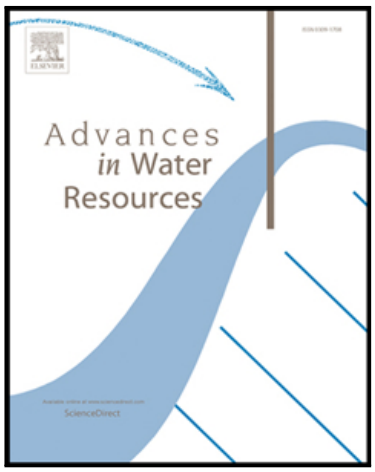


Development and evaluation of a hydrologic data-assimilation scheme for short-range flow and inflow forecasts in a data-sparse high-latitude region using a distributed model and ensemble Kalman filtering

Jos Samuel , Alain N Rousseau , Kian Abbasnezhadi , Stéphane Savary

PII: S0309-1708(18)30521-9
DOI: https://doi.org/10.1016/j.advwatres.2019.06.004
Reference: ADWR 3359

To appear in: *Advances in Water Resources*

Received date: 15 June 2018
Revised date: 11 May 2019
Accepted date: 7 June 2019

Please cite this article as: Jos Samuel , Alain N Rousseau , Kian Abbasnezhadi , Stéphane Savary , Development and evaluation of a hydrologic data-assimilation scheme for short-range flow and inflow forecasts in a data-sparse high-latitude region using a distributed model and ensemble Kalman filtering, *Advances in Water Resources* (2019), doi: https://doi.org/10.1016/j.advwatres.2019.06.004

Highlights

- A framework to merge NAEFS data into operational flow forecasting through EnKF
- EnKF data assimilation framework improved flow and inflow estimations
- Largest improvements were observed in late spring, summer, and early fall
- Raw NAEFS data overestimate precipitation and underestimate temperature
- Spatiotemporally corrected NAEFS data produce more accurate flow and inflow forecasts

Development and evaluation of a hydrologic data-assimilation scheme for short-range flow and inflow forecasts in a data-sparse high-latitude region using a distributed model and ensemble Kalman filtering

JOS SAMUEL[1*], ALAIN N ROUSSEAU[2], KIAN ABBASNEZHADI[2], STÉPHANE SAVARY[2]

[1] Yukon Research Centre, Yukon College, 500 College Drive, Whitehorse, Yukon Territory, Canada, Y1A 5K4
[2] Institut national de la recherche scientifique, Centre Eau Terre Environnement (INRS-ETE), 490, de la Couronne, Québec, Canada, G1K 9A9

Submitted to: *Advances in Water Resources*

*Corresponding author: Jos Samuel
e-mail: sjos@hotmail.ca

May 2019

Abstract

A forecasting system combining a physically-based distributed hydrological model (HYDROTEL), an Ensemble Kalman Filtering (EnKF) of Data Assimilation (DA), and forecasted meteorological data (obtained from the North American Ensemble Forecast System; NAEFS) is developed to forecast short-range (0-14 days lead) flows and inflows in the Aishihik and Mayo basins in Yukon Territory, Canada. The system was assessed at three sites, including at the outlet of the Sekulmun River subbasin of the Aishihik basin for river flow forecasting, as well as at Aishihik Lake and Mayo Lake for reservoir inflow forecasting. Model development and evaluation was performed systematically to ensure accuracy of forecasting outputs by: (i) investigating the use of coupled EnKF and HYDROTEL models for improved flow and inflow estimations, (ii) evaluating NAEFS data for short-range flow and inflow forecasts, and (iii) using probabilistic and deterministic criteria to evaluate the forecast performance of the HYDROTEL-EnKF-NAEFS model at each site. Results illustrate DA significantly improve flow and inflow forecasts, and raw NAEFS data need to be spatially and temporally corrected to be used for hydrological forecasts. Based on probabilistic and deterministic scores, it was found that the developed forecasting system can provide flow and inflow forecasts at the Sekulmun River subbasin, Mayo Lake, Aishihik Lake sites with high, medium, and low accuracies, respectively. Differences in forecast accuracies at each site are possibly associated with: (i) uncertainties of forecasted meteorological data, (ii) ability of HYDROTEL to capture daily flow and inflow variation, (iii) DA algorithm used, (iv) heterogeneity in basin attributes, and (v) limited availability of sources of information or data particularly in the lake areas.

Key words: Data assimilation, Ensemble Kalman Filter, ensemble weather prediction, flow forecasting, HYDROTEL, North American Ensemble Forecast System (NAEFS)

1. Introduction

Ensuring stability of power supply with demand is essential for sustainable hydro-electric generation. Energy supply and demand vary overtime which can challenge managers and operators of hydro-electric companies who do not have access to accurate, continuous, and real-time flow or inflow forecasts. A robust and reliable forecasting system made of a hydrological model and real-time meteorological forecast data can help provide real-time flow/inflow forecasts. However, such a framework is seldom deployed in high-latitude sub-arctic regions where snowmelt represents a major hydrologic process [73] and spatio-temporal sparsity of hydrometeorological data is a major obstacle [74, 2]. Given recent improvements in hydrological forecasting and availability of continuously improving meteorological forecasts, the objective of this study was to develop a reliable and robust flow/inflow forecasting system. This system was designed to forecast flows/inflows with a lead time of up to 14 days for the Mayo and Aishihik hydro-electric facilities in Yukon territory in Canada operated by the Yukon Energy Corporation (YEC); these two facilities by storage are the second and the third largest in the territory, respectively. The largest hydro-electric facility, the Whitehorse plant, is not included in the current study due to the inherent physical and hydrological complexity (such as the presence of glaciered areas), but is currently under development.

Several hydrological forecasting systems have been developed in the last few decades to improve flow/inflow prediction accuracies and at the same time to facilitate automation of simulations (i.e. automated and continuous computation without or very little human intervention). One of the well-known approaches is Data Assimilation (DA), which is popular due to its ability to merge a diverse range of observations with a dynamic model and provide a continuous correction of the forecast as observational data become available [21, 45]. It also offers the possibility to handle uncertainties associated with model parameters, observations, and forcing meteorological data by assimilating observations with hydrological models to correct the effects of model errors, update model states, and thus to improve flow/inflow forecasting accuracy [e.g. 13, 41, 44, 45, 62, 70, 80].

In its simplest form, a DA scheme can be configured in order to adjust initial states successively in time. This procedure is known as ‘filtering’. DA filtering is specifically effective in case of real-time operational flow forecasting, in which updating the initial states of a forecasting system is necessary in consecutive time intervals [48]. The most commonly applied

DA practice to increase the accuracy of flow forecasting is by adjusting (updating) the model's initial states based on observations. For instance, DeChant & Moradkhani [16] improved the estimation of snow depth quantities by designing a framework which takes into account the uncertainty of initial snow water storage conditions. In addition to initial states, DA can also update model parameters. Moradkhani et al. [46] presented a dual state–parameter estimation approach to update model states and parameters simultaneously. Vrugt et al. [78] designed a synchronized optimization and data assimilation scheme to reduce the uncertainty in hydrologic modeling. Pathiraja et al. [55] evaluated the performance of a DA-based framework (by adjusting time-varying model parameters) and explored if land cover changes in a mid-sized catchment in Vietnam were responsible for observed hydrologic variations. It is also possible to correct the mathematical structure of a hydrological model. For instance, Bulygina & Gupta [9] introduced a Bayesian DA procedure to adjust a deterministic hydrological model's structure. In addition, to estimate the uncertainty of the flow approximations, many DA approaches incorporate an ensemble of initial states. This ensemble technique is an efficient probabilistic way to quantify the uncertainty of flow estimation more accurately.

There are a number of deterministic and ensemble approaches available for data assimilation of linear and nonlinear systems, including flow, snow cover, snow water equivalent, water level, and soil moisture forecasting (to name a few), and any combination of these hydrological variables (e.g. [5, 7, 15, 32, 40, 41, 47, 54, 79, 81]). Available DA approaches include Kalman filter (KF; [31]), Extended Kalman filter (EKF; [21]), Ensemble Kalman filter (EnKF; [17]), and Particle filter (PF) [4], among others (e.g. variational filters, the maximum likelihood ensemble filter). Among these approaches, the EnKF has been by far one of the most popular ensemble DA techniques used for hydrological applications (see e.g. [22, 37, 39, 71]). There are several factors which have made EnKF an attractive DA scheme among hydrometeorologists, including ease of application, efficiency of the framework, and the explicit handling of uncertainties through an ensemble approach [48]. EnKF stems from the development of the KF approach and was extended from the EKF procedure. While the older versions (e.g. KF) were only applicable to linear systems, EKF and EnKF can be applied to non-linear dynamic systems (e.g. hydrological models) [1, 20, 37]. However, one of the main limitations of EnKF is its limited applicability to non-Gaussian state-space models [68]. Comparing between KF and PF, Nelly et al. [51] showed that KF and PF are both robust in that the KF is very fast in terms of

calculation time and convergence and the PF has advantages in handling nonlinear features of the model. Comparing the quantification of uncertainty for hydrologic forecasting between EnKF and PF, DeChant & Moradkahni [17] found similar abilities between two models to track the observations, but EnKF is more consistent in producing the results, while PF is more robust in parameter estimation technique.

In the current study, the EnKF is selected. The EnKF algorithm has been implemented for various configurations, such as for streamflow assimilation with lumped models [56, 72, 78, 80], and streamflow assimilation with distributed models [13, 30, 36]. The application of EnKF has also been applied for real-time water resource or hydrological applications [1]. EnKF incorporates statistical distributions to represent observation uncertainties of observations and model errors and to generate ensembles of model forcing and states [22, 37, 39]. In EnKF, an ensemble of initial conditions can be generated by perturbing various external and internal factors operating on and inside the model. For instance, the ensemble can be produced internally by re-running the hydrologic model while perturbing input data and/or model parameters [21, 62], or by accounting for model parameter errors, data errors, or structure errors [18, 35].

Applying EnKF to meet the degree of desired accuracy for the Aishihik and Mayo hydro-electric facilities is not a straightforward task. The main obstacles are the large uncertainties associated with limited availability of hydro-meteorological data, spatial heterogeneity of basin attributes, and the presence of lakes, which are subject to periodic water-level changes (see further discussion in the following sections). Selection of a hydrological model, which can be coupled to EnKF also requires additional considerations. Given the spatial heterogeneity of the study basins and possible land cover changes due to potential land developments, a physically-based distributed hydrological model was selected for operational hydrological forecasting in this study. A distributed model for operational hydrological forecasting provides an opportunity to forecast flows or inflows at various locations for a wide range of users [36]. However, compared to using lumped hydrological models, the application of EnKF with a distributed hydrological model is particularly challenging due to the existence of large degrees of freedom in both models [12, 33, 84]. Nevertheless, with increasing computational power, it is becoming the norm to apply distributed models in hydrologic data assimilation [36]. In this study, HYDROTEL [23, 24, 75, 76], a distributed hydrological model compatible with remote sensing and a Geographical Information System (GIS) was used (detailed description of HYDROTEL

will be presented in Section 3). This is the first time that HYDROTEL and EnKF have been coupled for operational hydrological forecasting, particularly in a high-latitude environment like the Yukon Territory, which is characterized by limited hydro-meteorological observation networks and climate conditions that continuously challenge hydro-electric managers.

Ensemble flow forecasting can be performed by incorporating weather ensemble forecasts as well. In fact, the number of studies in which ensemble streamflow forecasting for operational purposes is performed through the application of probabilistic weather products is increasing. The main benefit of using ensemble weather products in flow forecasting is that it provides an opportunity to account for forecast uncertainties (see e.g. [65, 66]). There are several major operational weather forecasting centers worldwide, which provide ensemble climate products, including the European Centre for Medium-Range Weather Forecasts (ECMWF), Environment and Climate Change Canada (ECCC), National Centers for Environmental Prediction, United Kingdom's Met Office, Australian Bureau of Meteorology, China Meteorological Administration, and Japan Meteorological Agency, among others. The ensemble weather products from a number of these centers have been used in several ensemble streamflow prediction studies. Gobena & Gan [25] used 0-3 months lead ensemble temperature and precipitation forecasts, including the Global Environmental Multiscale model (GEM) from the ECCC's Canadian Meteorological Centre (CMC) and the second generation atmospheric general circulation model (AGCM2) from the Canadian Centre for Climate Modeling and Analysis (CCCma) for seasonal streamflow prediction in two catchments in Alberta, Canada. Hopson & Webster [28] built a fully automated system to provide a 1-10 days lead ensemble flow forecasts for the major ungauged river basins in Bangladesh. The ECMWF ensemble weather products were bias-corrected based on satellite and rain gauge data and incorporated into ensemble flow prediction. The ECMWF ensemble weather products were also used by Renner et al. [65] for ensemble flow forecasting at various stations in the Rhine basin, Germany, and by Roulin & Vannitsem [66] to design a hydrological ensemble prediction system in two catchments in Belgium.

The North American Ensemble Forecasting System (NAEFS) is another ensemble weather product, which provides 0-14 days lead meteorological forecasts. NAEFS was developed under a joint project between the Meteorological Service of Canada (MSC), the United States National Weather Service (NWS), and the National Meteorological Service of

Mexico (NMSM). The result from this combination of products is higher in quality, more extensive, and consistent across the national boundaries of the three nations (https://weather.gc.ca/ensemble/naefs/info-semaine2_e.html, retrieved Feb 21, 2018). The combined system mixes different analyses, models, and perturbations, which increases the ensemble size and provide lower detection thresholds and earlier warnings in case of severe weather conditions. In addition, it provides a better assessment of the uncertainty and forecast skill for up to two weeks [10]. The system produces 20 ensemble members generated from the combination of two Ensemble Prediction Systems developed by MSC and NWS.

One benefit of using ensembles of meteorological data in hydrological/forecasting system is to account for forecast uncertainties (e.g. [65, 66]). For this and since its launch, the NAEFS ensembles have been used in a number of hydrological modeling and forecasting studies. For instance, Pomeroy et al. [57] used three-hourly NAEFS forecasts as driving meteorological variables in a model for the Little Smoky River in northern Canadian Rockies. However, to the best of our knowledge, there has not yet been any effort to merge NAEFS into an operational flow forecasting system by means of the merits that the ensemble Kalman filter DA approach provides. Therefore, the main objective of this study is to develop and evaluate a flow forecasting system, which integrates ensemble weather forecast data (here NAEFS) with a physically-based distributed hydrological model (here HYDROTEL) through an EnKF DA algorithm. This framework is designed in order to provide accurate and continuous short-range (0-14 days lead) flow and inflow forecasts. The framework will be developed and evaluated in three steps to ensure the accuracy of forecasting outputs, as follows:

(i) Investigate the improved performances of the coupled HYDROTEL-EnKF model over the stand-alone HYDROTEL model (referred to as "open loop" in this study),
(ii) Evaluate the NAEFS data for short-range (0-14 days lead) flow and inflow forecasts, and
(iii) Evaluate the performance of short-range flow and inflow forecasts in each study basin based on probabilistic and deterministic criteria.

The remainder of the paper is organized as follows. Data and basin characteristics are described in Section 2. The methodology is introduced in Section 3 and describes: (a) the HYDROTEL hydrological model, (b) NAEFS meteorological data, (c) the EnKF DA algorithm, and (d) the coupling of NAEFS data, HYDROTEL, and EnKF. In Section 4, evaluations of HYDROTEL open-loop, coupled HYDROTEL-EnKF model, and coupled HYDROTEL-EnKF-

NAEFS model in each study basin are discussed. A detailed discussion of the probabilistic and deterministic evaluation of the coupled HYDROTEL-EnKF-NAEFS model for short-range flow and inflow forecasting in each basin is also presented in Section 4. Conclusions are drawn in Section 5.

2. Study Areas, Basin Characteristics, and Data Availability

2.1. Study Areas and Basin Characteristics

The Aishihik and Mayo basins are located in the south-west and centre of Yukon Territory, respectively, as shown in Figure 1. The Aishihik basin covers an area of 4551 km². The climate is subarctic (based on Köppen climate classification) with the average annual precipitation of around 273 mm and average daily air temperature about -2.0°C (true for the 1976-2006 period). The maximum spring flow occurs in June, and the high flow season takes place from May to October [8]. The Aishihik Facility is a 37-MW generating station located south of Canyon Lake. Compared to the Aishhik basin, the Mayo basin, a subbasin of the Stewart River, covers a smaller area of roughly 2670 km². The average annual precipitation and daily air temperature are 316 mm and -2.6°C, respectively (true for the 1976-2006 period). The flow varies seasonally and the minimum and maximum flows occur, respectively, in December/January and June/July. The Mayo Facility has two generating stations: Mayo A, a 5-MW station located about 45 km north of Mayo, and Mayo B, a 10-MW station. These facilities use two lakes for water storage: Mayo Lake, 50 km north of Mayo, and Wareham Lake, 10 km north of Mayo.

Two digital Natural Resources Canada land cover maps, that is those from the Land Cover, circa 2000-Vector (LCC2000-V, http://www.nrcan.gc.ca/earth-sciences/geography/topographic-information/free-data-geogratis/download-directory-documentation/17215) and Moderate Resolution Imaging Spectroradiometer (MODIS, http://www.nrcan.gc.ca/node/9144) images were used to determine land cover characteristics of the study basins. The LCC2000-V images were used to classify the land cover characteristics, while the MODIS images were used to fill missing information where no data was found in the LCC2000-V images. It is noteworthy that the horizontal resolution of the LCC2000-V images (30 m) was better than those of the MODIS images (250 m). The LCC2000-V images included 43 different land cover classes. For hydrological modelling purposes, they were re-grouped into

seven land-cover classes, including water, short vegetation (shrub, herb, lichen, bare soil, and rock), wetlands, evergreen forest, deciduous forest, mixed forest, and burns. It was found that the dominant land covers in the Aishihik and Mayo basins were short vegetation and evergreen forest, respectively.

Using DEM obtained from Geomatics Yukon (ftp.geomaticsyukon.ca/DEMs/) and Natural Resources Canada (http://www.nrcan.gc.ca/home) at a resolution of 30 m, it was found that elevations of the Mayo and Aishihik basins were relatively similar. Minimum and maximum elevations for Mayo are 411 m and 2053 m, and those of Aishihik are 547 m and 2559 m, respectively (see Figure 1). These DEMs were then integrated into the stream and lake network data (watercourse and water body files are available at the Geomatics Yukon and DataBC portals (https://data.gov.bc.ca/) at a resolution of 1:50,000 to build the physiographic databases of the study basins. It was found that the number of lakes, and thus, the corresponding open water areas are smaller in Mayo than in Aishihik (Figure 1).

Information on soil texture and soil type for the study basins are limited. Therefore, soil type maps were derived using information provided by Rawls & Brakensiek [61], as follows. First, percentages of sand and clay available for three different soil horizons (0-10 cm, 10-25 cm, 25-375 cm) for each 1-km tile were determined following the same soil texture triangle approach proposed by Moeys [43]. Based on this triangulation, soil maps for the texture of second soil layer (10-25 cm) were derived and used as reference to estimate the textures of the first and third soil layers. However, in the presence of a non-mineral soil type, the information available for the first and third soil layers were used to substitute the non-mineral soil with the mineral soil information, whenever available. Finally, the ensuing 1-km resolution maps were subdivided into 30-m tiles to match the resolution of the DEMs. The estimated soil type maps for Aishihik and Mayo are introduced in Figures 2a and 2b, respectively. As shown, the dominant soil types in Aishihik is sandy-loam (mainly located in the western regions) and loam (located in the eastern regions). In Mayo, the dominant soil type is silty-loam.

2.2. Hydro-meteorological Data

A large number of meteorological stations have missing data or were discontinued in Yukon's high-latitude basins. In general, meteorological stations operated by the MSC and YEC are located within a 200-km radius of each study basin (see Table 1). For the 2010-2016 period, two

stations located within or in the vicinity of the Mayo basin and nine stations located within or in the vicinity of Aishihik basin have good data records (see Figure 3 and Table 1). Snow depth and SWE data were also used to calibrate HYDROTEL, and hence, data were retrieved from snow courses records (manual snow sampling sites, see Table 1). Available hydrometric stations providing river flows and water levels in each study basin are listed in Table 1. The inflows to Aishihik Lake (station #0000003, listed in Table 1) and those into Mayo Lake (station ##0000003) were 'reconstructed' based on water level and river flow changes surrounding the lake areas. Detailed methodology and the equations used to determine the reconstructed data are described in Appendix A. In brief, the reconstructed inflows were computed using a 3-day moving-average window. The reconstructed inflows to Mayo Lake were estimated using flow and water level data obtained from the YECMAYO and #09DC005 stations, while the inflows to Aishihik Lake were estimated using water level and/or flow data from the #08AA005, #08AA010, #08AA012, and #08AA008 stations. HYDROTEL was calibrated/validated at sites #08AA008 (Sekulmun River, Aishihik) and #0000003 (inflows to Aishihik Lake) in the Aishihik basin, and site ##0000003 (inflows to Mayo Lake) in the Mayo basin. Water level monitoring stations with poor records of data or those located downstream of reservoir inflows of the Aishihik and Mayo Lakes were not used in this study.

3. Models and Methodology

3.1. HYDROTEL

HYDROTEL is a semi-distributed physically-based model currently used for hydrological forecasting by the Division of Water Expertise, an organization in charge of management and safety of publicly-owned dams of the Québec Ministry of Sustainable Development, Environment, and Fight Against Climate Change [76]. The model simulates evapotranspiration, snow accumulation/melt, soil temperature/freezing depth, infiltration, recharge, surface flow, subsurface flow, and channel routing. The model is based on algorithms derived from physically-based equations (e.g. approximation of Saint-Venant's equations, one-dimensional (1D) Richards equation and Darcy's law), along with more conceptual or empirical approximations (e.g. base flow recession coefficient), and a net-radiation module based on the work of Archibald & Walter [6]. The computational domain is made of interconnected river segments, and either three-soil-layer elementary subbasins or hillslopes, referred to as relatively

homogeneous hydrological units (RHHU). The soil column is divided into three layers. The surface layer is relatively shallow (5-20 cm in thickness), so as to represent the soil layer affected by evaporation over bare ground. While the first layer controls infiltration, the second and the third layers are associated with interflow and base flow, respectively. Richards 1D transient equation is used to simulate flux exchange between the three layers. At each time step, the state variables calculated are the water contents of the three soil layers.

The major physical processes governing the water budget are simulated using different sub-models described in Table 2. Each sub-model generally offers more than one simulation option. For example, users have the option to compute potential evapotranspiration using either Thornthwaite, Linacre, Penman- Monteith, Priestly-Taylor, or the Hydro-Québec methods (see Table 2). The interested readers may refer to Fortin et al. [24] for additional details of each sub-models and simulation options. The bolded options listed in Table 2 represent those used in this study. In addition, state variables that influence hydrological processes include: (i) accumulation and snowpack melt, (ii) soil temperature and soil freezing, (iii) soil water content and vertical water budget, and (iv) stream flow towards the watershed outlet. Detailed list and description of each model state and parameter used in this study are provided in Appendix B.

The complete drainage structure of the basin and the associated RHHUs were determined in PHYSITEL, which is a specialized GIS software that supports the implementation of HYDROTEL [67, 75]. PHYSITEL converts river and lake networks into a raster format and then calculates the slope and flow direction of a DEM tile using the D8-LTD algorithm [53]. PHYSITEL also integrates land cover and soil texture maps based on percentages of sand, loam, and clay, along with corresponding hydrodynamic properties and wetland attributes [61]. In this study, soil type and land cover maps, DEM, and river and lake networks were taken from the data described in Section 2.1.

Input meteorological data to force HYDROTEL included precipitation and maximum and minimum air temperature data. In this study, the meteorological data were derived from the stations described in Section 2.2. Given the available meteorological data for the study basins, HYDROTEL was run on a daily time step. The hydrometric data described in Section 2.2 were used for model calibration. The model calibration was done manually with no optimization algorithm. The model was calibrated for the period of 2010-2016 through visual inspection and by maximizing the Nash-Sutcliffe Efficiency (NSE) and minimizing the percentage difference

between observed and simulated flows (PBIAS) and Root Mean Square Errors (RMSE) values. (the NSE and PBIAS formulas will be described in Section 3.5). For the Aishihik basin, the model calibration was performed for: (a) the specific model parameters for the entire Sekulmun River subbasin (i.e. upstream of the #08AA008 station), and (b) the specific model parameters for the remaining portion of the basin. For the Mayo basin, the model parameters were calibrated for the entire basin.

3.2. NAEFS Data

NAEFS data were obtained and downloaded on a grid resolution of 50 km × 100 km from the ECCC website (https://weather.gc.ca/ensemble/naefs/index_e.html). The acquired data were corrected with respect to spatial and temporal scales to provide more accurate flow/inflow forecasts. The raw NAEFS data were corrected in two steps. In the first step, the precipitation and minimum and maximum air temperature data were spatially distributed from the associated grid points to locations of the observed meteorological stations using the Inverse Distance Weighting (IDW) approach, given by Eq. (1a):

$$f_j = \sum_{i=1}^{n} w_i f_i \tag{1a}$$

where n is the number of NAEFS grid points surrounding the basin areas (here six NAEFS grid points in each study basin), f_i is the raw NAEFS meteorological data, f_j is the spatially distributed NAEFS meteorological data transferred from the NAEFS grid points to the meteorological station locations, and w_i is the interpolation weight, given as follows:

$$w_i = \frac{(h_i^{-2})}{\sum_{i=1}^{n} h_i^{-2}} \tag{1b}$$

where h_i is the distance from the NAEFS grid point to each meteorological station location.

In the second step, by using the observed data, the distributed NAEFS data were temporally corrected at each station through a scheme known as the delta change (DC) approach [38, 60]. This method is mainly used to deal with Global Climate Model (GCM) inadequacies, in which differences between current and future GCM simulations are computed and the changes are added to observations [27]. In this method, relative changes – or the "deltas" – in estimated rainfall amounts are translated to observed rainfall time series, generally by multiplicative factors [52]. Hay et al. [27] compared DC approach with statistical downscaling and used the approach

to compute the differences between current and future HadCM2 climate simulations and then added the changes to observed time-series of climate variables. Lenderink et al. [38] used the DC approach to perturb the HadRM3H time-series to project future climate flow rates.

By denoting P_{raw} as the spatially distributed raw NAEFS daily precipitation derived from Eq. (1a), $\bar{P}_{raw}$ as the spatially distributed raw NAEFS monthly average precipitation, and $\bar{P}_{obs}$ as the average observed monthly precipitation, the corrected NAEFS daily precipitation value (P_{corr}) was obtained by scaling P_{raw} by the ratio of $\bar{P}_{obs}$ to $\bar{P}_{raw}$, as follows:

$$P_{corr} = P_{raw} \times \frac{\bar{P}_{obs}}{\bar{P}_{raw}} \quad (2)$$

Temperature data were also corrected by first computing the difference between the average daily observed NAEFS minimum/maximum air temperature data and the spatially distributed raw NAEFS daily average air temperature data and then adding it to the raw NAEFS temperature value, as follows:

$$T_{corr} = T_{raw} + (\bar{T}_{obs} - \bar{T}_{raw}) \quad (3)$$

where T_{corr} is the corrected NAEFS daily minimum or maximum air temperature, T_{raw} is the spatially distributed raw NAEFS daily minimum or maximum air temperature derived from Eq. (1a), $\bar{T}_{raw}$ is the raw NAEFS daily average air temperature in each month, and $\bar{T}_{obs}$ is the average observed daily minimum or maximum air temperature in each month.

3.3. Ensemble Kalman Filtering (EnKF)

A brief description of the EnKF procedure used in this study is illustrated in Figure 4. Note that in this study, only model states were updated using flow or inflow observations, while model parameters were assumed constant and not updated. The superscript (-) sign indicates forecasted state, whereas the superscript (+) sign designates updated state derived by assimilating observed flows or inflows. The ensemble size *n* was set to 1000 (this value has been used in many EnKF studies, see for example [71]). The EnKF procedure can be summarized as follows:

(1) Compute an ensemble of *n* model states (x_{t+1}^{i-}) and predictions ($\hat{y}_{t+1}^{i}$) using calibrated model parameters (θ) and replicates of forcing data (u_t^i) through the hydrological model:

$$x_{t+1}^{i-} = f\left(x_t^{i+}, u_t^i, \theta\right) \tag{4}$$

$$\hat{y}_{t+1}^{i} = h\left(x_t^{i-}, \theta\right) \tag{5}$$

where x_{t+1}^{i-} is the *i*th ensemble member of forecasted states at time *t*+1 and x_t^{i+} is the *i*th ensemble member of updated states at time *t*.

The forcing data are perturbed by adding Gaussian random noise ς_t^i with covariance Σ_t^u at each time step t generated for all ensemble sizes, as follows:

$$u_t^i = u_t + \varsigma_t^i, \qquad \varsigma_t^i \sim N(0, \Sigma_t^u) \tag{6}$$

(2) Perturb observation at time *t*+1 when flow or inflow observation becomes available:

$$y_{t+1}^i = y_{t+1} + \varepsilon_{t+1}^i, \quad \varepsilon_t^i \sim N(0, \Sigma_{t+1}^y) \tag{7}$$

(3) Update the ensemble of model states according to the standard Kalman equation by assimilating observed flows or reconstructed inflows as follows:

$$x_{t+1}^{i+} = x_{t+1}^{i-} + K_{t+1}^x\left(y_{t+1}^i - \hat{y}_{t+1}^i\right) \tag{8}$$

where,

$$K_{t+1}^x = \Sigma_{t+1}^{xy}\left[\Sigma_{t+1}^{yy} + \Sigma_{t+1}^{y}\right]^{-1} \tag{9}$$

in which K_{t+1}^x is the Kalman gain for correcting the model state trajectories, Σ_{t+1}^{xy} is the cross covariance of model state ensemble (x_{t+1}^{i-}) and observed prediction ensemble, Σ_{t+1}^{yy} is the forecast error covariance matrix of the observed prediction ensemble ($\hat{y}_{t+1}^i$), and Σ_{t+1}^{y} is the covariance of the actual flow/inflow observation.

For the Aishihik basin, the EnKF procedure was applied in two steps sequentially. In the first step, EnKF was used to update model states of the Sekulmun River subbasin using observed flows at site #08AA008, and in the second step, model states for locations between upstream of site#0000003 and downstream of site #08AA008 were updated using reconstructed inflows at site #0000003. Note that site #08AA008 is located upstream of site #0000003, so that the states of the #08AA008 subbasin were updated earlier than those at site #0000003 (see Figure 1). This two-step DA simulation allows the users to obtain more accurate results at the upstream site

(#08AA008) if gauge observation in the downstream site (#0000003) is not available or missing. It should be noted that inflows at downstream site (#0000003) is estimated from different observational gauges (see Appendix A for computing the reconstructed inflows in the #0000003 site). For the Mayo basin, the EnKF procedure was applied in only one step, in which all model states located upstream of site ##0000003 were updated using reconstructed inflows at this station.

3.4. Integration of NAEFS into HYDROTEL through EnKF

NAEFS data were integrated into HYDROTEL through the application of the EnKF DA procedure. The integration was completed to form an automated forecasting system (see the illustration of the system in Figure 5), so that the system runs continuously for operational purposes. A Graphical User Interface was designed for this DA system (which, without loss in continuity, is not introduced herein). In short, the system first downloads, reads, and arranges real-time observed flows (obtained from the ECCC and YEC ftp servers) and the NAEFS forecast meteorological data (from the ECCC servers) at 3 am daily. The downloaded NAEFS data include a 0-14 days lead daily precipitation and maximum and minimum air temperature forecasts. The available observed flow/inflow data of previous day are also downloaded. The NAEFS data are then corrected and their ensembles are generated. The ensembles of errors of observed flows and inflows are also generated. In the next step, the corrected NAEFS data are forced into HYDROTEL to simulate model states and forecasted flows and inflows. Note that the NAEFS data has 20 ensemble members and each one is then perturbed 50 times (assuming to be Gaussian with 0-mean and 5%-standard deviation) to form a thousand ensembles of model runs. Subsequently, the EnKF algorithm is used to update states of HYDROTEL. Finally, the updated states are returned to HYDROTEL to be used in the next simulation run. This system runs once per day to produce short-range flow and inflow forecasts.

3.5. Probabilistic and Deterministic Model Evaluation

Evaluation of forecast accuracy is crucial to assess and quantify forecast performances and uncertainties, particularly in operational hydrological forecasting systems (see [25, 28, 85], among others). Forecast performances/skills are commonly assessed using deterministic and/or

probabilistic criteria. In brief, deterministic criteria measure performance of the average (or median) forecasts, while probabilistic criteria measure the performance of the ensemble forecasts. The evaluation results can be used to help the forecaster make unbiased judgements, for example, to issue warnings in probabilistic format [19]. In this study, the evaluation of short-range flow and inflow forecast performances was performed using both probabilistic and deterministic criteria. By accounting for uncertainties in decision making, economic benefits of forecasts may increase [17, 34, 63]. It also helps measure the ability of the model to track the observation and represent the inherent uncertainty in the prediction [16] and evaluate different sources of uncertainty during model development [63].

Model accuracies were evaluated using deterministic scores, including Nash-Sutcliffe Efficiency (NSE; Eq. 10), Root Mean Square Error-Observations Standard Deviation Ratio **(**RSR; Eq.11), and Percent Bias (PBIAS; Eq.12) [49], given as follows:

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}\left(\mathrm{x}_{obs,i}-\mathrm{x}_{sim,i}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{x}_{obs,i}-\bar{\mathrm{x}}_{obs}\right)^2}\right] \quad (10)$$

$$RSR = \frac{\sqrt{\sum_{i=1}^{n}\left(\mathrm{x}_{obs,i}-\mathrm{x}_{sim,i}\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(\mathrm{x}_{obs,i}-\bar{\mathrm{x}}_{obs}\right)^2}} \quad (11)$$

$$PBIAS = \frac{\sum_{i=1}^{n}\left(\mathrm{x}_{obs,i}-\mathrm{x}_{sim,i}\right)*100}{\sum_{i=1}^{n}\left(\mathrm{x}_{obs,i}\right)} \quad (12)$$

where $\mathrm{x}_{obs,i}$ and $\mathrm{x}_{sim,i}$ are the observed and simulated flows or inflows, respectively, for the *i*th day, and $\bar{\mathrm{x}}_{obs}$ is the mean of observed flows or inflows over *n* days of record. The closer are the values of RSR, PBIAS, PVE, and VE to 0, and NSE to 1, the better would be the model's performance.

Moriasi et al. [49] classified model simulation performance based on these three criteria from the highest to the lowest performance ratings as very good, good, satisfactory, or unsatisfactory; a simulation would be considered as unsatisfactory if NSE < 0.50, RSR ≥ 0.70, and PBIAS ≥ 25%. These three model criteria were used to evaluate HYDROTEL stand-alone and coupled HYDROTEL-EnKF model simulations (without using NAEFS).

For evaluating short-range flow and inflow forecast performances using the combined HYDROTEL-EnKF-NAEFS model, two more deterministic criteria were considered, namely the Percent change of absolute Volume Error (PVE; Eq. 13) and Volumetric Error (VE; Eq. 14) [49, 69], given by:

$$PVE_i = \frac{|x_{obs,i} - x_{sim,i}|}{(x_{obs,i})} \quad , \quad i = 1, \dots, m \tag{13}$$

$$VE_i = \left(x_{obs,i} - x_{sim,i}\right) \ , \quad i = 1, \dots, m \tag{14}$$

where *m* is the number of days.

The PVE criterion measures the relative changes in absolute observed and forecasted flow or inflow (in percent) on each day. The absolute measurement used in the PVE equation is to disregard the overestimation or underestimation of modeled flows or inflows over observations. On the other hand, VE which measures the difference between modeled and observed flows or inflows (i.e. modeled flow or inflow errors) (in m^3/s) on each day, does not use an absolute measurement. The VE criterion rather investigates the variation of underestimation or overestimation of modeled errors. We used PVE and VE criteria instead of NSE, PBIAS, and RSR criteria for evaluating the short-term flow/inflow forecast performance since the former criteria can measure model performances on a daily basis, which could be useful for operational forecasting applications, while the other criteria need a series of data (see the differences between Eqs. 10-12 and Eqs. 13-14).

Probabilistic skill scores, which are widely used to describe the quality of ensemble forecasts [25, 82] were also applied in this study. They include the ranked probability skill score (RPSS; [83]) and Receiver Operating Characteristic (ROC) Curves [25, 42], as follows:

$$RPSS = 1 - \frac{\overline{RPS}}{\overline{RPS}_{ref}} \tag{15a}$$

$$RPS = \sum_{m=1}^{J}\left(\sum_{j=1}^{m} f_i - \sum_{j=1}^{m} o_j\right)^2 \tag{15b}$$

where J is the number of categories (here J=3), f and o are, respectively, the forecast and observation probabilities (they equal to 1 if the event occurs in the jth category and zero if otherwise), and $\overline{RPS}$ is the average value of RPS for each forecast/event pair. For the 3-category forecast, the climatological forecast probability in the jth category is 1/3. Thus, the *RPS* for the climatological forecast is computed using f_j = 1/3 in Eq. (15b). The subscript "*ref*" in Eq. 15b

refers to the reference forecast (i.e. climatological forecast with no useful forecast information). For a perfect forecast with respect to climatology, *RPS* is equal to zero and *RPSS* is equal to 1. Forecasts that are worse than the reference receive negative *RPSS* values. Positive *RPSS* values indicate improvements relative to the reference forecast.

The *RPSS* measures the improvement of forecast probability (Eq. 15a) relative to the reference (or climatology) forecast (Eq. 15b). The steps to compute the *RPSS* are described as follows (see also [25]). The ensemble members were first transformed into binary or multi-category groups in order to discretize forecast probabilities and compute the probabilistic scores. In this study, the ensemble members were classified into three groups (*i.e.*, "below normal", "near normal" and "above normal" categories) based on the one-third percentiles of the cumulative distribution of the observed flows following Gobena & Gan [25]. If an event (i.e. an ensemble member) occurred in a specific category, it was marked as 1. Otherwise, it was marked as zero. The forecast probability (i.e. the *f*-value in Eq. 15b) is then computed based on the fraction of ensemble members falling in a specific category.

While *RPSS* measures the overall performance of the ensemble forecasts, the ROC Curves measure the performance of the ensemble forecast at different levels of probability [25]. The primary use of the ROC Curve is to compare probabilistic forecasts of events or non-events. In this study, event or nonevent categories were determined based on whether flow/inflow ensembles fell into below- or above-normal categories, respectively, based on one-third percentiles of the cumulative distribution of the observed flows. The probability levels (warming thresholds) used are from 10% to 90% (at increments of 10%). The ROC Curve can be represented by a plot of hit rate (HT; Eq. 16a) on the *y*-axis and false-alarm rate (FAR; Eq. 16b) on the *x*-axis at different warning thresholds. The hit rate and false-alarm rate are calculated by:

$$HT = \frac{h}{(h+m)} \tag{16a}$$

$$FAR = \frac{f}{(f+r)} \tag{16b}$$

where *h* and *m* specify the number of cases where a warning is issued correctly and incorrectly for an event, respectively, and *f* and *r* are the number of cases where a warning is issued correctly and incorrectly for a nonevent, respectively.

The plot of hit rate and false-alarm rate values is normally compared with no skill (random) performance in which the hit rate and false-alarm rate are equal, creating a diagonal line (1:1 line). The farther are the hit and false-alarm curves above the 1:1 line and closer to the ideal point (i.e. HT=1 and FAR=0), the higher would be the forecast skill. Moreover, the warning threshold having the closest distance to the perfect model indicates the optimal warning threshold. This optimal threshold corresponds to the optimal HT and FAR, which identifies the probability of the optimal detection ability of ensembles to occur in a specific category (either above- or below- normal).

4. Results and Discussion

This section discusses three analyses that are interconnected among their corresponding results. In the first analysis (Sections 4.1 and 4.2), observed and simulated flow/inflow obtained from the HYDROTEL open-loop simulation and the coupled EnKF-HYDROTEL model are compared. The main purpose of this analysis is to examine the benefit of using EnKF for flow/inflow estimation. In this analysis, the observed meteorological data were used as forcing in the coupled model. In the second analysis (Section 4.3), the forcing observed meteorological data were replaced with raw and corrected NAEFS data in the coupled HYDROTEL-EnKF model to investigate short-range (0-14 days) flow/inflow forecasts. This section discusses the improvements gained from using flow/inflow forecasts derived from the corrected NAEFS data compared to those obtained from the raw data. In the third analysis (Section 4.4), short-range flow/inflow forecasts are further assessed based on various probabilistic and deterministic criteria. These three analyses were evaluated at the different sites, including #08AA008 (the Sekulmun River subbasin in the Aishihik basin), #0000003 (Aishihik Lake in the Aishihik basin), and ##0000003 (Mayo Lake in the Mayo basin). The #08AA008 site records natural river flows without or very minor flow alterations, while the other two sites measure reconstructed inflows.

4.1. HYDROTEL Stand-alone

Results of HYDROTEL open-loop calibration for the period of 2010-2016 for each site are shown in Figure 6 and Table 4. In general, HYDROTEL can decently capture flow observations, while it is less skillful to capture reconstructed inflows. For example, as shown in Figure 6a, in the Aishihik basin, HYDROTEL can closely capture flow observations at site #08AA008 with the NSE score of as high as 0.84 (Table 4). The remaining errors are due to overestimation of winter low flows in some years and in part by underestimation of snowpack accumulation by HYDROTEL. As shown in Figure 6b, HYDROTEL has generally closely captured annual inflow variations at site #0000003. However, as shown in Figure 7b, the model has poorly simulated daily inflow variations. Daily inflows at site #0000003 show rapid fluctuations, which is associated with the uncertainty of combined water level and river flow data obtained from the four flow and water level stations surrounding Aishihik Lake (see Section 2.2). There were not enough data available to clearly identify the physical processes occurring in Aishihik Lake, such as the influences of wind speed/direction, lake bathymetry, water temperature, lake evaporation, and timing and depth of freezing and melting of snow and water on the lakes. The version of HYDROTEL used in this study did not have a reservoir module to account for in-depth physical processes occurring in lakes. This will be the subject of a future study to improve the accuracy of the model outputs and to understand the main cause for daily inflow variations.

In the Mayo basin, as shown in Figure 6c the model can closely capture annual inflow variability, while it generally underestimates summer and fall peak flows. It is partly due to the limited availability of meteorological stations having continuous and good data coverage in this area. There were only two meteorological stations available when the model was calibrated. It should be mentioned that since this study was started, another two new meteorological stations were installed (one in Mayo and one in Aishihik). However, since the period of data records for these stations were relatively short (when HYDROTEL was calibrated), they were not used in this study.

The reconstructed inflows to Mayo Lake show less variation than those of Aishihik Lake (see differences in Figures 7b, c). It was easier to reconstruct inflows to Mayo Lake (based on data from two flow/water level stations than those into Aishihik Lake (based on five stations combining natural and regulated flows and water level data, see Section 2.2). The Mayo basin is not as heterogeneous as the Aishihik basin, the size of the basin and its lakes is small, and there

are only a small number of lakes inside the basin. Detailed comparison of the hydrological processes between the Mayo and Aishihik basins will be presented in a separate manuscript.

The NSE, PBIAS, and RSR values of the HYDROTEL open-loop simulation for the period of 2010-2016 show that the model can capture flow observations very well at site #08AA008. However, the accuracies slightly degrade at sites #0000003 and ##0000003 (Table 3a). Based on the classification introduced by Moriasi et al. [49], the current values of NSE, PBIAS, and RSR indicate that the HYDROTEL stand-alone model can provide a "very good" simulation at site #08AA008 and a "satisfactory" one at sites #0000003 and ##0000003 (Table 3a). For the most part, these results indicate that the calibrated HYDROTEL model can adequately simulate observed flows and inflows in the Mayo and Aishihik basins.

4.2. HYDROTEL-EnKF Coupling

The calibrated HYDROTEL model was coupled with EnKF. Within this framework, the HYDROTEL-EnKF model was run in a daily sequence, which required a significant amount of time to complete. Therefore, the evaluation was performed using the 2016 data only. A one-year simulation is relatively enough to assess the performance of the coupled model in different seasons. However, the use of longer data would provide more detailed statistical analyses. We compared the flow/inflow estimates resulting from: (a) the HYDROTEL open-loop simulation, (b) the coupled HYDROTEL-EnKF model simulation, and (c) observed data. The same observed meteorological data that were used as forcing in the coupled HYDROTEL-ENKF model, were also used to calibrate HYDROTEL. It was found that at all three sites, the coupled HYDROTEL-EnKF model produces much more accurate flow and inflow estimates compared to the open-loop simulation (see Figure 7 and Table 3b). The coupled model can better capture daily variations of observed flows/inflows (Figure 7) with NSE values are twice as high as the open-loop simulation (Table 3b). For example, the NSE scores at sites #08AA008, #0000003, and ##0000003 derived from the HYDROTEL open-loop simulation are 0.7, -0.1, and 0.4, respectively. With the coupled HYDROTEL-EnKF model, these scores increased up to 0.99, 0.31, and 0.96, accordingly. The improvements generally take place during late spring, summer, and early fall, during which the coupled model has improved the underestimation of flows/inflows previously seen in the HYDROTEL open-loop simulation (Figure 7). Overall, the results demonstrate the

effective incorporation of the EnKF DA scheme into the forecasting system which is consistent with those of other similar studies [1, 46, 50, 69]. We can compare our results with those of similar studies performed in Canada. Abaza et al. [1] conducted streamflow data assimilation using HYDROTEL in two catchments in Quebec based on an EnKF sequential procedure, and found an improvement of NSE comparing between open-loop and DA simulations for various experiment setup within the range of 0.05-0.25. Comparing our results to those of other studies applied for operational water resources, Munier et al. [50] found an improvement of NSE to about 0.2-0.25 of using various DA compared to using open-loop simulation applied at the Selingue dam located in the Upper Niger River Basin in Nigeria.

As shown in Figure 7 and Table 3b, the evaluation of the coupled HYDROTEL-EnKF outputs shows that the flow estimates at sites #08AA008 (Figure 7a) and ##0000003 (Figure 7c) are more accurate than the inflow estimates at site #0000003 (Figure 7b). The difference in estimation accuracies is mainly associated with the ability of the HYDROTEL stand-alone model in simulating observed flows and inflows. The more accurate the HYDROTEL stand-alone simulation at capturing flow/inflow observation is, the more accurate the coupled HYDROTEL-EnKF simulation will be (Table 3b). It appears that to obtain more accurate estimates; the hydrological model needs to be first improved.

4.3. Evaluation of Raw and Corrected NAEFS Data

The applicability and evaluation of NAEFS data for a 0-14 day lead flow/inflow forecasts are discussed in this section. In this evaluation, both raw and corrected NAEFS data were used as input in the coupled model and the results were compared. The comparison is only shown for site #08AA008 (Sekulmun River at Aishihik), while similar conclusions can be drawn for the other sites. In general, the raw NAEFS data produced larger errors of flow and inflow forecasts than those derived from the corrected NAEFS data. For instance, as displayed in Figure 8, when the coupled model was forced with the raw NAEFS data, the flow estimates showed inaccuracies. After correcting the NAEFS data (discussed in Section 3.2), we found an improvement in the flow accuracies, particularly for longer lead times (see Figure 9a). When comparing raw NAEFS data against observations in both basins, the former generally indicates higher precipitation amounts and lower minimum and maximum air temperatures than the observations (not shown here). This is due in part to a coarser grid resolution (50 km × 100 km) of NAEFS data. Overall,

these results indicate that it is required to correct the raw NAEFS data spatially and temporally before using them in short-range flow/inflow forecasts.

The visual inspections revealed that on day 0 of the forecasting period (i.e. the date on which the forecast is issued), the accuracy is slightly lower when NAEFS data is used as input in the coupled HYDROTEL-EnKF model than in case the meteorological observation are used as forcing (the 0 day of the forecasting period is not plotted for sake of simplicity in Figure 9). The results indicate that there is a slight error when transferring/correcting raw NAEFS data from NAEFS grids to the location of observations on day 0 of the forecasting period. For longer lead times, flow/inflow inaccuracies build up gradually (Figure 9). Candille [10] investigated NAEFS performances of several atmospheric variables using objective verification tools developed at the Canadian Meteorological Centre and found that their dispersion and Continuous Ranked Probability Scores increase with increasing lead times. Since no additional DA was performed beyond the issue date, the uncertainty of the NAEFS meteorological forecasts for larger lead times affects the uncertainty of the associated flows/inflows. The flow/inflow uncertainty for larger lead times is also associated in part with HYDROTEL, albeit not quantified here. Figures 7 and 9 also indicate that the HYDROTEL-EnKF-NAEFS framework underestimates the flow/inflow observations (see Figure 9), which is in the same direction as the HYDROTEL stand-alone simulation (see Figure 7). Identification of the main factors causing this gradual decline of flow/inflow forecast accuracy at longer lead times requires additional investigation.

4.4. Forecast Evaluation of the Combined HYDROTEL-EnKF-NAEFS Model

The performance of short-range (0-14 day lead) flow/inflow forecasts obtained from the combined HYDROTEL-EnKF-corrected NAEFS approach was evaluated using the deterministic (PVE and VE) and the probabilistic (RPPS and ROC curve) criteria at each site. Although all 1000 ensemble values were used in the probabilistic analysis, only the average values were used in the deterministic analysis. Figure 10 shows the PVE and VE boxplots for all lead times. The boxplots are delimited by the 25^{th} and 75^{th} percentiles; the median is marked inside the box, and the whiskers are delimited by the 10^{th} and 90^{th} percentiles of the 1-366-day period (using the 2016 data). Figure 11 shows the ROC curves for the 1-, 3-, 7-, and 14-day forecasts (selected lead times are shown for simplicity). Table 4 shows the RPSS scores and the optimal warning

threshold on the ROC curves measured by the Euclidean distance from a perfect model (*i.e.* HR=1 and FAR=0) and their corresponding HRs and FARs for all forecast periods.

All deterministic and probabilistic results for the Sekulmun River subbasin (#08AA008) provide very good scores for all lead times. In general, the inaccuracies build up gradually at longer lead times (Table 4 and Figures 10a and 11a). For example, the 90th percentiles of the PVE scores are below 10% on day 0 and below 40% on day 14. The inter-quantile range of the VE values is relatively small, varying from 0 to -0.36 m^3/s (on day 0) and from 1.0 to -1.3 m^3/s (on day 14). The two evaluation criteria (RPSS and ROC curves) also show that the model can provide very good detection of the occurrence of the events (either above- or below-normal category) for up to 14 days. On day 14, the RPSS values are as high as 58% indicating an improvement relative to the reference forecast (Table 4). This indicates that the coupled model can provide a 58% improvement over a forecast with no skill up to day 14 of the forecast. The below- and above-normal forecast categories of the ROC curves on day 14 are above the 1:1 lines and bends towards the top left optimal corner point (HR=1 and FAR=0) for all warning thresholds (10%-90%) (Figure 11a). The optimum HR on day 14 is still very high (above 85%), while the FAR remains below 1% indicating very accurate forecast outputs (Table 4). The forecasting system can provide warning rates in the order of 85% correctly and 1% incorrectly that the event (either below- or above-normal categories) will occur on day 14. Overall, the results show that the forecasting system can be used to forecast flows at site #08AA008 up to 14 days with high accuracy.

Compared to the evaluation results at site #08AA008, the performance of the forecasting system at site #0000003 (Aishihik Lake) shows low quality skills (see Table 4 and Figures 10b and 11b). The model performance from day 0 to day 14 is also not much different. This under-performance is mainly due to a high variation of daily reconstructed inflows that cannot be captured well by the HYDROTEL open-loop simulation (see also the discussion in Section 4.2). For all forecast periods, the 90th percentiles of the PVE and VE values are relatively high (>100% and >5 m^3/s, respectively; see Figure 10b). The RPSS values for all forecasting time periods are also relatively low (around 0.2), indicating low forecasting skill. However, the ROC curves for all forecasting time periods still show relatively acceptable outcomes. The HRs and FARs are still above diagonal lines (see Figure 11b). For example, at day 1, the HRs of the below- and above-normal categories remain above 70% at the warning threshold of 20% and

40%, respectively. This indicates that HR exceeds FAR. The proportion of events (both for below- and above- normal forecast categories) for which a warning was issued correctly is still higher that the proportion of non-events for which a warning was issued incorrectly. ROC curves (Figure 11b) also indicate a better skill for above-normal forecasts than for below-normal ones; this indicates that the hydrologic model underestimates low flows much less than it overestimates high flows at this site. By evaluating the optimal HR and FAR (Table 4) and averaging the corresponding rates for above- and below-normal categories of all forecasting time periods, the values are in the order of 75% and 25%, respectively. It means that on average, the coupled model can issue a warning about 75% correctly and 25% incorrectly that the event (either below- or above-normal categories) occurs within the 0-14 day forecasting lead time. In general, it can be concluded here that even though the ensembles of the forecasting system can satisfactorily detect the occurrence of above- or below-normal events, it generally produces relatively high uncertainties and errors.

The performance of the forecasting system at site ##0000003 (Mayo) is better than those at site #0000003 (Aishihik), but it is lower than those at site #08AA008 (see Table 4 and Figures 10 and 11). For example, the inter-quantile range of the PVE values is between 0.5% and 60% at day 0 and between 11% and 74% at day 14 (see Figure 10c), which is larger than those found at site #0000003, but lower than those found at site #08AA008. Similar trends were also detected based on the VE and RPSS criteria. The RPSS values gradually decrease from 72% (on day 0) to 23% (on day 14) with a value greater than 50% on day 6 (Table 4). The below- and above-normal ROC curves showing good results are those above the 1:1 random line (on which there is a forecast with no skill) (Figure 11c). The HRs exceeds the FARs predominantly for almost all forecast lead times. For example, on day 7 and for a warning threshold of 60%, the HRs and FARs are respectively 83% and 19% for the below-normal category and 95% and 12% for the above-normal category (see Figure 10c and Table 4). The range of optimal HR and FAR for the below-normal category is 85%-92% and 12%-25%, respectively. For the above-normal category, the ranges are 86%-98% and 3%-18%, respectively (Table 4). The differences between the HRs and FARs are relatively high, indicating the model can almost properly detect the occurrence of above- or below-normal events for up to 14 days.

The following points summarize and compare probabilistic and deterministic model performances of the combined HYDROTEL-EnKF-NAEFS model at each site:

(1) Forecast accuracies decrease with increasing forecasting lead times (Table 4, Figures 10 and 11) at sites #08AA008 (Sekulmun River, Aishihik) and ##0000003 (Mayo), but the forecast accuracies are similar for the entire forecast lead time at site #0000003 (Aishihik Lake), which is mainly due to high daily inflow variation.

(2) The deterministic evaluation results (*i.e*. both the PVEs and VEs, Figure 10) for the 2016 period show that the system produced low, medium, and high errors at sites #08AA008 (Sekulmun River, Aishihik), ##0000003 (Mayo), and #0000003 (Aishihik Lake), respectively.

(3) The above-normal categories of the ROC curves (Figure 11) are generally higher than those of the below-normal categories for almost all cases and at all study sites. This indicates that the forecasting system overestimates high flows more often than it underestimates low flows.

(4) The above- and below-normal categories of ROC curves (Figure 11) are above the 1:1 random line for all cases and at all study sites. This indicates that compared to a forecast with no skill, the ensembles of the flow and inflow forecasts can satisfactorily detect the occurrence of below- or above-normal events within the forecast lead time.

(5) The average optimal HRs and FARs for all forecasting time periods at sites #08AA008 (Sekulmun River, Aishihik), #0000003 (Aishihik Lake), and ##0000003 (Mayo), as provided in Table 4 (written in the order of HRs [FARs]) are: 95% [5%], 67% [28%], and 86% [19%] for below-normal categories, and 95% [3%], 77% [20%], and 93% [11%] for above-normal categories, respectively. Based on these results, the ensembles of the forecasting system can provide the best detection of an event to occur in each above- or below-normal category at site #08AA008 (Sekulmun River, Aishihik), while the worst detection was identified at site #0000003 (Aishihik Lake).

(6) Based on all probabilistic and deterministic results, we can conclude that the forecasting system can be used to forecast flows and inflows at sites #08AA008 (Sekulmun River, Aishihik), ##0000003 (Mayo), and #0000003 (Aishihik) with a high, medium, and low degree of accuracy, respectively.

5. Summary and conclusion

In this study, an EnKF data assimilation approach was designed for ensemble flow prediction, in which ensemble weather prediction products were forced into a physically-based distributed hydrological modelling system (HYDROTEL). This framework was developed for short-range (0-14 day lead time) ensemble flow and inflow forecasting, in two major hydro-electric facilities (Aishihik and Mayo) in Yukon, Canada. The model was examined at various subbasins, including the Sekulmun River subbasin (site #08AA008) for flow forecasting, and the Aishihik Lake (site #0000003) and Mayo Lake (site ##0000003) for reservoir inflow forecasting. A step-by-step model development and evaluation was performed to ensure the accuracy of the forecast outputs at all sites.

The HYDROTEL model was first calibrated and validated for the period of 2010-2016. The EnKF approach was then coupled with the calibrated HYDROTEL model to investigate the benefits of using the coupled HYDROTEL-EnKF framework over the HYDROTEL stand-alone (open-loop) simulation. In this evaluation, observed meteorological data were used to feed HYDROTEL open-loop and the coupled models. Results showed that the EnKF data assimilation framework improved flow and inflow estimations. During the evaluation period (2016), the coupled model was able to improve the accuracy of the flow and inflow simulations compared to the open-loop simulation at sites #08AA008, ##0000003, and #0000003, increasing NSE values from 0.7 to 0.99, from -0.1 to 0.31, and from 0.4 to 0.96, respectively. The largest improvements were observed in late spring, summer, and early fall. The coupled model was able to do a better job at sites #08AA008 (Sekulmun River, Aishihik) and ##0000003 (Mayo) compared to site #0000003 (Aishihik Lake). This is mainly due to the high variation of lake water levels and river flows at the latter site. Additional information in the lake areas, future improvements in HYDROTEL, and application of other data assimilation approaches could help improve the inflow accuracy at site #0000003 in a future study. In addition, the model can better capture reconstructed inflows to Mayo Lake than to Aishihik Lake; this is partly due to the lower variation of the Mayo basin attributes, including size of the basin and lake areas.

In the second step of model development and evaluation, raw and corrected 0-14 day lead NAEFS weather prediction data were used as input in the coupled HYDROTEL-EnKF model to derive and examine 0-14 day lead flow and inflow forecasts. Raw NAEFS data were corrected spatially and temporally using the IDW and delta approaches, respectively. Our results showed that the raw NAEFS data overestimated precipitation and underestimated minimum and

maximum air temperatures when compared to station observations, causing inaccuracy of flow and inflow forecasts. On the other hand, application of the corrected NAEFS data produced more accurate flow and inflow forecasts, albeit with increasing errors at longer lead times. Since no data assimilation was performed beyond the forecast issue date, the gradual increase of flow/inflow inaccuracies with increasing lead times should be associated with inaccuracies of the NAEFS meteorological forecasts for longer lead times and also current difficulties in HYDROTEL to capture flow and inflow variations. We found that at longer lead times, the forecast outputs generally underestimated the observations, following the HYDROTEL open-loop simulation trends.

In the final step, the forecast performance of the combined HYDROTEL, EnKF, and corrected NAEFS data was evaluated using two deterministic (PVE and VE) and two probabilistic (RPSS and ROC curve) criteria at each site. In brief, the deterministic criteria were used to measure the performance of the average forecasts, while the probabilistic criteria quantified the performance of the ensemble forecasts. These criteria were used to evaluate the performance of the forecasting system for different lead times. The RPSS values on day 14 were 58%, 23%, and 18% at sites #08AA008, ##0000003, and #0000003, respectively, and the deterministic scores (i.e. the 90^{th} percentiles of the PVE values) were less than 40%, 70%, and 150%, and the inter-quantile range of the VE values were from -1.5 to 1 m^3/s, from -5 to 10 m^3/s, and from -5 to 8 m^3/s, respectively. In addition, the ROC curves for site #08AA008 (Sekulmun River, Aishihik) was closer to the perfect model than those for site ##0000003, while the ROC curves for site ##000003 (Mayo) were closer to the perfect model than those for site #0000003 (Aishihik Lake). These results indicated that at sites #08AA008 (Sekulmun River, Aishihik), ##0000003 (Mayo), and #0000003 (Aishihik Lake), the performance of the forecasting system show a high, medium, and low degrees of accuracy, respectively. The main possible reason for different accuracy levels among hydrometric sites is in all likelihood associated with uncertainties of the NAEFS meteorological data, ability of HYDROTEL to capture daily flow and inflow variations, the EnKF algorithm used, heterogeneity of basin attributes, and limited data availability, particularly in the lake areas.

This study contributes a hydrologic DA framework design by coupling a distributed hydrological model, corrected real-time meteorological forecasts, and EnKF that clarifies the integration of various types of datasets and accounts for spatial distribution of hydrological and

meteorological variables, and thus improves 0-14 days lead flow and inflow forecasting used for operational applications. The advance of automated DA tools in operational hydrologic forecasting has been hindered due to time consuming, incompatibility in customizing DA methods for every specific model, and transferring issues with the developed forecasting system replacing the existing system [41]. Our developed model has been currently used by YEC for their automated and continuous hydro-electric operations demonstrating no applicability issue in applying the model for operational applications. This systematic framework is also currently used to guide improvements for the development of the forecasting system in other YEC hydro-electric facilities, including the Whitehorse plant of the Upper Yukon River. This framework also includes model evaluation by applying probabilistic and deterministic statistical approaches to better understand model capability and forecasting uncertainty at different lead times and at each study site including the identification of possible sources of uncertainty and opportunities for future model improvement.

Even though we have shown the potential of EnKF for operational hydrologic DA in the Aishihik and Mayo basins, there are a few other topics that require further investigation, as described in the following:

(1) EnKF algorithm and its outputs, similar to any other DA approach, are directly influenced by the accuracy and density of flow observations. In this study, limitations were mostly related to the sparsity of observational information used to update the model states in EnKF, particularly in the lake areas. This limitation has a direct impact on the accuracy of the estimated model states and possibly on the parameters of hydrologic model, which can subsequently reduce the accuracy of flow/inflow estimations. Additional observations, such as lake water level, river flows in other river segments, continuous and accurate soil moisture and/or snowpack data could help improve inflow/flow forecasts as shown in other DA studies (see [2, 36, 71], among others). These additional data would be also useful for developing hydrological models, improving our understanding of the involving hydrological processes, and supporting the identification and resolution of basin processes.

(2) Other variations of EnKF, such as the dual EnKF approach [46], which has the flexibility to account for model adjustments through the time variation of parameters besides state

variables, may further improve model state and parameter estimations simultaneously and thus increase the inflow estimation accuracy.

(3) The selection of the noise settings was purely based on the random generation of spatially uniform white noise over the study area, which may not be realistic. Therefore, it would be a topic of great interest to investigate the effects of noise specification on the quality of hydrological forecasts. Goovaerts [26] performed a conditional geostatistical approach to match perturbations with the sample statistics of the observed rainfall fields. Hongli et al. [29] applied a hyper-parameter tuning procedure to fine-tune the parameters of the statistical model (used for generating errors to perturb input climate data) in such a way to improve the representativeness of forcing data uncertainty. Another popular approach is to conditionally generate precipitation time-series, which has been widely applied in various hydrometeorological studies (see e.g. [3, 14, 64, 77]). For example, Rakovec et al. [59] used the multivariate precipitation model developed by Rakovec et al. [58] to simulate an ensemble of correlated precipitation time-series and force them into a flow data assimilation scheme.

(4) Model states variations, including the impact of changing noise settings on states variations have not been investigated in the current study. An analysis of this kind can be considered in the future.

(5) Advanced methods to downscale and interpolate NAEFS data, such as a coupled regression and Kriging geospatial method [11] with additional data (climatological maps, etc.) can be applied in the future. This coupled method applies a combined digital elevation model (DEM) and supplementary information to relate precipitation and temperature with physical predictor variables, such as slope, curvature, distance to a coast, and others. The method appears to outperform regression methods and Kriging [11].

Acknowledgement

The authors wish to gratefully acknowledge the financial support from Natural Sciences and Engineering Research Council of Canada (NSERC) and Yukon Energy Corporation (YEC). This project would have not been possible without substantial contributions from staffs at NCE, Yukon College, INRS, and YEC. We would like to also acknowledge the Meteorological Survey

of Canada, Water Survey Canada, NAEFS, Environment and Climate Change Canada, and Geobase for providing data used in this study.

Appendix A. Procedure to reconstruct inflows in the Aishihik (#0000003) and Mayo (##0000003) sites

A simplified reservoir volume equation can be expressed as follows:

$$\Delta V = IN - Q_{out} \quad \text{(A.1)}$$

where ΔV is the variation of lake or reservoir volume in m³/s, between time $j-1$ and j, IN is the sum of inflows from upstream rivers and surrounding hillslopes, and Q_{out} is the sum of the entire lake or reservoir outflows. For both watersheds, V was determined by adopting a calculation procedure based on the three-day water level average, as follows:

$$\Delta V = V_j - V_{j-1} \quad \text{(A.2)}$$

For Aishihik Lake, the general volume calculation, which relates volume to level, is as follows:

$$V = \begin{cases} -38627.31L^6 + 678170.61L^5 - 3270008.00L^4 + 6352008.56L^3 - 3111511.38L^2 + 134383853.50 & \text{If } L < 915 \\ -50827494.83L^4 + 731085628.63L^3 - 3932429415.76L^2 + 9545583654.41L - 8448705648.59 & \text{If } L \geq 915 \end{cases} \quad \text{(A.3)}$$

where L represents the water level recorded at Site #08AA005 (Aishihik Lake near Whitehorse hydrometric station).

For Mayo Lake, the general volume calculation is as follows:

$$V = \frac{L-660}{0.00003814} \times 3600 \quad \text{(A.4)}$$

where L represents the water level recorded at Site #09DC005 (Mayo Lake near the outlet hydrometric station).

To determine Q_{out} for Aishihik Lake, we used the following equation:

$$Q_{\text{out}} = Q_{\text{08AA010}} - Q_{\text{08AA009}} \quad \text{(A.5)}$$

where Q_{08AA010} is the average daily flow at the Aishihik River below Aishihik Lake hydrometric station, and Q_{08AA009} is the average daily flow at the Giltana Creek near the mouth hydrometric station.

In order to accurately calculate the lake outflows, we subtracted the Giltana Creek (#08AA009) flow from the Aishihik River measurements, since the #08AA010 hydrometric station is located downstream of both Aishihik Lake and Giltana Creek and is the closest flow measurement downstream of the Lake.

To determine Q_{out} for Mayo Lake, we used the following equation:

$$Q_{\text{out}} = Q_{\text{YECMAYO}} \quad \text{(A.6)}$$

where $Q_{YECMAYO}$ is the average daily flow measurement made by YEC at the outlet of the Mayo Lake facility.

Care should be taken when determining the total inflow (IN) in Equation (A.1) where the volume variation must be divided by 86400 (s/day) to get the proper flow units in m³/s. Also, it is important to highlight that for both watersheds, the estimated total inflows may result in a false negative value since the water budget equation assumes a horizontal surface. However, large lakes act as large mechanical oscillator driven by wind forces, precipitations, ice, water management, among other factors. Such conditions can result in errors in total inflow calculations; including excessive variations and negative values. In the case of negative inflow

values, it was decided to substitute the negative values by a nominal value (0.01) which corresponds to minimum calculated positive inflow value.

Appendix B. List of model states and their description used in HYDROTEL

This appendix presents a list of HYDROTEL states updated by the EnsKF data assimilation scheme:

1. Snowpack

Modelling of spatio-temporal snowpack dynamics is performed for three types of land covers: deciduous forests, coniferous forests, and open areas (i.e. all other land covers). The three types of land covers could also be three forested areas with different tree densities. For each one of the aforementioned land cover type of each RHHU, five model states, that is snow depth (m), snow water equivalent (m), snowpack thermal deficit (J/m^2), snowpack liquid water content (m), and albedo are updated at each computational time step.

2. Soil temperature

Modelling of soil temperature (°C) is restricted to a user-specified depth of soil (5 cm for this study) and thus updated by the the EnsKF data assimilation scheme.

3. Soil moisture

For each RHHU, a water budget is computed over three soil layers. Their hydrodynamic properties are set with respect to their texture and thus soil type (from clays to sands). There is

one dominant soil type per RHHU. The soil water content of each layer (m^3/m^3) is updated at each computational time step. The effective porosity corresponds to the maximum soil water content.

4. Overland flow routing

The outflow of the water budget is routed to the river segment of a RHHU using a geomorphologic unit hydrograph based on the kinematic wave equation considering the physiographic characteristics (slope and surface roughness of the RHHU). This outflow is updated at each computational time step (m^3/s).

5. River flow routing

River flow is routed to the watershed outlet, segment by segment, using the kinematic wave equation. For each segment, upstream flow, outflow, and lateral flows (m^3/s) are updated at each computational time step.

References

[1] Abaza, M., Anctil, F., Fortin, V. & Perreault, L. (2017), Hydrological Evaluation of the Canadian Meteorological Ensemble Reforecast Product, Atmosphere-Ocean, doi: 0.1080/07055900.2017.1341384.

[2] Abbasnezhadi, K., Rousseau, A.N., Koenig, K.A., Zahmatkesh, Z., & Wruth, A.M. (2019a), Hydrological assessment of meteorological network density through data assimilation simulation. J. Hydrol., 569, 844–858, https://doi.org/10.1016/j.jhydrol.2018.12.027

[3] Abbasnezhadi, K., Rousseau, A.N., Wruth, A.M., & Zahmatkesh, Z. (2019b), Synchronized generation of high-resolution gridded precipitation and temperature fields. J. Hydrol., 573, 631–647, https://doi.org/10.1016/j.jhydrol.2019.03.096

[4] Abbaszadeh, P., Moradkhani, H., & Yan, H. (2018), Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo, Advances in Water Resources, 111, 192–204 https://doi.org/10.1016/j.advwatres.2017.11.011

[5] Andreadis, K.M. & Lettenmaier, D.P. (2006), Assimilating remotely sensed snow observations into a macroscale hydrology model, Adv. Water Resour, 29, 872–886.

[6] Archibald, J.A. & Walter, M.T. (2014), Do energy-based PET models require more input data than temperature-based models? – An evaluation at four humid FluxNet sites, Journal of the American Water Resources Association (JAWRA), 27, 3511–3515, doi: 10.1111/jawr.12137

[7] Biancamaria, S., Durand, M., Andreadis, K.M., Bates, P.D., Boone, A., Mognard, N.M., & Clark, E.A. (2011), Assimilation of virtual wide swath altimetry to improve Arctic river modeling, Remote Sens. Environ., 115 (2), 373–381, doi:10.1016/j.rse.2010.09.008

[8] Brabets, T.P. & Walvoord, M.A. (2009), Trends in streamflow in the Yukon River Basin from 1944 to 2005 and the influence of the pacific decadal oscillation, Journal of Hydrology, 371 (1-4), 108-119, doi: 10.1016/j.jhydrol.2009.03.018

[9] Bulygina, N. & Gupta, H. (2011) Correcting the mathematical structure of a hydrological model via Bayesian data assimilation. Water Resour. Res. 47(5), W05514 (2011). https://doi.org/10.1029/2010WR009614

[10] Candille, G. (2009), The Multiensemble Approach: The NAEFS Example, Mon. Weather Rev., 137 (5), 1655-1665.

[11] Cantet, P. (2015), Mapping the mean monthly precipitation of a small island using kriging with external drifts. Theoretical and Applied Climatology, Springer Verlag, doi: 10.1007/s00704-015-1610-z

[12] Carpenter, T.M., Georgakakos, K.P., & Sperflage, J.A. (2001), On the parametric and NEXRADradar sensitivities of a distributed hydrologic model suitable for operational use, J Hydrol, 253,169–93.

[13] Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., & Uddstrom, M.J. (2008), Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, Adv. Water Resour., 31, 1309–1324.

[14] Clark, M.P., Slater, A.G. (2006), Probabilistic quantitative precipitation estimation in complex terrain. J. Hydrometeorol. 7, 3–22. http://dx.doi.org/10.1175/JHM474.1

[15] DeChant, C.M. & Moradkhani, H. (2011a), Radiance data assimilation for operational snow and streamflow forecasting using the particle filter, Adv. Water Resour, 34 (3), 351–364.

[16] DeChant, C.M., & Moradkhani, H. (2011b), Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. Hydrol. Earth Syst. Sci. 15, 3399–3410, https://doi.org/10.5194/hess-15-3399

[17] DeChant, C.M., & Moradkhani, H. (2012), Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. Water Resources Research, 48, 1–15, https://doi.org/10.1029/2011WR011011

[18] Del Giudice, D., Löwe, R., Madsen, H., Mikkelsen, P.S., & Rieckermann, J. (2015), Comparison of two stochastic techniques for reliable urban runoff prediction by modeling systematic errors. Water Resour. Res. 51, 5004–5022, https://doi.org/10.1002/2014WR016678

[19] Diomede, T., Nerozzi, F., Paccagnella, T., & Todini, E. (2008), The use of meteorological analogues to account for LAM QPF uncertainty, Hydrol. Earth Syst. Sci., 12, 141-157, doi: 10.5194/hess-12-141-2008

[20] Dong, J., Walker, J.P., Houser P.R., & Sun C. (2007), Scanning multichannel microwave radiometer snow water equivalent assimilation, J. Geophys. Res. D: Atmos, 112 (D7).

[21] Evensen G. (1992), Using the extended Kalman filter with a multi-layer quasigeostrophic ocean model, J. Geophys. Res, 97, 17905–17924.

[22] Evensen G. (2003), The ensemble Kalman filter: theoretical formulation and practical implementation, Ocean. Dynam, 53, 343–367.

[23] Foulon, É., & Rousseau, A.N. (2018), Equifinality and automatic calibration, what is the impact of hypothesizing an optimal parameter set on modelled hydrological processes, Can.Water Resour. J., doi: 10.1080/07011784.2018.1430620

[24] Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J, & Villeneuve, J.P. (2001), A distributed watershed model compatible with remote sensing and GIS data, Part 1: description of the model, Journal of Hydrologie Engineering, 6 (2), 91-99.

[25] Gobena, A.K., & Gan, T.Y. (2010), Incorporation of seasonal climate forecasts in the ensemble streamflow prediction system, Journal of Hydrology, 385, 336–352.

[26] Goovaerts, P. (1997), Geostatistics for Natural Resources Evaluation. Oxford University Press, USA.

[27] Hay, L.E., Wilby, R.L., & Leavesley, G.H. (2000), A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States, Journal of the American Water Resources Association, 36, 387–397, https://doi.org/10.1111/j.1752-1688.2000.tb04276.x

[28] Hopson, T. & Webster, P. (2010), A 1-10 day ensemble forecasting scheme for the major river basins of Bangladesh: forecasting severe floods of 2003-07, J. of Hydrometeorol., 11 (3), 618-641.

[29] Hongli, L., Thiboult, A., Tolson, B., Anctil, F., & Maia J. (2019), Efficient treatment of climate data uncertainty in ensemble Kalman filter (EnKF) based on an existing historical climate ensemble dataset, J Hydrol, 568, 985–996. https://doi.org/10.1016/j.jhydrol.2018.11.047

[30] Ibbitt, R.P., Clark, M.P., Woods, R.A., Zheng, X., Slater, A.G., Rupp, D.E., *et al*. (2007), Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model, In: AGU fall meeting, San Francisco, 2007.

[31] Kalman, R.E. (1960), A new approach to linear filtering and prediction problems, J. Basic Eng, 82 (1), 35–45.

[32] Kitanidis, P.K., & Bras, R.L. (1980), Real-time forecasting with a conceptual hydrologic model, 2. Application and results, Water Resour. Res., 16 (6), 1034–1044.

[33] Koren, V.I., Finnerty, B.D., Schaake, J.C., Smith, M.B., Seo, D.J., & Duan, Q.Y. (1999), Scale dependencies of hydrologic models to spatial variability of precipitation, J Hydrol, 217, 285–302.

[34] Krzysztofowicz, R., Drzal, W.J., Rossi, D.T., Weyman, J.C., & Giordano, L.A. (1993), Probabilistic quantitative precipitation forecasts for river basins, Weather Forecasting, 8, 424–439.

[35] Kumar, S.V., Dong, J., Peters-Lidard, C.D., Mocko, D., & Gómez, B. (2016), Role of forcing uncertainty and model error background characterization in snow data assimilation.

Hydrol. Earth Syst. Sci. Discuss. 1–24. https://doi.org/10.5194/hess-2016-581

[36] Lee, H., Seo, D.J., & Koren, V. (2011), Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic model: effects of uncertainities in the data and initial model soil moisture states, Adv. Water Resour, 34, 1597–1615.

[37] Leisenring, M., & Moradkhani, H. (2011), Snow water equivalent estimation using Bayesian data assimilation methods, Stoch. Env. Res. Risk Assess, 253–270, http://dx.doi.org/10.1007/s00477-010-0445-5.

[38] Lenderink, G., Buishand, A., & van Deursen, W.V. (2007), Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, Hydrol Earth Syst Sci., 11 (3), 1145–1159. https://doi.org/10.5194/hess-11-1145-2007

[39] Liang, S. (2004), Quantitative remote sensing of land surfaces. Wiley, Hoboken, NJ.

[40] Liu, Y., & Gupta, H.V. (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, Water Resour. Res., doi:10.1029/2006WR005756

[41] Liu, Y., Weerts, A.H., Clark, M., Franssen, H.J.H., Kumar, S., Moradkhani, H., Seo, D.J., Schwanenberg, D., Smith, P., van Dijk, A.I.J.M., van Velzen, N., He, M., Lee, H., Noh, S.J., Rakovec, O., & Restrepo, P. (2012), Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, Hydrol. Earth Syst. Sci., 16, 3863–3887.

[42] Mason, S.J., & Graham, N.E. (1999), Conditional probabilities, relative operating characteristics, and relative operating levels, Weather and Forecasting, 14, 713–725.

[43] Moeys, J. (2009), The Soil Texture Wizard: R functions for plotting, classifying and transforming soil texture data. 99pp.

[44] Moradkhani, H., & Meskele, T.T. (2009), Probabilistic assessment of the satellite rainfall retrieval error translation to hydrologic response, in Satellite Applications for Surface Hydrology, Springer, Water Science and Technology Library, doi 10.1007/978-90-481-2915-7, 229-242.

[45] Moradkani, H., Hsu, K., Gupta, H., & Sorooshian S. (2005a), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resour Res., doi: 10.1029/2004WR003604

[46] Moradkhani, H., Sorooshian, S., Gupta, H.V., & Houser, P.R. (2005b), Dual state parameter estimation of hydrological models using ensemble Kalman filter, Adv Water Resour., 28, 135–147.

[47] Moradkhani, H., DeChant, C.M., Sorooshian, S. (2012), Evolution of Ensemble Data Assimilation for Uncertainty Quantification using the Particle Filter-Markov Chain Monte Carlo Method, Water Resources Research, doi:10.1029/2012WR012144

[48] Moradkhani, H., Nearing, G., Abbaszadeh, P., & Pathiraja, S. (2018), Fundamentals of Data Assimilation and Theoretical Advances. Handbook of Hydrometeorological Ensemble Forecasting. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39925-1_30

[49] Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., & Veith, T.L. (2007), Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Trans. ASABE, 50 (3), 885-900.

[50] Munier, S., Polebistki, A., Brown, C., Belaud, G., & Lettenmaier, D.P. (2015), SWOT data assimilation for operational reservoir management on the upper Niger River Basin, Water Resour. Res., 51, 554–575, doi:10.1002/2014WR016157

[51] Nelly, J.-B., Malaterre, P.-O., Dorée, C., Sau, J. (2011), Data assimilation for real-time estimation of hydraulic states and unmeasured perturbations in a 1D hydrodynamic model, Mathematics and Computers in Simulation, 81 (10), 2201 – 2214, https://doi.org/10.1016/j.matcom.2010.12.021

[52] Olsson, J., Berggren, K., Olofsson, M., & Viklander, M. (2009), Applying climate model precipitation scenarios for urban hydrological assessment: A case study in Kalmar City, Sweden, Atmospheric Research, 92, 364–375, https://doi.org/10.1016/j.atmosres.2009.01.015

[53] Orlandini, S., Moretti, G., & Franchini, M. (2003), Path-based methods for the determination of non-dispersive drainage directions in grid-based digital elevation models, Water Resources Research, 39 (6), 1114.

[54] Parrish, M.A, Moradkhani, H., & DeChant, C.M. (2012), Toward reduction of model uncertainty: Integration of Bayesian model averaging and data assimilation, Water Resources Research, 48, W03519, doi:10.1029/2011WR011116

[55] Pathiraja, S., Anghileri, D., Burlando, P., Sharma, A., Marshall, L., & Moradkhani, H. (2017), Time varying parameter models for catchments with land use change: the importance of model structure. Hydrol. Earth Syst. Sci. Discuss. https://doi.org/10.5194/hess-2017-382

[56] Pauwels, V.R.N., & De Lannoy, G.J.M. (2006), Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, J Hydrometeorol, 7 (3), 458–77.

[57] Pomeroy, J., Shook, K., Fang, X., Brown, T., & Marsh, C. (2013), Development of a Snowmelt Runoff Model for the Lower Smoky River, Centre for Hydrology, University of Saskatchewan, Report No. 13. 89pp.

[58] Rakovec, O., Hazenberg, P., Torfs, P.J.J.F., Weerts, A.H., & Uijlenhoet, R. (2012a), Generating spatial precipitation ensembles: impact of temporal correlation structure. Hydrol. Earth Syst. Sci. 16, 3419–3434. http://dx.doi.org/10.5194/hess-16-3419-2012

[59] Rakovec, O., Weerts, A.H., Hazenberg, P., Torfs, P.J.J.F., & Uijlenhoet, R. (2012b), State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy. Hydrol. Earth Syst. Sci. 16, 3435–3449, http://dx.doi.org/10.5194/hess-16-3435-2012

[60] Räty, O., Räisänen, J., & Ylhäisi, J.S. (2014) Climate Dynamics (2014), Evaluation of delta change and bias correction methods for future daily precipitation: intermodel cross-validation using ENSEMBLES simulations, 42, 2287, https://doi.org/10.1007/s00382-014-2130-8

[61] Rawls, W.J., & Brakensiek, D.L. (1989), Estimation of soil water retention and hydraulic properties, In Unsaturated Flow in Hydrologic Modeling: Theory and Practice, Morel-Seytoux HJ (ed), NATO ASI series. Series C: Mathematical and Physical Sciences 275. Kluwer Academic: Boston, 275–300.

[62] Reichle, R.H. (2008), Data assimilation methods in the Earth sciences, Advances in Water Resources, 31, 1411–1418.

[63] Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S.W. (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resources Research, 46 (5), https://doi.org/10.1029/2009WR008328

[64] Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G., & Franks, S.W. (2011), Toward a reliable decomposition of predictive uncertainty in hydrological modeling: characterizing rainfall errors using conditional simulation. Water Resour. Res., 47, http://dx.doi.org/10.1029/2011WR010643

[65] Renner, M., Werner, M.G.F., Rademacher, S., & Sprokkereef, E. (2009), Verification of ensemble flow forecast for the River Rhine, J. Hydrol., 376, 463-475.

[66] Roulin, E. & Vannitsem, S. (2005), Skill of medium-range hydrological ensemble predictions, J. Hydrometeorol., 6, 729-744.

[67] Rousseau, A.N., Fortin, J.P, Turcotte, R., Royer, A., Savary, S., Quévy, F., Noël, P., & Paniconi, C. (2011), PHYSITEL, a specialized GIS for supporting the implementation of distributed hydrological models, Water News – Official Magazine of the Canadian Water Resources Association, 31 (1), 18-20.

[68] Salamon, P. & Feyen, L. (2009), Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter, Journal of Hydrology, 376, 428–442, https://doi.org/10.1016/j.jhydrol.2009.07.051

[69] Samuel, J., Coulibaly, P., & Metcalfe, R. (2011), Estimation of continuous streamflow in Ontario ungauged Basins: Comparison of regionalization methods, ASCE J. Hydrol Eng, 16 (5), 447–459.

[70] Samuel, J., Coulibaly, P., & Metcalfe, R.A. (2012), Evaluation of future flow variability in ungauged basins: Validation of combined methods, Advances in Water Resources, 35, 121–140.

[71] Samuel, J., Coulibaly, P., Dumenah, G., Moradkhani, H. (2014), Assessing Model State Variation in Hydrologic Data Assimilation, Journal of Hydrology, 513, 127–141.

[72] Seo, D.J., Koren, V., & Cajina, N. (2003), Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting, J. Hydrometeorol, 4, 627–641.

[73] Spence, C., & Woo, M. (2008) Hydrology of the Northwestern Subarctic Canadian Shield. In: Woo M. (eds) Cold Region Atmospheric and Hydrologic Studies. The Mackenzie GEWEX Experience. Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-75136-6_13.

[74] Spence, C. & Hedstrom, N. (2018), Hydrometeorological data from Baker Creek Research

Watershed, Northwest Territories, Canada, Earth Syst. Sci. Data, 10, 1753-1767, https://doi.org/10.5194/essd-10-1753-2018

[75] Turcotte, R., Fortin, J.P., Rousseau, A.N., Massicotte, S., & Villeneuve, J.P. (2001), Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network, Journal of Hydrology, 240 (3–4), 225–242.

[76] Turcotte, R., Lacombe, P., Dimnik, C., & Villeneuve, J.P. (2004), Distributed hydrological prediction for the management of Quebec's public dams/Prevision hydrologique distribuée pour la gestion des barrages publics du Quebec, Canadian Journal of Civil Engineering, 31 (2), 308–320.

[77] Vischel, T., Lebel, T., Massuel, S., & Cappelaere, B. (2009), Conditional simulation schemes of rain fields and their application to rainfall–runoff modeling studies in the Sahel, J. Hydrol., 375, 273–286. http://dx.doi.org/10.1016/j.jhydrol.2009.02.028

[78] Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W., & Verstraten, J.M. (2005), Improved treatment of uncertainty in hydrologic modeling: combining the strengths of global optimization and data assimilation, Water Resour. Res. 41, W01017. http://dx.doi.org/10.1029/2004WR003059

[79] Vrugt, J.A., ter Braak, C.J.F., Diks, C.G.H., & Schoups, G. (2013), Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications, Advances in Water Resources, 51, 457–478.

[80] Weerts, A.H., & El Serafy, G.Y.H. (2006), Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42, W09403. http://dx.doi.org/10.1029/2005WR004093

[81] Weerts, A.H., El Serafy, G., Hummel, S., Dhondia, J., & Gerritsen, H. (2010), Application of generic data assimilation tools (DATools) for flood forecasting purposes, Comput. Geosci., 26, 453–463.

[82] Weigel, A.P., Liniger, M.A., & Appenzeller, C. (2007), The Discrete Brier and Ranked Probability Skill Scores, Monthly Weather Review, doi: 10.1175/MWR3280.1

[83] Wilks, D.S. (2006), Statistical Methods in the Atmospheric Sciences, Academic Press, Amsterdam.

[84] Winchell, M., Gupta, H.V., & Sorooshian, S. (1998), On the simulation of infiltration- and saturation excess runoff using radar-based rainfall estimates: effects of algorithm uncertainty and pixel aggregation, Water Resour Res, 34 (10), 2655–2670.

[85] Zapata, J.A.V. (2010), Evaluation of Hydrological Ensemble Prediction Systems for Operational Forecasting, Department de Genie Civil Et Genie Des Eaux, Faculte des Science e de Genie, Universite Laval, Quebec, Canada, 100pp.

Table captions

Table 1. Hydro-meteorological networks/stations in each study basin

Table 2. HYDROTEL sub-model and simulation Options

Table 3. Model performances of the (a) HYDROTEL alone simulation (2010-2016) and (b) the comparison of HYDROTEL alone and coupled HYDROTEL-EnKF simulations (2016)

Table 4. RPSS, optimum warning threshold and their corresponding hit rate and false alarm rates

Table 1. Hydro-meteorological networks/stations in each study basin

Data type	Code	Station Name	Station No.	Period	Operation	Measurement type	Source*)	Basin
Meteorology	1MA	BLANCHARD RIVER	2100163	1986-2012	DAILY	AUTO	MSC	Aishihik
Meteorology	2MA	BURWASH A	2100181	2011-PRESENT	HOURLY & DAILY	AUTO & MANUAL	MSC	Aishihik
Meteorology	3MA	BURWASH A	2100182	1966-2015	HOURLY & DAILY	AUTO	MSC	Aishihik
Meteorology	4MA	BURWASH AIRPORT AUTO BC	2100184	2013-PRESENT	HOURLY & DAILY	AUTO	MSC	Aishihik
Meteorology	5MA	CARMACKS CS	2100301	1999-PRESENT	HOURLY & DAILY	AUTO	MSC	Aishihik
Meteorology	6MA	HAINES JUNCTION	2100630	1944-PRESENT	HOURLY & DAILY	AUTO	MSC	Aishihik
Meteorology	7MA	OTTER FALLS NCPC	2100840	1980-2015	DAILY	MANUAL	YEC	Aishihik
Meteorology	8MA	PELLY RANCH	2100880	1898-2015	DAILY	MANUAL	MSC	Aishihik
Meteorology	9MA	TAKHINI RIVER RANCH	2101095	1980-2015	DAILY	MANUAL	MSC	Aishihik
Meteorology	1MM	MAYO A	2100700	1924-2013	HOURLY & DAILY	AUTO & MANUAL	MSC	Mayo
Meteorology	2MM	MAYO A	2100701	2013-PRESENT	HOURLY & DAILY	AUTO & MANUAL	MSC	Mayo
Snow sampling	1SA	AISHIHIK LAKE	08AA-SC03	1994-PRESENT	UP TO 5 days/year	MANUAL	WRB	Aishihik
Snow sampling	2SA	CANYON LAKE	08AA-SC01	1975-PRESENT	UP TO 5 days/year	MANUAL	WRB	Aishihik
Snow sampling	3SA	MACINTOSH	09CA-SC02	1976-2016	UP TO 5 days/year	MANUAL	WRB	Aishihik
Snow sampling	1SM	CALUMET	09DD-SC01	1975-PRESENT	UP TO 5 days/year	MANUAL	WRB	Mayo
Snow sampling	2SM	EDWARDS LAKE	09DD-SC02	1987-2016	UP TO 5 days/year	MANUAL	WRB	Mayo
Snow sampling	3SM	MAYO AIRPORT A	09DC-SC01A	1968-PRESENT	UP TO 5 days/year	MANUAL	WRB	Mayo
Snow sampling	4SM	MAYO AIRPORT B	09DC-SC01B	1987-PRESENT	UP TO 5 days/year	MANUAL	WRB	Mayo
Water Level	1QA	AISHIHIK LAKE NEAR WHITEHORSE	08AA005	1972-2017	5MIN	AUTO	WSC	Aishihik
Flow & water level	2QA	AISHIHIK RIVER BELOW AISHIHIK LAKE	08AA010	1980-2017	5MIN	AUTO	WSC	Aishihik
Flow & water level	3QA	GILTANA CREEK NEAR THE MOUTH	08AA009	1980-2017	5MIN	AUTO	WSC	Aishihik
Flow & water level	4QA	SEKULMUN RIVER AT OUTLET OF SEKULMUN LAKE	08AA008	1981-2017	5MIN	AUTO	WSC	Aishihik
Water Level	5QA	AISHIHIK LAKE NEAR AISHIHIK	08AA012	1995-2015	5MIN	AUTO	WSC	Aishihik
Flow	6QA	INFLOW TO AISHIHIK LAKE	#0000003	1980-2017	DAILY	Reconstructed	YEC	Aishihik
Water Level	1QM	MAYO LAKE NEAR THE OUTLET	09DC005	1979-2017	5MIN	AUTO	WSC	Mayo
Flow	2QM	MAYO LAKE AT THE OUTLET	YECMAYO	1979-2017	5MIN	AUTO	YEC	Mayo
Flow	3QM	INFLOW TO MAYO LAKE	##0000003	1979-2017	DAILY	Reconstructed	YEC	Mayo

*) WRB = Water Resources Branch, Environment Yukon, Government of Yukon, MSC=Meteorological Survey Canada, WSC = Water Survey of Canada, EC=Environment Canada.

Table 2. HYDROTEL sub-model and simulation options

Water budget component (sub-model)	Simulation options/modules
1. Interpolation of meteorological data	1.1 Thiessen polygons
	1.2 Weighted mean of nearest three stations
2. Snow accumulation and melt	2.1 **Mixed (degree-day) energy-budget method**
	2.2 Multi-layer model*
3. Soil temperature and soil freezing	3.1 Rankinen
	3.2 Thorsen
4. Glacier dynamics	4.1 Glacier model*
5. Potential evapotranspiration	5.1 Thornthwaite
	5.2 Linacre
	5.3 Penman
	5.4 Priestley-Taylor
	5.5 Hydro-Québec
	5.6 Penman-Monteith
6. Vertical water budget	**6.1 BV3C**
	6.2 CEQUEAU (modified)
7. Overland water routing	**7.1 Kinematic wave equation**
8. Channel water routing	**8.1 Kinematic wave equation**
	8.2 Diffusive wave equation

Notes: *) the modules are under development

Table 3. Model performances of the (a) HYDROTEL alone simulation (2010-2016) and (b) the comparison of HYDROTEL alone and coupled HYDROTEL-EnKF simulations (2016)

(a) Period of 2010-2016 (HYDROTEL model calibration)

Basin	Site	NSE	PBIAS(%)	RSR	Grade*
Aishihik	Sekulmun River (#08AA008)	0.877	(-9.04)	0.35	Very Good
	Aishihik Lake (#0000003)	0.607	4.61	0.63	Satisfactory
Mayo	Mayo Lake (##0000003)	0.596	26.56	0.64	Satisfactory

(b) Period of 2016 (HYDROTEL vs coupled HYDROTEL -EnKF)

Basin	Site	NSE	PBIAS(%)	RSR	Grade*
Aishihik	Sekulmun River (#08AA008)	0.697 vs 0.991	24.52 vs (-3.78)	0.55 vs 0.09	Satisfactory vs Very Good
	Aishihik Lake (#0000003)	(-0.063) vs 0.305	39.46 vs (-8.298)	1.03 vs 0.834	Unsatisfactory vs Poor**
Mayo	Mayo Lake (##0000003)	0.393 vs 0.958	38.24 vs (-1.967)	0.78 vs 0.206	Unsatisfactory vs Very Good

* Based on the work of Moriasi *et al.* (2007)

** we grade as "Poor" due to PBIAS is at very good grade; but NSE and RSR are at unsatisfactory grades

Table 4. RPSS, optimum warning thresholds and their corresponding hit rate and false alarm rates

Site	Forecast day	RPSS	Below-normal			Above-normal		
			optimal warning threshold	correspond to false alarm rate	correspond to hit rate	optimal warning threshold	correspond to false rate	correspond to hit rate
#08AA008	0	0.885	0.4	0.004	0.992	0.6	0.008	0.983
#08AA008	1	0.871	0.3	0.016	0.983	0.7	0.004	0.983
#08AA008	2	0.855	0.3	0.020	0.975	0.5	0.029	1.000
#08AA008	3	0.836	0.3	0.029	0.967	0.5	0.029	0.992
#08AA008	4	0.814	0.2	0.041	0.967	0.5	0.029	0.983
#08AA008	5	0.788	0.3	0.045	0.942	0.5	0.033	0.975
#08AA008	6	0.759	0.2	0.061	0.950	0.5	0.033	0.967
#08AA008	7	0.728	0.2	0.061	0.942	0.6	0.033	0.958
#08AA008	8	0.698	0.2	0.061	0.942	0.6	0.041	0.950
#08AA008	9	0.672	0.2	0.061	0.942	0.6	0.057	0.942
#08AA008	10	0.650	0.2	0.061	0.933	0.7	0.037	0.917
#08AA008	11	0.632	0.2	0.065	0.933	0.7	0.045	0.900
#08AA008	12	0.615	0.2	0.065	0.917	0.7	0.049	0.892
#08AA008	13	0.600	0.2	0.065	0.908	0.7	0.045	0.883
#08AA008	14	0.588	0.2	0.065	0.908	0.7	0.045	0.867
#0000003	0	0.146	0.2	0.355	0.717	0.7	0.173	0.754
#0000003	1	0.204	0.3	0.253	0.667	0.7	0.161	0.746
#0000003	2	0.211	0.4	0.233	0.633	0.7	0.144	0.721
#0000003	3	0.206	0.4	0.265	0.683	0.6	0.198	0.746
#0000003	4	0.196	0.4	0.314	0.683	0.5	0.231	0.795
#0000003	5	0.185	0.5	0.241	0.617	0.5	0.218	0.779
#0000003	6	0.181	0.5	0.265	0.633	0.5	0.206	0.762
#0000003	7	0.186	0.5	0.265	0.633	0.5	0.206	0.762
#0000003	8	0.187	0.5	0.286	0.642	0.5	0.193	0.746
#0000003	9	0.183	0.5	0.286	0.642	0.5	0.181	0.746
#0000003	10	0.180	0.5	0.302	0.675	0.3	0.268	0.844
#0000003	11	0.179	0.5	0.298	0.683	0.3	0.255	0.836
#0000003	12	0.184	0.5	0.298	0.675	0.4	0.193	0.771
#0000003	13	0.188	0.5	0.314	0.683	0.3	0.243	0.820
#0000003	14	0.187	0.5	0.306	0.675	0.3	0.247	0.820
##0000003	0	0.717	0.1	0.124	0.915	0.6	0.034	0.987
##0000003	1	0.698	0.2	0.141	0.881	0.5	0.032	0.971
##0000003	2	0.644	0.3	0.152	0.885	0.4	0.044	0.945
##0000003	3	0.610	0.4	0.152	0.864	0.2	0.083	0.954
##0000003	4	0.581	0.4	0.171	0.868	0.2	0.080	0.916
##0000003	5	0.554	0.6	0.141	0.826	0.2	0.074	0.916
##0000003	6	0.516	0.6	0.166	0.830	0.1	0.106	0.949
##0000003	7	0.483	0.6	0.187	0.834	0.1	0.121	0.954
##0000003	8	0.444	0.6	0.202	0.847	0.1	0.121	0.928
##0000003	9	0.405	0.7	0.177	0.804	0.1	0.144	0.920
##0000003	10	0.356	0.6	0.244	0.872	0.1	0.144	0.903
##0000003	11	0.314	0.6	0.253	0.894	0.1	0.152	0.907
##0000003	12	0.287	0.7	0.223	0.847	0.1	0.159	0.886
##0000003	13	0.256	0.7	0.225	0.864	0.1	0.163	0.878
##0000003	14	0.231	0.7	0.225	0.868	0.1	0.173	0.857

Figure captions

Figure 1. (a) Locations of the study basins (Aishihik and Mayo basins) in North America; (b) Aishihik (left Figure) and Mayo (right Figure) basins overlaid on DEM, major lakes, and locations of stations #08AA008 (Sekulmun River subbasin of Aishihik basin measuring river flows, left), #0000003 (measuring reservoir inflow to Aishihik Lake, left), ##0000003 (measuring reservoir inflows to Mayo Lake in the Mayo basin, right) sites.

Figure 2. Estimates of soil type maps of (a) Aishihik and (b) Mayo

Figure 3. Hydro-meteorological networks of (a) Aishihik and (b) Mayo. Labels shown on these Figures refer to Table 1 column "Code"

Figure 4. General Procedure of EnKF

Figure 5. Integration of NAEFS, HYDROTEL and EnKF for short-range flow and inflow forecasts.

Figure 6. HYDROTEL model calibration for the period of 2010-2016 at site (a) #08AA008 (Sekulmun River, Aishihik), (b) #0000003 (Aishihik), and (c) ##0000003 (Mayo).

Figure 7. Comparison of observed and modelled flows/inflows generated using HYDROTEL alone and using the coupled HYDROTEL-EnKF approach at sites (a) #08AA008 (Sekulmun River, Aishihik), (b) #0000003 (Aishihik), and (c) ##0000003 (Mayo).

Figure 8. Flow forecasts for 1-, 3-, 7-, and 14-day ahead using raw NAEFS data for the period of 2016 at site #08AA008 (Sekulmun River, Aishihik).

Figure 9. Flow/inflow forecasts for 1-, 3-, 7-, and 14-day ahead using corrected NAEFS data for the period of 2016 at site (a) #08AA008 (Sekulmun River, Aishihik), (b) #0000003 (Aishihik), and (c) ##0000003 (Mayo).

Figure 10. Percentage change of absolute daily volume errors (left) and daily volume errors (right) for different forecasting periods at sites (a)#08AA008 (Sekulmun River, Aishihik), (b)#0000003 (Aishihik), and (c) ##0000003 (Mayo)

Figure 11. ROC curves at 1-, 3-, 7- and 14-day ahead forecasts at sites (a)#08AA008 (Sekulmun River, Aishihik), (b)#0000003 (Aishihik), and (c) ##0000003 (Mayo)

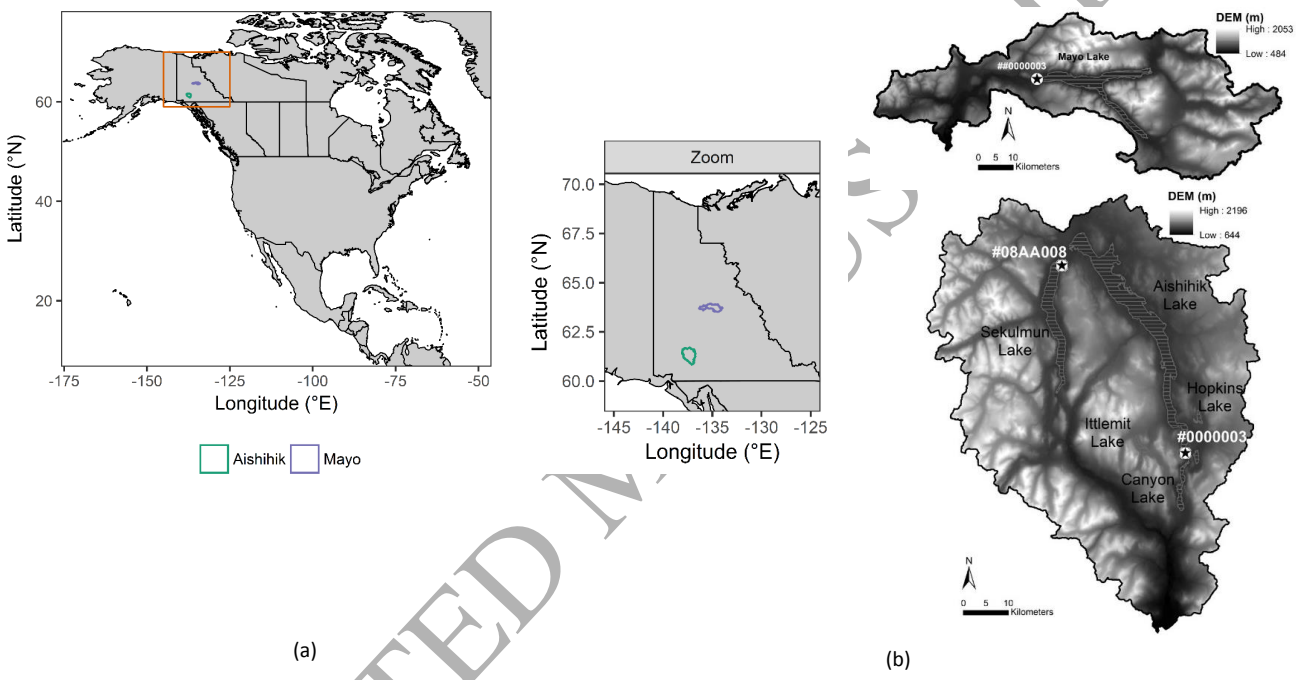


Figure 1. (a) Locations of the study basins (Aishihik and Mayo basins) in North America; (b) Aishihik (left Figure) and Mayo (right Figure) basins overlaid on DEM, major lakes, and locations of stations #08AA008 (Sekulmun River subbasin of Aishihik basin measuring river flows, left), #0000003 (measuring reservoir inflow to Aishihik Lake, left), ##0000003 (measuring reservoir inflows to Mayo Lake in the Mayo basin, right) sites

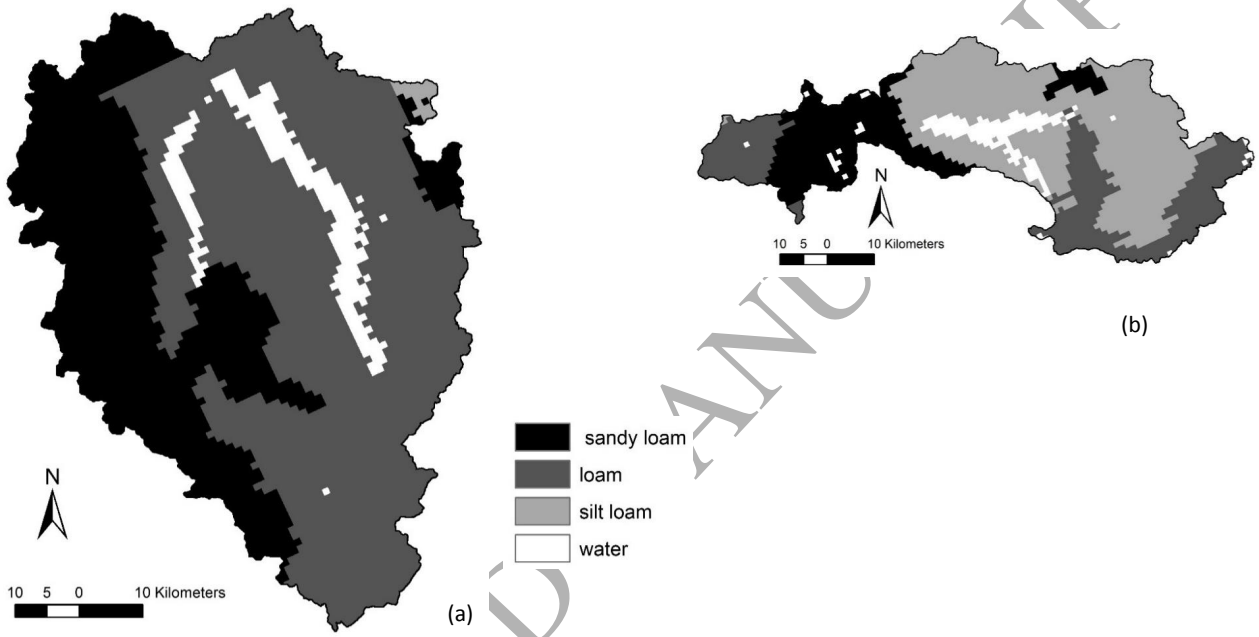
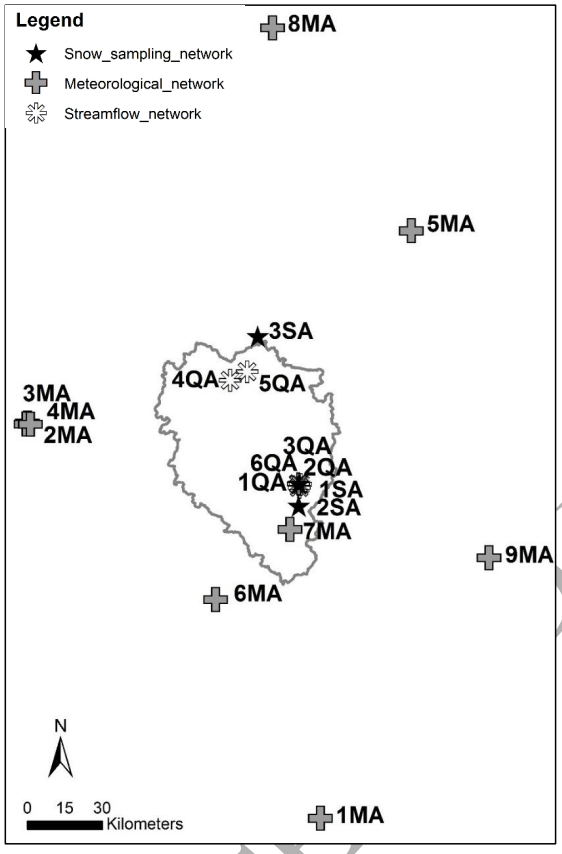
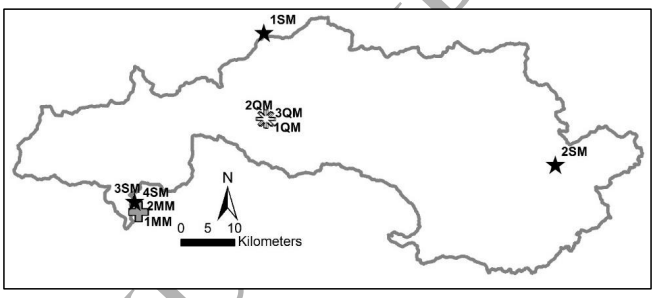


Figure 2. Estimates of Soil type maps of (a) Aishihik and (b) Mayo


Legend
Snow_sampling_network
Meteorological_network
Streamflow_network
8MA
5MA
3SA
4QA
5QA
3MA
4MA
2MA
3QA
6QA
2QA
1QA
1SA
2SA
7MA
9MA
6MA
1MA
N
0 15 30
Kilometers

(a)


1SM
2QM
3QM
1QM
2SM
3SM
4SM
2MM
1MM
N
0 5 10
Kilometers

(b)

Figure 3. Hydro-meteorological networks of (a) Aishihik and (b) Mayo. Labels shown on these Figures refer to Table 1 column "Code"

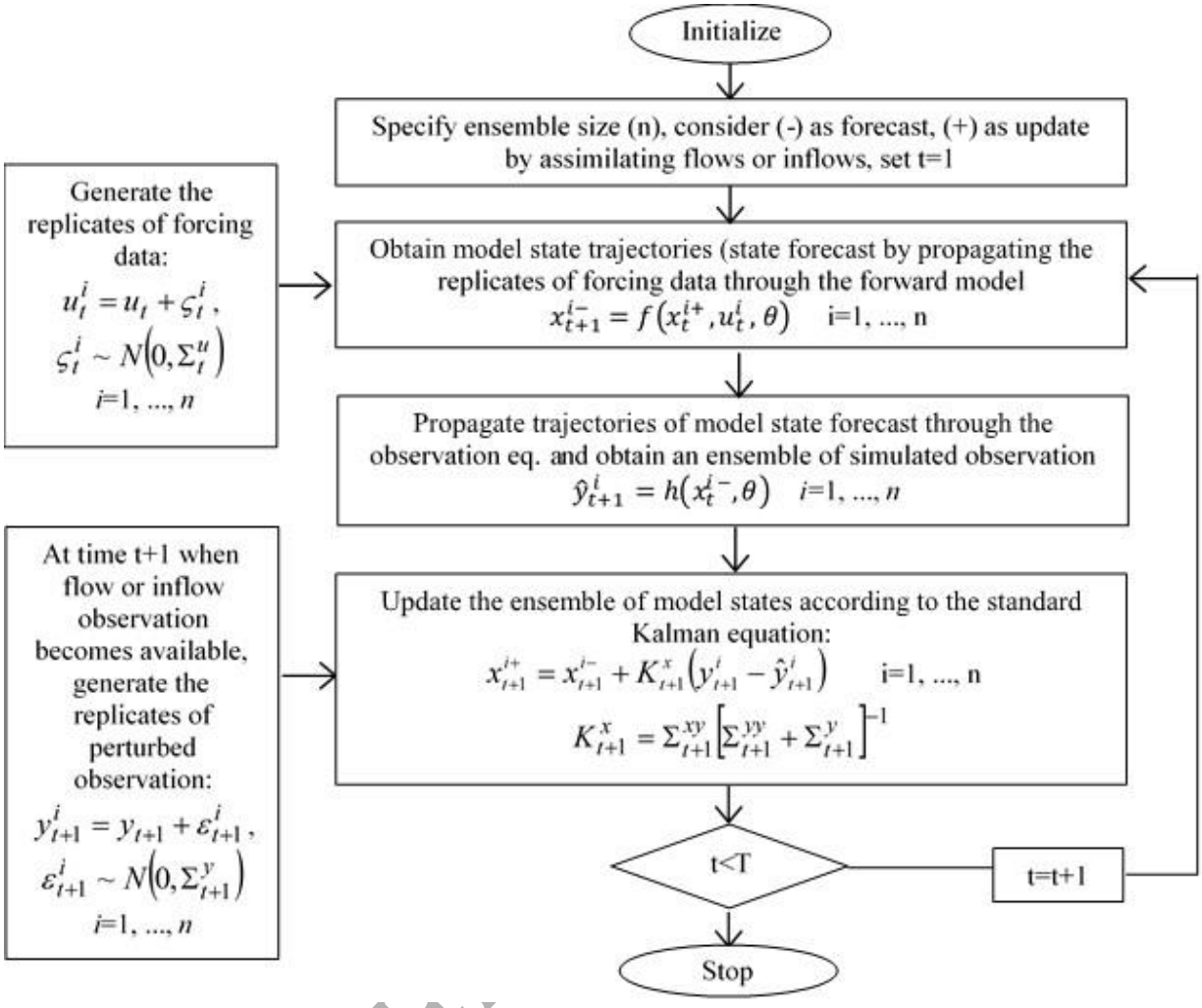


Figure 4. General Procedure of EnKF

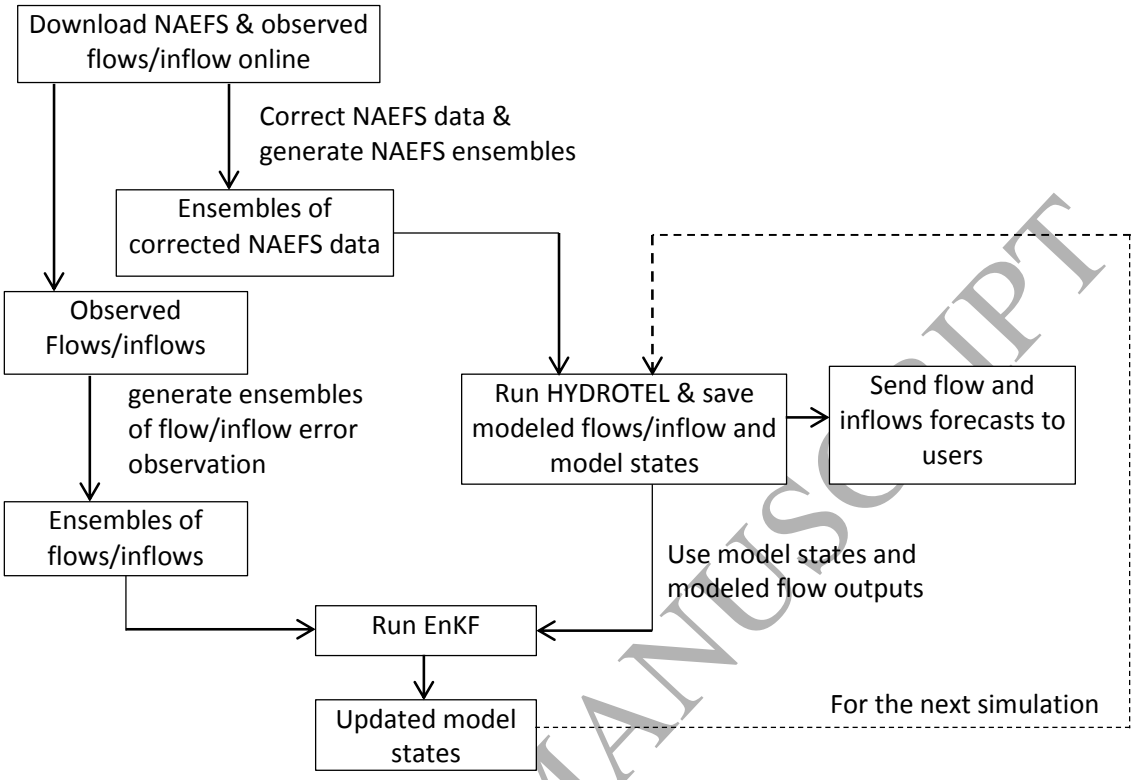
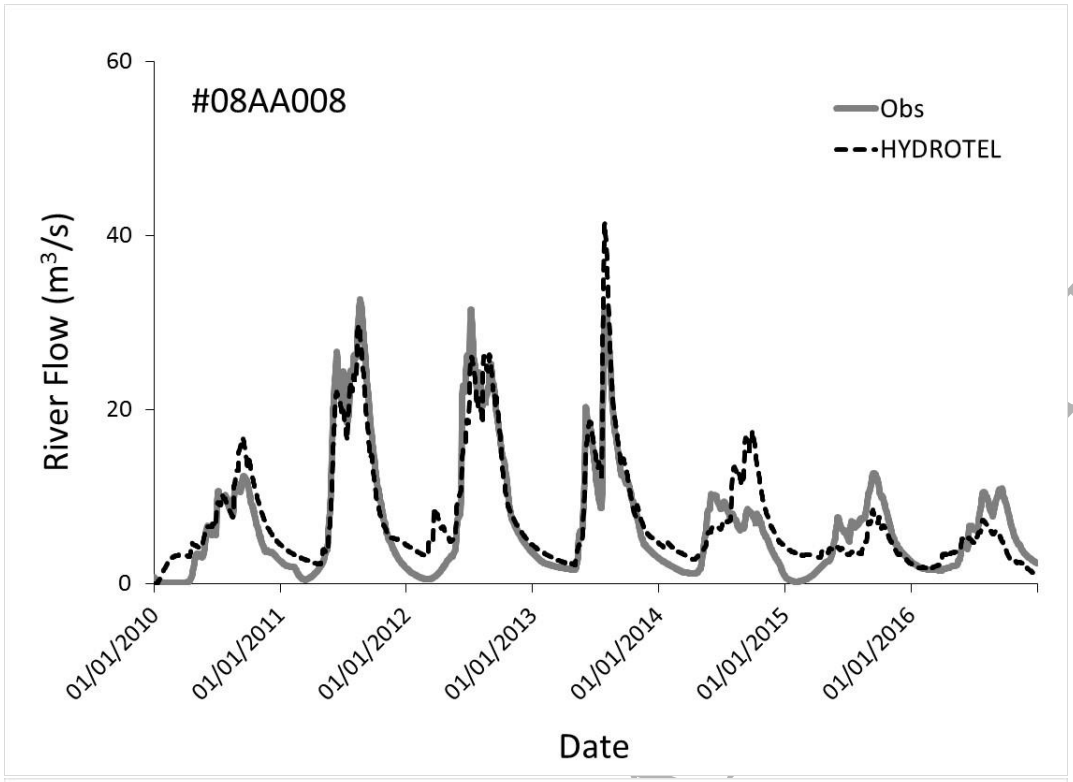
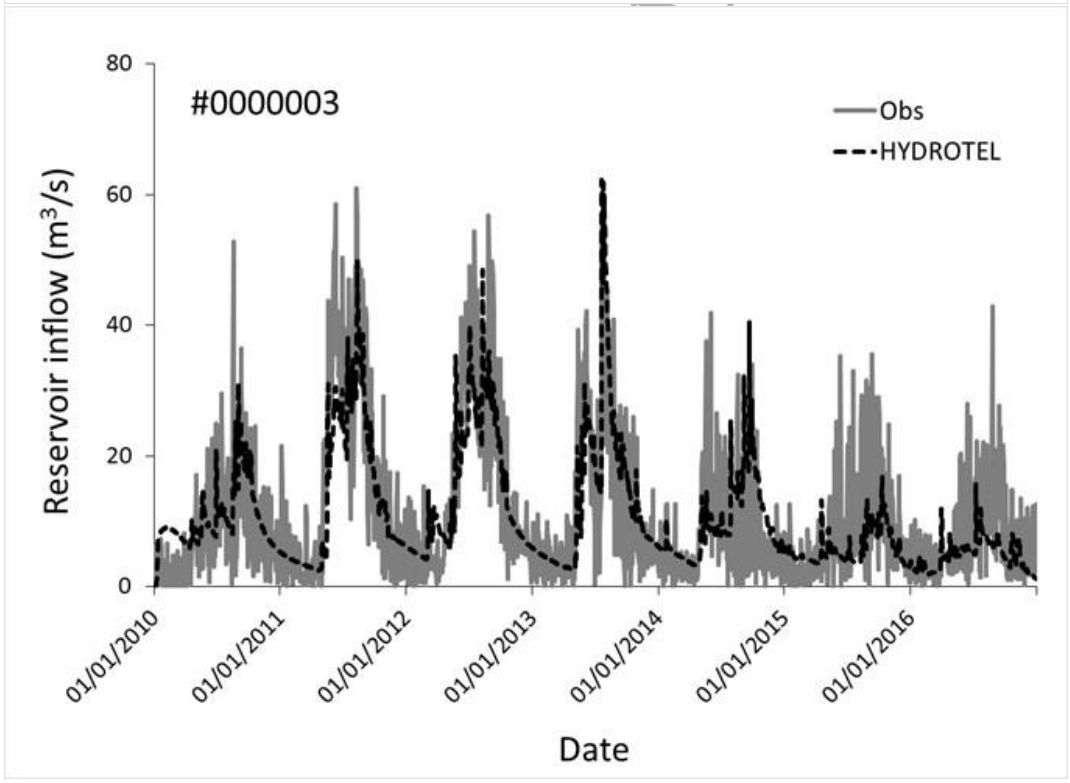


Figure 5. Integration of NAEFS, HYDROTEL and EnKF for short-range flow and inflow forecasts.

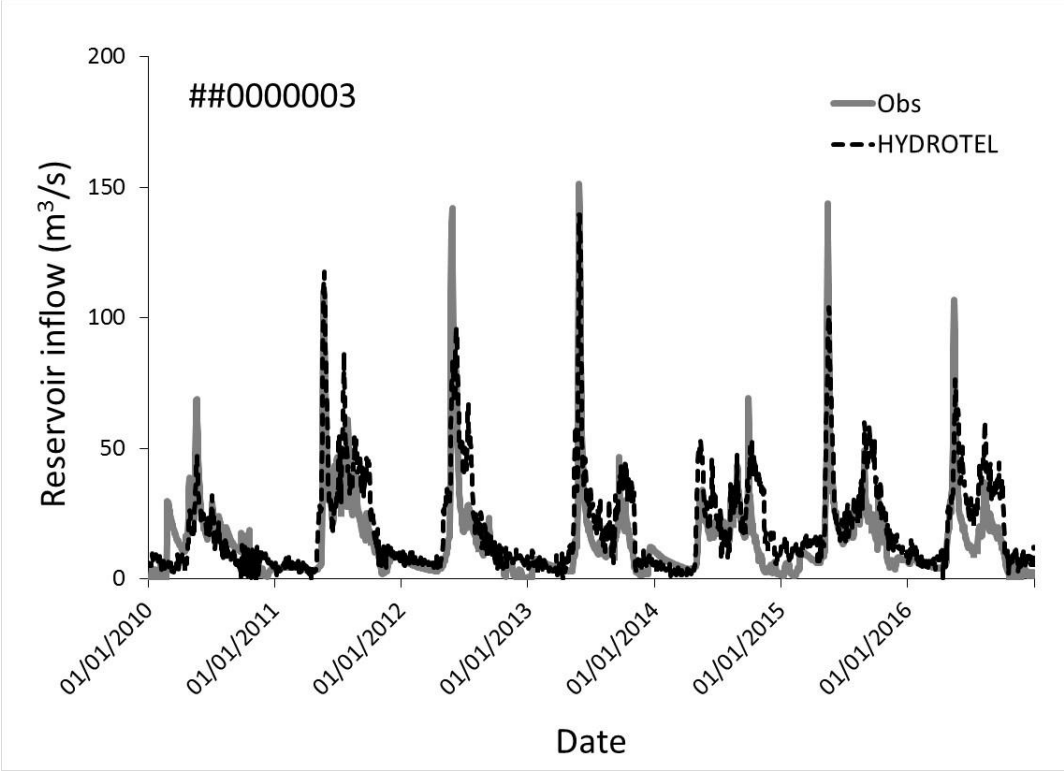


(a)



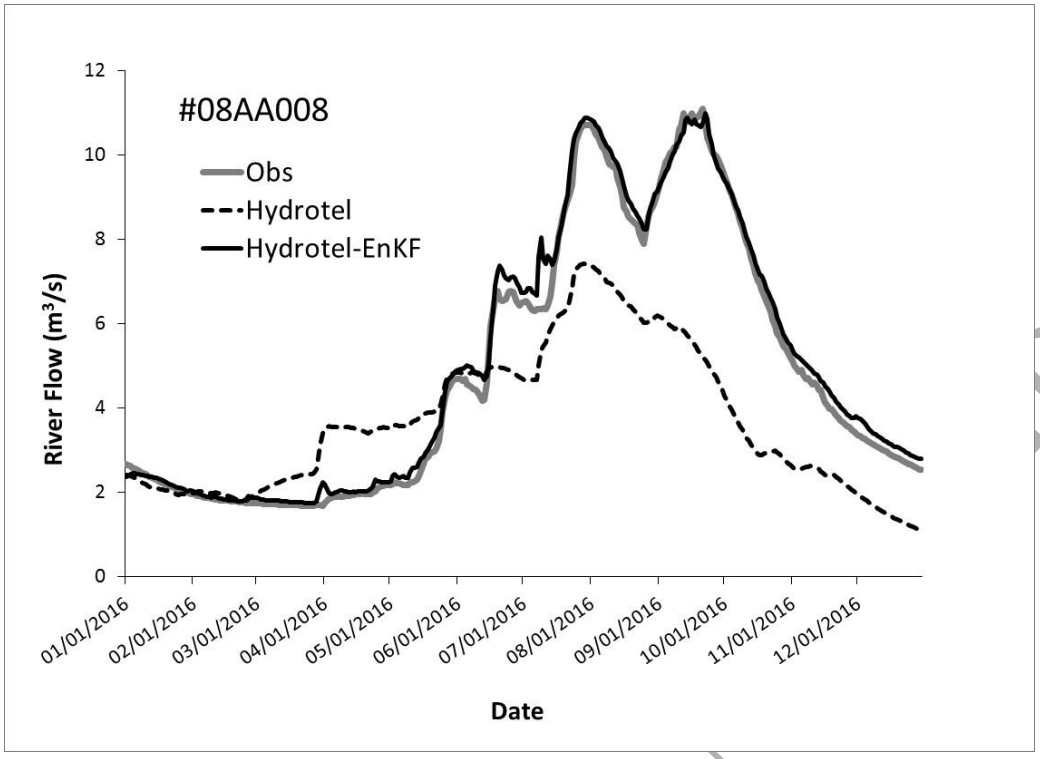
(b)

Continued below …

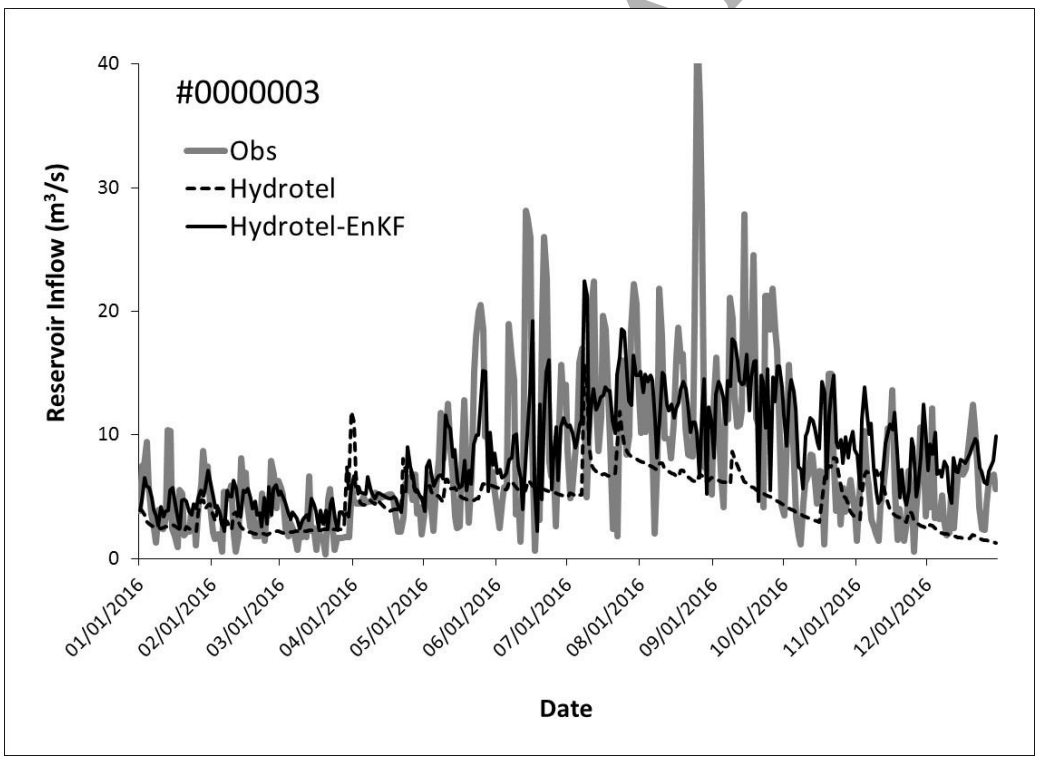


(c)

Figure 6. HYDROTEL model calibration for the period of 2010-2016 at site (a) #08AA008 (Sekulmun River, Aishihik), (b) #0000003 (Aishihik), and (c) ##0000003 (Mayo).

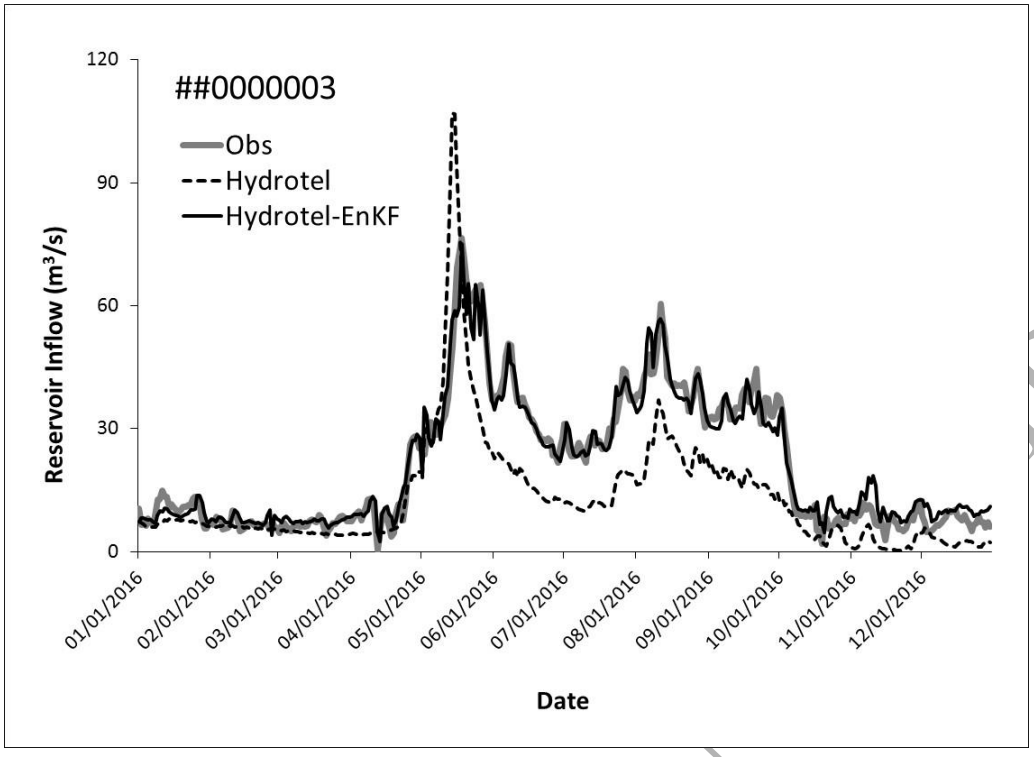


(a)



(b)

Continued below …



(c)

Figure 7. Comparison of observed and modelled flows/inflows generated using HYDROTEL alone and using the coupled HYDROTEL -EnKF approach at sites (a) #08AA008 (Sekulmun River, Aishihik), (b) #0000003 (Aishihik), and (c) ##0000003 (Mayo).

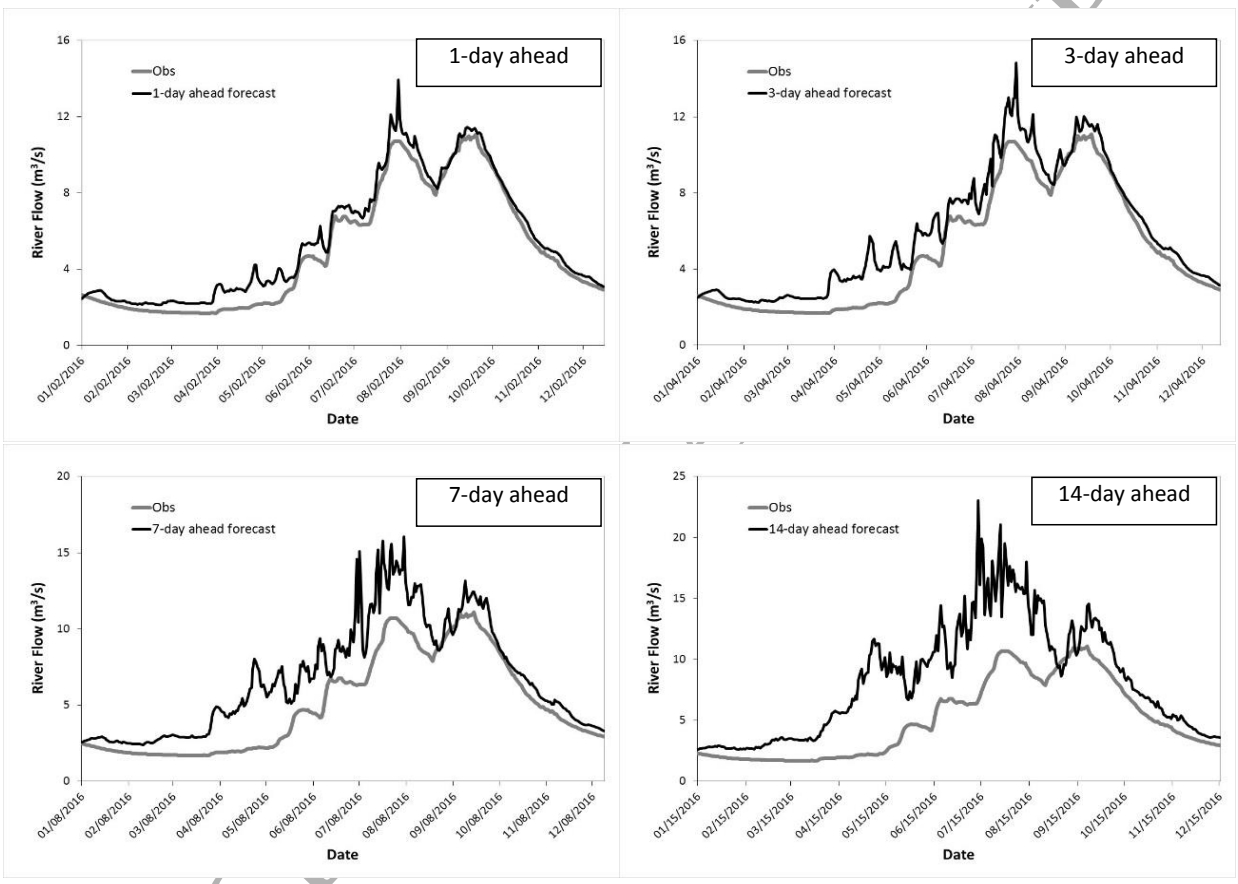
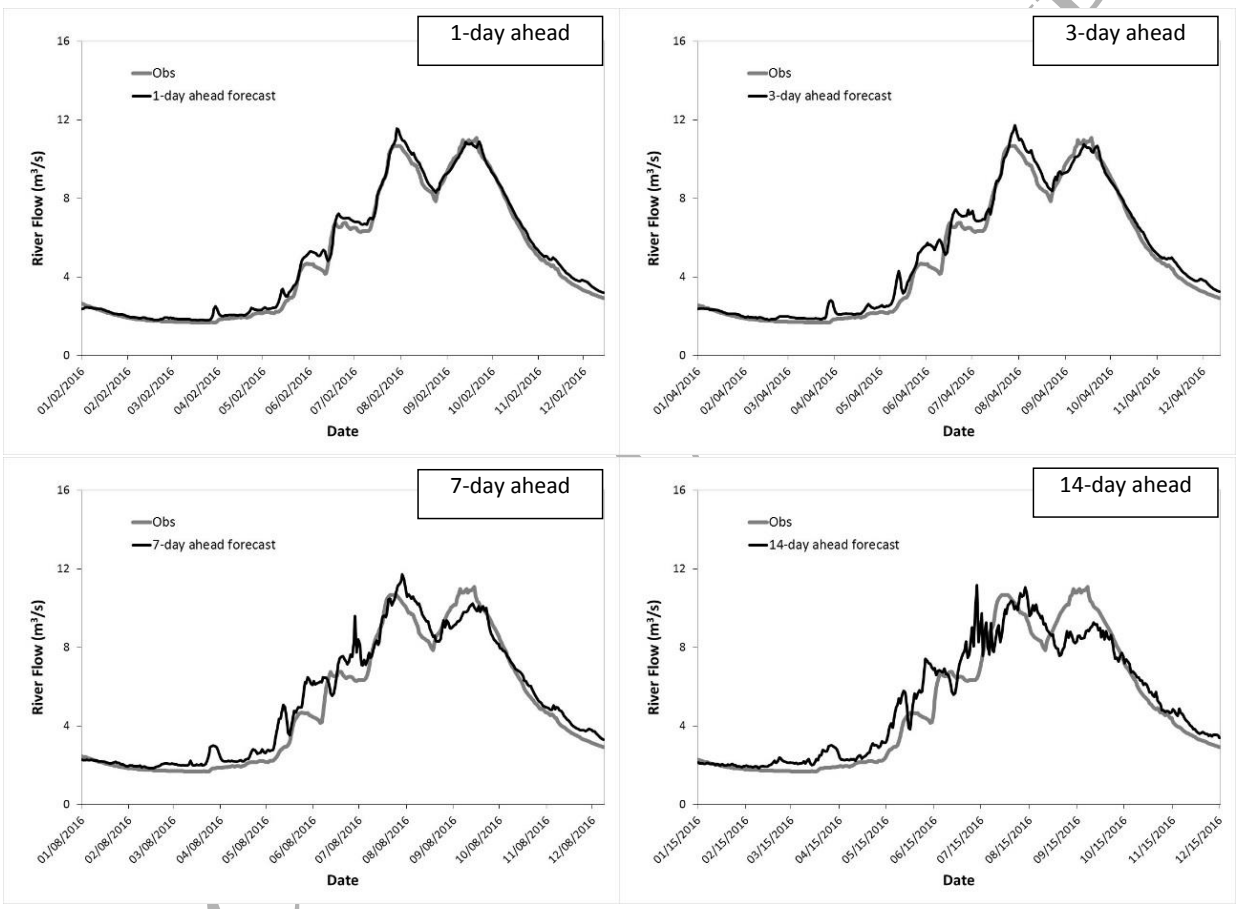
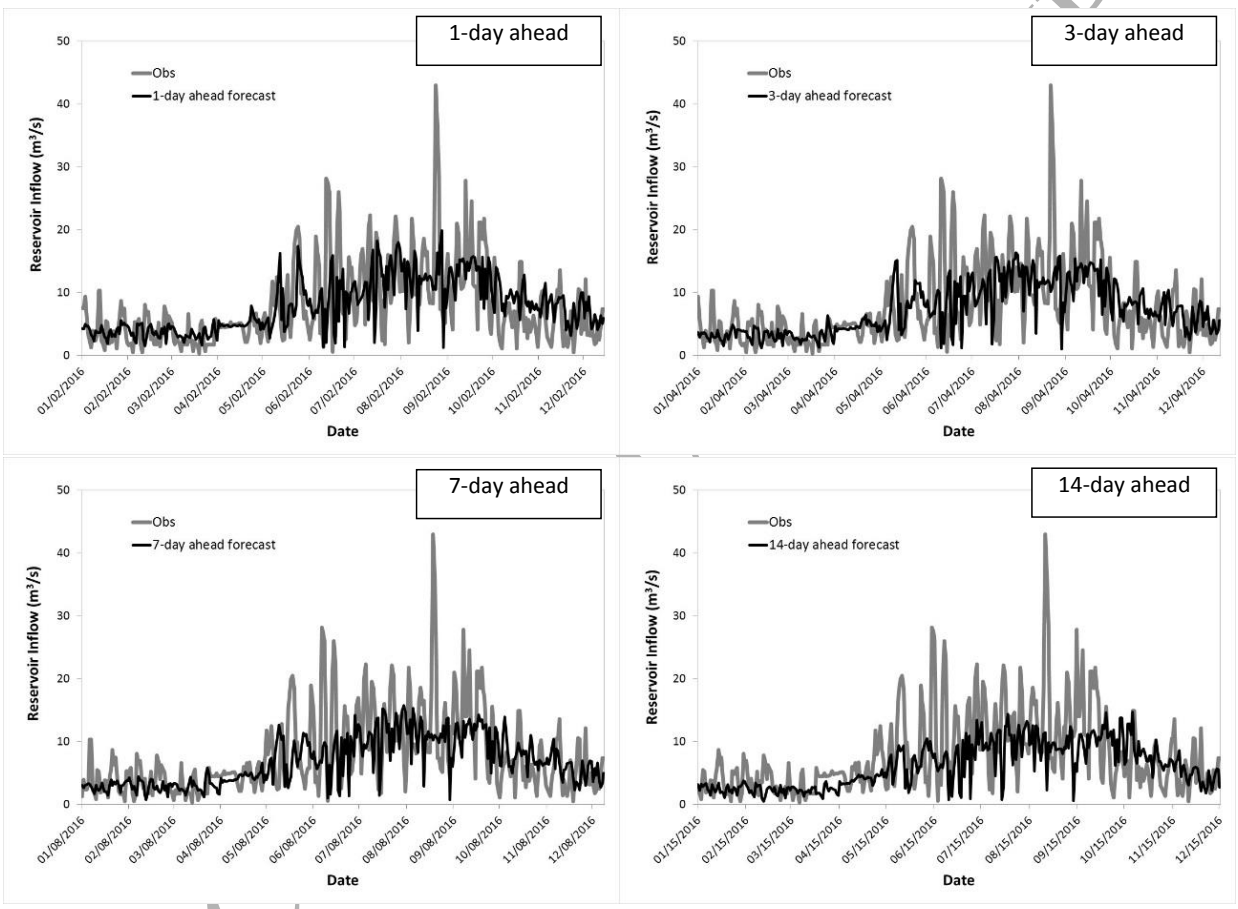


Figure 8. Flow forecasts for 1-, 3-, 7-, and 14-day ahead using raw NAEFS data for the period of 2016 at site #08AA008 (Sekulmun River, Aishihik).



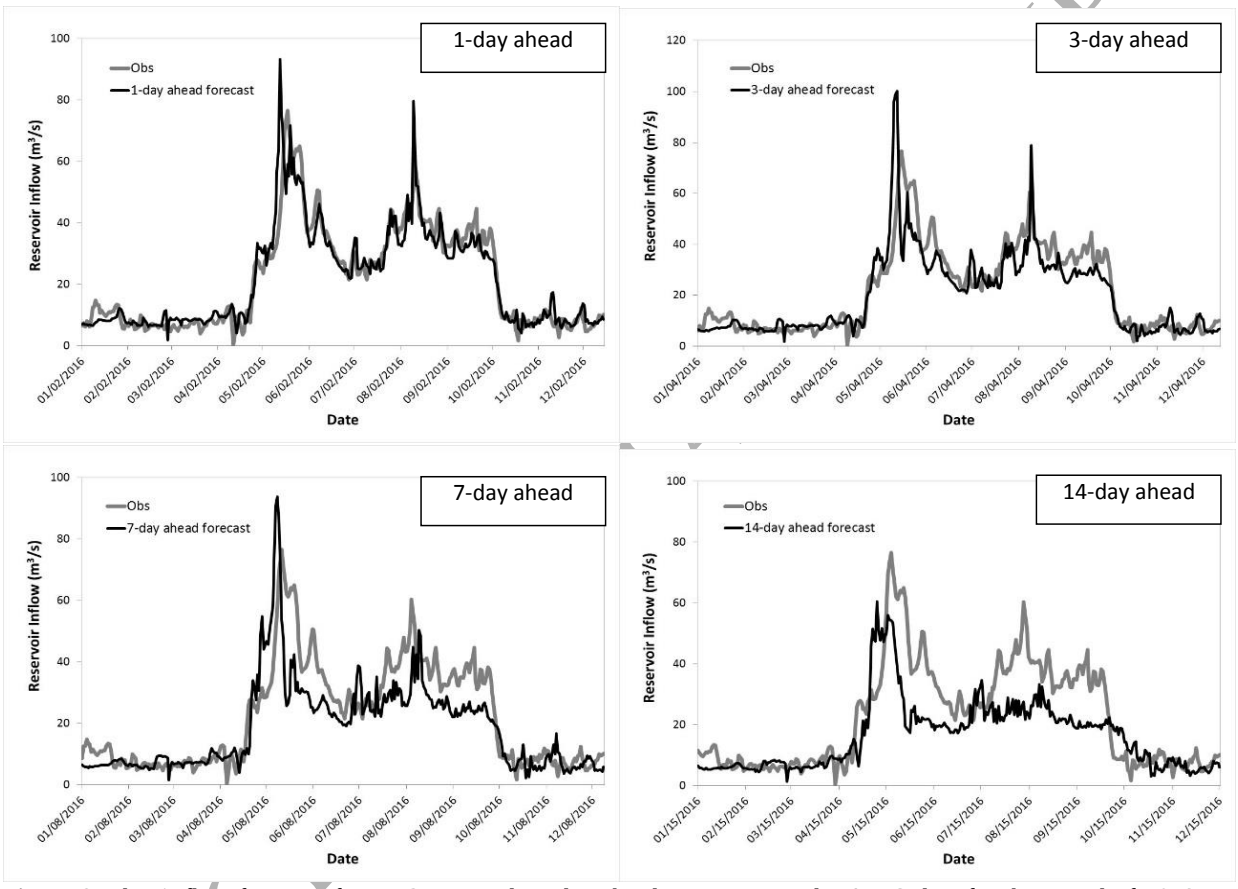
(a)

Continued below ….



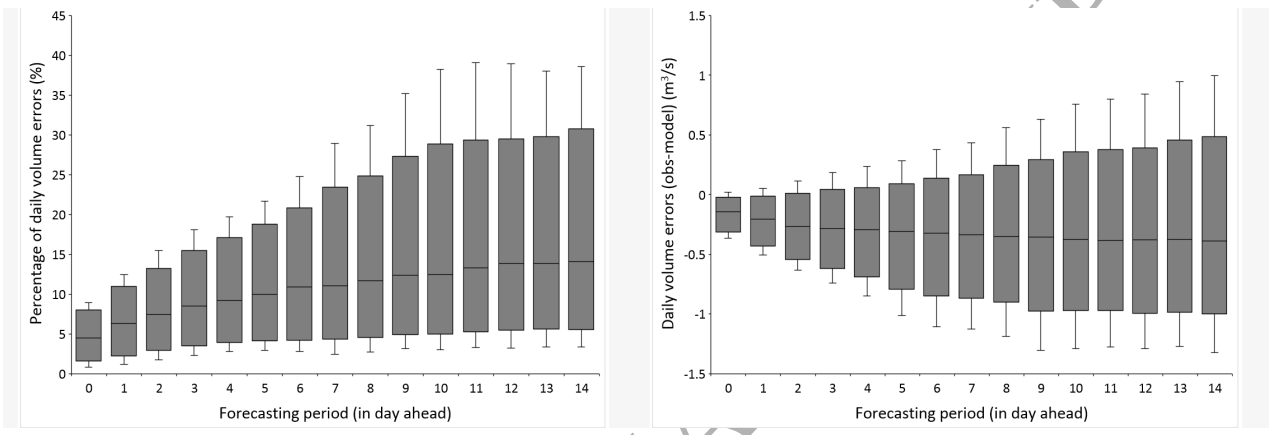
(b)

Continued below ….

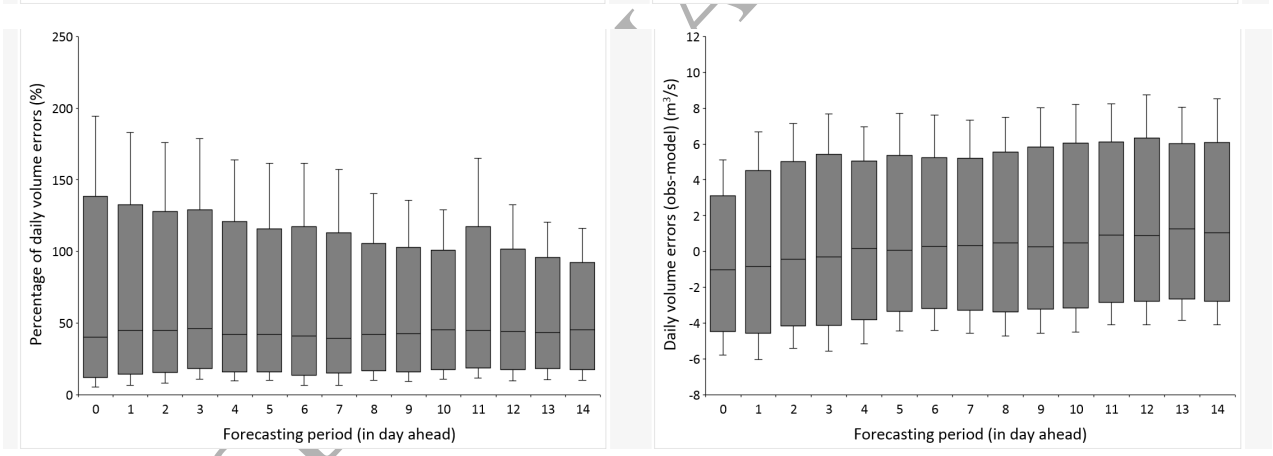


(c)

Figure 9. Flow/inflow forecasts for 1-, 3-, 7-, and 14-day ahead using corrected NAEFS data for the period of 2016 at site (a) #08AA008 (Sekulmun River, Aishihik), (b) #0000003 (Aishihik), and (c) ##0000003 (Mayo).

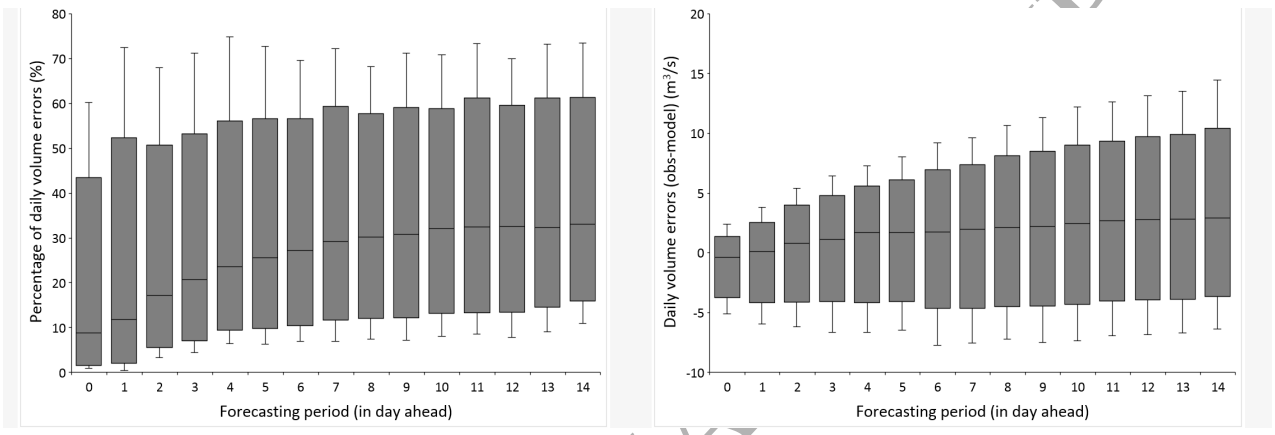


(a)



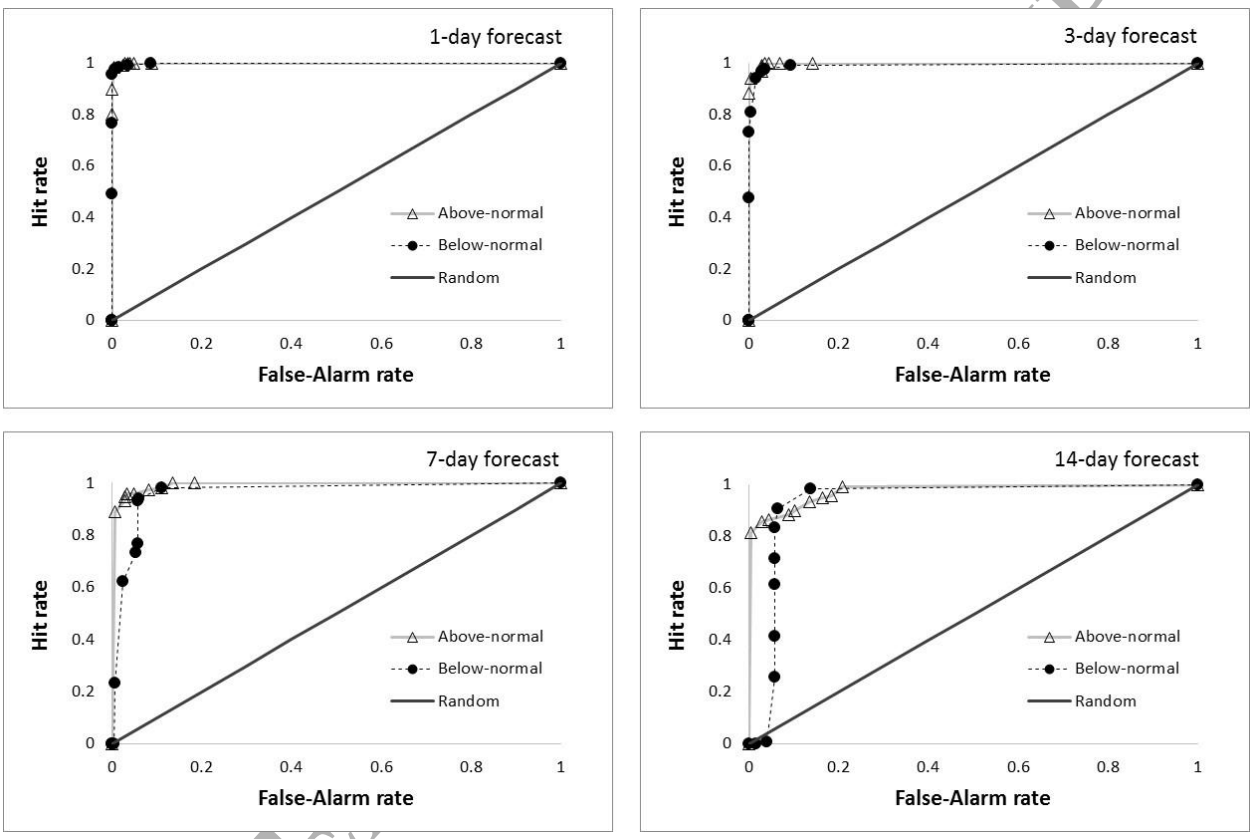
(b)

Continued below ….



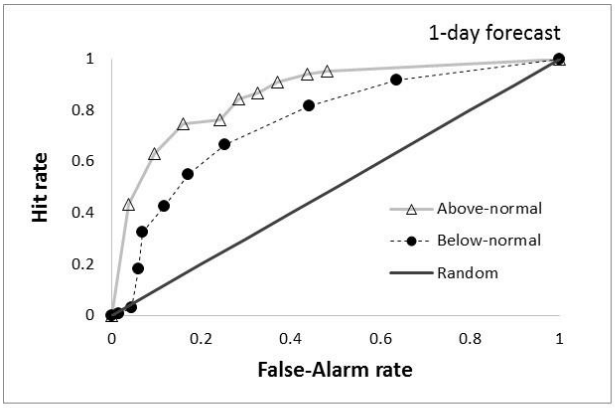
(c)

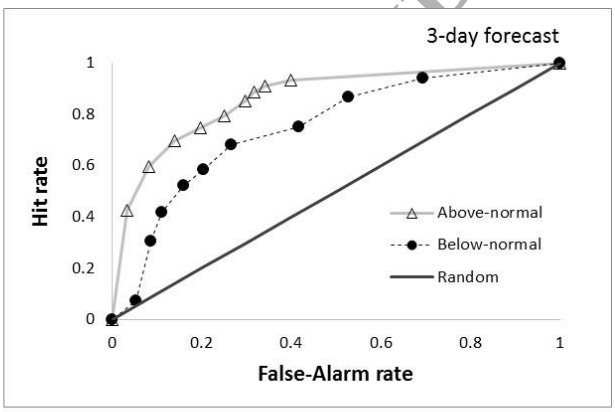
Figure 10. Percentage change of absolute daily volume errors (left) and daily volume errors (right) for different forecasting periods at sites (a)#08AA008 (Sekulmun River, Aishihik), (b)#0000003 (Aishihik), and (c) ##0000003 (Mayo)

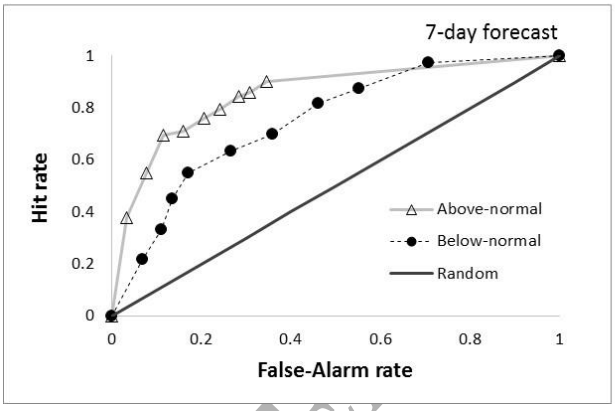


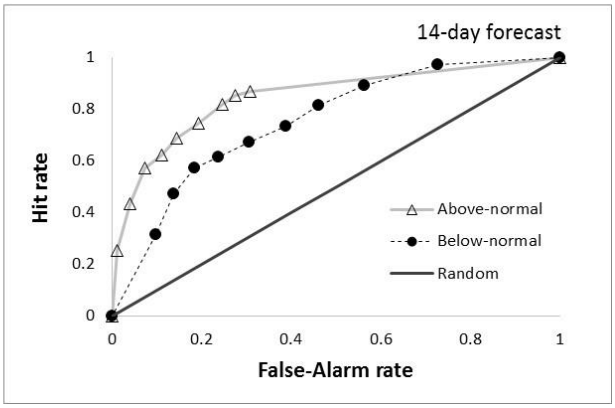
(a)

Continued below ….



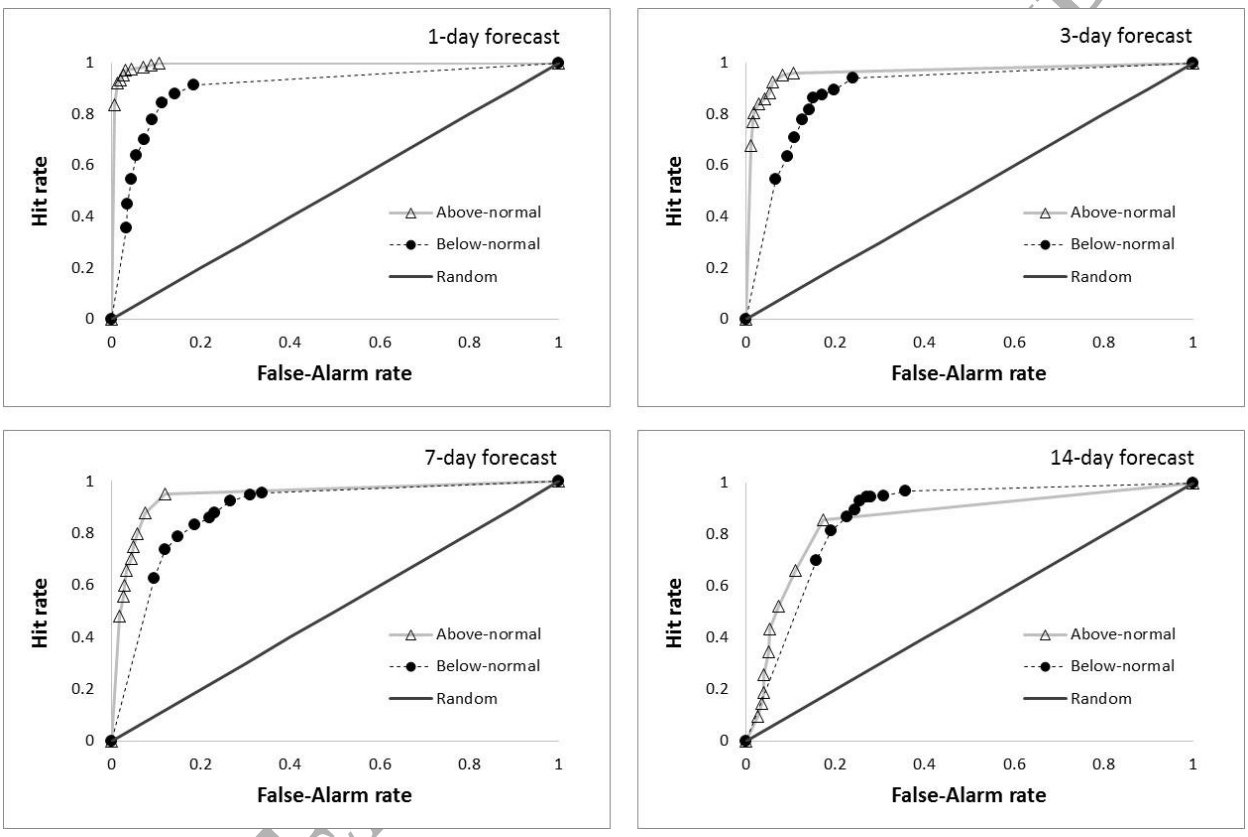






(b)

Continued below ….



(c)

Figure 11. ROC curves at 1-, 3-, 7- and 14-day ahead forecasts at sites (a)#08AA008 (Sekulmun River, Aishihik), (b)#0000003 (Aishihik), and (c) ##0000003 (Mayo)