jones D. (1971). Review of Hydrometric Schemes, Letter of Water Resources Board, Reading, to The Clerks of the River Authorities in England and Wales (Dated 8th February, 1971, unpublished).

Hamlin, M.J. (1970). A Study of Synthetic Flow Generation Techniques using Elan Valley Data, Proceedings of the Conference "Water Year 1970" held in Pretoria South Africa.

Hamlin, M.J. and N.T. Kottegoda (1971). Extending the Record of the Teme,Journal of Hydrology, No. 12.

Hardison, C.H. (1969). Accuracy of Streamflow Characteristics, US Geological Survey Professional Paper 650 - D.

Hayes, G.S. and R.A. Morill (1970). A Proposed Streamflow Data Program for Maine; US Geological Survey Open File Report.

Herfindahl, O.C. (1969). Natural Resource Information, published for Resources for the Future by John Hopkins Press, Baltimore, Maryland.

Hershfield, D.M. (1965). On the Spacing of Raingages, Symposium on the Design of Hydrological Networks, Quebec 1965,IASH Publication No. 6,Gentbrugge.

Institution of Water Engineers (1969). Symposium on River Flow Measurement, Proceedings of a Symposium held at the Loughborough University of Technology, September.

James, J.C., B.T. Bower and N.C. Matalas (1969). Relative Importance of Variables in Water Resources Planning, Water Resources Research Vol.V, No. 5.

Keyes, C.G. and R.J. Tefler (1973). Transmitting Water Resources Information by a Time Share System, Proceedings of the First International Conference on Transfer of Water Resources Knowledge, Fort Collins, Colorado, Water Resources Publications.

Kisiel, C.C. (1969). Time Series Analysis of Hydrologic Data, Advances in Hydrosciences, Vol. 5. Academic Press, edited by V.T. Chow.

Kisiel, C.C. and L. Duckstein (1972). Model Choice and Validation, General Report, Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resources Systems, Vol.III, University of Arizona, Tucson, Arizona.

Knutson, R.L. (1967). Quick Evaluation of Watershed Potentials, Journal of Soil and Water Conservation, January - Febraury.

Kottegoda, N.T. (1973). Flood Evaluation by Some Data Generation Techniques, Floods and Droughts, Proceedings of Second International Symposium in Hydrology, Fort Collins, Colorado, Water Resources Publications.

Langbein, W.B. (1946). Error in the Computation of Mean Areal Precipitation, U.S. Geological Survey (Processed).

Langbein, W.B. (1968). Hydrological Bench Marks, World Meteorological Organization, International Hydrological Decade, Report No. 8.

Langbein, W.B. and W.G. Hoyt (1959). Water Facts for the Nation's Future, New York.

Langford, R.H. and G.H. Davis (1970). National System for Water Data, PASCE, Journal of the Hydraulics Division, Vol. 96, No.HY7, July.

Langham, E.J. (1971). New Approach to Hydrologic Data Acquisition, PASCE, Journal of the Hydraulic Division, No.HY12, December.

Lund, J.A. (1970). A Monte Carlo Method for Testing the Statistical Significance of an Equation, Journal of Applied Meteorology, Vol. 9 No. 3.

Lund, J.A. (1971). Correlations Between Areal Precipitation and 850-millibar Geopotential Height, Monthly Weather Review, September.

Matalas, N.C. (1967). Mathematical Assessment of Synthetic Hydrology, Water Resources Research, Vol. 3, No. 4.

Matalas, N.C. and M.A. Benson (1968). Note on the Standard Error of the Coefficient of Skewness, Water Resources Research, Vol. 4 No. 1.

Matalas, N.C. and B. Jacobs (1964). A Correlation Procedure for Augmenting Hydrologic Data, U.S. Geological Survey Professional Paper 434-E.

Matalas, N.C. and W.B. Langbein (1962). Information Content of the Mean, Journal of Geophysical Research, Vol.67, No. 9, August.

Matheron, G. (1965). Les Variables regionalisées et leur estimation, Masson, Paris.

Matheron, G. (1970). Random Functions and their Application in Geology, Geostatistics, Edited by D.F. Merrian, Plenum Press.

McGarvey, H.A. (1968). Report on Review of Hydrometric Survey Data to 1960 for Lutteral Creek near Oustic, Station No.2GA-21, Department of Energy Mines and Resources, Water Survey of Canada, Ontario District.

McGuiness, J.L. and G.W. Vaughan (1969). Seasonal Variation in Rain Gage Catch, Paper presented at the 50th Annual Meeting, Washington, D.C.

McPherson, M.B., D.C. Taylor and L.S. Tucker (1969). An Analysis of National Basic Information Needs in Urban Hydrology, American Society of Civil Engineers.

Penmann, H.L. (1963). Vegetation and Hydrology, Technical Communication No. 53, Commonwealth Bureau of Soils, Commonwealth Agricultural Bureau.

Pinkayan,S. and P. Premchun (1973). Risks from Using Inadequate Flood Records, Proceedings of the Second International Symposium in Hydrology, Fort Collins, Colorado, Water Resources Publications.

Reimer, P.O. and J.B. Franzini (1971). Urbanization's Drainage Consequence, ASCE Journal of the Urban Planning and Development Division, Vol.97, No. UP2, December.

Riggs, H.C. (1970). Transfering Streamflow Characteristics, ASCE Meeting Preprint 1093, ASCE Meeting, Memphis, Tennessee.

Rodriguez-Iturbe, J. Data Network Design, General Report, Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resource Systems, Vol. III, University of Arizona, Tucson, Arizona.

Salas-La Cruz (1972). Information Content of the Regional Mean, Proceedings of the International Symposium on Uncertainties in Hydrology and Water Resource Systems, Vol. III, University of Arizona, Tucson, Arizona.

Sharp, A.L., A.E. Gibbs, and W.J. Owen (1966). Development of a Procedure for Estimating the Effects of Land and Watershed Treatment on Streamflow, Technical Bulletin No. 1352, US Government Printing Office, Washington, D.C.

Skelton, J. and A. Homyk (1970). A Proposed Streamflow Data Program for Missouri, U.S. Geological Survey, Open-file Report.

Sopper, W.E. and H.W. Lull, Editors, (1967). Forest Hydrology, Proceedings of a National Science Foundation for Advanced Science, held at Pennsylvania State University, Pergamon Press.

Tschannerl, G. (1971). Designing Reservoirs with Short Streamflow Records, Water Resources Research, Vol. 7, No. 3.

UNESCO (1969). International Conference on the Practical and Scientific Results of the International Hydrological Decade and on International Co-operation in Hydrology, Paris, 8-16, December 1969.

U.S. Geological Survey (1969). Concepts for the Design of Streamflow Data Programs, Open File Report, November 25.

U.S. Soil Conservation Service (1971). Engineering Handbook, Section 4 - Hydrology, U.S. Department of Agriculture, Washington, D.C.

Vinogradov, V.L. (1968). Accuracy of Flow Computations at Large Hydroelectric Plants (In Russian: O tochnosti ucheto stoka v stvorakh Krupnykh GES), Trudy GGI, Vol. 164.

Wilm, J.G. (1943). Efficient Sampling of Climatic and Related Environmental Factors. Transactions of the American Geophysical Union, Vol.24, pt.I.

Wong, S.T. (1963). A Multivariate Statistical Model for Predicting Mean Annual Flood in New England, Annals of the Association of American Geographers, Vol. 53, No. 3.

Woolhiser, D.A. (1972). Uncertainties in Deterministic Models, General Report, Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resources Systems, Vol.III, University of Arizona, Tucson, Arizona.

Woolhiser, D.A. Editor (1973). Decisions with Inadequate Hydrologic Data, Proceedings of the Second Hydrology Symposium, September, 1972, Fort Collins Colorado, Water Resources Publications.

Yen, B.C. (1970). Risks in Hydrologic Design of Engineering Projects, Journal of Hydraulic Division, ASCE, Vol. 96,No.HY4, April.

Yevjevich, V. (1972). Stochastic Processes in Hydrology, Water Resources Publications, Fort Collins, Colorado.

Young, G.K. and W.C. Pisano (1968). Operational Hydrology Using Residuals, Proceedings ASCE, HY4, July.