Journal of Hydrology 561 (2018) 688–701

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

# Application of SNODAS and hydrologic models to enhance entropy-based snow monitoring network design

Jongho Keum[a,*], Paulin Coulibaly[a,b], Tara Razavi[a], Dominique Tapsoba[c], Adam Gobena[d], Frank Weber[d], Alain Pietroniro[e]

[a] Department of Civil Engineering, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
[b] School of Geography and Earth Sciences McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
[c] Institut de Recherche d'Hydro-Québec (IREQ), 1800, Lionel-Boulet, Varennes, Québec J3X 1S1, Canada
[d] BC Hydro, Hydrology, 6911 Southpoint Drive, Burnaby, British Columbia V3N 4X8, Canada
[e] National Hydrology Research Centre, Meteorological Service of Canada, 11 Innovation Blvd, Saskatoon, Saskatchewan S7N 3H5, Canada

ARTICLE INFO

This manuscript was handled by G. Syme, Editor-in-Chief, with the assistance of Ashok Mishra, Associate Editor

*Keywords:*
Hydrometric network
Snow monitoring network
Network design
Information theory
Hot-spot map
Data quantization

ABSTRACT

Snow has a unique characteristic in the water cycle, that is, snow falls during the entire winter season, but the discharge from snowmelt is typically delayed until the melting period and occurs in a relatively short period. Therefore, reliable observations from an optimal snow monitoring network are necessary for an efficient management of snowmelt water for flood prevention and hydropower generation. The Dual Entropy and Multiobjective Optimization is applied to design snow monitoring networks in La Grande River Basin in Québec and Columbia River Basin in British Columbia. While the networks are optimized to have the maximum amount of information with minimum redundancy based on entropy concepts, this study extends the traditional entropy applications to the hydrometric network design by introducing several improvements. First, several data quantization cases and their effects on the snow network design problems were explored. Second, the applicability the Snow Data Assimilation System (SNODAS) products as synthetic datasets of potential stations was demonstrated in the design of the snow monitoring network of the Columbia River Basin. Third, beyond finding the Pareto-optimal networks from the entropy with multi-objective optimization, the networks obtained for La Grande River Basin were further evaluated by applying three hydrologic models. The calibrated hydrologic models simulated discharges using the updated snow water equivalent data from the Pareto-optimal networks. Then, the model performances for high flows were compared to determine the best optimal network for enhanced spring runoff forecasting.

## 1. Introduction

Freshwater is one of the most valuable natural resources, and spring freshet is a major source of freshwater in subarctic regions. Snow falls during the entire winter season, but water from snowmelt is delayed until the melting period in spring and may occur in a relatively short time. This snowmelt can cause severe flood events but also can be utilized to replenish hydropower reservoirs if well forecasted. The importance of water monitoring is well documented by researchers, water resources managers and decision makers (Mishra and Coulibaly, 2009; Yang and Burn, 1994). The purpose of snow monitoring networks is to provide a comprehensive picture of snow cover in the monitored areas, including extents, depths and the variability of snow cover. Such information is essential for flood forecasting, streamflow forecasting for reservoir operation planning, and other activities. Despite the importance of hydrologic information, declining trends in the spatial density of hydrometric networks were identified in many regions, such as the United States (Mishra and Coulibaly, 2009; U.S. Geological Survey, 1999), Canada (Brown et al., 2000; Mishra and Coulibaly, 2009; Pilon et al., 1996; Pyrce, 2004), and Africa (Rodda et al., 1993). Specifically, the limitation of the Canadian National Hydrometric Networks (CNHN) in terms of network density, has been well documented (Coulibaly et al., 2013; Mishra and Coulibaly, 2010). In general, given the large size of the country, the CNHN has been below the WMO guidelines (World Meteorological Organization, 2008) for the minimum density of monitoring stations. This implies either that there are not an adequate number of stations, and/or that their locations are not ideal, or both. Most of the stations are located in the southern parts of the country and at lower elevations (Brown et al., 2000; Coulibaly et al., 2013). Hence, hydrometric networks should be designed with careful

* Corresponding author.
*E-mail address:* jkeum@mcmaster.ca (J. Keum).

https://doi.org/10.1016/j.jhydrol.2018.04.037
Received 26 October 2017; Received in revised form 12 April 2018; Accepted 13 April 2018
Available online 17 April 2018

consideration of their adequacy and effectiveness.

While simple WMO guidelines exist, many have studied more scientific and practical network design methods, which are well reviewed in the literature (Behmel et al., 2016; Chacon-hurtado et al., 2017; Keum et al., 2017; Mishra and Coulibaly, 2009). Mishra and Coulibaly (2009) grouped the most widely-used methods into eight categories: statistically based methods, spatial interpolation techniques, entropy-based methods, optimization methods, methods based on basin physiographic characteristics, methods driven by sampling strategies, hybrid methods, and user survey approach. Among the diverse network design methods, the entropy-based methods have emerged as promising alternatives to traditional statistical methods (Mishra and Coulibaly, 2009). This study resorts to entropy-based method combined with multi-objective optimization.

Entropy in information theory was developed to measure the amount of uncertainty in a dataset (Shannon, 1948) and has often been applied to hydrometric network design by providing a measure of information content. Early studies (e.g., Husain, 1989; Yang and Burn, 1994; Yoo et al., 2008) utilized the entropy theory to evaluate the existing rain gauges or streamflow stations. The entropy applications have been extended to determine proper locations of new additional monitoring stations. For example, the entropy terms have been used as the objective functions of multi-objective optimization technique to search the optimal hydrometric networks (e.g., Alfonso et al., 2013, 2010a,b; Kornelsen and Coulibaly, 2015; Ridolfi et al., 2014; Samuel et al., 2013). These studies specified the multi-objective optimization problem as maximizing joint entropy (i.e., total information content from a network) and minimizing total correlation (i.e., redundant shared information within the network), which are the basic principles of the Dual Entropy and Multiobjective Optimization (DEMO). Also, Leach et al. (2016, 2015) assessed the applicability of physical parameters, such as hydrologic signatures, indicators of hydrologic alteration and groundwater recharge, as objectives in addition to the traditional entropy objectives. However, research regarding snow monitoring networks is limited while many entropy-based hydrometric network studies have focused on streamflow and rainfall as reviewed above.

The objectives of this research are (1) to determine optimal snow monitoring networks that add more stations to the existing network, (2) to evaluate the Pareto-optimal networks from the performance of hydrologic models, and (3) to investigate the use of the Snow Data Assimilation System (SNODAS) products for the time series data at potential stations. While this research identifies the optimal snow networks in La Grande River Basin (LGRB) in Québec and Columbia River Basin (CRB) in British Columbia, the second objective focuses on the LGRB and the third objective explores the CRB.

## 2. Study area and data

### 2.1. La Grande River Basin, Québec

The first study area is a combined basin including La Grande River Basin and some of its surrounding watersheds that have been diverted into it. The combined basin, which includes 12 watersheds, will be referred to herein as La Grande River Basin (LGRB). LGRB is located in north-central Québec as shown in the index map of Fig. 1. The total drainage area is approximately 209,000 km$^2$ while the drainage area excluding diverted basins (i.e., Caniapiscau, Eastmain-1, Rupert bief aval, Rupert bief amont, Lac Mesgouez au lac Mistassini, and Lac Mistassini) is approximately 99,000 km$^2$. La Grande River rises in central Québec and flows west to James Bay. From the 30-year of climate observations at La Grande Riviere A station (WMO ID:71827), which is located near the outlet of La Grande River, daily mean temperatures are below zero from November to April and the lowest in January (−22 °C). Mean daily snowfall is the highest in November (1.9 cm/day) while the total precipitation is the highest in September (see Fig. 2). Based on the Advanced Very High Resolution Radiometer (AVHRR) Land Cover Dataset for Québec, 97 percent of the basin area is covered by forest while water bodies occupy 3 percent, and nearly the entire basin is an undeveloped area.

Due to the icy, snowy fields and the numerous lakes covering most areas in northern territories including LGRB, hydropower generation in Canada is the second largest in the world (BP, 2015) and covers more than half of the total national electricity consumption. Specifically in Québec, hydropower supplies account for approximately 95 percent of the total electricity consumption (Hydro-Québec, 2015). Major hydropower plants in Québec are mostly located in LGRB, and the primary source of reservoir inflows for those plants is the snowmelt during spring to early summer. LGRB currently has 47 snow monitoring sites (hereafter, existing stations), of which 30 stations are Hydro-Québec operating snow course sites, and the other 17 stations are from Environment and Climate Change Canada. Each station is supposed to have monthly visits in winter months (i.e., January, February, March, and April); however, due to the availability and accessibility, maintaining regular measurements (e.g., 15th day in each winter month) is often limited and cause snow data deficiency. Based on the physiographic units that WMO categorized to guide the minimum network density, the LGRB falls under the Interior Plains unit and its minimum network density should be 575 km$^2$ per non-recording station or 5750 km$^2$ per recording station (World Meteorological Organization, 2008). Recalling that the total drainage area of LGRB is approximately 209,000 km$^2$, there should be 37 automatic stations or 364 manual stations to meet the WMO guideline for a minimum snow network. Therefore, it can be concluded that there is a large deficit in snow monitoring. To overcome sparse density of snow monitoring networks and difficulties in data samplings, there is a need to install automatic snowpack sensor stations in LGRB, which can measure, calculate, and transmit the snow water equivalent (SWE) with comparable accuracy to other ground-based techniques, such as snow courses, snow pillows, and weighing precipitation gauges.

The snow data for the existing stations were obtained from the sample observations during winter months for the years of 1970–2005 (36 years) so that the number of total data points for each station is 144. A gridded SWE dataset for the entire study area is also available from Hydro-Québec, the hydropower company of Québec province (Tapsoba et al., 2005). Specifically, the 14 km resolution of the grid cells provides 990 grid points in LGRB. In this study, the weather stations currently installed in LGRB are chosen as the potential locations of new SWE observations (hereafter, potential stations) following personal communications with Hydro-Québec about the snow monitoring needs and plans, and the number of potential stations is 47 (see Fig. 1). The snow data of the potential stations are obtained from the SWE of the closest grid point. It should be noted that the current monitoring sites (called existing stations) are mainly for the ground-based snow observations, such as snow courses, snow pillows or weighing precipitation gauges, while automatic snowpack sensor stations are being planned at the optimal locations of additional stations. Therefore, the potential station locations which are very close to the existing monitoring sites are also considered given that automatic stations could be placed at selected manual monitoring sites to ensure continuity of historical snow observations.

### 2.2. Columbia River Basin, British Columbia

The entire Columbia River Basin covers parts of British Columbia (Canada) and the northwestern United States. However, this study will consider the upper parts of the Columbia River Basin (Fig. 3), where the British Columbia Hydropower company (namely BC Hydro) is responsible for the water resources management. It should be noted that the southern part of the basin is extended to the states of Idaho and Montana, and the American side of the basin is not included in the network design in this study. The drainage area of the Columbia River above Arrow Lakes Dam and the Kootenay River above Kootenay Canal