HYDROLOGIC ENSEMBLE PREDICTIONS USING ENSEMBLE METEOROLOGICAL FORECASTS

HYDROLOGIC ENSEMBLE PREDICTIONS USING ENSEMBLE METEOROLOGICAL FORECASTS

By
SADIK AHMED
Master in Engineering (University of the Ryukyus)

A Thesis
Submitted to the School of Graduate Studies
In Partial Fulfillment of the Requirements
For the Degree
Master of Applied Science

McMaster University
Department of Civil Engineering

MASTER OF APPLIED SCIENCE (2010) **McMaster University**

Civil Engineering Hamilton, Ontario

TITLE : Hydrologic Ensemble Predictions using Ensemble Meteorological Forecasts

AUTHOR : Sadik Ahmed

Master in Engineering (University of the Ryukyus)

SUPERVISOR : Dr. Paulin Coulibaly

Dr. Ioannis Tsanis

NUMBER OF PAGES: xii, 91

Abstract

The objective of this thesis is to investigate the potential of ensemble meteorological forecasts (15 members for each day) in improving ensemble flow prediction up to 14 days ahead. Large scale ensemble meteorological forecasts generated by the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS) are used. The hydrologic model used in watershed analysis of the study area is Hydrologiska Byråns Vattenbalan-avdelning (HBV). The study area is located in the Saguenay-Lac-Saint Jean watershed in northeastern Canada and comprises the Serpent River and Chute-du-Diable basins and a reservoir in Chute-du-Diable.

The NCEP ensemble meteorological forecast data is initially used as input in the hydrological model HBV to simulate ensemble reservoir inflows and the Serpent River flows for 5 to 14 days ahead. The ensemble inflow and flow forecasts are compared with the case where only observed historical data are used. The study results show that there is a significant improvement in the model forecast performance when NCEP forecast data are used. The improvement for 5 to 14 day forecasts is revealed by an approximately 20% decrease in root mean square error (RMSE) for both reservoir inflow and river flow. A decrease in the Brier score (BS) and rank probability score (RPS) indicates considerable improvement and an increase in the correlation coefficient (r) and the Nash and Sutcliffe coefficient (R^2) is shown for reservoir inflow and the Serpent River flow respectively, indicating the advantage of using NCEP data. This improvement is also revealed by the visual inspection of scatter plots, hydrographs of ensemble mean and ensemble members. The hydrologic forecasts are also assessed on a seasonal basis indicating an improvement in forecasting indicated by a 30% decrease in RMSE during the spring season, and a decrease in BS and RPS values. For other seasons, specifically autumn and summer, the use of the ensemble meteorological forecasts do not provide significant improvement because of the poor skill of predicted precipitation. More accurate predictions of reservoir inflow and river flow with adequate lead time will assist in improving relevant issues in water resources management and planning.

Acknowledgements

I would like to thank my supervisors Dr. Paulin Coulibaly and Dr. Ioannis Tsanis for their constant guidance, encouragement and enthusiasm during this research work.

I would also like to express my sincere appreciation to my laboratory colleagues Dr. Jos Samuel, Xiaoli Liu, Zakia Sultana and Manu Sharma for their significant contribution to many aspects of my research.

The love, support and encouragement of my parents and my wife are invaluable; thanks must go to them.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ANN	Artificial Neural Network
BNN	Bayesian Neural Network
BS	Brier score
BSS	Brier skill score
CDC	Climate Diagnostic Center
CDD	Chute-du-Diable
CDP	Chute-des-Passes
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	Ensemble Prediction System
GFS	Global Forecast System
HBV	Hydrologiska Byråns Vattenbalan-avdelning
MF	Meteorological Forecasts
MRF	Medium Range Forecast
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
RMSE	Root Mean Square Error
RPS	Rank probability score
RPSS	Rank probability skill score

Chapter 1: Introduction

1.1 Background

Hydrologic forecasts play a key role in water resources management and planning. Accurate river flow forecasts are essential for management of extreme events such as floods and droughts, optimal design of water storage and drainage networks, hydropower generation, optimizing the use of water, as well as ensuring adequate supply of water for irrigation and recreational uses and the maintenance of aquatic ecosystems. Forecast of inflows into reservoirs used to generate hydropower results in improved management of water resources, increases the benefit from power generation and reduces the risks associated with spillway operation. Hydrologic forecasts contribute considerable economic benefits by providing valuable support for decision making and by reducing flood damages, providing greater efficiency in power generation as well as diminishing environmental problems associated with hydraulic structures. The last few years have seen an increasing demand emerging from the user community (e.g. hydro-power producers, water authorities) to extend the forecast horizon in order to plan smooth production planning and scheduling reservoir operation and implement plans for downstream flood-prone areas effectively.

Considering the necessity of accurate stream flow forecasting, a great deal of research has been devoted during the past few decades to the modeling and forecasting of river flow dynamics. Such efforts have led to the formulation of different approaches associated with different levels of uncertainty and the development of a large number of models (Sivakumar et al., 2002; Singh and Woolhiser, 2002). This wide variety of hydrological rainfall-runoff models included complex physically-based distributed models like the MIKE SHE (Johnson et al., 2003; Jasper et al., 2002; Feyen et al., 2000; Refsgaard et al., 1997), HRCDHM (Carpenter et al., 2006), SHE (Abbot et al., 1986 a, b).

These models provide a detailed description of physics involved in the watershed but require a large amount of input data. On the other hand, lumped models like TOPMODEL (Beven and Kirkby, 1979) or the Xinanjiang model (Zhao et al., 1991; Gan et al., 1997) and semi-distributed models like the ARNO model (Todini, 1996) or HEC-HMS (Feldman et al., 1981) are modeling the (sub-) basin response, ignoring spatial variability. Similarly, HBV (Bergstrom, 1991; Bergstrom et al., 2002; Lindstrom et al., 1997), a semi-distributed conceptual model, is easier to apply than the physically based distributed hydrological rainfall-runoff models, as the required amount of data and CPU time are significantly lower.

Hydrologic model is the heart of hydrologic forecast and several recent studies have demonstrated that there is a significant forecast improvement when meteorological information, for instance quantitative precipitation and temperature forecasts of numerical weather prediction systems are used as an input in the hydrologic model for streamflow forecasting (Coulibaly, 2003; Werner et al., 2004; Bartholmes and Todini, 2005). Moreover, over the last few years the operational and research streamflow forecasting systems around the world are increasingly moving towards ensemble prediction systems rather than deterministic/point streamflow forecasts due to the increasing demand from the user community (Boucher et al., 2009; Cloke and Pappenberger, 2009). A deterministic forecast provides a single value per time step and aims to implement a model that produces a point forecast that is as close as possible to the observed outcome, and on the other hand ensemble forecasting is different from the deterministic point of view by avoiding the assumption of existence of a perfect model (Boucher et al., 2009). Ensemble forecasting systems produce n members of forecasts for each lead time instead of producing a single value for each time step, and focus on issuing a type of forecast that accounts explicitly on the uncertainty inherent in the forecasting system as a whole. The information provided by the ensemble forecasting system informs the user about the uncertainty and allows the decision maker to determine the probability of exceeding certain thresholds (Velazquez, 2009) as well as the assessment of the confidence intervals to be associated with the forecast. In hydrologic modeling, uncertainty mainly originates from a) model error – model parameter error (due to the use of non-optimal parameters) and model structure error (relationships among the variables characterizing the behavior of the systems), b) errors in meteorological input (observed and forecasted meteorological variables) and hydrological input (flow data) (Coulibaly, 2003), c) others – data do not represent the required spatial and temporal average, measurement errors, human reliability , truncation errors and rounding errors etc (Beck, 1987).

In the context of hydrological forecasting, ensemble (n members) streamflow can be obtained by a) providing a model n sets of equally likely initial conditions or parameters b) running n different models in parallel and c) incorporating ensemble (n members) meteorological forecasts into a hydrologic model (Boucher et al., 2009). In this study, a semi-distributed conceptual hydrologic model HBV is chosen to simulate ensemble (15 members) river flow and reservoir inflow. Here the ensemble (15 members) meteorological forecast data, generated by the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS), collected from the re-forecast project of the Climate Diagnostic Center (CDC) are incorporated into the HBV model for ensemble hydrologic simulation.

1.2 Research Objectives

Ensemble stream flow predictions obtained by forcing hydrologic models with ensemble numerical meteorological forecasts are becoming more commonly used in operational hydrologic forecasting application. In recent years, several hydrological modeling studies for flow forecasting have been carried out on the Saguenay-Lac-Saint-Jean Watershed located in Northern Quebec, Canada (Coulibaly, 2003; Coulibaly et al., 2001; Coulibaly et al., 2000; Khan and Coulibaly, 2006; Liu, 2007). Liu 2007 reported the potential of NCEP meteorological forecasts for improved hydrologic simulation after forecast range 5 in the study region; however in that study only deterministic flows are simulated using downscaled meteorological forecasts and investigated in a deterministic manner. The specific objective of this study is to investigate the potential of ensemble meteorological forecasts (15 members for each day) for improved ensemble flow prediction for 5 to 14 days ahead. To achieve the objective, the following main activities are carried out in this research: simulation of reservoir inflow and Serpent River flow using the observed historical meteorological data, simulation of ensemble reservoir inflow and Serpent River flow using NCEP GFS ensemble meteorological forecasts, and the assessment of flow forecasts using conventional deterministic and probabilistic measures as well as visual inspection of hydrographs and scatter plots.

1.3 Structure of the Thesis

Chapter 1 of this thesis describes the background, research objective and scope of the thesis. This followed by Chapter 2 provides a description of the study area, collected observed meteorological data at CDD and CDP, observed reservoir inflow and the Serpent River flow data, and NCEP GFS 15-member ensemble meteorological forecasts. Chapter 3 provides an overview of hydrologic models, a review of ensemble hydrologic forecast studies, a brief description of the HBV hydrologic model, model calibration and validation and finally an outline of the hydrologic forecast experiment. In Chapter 4, the study results are presented and discussed and Chapter 5 provides the conclusions of this study and recommendations for further research work.

Chapter 2: Study Area and Data

2.1 Study Area

The two sub-basins selected for the ensemble hydrologic predictions, and the assessments of the hydrologic predictions, are the Chute-du-Diable and the Serpent River basin. Both basins are in the Saguenay-Lac-Saint Jean watershed located in northeastern Canada (shown in Figure 1). The sub-basin Chute-du-Diable, which has an area of approximately 9,700 km^2, located in the eastern part of the Saguenay-Lac-Saint Jean watershed is selected for reservoir inflow forecasts. This sub-basin contains a large hydropower reservoir managed by the Aluminum Company of Canada (ALCAN) for hydroelectric power production. The whole Chute-du-Diable sub-basin contributes to the reservoir inflow (Coulibaly et al., 2001). The Serpent River sub-basin, which has an area of approximately 1,700 km^2 is located in the middle part of the watershed and was selected for river flow forecasting. The main reason for choosing these particular basins is the availability of reliable historical hydro-meteorological record for a long period. The average annual precipitation in the study basins is about 950 mm and the mean daily temperature ranges between +30°C and -40°C.

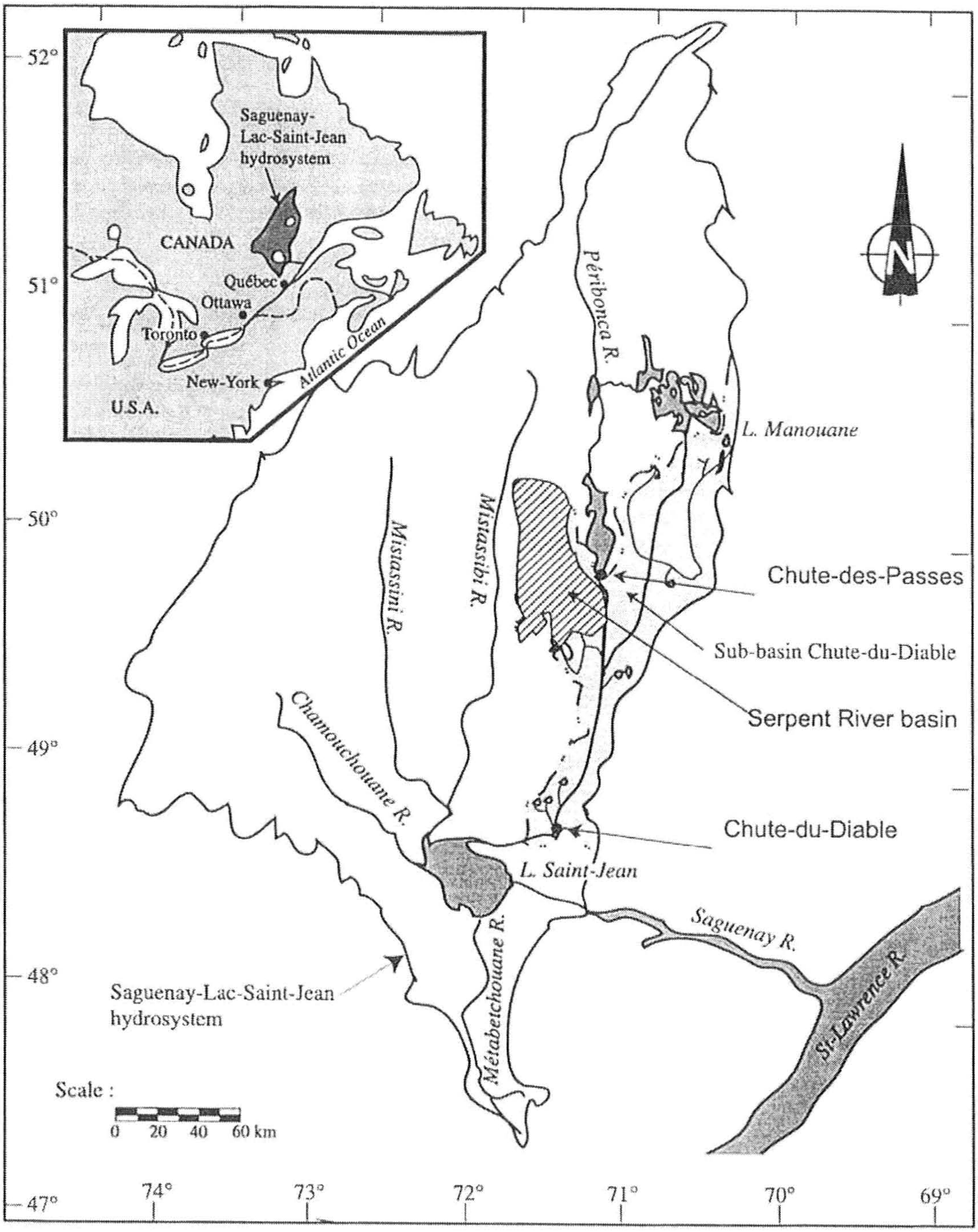

Figure 1: Map showing the study area within the Saguenay –Lac-Saint-Jean watershed (Coulibaly et al., 2001)

2.2 Data Collection

2.2.1 Observed Hydro-meteorological Data

The observed meteorological data used for this study are total daily precipitation (total precipitation in form of liquid and snow, measured in mm), and daily maximum and minimum temperatures (in °C). The observed precipitation and temperature data for twenty-three years, from 1979 to 2001, were obtained from two meteorological stations; namely the Chute-des-Passes (CDP) meteorological station with latitude and longitude of 49.9°N and 71.25°W and Chute-du-Diable (CDD) meteorological station with latitude and longitude of 48.75°N and 71.7°W respectively. Both Chute-des-Passes meteorological station (station ID 7061541) and Chute-du-Diable meteorological station (station ID 7061560) are shown in the study area map in Figure 1. The precipitation and temperature data obtained from Chute-des-Passes meteorological station are used for the Serpent River flow prediction and the precipitation and temperature data obtained from both Chute-des-Passes meteorological station and Chute-du-Diable meteorological station are used for Reservoir inflow prediction. For the Serpent River flow prediction, the observed daily flow data for eleven years, from 1991 to 2001, were obtained from a hydrometric station (station ID 062214) located at latitude and longitude of 49.41°N and 71.22°W respectively. Twenty-three year (from 1979-2001) reservoir inflow in the Chute-du-Diable catchment are used for reservoir inflow prediction. The hydrometeorological networks in the study sub-basins are maintained by the Aluminum Company of Canada (ALCAN) for a long time and all the observed hydrometeorological data used in this study were obtained from ALCAN company hydrometeorological networks.

2.2.2 NCEP Forecast Data

The meteorological forecast variables used in this research are accumulated precipitation (apcp, in mm) and temperature at 2 meter above ground level (t2m, in °C) collected from the re-forecast project ftp server of the Climate Diagnostic Center (CDC) at the Cooperative Institute for Research in Environmental Sciences (CIRES). The CIRES is a joint Institute of the National Oceanic and Atmospheric Administration (NOAA) and the University of Colorado at Boulder. The meteorological forecast variables were generated by an unchanged version of the National Centers for Environmental Prediction's Global Forecast System (NCEP GFS) model. This Global Forecast System (GFS) model was formally known as the Medium Range Forecast (MRF) model. 15-member ensemble forecasts (apcp and t2m) produced up to 15-days lead time for every day from 1979 to 2001 are used as predictors in the hydrologic model. The approach currently used at NCEP to generate ensemble members is the "breeding method" (Toth and Kalnay, 1997). The forecast model was run at T62 horizontal resolution and 28 sigmal levels to generate 15-member ensemble forecast (Hagedorn et al., 2007; Hamill and Whitaker 2007; Hamill and Whitaker 2006; Watson et al., 2002). These ensemble forecasts are initialized at 0000 UTC each day. The control run was initialized from a reanalysis data, while initial conditions for the other 14 ensemble members were produced from 7 pairs of bred modes (Whitaker et al., 2006; Hamill et al., 2004; Toth and Kalnay, 1997). The model forecast variables are saved at an interval of 12 hours, so the ensemble forecasts for near surface (2m AGL) temperature and accumulated precipitation are available from the reforecast dataset for 0000 UTC and 1200 UTC (Wilks and Hamill, 2007). The re-forecast data are available on global 144x 73 equally spaced latitude and longitude grid points with a horizontal resolution of 2.5° in both in the latitude and longitude. In this study, the forecast variables (apcp and t2m) obtained from grid point with latitude and longitude of 50°N and 72.5°W, and grid point with latitude and longitude of 47.5°N and 72.5°W are used for ensemble hydrological prediction. The forecast variables (apcp and t2m) obtained from grid point with latitude and longitude of 50°N and 72.5°W are used for the Serpent River flow prediction and forecast variables (apcp, t2m) from both grid points are used for reservoir inflow prediction. These grid points are chosen with the intent of providing representative coverage of the study area as the grid point with latitude and longitude of 50°N and 72.5°W is the closest grid point to the Chute-des-Passes meteorological station and grid point with latitude and longitude of 47.5°N and 72.5°W is the closest grid pint to the Chute-du-Diable meteorological station. In Figure 2, NCEP ensemble forecast grid points are shown in blue (where the black circled are selected grid points) and the two meteorological stations are shown in red.

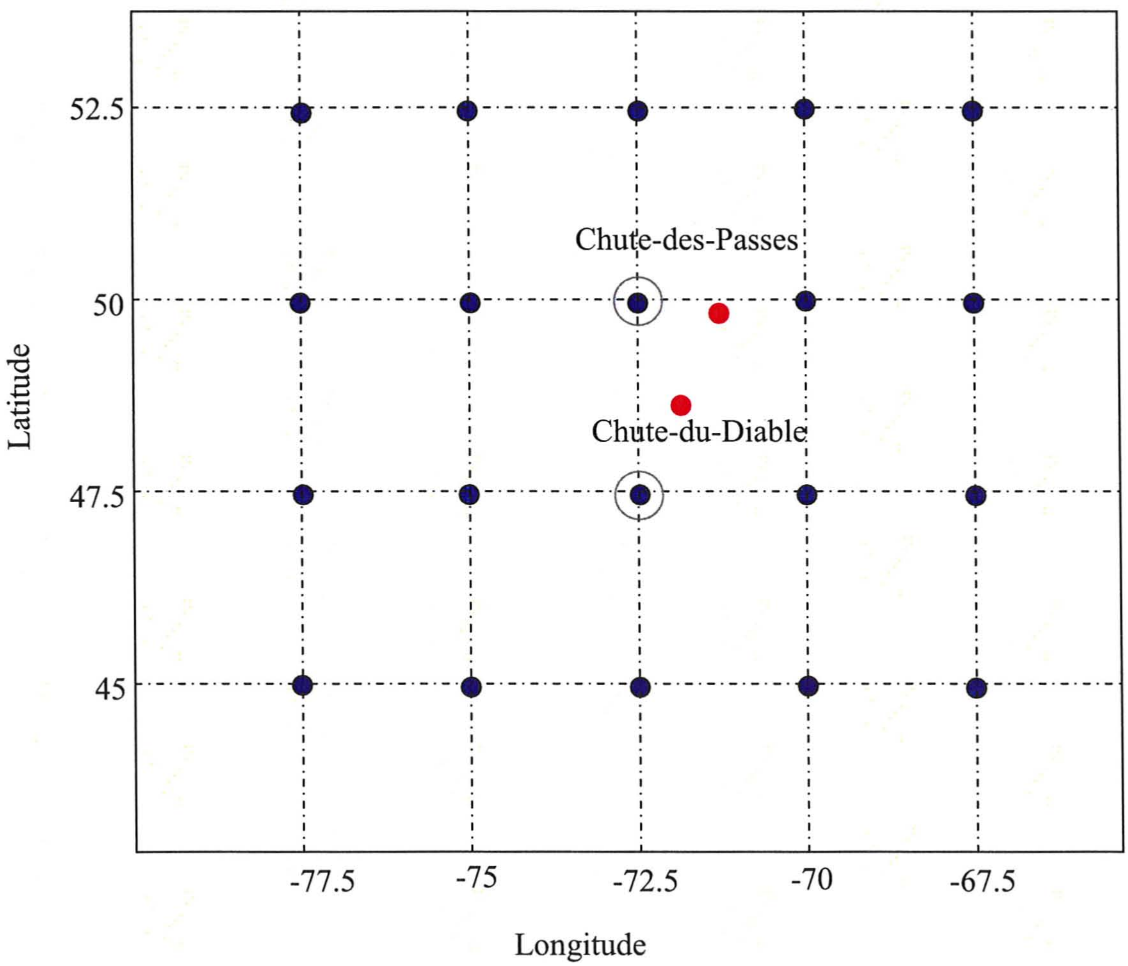

Figure 2: NCEP-GFS ensemble meteorological forecasts grid points (Liu, 2007)

The ensemble meteorological forecast variables collected from the reforecast project ftp server of the Climate Diagnostic Center (CDC) are in netCDF (network Common Data Format) file format. Figure 3 presents a 3-D description of the NCEP ensemble meteorological data. It illustrates that each sheet contains the data for each forecast range and contains 15-member time series data for each forecast range, and there are 15 forecast ranges for each variable (Liu, 2007).

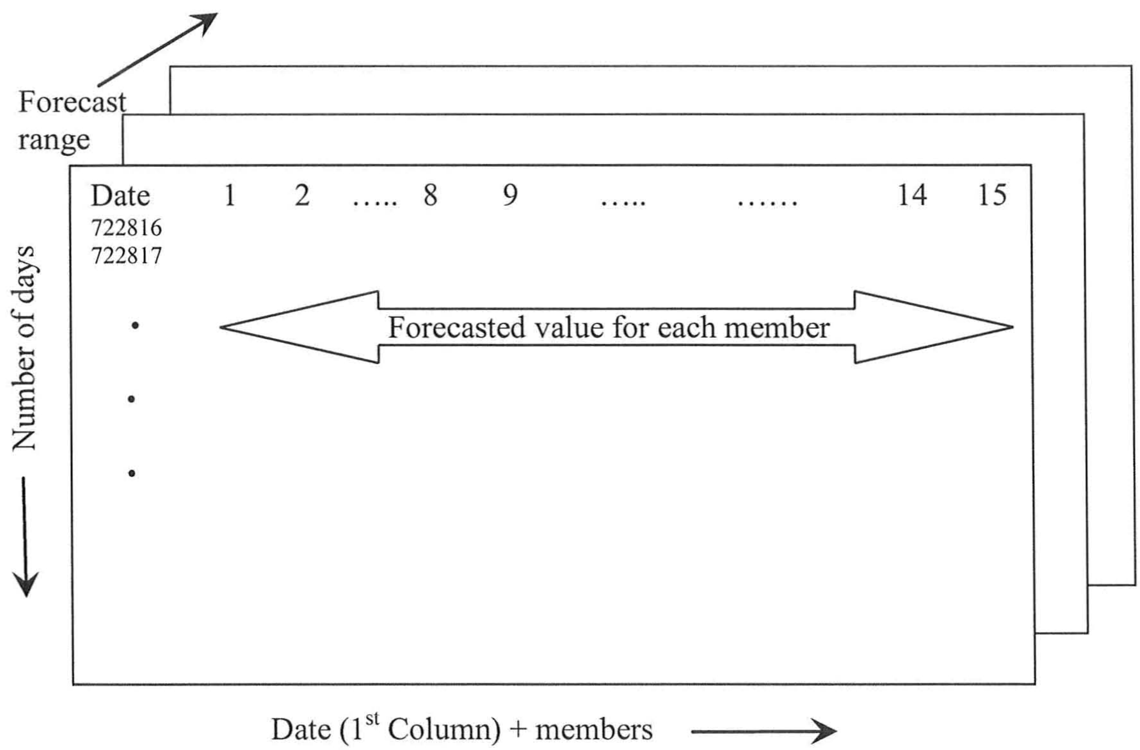

Figure 3: Description of 3-D ensemble meteorological data (courtesy Dr. Evora)

2.2.3 Evaluation of NCEP Ensemble Forecasts

In this section, Figures 4, 5, 6 and 7 represent observed daily mean temperature and NCEP temperature forecasts (all 15 member ensembles, ensemble mean, and upper and lower limit of 95% confidence interval). Figures 4 and 5 present observed mean temperature at CDD metrological station and NCEP temperature forecasts at grid point with latitude and longitude of 47.5°N and 72.5°W for the entire year and spring season of one year (2001). Figures 6 and 7 present observed mean temperature at CDP metrological station and NCEP temperature forecasts at grid point with latitude and longitude of 50°N and 72.5°W for the entire year and spring season of one year (2001). Figures 4 and 6 show that NCEP temperature forecasts are good over the entire year as the observed temperatures are close to the ensemble means and lie in the confidence intervals on most days. The high skill of NCEP temperature in the spring season can also be seen in Figures 5 and 7. These Figures also show that snow melting temperature starts at the beginning of April. The calculated correlation coefficient (r) shown in Tables 1 and 2 also revealed the high skill of NCEP forecasted temperature for the study basins. The correlation coefficients and root mean square error were calculated between two variables – observed temperature or precipitation (at CDD or CDP meteorological station) and NCEP forecasted temperature or precipitation (at 47.5°N, 72.5°W and 50°N, 72.5°W). It can be seen from Table 1 that for temperature, correlation coefficient is 0.92 for 5 day ahead forecasts and 0.83 for 14 day ahead forecasts at CDD, and it lies between 0.92 and 0.83 for other forecast ranges. It can also be seen from Table 2 that for temperature, correlation coefficient is 0.93 for 5 day ahead forecasts and 0.84 for 14 day ahead forecasts at CDP, and it lie between 0.93 and 0.84 for other forecast ranges. Table 1 and 2 also show that in case of temperature forecasts both the RMSE and r decrease with the increase of forecast lead time. This fundamental pattern was not found in case of precipitation forecasts as shown by RMSE and r in Tables 1 and 2. The poor skill of NCEP precipitation forecasts was revealed by very low correlation coefficient (for example, less than 0.2 for 5 days ahead) as shown in Tables 1 and 2. Figures 8, 9 and others (Figures A1 through A6 in Appendix) represent observed precipitation and NCEP precipitation forecasts (mean of 15 member ensembles, and upper limit of 95% confidence interval). The poor skill of NCEP precipitation forecasts are shown in these Figures as the observed precipitation has very poor agreement with NCEP ensemble mean and maximum values.

Clark and Hay (2004) reported that in the snowmelt dominated river basins when surface hydrology is strongly forced by temperature; the high skill in predictions of temperature translates into high skill in prediction of streamflows. They also reported that in a basin where streamflows are controlled by variation in temperature, difficulties in providing accurate precipitation forecasts are less important. Moreover, the study reported that overall skill of techniques of model output statistics (MOS) based precipitation forecasts is slightly lower than the raw NCEP precipitation forecasts. Moreover, an extensive investigation is required to find out an appropriate downscaling method to downscale precipitation forecasts. Khan and Coulibaly (2006) reported that the snowmelt runoff is responsible for high flows in the spring season as well as up to 40% of the annual flow volume in the study region. Considering the above mentioned facts, this study aimed to improve hydrologic prediction up to 14 day ahead using the raw NCEP meteorological forecasts.

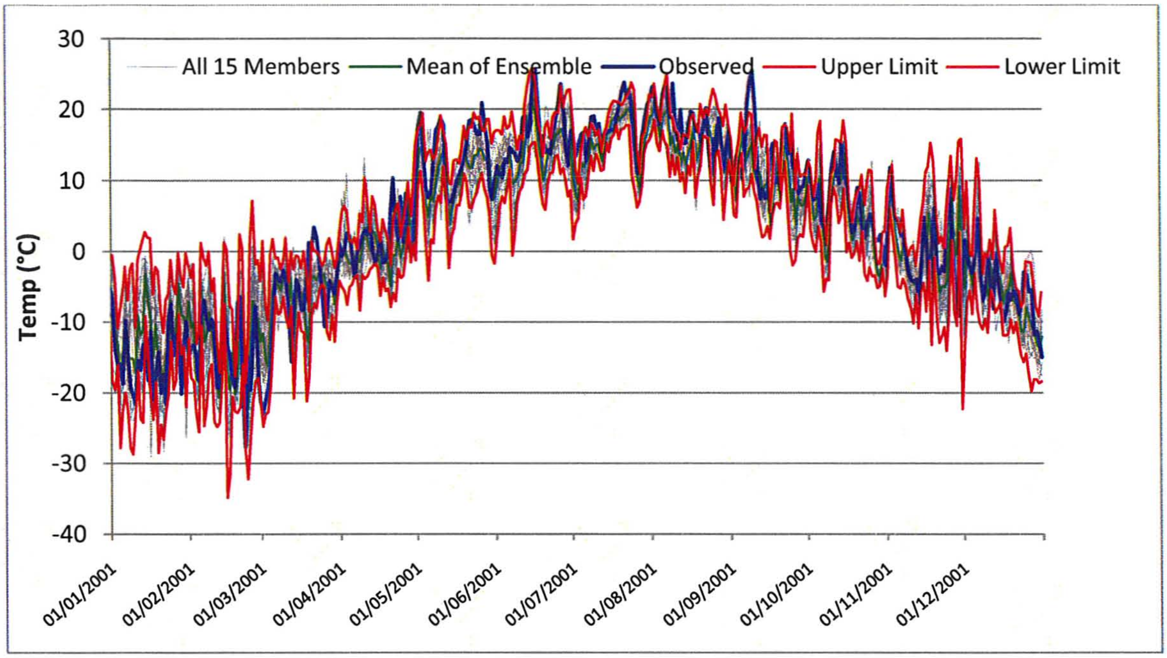

Figure 4: Plots of NCEP temperature forecasts (5 day ahead) and observed mean temperature at CDD for year 2001(limits show 95% confidence interval)

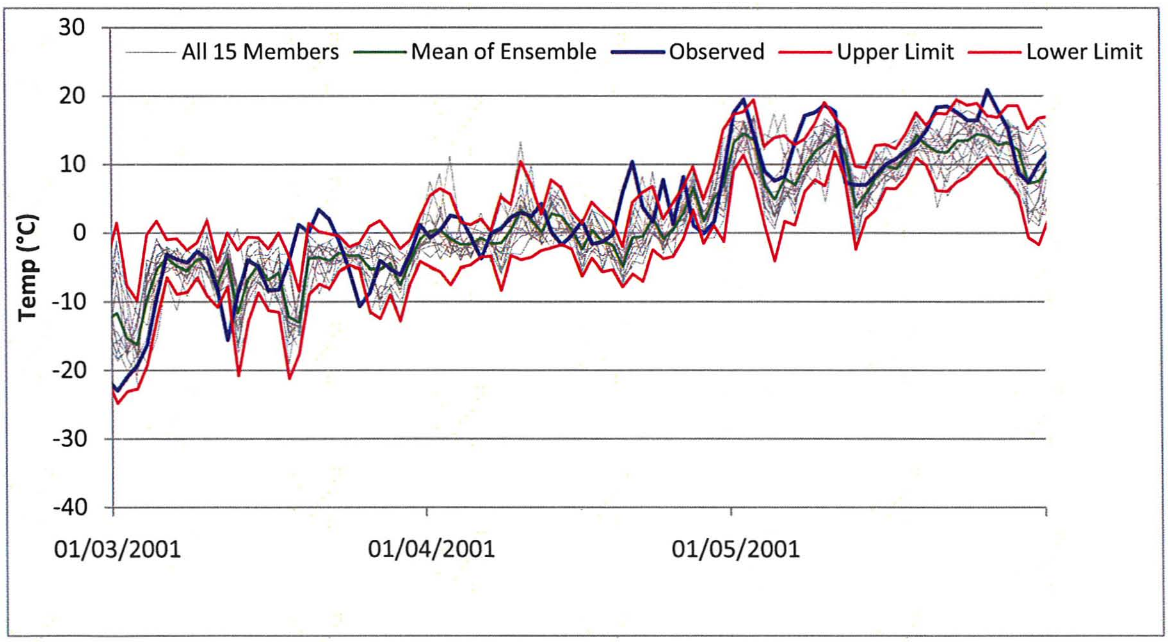

Figure 5: Plots of NCEP temperature forecasts (5 day ahead) and observed mean temperature at CDD for spring 2001(limits show 95% confidence interval)

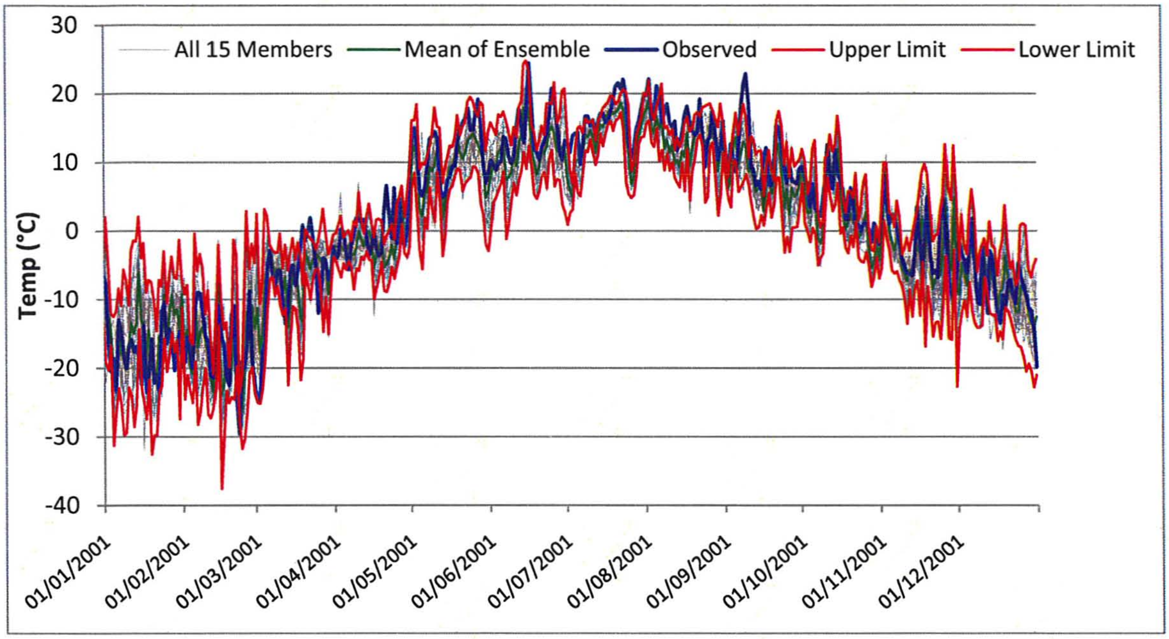

Figure 6: Plots of NCEP temperature forecasts (5 day ahead) and observed mean temperature at CDP for year 2001(limits show 95% confidence interval)

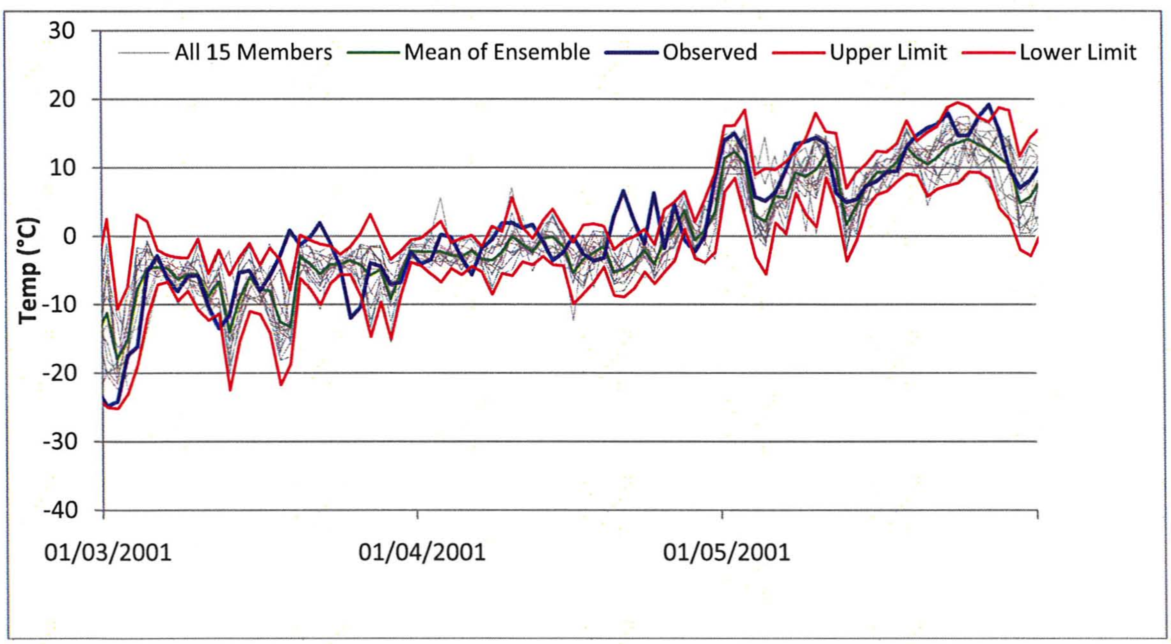

Figure 7: Plots of NCEP temperature forecasts (5 day ahead) and observed mean temperature at CDP for spring 2001(limits show 95% confidence interval)

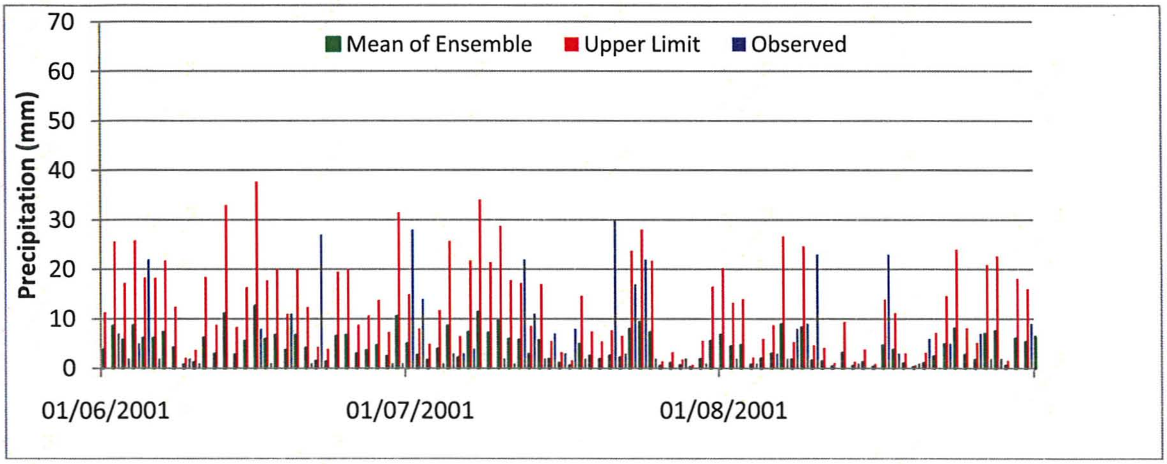

Figure 8: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDD for the summer season (limit shows 95% confidence interval)

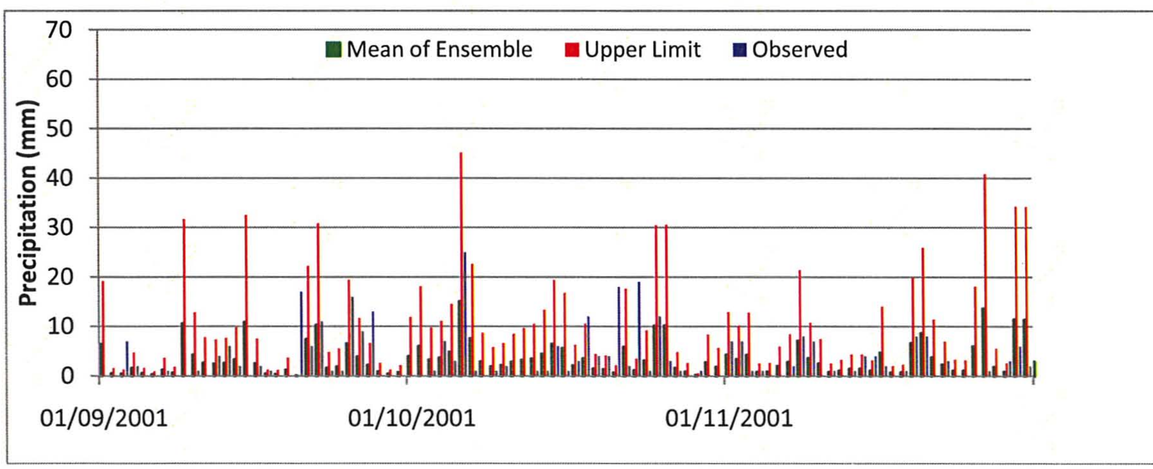

Figure 9: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDD for the autumn season (limit shows 95% confidence interval)

Table 1: Meteorological model performance statistics (RMSE, r) for member 1 at CDD (1997-2001)

Forecasting lags	**Variable - Tmean**		**Variable - Precipitation**	
	RMSE(°C)	r	RMSE(mm)	r
5-day-ahead	5.03	0.92	7.66	0.13
6-day-ahead	5.60	0.90	7.78	0.09
7-day-ahead	5.88	0.89	7.84	0.02
8-day-ahead	6.17	0.87	7.92	0.02
9-day-ahead	6.50	0.86	8.00	0.02
10-day-ahead	6.67	0.85	7.96	0.02
11-day-ahead	6.89	0.84	7.90	0.01
12-day-ahead	7.04	0.84	7.94	0.05
13-day-ahead	7.16	0.83	8.14	0.04
14-day-ahead	7.16	0.83	8.15	-0.02

Table 2: Meteorological model performance statistics (RMSE, r) for member 1 at CDP (1997-2001)

Forecasting lags	**Variable - Tmean**		**Variable - Precipitation**	
	RMSE(°C)	r	RMSE(mm)	r
5-day-ahead	4.78	0.93	7.27	0.16
6-day-ahead	5.35	0.91	7.24	0.17
7-day-ahead	5.59	0.90	7.52	0.08
8-day-ahead	5.99	0.88	7.54	0.03
9-day-ahead	6.31	0.87	7.86	0.04
10-day-ahead	6.47	0.86	7.66	0.06
11-day-ahead	6.50	0.86	7.67	0.04
12-day-ahead	6.68	0.85	7.48	0.08
13-day-ahead	6.92	0.84	7.99	0.03
14-day-ahead	6.89	0.84	7.75	0.02

Chapter 3: Hydrologic Forecasts

3.1 Overview of Hydrologic Models

A mathematical model describes a system of assumption, equations and procedures intended to describe the performance of a prototype system. Hydrological models are mathematical formulations of a particular phase in the hydrological cycle which attempts to describe the actual physical process of the hydrologic cycle so as to simulate actual hydrologic events such as the transformation of a series of rainfall inputs to the resulting streamflow. Now-a-days, hydrologic models have become crucial to water resources assessment, development and management. They are employed in a wide spectrum of areas ranging from watershed management to engineering design. They are used in varied purposes, such as analyzing the quantity and quality of streamflow, reservoir system operations, groundwater development and protection, surface water and groundwater conjunctive use management, water distribution systems, water use and water resources management activities (Wurbs 1998). Over the past few decades, a large number of models have been developed for different purposes and with different philosophical concepts. In the following section the fundamentals of the physically-based distributed model, lumped conceptual model and black box model are discussed.

Physically-based distributed models are very useful to our understanding of physical process involved in the hydrological process such as river flow. The physically based models are specially designed based on the underlying physical mechanism and the equations involved in the physics of the hydrological processes. The lumped models treat a whole catchment, or a significant portion of it, as if it was homogeneous in character and subject to uniform rainfall. Physically based distributed models provide a detailed representation of the physical characteristics (topography, soil conditions, land cover, etc.) of the watershed. These models feature the capability to incorporate the entire physiographic and hydrological variability of a study area by accounting for heterogeneities throughout the watershed (Sharma 2009). These models divide the catchment area into a large number of small grid systems, simulate each separately, and combine them to obtain catchment response. So, the distributed model deals with heterogeneities over a catchment more logically than lumped models. Fully distributed physically based models are very data intensive as they require large amounts of high quality physical, geographical and meteorological data to ensure accurate results (Cranmer et. al, 2001). Moreover the computations in the distributed model are usually too time-consuming for engineering applications. The enormous data requirements and computational time prevent the extensive use of the fully distributed model. On the other hand the conventional lumped models are less data intensive, less complex, relatively easy to use and the required input data are available for most of the applications.

The most recent advances in fully distributed hydrologic modeling were noted to have been the employment of Geographic Information Systems (GIS), and remotely sensed data to accurately account for the spatial variability in largely heterogeneous watersheds (Singh and Woolhiser, 2002). In the fully distributed models, mass, momentum and energy are calculated directly from the governing partial differential equations which are solved using numerical methods (Sharma, 2009). On the other hand, conceptual models are based on simplified representations of the hydrologic processes in a watershed and are normally run with point values of precipitation and temperature as the primary input (Liden & Harlen, 2000). The idea with conceptual modelling is to consider the catchment as a system whose components are precipitation, evapotranspiration, storage and runoff. The water balance equation for a catchment model can be written as

$$P - ET \pm \Delta S = Q \qquad (1)$$

where P is total precipitation on the catchment, ET is the evapotranspiration, ΔS is the change in water storage, and Q the runoff from the catchment (Liden & Harlen, 2000). The relationship between these variables is simple in conceptual models (Liden &

Harlen, 2000). The algorithms are usually simplified by the use of empirical relations in order to speed the solution and in order to cope with the point to point variations in the hydrologic processes within the catchment. Hydrologic models can be used for continuous and event-based simulations. Continuous models are capable of generating outflow hydrographs over a long period of time, and use finite time periods in the computation and the computed flows are discrete points in time. Event-based models are designed to simulate a single event such as the hydrograph of a single storm. Most event models were developed specially to design urban drainage systems and other small projects. Continuous models are useful for simulation of long flow records for use in design, evaluating the impact of changes in a catchment on streamflow and forecasting of streamflow.

The black-box models are designed to identify the connection between the input and the output without going into the analysis of the physical mechanisms involved in the hydrological process and are also capable of representing the complex non-linear river flow process, by relating the inputs and outputs of the underlying system (Sivakumar, et al., 2002).

There are a large number of hydrological models in current use in different countries and the models are used for different purposes. Singh and Woolhiser, 2002 reported that complex physically-based distributed models like the System Hydrologique Europeen SHE (Abbot et al., 1986 a, b) and conceptual model TOPMODEL (Beven and Kirkby, 1979) are standard for hydrologic analysis in many European countries; a physically-based semi-distributed event based runoff model Hydrologic Engineering Center-Hydrologic Modeling System (HEC-HMS) (Feldman, 1981) is widely used for the design of drainage systems, quantifying the effect of land-use change on flooding, etc. and a lumped continuous model National Weather Service-River Forecast System (NWS-RFS) (Houge et al., 2000) is the standard model for flood forecasting in the United States; a semi-distributed continuous flow simulation model Waterloo Flood System (WATFLOOD) (Kouwen, 2010; Kouwen et al., 1993) and lumped parameter continuous simulation University of British Columbia (UBC) model (Quick and Pipes, 1977) are popular in Canada for hydrologic simulation and a semi-distributed continuous streamflow simulation model HBV (Seibert, 1997; Bergstrom, 1991, Bergstrom et al., 2002, Lindstrom et al., 1997) is widely used for flood forecasting in Scandinavian countries. Recently black-box model artificial neural network (ANNs) have been used for modeling many of the nonlinear hydrologic processes such as rainfall-runoff, streamflow, ground water management, water quality simulation and precipitation (Gavindaraju, 2000). Gavindaraju, 2000 reported that Half et al., 1993 designed a three-layer feedforward ANN for rainfall-runoff modeling using rainfall hyetographs as inputs and hydrographs recorded by the U.S. Geological Survey (USGS) at Bellvue, Washington, as output; Karunanithi et al., 1994 studied ANN performance for river flow prediction at an ungauged location on the Huron River in Michigan using the data from USGS stream gauging stations located 30km upstream and 20km downstream of the prediction point; Kuligoski and Barros, 1998 used ANN approach for precipitation forecast based on upper atmospheric wind direction and antecedent precipitation data from a raingauge network.

'Although hydrological models have been around for quite some time, there is yet to be one exclusive model that can stand apart from the rest and be declared best at modeling all aspects of the hydrologic system' (Sharma, 2009). In this study, a semi-distributed conceptual model HBV is chosen to simulate river flow and reservoir inflow. The motivations for choosing this particular model are:

(a) snow routing component; Khan and Coulibaly, 2006 reported that snowmelt runoff is responsible for high flows in the spring season as well as up to 40% of the annual flow volume in the study region;

(b) the model was used in previous studies in the same study area and showed a good performance;

(c) Liu (2007) revealed the potential of the conceptual semi-distributed HBV model for improved flow forecasting in the study area when meteorological forecasts are included in the models; and

(d) computational time that is required for ensemble flow simulation is less than many distributed and black-box models.

3.2 Ensemble Hydrologic Forecast: Review

In the recent years, the operational and research in streamflow forecasting systems around the world have significantly moved towards ensemble prediction systems instead of a single deterministic forecast because of the increasing demand from the user community (Boucher et al., 2009; Cloke and Pappenberger, 2009). In response to the increasing demand from the users community for a better anticipation of hydrological events, the use of ensemble forecasts in hydrology is emerging as a key research area within the scientific community (Regimbeau et al., 2007).The development of ensemble hydrologic prediction systems has started in the late 1990s and the research is ongoing (Dietrich et al., 2008). There are different ways to produce ensemble hydrologic forecasts (Boucher et al., 2009), one approach is to use meteorological forecasts, for instance quantitative precipitation and temperature forecasts from numerical weather prediction systems as the input for the hydrological rainfall-runoff models to gain an ensemble of streamflow forecasts (Renner et al., 2009). In this study, the ensemble (15-member) meteorological forecast (precipitation and temperature) data, generated by the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS), collected from the reforecast project of Climate Diagnostic Center (CDC) are used as the input for a semi-distributed conceptual Hydrologic model HBV to produce ensemble reservoir inflows and the Serpent River flow forecasts. Therefore in the following sections, some findings from previous studies on ensemble hydrologic forecasts using numerical meteorological forecasts are reported.

Clark and Hey (2004) investigated the value of the National Center for Environmental Prediction's (NCEP) medium range forecast (MRF) modeling system output for the prediction of streamflow. The U.S. Geological Survey's Precipitation-Runoff Modelling System (PRMS) was used for hydrologic simulation. Streamflow forecasts using MRF output were generated for four basins in the United States – three snowmelt dominated basins (Animas River, Colorado; East Fork of the Carson River, Nevada and Cle Elum River, Washington) and one rainfall dominated basin (Alapaha River, Georgia). Model output statistics (MOS) technique was used to downscale the NCEP forecasts to improve the forecasts of precipitation and temperature to station location. However, the authors pointed out the fact that the overall accuracy of MOS-based precipitation forecasts is slightly lower than the raw NCEP forecasts. Streamflow forecasts produced using the downscaled MRF output as inputs were compared with those using station observations. Using RPSS, the study results revealed that increase in skill from MOS-based forecasts are most pronounced during the peak snowmelt season in the three snowmelt-dominated river basins, when streamflows are strongly forced by temperature, and the high skill in predictions of temperature translates into high skill in predictions of streamflow. In contrast, there was no improvement in predictions of streamflow in the rainfall-dominated basin. They decided that further improvement of streamflow would require more accurate local scale meteorological forecasts, more accurate specification of basin initial conditions and model simulation of streamflow.

Roulin and Vannitsem (2005) investigated the performance of an ensemble streamflow prediction system that uses 50 member precipitation forecasts from the Ensemble Prediction System (EPS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) and the study mainly evaluated the ability of the system for high flow forecasting for medium-range (9 days) lead time. Two main river basins (the Demer catchment in the River Scheldt basin and Ourthe catchment in the River Meuse basin) in Belgium were chosen as case studies. The Integrated Runoff Model F. Bultot (IRMB) was used for hydrologic simulation. The ensemble hydrologic forecasts were compared with climatology and a reference forecast based on historical records. Using the Brier score and the root mean square error, the study has revealed that the skill of ensemble hydrologic prediction is much better than the one based on historical precipitation inputs. For example, in case of Demer catchment, BSS showed about 10% improvement in summer and about 50% improvement in winter for high flow (80 percentile value) for 9 day ahead forecast when climatology was used as reference; and when historical record based streamflow prediction (HSP) was used as reference, there was about 5% improvement in summer and about 35% improvement in winter for high flow shown by BSS. The study results also revealed that in both basins the skill of the streamflow and precipitation forecasts in the winter is greater than in summer; and the skill of streamflow remains positive in the winter for the whole sreamflow forecast lead time. This study also investigated the impact of increased EPS model resolution for streamflow forecasts and revealed that skill is due to forecast resolution. They also suggested that further improvement can be achieved through processing EPS forecasts to remove biases and recovering space-time variability of precipitation.

Velàzquez et al. (2009) evaluated short range 20 member ensemble hydrologic predictions for short range (1-3 days) relying on the 20 member ensemble meteorological forecasts from Environment Canada's Meteorological ensemble prediction system. A spatially distributed hydrological model Hydrotel was used for hydrologic simulation. The study basins are located in the province of Quebec, Canada. The ensemble hydrologic forecasts were compared with deterministic forecasts that use output of the Global Environmental Multiscale Model (GEM). Results, based on a single rain storm, revealed that for all twelve watersheds and for prediction horizons, the ensemble hydrologic forecasts are better than its deterministic counterpart, moreover superiority is more remarkable for longer prediction horizon. They recommended that the results might be generalized through an extended simulation period, a selection of hydrologic models and use of high resolution meteorological input when available.

Renner et al. (2009) studied the performance of ensemble hydrologic forecasts for various river stations on the Rhine and its tributaries (for catchment area ranging from 4,000km^2 to 160,000km^2) up to nine days forecast horizon. A semi-distributed conceptual hydrologic model HBV was used for ensemble hydrologic simulation. Two meteorological ensemble forecasts were used as inputs in the hydrologic model to produce ensemble hydrologic forecasts: 1) forecasts from the Ensemble Prediction System of the European Centre for Medium-Range Weather Forecasts (ECMWF -EPS) (ensemble size 50, grid size 80 km, since February 2006: 50km) and 2) forecasts from the Consortium for Small-Scale Modeling based on a local area model (COSMO-LEPS) (ensemble size 16, grid size 10 km). The flow forecasts derived using the ECMWF –EPS forecasts were compared with the climatology using RPS, and the results show that ECMWF –EPS based forecasts has a positive skill over climatology for lead times up to nine days. The study also revealed that the skill increased with increasing size of the catchment area and deteriorates faster in smaller basins with increasing lead time. When RPSS were calculated for COSMO-LEPS flow forecasts, rather than using climatology the ECMWF-EPS flow forecasts were used as reference. The comparison of the flow forecasts using ECMWF-EPS and COSMO-LEPS demonstrated the clear improved performance of the higher resolution COSMO-LEPS based forecasts. The authors concluded that there is a need for downscaling ensemble weather prediction products to a more representative scale for sub-basins in the hydrologic model.

Roulin (2007) analysed a hydrologic ensemble prediction system based on the 50 member forecasts from the Ensemble Prediction System (EPS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) for two Belgian catchments with contrasted hydrological cycle. The Integrated Runoff Model F. Bultot (IRMB) was used for hydrologic simulation. The ensemble forecasts were compared with deterministic forecasts based on either the archives of the ECMWF operational deterministic runs or the mean of EPS ensembles or the control run of EPS. The skill of the ensemble forecasts was assessed with Brier Skill Score and the value of the system was assessed with a costloss decision model. The BSS values show positive skill (20% improvement) of ensemble hydrologic prediction up to 9 day ahead, whereas skill of deterministic forecasts using ECMWF operational diminishes on 7 day ahead. The study demonstrated the positive skill of the ensemble hydrologic forecasts as the author concluded: 'The hydrological ensemble predictions have greater skills than deterministic one' (Roulin, 2007).

In addition to the above-mentioned studies, some studies carried out in the study region for hydrologic predictions using meteorological forecast variables are mentioned in the following sections. It is notable that these studies only investigated the single deterministic hydrologic forecasts obtained using meteorological predictions as input in the hydrologic model, however the information and knowledge gained from these studies may support the findings of this study.

Liu (2007) investigated the use of meteorological forecasts from the National Center for Environmental Prediction (NCEP) medium range forecast (MRF) modeling system for forecasting streamflow. This study used the ensemble meteorological forecast variables for downscaling precipitation and temperature at two meteorological stations in the Saguenay watershed in the northeastern Canada. The hydrological models HBV and Bayesian neural networks BNN were used for hydrologic forecasts up to 14 day ahead. Downscaled meteorological forecasts were used to produce deterministic forecasts, and these deterministic forecasts were compared with other deterministic forecasts where no downscaled data are used. The study results revealed that when TLFN downscaled data were included in HBV model, the RMSE decreases by about 10-20% for 6 to 14 day ahead reservoir flow forecasting; and the models have good performance (R^2> 0.7) up to 6 day for entire year. The study results also revealed that the HBV model performs better when downscaled meteorological data are included, but Bayesian Neural Network (BNN) does not show significant improvement, and the performance of the models are more pronounced in the spring season.

Coulibaly (2003) assessed the usefulness of meteorological predictions on real time spring flow forecasting. The short range (up to 7 days) numerical weather predictions provided by the Canadian Regional Forecast System based on a regional finite-element (RFE) model were used as input in the hydrologic model for deterministic reservoir inflow forecasts in the spring season. A conceptual hydrologic model PREVIS and a data driven model artificial neural network (ANN) were used for hydrologic simulation. The study results revealed that using both models, there is a relative improvement of the model efficiency in flow forecasting for all forecast horizons when meteorological forecasts are used properly. The study also reported that the model's efficiency is good (Nash and Sutcliffe coefficient, $R^2 > 0.8$) up to 3 days ahead for reservoir inflow forecasting. The study also revealed that even meteorological predictions with large errors can produce improved flow forecasts. The author recommended the use of multiple models and appropriate approach for using meteorological predictions for improving real time spring flow forecasting.

3.3 HBV Hydrologic Model

The HBV model (Bergstrom, 1991) which includes conceptual numerical descriptions of hydrological processes at the catchment scale is best characterized as a semi-distributed conceptual hydrologic model. The model was named after the abbreviation of Hydrologiska Byråns Vattenbalan-avdelning (Hydrological Bureau Waterbalance-section). The HBV model was developed at the Swedish Meteorological and Hydrological Institute (SMHI) and its first application dates back to the early 1970s (Lindstorm et al., 1997). Since then different versions of HBV have been applied in some 45 countries with different climate conditions and in different catchments with varying size from small research basin to the continental scale (Sorman et al., 2009).

Originally the model was developed for runoff simulation and hydrological forecasting, but the scope of the applications has increased rapidly (Bergstrom et al., 2002). Over the time the changes in the model structure have been made, but the basic modeling philosophy has been unchanged (Lindstorm et al., 1997). The model was initially designed as a very simple lumped hydrological model and has gradually become more distributed. The model is best characterized as a semi-distributed hydrologic model because a catchment can be divided into a number of sub-basins and an area elevation distribution and a crude classification of land use (forest, open areas and lakes) are made.

The model is usually run on the daily values of precipitation, temperature and estimates of potential evopotranspiration. Flow observations are required to calibrate and validate the model. For most of the application the model is run on a daily time step, but it is possible to use shorter time steps. The evaporation values used are usually monthly averaged, but it is possible to use daily values. The potential evapotranspiration can be calculated using air temperature. The model consists of routines for snow accumulation and melt, soil moisture accounting, runoff generation and a routing procedure. A schematic structure of the HBV model for one sub-basin is shown in the Figure 10. The snowmelt routine of the HBV model is a degree-day approach. It is based on air temperature, with a water holding capacity of snow which delays runoff. The routine is described as:

$$\text{Snowmelt} = \text{CFMAX}\ (\text{T-TT}) \qquad (2)$$

where the variable T is the temperature in the elevation zone and CFMAX is the melting factor, TT is a threshold temperature below which precipitation is assumed to be snow. Another snow routine parameter DTTM is a value to be added to TT to give the threshold temperature for snow. The snow accumulation and melt routine can be ignored in the model when it is used in a catchment without snow. Glacier melting will occur in glacier zones and follows the same type of formula as used in the snow melting but with a different degree-day factor. The soil moisture routine of the HBV model is the main part controlling runoff formation, accounts for soil field capacity and change in soil moisture storage due to rainfall/snowmelt and evaptranspiration. . The routine is described as:

$$\Delta Q/\Delta P = (SM/FC)^{\beta} \qquad (3)$$

where the ratio $\Delta Q/\Delta P$ is often called the runoff coefficient, FC is the maximum soil moisture storage in the model, SM is the soil moisture storage and β is an empirical parameter that controls the contribution to the runoff response routine and increases in the soil moisture storage. This routine is also based on another parameter LP, value of soil moisture above which evapotranspiration reaches its potential level. So it is fundamental that the contribution to runoff from rain and snowmelt is small at dry condition of soil and large when the soil is wet. Also the actual evapotranspiration decrease with the decrease of moisture content in the soil. In the response routing, excess water from the soil moisture zone transforms to runoff. The response function of the model consists of two reservoir – one upper non linear, one lower linear, and one transform function. The outflow from the upper reservoir and lower reservoir is described by the following functions respectively:

$$Q_0 = k.UZ^{(1+\alpha)} \tag{4}$$

$$Q1 = k4.LZ \tag{5}$$

where Q_0 = reservoir outflow upper reservoir (mm), UZ = reservoir content upper reservoir (mm), k = recession coefficient upper reservoir, Q_1 = reservoir outflow lower reservoir (mm), LZ = reservoir content lower reservoir (mm), k4 = recession coefficient lower reservoir, α is measure of non-linearity. Finally, the runoff is computed by adding the contribution from upper and lower reservoir. Then the generated runoff is routed through a transformation function in order to get a proper shape of the hydrograph at the outlet of the subbasin.

RF	Rainfall
SF	Snowfall
SM	soil moisture storage
EA	Actual evapotranspiration
EI	Intercepting Evaporation
EL	Lake Evaporation
UZ	Content of upper reservoir
LZ	Content of lower reservoir
R	Recharge
CF	Capillary flow
Q_0	Outflow from upper reservoir
Q_1	Outflow from lower reservoir
Q	Outflow from transformation function
PERC	Percolation rate
Fc	Maximum soil moisture content
Lp	Limit of potential evapotranspiration

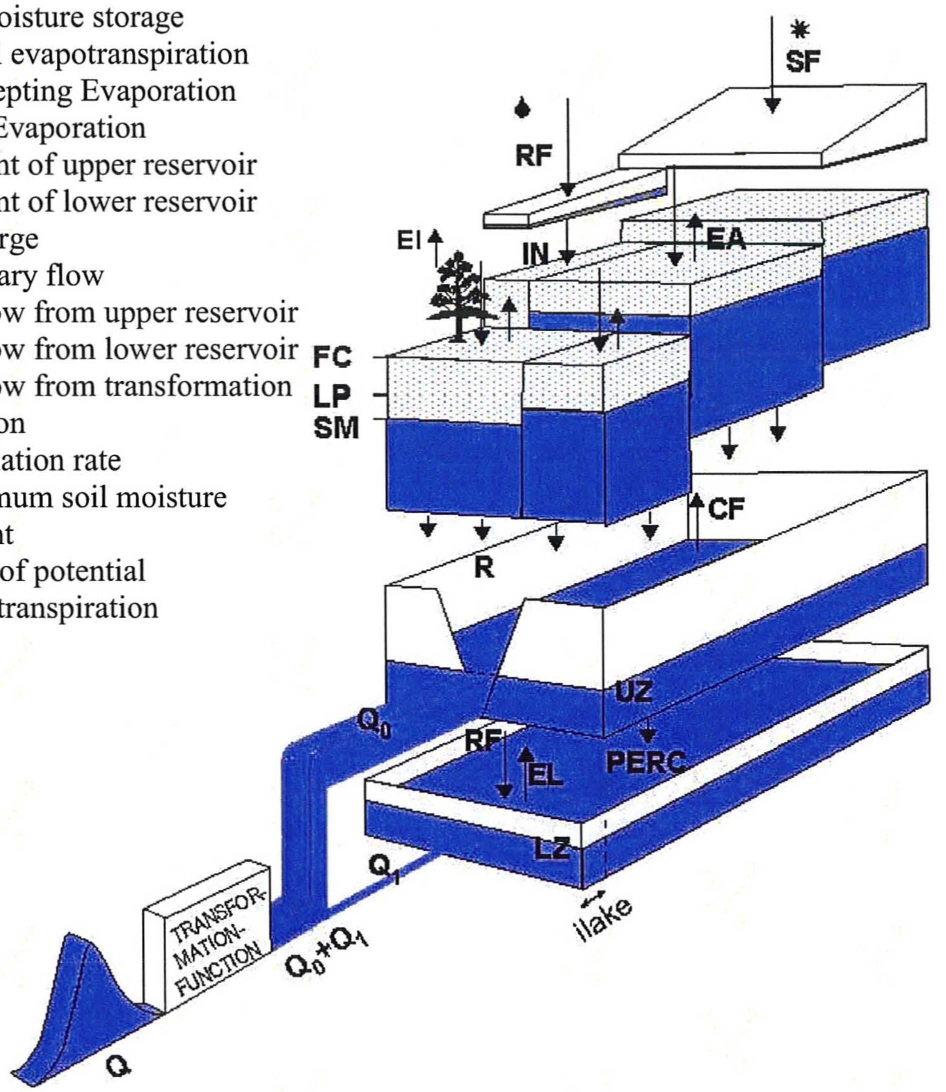

Figure 10: Schematic structure of the HBV model for one sub-basin (SMHI 2006)

3.4 Model Calibration and Validation

Being a conceptual model, the parameters of the HBV model need to be calibrated in order to provide model output that closely resembles observed data. According to the recommendation of the HBV manual, while calibrating the evaluation of the results is mainly done in three different ways – calculating the explained variance R^2, visually inspecting and comparing the simulated and observed hydrographs, and assessing the accumulated difference between the simulated and observed flow.

The HBV manual (SMHI, 2006) also recommends using 75% of the total data for calibration and 25% for validation. For the study basins, observed meteorological data and NCEP meteorological forecasts are available for the period 1979 to 2001, but the observed flow data at hydrometric station (station ID 062214) with latitude and longitude of 49.41°N and 71.22°W respectively are available for the period 1991 to 2001. Therefore, for the Serpent River flow simulation, the first 8 years data (from 1991 to 1998) were used to calibrate the hydrologic model and the last 3 years of data (from 1999 to 2001) were used to validate the model. In the case of the reservoir inflow simulation, a total of 23 years (from 1979 to 2001) are available, of which the first 18 years (from 1979 to 1996) were used for model calibration and the last 5 years (from 1997 to 2001) for model validation. In cases where only the observed meteorological data were used for hydrologic simulation, the model parameters used by Liu, 2007 are used in this study. When using NCEP ensemble (15-member) meteorological forecasts variables for ensemble (15-member) flow simulation, the model is calibrated for the first members for each forecast range (5 to 14 days ahead), and the parameters obtained for the first member are used for the other 14 members to simulate flows for each forecast range . In all cases, the calibrated models were used to simulate the reservoir inflow and the Serpent River flow for validation period.

3.5 Hydrologic Simulation

The calibrated HBV models are used to simulate the river flows for the Serpent River and reservoir inflows in Chute-du-Diable sub-basin for forecast ranges of 5-day to 14-day ahead. Recent studies revealed that the hydrologic model performance for flow forecasts mostly deteriorate after 5 days ahead (Liu, 2007). This is why this study aimed to improve hydrologic forecasts using meteorological forecasts for forecast range of 5-day to 14-day ahead. Both the Serpent River flows and reservoir inflows are simulated on a daily time step for two cases: 1) using observed historical meteorological data (without NCEP meteorological forecasts); and 2) using NCEP ensemble meteorological forecasts. Case 1 is used as reference forecast to assess the skill of the ensemble hydrologic forecast when NCEP ensemble meteorological forecasts are used. In case 1, a single deterministic hydrologic forecast is performed; while in case 2, an ensemble of 15 deterministic hydrologic forecasts is achieved.

Chapter 4: Results and Discussion

4.1 Forecasts Assessment Measures

In this section a number of scalar verification measures used for assessment of deterministic and probabilistic forecasts systems are discussed. In addition,, scatter plots, hydrographs and ensemble plots are presented in sections 4.3 and 4.5.

The traditional model verification measures, i.e., root mean square error (RMSE) and correlation coefficient (r), and Nash and Sutcliffe coefficient (R^2) (Nash and Sutcliffe, 1970), are used for evaluation of model performance/forecasts. The root mean square error and correlation coefficient are commonly used to assess any deterministic forecast, and can be also used to verify ensemble forecasts by assessing the ensemble mean (deterministic) (Brown, 2010). Root mean square error incorporates the random errors and biases in its calculations and shows global goodness of fit, and thus provides a general illustration of the overall accuracy of the prediction. The lower the value of RMSE, the better the model performance and forecast. The correlation coefficient measures the variability and it ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), and 0 means no correlation. The Nash and Sutcliffe coefficient is the variance around the mean explained by the model. The optimum value of the Nash and Sutcliffe coefficient is one (1), and a value less than 0.7 represents poor performance (Coulibaly et al., 2000).

The equations used to calculate root mean square error (RMSE) and correlation coefficient (r), Nash and Sutcliffe coefficient (R^2) are as follows:

$$RMSE = \sqrt{\frac{\sum_{i}^{N}(y_i' - y_i)^2}{N}} \tag{6}$$

$$r = \frac{\sum_{i}^{N}(y'_i - y'_{mean})(y_i - y_{mean})}{\sqrt{\sum_{i}^{N}(y'_i - y'_{mean})^2 . \sum_{i}^{N}(y_i - y_{mean})^2}} \tag{7}$$

$$NASH(R^2) = 1 - \frac{\sum_{i}^{N}(y_i - y'_i)^2}{\sum_{i=1}^{N}(y_i - y_{mean})^2} \tag{8}$$

where, y_i is the observed streamflow at time step i, y'_i is forecasted streamflow at time step i, y_{mean} is the mean of observed streamflow, y'_{mean} is the mean of the model simulated streamflow, and N is the number of data points.

Probabilistic verification methods have been used in the assessment of meteorological and climate forecasts (Murphy et al., 1989; Wilks 2000; Hartmann et al., 2002), however they have not been widely used in the field of hydrology (Franz et al., 2003). Two probabilistic verification measures, namely the Brier skill score (BSS) and the rank probability skill score (RPSS) are used in this study to assess the quality of reservoir inflow forecasts and the Serpent River flow forecasts. The Brier skill score (BSS) and the rank probability skill score (RPSS) are based on the Brier score (BS) and the rank probability score (RPS) respectively.

The most commonly used scalar measure for probability forecasts is the Brier score (Brier 1950), which is essentially the mean-square error of probabilistic forecasts. It is usually used for dichotomous predictands (Wilks, 1995). This score can be applied to continuous-valued forecast (Renner et al., 2009), in this case continuous valued forecasts have been converted into a binary event using a threshold filter which can either be exceeded or not (Renner et al., 2009; Roulin, 2007). Theis et al. (2005) also used this score for assessing deterministic forecasts by using a threshold filter. In this study for comparison purpose and consistency, the means (i.e. mean for the entire year, spring, summer, autumn and winter) of the observed flow are considered as BS thresholds. The Brier score BS is calculated by the following equation:

$$BS = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2 \tag{9}$$

where n represents the number of days, k is the number of the n forecast/event pair, y_k is the forecast probability and o_k is the observed probability (occurrence and nonoccurrence of the event being forecast). y_k is derived by the relative frequency of the ensemble members exceeding the chosen threshold. The observations o_k are translated similar to the forecasts, i.e. the observation o_k = 1 if the event occurs (if the threshold is exceeded) and o_k = 0 if the event does not occur. The Brier score ranges between 0 and 1 because the observation and probability forecasts are bounded by 0 and 1. It is negatively oriented with a perfect forecast exhibiting BS=0 and less accurate forecasts receive higher Brier score. In order to provide information on the accuracy of the forecasts relative to reference forecasts, Brier skill score (BSS) is computed as:

$$BSS = \frac{BS_{ref} - BS}{BS_{ref}} * 100 \tag{10}$$

In this study, Brier scores calculated for the forecasts without using NCEP meteorological forecasts are used as reference forecasts. So, the calculated Brier skill score (BSS) represents the percentage improvement of streamflow forecasts (when NCEP meteorological forecasts are used) over the reference forecasts.

The ranked probability score (RPS) (Wilks, 1995) is a generalization of the Brier score to the multi-category (Mullen et. al., 2000). The RPS is calculated by the following equation:

$$RPS = \sum_{m=1}^{j}(Ym - Om)^2 \tag{11}$$

where, Ym is the cumulative probability of the forecast for category m and Om is the cumulative probability of the observation for category m. For a group of n forecasts, the RPS is the average ($\overline{RPS}$) of the n RPSs:

$$\overline{RPS} = \frac{1}{n}\sum_{k=1}^{n} RPS_k \tag{12}$$

It is notable that the RPS values shown in Table 13 through Table 22 are the average of n RPSs ($\overline{RPS}$). The RPS is calculated in the same procedure followed by Clark and Hay, 2004 and Gangopadhyay et al., 2004: At first, the observed time series data are used to differentiate 10 possible categories (j) (i.e. the minimum value to the 10^{th} percentile, the 10^{th} percentile to the 20^{th} percentile, the 20^{th} percentile to the 30^{th} percentile up to the 90^{th} percentile to the maximum value). These categories are determined separately for the entire year, the spring, the summer, the autumn and the winter season. Then, the number of ensemble members forecast in each category is determined (out of 15 members) and their cumulative probabilities are calculated for each forecast-observation pair. Next, in the same way, the observation's cumulative probabilities are calculated. Here, all categories below the observation's position are assigned '0', and all categories equal to and above the observation's position are assigned '1'. Then, the RPS is computed as the squared difference between cumulative probabilities of the observed and forecast and the summation of squared differences over all categories. RPS is zero for a perfect forecast and positive otherwise (Weigel et al., 2006). The quality of forecasts is difficult to assess based on RPS alone (Franz et al., 2003); therefore, ranked probability skill score (RPSS) is calculated in order to provide information on the accuracy of the forecasts relative to reference forecasts:

$$RPSS = \frac{\overline{RPS}_{ref} - \overline{RPS}}{\overline{RPS}_{ref}} * 100 \tag{13}$$

In this study, RPSs calculated for the hydrologic forecasts without using NCEP meteorological forecasts are used as reference forecasts. So, the calculated RPSSs represent the percentage improvement of streamflow forecasts (when NCEP meteorological forecasts are used) over the reference forecasts.

4.2 Model Performance Assessment with Deterministic Measures

4.2.1 Reservoir Inflow Forecasting

The reservoir inflows were simulated for two cases 1) using observed historical meteorological data (without NCEP meteorological forecasts); and 2) using NCEP ensemble meteorological forecasts. The model performance statistics root mean square error (RMSE) and correlation coefficient (r), Nash and Sutcliffe coefficient (R^2) were calculated for the validation period (1997-2001) to evaluate model performance.

Table 3 presents the three model performance statistics for 5 to 14 days ahead forecasting for the entire year. It is shown in Table 3 that in both cases without NCEP meteorological forecasts and with meteorological forecasts, the model performance deteriorate with the increase of forecast lead time as the RMSE increase and r and R^2 decrease with the increase of lead time. Lower RMSE and higher r and R^2 values are indicative of better forecasts. It can be seen from Table 3 that when NCEP meteorological forecasts are not included in the model, performance of the model is not satisfactory even on the 5 day ahead as the R^2 value is 0.69, which is less than 0.7. When NCEP meteorological forecasts are used for reservoir inflow simulation, the R^2 value for 5 day ahead increased significantly from 0.69 to 0.81. It can also be seen from Table 3 that when meteorological forecasts are used, the model performs well up to 8 day ahead with a R^2 value 0.72. When meteorological forecasts are used, the correlation coefficient (r) increased from 0.85 to 0.91 for 5 day ahead and from 0.66 to 0.80 for 14 day ahead, increases of r in other forecast range is significant as forecast range 5 and 14 days ahead, which shows a clear improvement. Table 3 also shows that the RMSE increased from 113

m^3/sec for 5 day ahead to 158 m^3/sec for 14 day ahead when meteorological forecast data are not used, and increased from 89 m^3/sec for 5 day ahead to 123 m^3/sec for 14 day ahead when meteorological forecast data are used. It can also be seen from Table 3 that there is on average a 20% decrease in RMSE when meteorological forecast data are used in the model for reservoir inflow forecasting. So all the model verification measures show that there is a significant improvement in reservoir inflow forecasting when NCEP numerical meteorological forecast variables (accumulated precipitation and mean temperature) are used in the HBV model.

The hydrologic model forecasts were also investigated on a seasonal basis. The simulated reservoir inflows from the entire year were divided into four seasons – Spring (March to May), Summer (June to August), Autumn (September to November) and Winter (December to February). The model performance statistics were then calculated for each season. The model performance statistics for Spring, Summer, Autumn and Winter are presented in Tables 4, 5, 6 and 7 respectively. Table 4 shows that when meteorological forecasts are not used, the models do not perform well after 5 day ahead for the spring season, and the model performance deteriorates significantly with the increase of forecast lead time (as example from 5 day ahead forecasts to 14 ahead forecasts the r decreased from 0.87 to 0.64 and R^2 decreased from 0.74 to 0.35). It can also be seen from the Table 4 that when meteorological forecasts are used, the model performs well up to 13 day ahead with a R^2 value 0.71. Table 4 also shows that when meteorological forecasts are used, the correlation coefficient (r) increased from 0.87 to 0.96 for 5 day ahead and from 0.64 to 0.83 for 14 day ahead, increases of r in other forecast range is significant as forecast range 5 and 14 day ahead, this shows a clear improvement in the spring season. Moreover, when meteorological forecasts are used, the RMSE decreased from 171 m^3/sec to 98 m^3/sec and from 261 m^3/sec to 189 m^3/sec for 5 day ahead and 14 day ahead respectively. The calculated decrease in RMSE shows that there is on average a 30% decrease when NCEP meteorological forecast variables are used. This decrease in RMSE for the spring season is much higher than the decrease

(20%) in RMSE for the entire year. Tables 5, 6 and 7 show that the model's performance is very poor for summer, autumn and winter, which is revealed by the low values of Nash and Sutcliffe coefficient (R^2). Khan and Coulibaly, 2006 reported that the snowmelt runoff is responsible for high flows in the spring season as well as up to 40% of the annual flow volume in the study region. Close examination of Tables 4, 5, 6 and 7 indicates that the improvement of reservoir flow forecast in the spring season is much better than all other seasons, and there is also deterioration in forecast quality in other seasons when NCEP forecasts are used. The reason of this improvement in the spring season is that the good temperature forecasts (shown in section 2.2.3) by the NCEP up to 14 day ahead which influence the snowmelt dominated spring flow in the study region.

Table 3: Comparison of reservoir inflow forecasts from 5 to 14 day ahead for the entire year

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	*RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	113.16	0.85	0.69	88.80	0.91	**0.81**	22
6-day-ahead	125.50	0.81	0.62	104.49	0.86	**0.74**	17
7-day-ahead	130.81	0.78	0.59	105.71	0.86	**0.73**	19
8-day-ahead	132.74	0.78	0.57	107.98	0.85	**0.72**	19
9-day-ahead	133.90	0.77	0.57	113.23	0.84	0.69	15
10-day-ahead	141.61	0.74	0.52	119.36	0.82	0.66	16
11-day-ahead	147.35	0.71	0.48	114.58	0.83	0.68	22
12-day-ahead	149.79	0.70	0.46	118.48	0.82	0.66	21
13-day-ahead	153.91	0.67	0.43	117.90	0.82	0.66	23
14-day-ahead	158.90	0.66	0.39	122.52	0.80	0.64	23

*RMSE (m^3/s)

Table 4: Comparison of reservoir inflow forecasts from 5 to 14 day ahead for spring season

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	171.47	0.87	**0.74**	97.77	0.96	**0.92**	43
6-day-ahead	205.47	0.81	0.62	142.01	0.91	**0.82**	31
7-day-ahead	216.57	0.79	0.57	134.48	0.92	**0.84**	38
8-day-ahead	222.58	0.77	0.55	145.28	0.90	**0.81**	35
9-day-ahead	221.88	0.77	0.55	162.19	0.88	**0.77**	27
10-day-ahead	240.35	0.72	0.46	169.29	0.87	**0.75**	30
11-day-ahead	251.40	0.68	0.41	165.85	0.87	**0.76**	34
12-day-ahead	256.00	0.67	0.38	179.99	0.85	**0.71**	30
13-day-ahead	261.77	0.66	0.35	180.07	0.85	**0.71**	31
14-day-ahead	261.11	0.64	0.35	189.17	0.83	0.68	28

Table 5: Comparison of reservoir inflow forecasts from 5 to 14 day ahead for summer season

Forecasting lags	Without meteorological forecasts			With meteorological forecasts			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	104.94	0.46	0.06	104.37	0.49	0.17	1
6-day-ahead	102.01	0.44	0.09	103.08	0.45	0.19	-1
7-day-ahead	102.51	0.42	0.05	113.02	0.39	0.03	-10
8-day-ahead	101.52	0.41	0.06	110.89	0.40	0.06	-9
9-day-ahead	102.07	0.39	0.04	109.88	0.46	0.08	-8
10-day-ahead	100.93	0.41	0.07	119.30	0.39	-0.08	-18
11-day-ahead	103.78	0.39	0.02	109.03	0.45	0.10	-5
12-day-ahead	105.81	0.33	-0.02	103.25	0.47	0.19	2
13-day-ahead	110.85	0.26	-0.11	105.20	0.45	0.16	5
14-day-ahead	121.90	0.19	-0.34	105.49	0.43	0.15	13

Table 6: Comparison of reservoir inflow forecasts from 5 to 14 day ahead for autumn season

Forecasting lags	Without meteorological forecasts			With meteorological forecasts			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	91.81	0.67	0.11	94.50	0.47	0.06	-3
6-day-ahead	88.24	0.63	0.18	102.54	0.41	-0.11	-16
7-day-ahead	91.90	0.59	0.11	108.00	0.26	-0.23	-18
8-day-ahead	89.24	0.57	0.16	105.84	0.36	-0.18	-19
9-day-ahead	93.96	0.54	0.08	104.04	0.30	-0.14	-11
10-day-ahead	95.37	0.52	0.05	108.31	0.31	-0.23	-14
11-day-ahead	97.43	0.50	0.01	104.25	0.31	-0.14	-7
12-day-ahead	98.41	0.48	0.00	102.35	0.39	-0.10	-4
13-day-ahead	102.54	0.46	-0.08	99.28	0.46	-0.04	3
14-day-ahead	116.84	0.44	-0.40	103.42	0.39	-0.12	11

Table 7: Comparison of reservoir inflow forecasts from 5 to 14 day ahead for winter season

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	46.02	0.68	-0.13	45.31	0.72	0.23	2
6-day-ahead	47.30	0.66	-0.29	46.77	0.73	0.18	1
7-day-ahead	46.88	0.64	-0.45	44.73	0.74	0.25	5
8-day-ahead	47.53	0.61	-0.62	42.81	0.79	0.31	10
9-day-ahead	53.33	0.59	-1.18	42.96	0.78	0.31	19
10-day-ahead	52.01	0.58	-1.23	45.85	0.79	0.21	12
11-day-ahead	53.73	0.56	-1.59	44.87	0.77	0.25	16
12-day-ahead	52.92	0.55	-1.59	48.48	0.65	0.12	8
13-day-ahead	53.63	0.52	-1.83	44.43	0.72	0.26	17
14-day-ahead	61.30	0.49	-2.90	46.22	0.73	0.20	25

4.2.2 Serpent River Flow Forecasting

The Serpent River flows were also simulated for two cases: 1) using observed historical meteorological data (without NCEP meteorological forecasts) and 2) using NCEP ensemble meteorological forecasts. The model performance statistics root mean square error (RMSE) and correlation coefficient (r), Nash and Sutcliffe coefficient (R^2) were calculated for the validation period (1999-2001) to evaluate the model performance.

The three model performance statistics for 5 to 14-day ahead forecasting for the Serpent River flow for entire year are presented in Table 8. Table 8 shows that in both cases without NCEP meteorological forecasts and with NCEP meteorological forecasts, the model performance deteriorated with the increase of forecast lead time as the RMSE increase and r and R^2 decrease with the increase of lead time. It is fundamental that better forecasts exhibit lower RMSE and higher r and R^2. Table 8 also shows that when meteorological forecasts are not included in the model, it does not perform well even on the 5 day ahead as the R^2 value 0.60, which is less than 0.70. When NCEP meteorological forecasts are used for the Serpent River flow simulation, the R^2 value for 5 day ahead increased significantly from 0.60 to 0.73. It is also shown in Table 8 that when NCEP meteorological forecasts are used, the model performs well up to 7 days ahead with R^2 value 0.71, which is greater than 0.70. When meteorological forecasts are used, the correlation coefficient (r) increased from 0.82 to 0.87 for 5 day ahead and from 0.64 to 0.76 for 14 day ahead, increases of r in other forecast range are as significant as that in forecast range 5 and 14 day ahead; this shows a clear improvement. It can also be seen from Table 8 that the RMSE increased from 28.71 m^3/sec for 5 day ahead to 37.86 m^3/sec for 14 day ahead when meteorological forecast data are not used, it increased from 23.43 m^3/sec for 5 day ahead to 30.70 m^3/sec for 14 day ahead when meteorological data are used. Table 8 shows that there is on average about a 19% decrease in RMSE when meteorological forecast data are used in the model for the Serpent River flow forecasting. Therefore, all of the model verification measures show that there is a significant improvement in the Serpent River flow forecasting when NCEP numerical meteorological forecast variables (accumulated precipitation and mean temperature) are used in the HBV model.

The hydrologic model forecasts were also investigated on a seasonal basis. Similar to reservoir inflow forecasting, the simulated Serpent River flows from the entire year were divided into four seasons – Spring (March to May), Summer (June to August), Autumn (September to November) and Winter (December to February). Then the model performance statistics were calculated for each season. The model performance statistics for Spring, Summer, Autumn and Winter are presented in Tables 9, 10, 11 and 12 respectively. It can be seen from Table 9 that when NCEP meteorological forecasts are not used, the models do not perform well even 5 days ahead for the spring season, and the model performance deteriorates significantly with the increase of forecast lead time (as example from 5 day ahead forecasts to 14 ahead forecasts the r decreased from 0.84 to 0.63 and R^2 decreased from 0.66 to 0.29). Table 9 also shows that when meteorological forecasts are used, the model performs well up to 10 day ahead with an R^2 value 0.72. It is also shown in Table 9 that when meteorological forecasts are used, the correlation coefficient (r) increased from 0.84 to 0.93 for 5 day ahead and from 0.63 to 0.80 for 14 day ahead; increases of r in other forecast range is as significant as forecast range 5 and 14 day ahead; which shows a clear improvement in the spring season. Moreover, when meteorological forecasts are used, RMSE decreased from 44.64 m^3/sec to 28.79 m^3/sec and from 63.25 m^3/sec to 48.18 m^3/sec for 5 day ahead and 14 day ahead respectively. The calculated decrease in RMSE shows that there is on average about a 30% decrease when meteorological forecast variables are used. This decrease in RMSE for the spring season is much higher than the decrease (19%) in RMSE for the entire year . Tables 10, 11 and 12 show that the model performance is very poor for summer, autumn and winter revealed by the very small Nash and Sutcliffe coefficient (R^2) values. Close examination of Tables 9, 10, 11 and 12 indicates that improvement of the Serpent River flow forecast in the spring season is much better than all other seasons and there is a deterioration in forecast quality in other seasons when NCEP forecasts are used. The reason for this improvement in the spring season is that the spring flow of the Serpent River is dominated by snow melt, which is influenced predominantly by the temperature variations; and temperature forecasts by the National Centers for Environmental Prediction's (NCEP) Global Forecast System up to 14 day ahead are good.

Table 8: Comparison of Serpent River flow forecasts from 5 to 14 day ahead for the entire year

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	28.71	0.82	0.60	23.43	0.87	**0.73**	18
6-day-ahead	31.11	0.78	0.53	24.09	0.86	**0.72**	23
7-day-ahead	31.95	0.77	0.51	24.59	0.86	**0.71**	23
8-day-ahead	33.62	0.73	0.45	25.94	0.84	0.67	23
9-day-ahead	35.05	0.70	0.41	26.75	0.82	0.65	24
10-day-ahead	35.00	0.70	0.41	28.82	0.80	0.60	18
11-day-ahead	35.50	0.69	0.39	30.58	0.77	0.55	14
12-day-ahead	35.78	0.69	0.38	31.28	0.75	0.53	13
13-day-ahead	37.45	0.65	0.32	31.10	0.75	0.53	17
14-day-ahead	37.86	0.64	0.31	30.70	0.76	0.54	19

*RMSE (m^3/s)

Table 9: Comparison of Serpent River flow forecasts from 5 to 14 day ahead for spring season

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	44.64	0.84	0.66	28.79	0.93	**0.86**	36
6-day-ahead	50.09	0.80	0.58	31.66	0.91	**0.84**	37
7-day-ahead	51.02	0.80	0.56	27.85	0.93	**0.87**	45
8-day-ahead	55.38	0.76	0.48	31.63	0.92	**0.84**	43
9-day-ahead	58.99	0.72	0.40	37.49	0.88	**0.77**	36
10-day-ahead	58.34	0.73	0.41	41.12	0.86	**0.72**	30
11-day-ahead	59.22	0.71	0.39	47.18	0.81	0.63	20
12-day-ahead	59.30	0.69	0.39	50.01	0.78	0.59	16
13-day-ahead	62.84	0.65	0.31	50.24	0.78	0.58	20
14-day-ahead	63.25	0.63	0.29	48.18	0.80	0.62	24

Table 10: Comparison of Serpent River flow forecasts from 5 to 14 day ahead for summer season

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	20.83	0.53	0.17	22.68	0.41	**0.05**	-9
6-day-ahead	20.67	0.52	0.16	20.86	0.46	**0.19**	-1
7-day-ahead	21.53	0.48	0.06	24.87	0.25	**-0.15**	-16
8-day-ahead	20.72	0.49	0.13	25.16	0.24	**-0.17**	-21
9-day-ahead	20.48	0.50	0.15	21.25	0.44	**0.16**	-4
10-day-ahead	21.08	0.47	0.10	23.04	0.34	**0.02**	-9
11-day-ahead	21.58	0.45	0.07	21.75	0.39	0.12	-1
12-day-ahead	22.47	0.40	0.00	21.29	0.45	0.16	5
13-day-ahead	23.14	0.37	-0.06	21.65	0.42	0.13	6
14-day-ahead	23.65	0.34	-0.11	21.87	0.40	0.11	8

Table 11: Comparison of Serpent River flow forecasts from 5 to 14 day ahead for autumn season

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	21.84	0.58	0.24	22.49	0.50	**0.22**	-3
6-day-ahead	23.20	0.51	0.14	23.51	0.57	**0.15**	-1
7-day-ahead	24.57	0.46	0.03	26.45	0.40	**-0.08**	-8
8-day-ahead	24.69	0.43	0.01	27.04	0.36	**-0.13**	-10
9-day-ahead	24.93	0.41	-0.02	26.51	0.25	**-0.08**	-6
10-day-ahead	25.15	0.40	-0.05	27.90	0.19	**-0.20**	-11
11-day-ahead	25.63	0.38	-0.09	27.13	0.21	-0.13	-6
12-day-ahead	26.25	0.34	-0.15	25.47	0.27	0.00	3
13-day-ahead	26.68	0.30	-0.19	24.36	0.45	0.09	9
14-day-ahead	27.64	0.25	-0.28	25.49	0.44	0.00	8

Table 12: Comparison of Serpent River flow forecasts from 5 to 14 day ahead for winter season

Forecasting lags	**Without meteorological forecasts**			**With meteorological forecasts**			Decrease in RMSE (%)
	RMSE	r	R^2	RMSE	r	R^2	
5-day-ahead	19.54	0.43	-3.58	18.50	0.29	-2.85	5
6-day-ahead	19.52	0.42	-3.53	18.09	0.34	-2.68	7
7-day-ahead	19.93	0.37	-3.69	17.93	0.36	-2.61	10
8-day-ahead	19.94	0.36	-3.65	17.91	0.39	-2.61	10
9-day-ahead	19.26	0.38	-3.29	17.16	0.45	-2.31	11
10-day-ahead	19.92	0.34	-3.54	17.68	0.48	-2.51	11
11-day-ahead	19.71	0.34	-3.38	17.08	0.41	-2.28	13
12-day-ahead	19.69	0.35	-3.30	17.19	0.37	-2.32	13
13-day-ahead	19.67	0.36	-3.21	16.29	0.45	-1.98	17
14-day-ahead	19.58	0.36	-3.10	17.45	0.45	-2.42	11

4.3 Visual Inspection of Ensemble Mean

The scatter plot is the familiar and probably the simplest tool for visual inspection of simulated and observed data. It is simply a collection of points plotted on a graph whose ordinate and abscissa are the values of each member of the data pair. It allows examining features like trends in data, clustering of one or both variables, changes in one variable as a function of the other and extraordinary points or outliers. The ordinate and abscissa are plotted on the same scale, in which case perfection is represented by any point on the 45 degree line which is usually drawn to facilitate interpretation of the scatter plot. In this section scatter plots and hydrographs are presented for both reservoir flow forecasts and the Serpent River flow forecasts for 5 day ahead, 8 day ahead, 11 day ahead and 14 day ahead for the entire year. It is notable that in the case of flow forecasting using NCEP meteorological forecasts, the plots represent the mean of the ensemble (15 member) flow forecasts.

4.3.1 Reservoir Inflow Forecasting

The scatter plots for 5 day ahead, 8 day ahead, 11 day ahead and 14 day ahead for the entire year are presented in Figures 11, 13, 15 and 17 for reservoir inflow forecasts without meteorological forecasts and Figures 12, 14, 16 and 18 for reservoir inflow forecasts with NCEP metrological forecasts. These plots present all the data for the validation period (from 1997 to 2001). Figures 11 and 12 reveal that data are more clustered along the 45 degree line in case of using NCEP meteorological forecasts. For example, for observed reservoir inflow 900 m^3/sec, the simulated flows in Figure 11 are scattered between about 350 m^3/sec and 1500 m^3/sec and the simulated flows in Figure 12 are scattered between about 600 m^3/sec and 1050 m^3/sec. The scatter plots in Figures 11 and 12 reveal that there is an improvement in the reservoir inflow forecasts for 5 day ahead when NCEP meteorological forecasts are used in the hydrologic model. The overall examination of Figures 13, 14, 15, 16, 17 and 18 show a similar phenomenon for forecast ranges of 8, 11 and 14 day ahead. The scatter plots also show that in both cases without meteorological forecasts and with meteorological forecasts the data are more dispersed from the 45 degree line with increased forecast range. It can be seen from the plots that when the model underestimates the high flows, underestimations of flows are lower when the NCEP meteorological forecasts are used than without meteorological forecasts. For example, for observed reservoir inflow 1200 m^3/sec, the simulated flows in Figure 17 are scattered between about 300 m^3/sec and 800 m^3/sec and the simulated flows in Figure 18 are scattered between about 600 m^3/sec and 900 m^3/sec.

The hydrographs for reservoir flow forecasts for 5, 8, 11 and 14 day ahead for the entire year are presented in Figures 19, 20, 21 and 22 respectively. For better visualization, hydrographs are shown for 3 years (1999-2001) out of 5 years (1997-2001) of the validation period. In the hydrographs, the green line, red line and blue line represent the flow forecasts (mean of 15 members) with NCEP meteorological forecasts, flow forecasts without NCEP forecasts and the observed flows respectively. An overall examination of the hydrographs shows that simulated flow series using NCEP meteorological forecasts are better than those without using NCEP meteorological forecasts in the spring season. For 5 day ahead forecasting, peak flows in May 2000 and May 2001 were underestimated slightly when using NCEP meteorological forecasts and this underestimation is higher when NCEP meteorological forecasts are not used. It can be seen from Figure 19 that in May 1999, the peak flow is underestimated when using NCEP meteorological forecasts and overestimated when NCEP meteorological forecasts are not used. In all the spring seasons, the agreement between the forecasted flows using NCEP meteorological forecasts and the observed flows in both rising limb and the falling limb of the hydrographs are more accurate than that in case of flow forecasts without NCEP meteorological forecasts. This phenomenon can be also seen for other forecast ranges (8, 11 and 14 day ahead) as shown in Figures 20, 21 and 22.

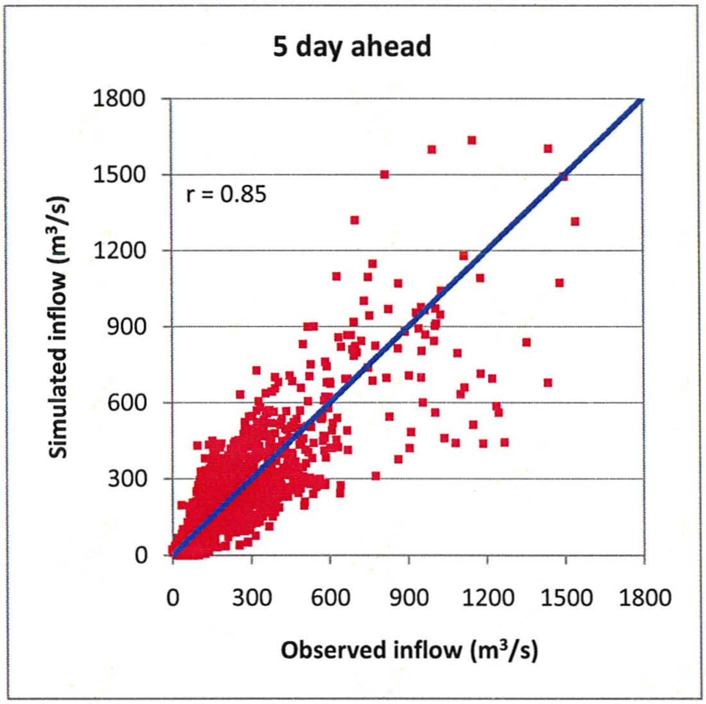

Figure 11: Scatter plot for reservoir inflows without meteorological forecasts (5 day ahead)

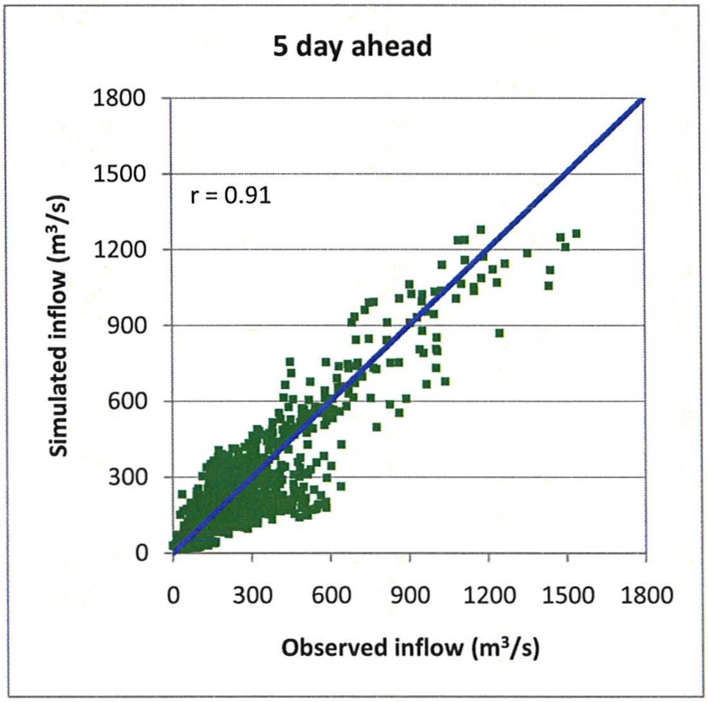

Figure 12: Scatter plot for reservoir inflows with meteorological forecasts (5 day ahead)

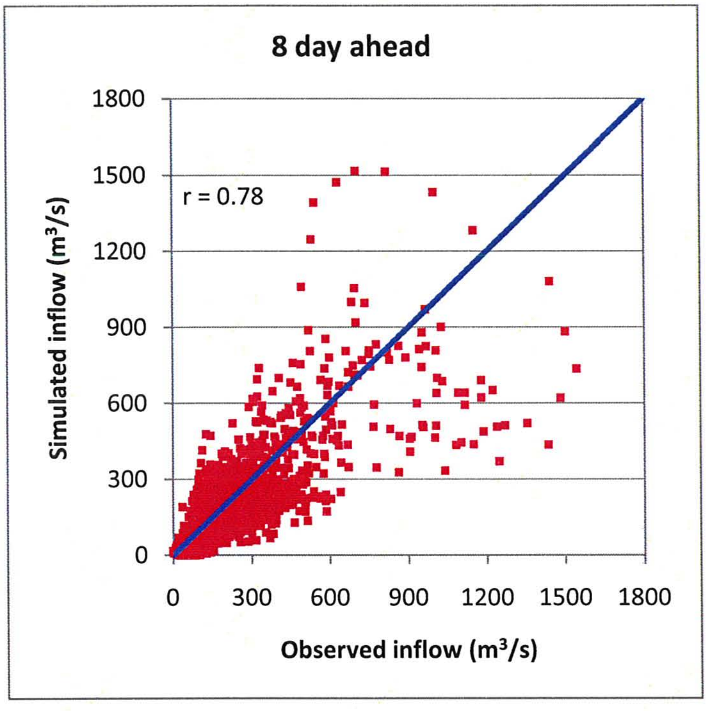

Figure 13: Scatter plot for reservoir inflows without meteorological forecasts (8 day ahead)

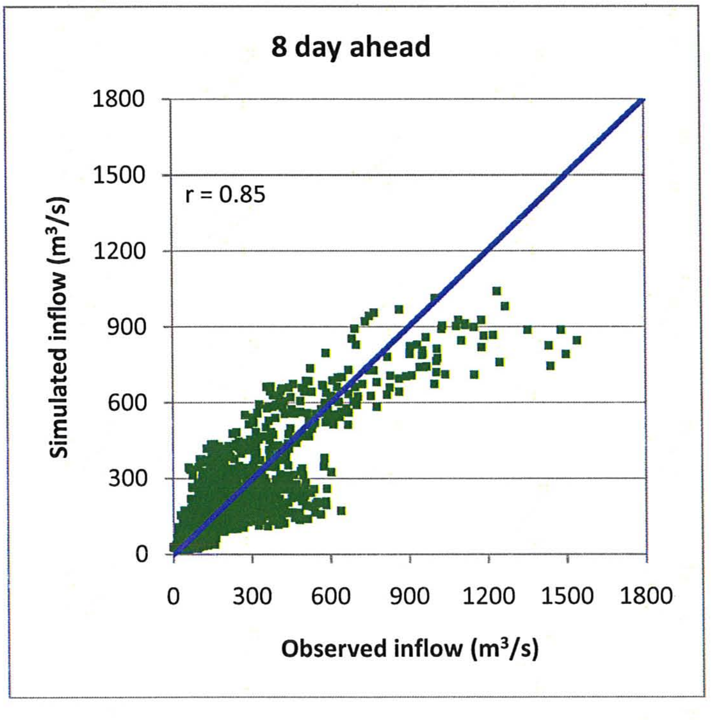

Figure 14: Scatter plot for reservoir inflows with meteorological forecasts (8 day ahead)

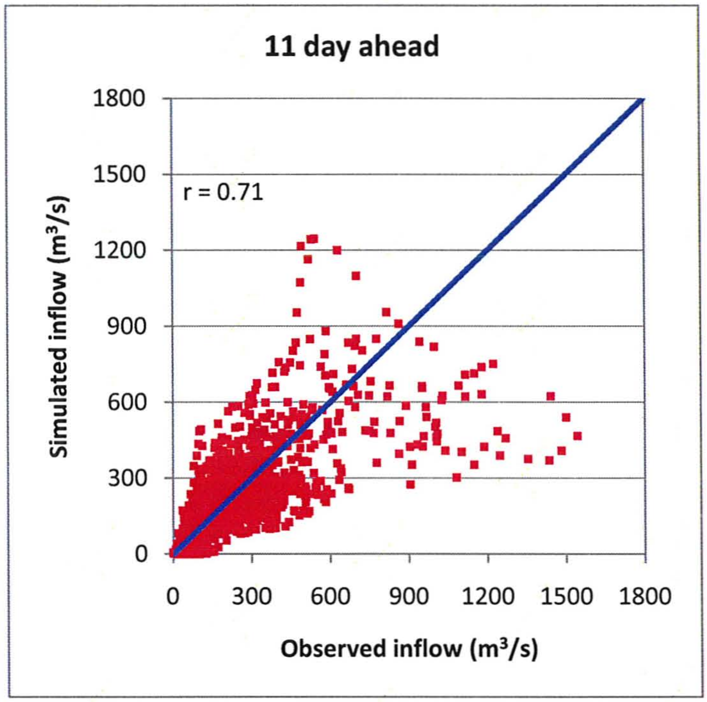

Figure 15: Scatter plot for reservoir inflows without meteorological forecasts (11day ahead)

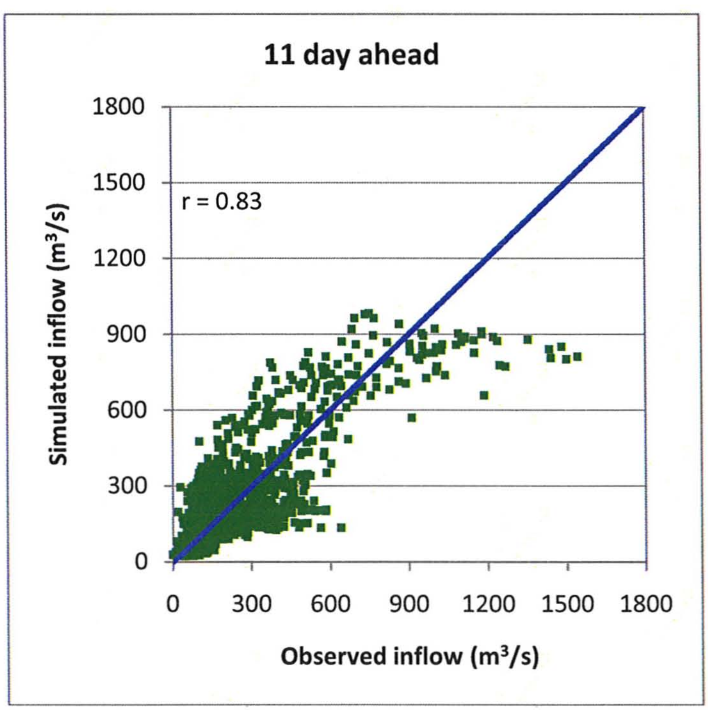

Figure 16: Scatter plot for reservoir inflows with meteorological forecasts (11 day ahead)

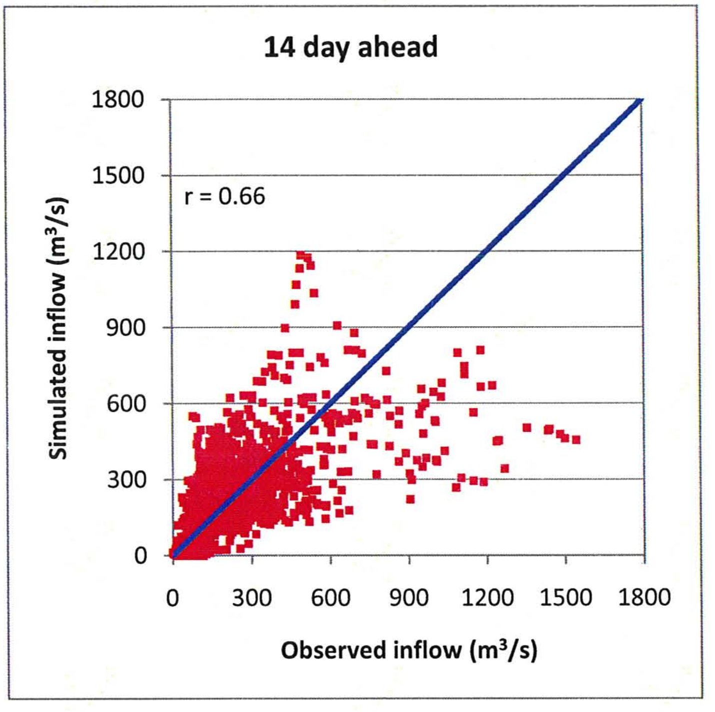

Figure 17: Scatter plot for reservoir inflows without meteorological forecasts (14 day ahead)

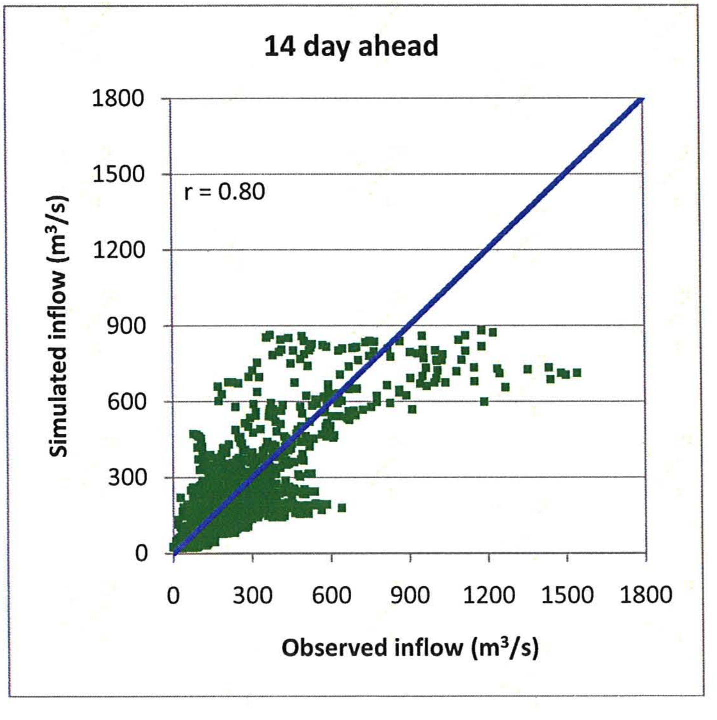

Figure 18: Scatter plot for reservoir inflows with meteorological forecasts (14 day ahead)

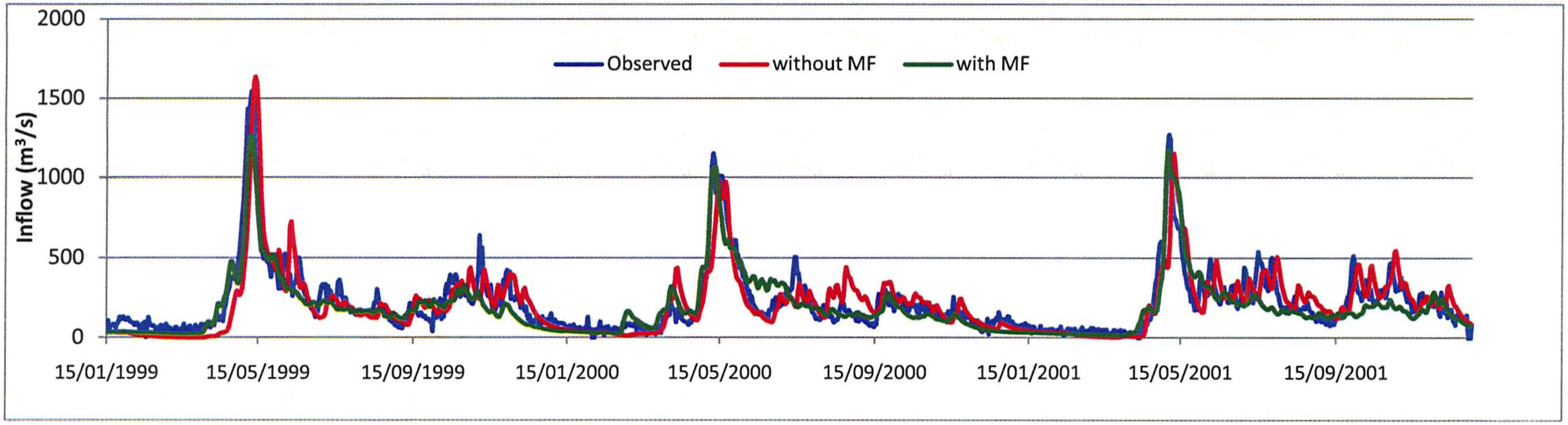

Figure 19: Hydrographs for 5 day ahead reservoir inflows for entire year

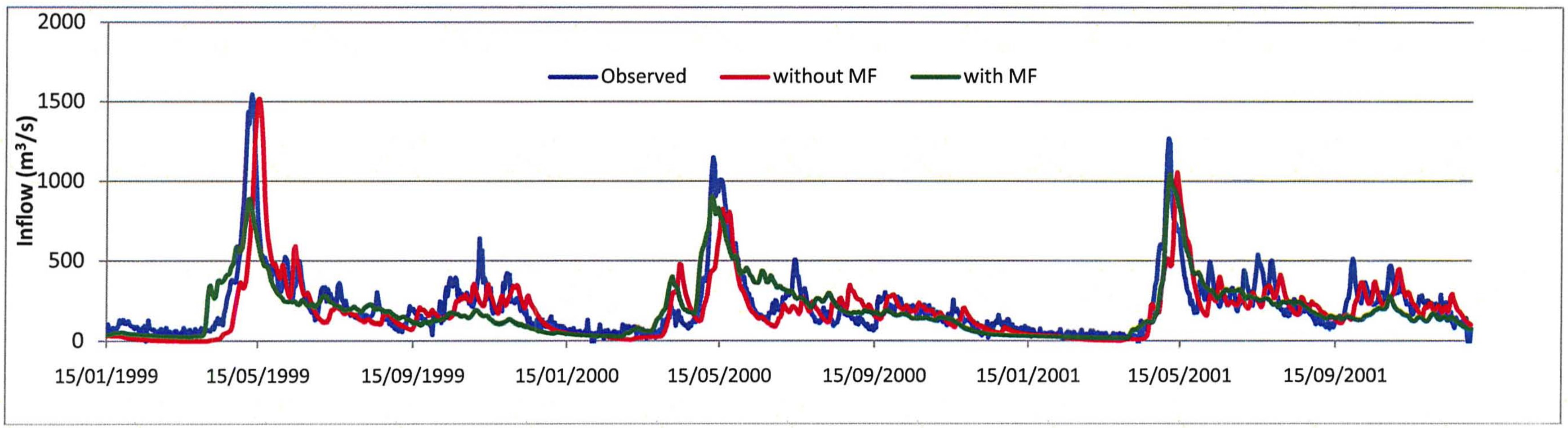

Figure 20: Hydrographs for 8 day ahead reservoir inflows for entire year

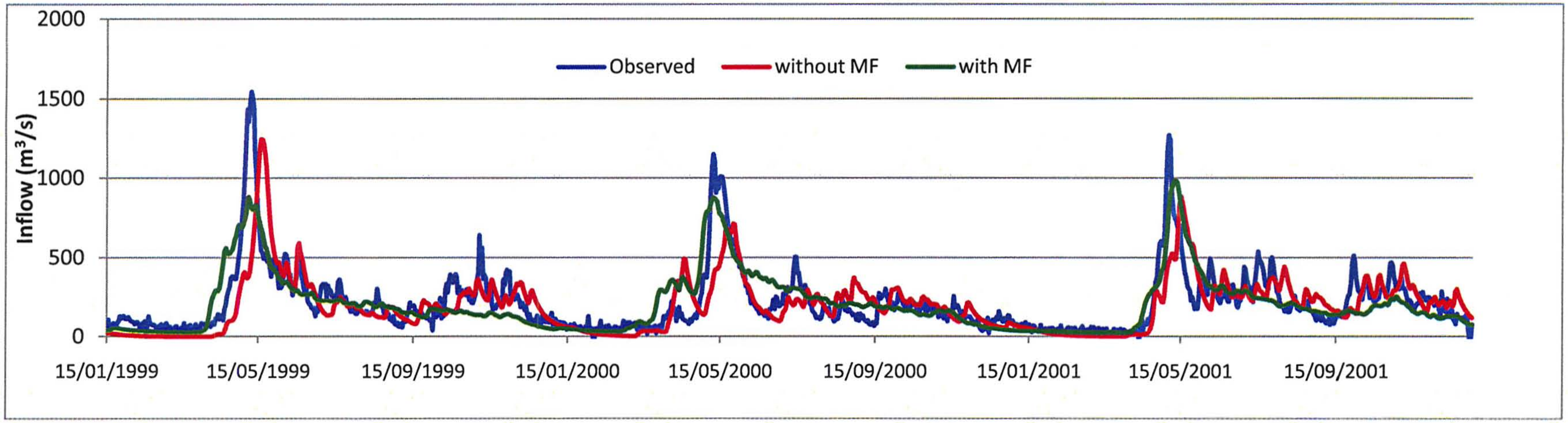

Figure 21: Hydrographs for 11 day ahead reservoir inflows for entire year

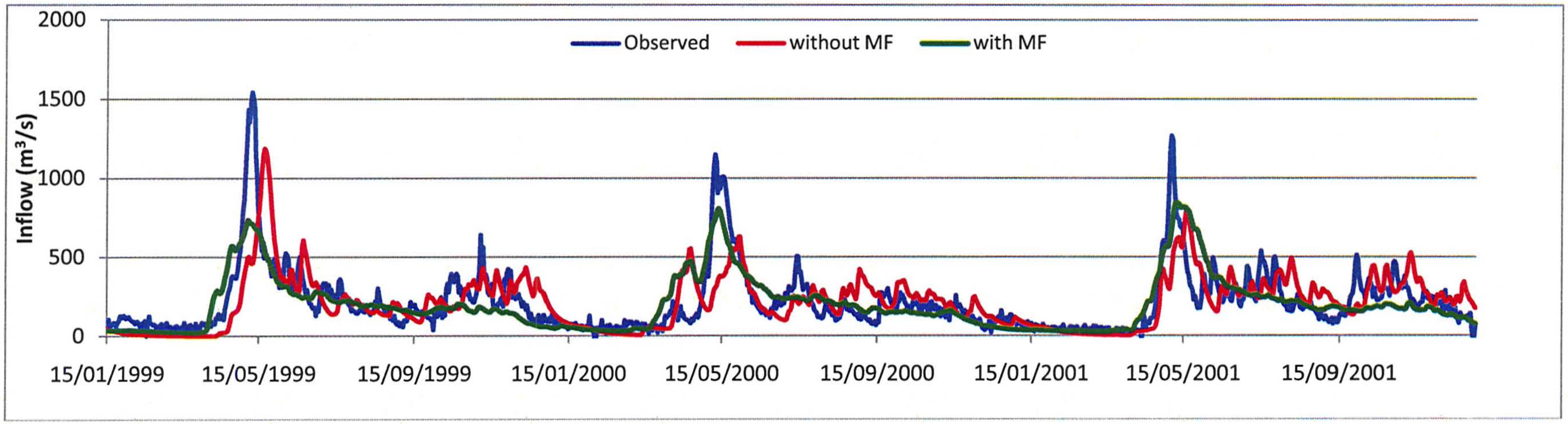

Figure 22: Hydrographs for 14 day ahead reservoir inflows for entire year

4.3.2 Serpent River Flow Forecasting

The scatter plots for 5, 8, 11 and 14 day ahead for the entire year are presented in Figures 23, 25, 27 and 29 for the Serpent River flow forecasts without meteorological forecasts and Figures 24, 26, 28 and 30 for the Serpent River flow forecasts with metrological forecasts. These plots present all of the data for the validation period (from 1999 to 2001). Figures 23 and 24 show that the data are more clustered along the 45 degree line in Figure 24 than that in the Figure 23. For example, for an observed Serpent River flow 150 m^3/sec, the simulated flows in Figure 23 are scattered between about 50 m^3/sec and 200 m^3/sec and the simulated flows in Figure 24 are scattered between about 50 m^3/sec and 150 m^3/sec. The scatter plots in Figure 23 and Figure 24 reveal that there is an improvement in the Serpent River flow forecasts for 5 day ahead when NCEP meteorological forecasts are used in the hydrologic model. The overall examination of Figures 25 & 26, Figures 27 & 28 and Figures 29 & 30 show a similar phenomenon for forecast ranges of 8, 11 and 14 day ahead. The scatter plots also show that in both cases without meteorological forecasts and with meteorological forecasts the data are more dispersed from the 45 degree line with an increase of forecast range. It can be seen from the plots that when the model underestimates the high flows, underestimations of flows are lower when the NCEP meteorological forecasts are used than those without meteorological forecasts. For example, for observed river flow 300 m^3/sec, the simulated flows in Figure 29 is about 150 m^3/sec and the simulated flows in Figure 30 is about 175 m^3/sec.

The hydrographs for river flow forecasts for 5 day ahead, 8 day ahead, 11 day ahead and 14 day ahead for the entire year are shown in Figures 31, 32, 33 and 34 respectively. The hydrographs are presented for the entire validation period (1999-2001). In the hydrographs, the green line, red line and blue line represent the flow forecasts (mean of 15 members) with NCEP meteorological forecasts, flow forecasts without NCEP forecasts and the observed Serpent River flows respectively. An overall examination of the hydrographs shows that simulated flow series using NCEP meteorological forecasts are better than that without using NCEP meteorological forecasts in the spring season. For 5 day ahead forecasting, peak flows in the spring 1999 are underestimated slightly when using NCEP meteorological forecasts and this underestimation is higher when NCEP meteorological forecasts are not used. It can be seen from Figure 31 that in the spring 2000, the peak flow is underestimated in both cases and in spring 2001 peak flow as well as flows in the rising and falling limb are captured very well when NCEP meteorological forecasts are used. In all the spring seasons, the agreement between the forecasted flows using NCEP meteorological forecasts and the observed flows in both the rising limb and the falling limb are much better than those in the case of flow forecasts without NCEP meteorological forecasts. This phenomenon can be also seen for other forecast ranges (8, 11 and 14 day ahead) as shown in Figures 32, 33 and 34.

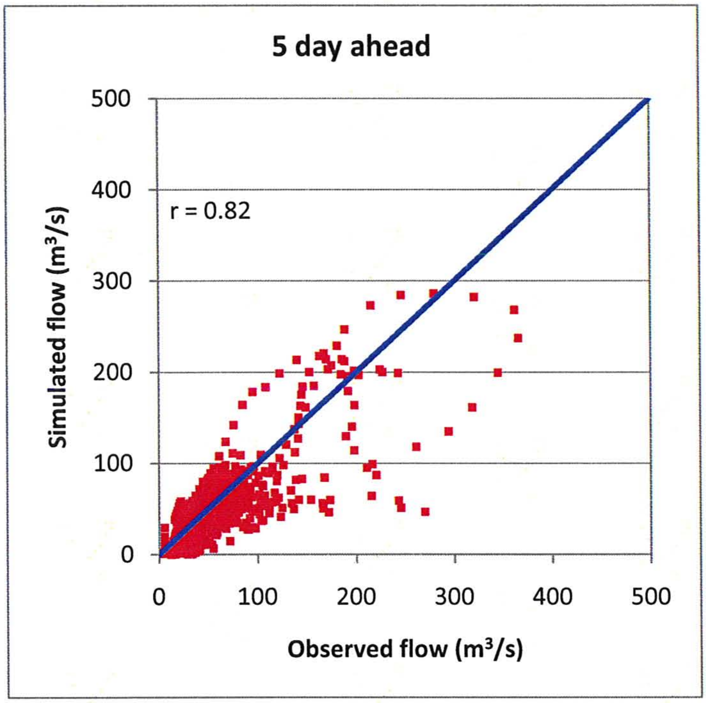

Figure 23: Scatter plot for Serpent River flow without meteorological forecasts (5day ahead)

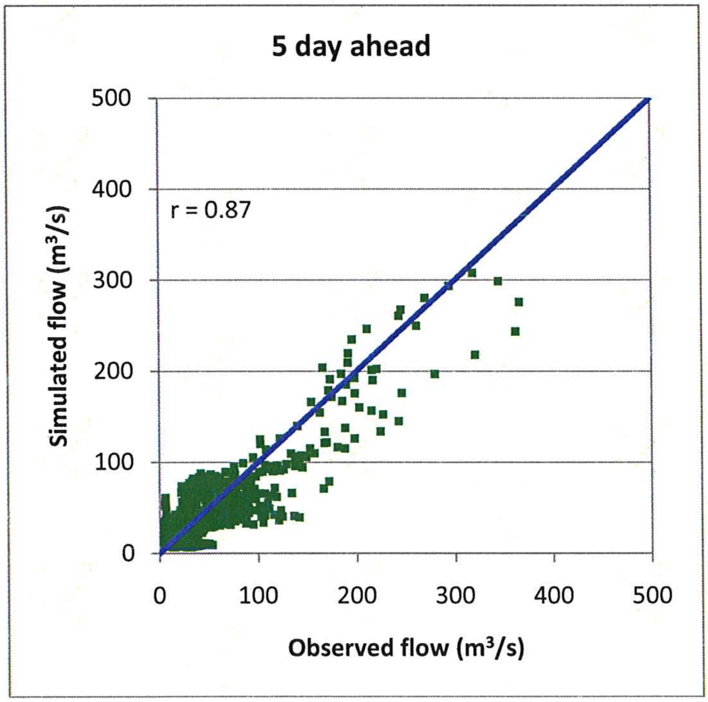

Figure 24: Scatter plot for Serpent River flow with meteorological forecasts (5 day ahead)

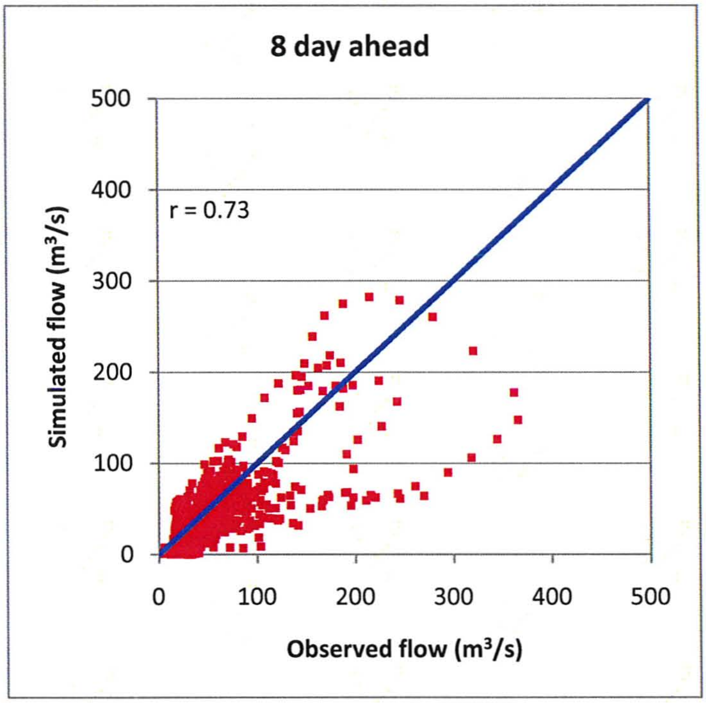

Figure 25: Scatter plot for Serpent River flow without meteorological forecasts (8 day ahead)

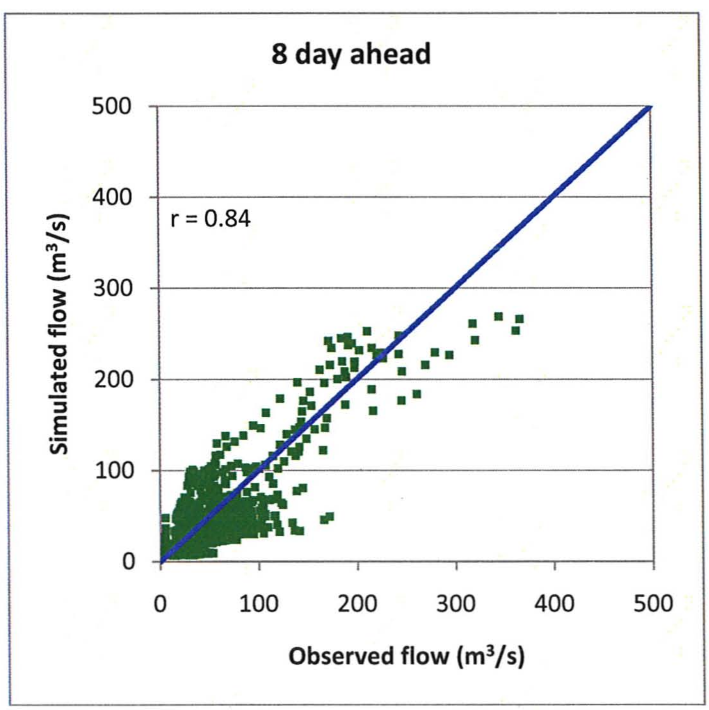

Figure 26: Scatter plot for Serpent River flow with meteorological forecasts (8 day ahead)

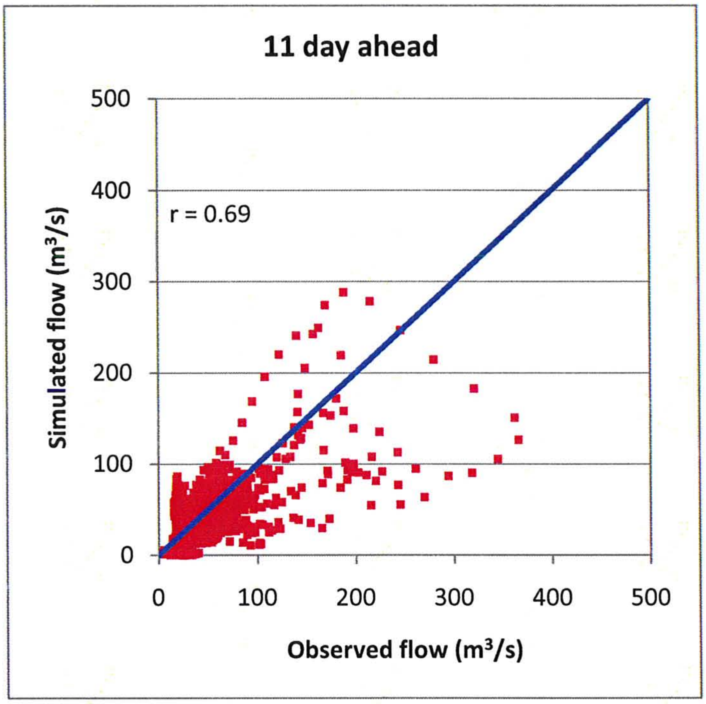

Figure 27: Scatter plot for Serpent River flow without meteorological forecasts (11day ahead)

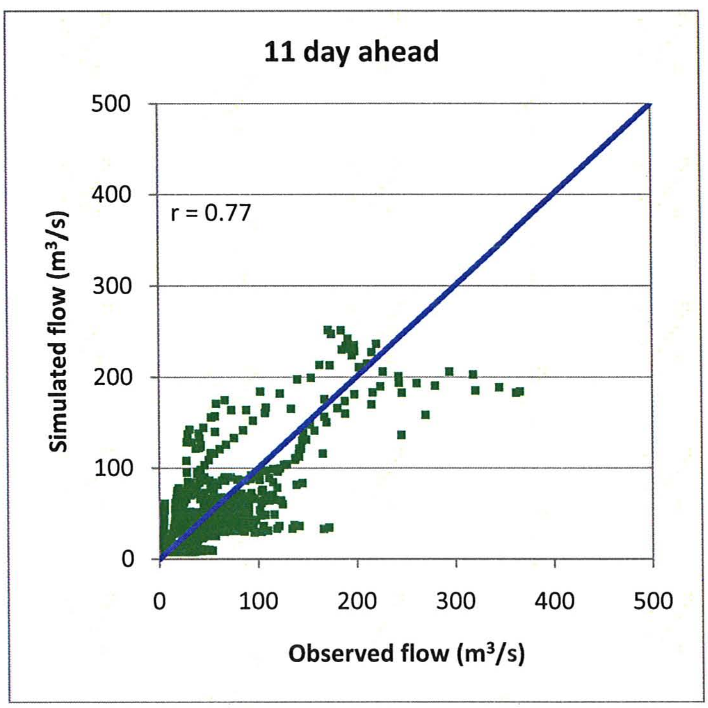

Figure 28: Scatter plot for Serpent River flow with meteorological forecasts (11 day ahead)

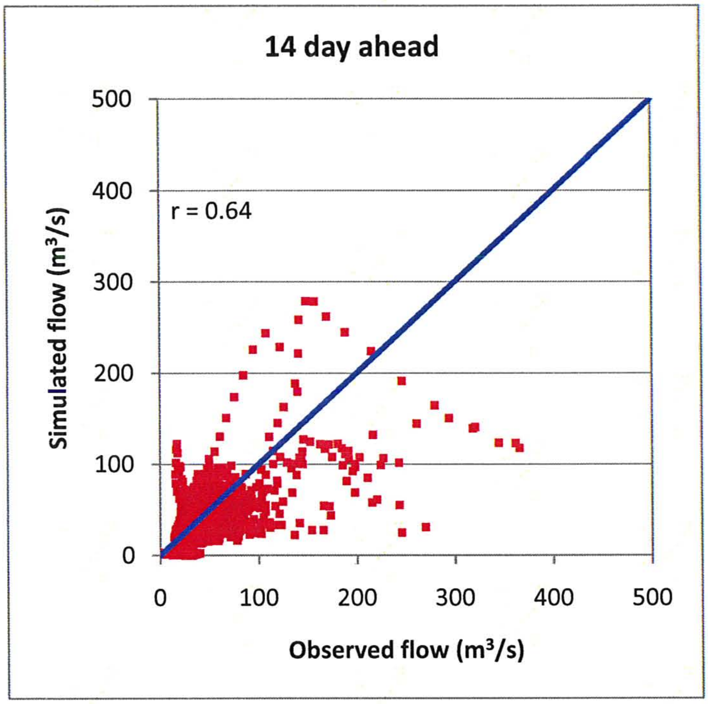

Figure 29: Scatter plot for Serpent River flow without meteorological forecasts (14day ahead)

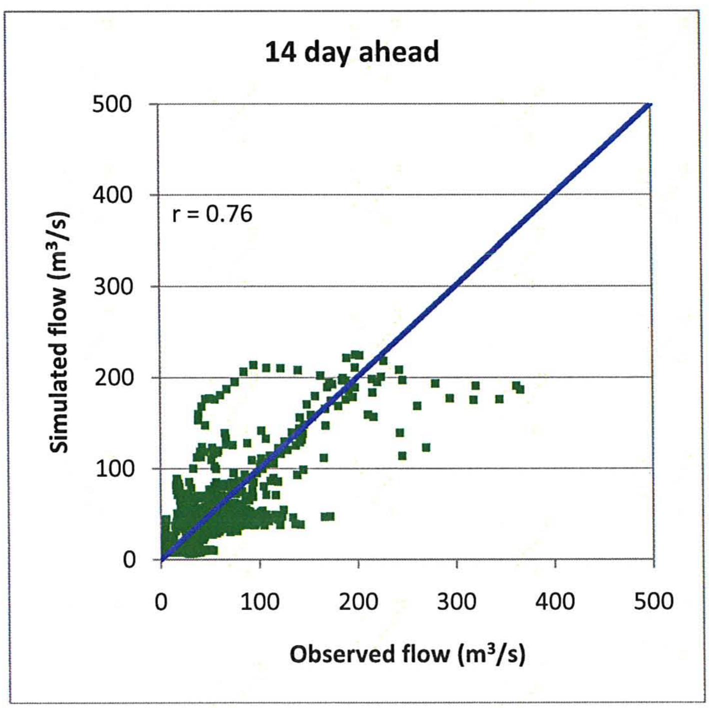

Figure 30: Scatter plot for Serpent River flow with meteorological forecasts (14 day ahead)

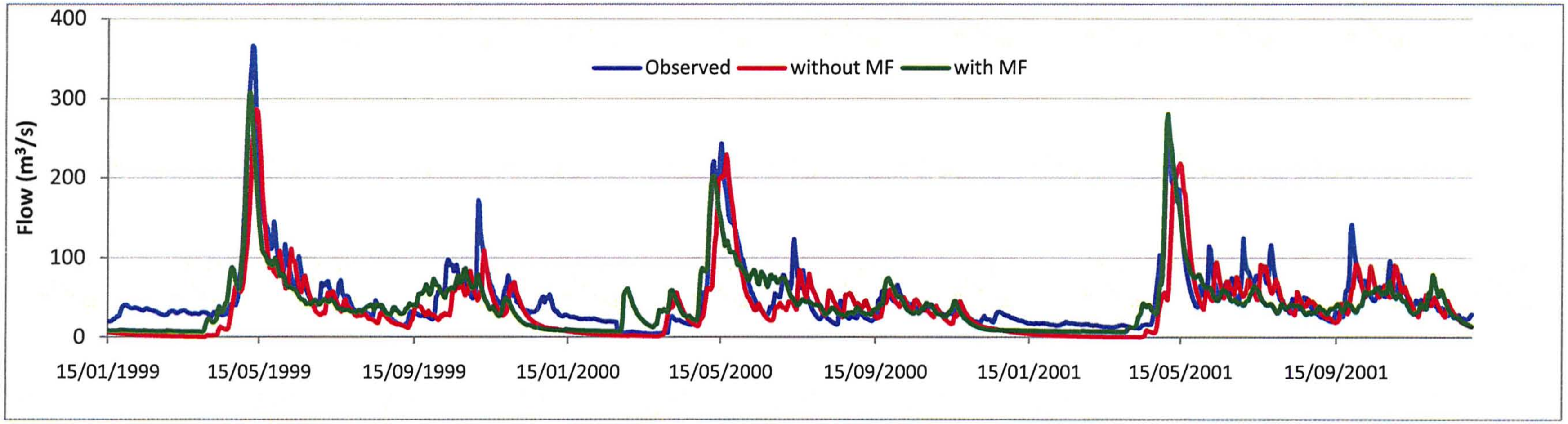

Figure 31: Hydrographs for 5 day ahead Serpent River flows for entire year

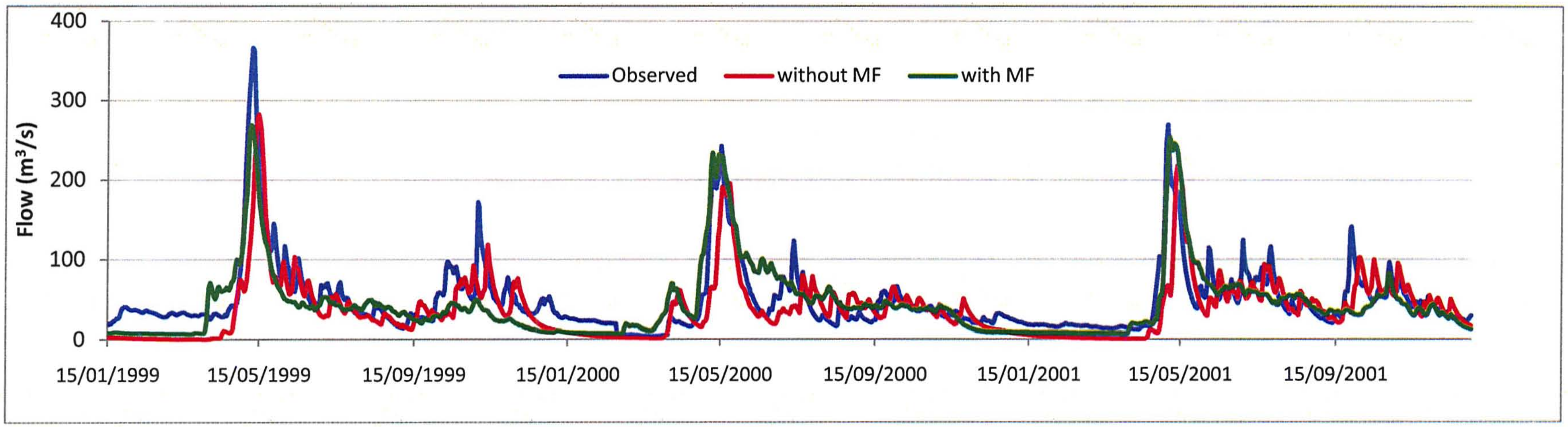

Figure 32: Hydrographs for 8 day ahead Serpent River flows for entire year

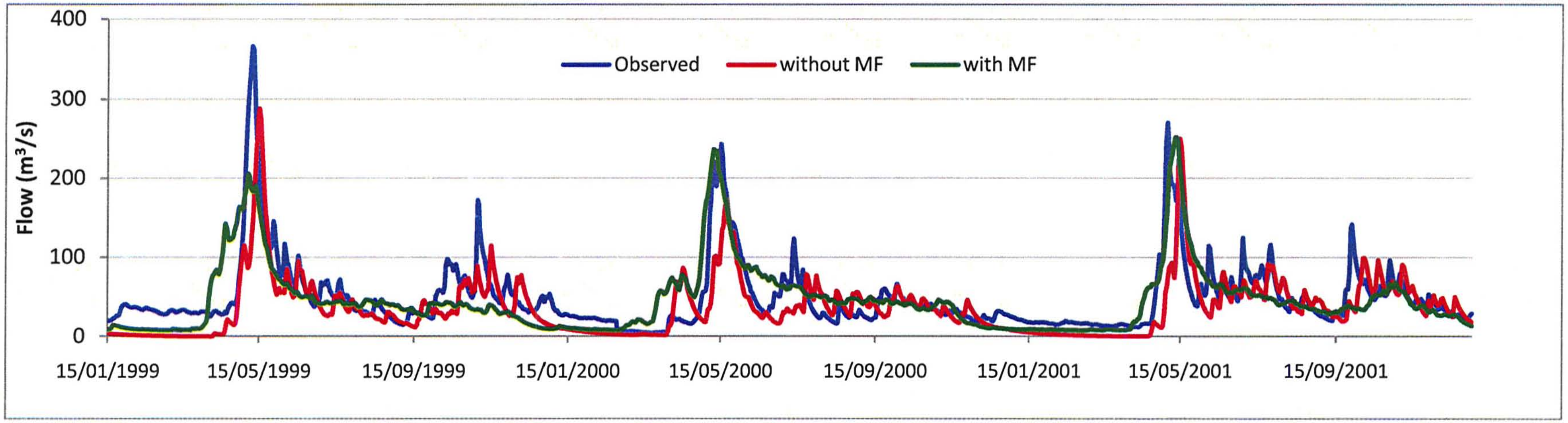

Figure 33: Hydrographs for 11 day ahead Serpent River flows for entire year

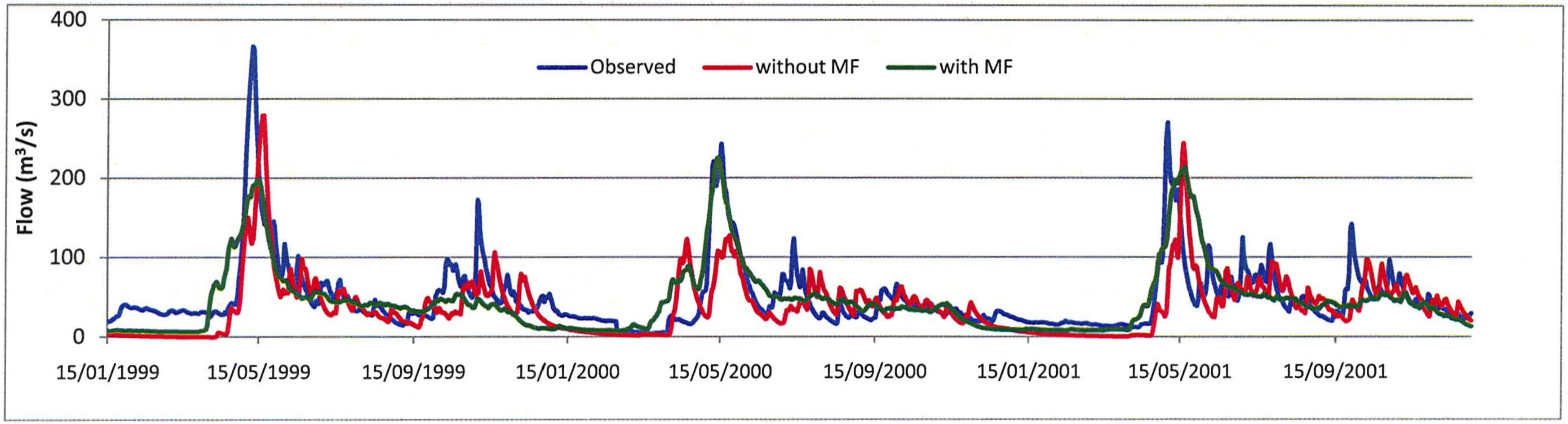

Figure 34: Hydrographs for 14 day ahead Serpent River flows for entire year

4.4 Model Performance Assessment with Probabilistic Measures

4.4.1 Reservoir Inflow Forecasting

The Brier skill score (BSS) and rank probability skill score (RPSS) as well as Brier score (BS) and rank probability score (RPS) were calculated for the entire year and four seasons - Spring (March to May), Summer (June to August), Autumn (September to November) and Winter (December to February). These scores were calculated for all data in the validation period (1997-2001). The calculated scores for the entire year, the spring, the summer, the autumn and the winter seasons are presented in Tables 13, 14, 15, 16 and 17 respectively. As mentioned in section 4.1, the calculated BS and RPSS represent the percentage improvement of streamflow forecasts (when NCEP meteorological forecasts are used) over the reference forecasts and here the reservoir inflow forecasts without NCEP meteorological are used as reference forecasts. In Table 13, both the calculated BSS and RPSS skill score show that there is a significant improvement (BSS on average 20.1% and RPSS 21.6%, calculated from the BSS and RPSS in Table 13) in the reservoir inflow forecasting when NCEP numerical meteorological forecasts variables (accumulated precipitation and mean temperature) are used in the HBV model. In Table 14, both the calculated BSS and RPSS values show that the improvement is higher (BSS on average 51.8% and RPSS 37%) in the spring season than that in the entire year. Table 13 also shows that for the entire year, the model performance deteriorates with the increase of forecast lead time as the BS and RPS values increase when the NCEP forecasts are not used, but there is no consistent deterioration when using NCEP meteorological forecasts. The reason may be the poor NCEP precipitation forecast with no significant change in variability after a forecast range 5 day ahead. The same phenomenon was found (as shown in Tables 15 and 16) for the summer and autumn season when the flow is dominated by rainfall. On the other hand, Table 14 presents the forecasts skill for the temperature influenced snowmelt dominated flows in the spring season which shows that in both cases without meteorological forecasts and with meteorological forecasts, the model performance deteriorates with the increase of forecast lead time as both the BS and RPS increase. It can be seen from Table 14 that the BS increased from 0.102 for 5 day ahead to 0.217 for 14 day ahead when NCEP meteorological forecast data are not used, it increased from 0.049 for 5 day ahead to 0.083 for 14 day ahead when meteorological data are used; the RPS increased from 1.152 for 5 day ahead to 1.496 for 14 day ahead when NCEP meteorological forecast data are not used, it increased from 0.761 for 5 day ahead to 0.973 for 14 day ahead when meteorological data are used. The spring flows are influenced predominantly by the temperature, thus show a consistent deterioration with the increase of forecast lead time when NCEP meteorological data are used. In Table 15 the calculated BSS and RPSS show that there is also a significant improvement in reservoir flow forecasting in the summer season when NCEP meteorological forecasts are used. However, a close examination of Tables 14 and 15 shows that when NCEP meteorological forecasts are used, the reservoir inflow forecasts in the spring season are more accurate than in the summer season (as example for 5 day ahead forecast, BS and RPS are 0.049 and 0.761 respectively in the spring season and those score are 0.245 and 1.58 respectively in the summer season). Similarly, an examination of Tables 14, 16 and 17 shows that when NCEP meteorological forecasts are used, the reservoir inflow forecasts in the spring season are more accurate than that in the autumn and winter seasons (for example for 5 day ahead forecast, BS and RPS are 0.049 and 0.761 respectively in the spring season and those score are 0.242 and 1.706 respectively in the autumn season and 0.246 and 2.498 in the winter season). Table 16 also shows that there is deterioration in forecasts quality in the autumn for some forecast ranges when NCEP meteorological forecasts are used. In the winter (as shown in Table 17) the improvement is much less than that in the spring season. Finally, it can be concluded that the quality of reservoir inflow forecasts in the spring season are more accurate than other seasons and the improvement of forecasts quality over the reference forecasts is significant in the spring season.

Table 13: Skill score for reservoir inflow forecasts from 5 to 14 day ahead for the entire year

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.164	1.198	0.139	0.975	15	19
6-day-ahead	0.171	1.229	0.154	1.070	10	13
7-day-ahead	0.187	1.282	0.166	1.088	11	15
8-day-ahead	0.183	1.271	0.163	1.037	11	18
9-day-ahead	0.198	1.329	0.160	1.042	20	22
10-day-ahead	0.203	1.354	0.177	1.095	13	19
11-day-ahead	0.210	1.369	0.161	1.043	24	24
12-day-ahead	0.216	1.410	0.156	1.019	28	28
13-day-ahead	0.233	1.458	0.159	1.034	32	29
14-day-ahead	0.256	1.503	0.161	1.060	37	29

Table 14: Skill score for reservoir inflow forecasts from 5 to 14 day ahead for spring season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.102	1.152	0.049	0.761	52	34
6-day-ahead	0.115	1.270	0.070	0.797	39	37
7-day-ahead	0.133	1.307	0.070	0.804	47	38
8-day-ahead	0.141	1.280	0.071	0.796	50	38
9-day-ahead	0.165	1.287	0.088	0.891	47	31
10-day-ahead	0.167	1.404	0.085	0.888	49	37
11-day-ahead	0.185	1.430	0.084	0.879	55	39
12-day-ahead	0.200	1.465	0.083	0.855	59	42
13-day-ahead	0.224	1.548	0.093	0.939	58	39
14-day-ahead	0.217	1.496	0.083	0.973	62	35

Table 15: Skill score for reservoir inflow forecasts from 5 to 14 day ahead for summer season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.328	2.083	0.245	1.580	25	24
6-day-ahead	0.324	2.078	0.219	1.583	32	24
7-day-ahead	0.333	2.172	0.226	1.706	32	21
8-day-ahead	0.335	2.204	0.217	1.645	35	25
9-day-ahead	0.354	2.350	0.213	1.586	40	32
10-day-ahead	0.346	2.374	0.245	1.705	29	28
11-day-ahead	0.348	2.441	0.222	1.604	36	34
12-day-ahead	0.370	2.557	0.202	1.512	45	41
13-day-ahead	0.389	2.652	0.217	1.570	44	41
14-day-ahead	0.441	2.780	0.210	1.552	52	44

Table 16: Skill score for reservoir inflow forecasts from 5 to 14 day ahead for autumn season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.211	1.774	0.242	1.706	-15	4
6-day-ahead	0.228	1.813	0.275	2.000	-21	-10
7-day-ahead	0.263	1.954	0.314	2.086	-19	-7
8-day-ahead	0.224	1.930	0.312	2.029	-39	-5
9-day-ahead	0.254	2.057	0.302	1.959	-19	5
10-day-ahead	0.276	2.109	0.320	2.060	-16	2
11-day-ahead	0.302	2.172	0.284	1.888	6	13
12-day-ahead	0.296	2.161	0.275	1.831	7	15
13-day-ahead	0.317	2.224	0.269	1.816	15	18
14-day-ahead	0.352	2.361	0.290	1.917	18	19

Table 17: Skill score for reservoir inflow forecasts from 5 to 14 day ahead for winter season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.270	2.613	0.246	2.498	9	4
6-day-ahead	0.298	2.754	0.250	2.571	16	7
7-day-ahead	0.285	2.776	0.237	2.442	17	12
8-day-ahead	0.304	2.754	0.220	2.236	28	19
9-day-ahead	0.311	2.887	0.231	2.221	26	23
10-day-ahead	0.315	2.983	0.226	2.282	28	23
11-day-ahead	0.274	2.843	0.222	2.162	19	24
12-day-ahead	0.278	2.898	0.239	2.224	14	23
13-day-ahead	0.276	2.748	0.229	2.136	17	22
14-day-ahead	0.263	2.715	0.233	2.286	11	16

4.4.2 Serpent River Flow Forecasting

The Brier skill score (BSS) and rank probability skill score (RPSS) as well as Brier score (BS) and rank probability score (RPS) were calculated for the entire year and four seasons - Spring (March to May), Summer (June to August), Autumn (September to November) and Winter (December to February). These scores were calculated for all data of the validation period (1999-2001) of the Serpent River flow forecasts. Tables 18, 19, 20, 21 and 22 present the calculated scores for the entire year, the spring, the summer, the autumn and the winter seasons respectively. As mentioned earlier, the calculated BS and RPSS represent the percentage improvement of streamflow forecasts (when NCEP meteorological forecasts are used) over the reference forecasts and here the reference forecasts are the Serpent River flow forecasts without NCEP meteorological forecasts. In Table 18, both the calculated BSS and RPSS skill score show that there is a significant improvement (BSS on average 20.7% and RPSS 17.7%, calculated from the BSS and RPSS in Table 18) in the Serpent River flow forecasting when NCEP numerical meteorological forecasts variables (accumulated precipitation and mean temperature) are used. In Table 19, both the calculated BSS and RPSS values show that the improvement is higher (BSS on average 34.2% and RPSS 28.2%) in the spring season than in the entire year. Table 18 also shows that the model performance deteriorates with the increase of forecast lead time as the BS and RPS values increase when the NCEP forecasts are not used, but there is no consistent deterioration when using NCEP meteorological forecasts. The reason might be the poor NCEP precipitation forecast with no significant change in variability after forecast range 5 day ahead. The same phenomenon is found (as shown in Tables 20 and 19) for the summer and autumn season when the flow is dominated by rainfall. On the other hand, Table 19 presents the forecasts skill for the temperature-influenced, snowmelt-dominated flows in the spring for both cases without meteorological forecasts and with meteorological forecasts, which shows in both cases the model performance deteriorated with the increase of forecast lead time as the BS increase. It can be seen from Table 19 that the BS increased from 0.109

for 5 day ahead to 0.217 for 14 day ahead when NCEP meteorological forecast data are not used, it increased from 0.063 for 5 day ahead to 0.114 for 14 day ahead when meteorological data are used. The spring flows are influenced predominantly by the temperature, and shows a consistent deterioration with the increase of forecast lead time when NCEP meteorological data are used. In Table 20 the calculated BSS and RPSS show that there is also a significant improvement in the Serpent River flow forecasting in the summer season when NCEP meteorological forecasts are used. However, an examination of Tables 19 and 20 shows that when NCEP meteorological forecasts are used, the Serpent River flow forecasts in the spring season are much better than that in the summer season (as example for 5 day ahead forecast, BS and RPS are 0.063 and 1.238 respectively in the spring season and those score are 0.226 and 1.577 respectively in the summer season). Similarly, an examination of Tables 19, 21 and 22 shows that when NCEP meteorological forecasts are used, the Serpent River flow forecasts in the spring season are much better than in the autumn and winter seasons (for example for 5 day ahead forecast, BS and RPS are 0.063 and 1.238 respectively in the spring season and those score are 0.157 and 1.611 respectively in the autumn season and 0.36 and 3.764 in the winter season). Similar to the reservoir flow forecasts, Table 21 also shows that there is a deterioration in forecast quality in the autumn for some forecast ranges when NCEP meteorological forecasts are used. It can be seen from Table 22 that the improvement over the reference forecasts is insignificant in the winter season. Finally, similar conclusion to the reservoir flow forecasts can be drawn for the Serpent River flow forecasting that the quality of forecasts in the spring season are much better than other seasons, and the improvement of forecast quality over the reference forecasts is significant in the spring season.

Table 18: Skill score for Serpent River flow forecasts from 5 to 14 day ahead for the entire year

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.134	1.598	0.120	1.398	11	12
6-day-ahead	0.143	1.648	0.128	1.483	11	10
7-day-ahead	0.167	1.759	0.145	1.528	13	13
8-day-ahead	0.172	1.776	0.158	1.532	8	14
9-day-ahead	0.182	1.814	0.152	1.486	16	18
10-day-ahead	0.192	1.891	0.152	1.568	21	17
11-day-ahead	0.204	1.924	0.153	1.521	25	21
12-day-ahead	0.218	1.963	0.156	1.511	28	23
13-day-ahead	0.229	1.959	0.146	1.465	36	25
14-day-ahead	0.233	1.974	0.144	1.503	38	24

Table 19: Skill score for Serpent River flow forecasts from 5 to 14 day ahead for spring season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.109	1.583	0.063	1.238	42	22
6-day-ahead	0.120	1.663	0.095	1.318	21	21
7-day-ahead	0.149	1.717	0.077	1.156	48	33
8-day-ahead	0.167	1.743	0.096	1.093	43	37
9-day-ahead	0.192	1.786	0.113	1.175	41	34
10-day-ahead	0.149	1.815	0.102	1.301	31	28
11-day-ahead	0.156	1.790	0.129	1.276	17	29
12-day-ahead	0.178	1.725	0.132	1.288	26	25
13-day-ahead	0.192	1.775	0.143	1.346	26	24
14-day-ahead	0.217	1.772	0.114	1.260	47	29

Table 20: Skill score for Serpent River flow forecasts from 5 to 14 day ahead for summer season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.250	1.975	0.226	1.577	10	20
6-day-ahead	0.257	1.996	0.192	1.451	25	27
7-day-ahead	0.308	2.279	0.224	1.746	27	23
8-day-ahead	0.286	2.120	0.247	1.756	14	17
9-day-ahead	0.293	2.152	0.200	1.412	32	34
10-day-ahead	0.312	2.217	0.211	1.561	32	30
11-day-ahead	0.322	2.322	0.195	1.535	40	34
12-day-ahead	0.333	2.402	0.219	1.535	34	36
13-day-ahead	0.360	2.412	0.206	1.494	43	38
14-day-ahead	0.381	2.442	0.223	1.486	41	39

Table 21: Skill score for Serpent River flow forecasts from 5 to 14 day ahead for autumn season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.178	1.681	0.157	1.611	12	4
6-day-ahead	0.192	1.815	0.207	1.633	-8	10
7-day-ahead	0.225	1.902	0.241	1.785	-7	6
8-day-ahead	0.236	1.986	0.274	1.859	-16	6
9-day-ahead	0.254	2.033	0.263	1.840	-4	9
10-day-ahead	0.257	2.091	0.259	1.875	-1	10
11-day-ahead	0.275	2.236	0.246	1.683	11	25
12-day-ahead	0.290	2.384	0.231	1.608	20	33
13-day-ahead	0.315	2.424	0.211	1.494	33	38
14-day-ahead	0.319	2.493	0.232	1.666	27	33

Table 22: Skill score for Serpent River flow forecasts from 5 to 14 day ahead for winter season

Forecasting lags	**Without meteorological forecasts**		**With meteorological forecasts**		**Skill Score (%)**	
	BS	RPS	BS	RPS	BSS	RPSS
5-day-ahead	0.370	3.917	0.360	3.764	3	4
6-day-ahead	0.370	3.924	0.364	3.770	2	4
7-day-ahead	0.370	3.880	0.354	3.728	4	4
8-day-ahead	0.391	3.924	0.375	3.870	4	1
9-day-ahead	0.399	3.942	0.371	3.779	7	4
10-day-ahead	0.391	3.960	0.387	3.875	1	2
11-day-ahead	0.391	3.964	0.374	3.743	4	6
12-day-ahead	0.395	3.953	0.379	3.749	4	5
13-day-ahead	0.402	3.993	0.353	3.510	12	12
14-day-ahead	0.409	4.000	0.379	3.789	7	5

4.5 Visual Inspection of Ensembles

In this section, hydrographs represent all 15 member ensembles, ensemble mean, observed flow and upper limit and the lower limit of ensembles for 5 day ahead forecasts. The hydrographs from the entire year of flows and spring flows of one year (2001) for the reservoir inflows are presented in Figures 35 and 36 respectively. The hydrographs from the entire year flows and spring flows of one year (2001) for the Serpent River flows are presented in Figures 37 and 38 respectively. The objective of this visual inspection of ensembles is to highlight the fact that the performance of the models is improved in the spring season more than other seasons and to reveal the reason behind this better performance. Figure 35 shows a very good agreement between the pattern of the forecasted ensemble reservoir inflows and observed inflows in the spring season both in capturing the flows in the rising limb and the falling limb as well as in the peak flow. No such agreement is found in the reservoir flow forecasting for other seasons. Figure 35 also revealed that all the peak flows in the summer months lie out of the upper limit (indicating the 95% confidence interval); peak flow of the summer month of July is highly underestimated. Although the ensemble spread in the autumn is little smaller than that in the summer, all the peak flows in the autumn season are highly underestimated. The reservoir flows in the winter month of December lies between the limits, but the ensemble spread is higher than other seasons; moreover the flows in the other winter months are always underestimated. In can also be seen from all the hydrographs that there is a significant overestimation of flow in the beginning of the spring and this overestimation is due to early snowmelt. The reason for this early snowmelt is the overestimation and fluctuation of NCEP forecasted temperature (shown in Figures 4, 5, 6 and 7) in the winter. As mentioned earlier, the snowmelt runoff is responsible for high flows in the spring season as well as up to 40% of the annual flow volume in the study region. It is also shown in Section 2.2.3 that NCEP temperature forecasts exhibit high skill over the year. Figure 35 also shows that the flow in the spring season started to increase at the beginning of April. Here the flow in the spring season is dominated by snowmelt, which is influenced predominantly by the temperature variation. This also confirms the findings of Clark and Hay (2004) that in the snowmelt dominated River basins when streamflows are strongly forced by temperature; the high skill in prediction of temperature translates into high skill in prediction of streamflow. On the other hand precipitation forecasts are poor compared to temperature forecasts as shown in Section 2.2.3 and Appendix. This is why the temperature-influenced snowmelt dominated spring flow forecasts are much better than rainfall-dominated flows in the summer and the autumn.

Similar to reservoir inflow forecasting, Figures 37 and 38 show very good agreement between the pattern of the forecasted ensemble Serpent River flows and observed flows in the spring season both in capturing the flows in the rising limb and the falling limb as well as peak flow. No such agreement is found in the Serpent River flow forecasting for other seasons. Figure 37 also revealed that all the peak flows in the summer months lie out of the upper limit indicating the 95% confidence interval; peak flows of the summer months are highly underestimated, showing the poor performance of the model using NCEP data for flow forecasting in summer. The peak flows in the autumn season are also highly underestimated. The Serpent River flows in the winter month of December lies between the limits, but the ensemble spread is higher than other seasons; moreover the flows in the other winter months are underestimated. The reason for better flow forecasts in the spring season for the Serpent River is same as forecasts for reservoir inflows.

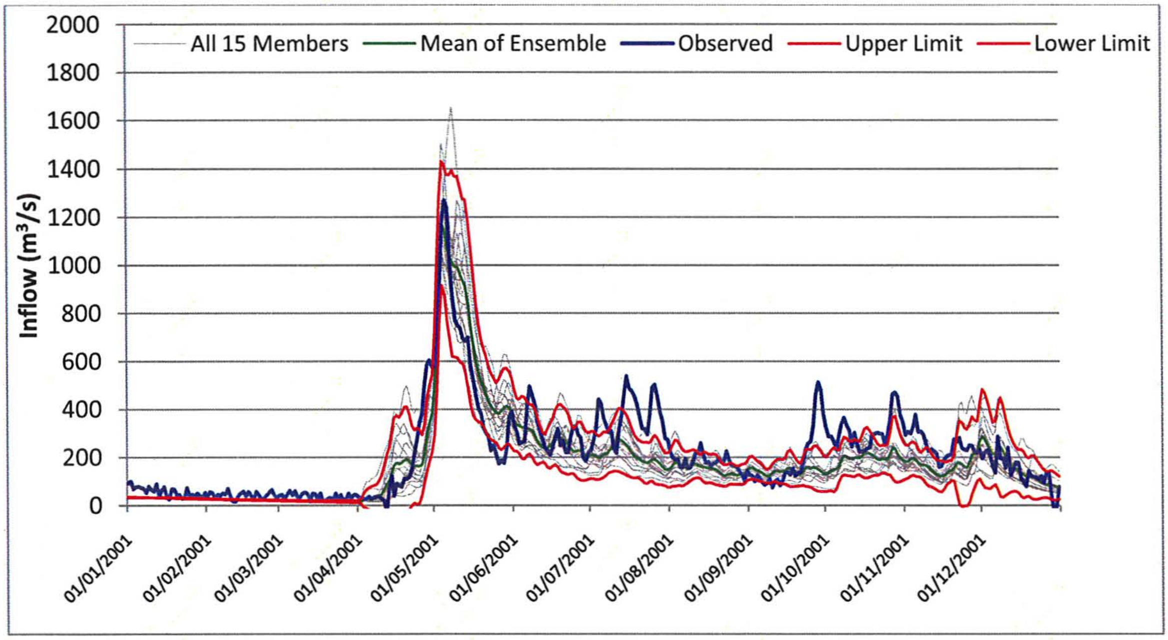

Figure 35: Ensemble reservoir inflow simulation for 5 day ahead for one year (limits show 95% confidence interval)

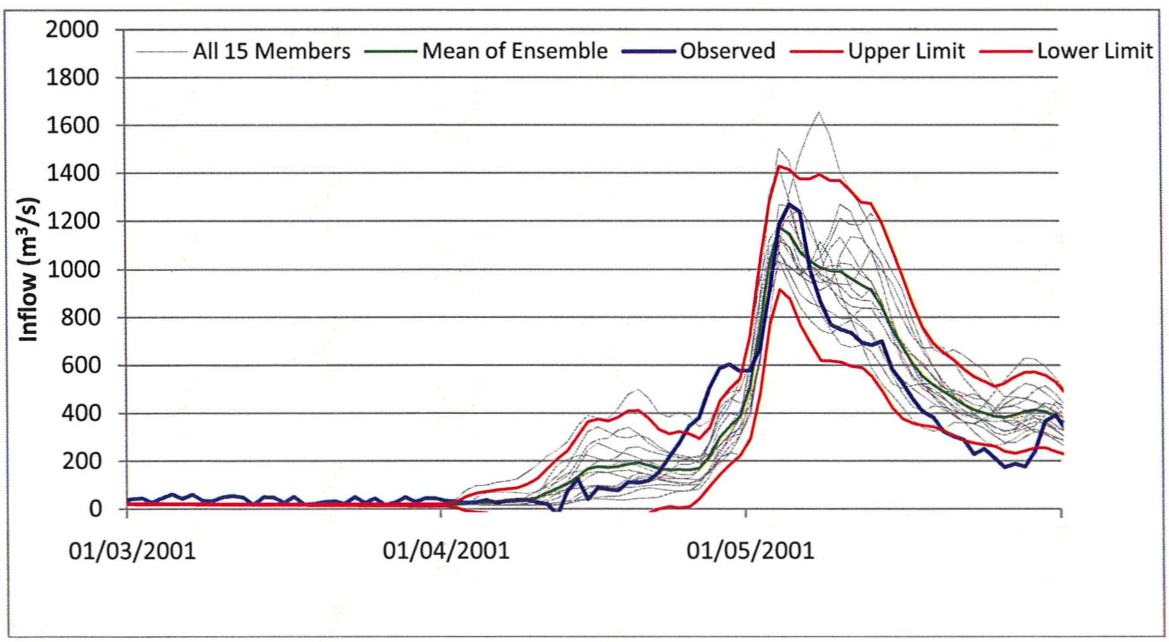

Figure 36: Ensemble reservoir inflow simulation for 5 day ahead for spring season (limits show 95% confidence interval)

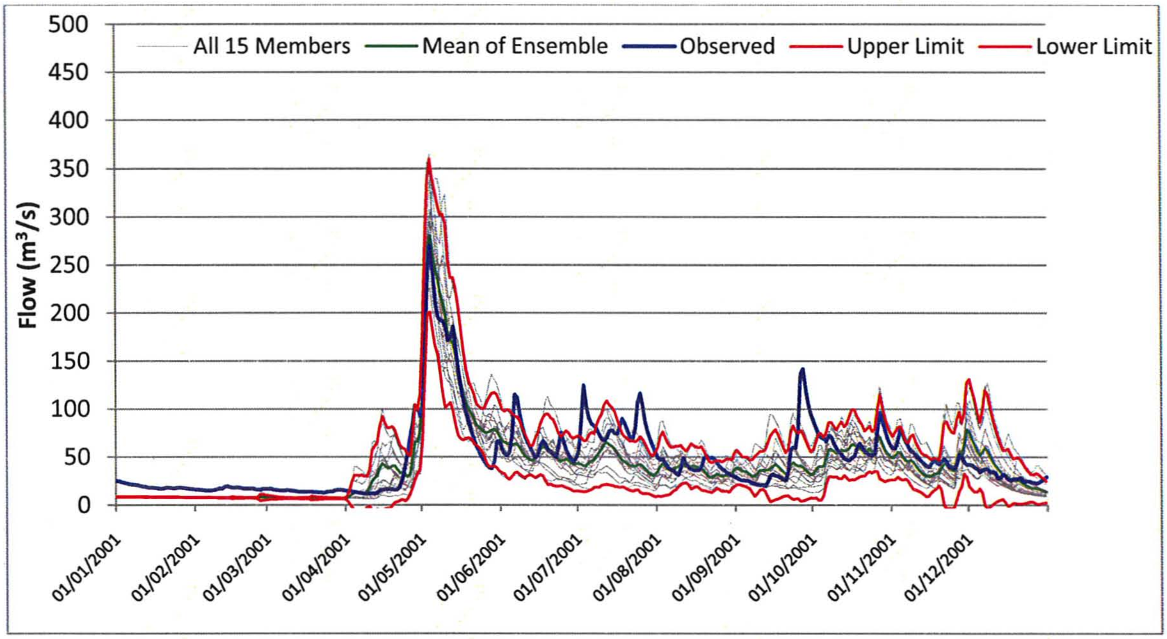

Figure 37: Ensemble Serpent River flow simulation for 5 day ahead for one year (limits show 95% confidence interval)

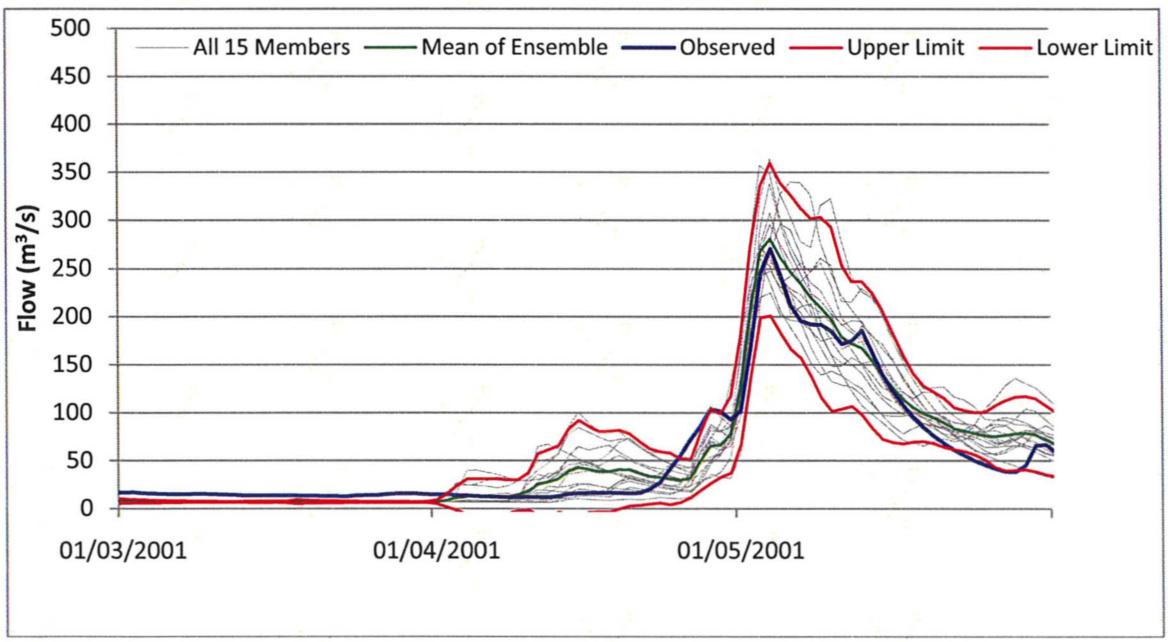

Figure 38: Ensemble Serpent River flow simulation for 5 day ahead for spring season (limits show 95% confidence interval)

Chapter 5: Conclusions and Recommendations

The study results revealed that there is a significant improvement in both reservoir inflow and the Serpent River flow forecasting when raw NCEP meteorological forecasts are used. A number of verification measures were used for assessing the ensemble hydrologic forecast skill.

The quality of hydrologic forecasts was initially assessed for the entire year. The results show that for both reservoir inflow and the Serpent River flow forecasts the model performance deteriorate with the increase of forecast lead time as the RMSE increases and r and R^2 decrease with the increase of lead time. More accurate forecast always indicates lower RMSE and higher r and R^2. When NCEP meteorological forecasts are used, the model performs well up to 8 day ahead and 7 day ahead for reservoir inflow forecasting and the Serpent River flow forecasting respectively. In the case where observed data are used instead of NCEP meteorological forecasts, the model performance is poor even on 5 day ahead for both reservoir inflow forecasting and the Serpent River flow forecasting. The increase of correlation coefficient (r) in all forecast ranges is significant, which shows a clear improvement. When NCEP meteorological forecasts are used, root mean square error (RMSE) also decreases on average 19.7% for reservoir inflow forecasting and 19.2% for the Serpent River flow forecasting. When comparing the models by scatter plots, it can be found that the data are more clustered along the 45 degree line in the case where NCEP meteorological forecasts are used for all forecast ranges for both reservoir inflows forecasting and the Serpent River flow forecasting. The hydrographs also show the improvement in both reservoir inflow forecasting and the Serpent River flow forecasting. The calculated probabilistic verification measures also show that there is about on average a 20% improvement in both reservoir inflow forecasting and the Serpent River flow forecasting for entire year flows as the average BSS and RPSS values are 20.1% and 21.6% respectively for reservoir inflow forecasting and 20.7% and 17.7% respectively for the Serpent River flow forecasting.

The hydrologic model forecasts are also investigated on seasonal basis and all the verification measure show that the forecast quality is much better in the spring season. When NCEP meteorological forecasts are used, the model performs well up to 13 day ahead for reservoir inflows forecasting. The calculated RMSE for spring flows also decreases on average by 32.7 %, which is 19.7% for entire year flows for reservoir inflow forecasting. The hydrographs of ensemble mean show that in all the spring seasons, the agreement between the forecasted reservoir inflows using NCEP meteorological forecasts and the observed inflows in both rising limb and the falling limb of the hydrograph are much better than that in case of inflow forecasts without NCEP meteorological forecasts. The calculated RPSS values show that the ensemble reservoir inflow forecasting improves on average by 37%, which is 21.6% for the entire year flows. The ensemble plots also show that the model give better forecasts in the spring season than other seasons as the reservoir inflows in the snow melting months are well captured in both rising limb and the falling limb of hydrographs as well as peak flows. The model performance in the Serpent River flow forecasting is quit close to that of the reservoir inflow forecasting. The reason of this improvement in the spring season is that the spring flow in the study basins is dominated by the snowmelt, which is influenced predominantly by the temperature variations; accurate temperature forecasts by the National Centers for Environmental Prediction (NCEP) Global Forecast System up to 14 day ahead translates into high skill in predictions of streamflow in the spring season. Because the peak flows in spring season are particularly important for scheduling reservoir operation, and responsible for flooding in the study region, therefore, spring peak flow forecasting using the raw NCEP meteorological forecasts could be a good alternative approach for obtaining more accurate flow forecasts up to 8 day-ahead.

Although the forecasting models show significant improvement in the spring flow forecasting, further investigation is required to improve the flow forecasting in the periods when the flows are dominated by precipitation. Further improvement in forecasting may be achieved by incorporating downscaled ensemble NCEP meteorological forecasts into hydrologic model and using another hydrologic model other than HBV.

References

Abbott, M., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986a. An introduction to the European Hydrological System – System Hydrologique Europeen, SHE; 1. History and philosophy of a physically based distributed modeling system, Journal of Hydrology, 87, 45-59.

Abbott, M., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986b. An introduction to the European Hydrological System – System Hydrologique Europeen, SHE; 1. Structure of a physically based, distributed modeling , Journal of Hydrology, 87, 79-102.

Bartholmes, J., Todini, E., 2005. Coupling meteorological and hydrological models for flood forecasting, Hydrology and Earth System Sciences, 9(4), 333-346.

Beck, M.B., 1987. Water quality modeling – a review of the analysis of uncertainty, Water Resources Res., 23(8), 1393-1442.

Bergstrom, S., 1991. Principles and confidence in hydrological modelling, Nordic Hydrology, 22, 123-136.

Bergstrom, S., Lindstrom, G., Petterson, A., 2002. Multi-variable parameter estimation to increase confidence in hydrological modeling, Hydrological Processes, 16, 413-421.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Science Bulletin ,24 (1), 43-69.

Boucher, M., Perreault, L., Anctil, F., 2009. Tools for assessment of hydrological ensemble forecasts obtained by neural networks, Journal of Hydroinformatics, 297-307.

Brier, G.W., 1950. Verification of forecasts expressed in terms of probabilities, Monthly Weather Review, 78, 1-3.

Brown, J.D., Demarge, J., Seo, D., Liu, Y., 2010. The ensemble verification system(EVS): a software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations, Environmental Modelling and Software, 25, 854-872.

Carpenter, T.M., Georgakakos, K.P., 2006. Discretization scale dependencies of the ensemble flow range versus catchment area relationship in distributed hydrologic modeling, Journal of Hydrology, 328, 242-257.

Clark, M.P., Hay, L.E., 2004. Use of medium-range numerical weather prediction model output to produce forecasts of streamflow, Journal of Hydrometeorology, 5, 15-32.

Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: A review, Journal of Hydrology, 375, 613-626.

Coulibaly, P., Anctil, F., Bobee, B., 2001. Multivariate reservoir inflow forecasting using temporal neural networks, Journal of Hydrologic Engineering, 6, 367-376.

Coulibaly, P., 2003. Impact of meteorological predictions on real-time spring flow forecasting, Hydrological Processes, 17, 3791-3801.

Coulibaly, P., Anctil, F., Bobee, B., 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach, Journal of Hydrology, 230, 244-257.

Cranmer A.J., Kouwen N., Mousavi S.F. 2001. Proving WATFLOOD: modeling the nonlinearities of hydrologic response to storm intensities, Canadian Journal of Civil Engineering, 28, 837-855.

Dietrich, J., Trepte, S., Wang, Y., Schumann, A., Vob, F., Hesser, F.B., Denhard, M., 2008. Combination of different types of ensembles for the adaptive simulation of probabilistic flood forecasts: hindcasts for the Mulde 2002 extreme event, Nonlinear Processes Geophysics, 15, 275-286.

Feldman, A.D., Ely, P.B., Goldman, D.M., 1981. The new HEC-1 Flood hydrograph package, Technical paper, TP-82, US Army Corps of Engineers, Institute of Water Resources, Hydrologic Engineering Center, Davis, CA, USA.

Feyen, L., Vázquez, R., Christiaens, K., Sels, O., Feyen, J., 2000). Application of a distributed physically-based hydrological model to a medium size catchment, Hydrology and Earth System Sciences, 4(1), 47-63.

Franz, K.J., Hartmann, H.C., Soroshian, S., Bales, R., 2003. Verification of National Weather Service ensemble streamflow predictions for water supply forecasting in the Colorado river basin, Journal of Hydrometeorology, 4, 1105-1118.

Gan, T.Y., Dlamini, E.M., Bifu, G.F., 1996. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling, Journal of Hydrology, 192, 81-103.

Gangopadhyay, S., Clark, M., Brandon, D., Wener, K., Rajagopalan, B., 2004. Effects of spatial and temporal aggregation on the accuracy of statistically downscaled precipitation estimates in the upper Colorado river basin, Journal of Hydrometeorology, 5, 1192-1206.

Govindaraju, R.S., 2000. Artificial neural networks in hydrology. II:Hydrological applications, by the ASCE task committee on application of artificial neural networks in hydrology, Journal of Hydrological Engineering, 5, 124-137.

Hagedorn, R.T., Hamill, T.M., Whitaker, J.S., 2007. Probabilistic forecast calibration using ECMWF and GFS ensemble forecasts. Part I: 2-meter temperature, Monthly Weather Review, 136, 2608-2619.

Half, AH., Half, H.M., and Azmoodeh, M., 1993. Predicting runoff from rainfall using neural networks, Proceedings, Engineering Hydrology, American Society of Civil Engineers, Ney York, 760-765.

Hamill, T.M., Whitaker, J.S., 2007. Ensemble calibration of 500 hPa geopotential and 850 hPa and 2-meter temperatures using reforecasts, Monthly Weather Review, 135, 3273-3280.

Hamill, T.M., Whitaker, J.S., 2006. Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application, Monthly Weather Review, 134, 3209-3229.

Hamill, T. M., Whitaker, J. S., Wei, X., 2004. Ensemble re-forecasting: improving medium-range forecast skill using retrospective forecasts, Monthly Weather Review, 132, 1434-1447.

Hartmann, H.C., Pagano, T.C., Bales, R., Sorooshian, S., 2002. Confidence builders: Evaluating seasonal climate forecasts from user perspectives, Bull. American Meteorological Society, 83,683-698.

Hogue, T.S., Soroshian, S., Gupta, H., 2000. A multistep automatic calibration for river forecasting models, Journal of Hydrometeorology, 1, 524-542.

Jasper, K., Gurtz, J., Lang, H., 2002. Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model, Journal of Hydrology, 267, 40-52.

Johnson, M.S., Coon, W.F., Mehta, V.K., Steenhuis, T.S., Brooks, E.S., Boll, J., 2003. Application of two hydrologic models with different runoff mechanisms to a hillslope dominated watershed in the northern US: a comparison of HSPF and SMR, Journal of Hydrology, 284, 57-76.

Karunanithi, N., Grenney, W.J., Whitley, D., Bovee, K., 1994. Neural network for river flow prediction, Journal of Computing in Civil Engineering, 8(2), 201-220.

Khan, M.S., Coulibaly, P., 2006. Bayesian neural network for rainfall-runoff modeling, Water Resources Research, 42, W07409.

Kuligowski, R.J., Barros, A.P., 1998. Experiments in short-term precipitation forecasting using artificial neural networks, Monthly Weather Review, 126(2), 470-482.

Kouwen, N., 2010. WATFLOOD^{TM} / WATROUTE Hydrological Model Routing and Flow Forecasting System, Department of Civil Engineering, University of Waterloo,Waterloo, Ontario, Canada.

Kouwen, N., Soulis, E.D., Pietroniro, A., Donald, J., Harrington, R.A., 1993. Group response units for distributed hydrologic modeling, Journal of Water Resources Planning and Management, 119 (3), 289-305.

Liden R, Harlin J. 2000. Analysis of conceptual rainfall–runoff modelling performance in different climates. Journal of Hydrology 238, 231–247.

Lindstrom, G., Johansson, B., Magnus, P., Gardelin, M., Bergstrom, S., 1997. Development and test of the distributed HBV-96 hydrological model, Journal of Hydrology, 201, 272-288.

Liu, X., 2007. Downscaling Meteorological Predictions for Short-term Hydrologic Forecasting, M.A.Sc Thesis, Department of Civil Engineering, McMaster University, Hamilton, Ontario, Canada.

Mullen, S.L., Buizza, R., 2000. Quantitative precipitation forecasts over the United States by the ECMWF ensemble prediction system, Monthly Weather Review, 129, 638-663.

Murphy, A.H., Brown, B.G., Chen, Y., 1989. Diagnostic verification of temperature forecasts, Weather Forecasting, 4, 485-501.

Murphy, A.H., 1971. A note of the ranked probability score, Journal of Applied Meteorology, 10, 155-156.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I – a discussion of Principals, Journal of Hydrology, 10, 282-290.

Quick, M.C., Pipes, A., 1977. UBC watershed model. Hydrology Science Bull., XXI (1/3), 285-295.

Refsgaard, J.C., 1997. Parameterization, calibration and validation of distributed hydrological models, Journal of Hydrology, 198, 69-97.

Regimbeau, F.R., Habets, F., Martin, E., Noilhan, J., 2007. Ensemble streamflow forecasts over France, ECMWF Newsletter, 111, 21-27.

Renner, M., Werner, M.G.F., Rademacher, S., Sprokkereef, E., 2009. Verification of ensemble flow forecasts for the river Rhine, Journal of Hydrology, 376,463-475.

Roulin, E., 2007. Skill and relative economic value of medium-range hydrological ensemble predictions, Hydrology and Earth System Sciences, 11, 725-737.

Roulin, E., Vannitsem, S., 2005. Skill of medium-range hydrological ensemble predictions, Journal of Hydrometeorology, 6, 729-744.

Seibert, J., 1997. Estimation of parameter uncertainty in the HBV model. Nordic Hydrology 28 (4/5), 247-2627.

Sharma, M., 2009. Comparison of Downscaled RCM and GCM data for hydrologic impact assessment, M.A.Sc Thesis, Dept. of Civil Engineering, McMaster University, Hamilton, Ontario, Canada.

Singh, V.P., Woolhiser, D.A., 2002. Mathematical modeling of watershed hydrology. Journal of Hydrological Engineering, 7, 270-292.

Sivakumar, B., Jawawardena, A.W., Fernando, T.M.K.G., 2002. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches, Journal of Hydrology, 265, 225-245.

SMHI, 2006. Integrated Hydrological Modelling System Manual, Version 5.10, Swedish Meteorological and Hydrological Institute, Sweden.

Sorman, A.A., Sensoy,A., Tekeli, A.E., Sorman, A.U., Akyurek, Z., 2009. Modelling and forecasting snowmelt runoff process using the HBV model in the eastern part of Turkey, Hydrological Processes, 23, 1031-1040.

Theis, S., Hense, A., and Damrath, U., 2005. Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. Meteorol. Appl., 12, 257-268.

Todini, E., 1996. The ARNO rainfall-runoff model, Journal of Hydrology, 175, 339-382.

Toth, Z., Kalnay, E., 1997. Ensemble forecasting at NCEP and the breeding method, Monthly Weather Review, 125, 3297-3319

Velàzquez, J.A., Petit, T., Lavoie, A., Boucher, M.A., Turcotte, R., Fortin, V., Anctil, F., 2009. An evaluation of the Canadian global meteorological ensemble prediction system for short-term hydrological forecasting, Hydrology and Earth System Science, 13, 2221-2231.

Watson, J.S., Colucci, S., 2002. Evaluation of ensemble predictions of blocking in the NCEP global spectral model, Monthly Weather Review, 130, 3008-3021.

Weigel, A.P., Liniger, M.A., Appenzeller, C., 2006. Notes and Correspondence: Generalization of the discrete brier and ranked probability skill scores for weighted multimodel ensemble forecasts, Monthly Weather Review, 135, 2778-2785.

Werner, K., Brndon, D., Clark, M., Gangopandhyay, S., 2004. Climate index weighting schemes for NWS ESP-based seasonal volume forecasts, Journal of Hydrometeorology, 5, 1076-1090.

Whitaker, J.S., Vitart, F., Wei, X., 2006. Improving week two forecasts with multi-model re-forecast ensembles, Monthly Weather Review, 134, 2279-2284.

Wilks, D.S., 2000. Diagnostic verification of the Climate Prediction Center long-lead outlooks, 1995-1998, Journal of Climate, 13, 2389-2403.

Wilks, D.S., Hamill, T.M., 2007. Comparison of Ensemble-MOS methods using GFS reforecasts, Monthly Weather Review, 135, 2379-2390.

Wilks, D.S., 1995. Statistical methods in the atmospheric sciences: an introduction, International Geophysics Series, Volume 59, Academic Press, San Diego.

Wurbs, R.A., 1998. Determination of generalized water resources models in the United States, Water International, 23, 190-198.

Zhao, R.J., 1991. The Xinanjiang model applied in China, Journal of Hydrology, 135, 371-381.

Appendix

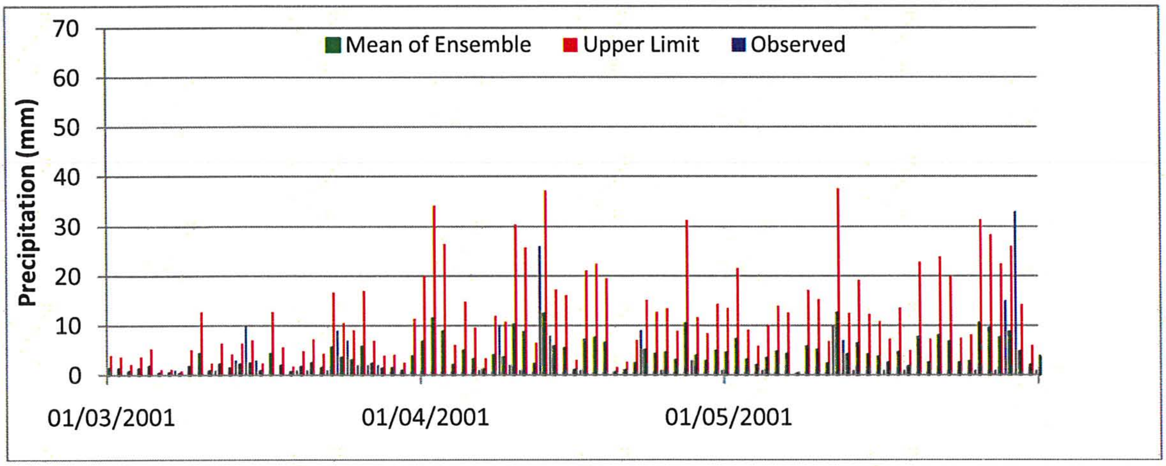

Figure A1: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDD for the spring season (limit shows 95% confidence interval)

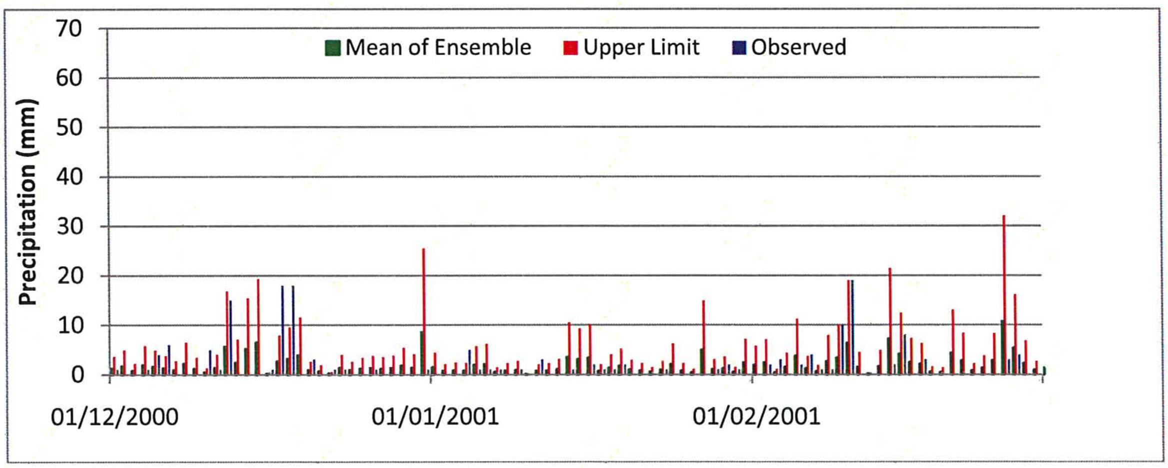

Figure A2: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDD for the winter season (limit shows 95% confidence interval)

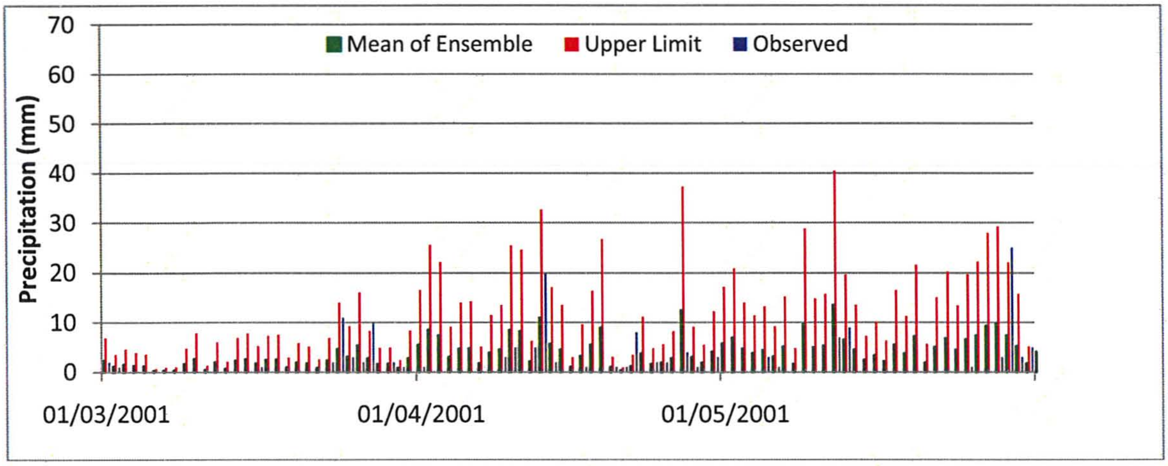

Figure A3: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDP for the spring season (limit shows 95% confidence interval)

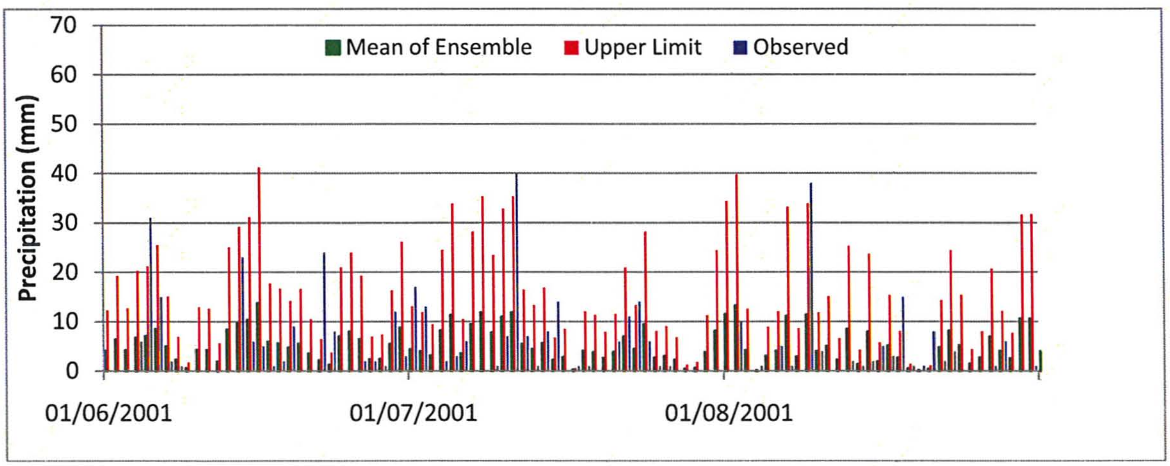

Figure A4: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDP for the summer season (limit shows 95% confidence interval)

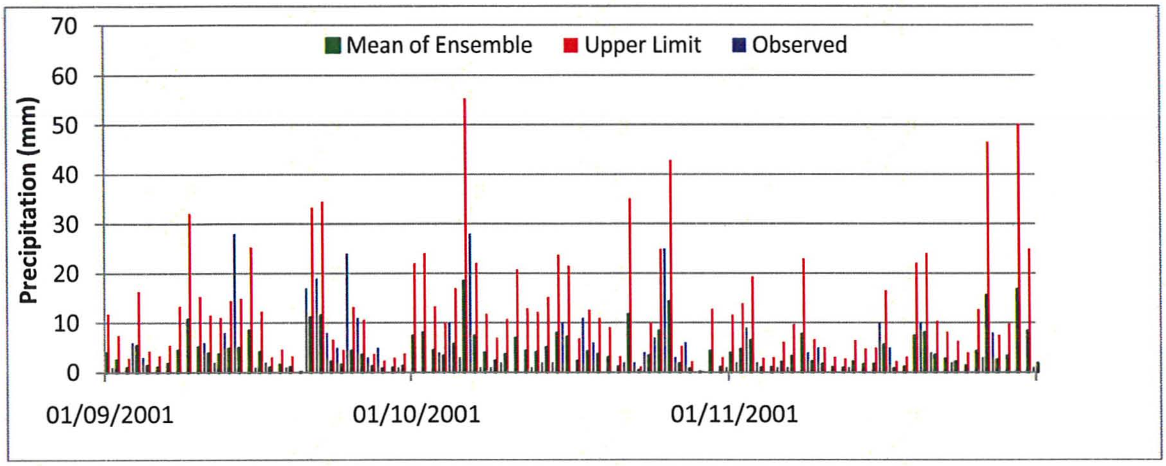

Figure A5: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDP for the autumn season (limit shows 95% confidence interval)

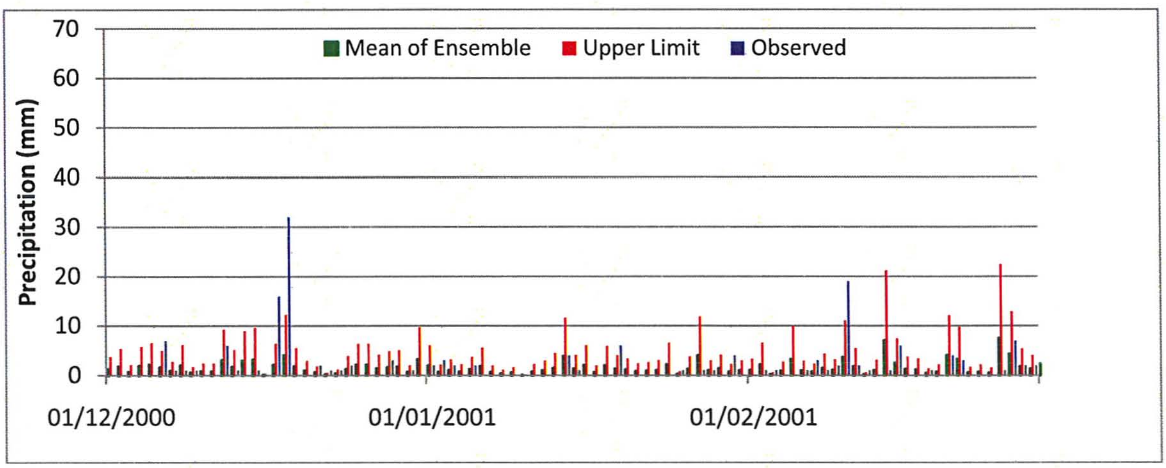

Figure A6: Plots of NCEP precipitation forecasts (5 day ahead) and observed precipitation at CDP for the winter season (limit shows 95% confidence interval)