Utilizing North American Regional Reanalysis for Climate Change Impact Assessment on Water Resources in Central Canada

Sung Joon Kim, Mark Lee, Woonsup Choi, and Peter F. Rasmussen
Department of Civil Engineering, University of Manitoba, 15 Gillson St., Winnipeg, Manitoba, Canada R3T 5V6

ABSTRACT

The province of Manitoba, Canada, is blessed with abundant surface water resources but lacks in weather stations. As a result, hydrological modeling and climate change impact assessments for water resources management face difficulties due to limited input data. Recent studies suggest that the North American Regional Reanalysis (NARR) has high potential for use as input data for hydrological modeling and statistical downscaling of GCM data.

The objective of this study is (1) to utilize the NARR data for hydrologic modeling and statistical downscaling of GCM data, and (2) to assess the climate change impact on Manitoba water resources. Two river basins in north-western Ontario were selected for this study. The SLURP model was set up with meteorological input data from NARR and calibrated for each catchment against the observed streamflow data. K-Nearest Neighbour (*k-NN*) resampling, a statistical downscaling technique, was used to downscale the output from the Canadian global climate model CGCM3 under the SRES A2 and B1 emission scenarios (2081-2100). The downscaled CGCM3 data were used as input to the calibrated SLURP model to assess the future climate change impact on water resources.

The results indicate that (1) the SLURP hydrological model can be reasonably calibrated with the meteorological input data from NARR, (2) the results from the statistical downscaling for baseline climate with NARR are comparable to the NARR data, (3) the warmer and wetter climate in the future under the A2 scenario is likely to lead to an increase in runoff, and (4) the B1 scenario resulted in different runoff changes in two catchments. NARR is found to be a good alternative to weather station data for climate change impact studies in data-scarce central Canada, where higher risk of flooding and lower risk of extended droughts are projected due to climate change.

Key words: climate change; hydrological modelling; North American Regional Reanalysis; nearest neighbour resampling

INTRODUCTION

The province of Manitoba, Canada, has an abundance of surface water. Manitoba Hydro, the sole provider of energy in the province, produces more than 95% of electricity from hydropower generating stations on the Nelson and the Winnipeg River basins, located in northern Manitoba and north-western Ontario, respectively. Hydrological modelling is believed to be a useful tool for water resource management, but is difficult in those regions due to lack of reliable meteorological data.

North American Regional Reanalysis (NARR) is “a long-term, dynamically consistent, high-resolution, high-frequency, atmospheric and land surface hydrology dataset” (Mesinger et al., 2006), and its spatial domain covers United States, Canada, and Mexico.

The NARR data comes with 3 hourly temporal and 32km spatial resolutions (see Figure 2) and assimilated precipitation (Mesinger et al., 2006) which makes it superior to previous global reanalysis data sets from the National Center for Environmental Prediction–National Center for Atmospheric Research (Kalnay et al., 1996) and the NCEP–Department of Energy (Kanamitsu et al., 2002).However, it has rarely been used for hydrological modeling.

Choi et al. (2007) evaluated the temperature and precipitation data from NARR by comparison with selected weather stations in Manitoba and concluded that they have good potential for use as input data for hydrological models. Kim et al. (2007) conducted a pilot study on evaluating the reliability of NARR data for hydrologic modelling. They applied NARR data to calibrate a hydrologic model and compared it to the calibration obtained with observed weather station data in northern Manitoba. In their study, the use of NARR data for hydrological modelling was found to be promising.

The objectives of this study are to use NARR data for downscaling GCM data and for hydrological modeling and to assess climate change impacts on future runoff to the study area. Two catchments in Winnipeg River basin located in north-western Ontario, Canada were modeled. This study may provide insight into climate change impact on water resources in data scarce regions.

STUDY AREA

Two catchments were selected for this study: the Sturgeon River and the Troutlake River in north-western Ontario (Figure 1 and Table 1). Trout River and Sturgeon River gauging stations are located within 50 km from the Red Lake Airport weather station and Sioux Lookout Airport weather station, respectively.

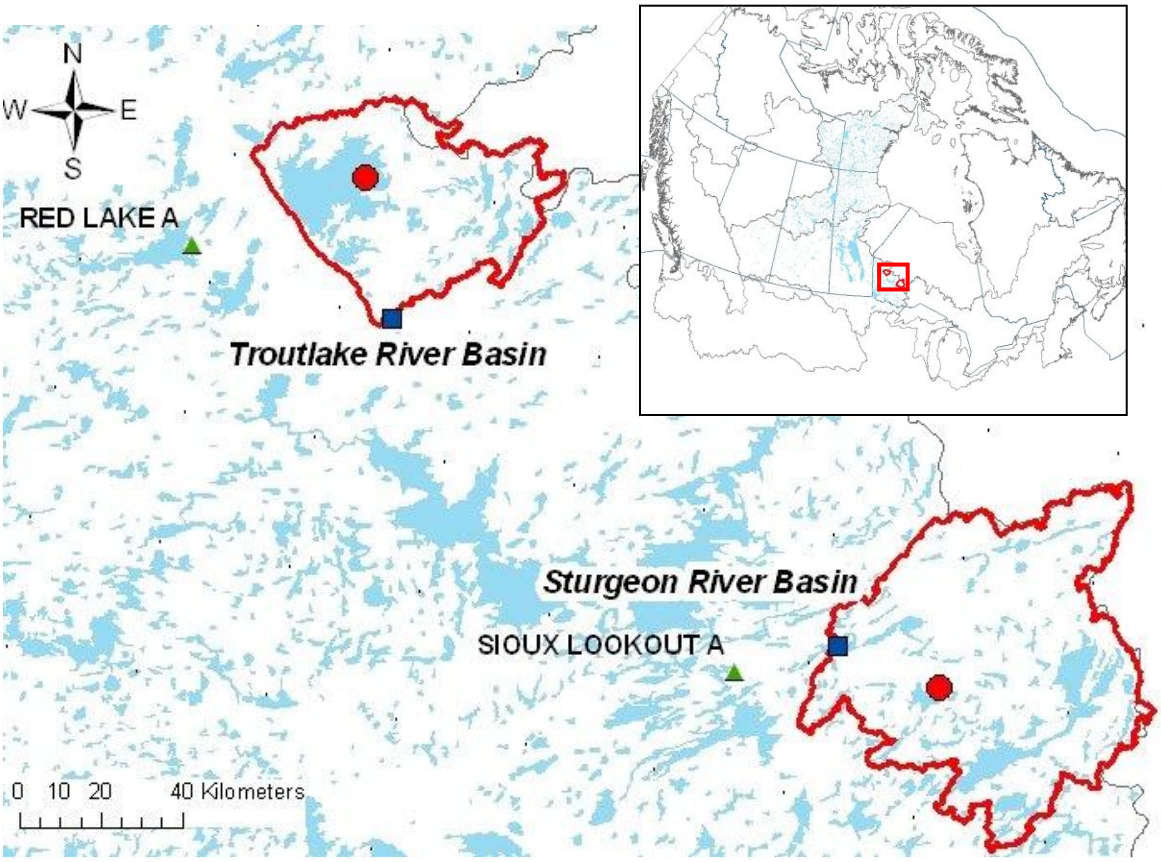

Figure 1. Study area and NARR grid points for each basin.

Table 1. **Streamflow gauges in the study area.**

Station No.	Name	Period of Record	Drainage Area(km^2)
05QC004	Sturgeon River at McDougall Mill	1961 – present	4450
05QC003	Troutlake River above big fall	1970 – present	2370

DATA

The NARR data was downloaded by the NOAA-ESRL Physical Sciences Division, Boulder Colorado from their Web site at http://www.cdc.noaa.gov/. NARR data have not been widely evaluated, primarily because NARR is a recent product. As show in Figure 2, the spatial resolution of NARR is significantly higher than the resolution of GCM, which is directly applicable to hydrological assessments. Since the NARR data cover all over the North American domain distributed, it can be applied for the hyrological assessment of remote areas located in a distance from a weather station.

For hydrologic modelling and statistical downscaling, time series data of daily mean temperature, relative humidity, solar radiation, and daily accumulated precipitation at a NARR grid point for each catchment were provided.

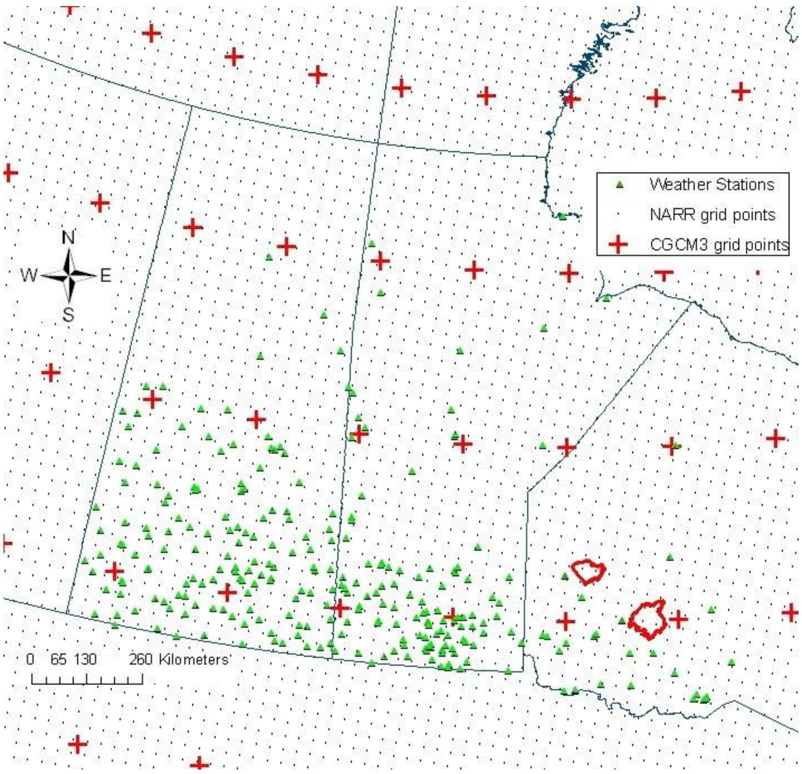

Figure 2. Location of Environment Canada weather stations, NARR grid points, and CGCM grid points in central Canada.

Prior to the hydrologic simulations using observed and NARR data, the two data sets were compared to evaluate the reliability of NARR for hydrological modelling. The observed daily mean temperature and total precipitation were obtained from two Environment

Canada weather stations in study area and compared with NARR data for the period of 1979-2004. The stations are located in Sioux Lookout (WMO ID 71842) and Red Lake (WMO ID 71854).

Mean monthly observed and NARR data were compared for precipitation and temperature. NARR precipitation during summer and autumn months is lower than that observed at the weather stations, while NARR temperature is higher than observed temperature (Figure 3).

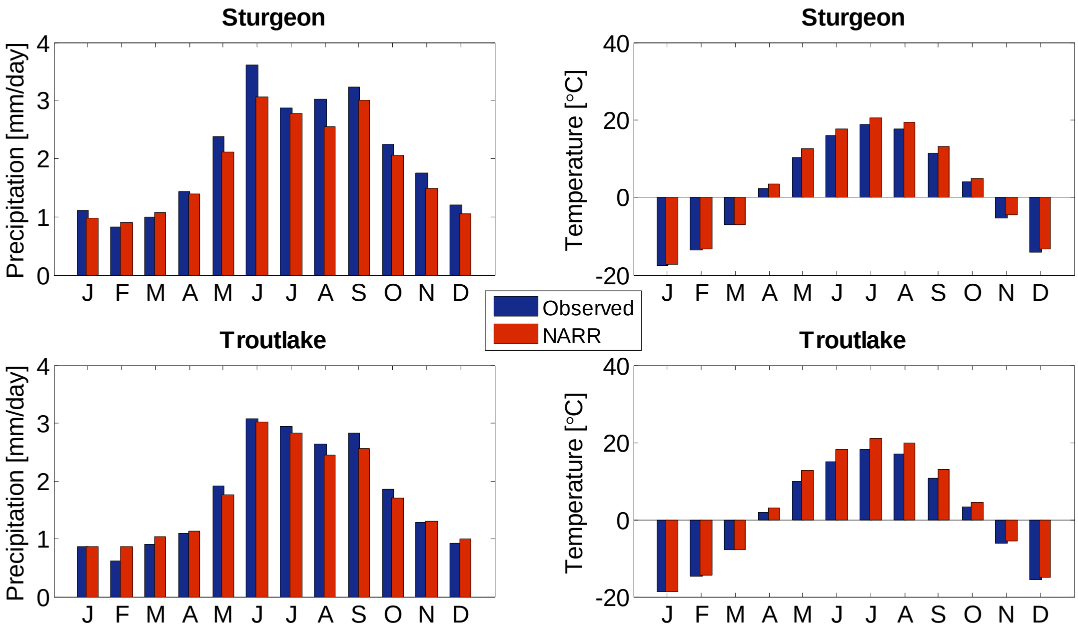

Figure 3. Mean monthly precipitation and temperature from weather stations and NARR in each basin.

HYDROLOGICAL MODELLING

The SLURP model

The SLURP (Semi-distributed Land Use-based Runoff Processes) model (Kite, 1995) is a semi-distributed conceptual hydrological model that was initially developed for modelling meso-scale Canadian watersheds as an alternative to the use of larger and more complicated hydrological models. The SLURP model simulates runoff based on daily precipitation, mean temperature, relative humidity, and bright sunshine hours, and physiographic data such as land cover and elevation.

In SLURP, a basin is divided into a number of aggregated simulation areas (ASAs). An ASA contains certain types of land cover and the vertical water balance is calculated in each land cover in each ASA. The water is routed to the outlet of each ASA and then to the outlet of the basin. SLURP simulates the vertical water balance with four storage tanks in each land cover in each ASA: canopy store, snow store, fast store, and slow store (Figure 4). Precipitation is provided as input of water to ASAs, and fluxes such as interception, sublimation, evapotranspiration (ET), surface runoff, interflow, and base flow are calculated from the storage tanks.

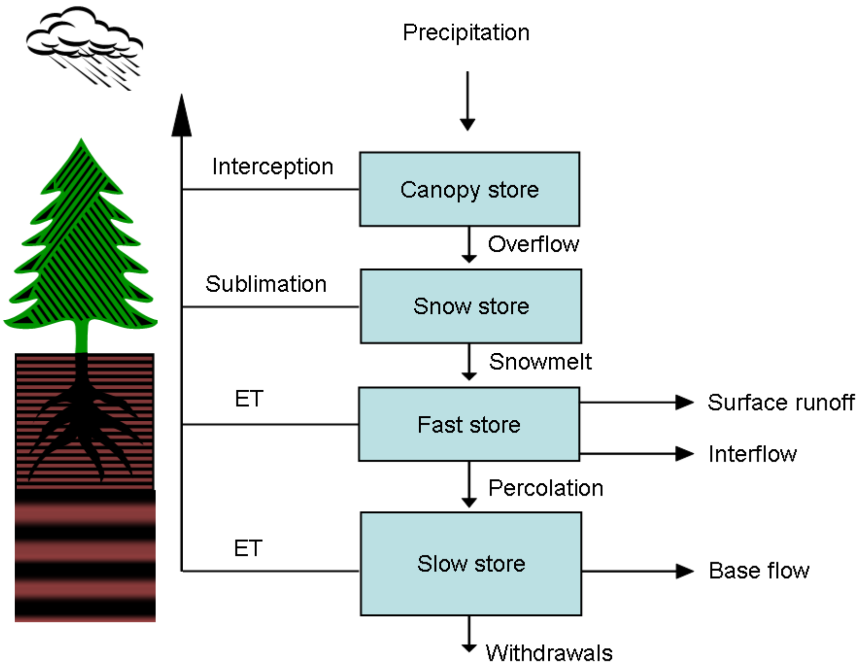

Figure 4 Vertical water balance in SLURP (adopted from Kite, 2000).

The SLURP model was set up for each catchment with digital land cover and elevation data. The digital elevation data were obtained from the National Aeronautics and Space Administration (NASA) Shuttle Radar Topography Mission (SRTM), and land cover data sets were derived from the Advanced Very High Resolution Radiometer (AVHRR) and the Forest Resources Inventory (FRI) of Manitoba data.

SLURP calibration and validation using NARR data

Each model was calibrated using streamflow data measured at the Sturgeon River and the Troutlake River. For comparison purpose, hydrologic simulations with observed weather station data were also considered. Since all weather stations have missing data, the calibration and validation were conducted over the periods of most complete records for each basin. The Sturgeon-model was calibrated first for the period of 1992-1995 and validated over the period 2000-2004. The Troutlake-model was calibrated for the period 1994-1997 and validated over 2000-2004. Since the NARR does not contain any missing data, the calibration and validation with NARR data can be conducted using any period. However, for fair comparison with observed data, the same validation periods were used (1989-1992 for calibration and 2000-2004 for validation).

The key parameters adjusted during the calibration were maximum infiltration rate (mm/d), retention constant for fast store (RCFS; in days), maximum capacity of fast store (MCFS; in mm), retention constant for slow store (RCSS; in days), maximum capacity of slow store (MCSS; in mm), rain/snow division temperature (in °C), canopy capacity (in mm), albedo, snowmelt rate (in mm/day), and evaporation-related parameters such as wilting point and field capacity.

The calibration criteria include deviation of volume (D_v) of mean runoff and Nash-Sutcliff Efficiency (E) of the daily runoff series. These statistics are explained by Legates and McCabe Jr. (1999) and measure volumetric error, goodness-of-fit, and daily average error between simulation and observation, respectively. Table 2 presents a summary of model performances of observed data and NARR data for the validation periods. As seen from the table, the performance statistics using NARR data are similar to the statistics obtained using observed data. In both cases, the average simulated runoff is close to the streamflow record and daily E-values are at an acceptable level.

Table 2. **Results from the SLURP model validation using observed and NARR data for each basin.**

	Sturgeon		Troutlake	
	Observed	NARR	Observed	NARR
Observed mean runoff (m^3/s)	46.13		20.68	
Simulated mean runoff (m^3/s)	49.14	45.04	19.86	19.86
D_v of mean runoff (%)	6.51	-2.38	-3.96	-3.99
E of daily runoff series	0.77	0.64	0.65	0.61

Observed and simulated monthly mean runoff for each basin is depicted in Figure 5. There appears to be a tendency that months with high runoff are underestimated and months with low runoff are overestimated.

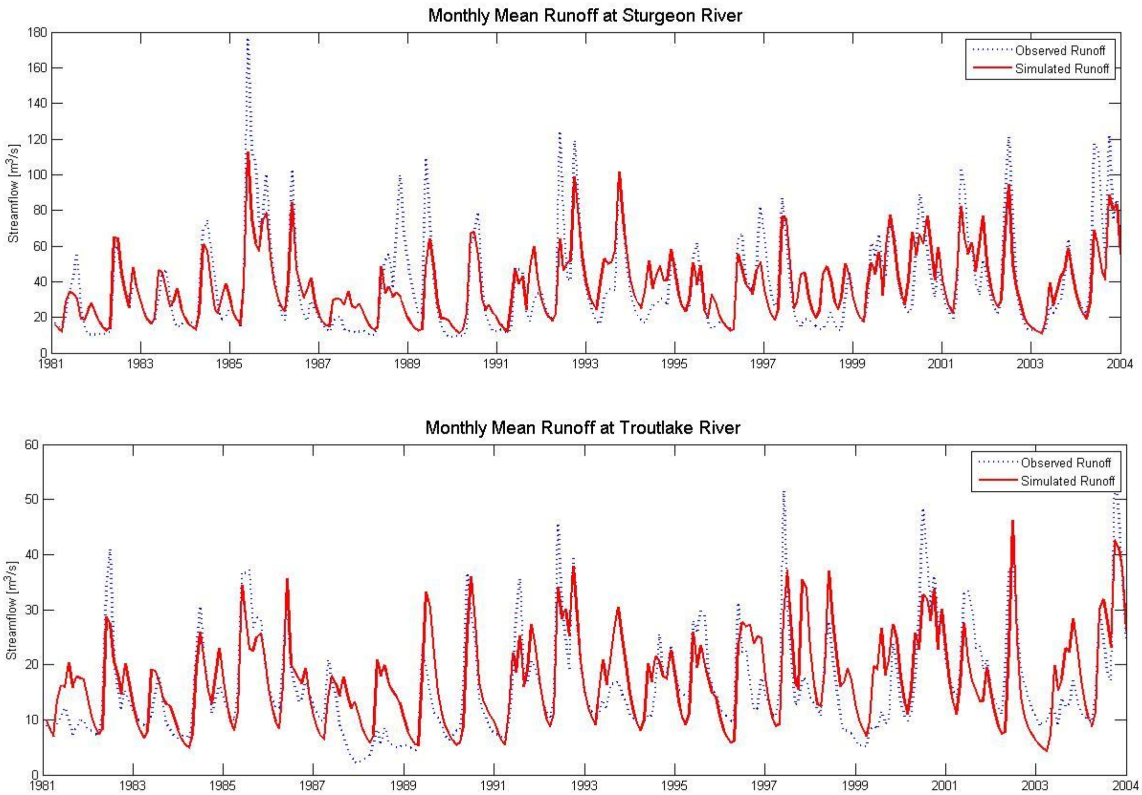

Figure 5. Observed and simulated monthly runoff series for each catchment.

SIMULATION USING DOWNSCALED GCM DATA

K-nearest neighbour (*k-nn*) downscaling

Nearest neighbor resampling is a nonparametric statistical downscaling method that has the primary advantage of avoiding the complex parameterization process of other statistical downscaling models. Local weather data are produced by strategically resampling from a historical record based on similarity of the daily large-scale atmospheric patterns of the GCM. Reanalysis data from the National Centre for Environmental Prediction (NCEP) provides the historical record of large scale atmospheric data while NARR data will provide the historical local weather. The nearest neighbors, or statistically most similar days, to the simulation day in the historical record are determined. One of the nearest neighbors is selected by random sampling. Since the resampled day has similar large-scale weather conditions, which are correlated to local weather conditions, this day provides the desired local weather variables for the simulated atmospheric conditions. The process is repeated for each day.

The comparison of the simulated atmospheric conditions to the historical record is made by using a feature vector, D_t, of the simulated day,

$$D_t = [v_1 \; v_2 \; v_3 \ldots v_n]$$

Components of the feature vector, v_i, are the *n* variables selected to describe the large-scale atmospheric circulation on day *t*. Since variables may have ranges that differ in orders of magnitudes, the variables are standardized by subtracting the mean and dividing by the standard deviation. To reduce the effects of seasonal variation the variables are standardized using the mean and standard deviation specific to the calendar day *t*. Smoothed estimates of daily mean and standard deviation were obtained through Fourier series analysis. To further reduce the effect of seasonal variation, the nearest neighbors are selected from within a moving window of width *W* centered on the calendar day of the simulated day.

The feature vector of the simulated day is compared to the same feature vector of days in the historical record by calculating the Euclidean distance (Gangopahyay et al., 2005; Buishand and Brandsma, 2001; Rajagopalan and Lall, 1999),

$$\delta_{tu} = \sqrt{\sum_{i=1}^{n} w_i \left(v_{ti} - v_{ui}\right)^2}$$

where δ_{tu} is the distance between the simulated feature vector and the feature vector of the historical day, D_u, and w_i are weights applied to the individual feature vector variables v_{ti} and v_{ui}. The value of the weights can be adjusted to optimize the performance of the resampling procedure. Days with the smallest distance are the nearest neighbors to the simulated day and exhibit the most similar atmospheric conditions.

For a given simulation day, a number of nearest neighbors, *k*, are retained from the historical record. One of these days is randomly selected to provide the local climate variables for the simulated day. A probability weighting scheme is used to give more weight to the closer neighbors. A decreasing kernel density function (Lall and Sharma, 1996; Brandsma and Buishand 1998; Buishand and Brandsma, 2001; Wojcik and Buishand, 2003),

$$p_j = \frac{1/j}{\sum_{i=1}^{k} 1/i}$$

is used to assign probability weights to the nearest neighbors based on their ranked distances. In the above equation, p_j is the probability that day j is resampled. Once a nearest neighbor has been selected, the desired station data for the simulated day is provided by the station data recorded at the resampled day.

The process of selecting a nearest neighbour is repeated for each GCM-simulated day.

K-nn **model set-up**

In the application presented here, GCM data were downscaled to produce time series data of mean daily temperature, relative humidity, solar radiation, and daily accumulated precipitation at five NARR grid points. Twenty-six years (1979-2004) of NARR data were available to use as historical data base of local weather. The NCEP Reanalysis 1 (Kalnay et al., 1996) supplied the historical large scale atmospheric data. The CCCma third generation coupled GCM, CGCM3.1/T47 was selected to provide the simulation data for a 20th century control run (20c3m, 1961-2000) and the IPCC SRES A2 and B1 emission scenarios.

To adequately capture the large-scale circulation patterns, a large spatial area was selected over the region. The average surface temperature, 500mb temperature, 850mb temperature, 500mb geopotential height, and 850mb geopotential height variables were used as the large-scale variables. The grids for the NCEP and CCCma data sets have slightly different resolutions, 2.5° x 2.5° and 3.75° x 3.75°, respectively. To make the data sets consistent, the NCEP data was linearly interpolated onto the CCCma grid points. The data covers the region in Figure 2 and consists of 42 data points.

NCEP and the 20c3m experiment data were standardized using the mean and standard deviation from each data set to remove slight biases in the model for the current climate. The A2 and B1 scenario data was standardized using the means and standard deviations from the 20c3m data to preserve the biases created in the data due to changed atmospheric loadings.

Since each large scale climate variable contained 42 data points, the total number of data to compare between NCEP and CCCma data was 210. Principal component analysis was used to reduce the number of variables in the feature vector by removing the redundancy of information (Gangopadhyay et al., 2005; Buishand and Brandsma, 2001; Young, 1994).

To optimize the model, a cross validation method was set up in which the model was used to predict the historical station data. The NCEP data for one year was considered as simulation data and removed from the historical record. Local weather was then generated for this year of NCEP data by resampling NARR data from the historical record. This process was repeated for each of the 26 of years of historical data. The cross validation was repeated multiple times using optimization software to obtain the parameters, W (26 days), k (10), and w_i, that provided the best reproduction of historical local weather.

The cross validation showed the *k-nn* model was able to adequately reproduce the local weather variables for historical data. Figure 6 displays the results for one of the two grid points from downscaling the historical large scale climate variables to reproduce the historical local weather. Figure 6 (left) shows the estimated and observed accumulated precipitation over the winter months. Figure 6 (right) shows the estimated and observed daily temperature. A correlation of 0.96 was achieved for daily temperature.

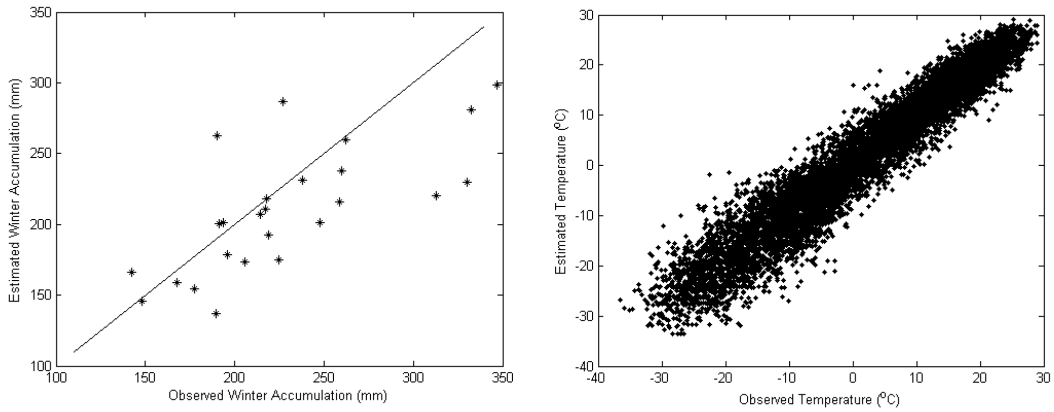

Figure 6. Estimated and observed winter precipitation accumulation (left) and winter daily temperature (right) for a NARR grid point in the Sturgeon River catchment.

Downscaled CGCM3 20C3M precipitation and temperature data were compared to those of NARR (Figure 7). Downscaled mean precipitation are generally lower than NARR except for the months of July and August for both catchments. The precipitation in June appears significantly lower than other months. Downscaled temperature data were fairly close to NARR data over the whole year.

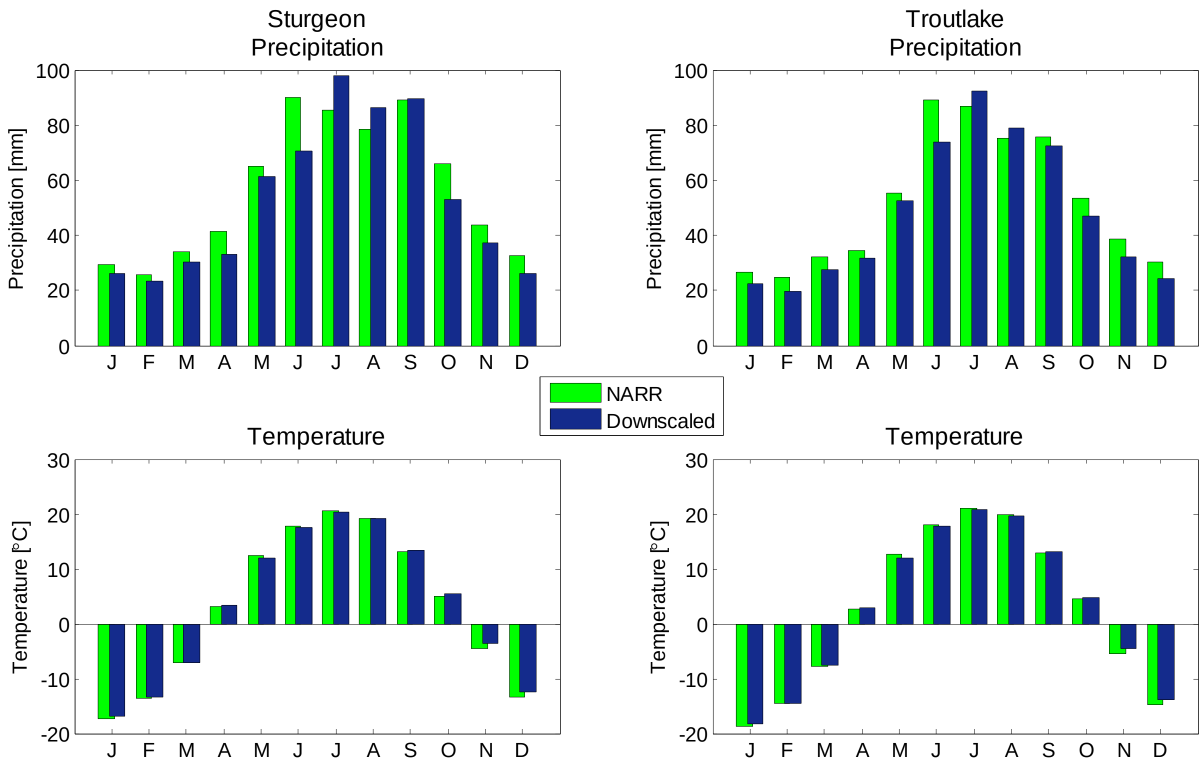

Figure 7. Mean monthly precipitation and temperature from NARR and downscaled CGCM3 data (20C3M) for historical period (1961-2000) in each catchment.

Downscaling of GCM data

With the *k-nn* model showing good performance in cross validation, the model was then used to downscale the IPCC SRES A2 and IPCC SRES B1 climate change scenarios. The *k-nn* model used the large scale climate variables provided by the GCM to resample days from the historical NARR data at two grid points located in the center of the Sturgeon River and the Troutlake River catchments, respectively.

Table 4. Downscaled annual precipitation and temperature for 20C3M, A2, and B1. Changes from 20C3M are shown in parenthesis.

	Sturgeon			Troutlake		
	20C3M	A2	B1	20C3M	A2	B1
Annual Precipitation (mm)	634	687 (+8.4%)	622 (-1.9%)	576	597 (+3.6%)	585 (+1.6%)
Average Temperature (°C)	3.22	5.67 (+2.45°C)	5.63 (+2.41°C)	2.77	5.62 (+2.85°C)	5.30 (+2.53°C)

Both scenarios show increase of annual precipitation for Troutlake River, while annual precipitation in Sturgeon River increases for the A2 scenario but decreases for the B1 scenario. As expected, both scenarios show increases in temperature. Average temperature in both catchment increased by aproximately 2.5°C (see Table 4). As shown in Figure 8, both scenarios show higher increases of temperature in the winter than in the summer.

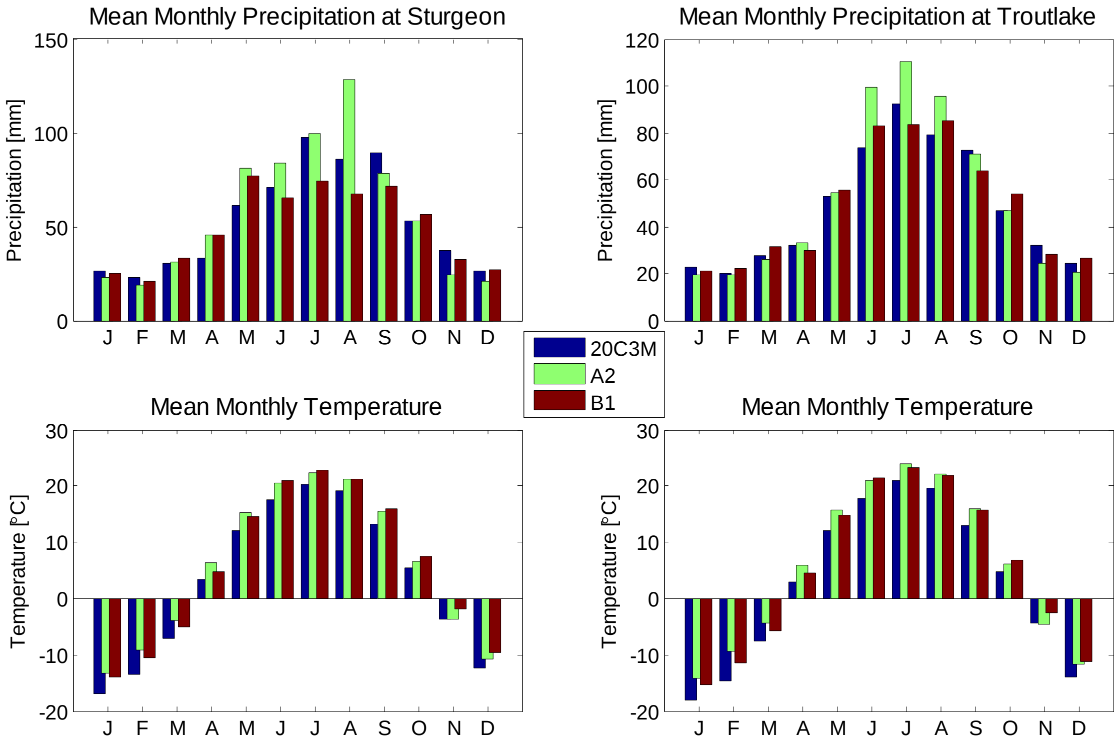

Figure 8. Mean monthly precipitation and temperature obtained from CGCM3 and downscaled by k-nn for 20C3M, A2, and B1.

Hydrological modelling using downscaled GCM data

Downscaled CGCM3 data for the two scenarios and the control run were applied in the SLURP model for each catchment. As expected, simulated future runoffs reflects the projected change in future precipitation. In B1 scenario for Sturgeon River, as precipitation decreases runoff also decreases. However, runoff increases as precipitation increases in A2 scenario for Sturgeon River and both scenarios for Troutlake River. The increasing rate of runoff is proportional to the rate of precipitation. For instance, runoff increases the most (Table 5) in A2 scenarios for Sturgeon River as precipitation does (Table 4).

Table 5. Mean annual runoff (in m^3/s) simulated by SLURP for 20C3M, A2, and B1. Changes from 20C3M are shown in parenthesis.

	20C3M	A2	B1
Sturgeon	31.2	34.7 (+11.2%)	28.9 (-7.4%)
Troutlake	14.9	16.1 (+8.1%)	15.3 (+2.7%)

The temperature increases in the winter advance the spring snowmelt runoff by almost a month. The increased summer precipitation in the A2 scenario results in increased runoff in the summer and in the early autumn. On the other hand, the B2 scenario shows decrease in summer runoff for the Sturgeon River and no change for the Troutlake River (Figure 9).

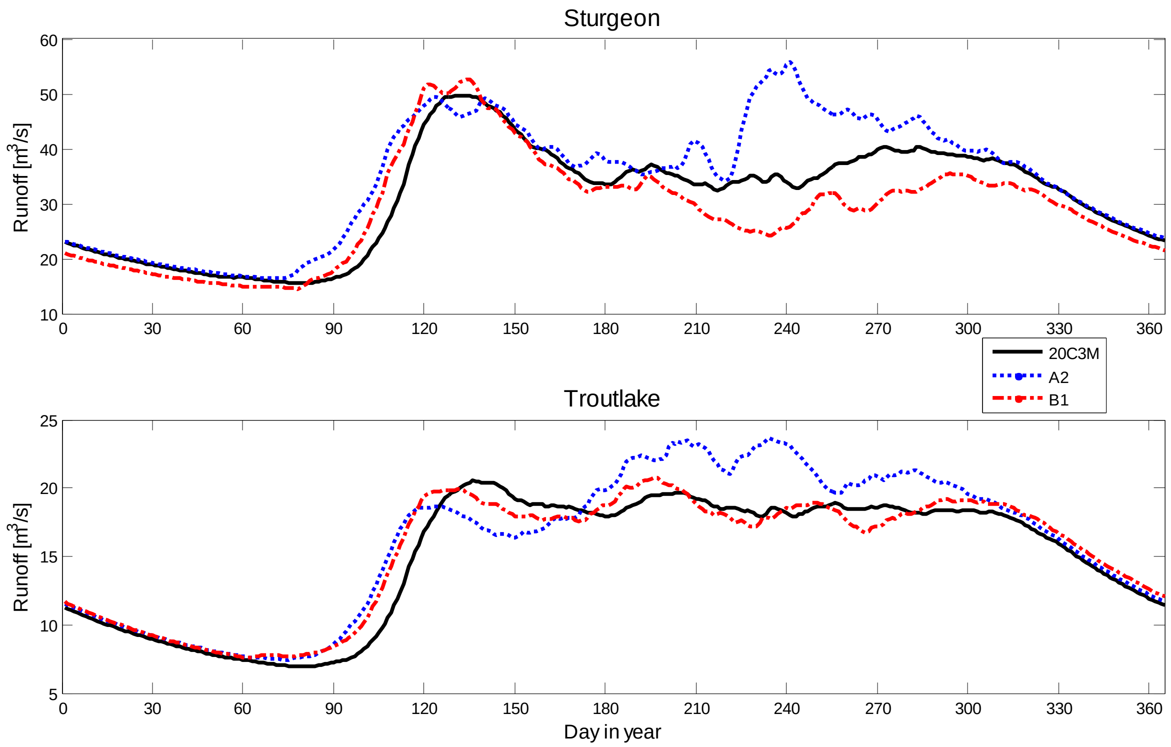

Figure 9. Mean daily runoff simulated by SLURP with the downscaled CGCM3 output for 20C3M, A2, and B1.

DISCUSSION AND CONCLUSIONS

We utilized the NARR data for hydrological modelling and statistical downscaling of GCM data to assess the hydrological changes in the Winnipeg River basin under climate change. Future climate scenarios were generated by statistically downscaling GCM data under the A2 and B1 emission scenarios with the *k-nn* method for which the NARR data were used as historical records.

Since the *k-nn* method is a nonparametric statistical downscaling method, it can be relatively easily implemented without complex parameterization processes required by parametric models such as SDSM (Wilby, 2002) and LARS-WG (Semenov and Barrow, 1997). However, the method produces the local weather by resampling using historical records, which means that extreme values are conditioned by the historical records. Although the *k-nn* method underestimated precipitation, generated local weather data were quite acceptable for hydrologic modelling in overall.

In this study, we obtained reliable results for the current hydrological and climatic conditions by using the NARR data for hydrological modelling and the *k-nn* statistical downscaling. Since NARR contains almost every atmospheric and surface variable required for climate change impact studies, the results are especially promising remote regions where climate data are scarce. The results from hydrological simulations are comparable to those from previous studies (e.g. Woo and Thorne, 2006; Kim et al., 2007) and statistical downscaling with the NARR data is an unprecedented attempt.

Hydrological simulations with future climate scenarios demonstrate that March and April runoff is likely to increase substantially under both A2 and B1 scenarios and summer and autumn runoff is likely to increase under the A2 scenario while decreasing under the B1 scenario. It should be noted that the average magnitude of spring peak flows does not change much in any case. Therefore, analyses on the changes in the variability of spring peak flows and the magnitudes of extreme hydrological events are anticipated in a subsequent study.

ACKNOWLEDGEMENT

This study was funded by the Natural Sciences and Engineering Research Council of Canada, Manitoba Hydro, and the Prairie Adaptation Research Collaborative. GCM data was obtained from the Canadian Centre for Climate Modeling and Analysis and NARR data was provided by the NOAA-ESRL Physical Sciences Division, Boulder Colorado.

REFERENCES

Brandsma T. and T.A. Buishand (1998) Simulation of extreme precipitation in the Rhine basin by nearest-neighbor resampling, *Hydrology and Earth System Sciences*, 2(2-3), 195-209.

Buishand, T.A. and T. Brandsma (2001) Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling, *Water Resource Research*, 37(11), 2761-2776.

Choi, W., Moore, A. and Rasmussen, P. F. (2007) Evaluation of temperature and precipitation data from NCEP-NCAR global and regional reanalyses for hydrological modeling in Manitoba. *Proceedings of 18th Canadian Hydrotechnical Conference*, Winnipeg, 10 pages

Gangopadhyay, S., M. Clark, B. Rajagopalan (2005) Statistical downscaling using k-nearest neighbors, *Water Resource Research*, 41(2), 1-23.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. (1996) The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society,* Vol. 77**,** 437-471.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M. and Potter, G. L. (2002) NCEP–DOE AMIP-II reanalysis (R-2). *Bulletin of the American Meteorological Society,* Vol. 83**,** 1631-1643.

Kim, S. J., Choi, W. and Rasmussen, P. (2007) Calibration of the SLURP hydrological model using North American Regional Reanalysis (NARR) data. *Proceedings of 18th Canadian Hydrotechnical Conference,* Winnipeg, 10 pages

Kite, G. W., 1995. The SLURP Model, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, chap. 15, pp. 521- 562, Water Resources Publications.

Kite, G. (2000) *Manual for the SLURP Hydrological Model, V. 11.4*. Colombo, Sri Lanka: International Water Management Institute.

Lall, U., and A. Sharma (1996) A nearest neighbor bootstrap for resampling hydrologic time series, *Water Resource Research*, 32(3), 679-693.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D. and Shi, W. (2006) North American Regional Reanalysis. *Bulletin of the American Meteorological Society,* Vol. 87**,** 343-360.

Rajagopalan, B. and U. Lall (1999) A k-nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resource Research*, 35(10), 3089-3101.

Semenov, M. A., and E. M. Barrow, Use of a stochastic weather generator in the development of climate change scenarios, *Climate Change*, *35* , 397-414, 1997.

Wilby, R. L., C. W. Dawson, and E. M. Barrow, SDSM - a decision support tool for the assessment of regional climate change impacts, *Environmental Modelling And Software*, *17* , 147-159, 2002.

Woo, M.-k. and Thorne, R. (2006) Snowmelt contribution to discharge from a large mountainous catchment in subarctic Canada. *Hydrological Processes,* Vol. 20**,** 2129-2139.

Wójcik, R., T.A. Buishand (2003) Simulation of 6-hourly rainfall and temperature by two resampling schemes, *Journal of Hydrology*, 273(1-4), 29-80.

Young, K.C. (1994) A multivariate chain model for simulating climatic parameters from daily data, *Journal of Applied Meteorology*, 33, 661-671.