Long-term weather, streamflow, and water chemistry datasets for hydrological modelling applications at the upper La Salle River watershed in Manitoba, Canada

Marcos R. C. Cordeiro1 | Jason A. Vanrobaeys2 | Henry F. Wilson3

1Science and Technology Branch, Agriculture and Agri-Food Canada, Winnipeg, Manitoba, Canada
2Science and Technology Branch, Agriculture and Agri-Food Canada, Morden, Manitoba, Canada
3Science and Technology Branch, Agriculture and Agri-Food Canada, Brandon, Manitoba, Canada

Correspondence
Jason A. Vanrobaeys, Science and Technology Branch, Agriculture and Agri-Food Canada, Morden, MB, Canada. Email: jason.vanrobaeys@agr.gc.ca

Funding Information
This research was supported by A-Base funding under Agriculture and Agri-Food Canada’s Growing Forward 2 program.

Abstract
Long-term weather (1990–2013), streamflow (1990–2013; excluding 7 years with no or poor data), and water chemistry (2009–2013) datasets for hydrological modelling applications were developed using simple methods for the upper La Salle River watershed, in Canada, to address the lack of such datasets in the northern Red River Basin. Weather variables consist of temperature, relative humidity, wind speed, solar radiation, and precipitation disaggregated to an hourly time-step. The only hydrometric variable included in the dataset is stream discharge in a daily time-step. Water chemistry data consisted of total nitrogen (TN), total dissolved nitrogen (TDN), total phosphorus (TP), and total dissolved phosphorus (TDP). Samples were collected weekly during the open water season at the same site as the hydrometric gauging station from August 2009 to October 2012 with some gaps (i.e. Fall 2011, Spring 2012, September 2012). In 2013 the sampling frequency was increased to daily or sub-daily during high stream discharge and weekly during low stream discharge. A data overview indicates values within ranges reported for the area. Mean annual, winter, and summer temperatures were 3.5, −10.7, and 17.2°C, respectively. Annual relative humidity averaged 73.1% but was higher and more homogenous in cold seasons. Wind speed was similar over the year with annual average of 4.3 m/s. Solar radiation followed the typical curve reported for western Canada, with peak daily average values around 250 W/m² in July. Precipitation records were mostly comprised of dry hours with 75.3% of the events being equal or less than 2 mm/h. Stream discharge was typical of the Canadian Prairies; the average peak discharge over the entire period was larger in April (2.3 m³/s) due to large amounts of snowmelt runoff. Average concentrations of TN, TDN, TP, and TDP of 1.54, 1.35, 0.56, and 0.49 mg/L, respectively, were in agreement with values found in previous studies at the site.

Keywords
hydrometric, gap-filling, uncertainty, water chemistry, weather
The datasets presented in this document come as three separate comma separated (csv) files:

- Identifier: https://doi.org/10.23684/odi-2017-00957
  Creator: Cordeiro, M.R.C., J.A. Vanrobaeys and H.F. Wilson
  Title: La Salle river watershed 05OG008 hourly weather 1990 to 2013
  Publisher: Government of Canada
  Publication year: 2017
  Resource type: Metadata document
  Version: 1.0
  Identifier: https://doi.org/10.23684/odi-2017-00959
  Creator: Cordeiro, M.R.C., J.A. Vanrobaeys and H.F. Wilson
  Title: La Salle river near Elie 05OG008 daily discharge 1990 – 2013
  Publisher: Government of Canada
  Publication year: 2017
  Resource type: Metadata document
  Version: 1.0
- Identifier: https://doi.org/10.23684/odi-2017-00958
  Creator: Vanrobaeys, J.A., H.F. Wilson and M.R.C. Cordeiro
  Title: La Salle river 05OG008 water chemistry 2009 – 2015
  Publisher: Government of Canada
  Publication year: 2017
  Resource type: Metadata document
  Version: 1.0.

1 | INTRODUCTION

Lake Winnipeg, the tenth largest freshwater lake in the world, has experienced rapid eutrophication in the recent past due to increased nutrient input (McCullough et al., 2012). Due to the prominence of the Red River Basin as the primary source of increasing nutrient loading to Lake Winnipeg (Mayer and Wassenaar, 2012), recent hydrologic modelling efforts have focused on this basin and its major tributary, the Assiniboine River (Shrestha et al., 2012a; 2012b; Yang et al., 2014). The geographical extent of this basin also increases its importance from a trans-boundary perspective, as it encompasses portions of two Canadian provinces (Manitoba and Saskatchewan) and three US states (Minnesota, North Dakota, and South Dakota; Figure 1). While these modelling exercises represent an important step towards hydrological simulations in the Red-Assiniboine Basin, they were performed using a daily time-step, which is not adequate (a) to represent the hydrology of small catchments because of their short storm response times (Beven, 2011) or (b) to force process-based hydrological models (Fang and Pomeroy, 2008; Ellis et al., 2010; Fang et al., 2010, 2013; Skaggs et al., 2012; Zhou et al., 2014).

Physically based simulations of hydrological processes with focus on finer spatial scale may overcome some of these inadequacies, but they often require input data at sub-daily time-steps, which remains one of the major limitations for this type of modelling in this basin. Sub-daily weather data records have become more commonly available only with the relatively recent expansion of automated weather station networks (Meyer and Hubbard, 1992; Fiebrich, 2009; Estévez et al., 2011). As a result, long-term simulations using sub-daily time steps are often hindered due to lack of sub-daily data (Gaume et al., 2007). Moreover, the lack of sub-daily data is emphasized in regions where the weather station density is relatively sparse or where daily data are more widely available, which is the case in much of Canada (Hutchinson et al., 2009). Even when sub-daily records can be obtained, data gaps are a frequent limitation (Kim and Pachepsky, 2010) due to loss of older paper records (e.g. fire, accidents) or interruption of automated stations due to calibrations, malfunctioning, or relocation (Simolo et al., 2010).

Streamflow data comprise another important input for assessment of hydrological simulations. Streamflow information for a given watershed or region is crucial for hydrological studies (Mishra and Coulibaly, 2010). In Canada, daily hydrometric data such as stream discharge and stream level are usually available for gauged streams (Environment and Climate Change Canada, 2013). While daily data are usually adequate to assess long-term, process-based modelling due to simulation results being summarized at this time-step, hydrometric records in Canada are plagued by large data gaps (Mishra and Coulibaly, 2010). An inspection of the HYDAT database (Environment and Climate Change Canada, 2013) also indicates that most of the hydrometric stations located in the Canadian portion of the Lake Winnipeg Basin, whose largest area is comprised by the prairie provinces of Alberta, Saskatchewan, and Manitoba, operate only seasonally (i.e. from March to October) due to no-flow conditions caused by negligible discharge or river ice cover (e.g. Corriveau et al., 2013). Recent analysis during the flow period also indicates that the presence of in-channel control structures and river ice constitute an uncertainty factor for hydrologic simulations in the region (Rasouli et al., 2014; Cordeiro et al., 2017c; Mahmood et al., 2017).

Water chemistry data are also of critical importance to identify nutrient sources and loads to Lake Winnipeg. Long-term monthly sampling has been carried out near the mouth of major rivers discharging to Lake Winnipeg (McCullough et al., 2012). Long-term monitoring within the Lake Winnipeg Basin has occurred on major rivers and at the provincial borders (Environment and Climate Change Canada,
2015), but water chemistry sampling at lower-order streams is less frequent.

The objective of this work was to compile long-term datasets using simple methods to be used for hydrological simulations in a sub-catchment of the La Salle River watershed, which is a tributary of the Red River Basin. This watershed has been selected due to its unique characteristics and importance from a nutrient export perspective. The very high proportion of the watershed used as cropland (87%), the extremely level topography (slopes varying between 0.004% and 0.02%), soil texture mostly comprised of clays (as opposed to clay loam and loam textures to the east), the modest depressional storage (which contrasts to the “prairie pothole” region), and the intensive surface drainage in farmland, are unique features of the La Salle River watershed and contrast with other areas in the Red River Basin with regard to land use proportions and topographic relief. The watershed is also a prominent source of phosphorus in the Red River Basin, with reported concentrations of total phosphorus as high as 2.0 mg/L (McCullough et al., 2012) and total dissolved phosphorus as high as 1.2 mg/L (Corriveau et al., 2013). The dataset presented and discussed here is comprised of three major components: weather (1990–2013), streamflow (1990–2013 except years with no or poor data), and water chemistry (2009–2013) data. Weather parameters included in the dataset are temperature, relative humidity (RH), wind speed (WS), precipitation (PPT), and solar radiation (SR). The only hydrometric data included are stream discharge, while water chemistry data included total dissolved phosphorus (TDP), total phosphorus (TP), total dissolved nitrogen (TDN), and total nitrogen (TN).

The importance of these datasets lies on several aspects. First, it offers ready-to-use, long-term hourly input data usually required to force physically based hydrological models in a relevant watershed for non-point sources of P to Lake Winnipeg. Although compiled from available data sources, these input datasets underwent thorough quality control and gap-filling procedures that represent added value compared to the original data in terms of completeness. They also addressed lack of precipitation data in hourly time-steps using downscaling techniques that required parameter estimation and comparison to hourly observations collected nearby. Second, it presents novel water chemistry data during important hydrological periods in the region. Such datasets were not previously available. Finally, the present work describes the methods that could be used to compile similar datasets in other areas of the Red River Basin. This aspect cannot be overemphasized as it hinders the application of physically based hydrological modelling in this region. The techniques used and discussed to derive continuous time-series are also important for other research fields besides hydrological modelling that rely on assessment at fine temporal scale, such as hillslope hydrology, eco-hydrology, and limnology.

While the datasets present here differ in their temporal resolution, this difference does not minimize the value of the datasets because of the purpose of each respective dataset as model input or for model assessment. Input (e.g. weather) and assessment (e.g. streamflow) datasets do not require the same temporal resolution. The temporal resolution of input datasets for modelling purposes depends on algorithm requirements, especially for physically based models. For example, physically based energy-balance and snow redistribution models usually require sub-daily input data (e.g. Gray and Landine, 1988). On the other hand, model simulations can either be assessed sub-daily or aggregated at coarser resolution, according to the datasets available for assessment or the objectives of the modelling exercise (Cordeiro et al., 2017c; Mahmood et al., 2017). For short-term assessments (i.e. event-based hydrological responses), the present streamflow dataset is certainly not appropriate. However, this type of assessment typically relies on custom monitoring carried out for research purposes. For long-term assessments, which are the goal of the dataset presented and discussed in the manuscript, model assessment is usually done at coarser temporal scales, either daily or monthly. The streamflow data presented are adequate for such assessments. Thus, physically based simulations at sub-daily time-steps can still be aggregated and assessed at daily resolution and be used to improve our knowledge about hydrology. A few examples of hydrological modelling in the area forced at daily or sub-daily time-steps but assessed at coarser temporal resolution are available in the literature (Yang et al., 2014; Cordeiro et al., 2017c; Mahmood et al., 2017).

2 | STUDY AREA

The data collection and analysis focused on a 189 km² sub-catchment of the La Salle River watershed (Figure 2). This watershed, located in the central plains region of Manitoba, Canada (Graveline and Larter, 2006), is a tributary of the larger Red River. Thus, it is representative of the Red River Basin and ideal for long-term and physically based simulation of cold region hydrological processes and nutrient dynamics. The surface geology consists of lacustrine clay deposited in glacial Lake Agassiz characterized by a lower, dark grey clay and a thinner upper unit of lighter coloured, calcareous silty clay, with surface texture being predominantly clayey (La Salle Redboine Conservation District, 2007). The watershed is located in the Prairie Ecozone and has mean annual temperature around 2.5°C, mean summer temperature of 16°C and mean winter temperature of −13°C; the mean annual precipitation is 560 mm, out of which around 25% takes place.
as snow, while the potential mean annual gross evapotranspiration is about 834 mm (La Salle Redboine Conservation District, 2007).

3 | DESCRIPTION OF DATASETS

3.1 | Weather dataset

3.1.1 | Selection of parent stations

Parent stations were selected from the available stations around the study area as the primary source of data to compile the weather dataset. The stations screened out during the selection of the parent station (i.e. primary data source) were still used for gap-filling purposes (section 3.1). The closest weather stations with long-term records of sub-daily (i.e. hourly) data belonged to Environment and Climate Change Canada (ECCC) and Manitoba Agriculture, Food and Rural Development (MAFRD) (Table 1). The MAFRD station (Figure 2, station A) did not come into operation until the second quarter of 2007; thus, this station was not used in the analysis since it did not cover the period of interest (i.e. 1990–2013). In addition, priority was given to stations within the same weather network for improved data consistency. The Portage La Prairie CDA (Canadian Department of Agriculture; Figure 2, station B) was also excluded because data were only available at a daily time-step. The Portage Southport Airport station (Figure 2, station D) was the source of temperature, relative humidity, and wind speed data since this was the closest station with hourly data available. The Winnipeg International Airport station (Figure 2, station F) was the only station measuring solar radiation and was selected for this weather element. The only stations equipped with both tipping buckets and weighing gauges capable of measuring precipitation in both liquid and solid forms were located at the Portage Southport Airport (Figure 2, station D), at the University of Manitoba Research Station in Carman (Figure 2, station E), and at The Forks in Winnipeg (Figure 2, station G). However, none of these stations were selected because the records started either in 1999 (The Forks) or in 2004 (Portage Southport Airport and University of Manitoba Research Station). As a result, precipitation available in a daily time-step had to be disaggregated to an hourly time-step (sub-section 3.1.3). Among the candidate stations with

![Figure 1: The Red River Basin and those of its major tributaries, namely, the Assiniboine River and the Souris River](image_url)
daily precipitation, the Marquette station was selected due to the close proximity to the study area and measurement of precipitation as both rain and snow. Proximity was considered the most important criteria for selecting the weather station because of the inherent spatial variability in precipitation (Ramos-Calzado et al., 2008). The adjusted precipitation dataset for Canada, which corrects known measurement issues and inhomogeneities such as wind undercatch, evaporation and wetting losses, as well as trace observations, change in site location, change in observing procedure, and instrument deficiencies (Mekis and Hogg, 1999; Mekis and Vincent, 2011), was used in this case since it is available at a daily time-step.

3.1.2 | Gap-filling

The presence of gaps in meteorological time series is a very common problem for long term studies (Tardivo and Berti, 2014). The records for all the variables in the weather dataset but adjusted precipitation had some gaps that had to be infilled. The gaps in the temperature, relative humidity, wind speed, wind direction, and solar radiation data corresponded to 27.3%, 29.8%, 27.4%, 56.1%, and 63.0% of the records, respectively. Wind direction was not included in the present dataset because this variable is not commonly required as an input for hydrological modelling and more than half of the records were missing, resulting in an unacceptable uncertainty for modelling exercises after gap-filling. Most of these data gaps occurred from 18:00 to 3:00 hr (62% of the missing records) and during weekend days (45% of the missing records). The gaps were negligible (<0.4%) in the first 2 years of the records but became more prevalent after the beginning of automated measurements in 1994, being evenly distributed from 1994 to 2013 and averaging 4.6% in each year. The gaps in temperature and wind speed records were usually short (few hours) and distributed over the entire time series. Data gaps in RH records occurred systematically from 18:00 to 3:00 hr and during weekends until 1993, and then occurred only sporadically from 1994 onwards, indicating the beginning of automated measurements. Similar to temperature and wind speed, gaps in solar radiation records were short. However, they were mostly found between years 1992 and 2000. The absence of gaps in precipitation is due to the time-step used (i.e. daily) and the origin of the dataset (i.e. adjusted; c.f. section 3.1).

A large number of methodological approaches are available for data gap-filling such as within-station (e.g. interpolation), between-station, and regression-based (simple, multivariate, nonlinear) methods (Kemp et al., 1983; Allen...
and DeGaetano, 2001; Tardivo and Berti, 2012, 2014), as well as downscaled and other reanalysis (e.g. gridded) datasets (Hutchinson et al., 2009; Hopkinson et al., 2011; Thornton et al., 2012; Wang et al., 2016). However, preference was given to simple linear regression due to the sparse density of weather stations in the area, the length of the dataset, and potentially large gaps for which interpolation techniques are not adequate. When simple linear regression was not sufficient to address all the data gaps, different gap-filling strategies were used to reconstruct the datasets, such as data transplanting and long-term averages. A summary of the datasets used in the analysis, their sources, record length, and gap-filling methods used are presented in Table 2.

The uncertainty of gap-filled records was assessed for temperature, relative humidity, wind speed, and solar radiation by calculating the 95% confidence interval for time step (i.e. hour) in the observed records and comparing the missing records to the confidence interval range. The confidence interval for each hourly record was estimated by (a) sub-setting the entire dataset for each given hour (e.g., selecting the records from all years for 1 Jan. 01:00), (b) calculating the mean and the standard error at 95% confidence level for the subset, and (c) calculating the lower and upper limit of the confidence interval by subtracting and adding the standard error to the mean, respectively (Helsel and Hirsch, 2002; Coleman and Steele, 2018). For precipitation, the downscaled dataset was assessed using observed hourly precipitation from the closest station (i.e. Portage Southport Airport) for the period available (i.e. 2004 onwards).

**Relative humidity**

Similar to temperature, gaps in the RH records were infilled using linear regression between the Portage Southport Airport station and the Winnipeg International Airport or The Forks stations. The coefficient of determination $R^2$ between the target station and both stations in Winnipeg was decreased from 27.3% to 1.1% using the Winnipeg International Airport. The remaining 1.1% of measurements were infilled using the regression between the target station and the station at The Forks to achieve a complete dataset.

**Temperature**

Linear regression between the Portage Southport Airport station (target station to be gap-filled) and the Winnipeg International Airport and The Forks stations (data sources) was used to reconstruct the temperature time series. Regression-based techniques are usually used for reconstructing temperature records (Tardivo and Berti, 2014). The method was chosen because it is robust with regard to extreme events (i.e. very high or very low temperatures) or local effects (i.e. spatial variability) (Ramos-Calzado et al., 2008; Hutchinson et al., 2009). Potential problems with temperature lapse rate due to elevation changes (Henn et al., 2013) were negligible in this area due to its flat topography (Graveline and Larter, 2006). The coefficient of determination ($R^2$) between the Portage Southport Airport and either station in Winnipeg was 0.98. Due to the similarity in $R^2$, both neighbouring stations in Winnipeg (which were 8.3 km apart from each other) were considered mutually equivalent. However, the station at the Winnipeg International Airport was given priority due to the shorter distance to the target station (Table 1). The proportion of missing temperature data in the target station was decreased from 27.3% to 1.1% using the Winnipeg International Airport. The remaining 1.1% of measurements were infilled using the regression between the target station and the station at The Forks to achieve a complete dataset.

**Table 1** Weather stations surrounding the La Salle River watershed shown in Figure 2

| Station | WMO ID | Network | Location | Distancea (km) |
|---------|--------|---------|----------|---------------|
| A       | N/A    | MAFRD   | Portage La Prairie East | 24.9 |
| B       | N/A    | ECCC    | Canadian Dept. of Agric., Portage La Prairie | 26.0 |
| C       | N/A    | ECCC    | Marquette | 9.9 |
| D       | 71851  | ECCC    | Southport Airport, Portage La Prairie | 26.6 |
| E       | 71147  | ECCC    | University of Manitoba Research Station, Carman | 51.0 |
| F       | 71852  | ECCC    | International Airport, Winnipeg | 47.9 |
| G       | 71579  | ECCC    | The Forks, Winnipeg | 56.1 |

WMO, World Meteorological Organization; N/A, not available; ECCC, environment and climate change Canada; MAFRD, manitoba agriculture, food and rural development.

*aDistances measured from the geometric center of the study area.*
### Table 2

Description of the datasets used in the analysis, their sources, record length, frequency, format, and gap-filling method used

| Weather element | Data use          | Station*          | Network | Climate ID | Period available       | Frequency | Availability | File type | Gap-filling method |
|-----------------|-------------------|-------------------|---------|------------|------------------------|-----------|--------------|-----------|-------------------|
| Temperature     | Parent station    | PLP-SP            | ECCC    | 5012320    | January 1953—Present   | Hourly    | Websiteb     | CSV       | LR                |
|                 | Gap-filling       | WNP-AR            | ECCC    | 5023222    | January 1953 — Present | Hourly    | Websiteb     | CSV       | —                 |
|                 | Gap-filling       | WNP-FK            | ECCC    | 5023262    | November 1999 — Present| Hourly    | Websiteb     | CSV       | —                 |
| RH              | Parent station    | PLP-SP            | ECCC    | 5012320    | January 1953 — Present | Hourly    | Websiteb     | CSV       | LR                |
|                 | Gap-filling       | WNP-AR            | ECCC    | 5023222    | January 1953 — Present | Hourly    | Websiteb     | CSV       | —                 |
|                 | Gap-filling       | WNP-FK            | ECCC    | 5023262    | November 1999 — Present| Hourly    | Websiteb     | CSV       | —                 |
| Wind speed      | Parent station    | PLP-SP            | ECCC    | 5012320    | January 1953 — Present | Hourly    | Websiteb     | CSV       | LR                |
|                 | Gap-filling       | WNP-AR            | ECCC    | 5023222    | January 1953 — Present | Hourly    | Websiteb     | CSV       | —                 |
|                 | Gap-filling       | WNP-FK            | ECCC    | 5023262    | November 1999 — Present| Hourly    | Websiteb     | CSV       | —                 |
| Solar radiation | Parent station    | WNP-AR            | ECCC    | 5023222    | January 1953 — Present | Hourly    | Requestc     | CSV       | T/LTM             |
|                 | Gap-filling       | The Point         | UoM     | —          | January 1995 — Present | Hourly    | Websiteb     | ZIP       | —                 |
| Precipitationf  | Parent station    | Marquette         | ECCC    | 5021695    | November 1969 — February 2016 | Daily    | Websitee     | TAR       | —                 |

ECCC: Environment and Climate Change Canada; UoM: University of Manitoba; CSV: comma separated values; ZIP: compressed format; TAR: compressed format; LR: Linear regression; T: transplanting; LTM: Long-term daily average.

*Station acronyms are PLP-SP: Southport Airport, Portage La Prairie; WNP-AP: International Airport, Winnipeg; WNP-FK: The Forks, Winnipeg.

http://climate.weather.gc.ca/historical_data/search_historic_data_e.html.

Data can be purchased from Environment and Climate Change Canada, Climate Services.

https://home.cc.umanitoba.ca/~adam/index.html.

https://open.canada.ca/data/en/dataset/d8616c52-a812-44ad-8754-7bco8d8ec305.

Precipitation did not need to be gap-filled since adjusted and homogenized Canadian climate (AHCC) data were used.
variability when compared to other weather elements such as precipitation. Using the station at the Winnipeg International Airport in the first gap-filling step, the missing records decreased from 29.8% to 0.03%. The remaining missing records were infilled using the station at The Forks.

**Wind speed**
Linear regression was also employed to reconstruct the wind speed dataset using the same stations used for temperature and relative humidity. However, the correlation between those stations for wind speed was weaker than those found for temperature and relative humidity (i.e. $R^2 = 0.48$ between Portage Southport Airport and the Winnipeg International Airport; $R^2 = 0.34$ between Portage Southport Airport and The Forks station). Despite the weaker correlations, this method was preferred over the typical approach used to address missing data in weather records which is to transplant data from a nearby region to the area of interest (Liu et al., 2013; Pomeroy et al., 2013) or using multivariate regression including stations further away from the target station. The missing records decreased from 27.4% to 1.2% after the infilling using the Winnipeg International Airport. The dataset was completed by gap-filling the remaining missing records using the station at The Forks.

**Solar radiation**
Since the station at the Winnipeg International Airport was the only location with long-term measurement of solar radiation, data used for gap-filling had to be acquired from The Point research station in the University of Manitoba. This station is located 13.7 km from the Winnipeg International Airport. The missing data were replaced directly with data from the Point station due to proximity (Liu et al., 2013; Pomeroy et al., 2013). After gap-filling, there were 6% of the records still missing, which were replaced with the long-term average (1990–2013) for that particular Julian day. This approach was preferred over more complex gap-filling methodologies for solar radiation that rely on derivation of coefficients as well as temperature and precipitation information (Hunt et al., 1998; Aladenola and Madramootoo, 2013). The long-term average was deemed suitable due to the small proportion of the dataset left to be infilled and because solar radiation is a very predictable variable for specific days. Besides, the accuracy of complex methods can be limited during specific seasons. For example, the performance of a solar radiation estimation method in Canada that incorporates temperature and precipitation was poor during the late fall and winter periods and acceptable during the spring and summer (Jong and Stewart, 1993). In fact, the performance of temperature and/or precipitation-based methods was poorer in Canada ($R^2 = 0.57$) when compared to Australia ($R^2 = 0.79$) and Europe ($R^2$ ranging from 0.81 to 0.85) (Aladenola and Madramootoo, 2013).

**Precipitation**
There were no missing records in the precipitation dataset since the adjusted precipitation dataset at a daily time-step was used to derive hourly precipitation. Thus, no gap-filling procedure was applied to this variable. However, the downscaled dataset was graphically assessed against hourly observations from the closest station (i.e. Portage Southport Airport) for the period available (i.e. 2004 onwards).

### 3.1.3 Precipitation disaggregation
Disaggregation of precipitation to an hourly time-step was performed using HyetosMinute (Kossieris et al., 2016), which is an R package for the temporal stochastic simulation of rainfall process at fine time scales based on Bartlett–Lewis rectangular pulses rainfall model (Koutsoyiannis and Onof, 2001). Poisson-cluster models such as the Bartlett–Lewis can be used for point-precipitation simulation while keeping the statistical properties of the process through a wide range of aggregation levels (Velghe et al., 1994). A detailed description of the model including its parameters is given by Velghe et al. (1994) and Koutsoyiannis and Onof (2001).

The parameters needed for model disaggregation have to be estimated from hourly records. A six-parameter model was used for disaggregation (the model can also be run using seven parameters). Since the Marquette weather station did not have precipitation records in an hourly time-step, the hourly records from the Portage Southport Airport station were used for parameter estimation and assessment of the downscaled dataset. This station was selected because it was the closest station with available data. Monthly parameters were estimated using the evolutionary annealing-simplex method (Efstratiadis and Koutsoyiannis, 2005) in HyetosMinute. The estimated parameters were used as inputs to the DisagSimul function in HyetosMinute to disaggregate the daily precipitation records into an hourly time-step.

### 3.2 Streamflow data
Daily streamflow observations between 1990 and 2013 were obtained from the hydrometric data (HYDAT) database (Environment and Climate Change Canada, 2013) for the Water Survey of Canada (WSC) gauging station 05OG008 (La Salle River near Elie; Figure 2) located at the outlet of the watershed’s sub-catchment. Data collection at this location was seasonal from 1990 to 1996 and has been continuous from 2002 to present. Only flow data were available from HYDAT for the period prior to 1996, while flow and water level were both recorded from 2002 onwards. Water level data were not included in the present...
dataset because they are not usually required for hydrological modelling, but the data are available in the publically available HYDAT database. The annual monitoring period for this station spans from 1 March to 31 October, with no data available during winter months. A gap in available flow data exists between flooding in 1997 and instrument replacement in 2001. Notes in the HYDAT metadata pertaining to 2004 and 2008 indicate equipment malfunctions resulting in loss of data. For this reason, the periods from 1997 to 2001, 2004, and 2008 are not included in the dataset presented here.

3.3 Water chemistry data

Prior to the initiation of sampling at higher temporal frequency in 2013, grab water samples were collected weekly at a water control structure located at the hydrometric gauging site for the watershed. In 2013 samples were collected during snowmelt and storm events at a higher frequency using an auto sampler (Sigma 900). Timing of sample collection from 2009 to 2012 was designed to provide seasonal coverage (multiple samples monthly) with some higher frequency sample collection during high flow events. Frequency was increased in 2013 to provide coverage of each runoff event hydrograph with samples on rising, falling, and near peak.

From 2009 to 2012, grab samples were collected, placed on ice, and shipped to the Environment and Climate Change Canada National Laboratory for Environmental Testing (NLET) in Saskatoon, Saskatchewan for analysis using standard analytical techniques at this accredited laboratory. Samples were filtered (0.45 μm pore size) on arrival at the laboratory (within 4 days of collection) and the particulate material collected on the filter was analysed for dissolved N and P. Resulting filtered and unfiltered samples were kept refrigerated until being analysed for P (within 28 days of collection) and N (within 20 days of collection). Total and dissolved N were determined at NLET as nitrate in solution following alkaline potassium persulphate digestion. Total and dissolved P were measured as orthophosphate in solution following sulphuric acid/persulfate digestion.

Samples collected in 2013 were kept on ice until filtered (0.7 μm nominal pore size pre-combusted glass fiber filter; GFF) and frozen as filtered or unfiltered aliquots within 48 hr of sample collection. Water samples were analysed in the hydrology laboratory at the Brandon Research and Development Centre of Agriculture and Agri-Food Canada in Brandon, Manitoba. Comparison of dissolved N and P for a variety of samples filtered to 0.45 and 0.7 μm indicated no significant difference (unpublished data). Analyses for TP were completed by sulfuric acid/persulfate digestion followed by colorimetric analysis using the ascorbic acid method. TN calculated as the sum of particulate and dissolved N. Analyses for dissolved N were completed by the combustion method using a Shimadzu TOC-VCSH analyzer and of particulate N by combustion using a Thermo Scientific Flash 2000 CHNS/O elemental analyzer. Coefficient of variation for replicates with each analysis for TP, dissolved N, and particulate N was generally less than 5%, internal check standards created over the range of observed concentrations were within 10% of expected values, and external quality control standards are run periodically in the AAFC laboratory to ensure values fall within range stated on certificate of analysis.

4 DATASET OVERVIEW

4.1 Weather dataset

4.1.1 Analysis of variables

Temperature

The overall temperature distribution followed the expected range of the Canadian Prairies (Figure 3). The seasonal temperature values for the 1990–2013 period were also in general agreement with published values for the La Salle River watershed (La Salle Redboine Conservation District, 2007). However, there seems to be a slight trend towards warmer temperatures when the data are analysed annually and seasonally. The reported mean annual temperature is around 2.5°C, while this value for the present dataset was 3.5°C. Similarly, reported mean annual temperatures during winter and summer were, respectively, −13 and 16°C, while those values calculated from the dataset were −10.7 and 17.2°C, respectively. This small discrepancy between the present dataset and reported values likely arise from the larger geographic area encompassed by the former, which refers to the entire La Salle River watershed that extends eastward and is inherently more spatially variable due to the inclusion of a number of weather stations. The Mann–Kendall trend tests performed using the R package Kendall (McLeod, 2011) indicated no trend in either annual minimum (p = 0.58), annual maximum (p = 0.46), or annual average (p = 0.60) temperatures, possibly due to the short period analyzed, which contrasts to the trends found for longer periods (i.e. 1950–2003 and 1900–2003) (Vincent and Mekis, 2006).

Relative humidity

Relative humidity averaged 73.1% (SD = 16.8%) over the 1990–2013 period. Seasonally, RH tends to be higher and more homogenous (i.e. narrower range) in cold seasons (Table 3) due to cold temperatures that lower the saturation capacity of the atmosphere. For example, 46.8% of the RH values in cold seasons (i.e. winter and fall) were above 80%, while only 36.6% of the values were above this threshold in
warm seasons (i.e. spring and summer). The boxplot of the annual RH average for different seasons illustrate this difference, with the seasonal RH average being the lowest in the spring, increasing in the summer to reach its maximum at fall and winter (Figure 4).

Wind speed
The statistical properties of wind speed were quite similar over the different seasons (Figure 5a). The annual average wind speed was 4.3 m/s, while it was 4.5 m/s during the winter, spring, and fall but dropped to 3.9 m/s during the summer. The Mann–Kendall test indicated no trend in annual average wind speed ($p = 0.67$; Figure 5b), although studies in the Canadian Prairies indicated decreasing trends in most station between April and October (Burn and Hesch, 2007), which is in agreement with other studies that also suggest a decrease in annual wind speed in the region (Hugenholtz and Wolfe, 2005). Restricting the present analysis to those months only also resulted in no trend despite a decrease in $p$ value ($p = 0.21$). Wind direction was not included in this analysis due to the large uncertainty, as discussed in section 3.2. However, the long-term prevailing wind directions in the area is northwest or north; the most frequent wind direction year-around is northwest, except between March and May (north prevailing winds) and June (northeast prevailing winds) (Page, 2011).

Solar radiation
The long-term trend of solar radiation data followed the typical curve reported for Western Canada (Hare, 1997), with peak daily average daily values around 250 W/m$^2$ in July (Figure 6). However, hourly solar radiation reached values as high as 1,003 W/m$^2$. During the winter, values ranged from

---

**TABLE 3** Statistical properties of relative humidity

| Season | Relative humidity (%) | Mean (SD) | Range | $n$ |
|--------|-----------------------|-----------|-------|-----|
| Winter | 75.8 (12.8)           | 80        | 51,983|
| Spring | 67.0 (20.1)           | 86        | 52,416|
| Summer | 72.2 (17.4)           | 86        | 52,992|
| Fall   | 77.2 (14.6)           | 84        | 52,992|

**FIGURE 3** Graphical description of the temperature dataset showing the boxplot of temperature values between 1990 and 2013 (a), annual mean temperature (b), annual minimum temperature (c), and annual maximum temperature (d).

**FIGURE 4** Statistical properties of relative humidity showing the boxplots of annual and seasonal values.
40 to 50 W/m², which is in agreement with ranges in southern Canada between 30 and 50 W/m² (Hare, 1997).

Precipitation
The majority of the 210,335 records (i.e. 94.5%) corresponded to dry hours (i.e. no precipitation recorded). These records were removed from the dataset and statistics for the dataset were computed; thus, the data presented here pertain to censored data (i.e. wet days), which is a common procedure used in assessments of precipitation. Investigations of precipitation trends in the Canadian Prairies have defined wet days as those with precipitation above 1 mm/day (Shook and Pomeroy, 2012). Although smaller precipitation events are possible, such events are usually censored to avoid possible inhomogeneity in the definition of trace precipitation or lack of correction of trace amounts (Akinremi et al., 1999). Thus, a conservative approach was adopted in this study and a threshold of 1 mm/day or 0.042 mm/hr was used to select rain events used in the statistical calculations. The characteristic precipitation pattern of the Canadian Prairies with high frequency of small precipitation events (Akinremi et al., 1999; Shook and Pomeroy, 2012) was observed in the dataset, with 81.3% of the hourly precipitation being equal or less than 2 mm/hr. The average precipitation was 1.35 mm, while the median was 0.57 mm. Out of 10,356 wet hours, only 128 and 28 events were larger than 10 and 20 mm/hr, respectively. The Mann–Kendall test indicated a decreasing trend in precipitation amounts (p < 0.05). This result is consistent with other studies in the Canadian Prairies that report an increase in the number of low-intensity events (Akinremi et al., 1999). The prevalence of small-magnitude events was also observed seasonally (Figure 7a). The interquartile range was consistently less than 3 mm/hr (Figure 7b).

4.1.2 Uncertainty and dataset assessment
The uncertainty analysis for hourly temperature indicated that nearly half (i.e. 47.3%) corresponded to dry hours (i.e. no precipitation recorded). These records were removed from the dataset and statistics for the dataset were computed; thus, the data presented here pertain to censored data (i.e. wet days), which is a common procedure used in assessments of precipitation. Investigations of precipitation trends in the Canadian Prairies have defined wet days as those with precipitation above 1 mm/day (Shook and Pomeroy, 2012). Although smaller precipitation events are possible, such events are usually censored to avoid possible inhomogeneity in the definition of trace precipitation or lack of correction of trace amounts (Akinremi et al., 1999). Thus, a conservative approach was adopted in this study and a threshold of 1 mm/day or 0.042 mm/hr was used to select rain events used in the statistical calculations. The characteristic precipitation pattern of the Canadian Prairies with high frequency of small precipitation events (Akinremi et al., 1999; Shook and Pomeroy, 2012) was observed in the dataset, with 81.3% of the hourly precipitation being equal or less than 2 mm/hr. The average precipitation was 1.35 mm, while the median was 0.57 mm. Out of 10,356 wet hours, only 128 and 28 events were larger than 10 and 20 mm/hr, respectively. The Mann–Kendall test indicated a decreasing trend in precipitation amounts (p < 0.05). This result is consistent with other studies in the Canadian Prairies that report an increase in the number of low-intensity events (Akinremi et al., 1999). The prevalence of small-magnitude events was also observed seasonally (Figure 7a). The interquartile range was consistently less than 3 mm/hr (Figure 7b).
interval, with 55.5% of the outside values below the lower confidence limit.

For wind speed, 46.5% of the gap-filled values were within the 95% confidence interval, resulting in 14.6% of the complete time series having higher uncertainty. Values outside the confidence interval tended to be more concentrated in the spring and summer (i.e. 30.0% and 29.6% of the records, respectively), with lower prevalence in the winter and fall (i.e. 21.4% and 19.0% of the records, respectively). There was also a slight trend to overestimation with 65.7% of the values outside the confidence interval occurring above the upper confidence limit.

For solar radiation, 52.2% of the gap-filled values were within the 95% confidence interval, resulting in 30.1% of the reconstructed time series having higher uncertainty. Values outside the confidence interval tended to be more concentrated in the spring and summer (i.e. 30.0% and 29.6% of the records, respectively), with lower prevalence in the winter and fall (i.e. 21.4% and 19.0% of the records, respectively). There was also a slight trend to overestimation with 55.7% of the values outside the confidence interval occurring above the upper confidence limit.

In general, the uncertainty for temperature, relative humidity, wind speed, and solar radiation was not unexpected because there is a 5% chance of each given record to occur outside the 95% confidence interval as the interval was estimated for every hour of the year. As a result, the uncertainty compounds as the time series length increases. While it is acknowledged that this uncertainty could be potentially constrained by using other methodological approaches (e.g., multivariate regression), the impact of such approaches is unclear due to (a) gaps in the records of the other variables used in the process, which could be likely missing for the same station and time period, and (b) the lack of stations with similar datasets in close proximity.

The assessment of the downscaled precipitation indicated that the downscaling process was able to satisfactorily capture the characteristics of the observed precipitation dataset. Small, moderate, and large precipitation events were well represented, except the very large events (Figure 7a). The median and interquartile ranges were generally well represented for all the seasons, indicating that most of the events in the region are of small magnitude (Figure 7b). There was no added uncertainty due to gap filling as the adjusted precipitation dataset at daily time-step used for downscaling was complete.

4.2 | Streamflow data

The characteristics of the streamflow dataset were typical of the Canadian Prairies with peak discharge during the spring due to snowmelt runoff (Shook and Pomeroy, 2010). Peak discharges occurred in April in 12 out of 17 years with good data (Table 4). The average discharge between 1990 and 2013 for years with good data is also higher in April (i.e. 2.3 m³/s; Figure 8a). Two peak discharges occurred in May, while one peak discharge occurred in March and one in June. An unusual peak discharge occurred in July, which is not typical for this region. Inspection of the streamflow data in July of 2005 suggests an anomaly in the hydrograph, with a very sharp rise and a ‘flat top’, which resembles a culvert outflow hydrograph or some other form of upstream flow restriction (Figure 8b). This type of behaviour is not expected and indicates potential issues with the streamflow
data since years with larger peak flows such as 2006 did not show these anomalies (Figure 8c). Removing the year of 2005 from the dataset results in an average monthly hydrograph that resembles the expected trend in the prairies, where peak flow occurs in April and lower flows occur over the summer months (Figure 8a). This anomaly in July of 2005 actually represented disinformation for model assessment (Beven, 2011) and was removed from model assessments in the sub‐catchment (Cordeiro et al., 2017c). Since the datasets presented here were specifically developed for forcing and assessing hydrological modelling, the period between June 28th and July 31st of 2005 has been removed from the streamflow dataset. This period has been flagged as ‘Removed’ in the respective dataset accompanying this manuscript. Readers interested on the complete streamflow time‐series are referred to the HYDAT database (Environment and Climate Change Canada, 2013).

Another feature of the streamflow data is the strong correlation between peak discharge and annual discharge. The overall correlation between these two variables (including the year of 2005) is very good ($R^2 = 0.68$), despite several aspects influencing stream discharge such as antecedent conditions (e.g. soil moisture status in the fall), snow water equivalent (SWE), which integrates winter precipitation and snow dynamics by snow blowing and sublimation, water withdrawals and flow diversions, as well as spring snowmelt dynamics (e.g. quick vs slow melt). When the year of 2005 is excluded, the correlation improves even more ($R^2 = 0.90$), indicating that most of the annual discharge occurs during spring and is associated with snowmelt runoff. Assessment of water yield for different years confirms these results (Figure 8d; July 2005 removed). The water yield during snowmelt corresponds to most of the annual water yield in the study area. The exceptions to this trend are two snowfall dry years (e.g. 1994, 2012) and years with excessively wet summers (i.e. 2010). The average water yield in the study area is 64 mm, out of which 72% occurs during snowmelt.

4.3 | Water chemistry data

The average concentrations of TN, TDN, TP, and TDP between 2009 and 2013 were, respectively, 1.54, 1.35, 0.56, and 0.49 mg/L. The TN and TP concentrations were in agreement with values found in previous studies at the same location (WSC gauging station 05OG008) between 1995 and 1996, which report annual TN concentrations of 1.67 mg/L and annual TP concentrations of 0.56 mg/L (Corriveau et al., 2013). On average, 88% and 84% of the total nitrogen and phosphorus were in dissolved form, which explains the similar temporal trend between total and dissolved forms of both nitrogen and phosphorus (Figure 9). High proportion of dissolved forms of nutrient is in agreement with water chemistry results published for the La Salle River and other watersheds in Manitoba. The TDP/TP ratio in the La Salle watershed ranged from 0.25 to 0.99 (Corriveau et al., 2013), although values in the higher end of the spectrum were more frequent. McCullough et al. (2012) also report TDP corresponding to 81% of TP in the lower reaches of the main stem of the La Salle River over the course of a large snowmelt flood event in April–June 2009.

### Table 4 Streamflow characteristics of the study area

| Year | Discharge Annual$^1$ (m³) | Peak (m³/s) | Peak date | Ice conditions |
|------|---------------------------|------------|----------|----------------|
| 1990 | 5.32 × 10⁶                | 2.4        | April 4th| March 1st — April 21st |
| 1991 | 5.91 × 10⁶                | 2.1        | April 7th| March 1st — April 13th |
| 1992 | 1.21 × 10⁷                | 6.7        | April 10th| March 1st — April 12th |
| 1993 | 1.25 × 10⁷                | 5.6        | April 07th| March 1st — April 10th |
| 1994 | 3.33 × 10⁶                | 0.7        | April 10th| March 1st — April 16th |
| 1995 | 1.15 × 10⁷                | 5.0        | March 31st| March 1st — April 15th |
| 1996 | 1.87 × 10⁷                | 13.5       | April 29th| March 1st — April 28th |
| 2002 | 1.94 × 10⁶                | 1.6        | April 16th| March 1st — April 16th |
| 2003 | 3.49 × 10⁶                | 2.1        | April 2nd | March 1st — April 09th |
| 2005 | 3.96 × 10⁷                | 11.1       | July 1st | March 1st — April 04th |
| 2006 | 2.18 × 10⁷                | 16.5       | April 10th| March 1st — April 07th |
| 2007 | 7.26 × 10⁶                | 4.6        | April 12th| March 1st — April 05th |
| 2009 | 1.69 × 10⁷                | 13.3       | April 17th| March 1st — April 16th |
| 2010 | 2.09 × 10⁷                | 10.7       | June 1st | March 1st — April 05th |
| 2011 | 2.84 × 10⁷                | 15.7       | April 13th| March 1st — April 13th |
| 2012 | 4.89 × 10⁶                | 2.5        | May 29th | March 1st — March 24th |
| 2013 | 1.44 × 10⁷                | 9.4        | May 04th | March 1st — May 1st |

$^1$Annual discharge from March 1st to October 31st.
The plot of the concentrations over the monitoring period (Figure 9) indicates that concentrations of all analytes increased in wet years (2010 and 2011) and decreased in dry years (2012). When considered across years, a wide range of TN and TP concentrations were observed for lower flows, but concentrations at high flow were generally elevated. Peak values in 2009 were missed since monitoring started in August, after the spring snowmelt. The monthly trends show that concentration of TN peak in March (3.3 mg/L), increasing from lower values in October (1.32 mg/L; Figure 10a). When streamflow discharge is at its peak in April, TN concentrations are already decreasing (2.6 mg/L). Concentrations of TP are also lower in October (0.37 mg/L) than in March (0.71 mg/L) and start to decrease...
by the time of peak discharge (0.54 mg/L). However, different that TN concentrations that peaks in March, TP concentrations peak in June (0.72 mg/L), remain relatively stable until July (0.68 mg/L), and start to decrease towards the fall. Concentration-discharge (C-Q) relationships are not very distinct in the study area (Figure 10b,c).

5 | CONCLUSIONS

The weather, streamflow, and water chemistry datasets presented and discussed in this work represent an effort to develop a long-term record not usually available in the Canadian Prairies and particularly in the Red River Valley within the Lake Winnipeg Basin, where the monitoring networks are sparse, and records contain frequent gaps. Such datasets represent a crucial input for physically based simulations of hydrological processes, which remains one of the major limitations for this type of modelling in this basin. Although compiled from available data sources, thorough quality control and gap-filling procedures performed on the original datasets represent added value in terms of completeness and accuracy. The methodology used to develop complete datasets consisted of drawing the best data available for specific weather elements from the closest stations and using linear regression and direct transplantation from close-by stations to infill gaps in the records. While more complex gap-filling methods were available, a deliberate attempt was made to keep the methods simple enough to promote replication of this effort in other regions of the Lake Winnipeg Basin. The uncertainty introduced by the gap-filling procedure was relatively higher for less than 15% of the record length for all variables except solar radiation, which showed higher uncertainty for 30% of the records. Due to the lack of hourly data that met the quality criteria, daily precipitation data had to be disaggregated into an hourly time step using a Poisson-cluster model. Streamflow data were only available in a daily time-step and no attempt was made to develop records in an hourly time-step since results of long-term environmental and modelling studies are usually summarized at this time scale or coarser. Overall, the streamflow data were much more consistent than the weather data, but entire years were missing in the records due to no data collection. Years with data of dubious quality were also removed from the dataset due to the uncertainty created for long-term environmental and modelling studies. Although short in length compared to weather and streamflow data (i.e. 1990–2013 vs 2009–2013), water chemistry records represent an important source of data for modelling studies with focus on nutrient export and impact to downstream water bodies in a region where eutrophication is of significant environmental concern.

OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at https://doi.org/10.23684/odi-2017-00957. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki.

ACKNOWLEDGEMENTS

Collaboration in data preparation and tabulation with Dr. Zhiqiang Yu and Dr. Glenn Benoy is greatly appreciated.

ORCID

Marcos R. C. Cordeiro https://orcid.org/0000-0002-1535-5096
Jason A. Vanrobaeys https://orcid.org/0000-0002-0039-6397

REFERENCES

Akinremi, O.O., McGinn, S.M. and Cutforth, H.W. (1999) Precipitation trends on the Canadian Prairies. Journal of Climate, 12, 2996–3003.
Aladenola, O.O. and Madramootoo, C.A. (2013) Evaluation of solar radiation estimation methods for reference evapotranspiration estimation in Canada. *Theoretical and Applied Climatology*, 118, 377–385.

Allen, R.J. and DeGaetano, A.T. (2001) Estimating missing daily temperature extremes using an optimized regression approach. *International Journal of Climatology*, 21, 1305–1319.

Beven, K.J. (2011) *Rainfall-Runoff Modelling: The Primer*. West Sussex, UK: Wiley.

Burn, D.H. and Hesch, N.M. (2007) Trends in evaporation for the Canadian Prairies. *Journal of Hydrology*, 363, 61–73.

Coleman, H.W. and Steele, W.G. (2018) *Experimentation, Validation, and Uncertainty Analysis for Engineers*. Hoboken, NJ, USA: John Wiley & Sons.

Cordeiro, M.R.C., Wilson, H.F., Vanrobaeys, J., Pomeroy, J.W. and Fang, X.; The Red-Assiniboine Project Biophysical Modelling, T. (2017c) Simulating cold-region hydrology in an intensively drained agricultural watershed in Manitoba, Canada, using the Cold Regions Hydrological Model. *Hydrology and Earth System Sciences*, 21, 3483–3506.

Corriu, J., Chambers, P. and Culp, J. (2013) Seasonal variation in nutrient export along streams in the Northern Great Plains. *Water, Air, & Soil Pollution*, 224, 1–16.

Efstratiadis, A. and Koutsoyiannis, D. (2005) The multiobjective evolutionary annealing-simplex method and its application in calibrating hydrological models. *European Geosciences Union General Assembly*, 7, 04593. Geophysical Research Abstracts, Vienna.

Ellis, C.R., Pomeroy, J.W., Brown, T. and MacDonald, J. (2010) Simulation of snow accumulation and melt in needleleaf forest environments. *Hydrology and Earth System Sciences*, 14, 925–940.

Environment and Climate Change Canada. (2015) About fresh water resources techniques of water resources investigations, Book 4, Chapter A3. 522.

Henn, B., Raleigh, M.S., Fisher, A. and Lundquist, J.D. (2013) A comparison of methods for filling gaps in hourly near-surface air temperature data. *Journal of Hydrometeorology*, 14, 929–945.

Hopkinson, R.F., McKenney, D.W., Milewski, E.J., Hutchinson, M.F., Papadopol, P. and Vincent, L.A. (2011) Impact of aligning climatological day on gridding daily maximum-minimum temperature and precipitation over Canada. *Journal of Applied Meteorology and Climatology*, 50, 1654–1665.

Hugenholtz, C.H. and Wolfe, S.A. (2005) Recent stabilization of active sand dunes on the Canadian prairies and relation to recent climate variations. *Geomorphology*, 68, 131–147.

Hunt, L.A., Kuchar, L. and Swanton, C.J. (1998) Estimation of solar radiation for use in crop modelling. *Agricultural and Forest Meteorology*, 91, 293–300.

Hutchinson, M.F., McKenney, D.W., Lawrence, K., Pedlar, J.H., Hopkinson, R.F., Milewski, E. et al. (2009) Development and testing of canada-wide interpolated spatial models of daily minimum–maximum temperature and precipitation for 1961–2003. *Journal of Applied Meteorology and Climatology*, 48, 725–741.

Jong, R.D. and Stewart, D.W. (1993) Estimating global solar radiation from common meteorological observations in western Canada. *Canadian Journal of Plant Science*, 73, 509–518.

Kemp, W.P., Burnell, D.G., Everson, D.O. and Thomson, A.J. (1983) Estimating missing daily maximum and minimum temperatures. *Journal of Climate and Applied Meteorology*, 22, 1587–1593.

Kim, J.-W. and Pachepsky, Y.A. (2010) Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. *Journal of Hydrology*, 394, 305–314.

Kossieris, P., Tyralis, H., Koutsoyiannis, D., Makropoulos, C. and Efstratiadis, A. (2016) HyetosMinute: A package for temporal stochastic simulation of rainfall at fine time scales. R package version 2.0. Available at: http://www.itia.ntua.gr/en/softinfo/3/ [Accessed 25 November 2017].

Koutsoyiannis, D. and Onof, C. (2001) Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *Journal of Hydrology*, 246, 109–122.

La Salle Redboine Conservation District. (2007) La Salle river watershed - state of the watershed report, Holland, Manitoba; 296.

Liu, S., Mo, X., Liu, C., Xia, J. and Zhao, W. (2013) How to maximize the predictive value of available data in ungauged basins? – Chinese lesson. In: Pomeroy, J.W. and Whitfield, P.H. (Eds.) *Putting Prediction in Ungauged Basins into Practice*. Ottawa, ON: Canadian Water Resources Association.

Mahmood, T.H., Pomeroy, J.W., Wheater, H.S. and Baulch, H.M. (2017) Hydrological responses to climatic variability in a cold agricultural region. *Hydrological Processes*, 31, 854–870.

Mayer, B. and Wassenaar, L.I. (2012) Isotopic characterization of nitrate sources and transformations in Lake Winnipeg and its contributing rivers, Manitoba, Canada. *Journal of Great Lakes Research*, 38 (Supplement 3), 135–146.

McCullough, G.K., Page, S.J., Hesslein, R.H., Stainton, M.P., Kling, H.I., Saik, AG et al. (2012) Hydrological forcing of a recent trophic
Cordeiro, R.C., Vanrobaeys JA, Wilson HF. Long-term weather, streamflow, and water chemistry datasets for hydrological modelling applications at the upper La Salle River watershed in Manitoba, Canada. Geosci Data J. 2019;6:41–57. https://doi.org/10.1002/gdj3.67