A Comparative Study of Flow Forecasting in the Humber River Basin using a Deterministic Hydrologic Model and a Dynamic Regression Statistical Model

by

Robert C. Picco

A thesis submitted to the
School of Graduate Studies
in partial fulfilment of the
requirements for the degree of
Master of Engineering

Faculty of Engineering and Applied Science
University of Newfoundland
August 1997

St. John's Newfoundland

Table of Contents

APPENDICES

List of Tables

List of Figures

Abstract

Since 1995 the Water Resources Management Division of the Department of Environment and Labour has generated flow forecasts for the Humber River Basin. These forecasts are required to provide a flood warning for residents living in the downstream sections of the basin. Information is also useful to the Deer Lake Power Company for the safe and efficient operation of the hydroelectric development that controls over two thirds of the basin.

Flow forecasts for river systems can be generated using two approaches. One approach is to use a deterministic model that tries to construct a mathematical model by accounting for some, or all of, hydrologic factors responsible for runoff in the basin. The second approach is to develop a statistically based model that uses the historic flows and climate data from the basin to generate a forecast. For any model to be effective in an operational environment, it must be compatible with hydrometeorologic input data collected in near real time.

The current method of generating forecasts uses the SSARR (Streamflow Synthesis and Reservoir Regulation) hydrologic model, a continuous simulation model that has reservoir routing capabilities, plus the ability to account for the areal distribution of meteorologic inputs, including snowmelt. This model defines the water budget using hydrometeorological inputs.

In this study a forecast method based on the dynamic regression technique was developed to produce one, two and three day forecasts for the five gauged sub basins results. Also, forecasts for the same time periods were produced using the SSARR model. The results of both methods were compared using the mean absolute percentage error (MAPE) criterion.

The results of the comparison showed that the dynamic regression performed better than the SSARR model for all basins, particularly for the larger basins.

Acknowledgements

I would like to take this opportunity to thank a number of people who provided advice, assistance and support during my pursuit of a Masters of Engineering degree. First of all, I would like to thank the Newfoundland Department of Environment and Labour for providing access to the data used in the analysis. In particular, I would like to thank Dr. Wasi Ullah (now retired) for his encouragement to start the program.

My sincere thanks are also extended to Dr. Leonard Lye, my supervisor, for his academic guidance and suggestions during the program.

Last, but not least, I sincerely thank my wife Brenda for her patience and moral support.

Abbreviations

ACF	Autocorrelation Function
AES	Atmospheric Environment Service of Environment Canada
ARIMA	Autoregressive integrated moving average
BII	Baseflow Infiltration Index
DCP	Data Collection Platform
ETI	Evapotranspiration Index
NDOEL	Newfoundland Department of Environment and Labour
SMI	Soil Moisture Index
SSARR	Streamflow Synthesis and Reservoir Regulation Model
S-SS	Sub Surface Infiltration
WSC	Water Survey of Canada, branch of Environment Canada

1. INTRODUCTION

1.1 Background

The Humber River Basin is the second largest river system on the island of Newfoundland with a drainage area of over 8000 km^2. Over half the basin is regulated for hydroelectric power generation by the Deer Lake Power Company (DLPC). Since the early 1900's, the communities, including Deer Lake and Steady Brook, have developed along the Humber River. As part of this development many residents were drawn to the scenic areas and flat lands along the flood plain areas of the Humber River and Deer Lake. Each year some flooding occurs at Steady Brook and Deer Lake. While the flooding is usually minor, the potential for flooding that affects large areas of the communities exists, as occurred in 1969 and 1981.

In an effort to reduce flood damage caused by this annual problem, these areas were included in the Canada-Newfoundland Flood Damage Reduction Program. A hydrotechnical study completed in 1984 (Cumming Cockburn, 1984) identified the flood risk areas corresponding to the 1:20 and 1:100 year interval floods. The hydrologic modelling was carried out using the Streamflow Synthesis and Reservoir Regulation (SSARR) model. Additional studies (Cumming Cockburn, 1985 & 1986) recommended that a flood forecasting system be developed to warn residents of impending flood events and did a preliminary evaluation of a forecasting system. This would allow residents to take preventative action to move valuables or take other appropriate actions. Such a system would also provide the

Deer Lake Power Company with information to operate the hydroelectric generating station with additional efficiency and safety.

Even before the completion of these studies, the need for an accurate 'near real time' flow forecasting and flood warning system for this basin was recognized by the Department of Environment and Labour (NDOEL) and the Power Company. The Department's goal is to provide a flood warning system with up to 72 hours advance warning of an impending flood. The Power Company requires accurate flow forecasts for safe and efficient operation of the hydroelectric development at Deer Lake. Based on this common interest, the Department of Environment and Labour and the Power Company cost shared the expansion of the data collection network required to implement the flow forecasting and flood warning system for the basin.

1.2 Objectives of the Study

The primary objective of this study is to develop and evaluate an alternate method of forecasting based on dynamic regression modelling to forecast flows for the Humber River Basin in western Newfoundland. This model will be compared, using a statistically sound method, with the deterministic model currently in use. As well, descriptive and other background information on the basin and the development of the SSARR method will be presented.

1.3 Outline of Thesis

This thesis document is composed of six chapters. The first chapter provides the background and explains the need for flow forecasting in the study area. Chapter 2 describes the study area in terms of the physiographic and hydrologic parameters relevant to the study objectives and lists the hydrometric and climate data available for the study area. The procedure used to describe, analyse and prepare the data for development of the two forecasting methods used in the study is presented in Chapter 3. The first section of Chapter 4 discusses the deterministic hydrologic model presently used by NDOEL while the second part discusses the background and development of the dynamic regression model. In Chapter 5, the forecasted flows calculated by each of the modelling methods are presented and compared. The conclusions and recommendations are presented in Chapter 6.

2. DESCRIPTION OF THE STUDY AREA

2.1 General

The Humber River Basin is located on the western side of the island of Newfoundland as shown in Figure 1. The total drainage area measured at the hydrometric station on the Humber River at Humber Village Bridge is 7860 km^2. The watershed is drained by two main branches - the Upper Humber River and Grand Lake, both of which drain into Deer Lake and then into the Bay of Islands via the Humber River. The Upper Humber River basin, with a drainage area of 2110 km^2 measured at the hydrometric station near Reidville, originates in Gros Morne National Park and flows in a southerly direction. This section of the basin is in a relatively natural state and flows uncontrolled. The Grand Lake section of the basin has a drainage area of over 5000 km^2. This section of the basin is regulated to produce hydroelectricity with a generating plant that operates at Deer Lake. A second, smaller hydroelectric development operates within the basin at Hinds Lake.

The areas in the basin with concentrations of residential development are located near the confluence of the two main subbasins at Deer Lake, or downstream along the Humber River between Deer Lake and Corner Brook. The presence of residential areas with a history of flooding in vulnerable downstream reaches, coupled with the presence of the hydroelectric developments, provides the basis for the necessity of accurate flow forecasting in the basin.

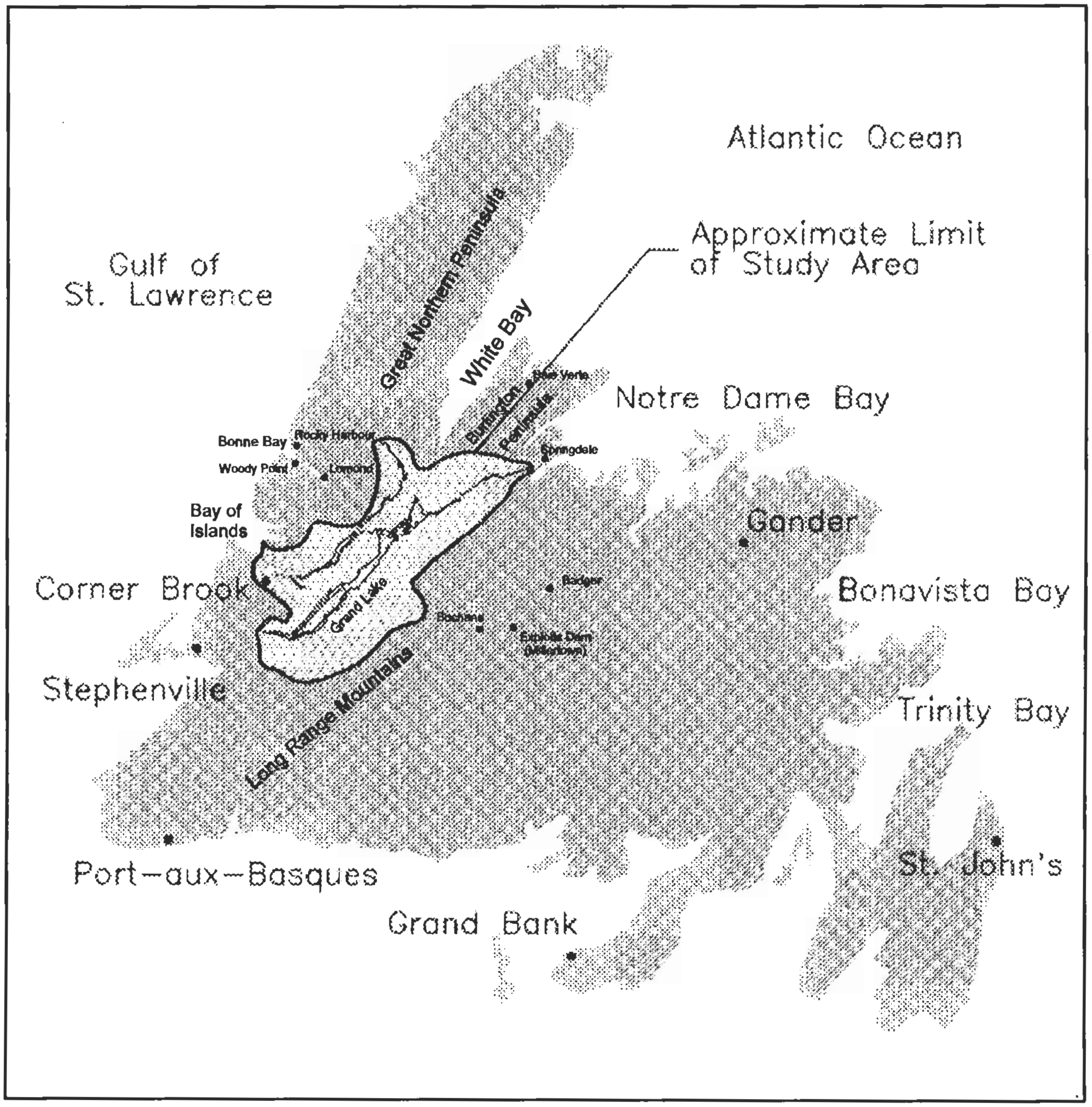


Figure 1- General Location of Study Area

2.2 Physiography

The Humber River Basin is, for the most part, situated in a physiographic region called the Newfoundland Highlands (Bostock, 1970). This area is further broken down into four sub-regions called the Great Northern Highlands, Blow Me Down Mountains, Atlantic Upland of Newfoundland and the Grand Lake Lowlands. These areas are shown on Figure 2. (Sanford et al, 1976)

The Great Northern Highlands form a barren mountainous plateau that extends northward along the Great Northern Peninsula from Bonne Bay and the Lomond River Valley with elevations ranging to 180 to 800 m. The inland surface of the highlands slopes gently in a southeasterly direction towards the Humber Valley and White Bay. The Upper Humber River and the Main River flow along this slope flowing in a southeastwardly direction.

The Blow Me Down Highlands form an area of dissected plateaus and mountains that rise abruptly from the coast along the Gulf of St. Lawrence in the Bay of Islands to Trout River area. The elevations in the area range from 550 m to 700 m. Drainage from this area flows in a westerly direction to the Gulf of St. Lawrence.

The Atlantic Upland of Newfoundland is a barren to sparsely forested plateau located between Grand Lake and Red Indian Lake and extending northward to the Burlington Peninsula. Elevations in this area range from 400 m to 600 m. The area forms the drainage divide for the two largest drainage basins on the island - the Humber River to the west and the Exploits River to the east.

The Grand Lake Lowlands are located towards the centre of the Humber River basin. These lowlands include the largest water bodies in the basin: Grand Lake, Sandy Lake and Deer Lake.

2.3 Surficial Geology

The surficial geology within the Humber River watershed ranges from bed rock outcropping, areas of glacial tills, and areas of sands and gravels to areas of organic soils. Based on the available information, (Golder, 1983) bedrock is the most common surficial geology classification type. Exposed bedrock forms extensive rock plains, knolls and ridges throughout much of the watershed. The rock, in most cases, is covered by a thin veneer of till soils or concealed by vegetation of either forest, scrub or peat bog. The area that comprises the highlands of the Long Range Mountains, Topsail Uplands and Burlington Peninsula is mostly exposed with rock talus or alluvium occurring on the mountain slopes. The soil types in the watershed, with some exceptions along the Humber River near Deer Lake and along the Humber River, are generally not suitable for agricultural uses due to excessive soil moisture, and adverse relief because of steepness or pattern of slopes, stoniness and shallowness to bedrock.

The glacial tills are found throughout the watershed with large variability in thickness from one area to another. The soils may be found as a thin veneer or extensive moraine deposit overlying the bedrock. The composition of the tills vary from grey silty sand or sand silt within the Long Range Mountains and Topsail Uplands, to red clayey silt within the

Humber River Valley. There are various local areas of till cover which are comprised of unsubdivided deposits of ice contact sand and gravel. In general, the composition of all tills closely resemble the lithology of the underlying bedrock.

The sand and gravel deposits found within the study area are outwash and fluvial in origin and are generally confined to stream and river valleys, The major sand and gravel deposits are found in the Deer Lake, Upper Humber River Valley, and the Sandy Lake-Birchy Lake areas. Additional buried deposits of sand and gravel occur at various points interstratified within the till deposits.

Peat deposits are found commonly throughout the watershed in areas with poor surface drainage. These deposits occur extensively on the barren Topsail Uplands just south of Grand Lake. Peat accumulations in the form of high moor bogs and string bogs are found on the highland plateaus of the Long Range Mountains. Peat deposits can attain thicknesses of several metres overlying bedrock or till deposits. Figure 2 shows the surficial geology and the major soil types found within the watershed.

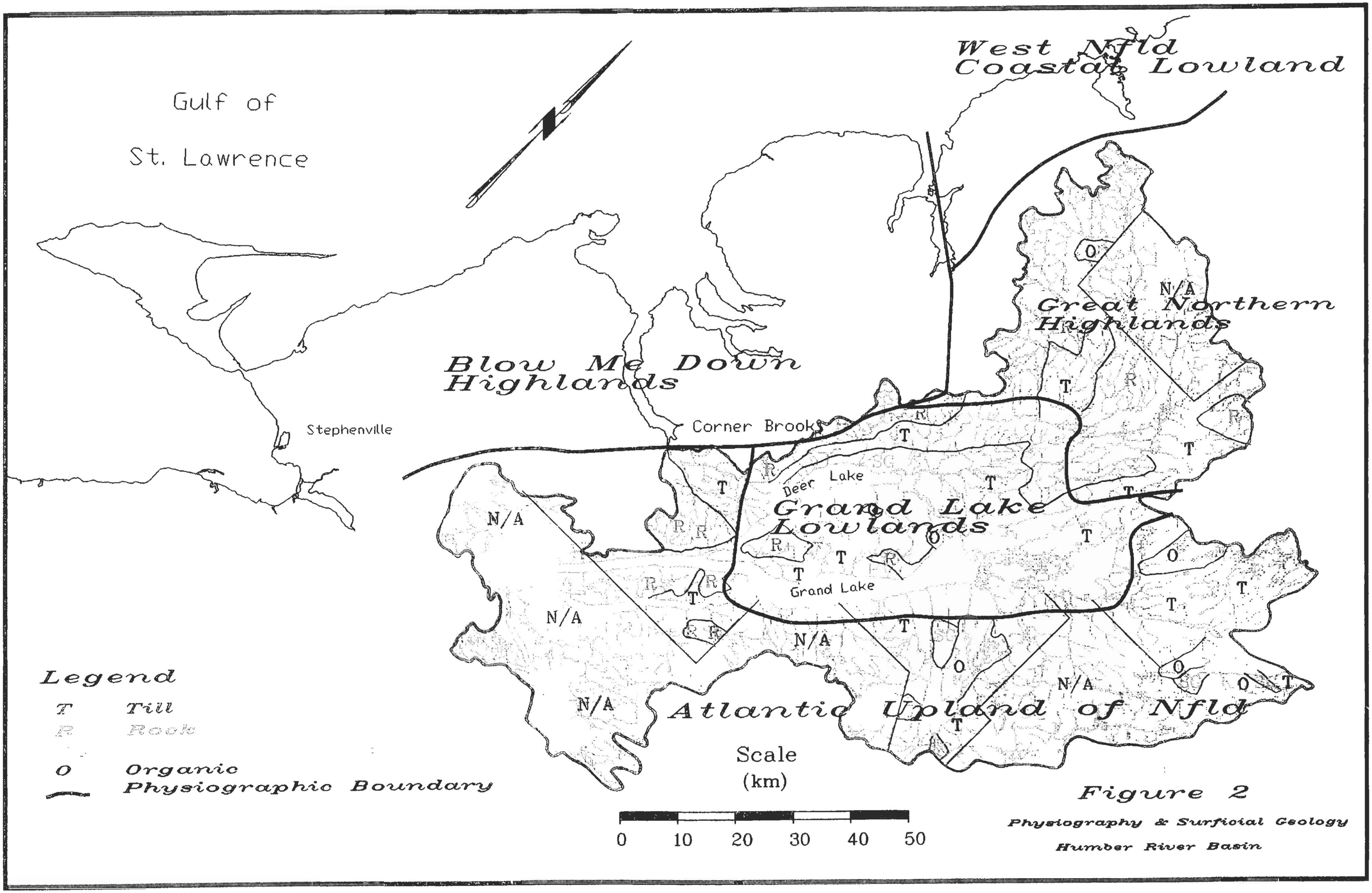


Figure 2
Physiography & Surficial Geology
Humber River Basin

2.4 Climate

The Humber River Basin has a temperate, marine climate that is influenced by the Gulf of St. Lawrence to the west and the Long Range Mountains to the east. Based on data from the long term climate stations in the basin (Atmospheric Environment Service, 1981& 1991) the average annual temperatures range from a low of 2.9°C at Baie Verte to a high of 5.1°C at Corner Brook. A summary of the average daily temperatures and average annual precipitation for the long term stations in and around the basin are shown in Table 1.

The annual precipitation ranges from a low of 943.5 mm at Badger, to a high of 1470 mm at Woody Point. Woody Point is located on the western coast while Badger is located far inland away from the coastal influence. The upland areas of the Long Range Mountains and the Topsails have greater precipitation due to orographic effects. The amount of precipitation decreases in the northeastern direction as indicated by Badger and Springdale with 943.5 and 967.1 mm respectively. For all stations, precipitation is lowest during the months of April and May while the highest amounts are recorded in the fall.

Table 1 - Long Term Climate Data - Humber River Basin

Station	Annual Precipitation (mm)		Average Daily Temperature (°C)	
	1951-1980	1961-1990	1951-1980	1961-1990
Badger	943.5	n/a	3.9	n/a
Baie Verte	1064.7	1100.8	2.9	3.0
Buchans	1071.5	1128.8	3.5	3.4
Corner Brook	1133.5	1186.0	5.1	5.2
Deer Lake	1033.1	1067.5	4.1	4.1
Deer Lake Airport	1023.2	1034.3	3.3	3.4
Exploits Dam	1099.0	1099.5	3.2	3.2
Rocky Harbour	1199.7	n/a	4.2	n/a
Springdale	967.1	1010.3	3.9	3.8
Stephenville	1166.5	1272.1	4.8	4.7
Woody Point	1470.0	n/a	4.6	n/a

(Station locations are shown on Figures 1 or 9)

2.5 Forest Cover

The Humber River Basin is within a classification band called the Boreal Forest Region of Canada (Rowe, 1972). This general region is subdivided into two subregions, Predominantly Forest Boreal Region and Forest and Barren Boreal Region. The Humber River watershed contains areas within both of these subregions. Figure 3 illustrates the forest and vegetative cover within the basin.

The dominant tree species include White Spruce, Black Spruce, Balsam Fir with areas of Tamarack, White Pine, White Birch, Yellow birch, Trembling Aspen and Balsam Poplar.

The forest cover in the Humber Basin is divided into four major classes: mature forest, scrub land, barren and peat bog, Mature forest is the most common forest cover class within the study area. This forest classification generally occurs in the areas that have the thickest glacial till layer above the bedrock. However, most of the tree cover is fairly sparse due to the thin veneer till soils and exposed bedrock. The dominant tree species within this class are Balsam Fir and Black Spruce, with White Birch, Trembling Aspen and Balsam Poplar also being found. Overall, sparse, mature forest cover is found in all subwatersheds. The most densely treed area within the watershed is located downstream of Deer Lake to the Community of Steady Brook where Balsam Fir and Black Spruce dominate with softwood scrub land, White Birch and peat bog also being found. Extensive logging has been carried out, and continues to be carried out, in the basin to support the paper mill in Corner Brook.

The remainder of the study area contains softwood and hardwood scrub lands, rock barrens and peat bogs. As can be expected, the vegetation classes are closely related to the soil type, soil depth and moisture regime. Well vegetated areas are usually small in area and scattered. Peat bogs are common in the highland plateaus, along river valleys and around lakes within the lowlands.

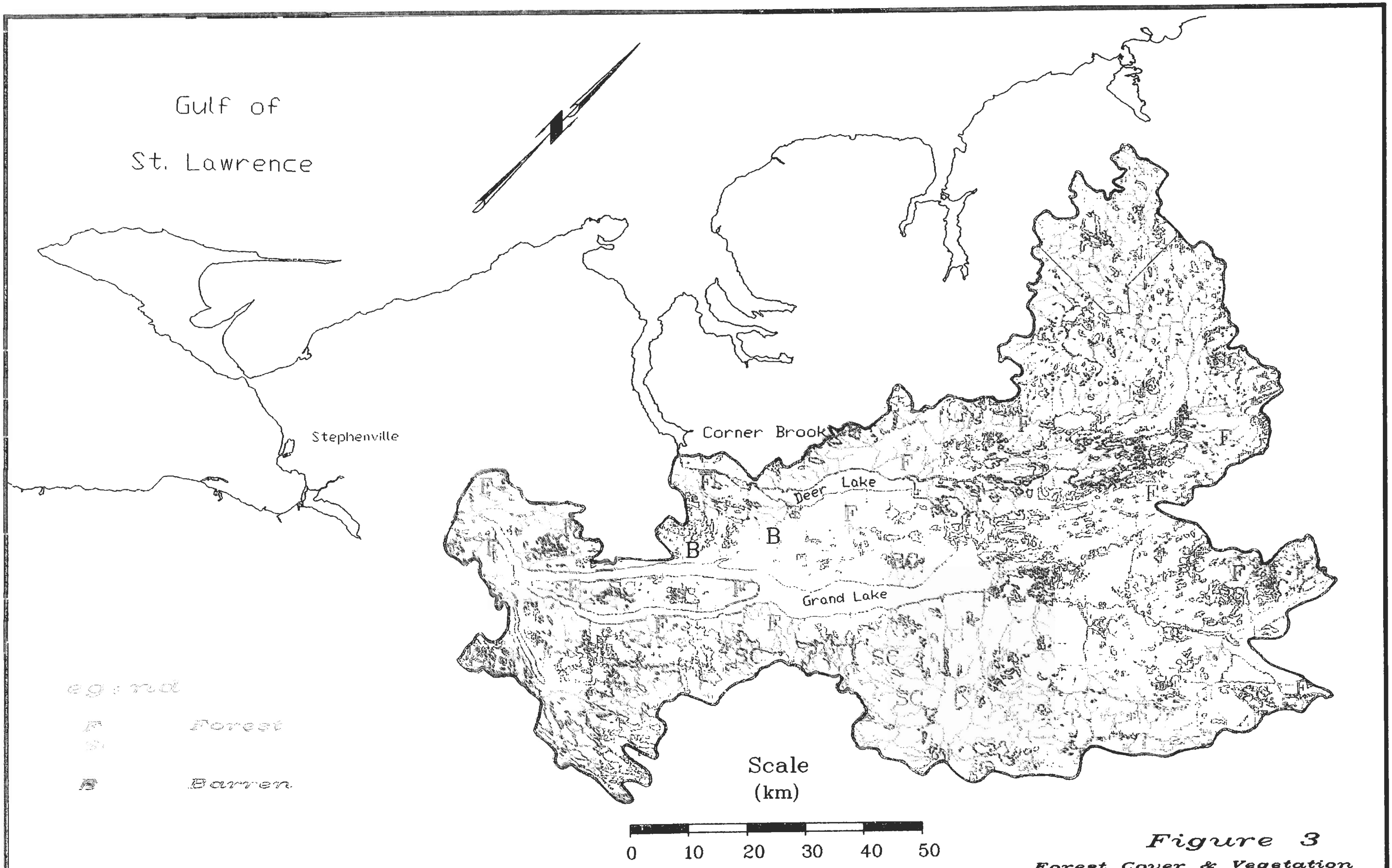


Figure 3

Forest Cover & Vegetation

3. DATA PREPARATION

3.1 Available Data

3.1.1 AES Climate Data

The long term climate stations in and near the basin are listed in Table 1 in Chapter 2.4. These stations are operated by Environment Canada's Atmospheric Environment Service (AES), and for the most part, are located outside the boundaries of the basin and consequently are not suitably located for short term flow forecasting. Also, the data available from these stations is not available in near real time since the gauges, except the Airport sites, are read manually by observers. Arrangements can be made with AES to obtain the data but the data must be manually entered into the computer.

3.1.2 WSC Flow/Stage

Environment Canada, through its Water Survey of Canada (WSC) branch operates a series of hydrometric stations in and around the basin under a cost sharing agreement with NDOEL. The flow data is published yearly in CD-ROM format following a comprehensive quality control review. The data is not readily available through WSC in near real time. A summary of the long term stations in the basin is presented in Table 2. One important observation from this data is that all of the maximum instantaneous flows occurred during the spring snowmelt runoff period.

Table 2 - Summary of Hydrometric Data - Humber River Basin

Station Name	WSC Code	Area (km^2)	Period of Record	Mean Annual Discharge (m^3/s)	Maximum Instantaneous Flow (m^3/s)
Lewaseechjeech Brook **(LEWA)**	02YK002	470	1952-1967 1972 - present	17.6	175 (04/93)
Hinds Lake at Outlet **(HIND)**	02YK004	630	1956-1979	16.5	136 (04/71)
Sheffield Brook at Sheffield Lake (SHEF)	02YK003/	380	1955-1966	9.9	93.4 (05/60)
Sheffield Brook near TCH **(SHEF)**	02YK005	391	1972-present	12.0	121 (05/87)
Indian Brook Diversion to Birchy Lake **(INDI)**	02YM002	238	1963-1978	5.7	46.7 (05/75)
Indian Brook above Birchy Lake **(INDI)**	02YM004	238	1990-present	6.5	51.4 (05/93)
Upper Humber River near Reidville **(REID)**	02YL001	2108	1940-present	82.2	1060 (06/84)
Upper Humber River above Black Brook **(BLAC)**	02YL008	471	1988-present	26.4	302 (06/95)
Humber River at Humber Village **(VILL)**	02YL003	7860	1982-present	257	887 (06/95)

(See Figure 9 for station locations)

3.1.3 Near Real Time Data

Near real time data are collected at the stations using Data Collection Platforms (DCP)'s. The data are telemetered via satellite to an earth receiving station where they are retrieved by NDOEL via modem. These stations are operated and maintained by Environment Canada under a cost sharing agreement between Environment Canada, NDOEL and the Deer Lake Power Company. The data retrieval, processing, and dissemination of the near real time data is carried out by NDOEL.

The raw data from the sites are given only a brief quality control review prior to archiving in database files.

3.2 Data Review

As noted above, the data available for the Humber River Basin originate from three sources, namely AES for long term climate data, WSC for long term hydrometric data, and NDOEL for near real time climate and hydrometric data. The AES data is not readily accessible on a real time basis and WSC does not publish the data until about six months after the data are collected.

Since any flow forecasting system for the basin will require near real time data, the data from NDOEL was selected as the basis for the model development. The main cause for concern in using this data is that it is not subject to the extensive quality control review and correction like the AES and WSC data. For this reason, the data review for the NDOEL data will concentrate on identifying any problems with the data

The first step in reviewing any data set is to plot the data. A sample plot of the data for the Upper Humber River near Reidville is shown in Figure 4 and a plot of the precipitation and temperature data for Grand Lake on Glover Island is shown in Figure 5. This initial review aided in the task of identifying incorrect data due to instrument error, outliers or the presence of missing data. Any problems noted were corrected in the data files. Figures 4 and 5 are illustrative of the nature of the data. The data review was done primarily using the computer display screen. Plots for the remainder of the stations used in this study are presented in Appendix A.

The next step in the data review was to calculate the descriptive statistics for the data. SYSTAT software (SYSTAT Inc, 1991) was used for this purpose. The results of this analysis is given in Tables 3-6.

The next step in the review of the data was to compare the quality controlled data from the WSC with NDOEL's near real time data. As shown in the example in Figure 6, there is a definite difference in the two data sets that occurs during the winter season. The apparent higher flows during the winter season were believed, and were later confirmed, (Baker, 1996) to be a result of an ice cover in the stream. The instrument at the site measures water level which is converted to a flow based on a rating curve developed using open water conditions. This causes an erroneously high water level and consequently an erroneously higher flow. The effect was noted at all of the gauges, with the exception of two, the

Lewaseechjeech Brook and Humber River at Humber Village Bridge. The comparison plots for the remainder of the stations is presented in Appendix B.

At this point the possibility of correcting for the ice cover was investigated. The procedure used by Environment Canada to correct for this backwater effect involves a graphical technique supported by on site reconnaissance and knowledge of nearby basins to adjust for the ice effect. Since this type of correction cannot be carried out on real time data due to the unavailability of field information and time constraints, the only choice was to proceed with the analysis using the real time data, with the knowledge that there is an inaccuracy in the data.

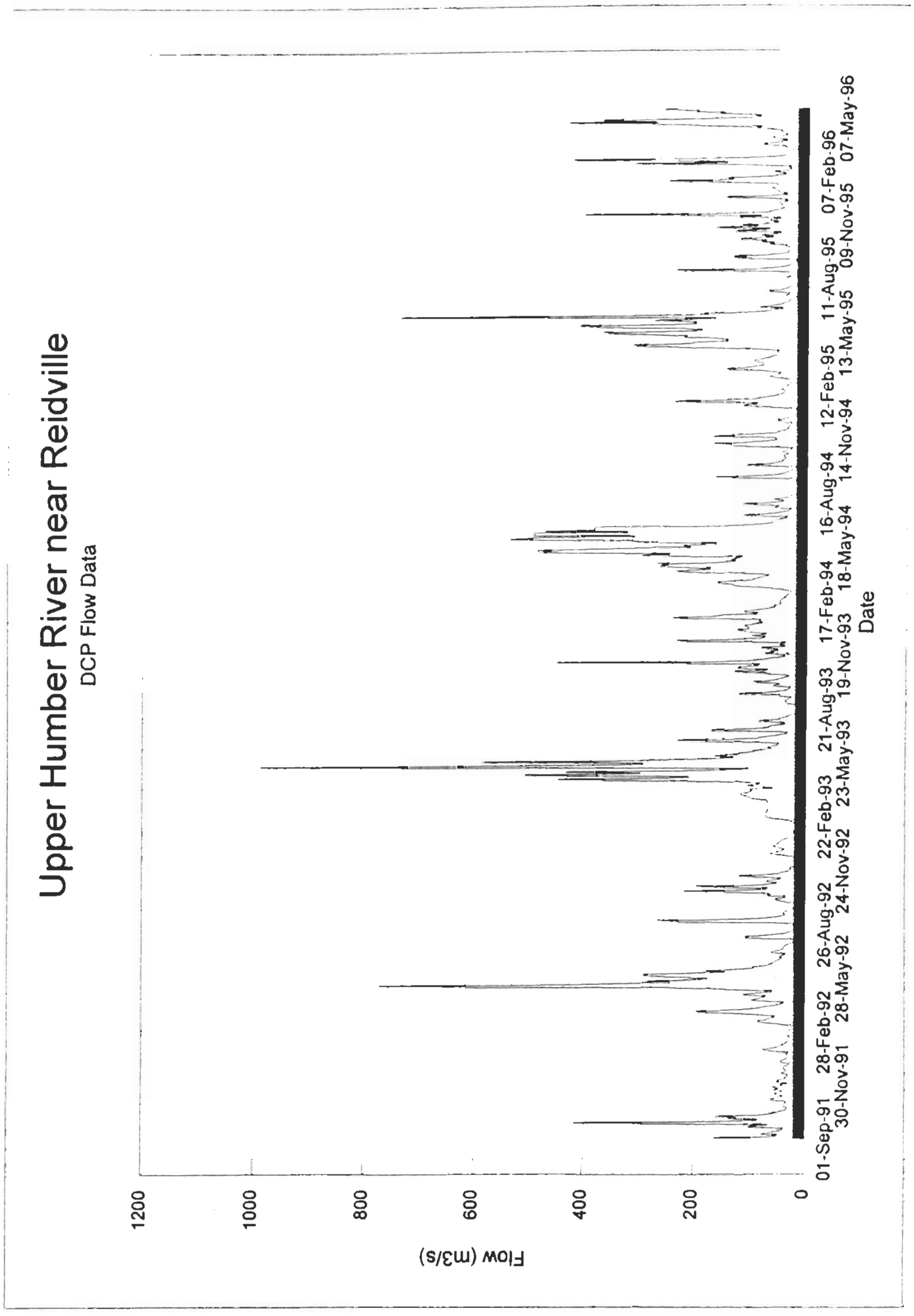


Figure 4 - Sample Flow Data from DCP Station

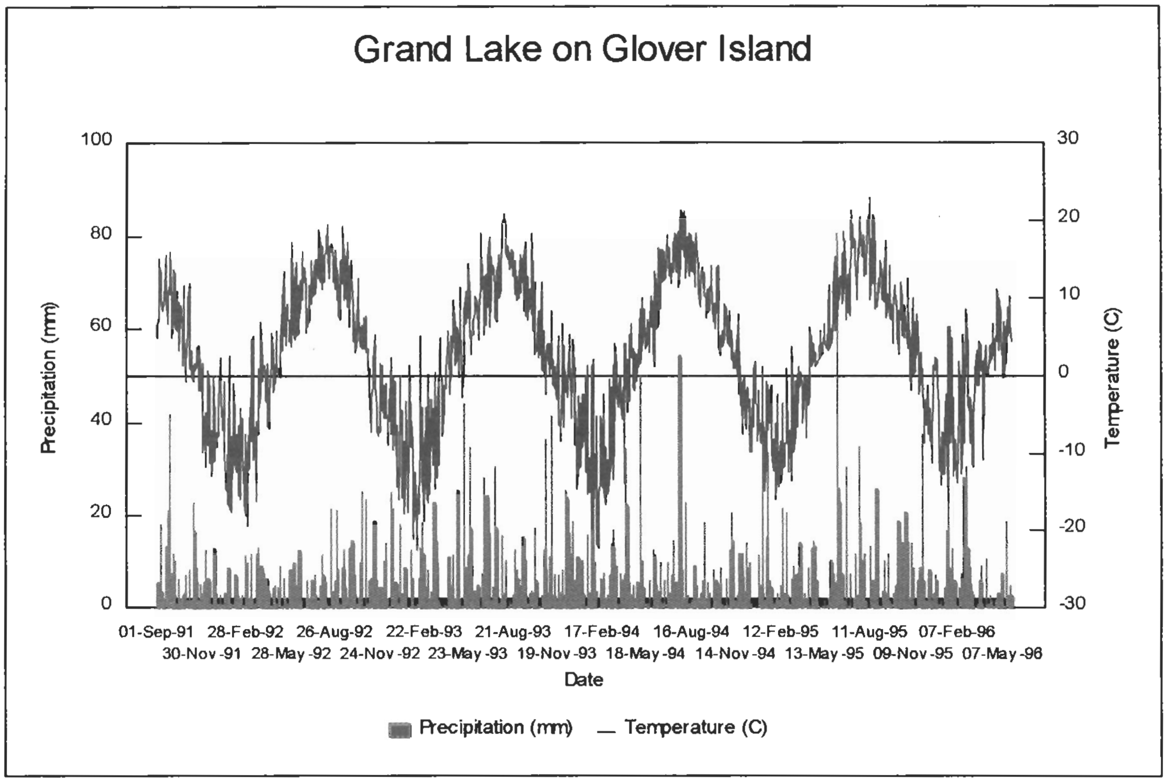


Figure 5 - Sample Temperature and Precipitation Data

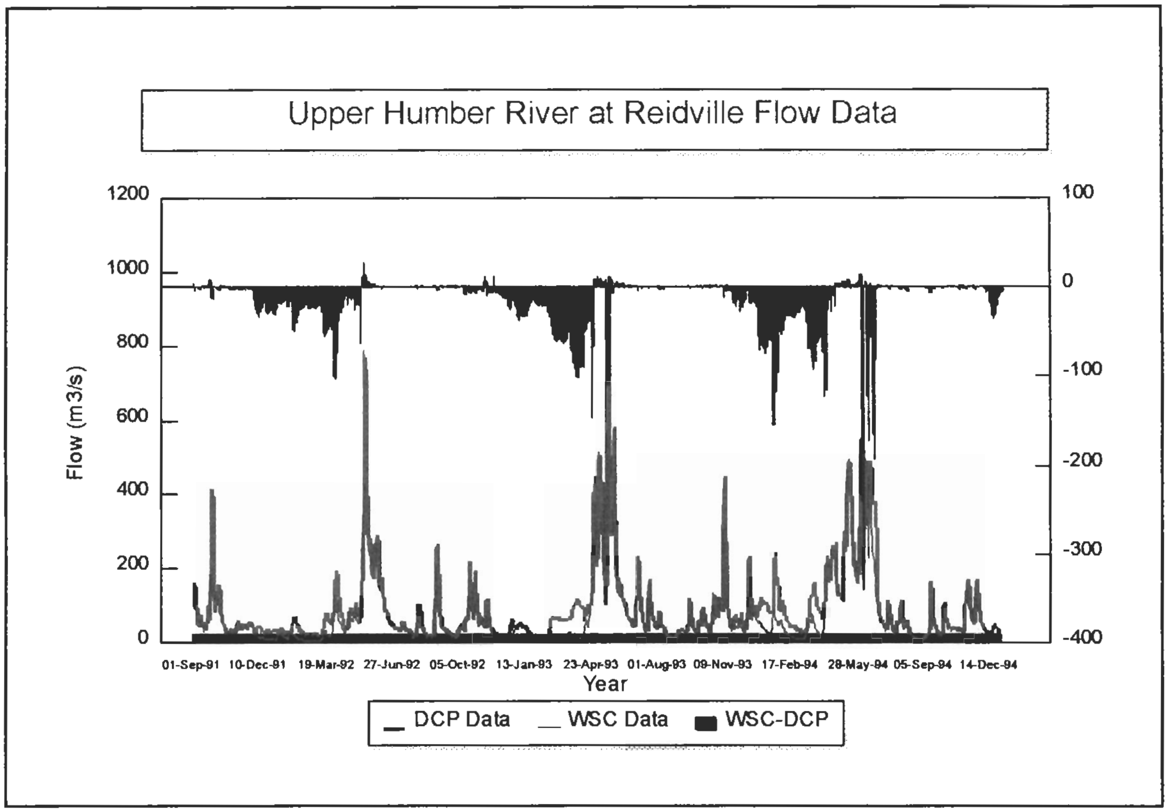


Figure 6 - WSC and DCP Data Comparison

Table 3 Descriptive Statistics - Flow and Climate Data

Upper Humber River above Black Brook

	FLOW	TOTPRE	TEMPAVG	PRERAIN	PRESNOW	TEMPMAX	TEMPMIN
N OF CASES	1666	1659	1659	1659	1659	1659	1659
MINIMUM	1.140	0.000	-30.280	0.000	0.000	-22.700	-39.500
MAXIMUM	275.290	70.880	21.210	70.880	55.410	28.900	18.700
RANGE	274.150	70.880	51.490	70.880	55.410	51.600	58.200
MEAN	40.717	3.854	0.197	2.350	1.502	5.110	-5.262
STANDARD DEV	37.891	6.769	10.272	5.889	3.880	10.516	11.206
STD. ERROR	0.928	0.166	0.252	0.145	0.095	0.258	0.275
SKEWNESS(G1)	1.877	3.741	-0.223	4.960	5.175	0.021	-0.503
KURTOSIS(G2)	5.052	20.418	-0.604	34.448	39.916	-0.692	-0.341
C.V.	0.931	1.756	52.026	2.506	2.583	2.058	-2.130

Indian Brook Diversion to Birchy Lake

	FLOW	PRECRAIN	PRECSNOW	TEMPAVG	TEMPMAX	TEMPMIN	TOTPREC
N OF CASES	1728	1728	1728	1728	1728	1728	1728
MINIMUM	0.490	0.000	0.000	-27.650	-19.700	-36.800	0.000
MAXIMUM	72.560	117.710	53.680	22.730	32.200	18.400	117.710
RANGE	72.070	117.710	53.680	50.380	51.900	55.200	117.710
MEAN	11.192	2.238	0.982	2.125	7.321	-3.561	3.220
VARIANCE	119.828	38.638	11.410	101.405	110.612	115.657	48.453
STANDARD DEV	10.947	6.216	3.378	10.070	10.517	10.754	6.961
STD. ERROR	0.263	0.150	0.081	0.242	0.253	0.259	0.167
SKEWNESS(G1)	1.820	7.692	6.663	-0.322	-0.034	-0.619	6.001
KURTOSIS(G2)	3.596	103.787	63.756	-0.550	-0.751	-0.272	66.345
C.V.	0.978	2.778	3.439	4.738	1.437	-3.020	2.162

Table 4 - Descriptive Statistics - Climate Stations

Burgeo Road near Buchan's Access

	PRECRAIN	PRECSNOW	TEMPAVG	TEMPMAX	TEMPMIN	TOTPREC
N OF CASES	1728	1728	1728	1728	1728	1728
MINIMUM	0.000	0.000	-26.600	-23.000	-35.600	0.000
MAXIMUM	110.030	56.140	22.010	29.800	19.300	110.030
RANGE	110.030	56.140	48.610	52.800	54.900	110.030
MEAN	2.750	1.183	1.769	6.057	-2.661	3.932
VARIANCE	46.443	14.552	94.482	102.742	100.641	57.887
STANDARD DEV	6.815	3.815	9.720	10.136	10.032	7.608
SKEWNESS(G1)	5.312	6.582	-0.274	-0.139	-0.418	4.321
KURTOSIS(G2)	48.544	58.099	-0.694	-0.695	-0.532	31.983
C.V.	2.478	3.225	5.495	1.674	-3.770	1.935

Corner Brook Lake at Outlet

	PRECRAIN	PRECSNOW	TEMPAVG	TEMPMAX	TEMPMIN	TOTPREC
N OF CASES	1706	1706	1706	1706	1706	1719
MINIMUM	0.000	0.000	-28.910	-22.400	-37.400	0.000
MAXIMUM	106.230	76.580	20.950	28.600	17.500	106.230
RANGE	106.230	76.580	49.860	51.000	54.900	106.230
MEAN	3.120	1.463	1.277	5.384	-3.254	4.617
VARIANCE	61.745	13.885	94.869	98.437	107.898	72.686
STANDARD DEV	7.858	3.726	9.740	9.922	10.387	8.526
SKEWNESS(G1)	4.725	7.131	-0.309	-0.119	-0.534	4.067
KURTOSIS(G2)	33.871	104.804	-0.612	-0.713	-0.283	26.187
C.V.	2.518	2.547	7.628	1.843	-3.193	1.847

Grand Lake on Glover Island

	PRECRAIN	PRECSNOW	TEMPAVG	TEMPMAX	TEMPMIN	TOTPREC
N OF CASES	1728	1728	1728	1728	1728	1728
MINIMUM	0.000	0.000	-22.360	-19.100	-26.900	0.000
MAXIMUM	80.740	33.600	23.010	31.000	17.500	80.740
RANGE	80.740	33.600	45.370	50.100	44.400	80.740
MEAN	2.340	0.737	2.685	6.888	-1.265	3.077
VARIANCE	32.849	4.886	82.353	97.715	77.139	37.103
STANDARD DEV	5.731	2.210	9.075	9.885	8.783	6.091
SKEWNESS(G1)	5.152	5.901	-0.220	-0.059	-0.373	4.548
KURTOSIS(G2)	39.735	52.572	-0.679	-0.731	-0.548	31.727
C.V.	2.449	2.998	3.379	1.435	-6.941	1.979

Table 5 - Descriptive Statistics Climate Stations

Grand Lake at Southwest End

	PRECRAIN	PRECSNOW	TEMPAVG	TEMPMAX	TEMPMIN	TOTPREC
N OF CASES	1725	1725	1725	1725	1725	1725
MINIMUM	0.000	0.000	-25.510	-21.200	-34.400	0.000
MAXIMUM	78.230	63.550	22.600	32.800	17.200	78.230
RANGE	78.230	63.550	48.110	54.000	51.600	78.230
MEAN	2.197	0.841	2.320	7.355	-2.847	3.038
VARIANCE	30.647	8.799	94.501	109.767	99.611	37.344
STANDARD DEV	5.536	2.966	9.721	10.477	9.981	6.111
STD. ERROR	0.133	0.071	0.234	0.252	0.240	0.147
SKEWNESS(G1)	5.553	10.118	-0.296	0.011	-0.552	4.900
KURTOSIS(G2)	48.132	163.922	-0.580	-0.693	-0.366	37.590
C.V.	2.520	3.529	4.190	1.425	-3.506	2.012

Sandy Lake at Howley Road

	PRECRAIN	PRECSNOW	TEMPAVG	TEMPMAX	TEMPMIN	TOTPREC
N OF CASES	1725	1725	1725	1725	1725	1725
MINIMUM	0.000	0.000	-25.510	-21.200	-34.400	0.000
MAXIMUM	78.230	63.550	22.600	32.800	17.200	78.230
RANGE	78.230	63.550	48.110	54.000	51.600	78.230
MEAN	2.197	0.841	2.320	7.355	-2.847	3.038
VARIANCE	30.647	8.799	94.501	109.767	99.611	37.344
STANDARD DEV	5.536	2.966	9.721	10.477	9.981	6.111
STD. ERROR	0.133	0.071	0.234	0.252	0.240	0.147
SKEWNESS(G1)	5.553	10.118	-0.296	0.011	-0.552	4.900
KURTOSIS(G2)	48.132	163.922	-0.580	-0.693	-0.366	37.590
C.V.	2.520	3.529	4.190	1.425	-3.506	2.012

Table 6 - Descriptive Statistics - Flow Only Stations

Sheffield River near TCH

```
                           FLOW

N OF CASES                 1365
MINIMUM                   1.490
MAXIMUM                 107.000
RANGE                   105.510
MEAN                     12.622
VARIANCE                119.756
STANDARD DEV             10.943
STD. ERROR                0.296
SKEWNESS(G1)              2.581
KURTOSIS(G2)             10.055
C.V.                      0.867
MEDIAN                    8.490
```

Lewaseechjeech Brook

```
                           FLOW

N OF CASES                 1712
MINIMUM                   1.590
MAXIMUM                 174.630
RANGE                   173.040
MEAN                     20.167
VARIANCE                497.957
STANDARD DEV             22.315
STD. ERROR                0.539
SKEWNESS(G1)              2.953
KURTOSIS(G2)             11.035
C.V.                      1.107
MEDIAN                   12.130
```

Upper Humber River near Reidville

```
                           FLOW

N OF CASES                 1728
MINIMUM                   9.910
MAXIMUM                 989.000
RANGE                   979.090
MEAN                     97.733
VARIANCE              10856.951
STANDARD DEV            104.197
STD. ERROR                2.507
SKEWNESS(G1)              2.905
KURTOSIS(G2)             11.535
C.V.                      1.066
MEDIAN                   61.160
```

Humber River at Village Bridge

```
                           FLOW

N OF CASES                 1727
MINIMUM                 132.540
MAXIMUM                 837.580
RANGE                   705.040
MEAN                    279.265
VARIANCE              14224.584
STANDARD DEV            119.267
STD. ERROR                2.870
SKEWNESS(G1)              1.923
KURTOSIS(G2)              3.778
C.V.                      0.427
MEDIAN                  241.290
```

4. METHODOLOGY

This section will discuss the two methods used in this study to forecast flows for one, two and three days ahead for the Humber River Basin. Part one will focus on the deterministic model (SSARR), its history, application to the study basin and the flow forecasts for specific events.. The second section will provide similar information for the dynamic regression method.

4.1 Deterministic Model (SSARR)

4.1.1 Model Description

4.1.1.1 History

The SSARR model was developed in 1956 by the US Corps of Engineers to analyse and forecast flows for the natural hydrologic and controlled reservoir systems in the North Pacific Division area. This area included the Columbia River System in the US and Canada, coastal rivers in Western Oregon and Washington, and the State of Alaska. Since that time, the SSARR Model has been updated and made compatible for PC use and has been widely applied throughout the world.

One application that is similar to the Humber Basin in terms of geographic, climatic and operational characteristics is the St. John River Basin in New Brunswick. SSARR was selected for this basin in 1973 because it was believed to be the best available at that time

(Tang & Lockhart, 1983). The model continues to be used for flow forecasting for that basin up to the present day.

The model was selected for the Humber River basin for the following reasons: (Cumming Cockburn, 1984)

- The model is simple in structure and uses readily available data compared with many other continuous simulation models;
- The relatively fast simulation time which allows for low computational time compared with other models;
- The benefit of transferring model parameters and experience from other Canadian applications such as the St. John River noted above;
- The excellent reservoir routing capabilities;
- Variable computational time steps are possible, ie. a mixture of weekly and/or daily;
- The capability to account for the areal distribution of snowmelt;
- The model is non-proprietary and has been well proven in a number of practical applications;
- The model is widely used in flood forecasting; and
- An interactive version of the model is available.

4.1.1.2 General Description

The SSARR model has been developed to describe the main components of the hydrologic cycle. It was conceived as a closed hydrologic system in which the water budget is defined by meteorologic inputs (rainfall and/or snowmelt) and hydrologic outputs such as runoff, soil storage and evapotranspiration losses. The model is derived so that the main components of the hydrologic cycle are represented in a simplified but rigorously applied manner. The parameters used in the model allow for an extremely flexible means of representing the various hydrologic components. This unique feature allows the SSARR model to be applied to virtually any drainage basin or hydrologic system. A detailed discussion of the algorithms and data processing techniques utilized in the SSARR model is found in the Users Manual (Davis, 1991).

A schematic representation of the basic elements of the SSARR watershed model is presented in Figure 7.

The SSARR program carries out the following three distinct functions:

- It calculates the natural discharges for each of the elementary sub-basins within a subwatershed that is selected by the user for its relatively homogeneous hydrological characteristics.
- It calculates the natural discharges as well as routing and adding hydrographs through river reaches and natural lakes, up to the exit point of the entire watershed.

- It calculates the variations in discharge caused by the regulation of artificial reservoirs for hydroelectric generation or flood control operations.

4.1.1.3 Hydrologic Principles Utilized in the SSARR Model

The representation of hydrologic runoff processes in a watershed model is highly subjective. No two hydrologists look at watershed runoff processes in exactly the same light. Nevertheless, there are some underlying principles that must be preserved in the formulation of a deterministic hydrologic watershed model. These include the logical accounting of each of the basic elements in the hydrologic cycle, including rainfall, snowmelt, interception, soil moisture, interflow, groundwater recharge, evapotranspiration, and the various time delay processes. These elements must be accounted for while including the ability to maintain continuity of each of the processes and to represent each by objective functions that relate them to observed hydrometeorological parameters. The main difference between the ways various models account for these processes lies in the level of complexity that each model uses to represent a particular process.

The streamflow routing functions contained in the SSARR model provide a generalized system for solving the unsteady flow conditions in river channels where streamflow and channel storage effects are related, either at one point or at a series of points along a river system. In principle, the method involves a direct solution of a storage-flow relationship involved in maintaining continuity of streamflow and storage in each element of the river, using a procedure that solves the relationships in finite elements of time and river reach. This involves a completely general and flexible method for solving the flow routing equations which can be applied in many ways depending upon the type of basic data available, and the conditions of the river system with respect to backwater effects from variable stage discharge effects, such as tidal fluctuations or reservoir fluctuations.

The SSARR model was designed to include the effects of reservoirs or other water control elements within the streamflow simulation process. Reservoirs may be described for any location in a river system, whereby inflows are defined from single or multiple tributaries, derived either from watershed simulation for river basins upstream, or from specified flows as a time series, or a combination of the two.

Outflows from reservoirs are determined on the basis of specified operating conditions. In order to provide a once-through process for the system as a whole, including all natural effects and those related to human intervention, the processing of hydraulic conditions at reservoirs is performed sequentially with all other elements in the river basin simulation.

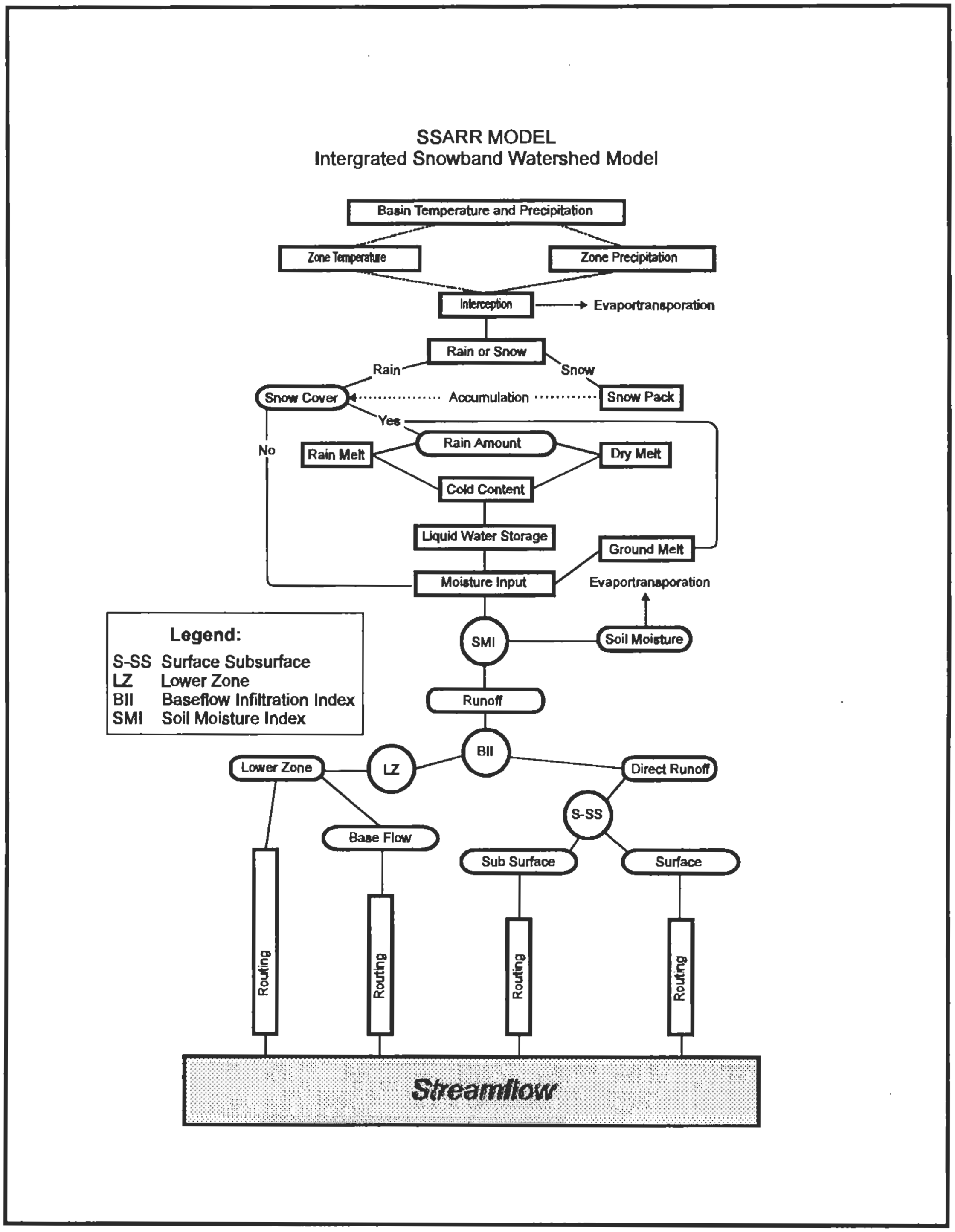


Figure 7 - SSARR Schematic with Snowband Option

4.1.1.4 Structure

This section defines and discusses hydrologic input parameters in general terms. The specific data input and parameter values used for various sub-watersheds of the Humber River in this study are presented and discussed in Chapter 4.1.2.

Net precipitation input (WP) - The average net precipitation value for a drainage basin or hydrologic unit is derived as a weighted daily or period amount from a series of individually reported or observed values. The weighting is generalized, and each station may be assigned its individual weighting value. The station weights may be determined on the basis of previously derived relationships between station and basin normal annual or normal seasonal precipitation; or the weighting values may be derived on an areal basis using the Thiessen Polygon or other similar technique.

The total precipitation on a watershed is computed from period precipitation amounts at one or more stations by:

$$WP_n = \frac{P^1 * W_1 + P_2 * W_2 + \ldots\ldots\ldots\ldots P_n * W_n}{n} \qquad \textbf{(1)}$$

Where

$P_1, P_2 \ldots\ldots P_n$ = Period precipitation amounts at Stations 1, 2, ... n respectively, in inches

$W_1, W_2 \ldots\ldots W_n$ = Weights applied to the station precipitation n = Total number of stations, a maximum of 30 stations

Wp_n = weighted period Net Watershed Precipitation inches

Basically, there are two general options for the of snow melt, namely:

- the temperature index method; and
- the use of generalized equations of snowmelt as determined by the thermal budget of heat loss and gain to the snowpack.

The temperature index method is usually used for daily forecasting applications, whereas the detailed energy budget approach is more appropriate for design flood calculations when extensive watershed data is available. By either method, daily or period values of effective snowmelt runoff values are computed as a time series, as a function of appropriate meteorological values.

The second option has the advantage of being more precise and detailed, but also has the disadvantage of requiring some data which is available at only a very limited number of stations. Examples of such data are:

- Difference between air temperature measured at 3 metres above the snow surface and snow surface temperature in degrees Celsius (°C). The snow surface temperature is assumed to be 0°C;
- Difference between the dewpoint temperature measured 3 metres above the surface of the snow and the temperature of the snow surface (0°C);
- Above the surface of the snow and the temperature of the snow surface (0°C);

- Wind velocity at 15 metres above the snow, in kilometres per hour;
- Solar radiation on a horizontal surface, in langleys;
- Average snow surface albedo;
- Basin shortwave radiation melt factor;
- Average forest canopy cover; and
- Convection-condensation melt factor.

In addition to the two options of snowmelt equations available to the user, two additional options for the simulation of snow cover are also available. These are:

- Snow Cover Depletion; and
- Snow Band Option.

With the snow cover depletion option, snow cover is diminished in thickness, and in surface area during the snowmelt season. However, two disadvantages of this option are that snowfall during the calculation period is not added to the existing snow cover and secondly, it is necessary to enter at the beginning of the period of calculation, the quantity of snow which will effectively run off after the melt. However, this option has been shown to be useful in the flood forecasting mode. The snow band option allows the user to separate the subwatershed into different "bands" according to elevation with the temperatures being lowered systematically with increased elevations, as illustrated in Figure 7

Soil Moisture Index (SMI) - The soil moisture index used in the SSARR model represents a weighted mean basin value of the water stored in the soil mantle that can be removed by plant roots through transpiration and natural evaporation. It does not include the part of the soil moisture content that exists at the permanent wilting point. The computation of the changes in soil moisture index values are based on the increases resulting from rainfall snowmelt, and the decreases by the evapotranspiration process. Increases in the soil moisture index values result in a "permanent" loss to runoff in the water balance for the basin as a whole. The upper limit of the soil moisture index is considered to be its field capacity, which is equivalent to the capillary moisture holding capacity, or the total amount of water which can be held under the force of gravity under natural conditions. Thus, the soil moisture index is a continuously varying parameter that may range from a value of zero when the soil moisture has been reduced to the "wilting point" by the evapotranspiration process, to a maximum value represented by the field capacity of the soil for the basin as a whole.

Evapotranspiration Index (ETI) - The evapotranspiration index (ETI) used in the SSARR model is a weighted basin mean daily value of the water lost to the atmosphere by the evapotranspiration process. Transpiration, soil evaporation and evaporation of free water from the plant or forest cover are considered to act together to produce the losses by evapotranspiration. Since the evapotranspiration loss is physically the result of change of state of water from the liquid to vapour phase, the process of evapotranspiration requires energy for the transformation, and is, therefore, dependent upon a source of energy, from the atmosphere.

In the model, the potential evapotranspiration may be computed by either of two basic methods, namely:

- mean daily amounts based on mean monthly values which are typical for a given hydrologic regime; or
- mean daily air temperature or dew point temperatures, or daily solar radiation amounts.

The daily computed amounts are adjusted in either case by a function to account for daily or the selected period rainfall that would reduce the potential rate of evapotranspiration. In the application of the SSARR model, the mean daily amounts computed through use of mean monthly amounts are most commonly used.

Baseflow Infiltration Index (BII) - The Baseflow Infiltration Index used in the SSARR watershed model provides a means for computing the relative proportion of the water available in the surface layers of the soil mantle that enters the ground water aquifers as deep percolation. Under the principle of "generated runoff", as defined above, all water which is not lost to the atmosphere by evapotranspiration, or the permanent loss by soil moisture increase, is available to runoff with a time delay function. Conceptually, the model considers the time delay to occur in three zones, namely surface, sub-surface and baseflow. The long time delay caused by base flow infiltration represents that portion of the water which is in transitory storage for several months (or possibly years under certain circumstances).

Surface-Sub-Surface Flow Index (S-SS) - In the model, the surface sub-surface (S- SS) flow separation index deals with the water excess which is generated from the residual after soil moisture and transpiration losses and base flow infiltration have been satisfied. Normally, the "direct" runoff is considered to be the result of the percolation of "free" water through the upper layers of the soil mantle. This could be in the zone up to a maximum of depth of 50 cm below the ground surface termed as sub-surface flow. When the water input rate exceeds the capacity of the sub-surface zone to transmit water under gravitational force, the residual water excess amount is considered to occur directly on the ground surface or the upper few centimetres of the soil mantle. The surface-sub-surface (S-SS) flow separation is a means for defining the relative portion of the direct runoff that contributes to each portion, as a function of input rate.

The S-SS function is usually specified as a nonlinear function whereby the lower rates of input provide water excess primarily in sub-surface zones while high input rates are predominantly on the surface runoff. The time delay functions for routing surface and sub-surface flow are specified to represent the difference in storage times for each of the two zones.

Watershed outflow transformation by polyphase routing for each flow component input (NP, TS) - The water excess values computed for each time period in each of the three flow components (surface, sub-surface and baseflow) must be transformed from values computed as input rates to time-distributed values of streamflow. In the SSARR model, this transformation is accomplished by polyphase routing, whereby the input rates expressed as cm per period are converted to equivalent values of steady-state outflow, expressed as cubic metres per second for the particular drainage area. The routing is performed through the use of the flow continuity equations set forth in the SSARR users manual (Davis, 1991).

The use of polyphase routing for this type of transformation has several advantages for computerized simulation of streamflow from watersheds. These include:

- simplicity of computation;
- ease of application in trial and error reconstitution studies for determining basin runoff characteristics;
- the relatively small amount of information required to store in the computer, in order to represent the basin runoff characteristics;
- the completely flexible means for representing time delays to runoff either for short term flood runoff on relatively small tributaries, or for long-term base flow on ground water discharge for large river system;
- the convenient means for preserving the continuity of flow at any specified point in time, for "stop action" or "instant replay" capabilities, particularly in its use for day-to-day streamflow forecasting;
- the flexibility of providing virtually any desired shape of runoff characteristics such as a unit distribution of known characteristics;
- the assured preservation of continuity of flow for any computed runoff excess, and;
- the ability to represent non-linear response for runoff from a watershed.

4.1.1.5 Methodology for Calibration and Validation

When using the snow band option in the SSARR model a typical subbasin requires up to 68 parameters to describe its hydrologic characteristics. Of these parameters, up to 47 are permanent characteristics of the model and include the station name and other parameters that can be measured, such as the drainage area. The remaining parameters must be calibrated or the defaults in the model accepted.

In general, the optimization of the various parameters is normally accomplished by trial-and-error reconstitution studies of historical streamflow data. These studies are generally performed using several years of historical data. The various parameters are tested to achieve the best fit of computed and observed streamflow. It is normally assumed that the physical factors affecting runoff are nonchanging over a period of years, so that the parameters and functions used in the model are fixed as a given set of values for the entire study period. The degree of fit between computed and historical streamflow is determined either visually by inspection of graphical plots of the data or by graphical or statistical methods. The principal objective is to achieve consistency over a wide range of hydrologic conditions to eliminate bias between high and low periods of streamflow and to achieve relatively uniform consistency for the years being studied.

The overall water balance for particular study areas should represent as closely as possible the known or expected values of precipitation, evapotranspiration, soil moisture and ground water condition that are characteristic for the climatological and hydrological regime of the area. The main objective of reconstitution studies is first to achieve a water balance by adjusting the following parameters:

- precipitation weighting for estimating basin precipitation from index station values;
- SMI function, in terms of total soil moisture index values and the shape of the SMI function;
- ETI values, based on observed or estimated amounts which properly reflect the seasonal or daily variation; and
- BII function, representing the portion of water input which contributes to base flow.

When the overall water balance is achieved, the refinements in timing can be taken into account by adjusting the polyphase routing parameters for each component (surface, subsurface and base flow), and by adjusting the S-SS flow separation function.

Once a simulated flow series is obtained using this procedure it is compared with the actual hydrograph for the basin. This comparison is then reviewed and the parameters are adjusted using the following guidelines:

- If the simulated hydrograph is consistently higher (or lower) than the actual and the difference is more pronounced over time, then the estimation for evaporation may be too high (or low). The slope of the SMI curve should be increased (or decreased).

- If the slope of the simulated hydrograph recession curve is higher (or lower) than the actual hydrograph, then the base flow is too low (or high). The slope of the BII curve should be reduced (or increased).
- If the simulated maximum discharge is higher (or lower) than the recorded maximum discharges and, if the runoff volumes are close (indicating that the SMI curve is good), then surface runoff is too large (or too small) with respect to the subsurface flow. The slope of the SS-S curve should be reduced (or increased).
- If the simulated hydrograph is obviously higher than the recorded hydrograph following a light rainfall in the simulation period, then the initial SMI should be reduced.
- If the shape of the simulated hydrograph is sharper than the recorded hydrograph following a light rainfall in the simulation period, then the maximum for the BII curve should be increased.

Once the model has been calibrated for rainfall only events, the snowmelt parameters are adjusted using the following additional guidelines:

- If the upward slope of the simulated hydrograph is too flat (or too sharp), then the melt rate should be increased (or decreased).
- If the simulated flood starts later than the recorded flows, then the base temperature should be increased.

- If the volume of the simulated hydrograph is too large (or too small), then the weighting of the snow course or precipitation stations may need to be decreased (or increased).

This description is a brief description of the steps necessary to calibrate the SSARR model. However, it is obvious that an adequate calibration of the model requires a great deal of knowledge about the model and about the watershed. It should also be noted that the calibration is user driven and requires frequent adjustments to be used effectively in an operational forecasting mode.

4.1.1.6 Model Setup for the Humber River Basin

The initial setup for the SSARR model was carried out as part of the study to develop flood risk areas (Cumming Cockburn, 1984). Further development was carried out to assess the possibility of using the SSARR model to forecast flows on the Humber River using the high flow event that occurred in June 1984 (Cumming Cockburn, 1985). Both of these studies recommended that the data collection network required improvement if the model was to produce accurate flow forecasts. In 1990 under a cost sharing agreement between Environment Canada, Newfoundland Department of Environment and Lands and the Deer Lake Power Company the data collection network was improved. These improvements included the installation of additional stations, the addition of temperature and precipitation sensors to existing stations and the installation of transmitters for near real time data acquisition. Once data collection network expansion was completed, additional work was carried out to develop the procedures to implement the SSARR model for flow forecasting (Cumming Cockburn, 1995). The study also used the April 1991 version of the SSARR model. The changes in this version included microcomputer processing and the use of metric units. Also included was a recalibration of the model to utilise the new data provided by the expanded data collection network. The following paragraphs summarize the setup procedures carried out through all of these studies.

The first step in applying the SSARR model to the basin was to identify and isolate the various subwatersheds, reservoirs and channel reaches required to facilitate model calibration and validation using observed discharge data. Subwatersheds were selected by separating the total watershed into relatively homogeneous hydrologic units. Consideration was also given to the meteorologic and hydrometric data available for a given subwatershed to aid in the calibration and validation of the model. In addition to the selection of the sub-watersheds, the reservoirs and river reaches which have a significant impact on the flows to downstream hydrometric gauges were also identified and modelled in order to give an accurate representation of the hydrologic and hydraulic response of the system. Figure 8 is the basin map with the SSARR subbasin discretization.

The main elements of the Humber River watershed are represented in the SSARR model by 11 subbasins, two reservoirs and one lake. Figure 10 shows a schematic representation of the basin.

Meteorologic Input Data

One of the useful features of the SSARR model is that it allows the distribution of data from a number of meteorologic stations to the subbasins defined in the model. This capability is particularly helpful in representing the hydrologic regimes of large basins, like the Humber River, since it accounts for the spatial variations of the meteorologic parameters.

The initial hydrotechnical study for the basin found that there was "a marginally sufficient amount of meteorologic data available for the Humber River watershed which is considered to be suitable for hydrologic modelling via the SSARR model." (Cumming Cockburn, 1984) At that time none of the data was telemetered to a central location requiring the upgrading of the system to permit using the SSARR model for operational flow forecasting. The climate station locations are listed in Table 7 and shown on Figure 9.

Table 7 - Climate Station Locations - Humber River Basin

Atmospheric Environment Service			**Department of Environment and Labour**		
Station Name	AES ID No.	SSARR Code	Station Name	DCP Platform ID	SSARR Code
Deer Lake	8401500	1500	Upper Humber above Black Brook	4812C702	BLAC
Buchans	8400698	0698	Grand Lake on Glover Island	480FC3FA	GLGI
			Corner Brook Lake at Outlet	480D1066	CORN
			Burgeo Road near Buchan's Access	480FB56A	BURG
			Grand Lake at Southwest End	4812B192	GLSW
			Indian Brook Diversion	480CF16E	INDI
			Sandy Lake at Howley Road	4812A2E4	SAND

The mean daily evaporation amounts were based on the mean monthly values. The mean monthly amounts were obtained from isohyetal maps. (Environment Canada, 1970). The values ranged from 0.25 mm/day to 3.3 mm/day with an average of 1.4 mm/day. These values were adjusted as recommended in the SSARR manual using a function to account for daily rainfall which would tend to reduce potential evaporation amounts.

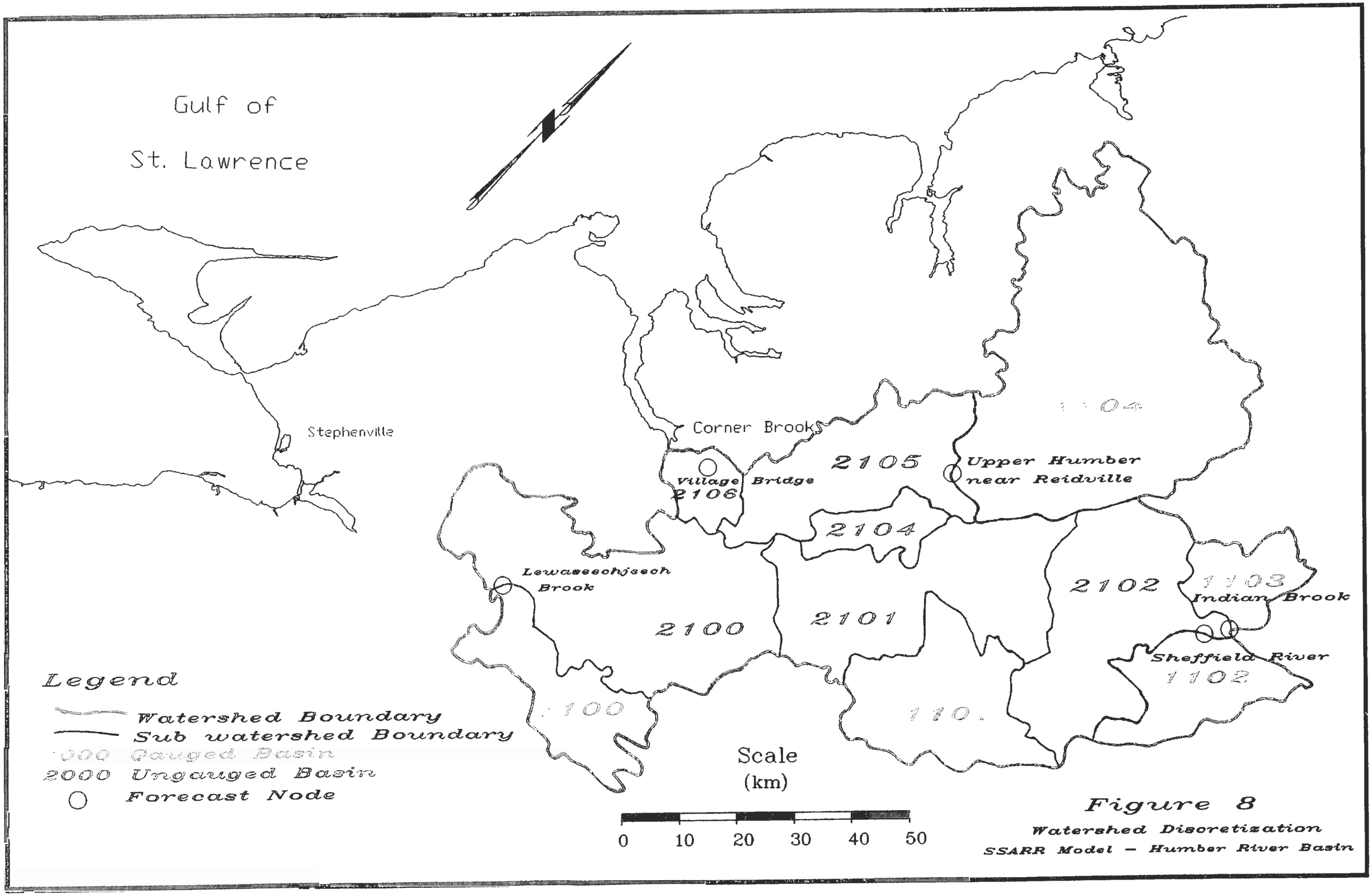


Figure 8

Watershed Discretization

SSARR Model – Humber River Basin

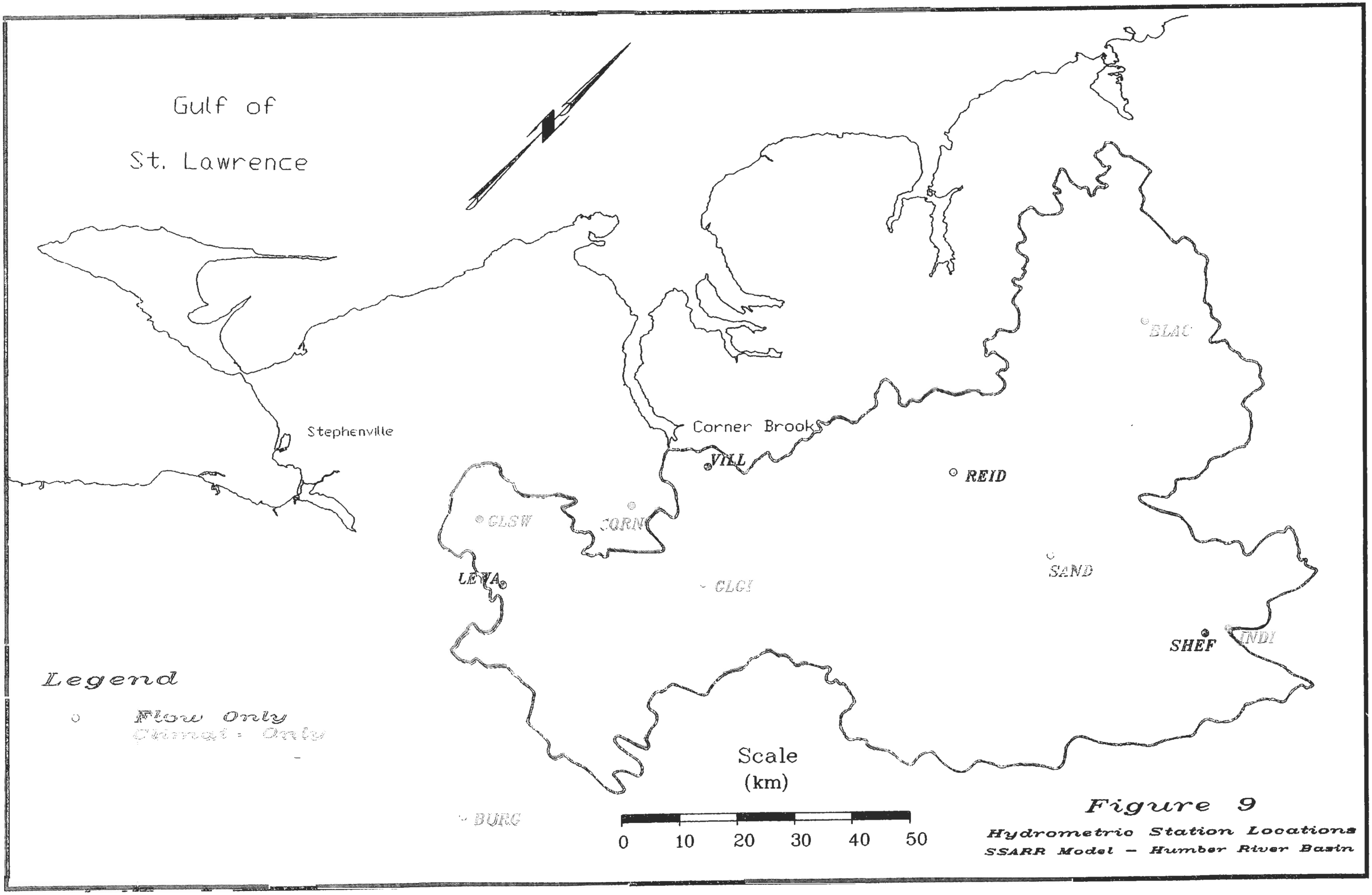


Figure 9

Hydrometric Station Locations

SSARR Model – Humber River Basin

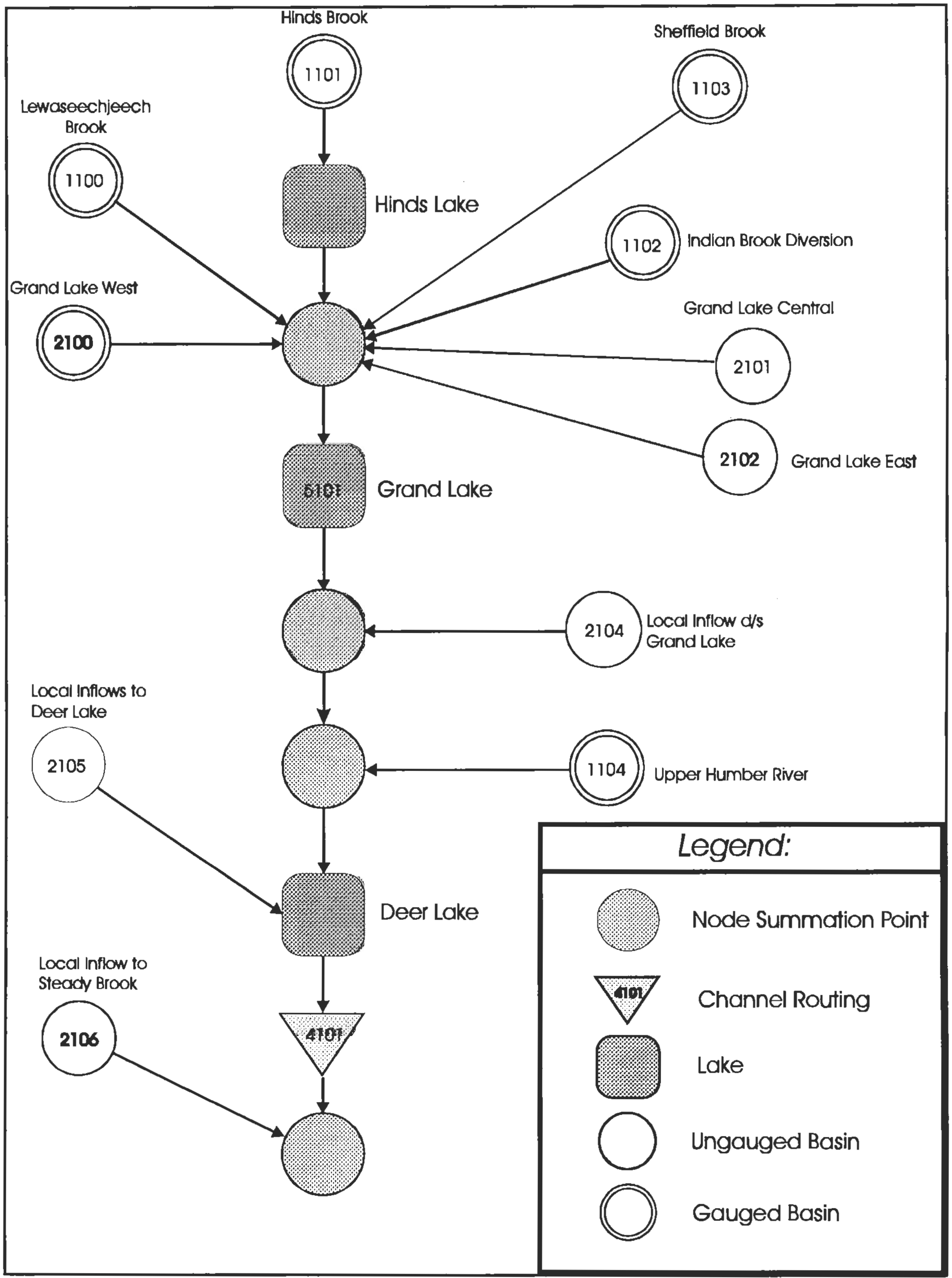


Figure 10 - Basin Schematic - SSARR Model

Table 8 - Subbasin Characteristics

Sub-basin Code and Name	Area (km^2)	Elevation Range (m)
W1100 - Lewaseechjeech Brook	470	145 - 700
W1101 - Hinds Lake	630	198 - 650
W1102 - Sheffield River	380	99 - 540
W1103 - Indian Brook Diversion	238	122 - 390
W1104 - Upper Humber River above Reidville	2108	15 - 685
W2100 - Grand Lake West	1298	85 - 640
W2101 - Grand Lake Central	690	85 - 640
W2102 - Grand Lake East	1180	85 - 568
W2104 - Local Inflow Downstream of Grand Lake	199	50 - 108
W2105 - Local Inflow to Deer Lake	640	4.5 - 458
W2106 - Local Inflow to Corner Brook	148	2 - 546

Area-Elevation Relationship

The area elevation relationship is input to the SSARR model to allow adjustments in the model for changes in the temperature as elevations in the basin change relative to climate stations. The relationship is shown in Figure 11.

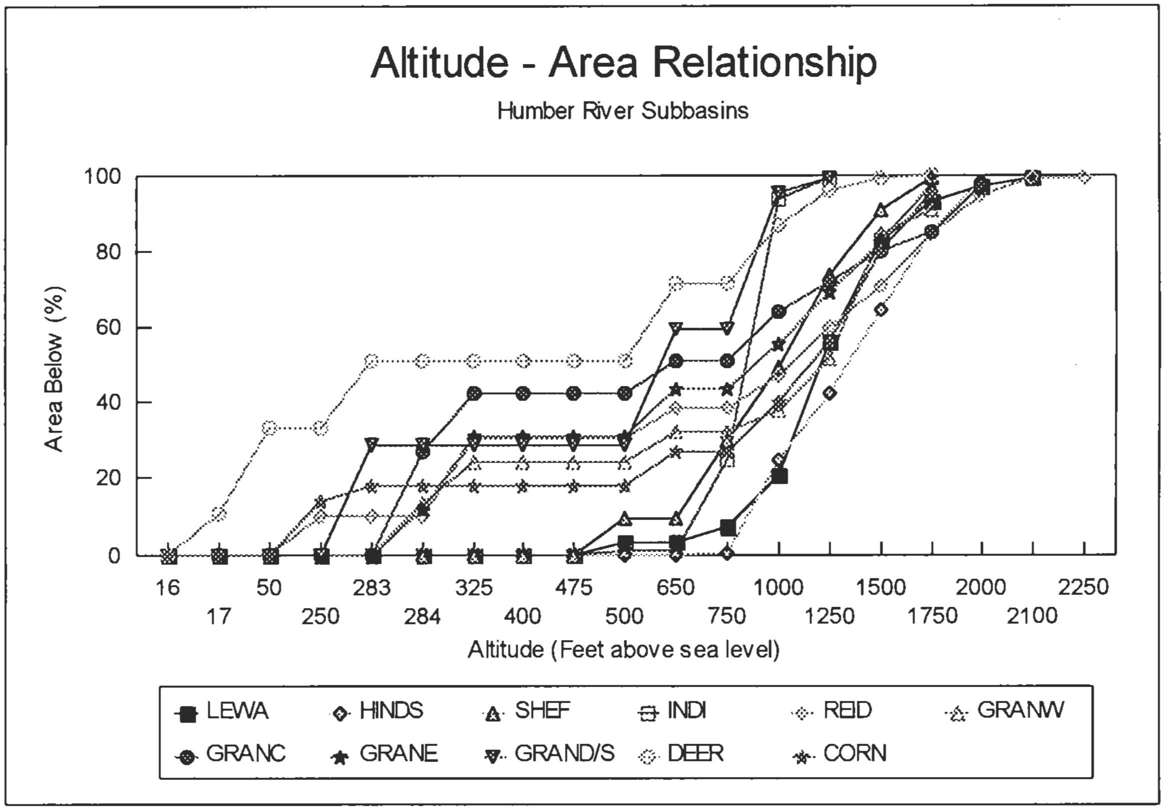


Figure 11 - Altitude-Area Relationship

4.1.1.7 Calibration and Verification

The calibrated parameters developed by Cumming Cockburn (Cumming Cockburn, 1995) were used as a basis for this study. The calibration procedure used data for the period from August 1991 to August 1992 using the available climate and hydrometric data. The model was verified using data from the period August 1992 to June 1993. Due to the problem with the data related to the ice effect on winter flows discussed in Chapter 3, no further calibration was undertaken for this study. However, a simulation run using the first two years of the data was carried out to visually evaluate the calibration. A sample of the results is presented in Figure 12.

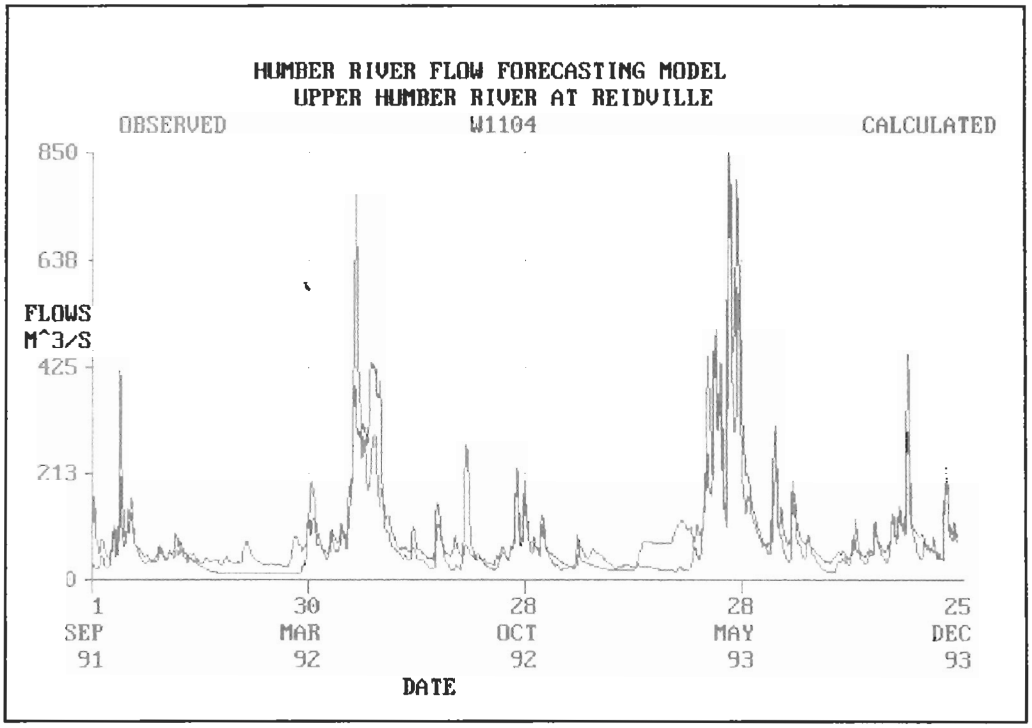


Figure 12 - Sample SSARR Calibration Run

Weighting of Meteorologic Stations

The SSARR model allows the user the option of assigning data available for several meteorologic stations to any of the subbasins defined in the model. Also, the user has the option of weighting meteorologic stations applied to each subbasin. The initial weights were determined by the consultant for the initial study (Cumming Cockburn, 1984) using Theissen polygons. In this process the temperature and precipitation data were first given the same weights. The weightings used in this study are based on the (Cumming Cockburn, 1995) study and are shown in Table 9.

Snowmelt Coefficients

Once the model yields good results for rainfall only events the calibration spring snowmelt events can be initiated. The parameters that require initial estimates are:

- rain freeze temperature;
- base temperature;
- lapse rate; and
- melt rate (for degree day method only).

The snowmelt coefficients used in this study are given in Table 10.

Routing Coefficients

The routing parameters were determined for the gauged watersheds then these values were transferred to the ungauged basins. The initial values were determined using runoff conditions that were independent of snowmelt. The routing coefficients are given in Table 11.

Table 9 - Precipitation Gauge Weights Applied to Sub-basins (Cumming Cockburn, 1995)

Sub-basin Code and Name	Climate Station								
	BLAC	**GLSW**	**BURG**	**GLGI**	**SAND**	**INDI**	**0698**	**CORN**	**1500**
W1100 - Lewaseechjeech Brook	n/a	150	150	175	n/a	n/a	n/a	n/a	n/a
W1101 - Hinds Lake	n/a	n/a	n/a	n/a	150	n/a	130	n/a	n/a
W1102 - Sheffield Brook	n/a	n/a	n/a	n/a	n/a	150	n/a	n/a	n/a
W1103 - Indian Brook Diversion	n/a	n/a	n/a	n/a	n/a	125	125	n/a	n/a
W1104 - Upper Humber River above Reidville	175	n/a	n/a	n/a	150	n/a	n/a	n/a	175
W2100 - Grand Lake West	n/a	150	n/a	140	n/a	n/a	n/a	n/a	n/a
W2101 - Grand Lake Central	n/a	n/a	n/a	140	150	n/a	n/a	n/a	135
W2102 - Grand Lake East	n/a	n/a	n/a	n/a	140	150	130	n/a	n/a
W2104 - Local Inflow Downstream of Grand Lake	n/a	n/a	n/a	150	140	n/a	n/a	n/a	120
W2105 - Local Inflow to Deer Lake	n/a	n/a	n/a	140	120	n/a	n/a	n/a	140
W2106 - Local Inflow to Corner Brook	n/a	n/a	n/a	140	n/a	n/a	n/a	140	n/a

(See Figure 9 for locations)

Table 10 - Snowmelt Coefficients (Cumming Cockburn, 1995)

Sub-basin Code and Name	Melt Rate (in/°F days)	Base Temperature (°F)	Freezing Temperature (°F) of rain	Lapse Rate (°F/1000 ft)
W1100 - Lewaseechjeech Brook	0.08	32	30	3
W1101 - Hinds Lake	0.09	30	32	3
W1102 - Sheffield River	0.07	28	35	3
W1103 - Indian Brook Diversion	0.05	28	35	3
W1104 - Upper Humber River above Reidville	0.07	32	35	3
W2100 - Grand Lake West	0.05	28	35	3
W2101 - Grand Lake Central	0.05	28	35	3
W2102 - Grand Lake East	0.05	28	35	3
W2104 - Local Inflow Downstream of Grand Lake	0.05	28	35	3
W2105 - Local Inflow to Deer Lake	0.06	28	35	3
W2106 - Local Inflow to Corner Brook	0.06	28	35	3

(See Figure 9 for locations)

Table 11 - Routing Coefficients (Cumming Cockburn, 1995)

Sub-basin Code and Name	Surface		Subsurface		Baseflow	
	NP	TS	NP	TS	NP	TS
W1100 - Lewaseechjeech Brook	3	15	2	40	2	300
W1101 - Hinds Lake	3	20	5	50	2	400
W1102 - Sheffield River	3	20	5	50	2	450
W1103 - Indian Brook Diversion	4	15	3	20	2	200
W1104 - Upper Humber River above Reidville	4	15	3	20	2	200
W2100 - Grand Lake West	3	15	3	20	2	200
W2101 - Grand Lake Central	3	15	3	20	2	200
W2102 - Grand Lake East	3	15	3	20	2	200
W2104 - Local Inflow Downstream of Grand Lake	3	10	2	20	2	200
W2105 - Local Inflow to Deer Lake	3	15	3	20	2	200
W2106 - Local Inflow to Corner Brook	3	10	3	15	2	100

(See Figure 9 for locations)

4.1.1.8 Data Format

The input data files for the SSARR model use a "card image" format. The format is based on the punch cards used when the model was developed in the 1950's. The card is divided into groups of columns specific to the type of card. For example, precipitation data is entered on a "Z4" card. In the example, in Table 12 the "Z4" identifies the card type, "BLAC" identifies the station, the next six columns the data, the "3" the time period for the data (daily in this case) and the remaining columns the actual data.

With the PC version of the program the control cards are placed in a separate data file in card format. The organization of the control card file is shown in Table 13.

Table 12 - SSARR Input Data

```
Z4 BLAC 010795 3   0.054  0.324  0.378  0.324  0.000  0.108  0.054  0.000
Z4 BLAC 090795 3   0.000  1.403  0.000  0.000  0.054  2.483  0.000  0.270
Z4 BLAC 170795 3   0.108  0.054  0.593  0.648  0.000  1.349  0.108  0.216
Z4 BLAC 250795 3   0.432  0.000  0.324  0.000  0.054  0.378  0.215
```

4.1.2 Forecasts for Specific Events

The selection of events was based on the period of record available for the DCP stations and the presence of suitable events to forecast. The spring snowmelt season was selected for the three years with data records outside of the calibration period.

The input files for each day were prepared and the model was run. Each of the input files was structured to provide a forecast three days beyond the forecast date for each of the five gauged subbasins. The results for one station are presented graphically in Figure 13.

Appendix C contains sample input and output files for the SSARR model as well as sample plots of the forecasts for each of the basins.

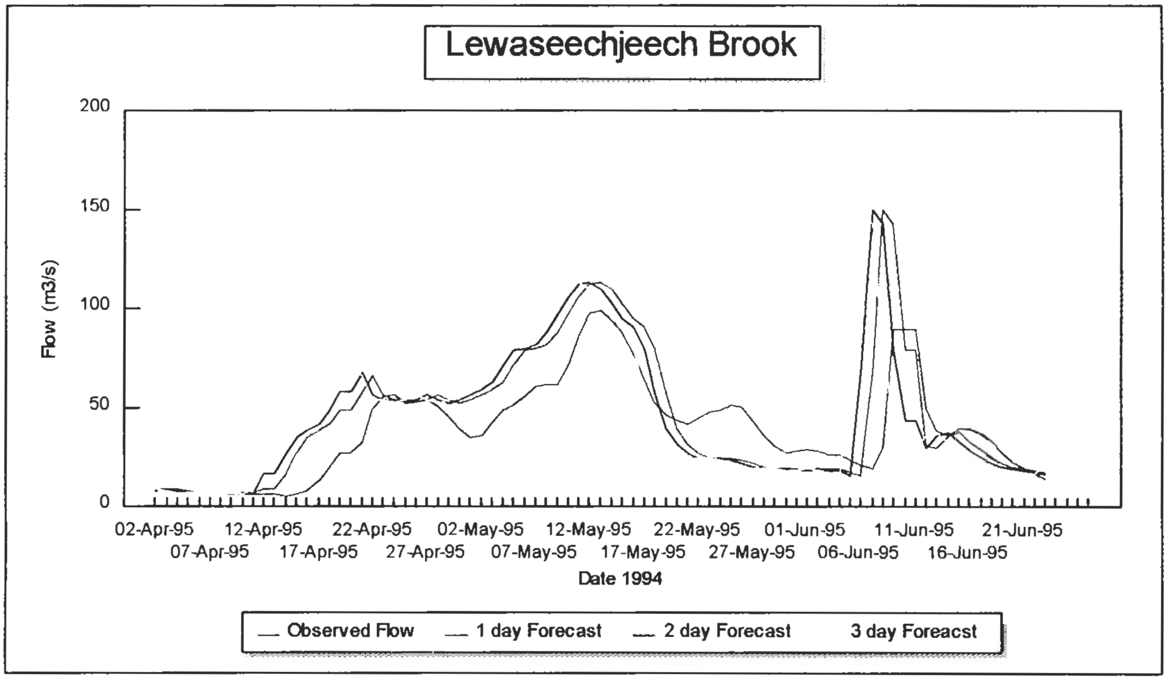


Figure 13 - Sample SSARR Flow Forecast Output

Table 13 - SSARR Control Card Organization

	Card	Description
Permanent characteristics of the model	J	Run characteristics
	CP	Characteristics of meteorologic stations
	CT	Tables of variations of basin parameters
	CB	Basin characteristics
	C3	Altitude-Area relationship
	CR	Reach characteristics
	CL	Lake and river characteristics
	C1	Reservoir storage curves
	CC	Node characteristics
	P	Network configuration
Initial conditions and input data	T	Duration of the simulation
	2B	Initial conditions of the basins
	2R	Initial conditions of the reaches
	3B	Revision of initial basin conditions
	4D	meteorologic data (observed and forecast)
	5	Temporal distribution of rain and temperature
	6	Hydrometric data
Control of Output	PR	Printing format for numeric output
	PQ	Printing format for graphic output
	END	end of file

4.2 Statistical Model (Dynamic Regression)

Dynamic regression models denote single equation regression models that combine time series oriented dynamic features with the effects of explanatory variables (Goodrich, 1989). The ARIMA type of model is purely dynamic. Its forecasts depend solely on the propagation of random shocks forwards in time. In addition to modelling this propagation process, a dynamic regression model must also account for the casual influences of the explanatory variables.

Dynamic regression may be used when:

- the datasets are long enough and stable enough to support a correlational model, and
- the explanatory variables increase the accuracy of fit in a meaningful way

4.2.1 Model Background

Much of the current theory and development of time series modelling goes back to Box and Jenkins (Box and Jenkins, 1976). In this text, Box and Jenkins referred a technique called the "combined transfer function-disturbance" model. Pankratz (Pankratz, 1991) used Box-Jenkins modelling technique calling it "dynamic regression".

The ordinary least squares dynamic regression model takes the form shown in Equation 2:

$$\phi(b)Y_t = \beta Z_t + \epsilon_t \tag{2}$$

where $\phi(b)$ = autoregressive polynomial
Y_t = dependant variable at time t
β = coefficient of i'th exogenous variable $Z_t^{(I)}$
Z_t = vector of exogenous variables at time t
ϵ_t = errors where the errors are $NID(0, \sigma^2)$, ie. normally and independently distributed with variance σ^2

Often the residuals from Eq (1) are correlated, contrary to the assumption of independence. A significant correlation in the residuals indicates that the historical data are related to current data and may be useful in predicting future values. With streamflows, it is quite conceivable that tomorrows flows are related to flows and precipitation that occurred one, two or three days ago. For this reason, error autocorrelation must be seriously considered in model construction.

Error autocorrelation can be detected by examining the autocorrelation function (ACF) results, using the Ljung-Box Q-test, the Durbin-Watson test, or by using other tests. The ACF determines the If autocorrelations are found this may indicate that one or more lags should be added to the model or that additional exogenous variables should be added. In the modelling process exogenous variables are variables, such as temperature or precipitation that may not be initially included in the model. In some cases, both may be required.

One method to improve the model dynamics is to use the Cochrane-Orcutt (Goodrich, 1989) model to add new parameters. With this method. Equation 1 is replaced by the following pair of equations:

$$\phi(b)Y_t = \beta Z_t + \omega_t \tag{3}$$

$$R(b)\omega_t = \epsilon_t \tag{4}$$

where $\phi(b)$ = autoregressive polynomial

Y_t = dependant variable at time t

β = coefficient of i'th exogenous variable $Z_t^{(i)}$

Z_t = vector of exogenous variables at time t

R(b) = polynomial in the backward shift operator

ω_t = raw residual at time t

ϵ_t = errors where the errors are $NID(0, \sigma^2)$, ie. normally and independently distributed with variance σ^2

These two equations can also be written as a single equation as shown in Equation 5:

$$R(b)(\phi(b)Y_t - \beta Z_t) = \epsilon_t \tag{5}$$

4.2.2 Forecast Model Development

The objective of this section is to create a dynamic regression based model for each of the five gauged basins within the Humber River Basin. Each of the basin models, in addition to being able to provide a reasonably accurate flow forecast, must conform to the theory of parsimony. In other words, the model should use the fewest number of coefficients that provide an adequate explanation of the data.

This section will describe the procedure used to develop the dynamic regression models. For this model process, the basin was represented by the schematic shown in Figure 14. The procedure was carried out using the Forecast Pro Software package (Business Forecast Systems Inc, 1993) using data prepared as described in Chapter 3. The step by step procedure used to develop the models is shown in Figure 15. The first step was to select an initial model. Since runoff is related to precipitation, the closest precipitation station was selected as the first independent variable. The next step is to fit the regression coefficients. The significance level of the variables are listed, with insignificant correlation flagged. If all the variables are significant the diagnostics are run. The diagnostics are checked for lagged variables and autoregressive terms. As part of the diagnostics, the software suggests a new lagged variable or autoregressive terms to add to the model. The procedure is continued until a satisfactory results is achieved. Figure 16 shows the change in the ACF using from: 1) just 0the precipitation term, 2) adding the flow lagged by one time step; and then 3) adding the precipitation lagged by one time step and an autoregressive term.

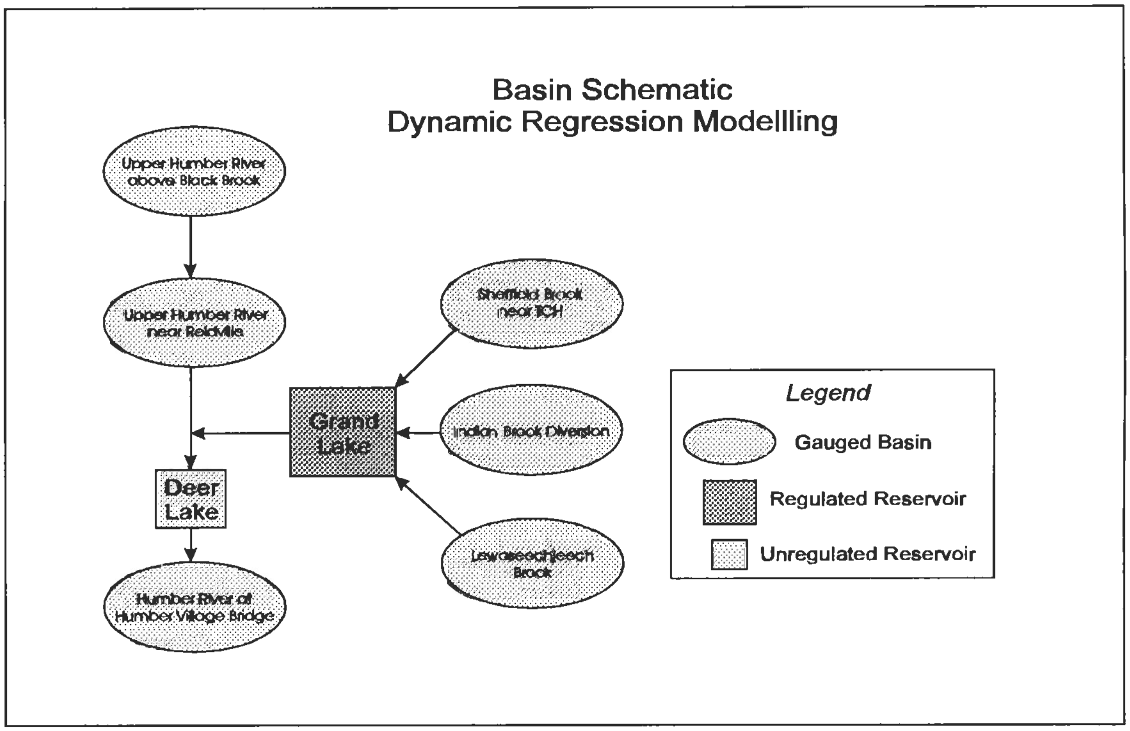


Figure 14 - Basin Schematic - Dynamic Regression Model

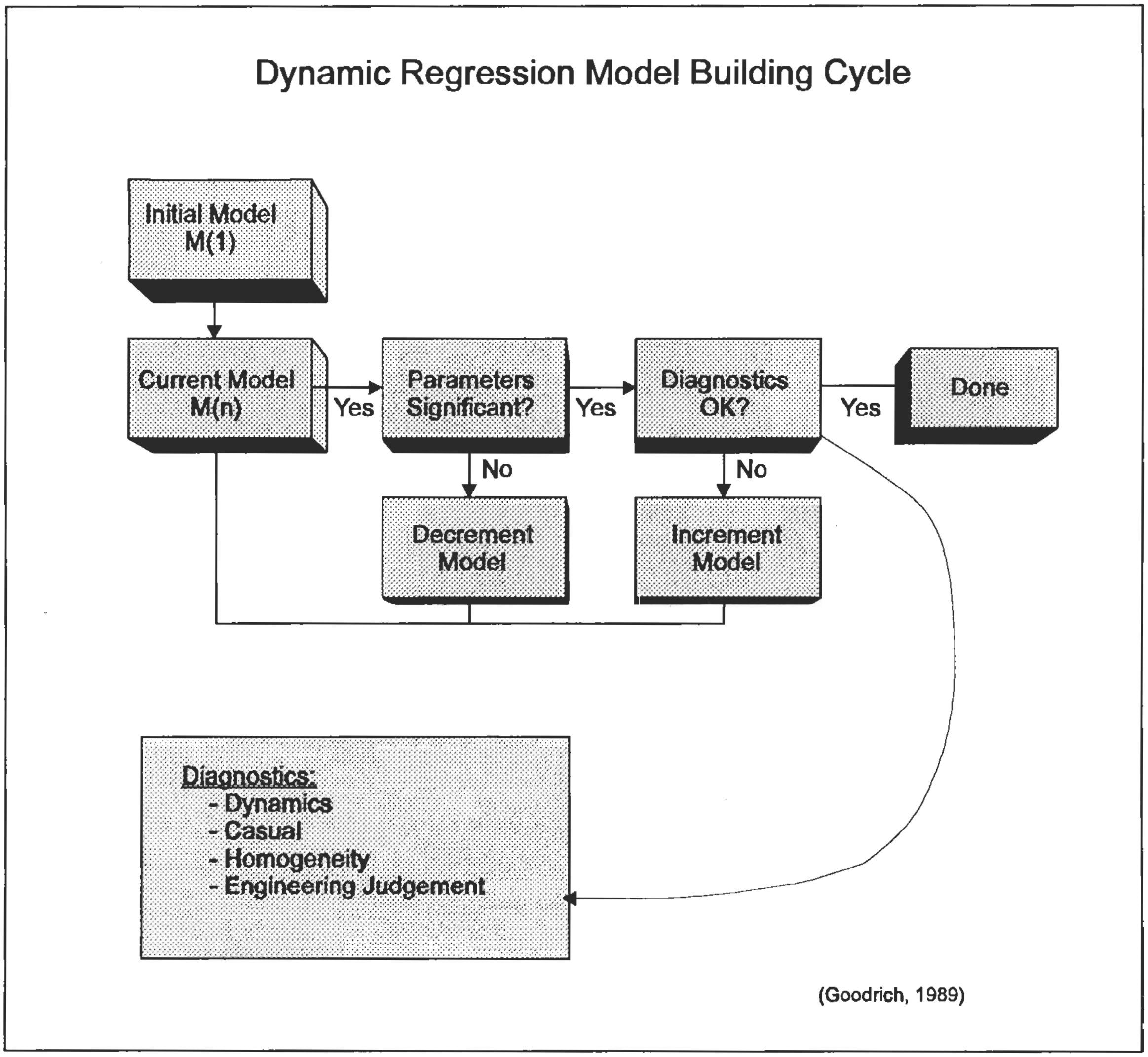


Figure 15 - Dynamic Regression Model Building Cycle

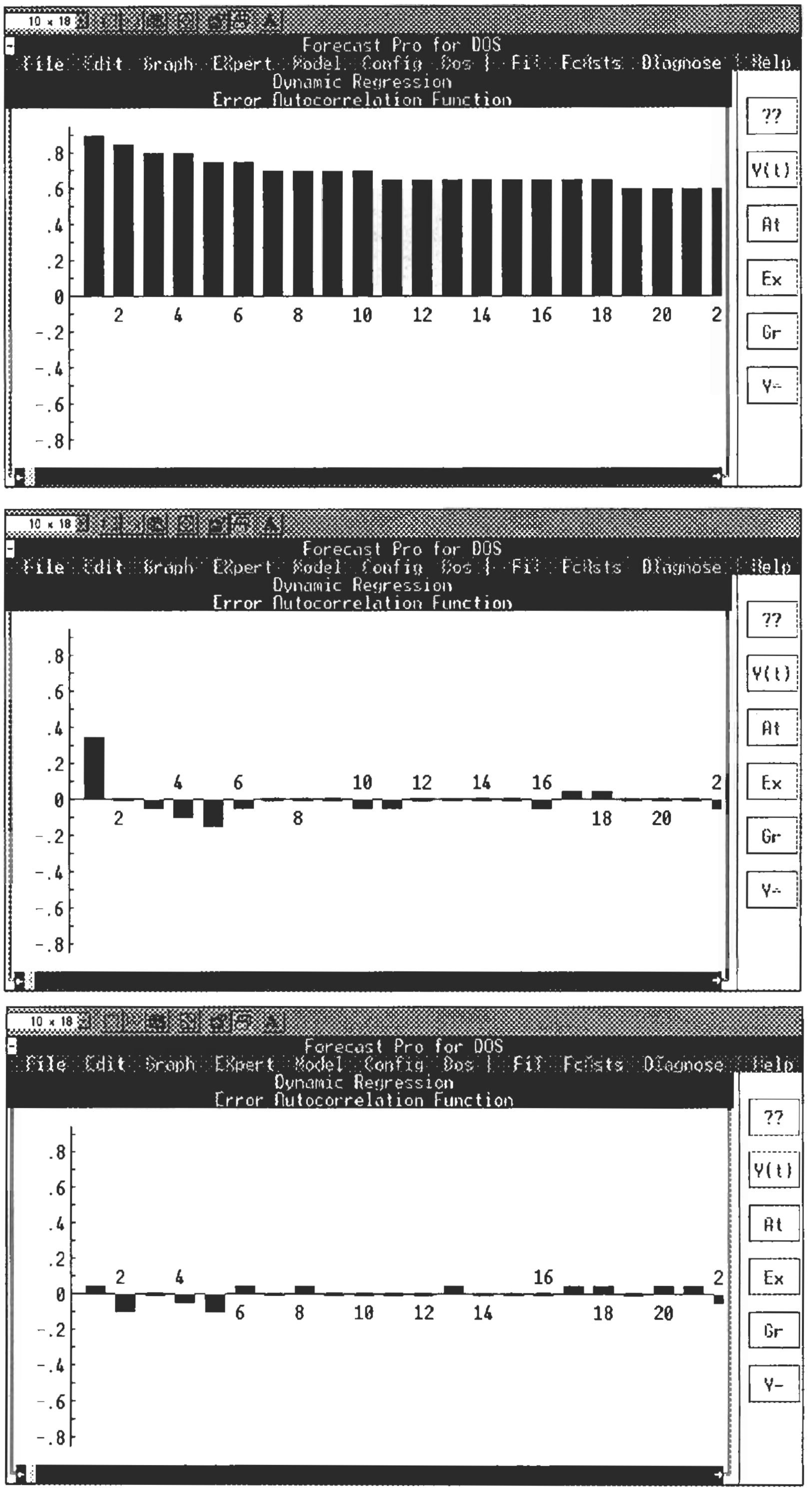


Figure 16 - Sample ACF Changes

4.2.3 Model Diagnostics

This section briefly describes the diagnostic tests that are performed by the Forecast Pro software during model development. The test can be performed each time a model is fitted to the data and the results can be use to add or remove terms from the model. The first group looks at the terms related to the dynamics of the model - the lags and the autoregressive terms. The second group looks at the inclusion or exclusion of variables. The variables can be either the inactive independent variables, the time related variables or a constant term.

4.2.3.1 Dynamics Specifications:

_AUTO[-n]	This test determines whether a Cochrane-Orcutt autocorrelation error term of lag **n** should be added to the model. The test is performed for the first twelve lags and the first two seasonal lags. If the term is already in the model the test is omitted.
Y[-n]	This test determines **n'**th lag of the dependent variable term should be added to the model. The test is performed for the first twelve lags and the first two seasonal lags. As with the previous test, the test is omitted if the term is already in the model.

The software provides a recommendation to add a specific new term to the model. The process can be repeated until the modeller is satisfied with the results.

4.2.3.2 Variable Specification

Excluded Variables: Each inactive variable in the tableau is evaluated tested with a Lagrange multiplier test.

Time Trend: This test uses alternative hypothesis testing to determine whether a linear time trend improves the fit of the model. A significant test result does not necessarily indicate that a time trend variable should be added but may indicate that there is a problem with the model's dynamics or with an excluded term.

Constant Term This test evaluates whether a constant term improves the fit of the model.

Lagged independent variables Each of the independent variables in the current model is tested.

4.2.3.3 Custom excluded variable tests

This test allows the user to test whether groups of excluded terms will improve the fit of the model. Sometimes combinations of excluded variables will be significant even if they are insignificant separately.

4.2.3.4 Standard diagnostics

Each time a model is fitted, a group of standard diagnostics are displayed. A typical set of data is displayed in Table 14. Included in the this group are the mean, standard deviation. Others are described below:

Table 14 - Sample Diagnostics from Dynamic Regression Model

```
Forecast Pro for DOS Version 2.00A
Sun Jul 06 19:03:30 1997

Forecast Model for FLOW
Regression(3 regressors, 0 lagged errors)

Term           Coefficient  Std. Error  t-Statistic  Significance
------------------------------------------------------------------
----
FLOW[-1]         1.344443     0.021865    61.488101      1.000000
FLOW[-2]        -0.387660     0.021795   -17.786622      1.000000
TOT_PREC         0.113747     0.009149    12.432964      1.000000

Standard Diagnostics
------------------------------------------------------------------
----
Sample size 1626                  Number of parameters 3
Mean 11.25                        Standard deviation 11.19
R-square 0.9398                   Adjusted R-square 0.9398
Durbin-Watson 1.979               ** Ljung-Box(18)=53.38 P=1
Forecast error 2.746              BIC 2.762 (Best so far)
MAPE 0.1313                       RMSE 2.743
MAD 1.32
```

Mean Absolute Deviation: This value is the measure of the average of the absolute discrepancies between the actual and fitted values in a given time series. The value is calculated using Equation 6:

$$MAD = \frac{\sum_{i=1}^{n} |Y - \hat{Y}|}{n} \qquad \textbf{(6)}$$

Standard Forecast Error: The standard error is the root mean square of the actual data minus the fitted values given by Equation 7:

$$SE = \sum_{i=1}^{n} (Y - \hat{Y})^{2} \quad \textbf{(7)}$$

R-square: This is a measure of the variance that is explained by the model. In the example, 93.98% of the variance is explained by the model.

Bayesian Information Criterion : The BIC statistic is a measure that is used to evaluate the parsimony of a model. This statistic reward for goodness of fit based on the mean square error and penalizing for complexity based on the number of parameters. The object of model building is to minimize the BIC. The form of the equation used in Forecast Pro is given in Equation 8:

$$BIC = sT^{\frac{n}{2T}} \quad \textbf{(8)}$$

where T = number of sample points

s = mean square error

n = number of parameters

Durbin-Watson test: The Durbin Watson statistic is a standard test for the presence of autocorrelation in regression residuals. Some of the assumptions for the use of this statistic (Pankratz, 1991) are:

- a constant term is present in the model,
- the series follows a AR(1) process

- the regression does not include any time lagged *Y* values
- there are no missing values

This statistic was not very useful for this study since a number of the assumption were violated. The form of the equation for the Durbin Watson *d* use in Forecast Pro is shown in Equation 9:

$$d = \frac{\sum_{t=2}^{r} (e_t - e_{t-1})^2}{\sum_{t=1}^{r} e_t^2} \tag{9}$$

Ljung-Box test: The Ljung-Box *Q* statistic (Pankratz, 1991) is used to test the overall autocorrelation of the fitted errors of a model. This test is an improvement of the Box-Pierce (portmanteau) test. The statistic is a weighted sum of squared autocorrelations and consequently zero only when every autocorrelation is zero. The value of *Q* increases with more autocorrelation. The weights are so that *Q* approximates $X^2(L - n)$, which is the Chi-square with $L - n$ degrees of freedom. The form of the equation used in Forecasts Pro is given in Equation 10:

$$Q = T(T + 2)\sum_{i=1}^{L} \frac{r_i^2}{(T - i)} \tag{10}$$

where T = number of sample points

r_i = i'th autocorrelation coefficient

L = number of correlation coefficients

4.2.4 Model Results

As discussed in Chapter 3.1, the data used in the analysis was obtained from NDOEL files. For this part of the analysis, the data was formatted in a Lotus 1-2-3 spreadsheet for each basin so that they were compatible with the Forecast Pro software. The procedure used was to start with the simplest form of the regression relationship and build on that relationship until the best fit to the data was obtained. Data for the period of record from 1991 to 1994 was used in the analysis to develop the dynamic regression relationship for each basin. The final set of coefficients developed in the analysis are presented in Table 15. A sample of the historic data and the data fitted with using the dynamic regression model for one of the gauged basins is presented in Figure 17.

4.2.5 Generation of Forecasts

Once the coefficients for the dynamic regression models were selected as described in the previous section, forecasts were generated. The nature of the basin had to be taken into account to set up the order for generation. The Sheffield, Lewaseechjeech and Indian Brook Diversion basins could be generated independently. Since three of the gauges, the Upper Humber River above Black Brook, the Upper Humber River near Reidville and the Humber River near Village Bridge are linked physically as described in the basin schematic

Table 15 - Form of Dynamic Regression Models

Sub-basin Name	**Form of Dynamic Regression Equation**
Lewaseechjeech Brook	_CONST + a PREGLGI + b FLOW[-1] + c FLOW[-2] + d FLOW[-3]; where: _CONST = 0.149490 a = 0.244776; b = 1.664595; c = -1.046278 d = 0.336051
Sheffield Brook	_CONST + a PREIND1 + b FLOW[-1] + c FLOW[-2]; where: _CONST = 0.348126 a = 0.041432; b = 1.432156 c = -0.462627
Indian Brook Diversion	_CONST + a PRECIND1 + b FLOW[-1] + c FLOW[-2]; where _CONST = 0.448615 a = 0.118321; b = 1.297039; c = -0.362822
Upper Humber River above Reidville	_CONST + a PRESAND + b FLOW[-1] + c FLOW[-2] + d FLOBLAC + e _AUTO[- 1]; where: _CONST =14.380586 a = -0.210896; b = 0.739059; c = -0.376750 d = 1.055177; e = 0.884608
Upper Humber River above Black Brook	_CONST + a PRECBLAC + b FLOW[-1] + c FLOW[-2]; where: _CONST = 1.254866 a = 0.558944; b = 1.238943; c = -0.315646
Humber River at Humber Village Bridge	_CONST + a PREBLAC + b FLOW[-1] + c FLOREID + d FLOBLAC + e _AUTO[- 1]; where: _CONST = 21.697851 a = 0.187210; b = 0.859492; c = 0.196181 d = -0.055710; e = 0.525336

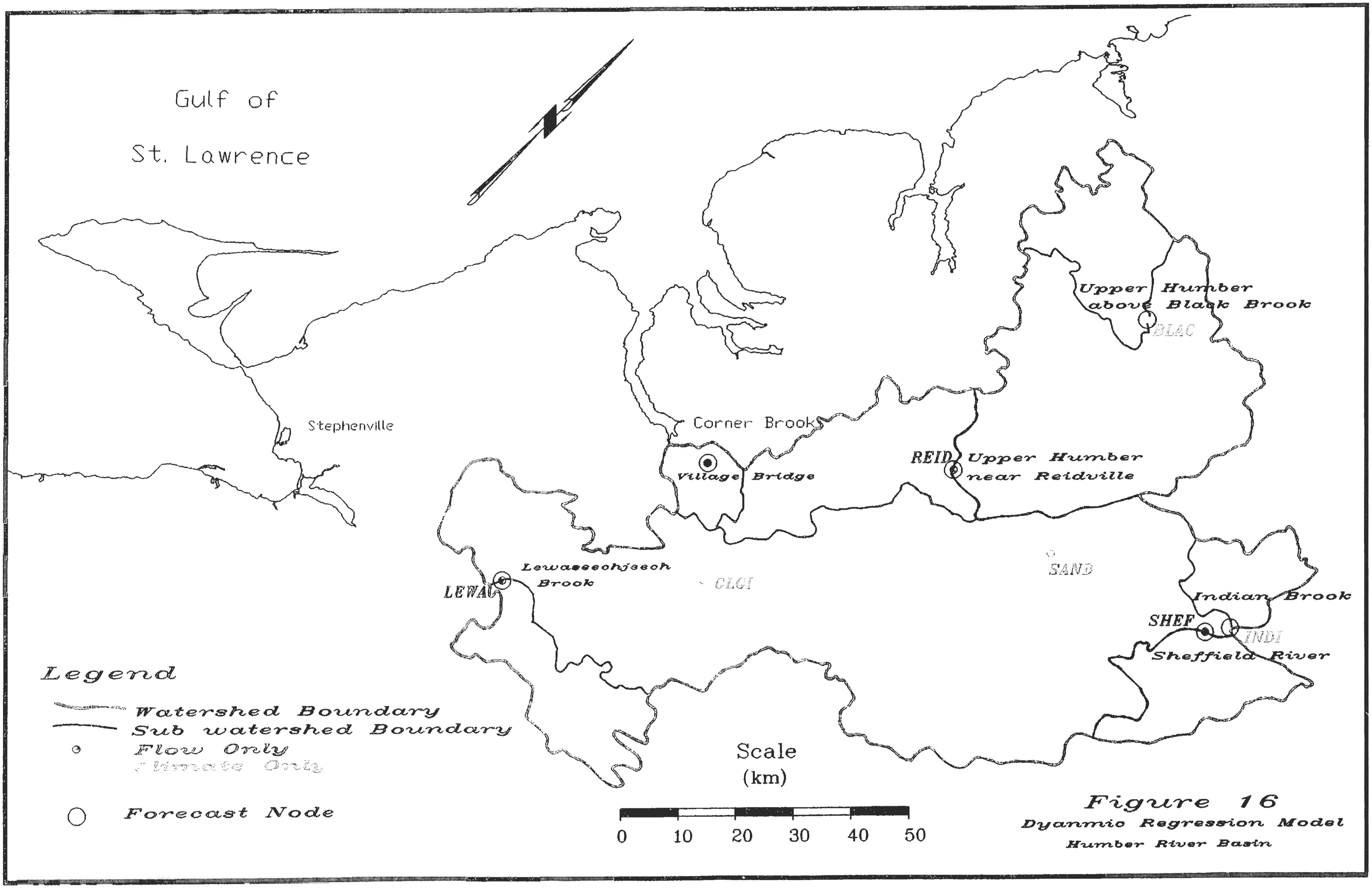


Figure 16
Dyanmic Regression Model
Humber River Basin

and basin map, Figures 14 and 17 respectively, forecasts for these stations had to be generated in an upstream to downstream order. That is, the forecasts for the Black Brook basin were used to generate forecasts for the Reidville basin and the Reidville forecasts were then used to generate forecasts at Humber Village Bridge.

Forecasts were generated for the same time periods used for the SSARR forecasts discussed in Section 4.1. However, for this analysis, the Upper Humber River above Black Brook basin was added. This basin was not set up separately in the SSARR model since a forecast was not required for this area. Since the flow and climate data was available and the model development for the dynamic regression equations described earlier in this chapter showed that flows from this basin were significant in forecasting downstream flows, it was included here.

The forecasts were generated for one day, two day and three day lead times using data up to the forecast data. The model was setup so that it did not update. That is, the forecasts for the one two and three day lead times used data up to the forecast date only. The results for the one, two and three day forecasts for the Humber River at Humber Village Bridge and the Upper Humber River at Reidville sites are shown in Figures 18-21. A sample output file from Forecast Pro as well as forecast plots from the other stations are presented in Appendix D.

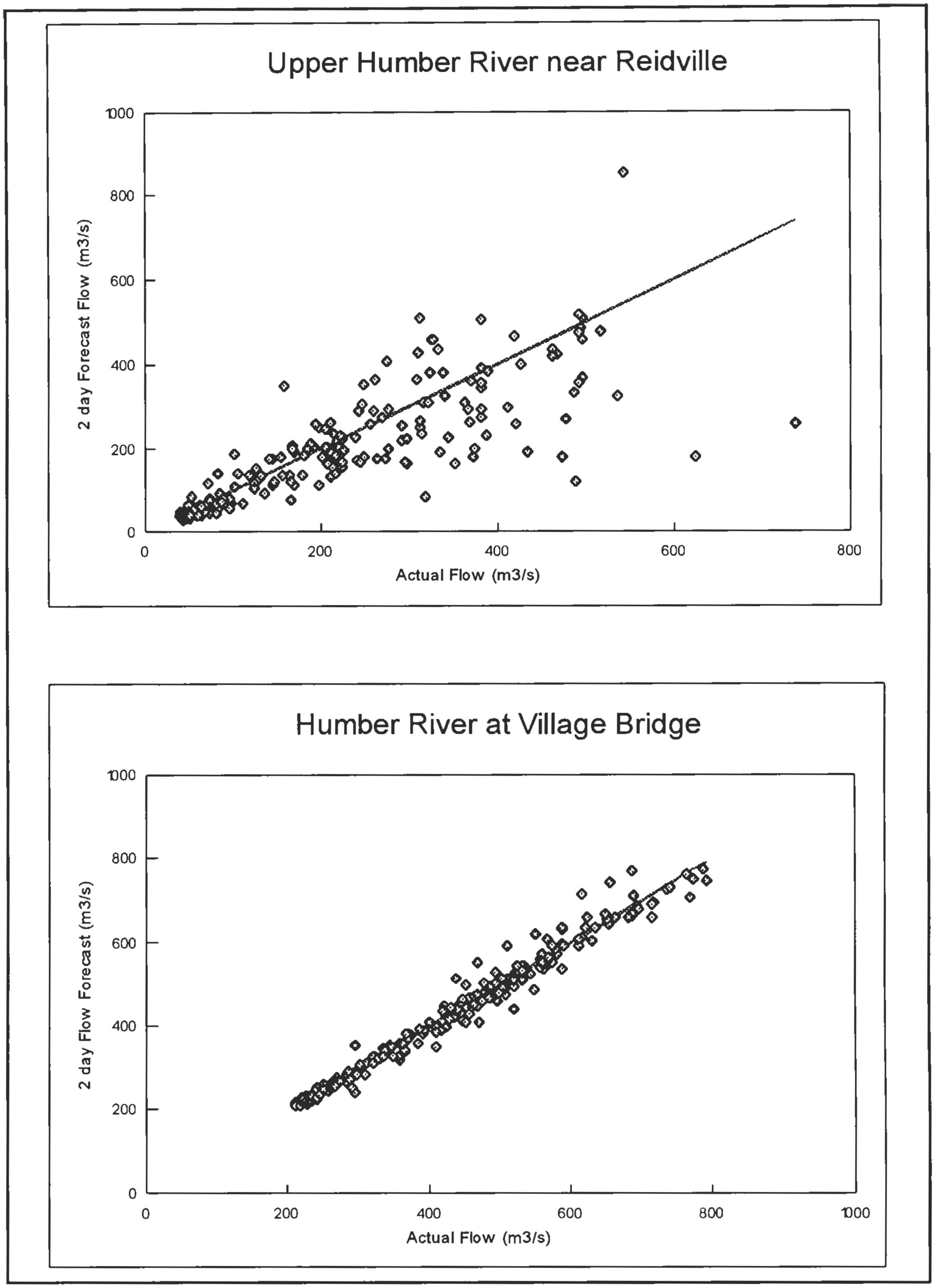


Figure 18 - Dynamic Regression Model Results vs Recorded Flows

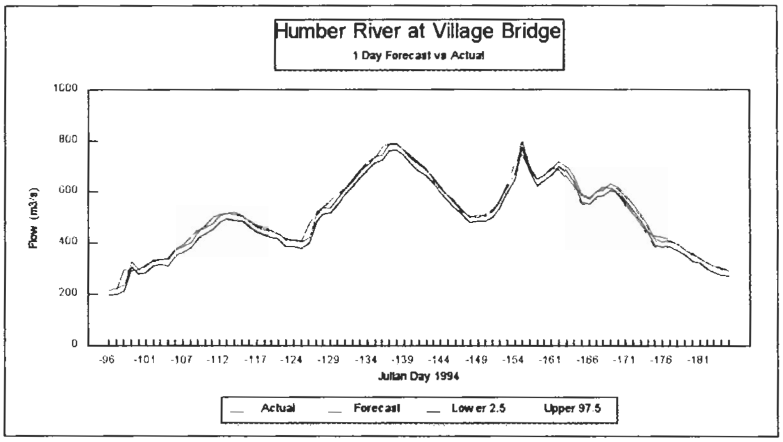


Figure 19 - Sample One day Forecast

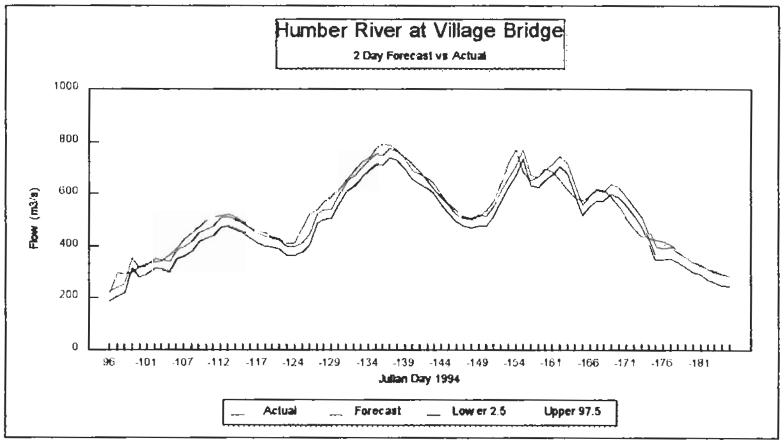


Figure 20 - Sample Two day Forecast

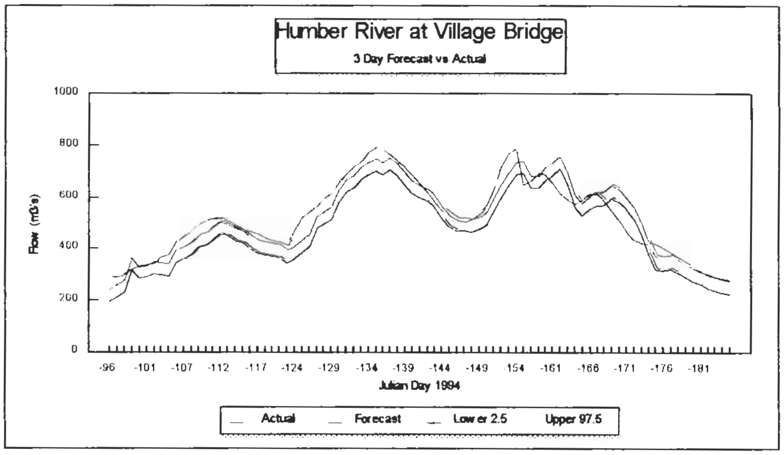


Figure 21 - Sample three Day Forecast

5. COMPARISON OF MODELLING RESULTS

The main objective of this study was to develop a flow forecast model based on dynamic regression and to compare the results with forecasts from the SSARR model. The previous chapter described the methodology used to develop the flow forecasts by the two methods under comparison. The chapter compares the results of the two methods, discusses the significance of the results and outlines some of the reasons why the SSARR model may have performed poorly compared with the dynamic regression model

5.1 Analysis of Results

A detailed survey (Mahmoud, 1984) reviewed the relevant literature and tested many of the accuracy measures used to compare forecasts, including the mean square error, the mean percentage error, the mean absolute percentage error, Theil's U-statistic, the root mean square error, the mean error and others. Since it is important to consider forecast bias in time series analysis, the mean absolute percentage error (MAPE) method was selected to measure the accuracy of the two forecasting methods. The equation describing this method is presented in Equation 11. The results from the two methods are presented in Tables 16 and 17.

$$MAPE = \frac{100}{m}\sum_{t=1}^{m} \left|\frac{e_t}{z_t}\right|, \qquad t = 1, 2, \ldots\ldots m \tag{11}$$

Where e_t = forecast error

z_t = observed value

m = total number of observed values

Table 16 - MAPE Dynamic Regression Model

MAPE for Dynamic Regression Method

Upper Humber River Above Black Brook

	t (Julian Day)	One Day	Two Day	Three Day
1994	121-180	17.8	32.7	40.2
1995	121-176	14.0	29.1	37.6
1996	80-140	18.1	33.4	41.2

Lewaseechjeech Brook Flow Data

	t (Julian Day)	One Day	Two Day	Three Day
1994	66-166	9.2	17.4	23.4
1995	96-176	7.5	17.2	26.0
1996	36-71	9.4	19.6	27.8

Sheffield Brook Flow Data

	t (Julian Day)	One Day	Two Day	Three Day
1994	96-176	5.9	10.4	14.9
1995	101-181	5.9	11.8	17.1
1996	36-73	6.3	12.1	17.7

Indian Brook Diversion Flow Data

	t (Julian Day)	One Day	Two Day	Three Day
1994	56-176	10.5	18.9	25.9
1995	121-171	8.6	17.1	24.3
1996	46-81	10.5	19.8	27.9

Upper Humber River at Reidville Flow Data

	t (Julian Day)	One Day	Two Day	Three Day
1994	121-179	13.8	25.0	33.6
1995	91-175	9.0	19.3	28.5
1996	81-140	11.0	22.1	31.1

Humber River at Village Bridge Flow Data

	t (Julian Day)	One Day	Two Day	Three Day
1994	96-185	2.4	4.6	6.4
1995	91-180	1.2	2.8	4.3
1996	80-140	1.7	3.5	5.4

Table 17 - MAPE Results for SSARR Model

MAPE for SSARR Model				
Lewaseechjeech Brook Flow Data				
	t (Julian Day)	One Day	Two Day	Three Day
1994	66-166	36.5	30.0	27.1
1995	96-176	46.1	47.9	50.3
1996	36-71	109.1	94.2	84.4
Sheffield Brook Flow Data				
	t (Julian Day)	One Day	Two Day	Three Day
1994	96-176	46.7	46.8	47.2
1995	101-181	29.0	29.1	28.1
1996	36-73	50.5	49.0	48.6
Indian Brook Diversion Flow Data				
	t (Julian Day)	One Day	Two Day	Three Day
1994	56-176	37.8	38.5	40.7
1995	121-171	85.7	88.1	90.5
1996	46-81	71.1	67.5	63.4
Upper Humber River at Reidville Flow Data				
	t (Julian Day)	One Day	Two Day	Three Day
1994	121-179	43.1	39.1	38.0
1995	91-175	51.1	52.0	53.4
1996	81-140	107.1	95.8	87.8
Humber River at Village Bridge Flow Data				
	t (Julian Day)	One Day	Two Day	Three Day
1994	96-185	22.8	23.3	23.5
1995	91-180	22.0	22.4	22.6
1996	80-140	31.8	31.9	32.4

5.2 Discussion of Results

The most obvious observation based on the MAPE criterion is that the dynamic regression model was consistently more accurate. The SSARR results, especially for the 1996 data set at Humber River near Reidville and Lewaseechjeech Brook, were far less accurate. However, as shown in Figure 22, the SSARR model was reasonably close in estimating the magnitude of the maximum flood flow, still the time of the peak is off by several days.

While the results from the SSARR model were disappointing, other studies have produced similar, and often variable, results. Figure 23 shows the results from a study (Manitoba, 1985) that compared several different deterministic models, including SSARR. In the top graph, the results are very good, while in lower graph the forecasted flows are not at all close to the actual flows.

The dynamic regression model has an advantage during time periods with consistent flows dynamic updating capabilities. This is the time when the SSARR model is at its greatest disadvantage since an inaccuracy in the SMI will carry through the whole time period.

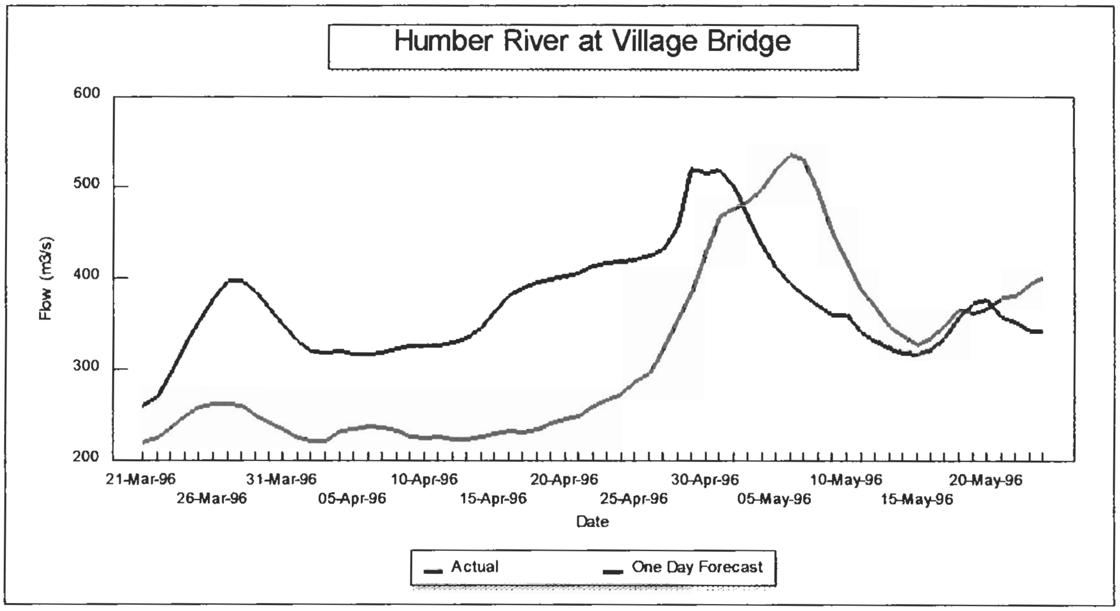
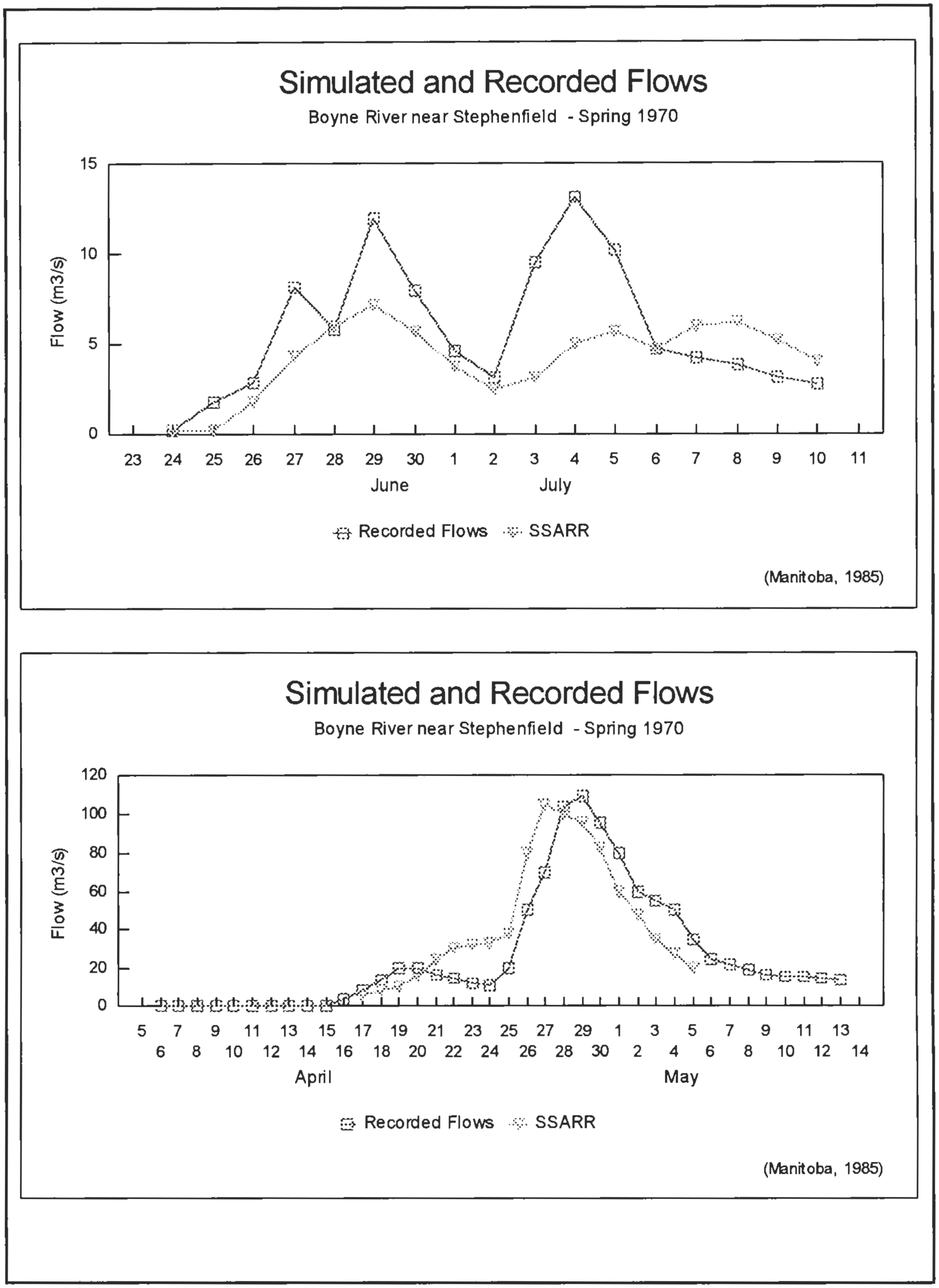


Figure 22 - SSARR Results for 1996 Data



Simulated and Recorded Flows

Boyne River near Stephenfield - Spring 1970

(Manitoba, 1985)

Simulated and Recorded Flows

Boyne River near Stephenfield - Spring 1970

(Manitoba, 1985)

Figure 23 - Results of Manitoba Model Comparison

Another problem with the SSARR model, and to some extent all deterministic models, is the lack of calibration data. In this study, the basin was divided into eleven subbasins. Of these subbasins, six have no flow records at all. The combined drainage area of these subbasins totals 4150 km^2, or about 50 percent of the total drainage area. Also, many of the parameters, such as the ETI, are based on empirical relationships that have very little field data on which to base a calibration. The real problem with these empirical relationships is separating the effects of one parameter from another during the calibration process. As discussed in Chapter 4.1.1, the calibration process is essentially a trial and error process to fit the model to measured hydrologic events. Usually only the flow, the temperature and precipitation are known at a small number of points in or near the basin. The model must take these point measurements and distribute the values over some or all of a subbasins. The location of the climate stations may not be in the optimal location to distribute the data. Many of the climate stations, such as the Indian Brook Diversion site and the Upper Humber River above Black Brook were installed at the hydrometric station for convenience of maintenance and operation. These locations are at the extreme downstream end of the basin at the lowest elevation in the basin so the climate data collected at these sites may not accurately represent the conditions upstream. Generally lower elevations tend to receive more precipitation than higher elevations. This is compensated for in the model using the meteorologic weighting factors. Adjustments in the SMI or the other parameters could achieve the same result for some events.

The dynamic regression model seemed to handle the problem with the erroneous flow values due to the ice effect discussed in Chapter 3. The initial data analysis in the dynamic regression model development identified the data as stationary but seasonal. Since the ice cover is a seasonal phenomenon the fitted model compensated for the differences. The SSARR model may also forecast for this season provided that calibration is carried out during the same season.

The modeller requires a thorough knowledge of the basin's hydrology to calibrate the SSARR model. It is essential to determine which parameters in the model to adjust to improve the fit. While the SSARR manual presents a step by step procedure for parameters adjustments, some require "dry" periods or rainfall events following a dry period to isolate specific parameters. With the variability of Newfoundland weather conditions, this is not an easy task. With the dynamic regression model, a reasonable knowledge of the basin's hydrology and a good statistical software package can provide good results with ease and economy. Fewer parameters are needed and the calibration can be easily updated as new data becomes available.

One factor that was not explored with the dynamic regression model was a method to account for spills that may occur from the Grand Lake Reservoir. These areincluded in the SSARR model and may present a difficulty with the regression model.

6. CONCLUSIONS AND RECOMMENDATIONS

The results of the dynamic regression model and the deterministic model were compared in the preceding Chapter. This Chapter presents the conclusions and recommendations.

6.1 Conclusions

The conclusions of the study are as follows:

1. The dynamic regression model provides a good alternative to the SSARR model for near real time flow forecasting for the Humber River Basin based on the results of this study. The results indicate that the dynamic regression model should be further evaluated under operational conditions.
2. The SSARR model provided a reasonable estimate of peak discharge at the Reidville and Village Bridge sites where these discharges are important for flood forecasting
3. More work needs to be done to improve the calibration of the SSARR model. Initially the present calibration parameters should be reevaluated to determine the areas that need improvement. Possibly more data collection stations, or the relocation of the existing stations would improve the model.

4. The SSARR model requires a higher level of knowledge in terms of the basins hydrology and the model application to provide good results compared with the dynamic regression model.
5. The dynamic regression models provided accurate forecasts in spite of the problem with the flow data related to ice effects discussed in Chapter 3.
6. The dynamic regression model requires less data and effort in calibration.
7. Confidence limits are given in the dynamic regression model.

6.2 Recommendations

The results of this study show that the dynamic regression model can produce accurate flow forecasts for the Humber River Basin. Since the dynamic regression model is less complicated and requires less effort to calibrate, the Department of Environment and Labour should include use of a dynamic regression model in its operational forecasting activities. Initially the dynamic regression model could be added on a trial basis to confirm that the model can reproduce results similar to this study. Once its effectiveness is confirmed in actual operations the model can be implemented, either on its own or concurrently with the SSARR model.

7. REFERENCES

Atmospheric Environment Service (1970) 10-Year Mean Annual Lake Evaporation, Environment Canada, Ottawa, ON

Atmospheric Environment Service (1981) Canadian Climate Normals 1951-1980, Environment Canada, Ottawa, ON

Atmospheric Environment Service (unpublished) Canadian Climate Normals 1961-1990, Environment Canada, Ottawa, ON

Baker, Calvin (1996), personal conversation, Water Resoucres Branch, Environment Canada, St. John's, NF

Bostock, H.S. (1970) Physiographic Subdivisions of Canada, Geology and Economic Minerals of Canada, Geologic Survey of Canada, Energy, Mines and Resources Canada, Ottawa, ON

Box, G.E.P. and Jenkins, G.M. (1976), Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, CA

Cumming Cockburn and Associates (1984) Hydrotechnical Study of the Steady Brook Area, Main Report, Canada-Newfoundland Flood Damage Reduction Program, St. John's, NF

Cumming Cockburn and Associates (1985) Preliminary SSARR Flood Forecasting on the Humber River System - Event of June 1984, Canada-Newfoundland Flood Damage Reduction Program, St. John's, NF

Cumming Cockburn and Associates (1995) Streamflow Forecasting for the Humber River Basin - Phase III Report Canada-Newfoundland Water Resources Management Agreement, St. John's, NF

Davis, Ed, (January 1991) SSARR Model User Manual (draft) US Army Corps of Engineers, North Pacific Division, Portland, OR

Golder Associates (1983), Hydrogeology of the Humber Valley Area, Water Resources Report 2-5 Groundwater Series, Department of Environment, St. John's, NF

Goodrich, Robert L. (1989) Applied Statistical Forecasting, Business Forecast Systems Inc. Belmont, MA

Forecast Pro for DOS (1993), Business Forecast Systems Inc., Belmont, MA

Inland Waters Directorate, (1979) Flood Forecasting in the Saint John River Basin - Mathematical Modelling Activities, Proceedings of the Technical Workshop on Streamflow Forecasting, Environment Canada, Ottawa, ON.

Pankratz, Alan, (1991) Forecasting with Dynamic Regression Models, John Wiley and Sons, New York, NY

Tang, P.W., and Lockhart, J.G. (1983), The Saint John River Forecast System, Proceedings of the Technology Transfer Workshop, Technical Workshop Series No.5, Inland Water Directorate, Environment Canada, Ottawa, ON.

Lockhart, J.G. (1987), The Saint John River Forecast System Update, Proceedings of the 3rd Flow Forecasting Workshop, Inland Water Directorate, Environment Canada, Ottawa, ON.

Mahoud, E., (1984) Accuracy in Forecasting: A Survey, Journal of Forecasting, Vol. 3

Manitoba Natural Resources and Environment Canada (1985) Manitoba River Forecast Development Phase 1 Planning and Design, Winnipeg Manitoba

Rowe, J.S. (1972) Forest Regions of Canada, Canadian Forestry Service Publication No. 1300, Ottawa, ON

Sanford, B.V., Grant, J.C., Ross, D.I., Van der Linden, W.J.W. (1976) Physiography of Eastern Canada and Adjacent Areas, Map 19399A(SE), Geologic Survey of Canada, Energy, Mines and Resources Canada, Ottawa, ON

SYSTAT Version 5.03 for DOS (1991), SYSTAT Inc., Evanston, IL