Research papers

Streamflow depletion from groundwater pumping in contrasting hydrogeological landscapes: Evaluation and sensitivity of a new management tool

Qiang Li[a,*], Samuel C. Zipper[b], Tom Gleeson[a]

[a] Department of Civil Engineering, University of Victoria, 3800 Finnerty Road, Victoria, British Columbia V8P 4C2, Canada
[b] Kansas Geological Survey, University of Kansas, Lawrence, KS, USA

ARTICLE INFO

ABSTRACT

Groundwater pumping can reduce streamflow by reducing groundwater discharge and/or inducing streamflow infiltration, which together are referred to as streamflow depletion. Recently, analytical depletion functions (ADFs) have been suggested as rapid and accurate tools for streamflow depletion assessment, but their performance has only been tested in a few hydrogeological settings. To evaluate whether they will be useful tools for other regions with contrasting stream network and hydrogeological characteristics, we compared ADFs to calibrated MODFLOW models in BX Creek and Peace region with distinct hydrogeological settings (interior plateaus & highlands, and boreal plains, respectively) and spatial scales (165 km^2 and 1952 km^2, respectively) in British Columbia, Canada. Results showed that ADFs can accurately identify most affected streams by pumping for 100% and 83% of wells in the BX Creek and Peace region, respectively, and had small prediction errors compared with MODFLOW. Specifically, the mean absolute error of predicted depletion ranged from 2% to 14% of the highest simulated pumping rate over the study period of 30 years, with improved accuracy during the pumping season. We also found different responses of ADF performance to hydrostratigraphic properties such as hydraulic conductivity, aquifer thickness, streambed conductance, and well depth across two domains, indicating that different factors control ADF accuracy in different hydrogeological settings. Therefore, we conclude that ADFs are useful tools for conjunctive water management, but a good understanding of local hydrogeological conditions is needed to address the potential uncertainty of ADFs for decision-making.

1. Introduction

Groundwater is a critical resource for society and aquatic ecosystems (Gleeson and Richter, 2018; de Graaf et al., 2019; Kurylyk et al., 2015; Power et al., 1999). Groundwater pumping can negatively affect rivers, lakes, wetlands, and other surface water bodies by reducing groundwater discharge into them or, in severe cases, inducing infiltration through the streambed. Combined, these impacts are known as streamflow depletion (Barlow and Leake, 2012; Zipper et al., 2019A). With anticipated growth in demand for agricultural irrigation due to climate change (Aeschbach-Hertig and Gleeson, 2012; Rodell et al., 2009; Wada et al., 2014) combined with other stressors, e.g. forest logging and fires (Kiffney et al., 2002; Wei et al., 2018), understanding the impacts of streamflow depletion is critically important for regional and global water management.

Unfortunately, streamflow depletion is impractical to measure directly because of limited streamflow monitoring data and because pumping-induced reductions in streamflow are superimposed on top of weather-driven flow variability (Barlow and Leake, 2012; Gleeson and Richter, 2018). As a result, streamflow depletion is typically estimated using numerical and/or analytical models (Chen, 2000; Huang et al., 2018; Schneider et al., 2017; Zipper et al., 2019B). Numerical models are often used for site-specific investigations and require substantial effort, data, and knowledge to calibrate and validate. Sometimes, water managers cautiously transfer the knowledge gained from one location to another as the differences in hydrogeological and climate conditions create significant uncertainty, which hinders decision-making in locations that have not been extensively studied. In contrast, analytical models can provide a quick assessment of streamflow depletion with relatively low data requirements (Huang et al., 2018; Huggins et al., 2018). Several studies have compared the performance between the numerical and analytical models and concluded that analytical models

---

are conservative tools for water managers to assess the pumping impacts on streamflow depletion (Huggins et al., 2018; Rathfelder, 2016; Zipper et al., 2018; Zipper et al., 2019A). However, analytical models have many simplifying assumptions, for instance, the commonly-used Glover analytical model assumed an infinite horizontal aquifer bounded by a single linear stream. Additionally, streams and their underlying aquifer are hydraulically connected and do not become dry over the pumping period (Glover and Balmer, 1954). The details of other analytical depletion models can be found in a review paper by Huang et al. (2018). These assumptions could lead to a biased assessment of streamflow depletion, which further hinder their application in real-world hydrogeological settings with complex stream networks of multiple, meandering, and different types (i.e., perennial and ephemeral streams) of stream segments.

To advance the application of analytical models in real-world settings with multiple stream segments and complex stream networks, Zipper et al. (2019A) developed analytical depletion functions (ADFs), which combine (1) *stream proximity criteria*, to determine which stream segments are most likely to be affected by a pumping well; (2) *a depletion apportionment equation*, a geometric method to distribute depletion among the affected stream segments; and (3) *an analytical model*, to calculate the amount of depletion for all impacted stream segments based on the previous two components. Zipper et al. (2018) conducted a systematic evaluation of depletion apportionment equations under steady-state conditions for Nanaimo, British Columbia. Zipper et al. (2019A) evaluated 50 ADFs in the Navarro River watershed, California, by comparing ADFs against a numerical groundwater model and found the combination of stream proximity criteria, a depletion apportionment equation, and analytical model which best matched numerical model results under transient pumping. The same ADF also performed the best in a comparison with the Republican River Compact Administration groundwater model, a regional-scale calibrated groundwater model used for streamflow depletion decision-making in a heavily stressed aquifer (Zipper et al., 2020).

These previous studies (e.g., Zipper et al., 2018, 2020; Zipper et al., 2019A) have found that the streamflow depletion predicted by the best performed ADFs provide comparable estimate to numerical groundwater models. However, to date ADFs have only been tested in these three limited domains, and there have been no systematic cross-regional comparisons to determine how ADF performance varies across different hydrogeological conditions and what factors most strongly influence ADF performance. Therefore, additional testing against calibrated numerical models is needed to evaluate whether ADFs will perform similarly across regions with different stream network geometries, hydrogeological characteristics, and bioclimatic conditions, which is necessary to evaluate whether ADFs are suitable for decision-making in real-world settings. To address these knowledge gaps, the goals of this study are to: 1) evaluate the accuracy of ADFs for estimating impacts of groundwater pumping on streamflow in two contrasting hydrogeological settings in British Columbia, Canada; and 2) understand the accuracy and sensitivity of ADFs in evaluating streamflow depletion to advance the application of analytical depletion function in real world settings.

2. Methods

To assess the performance of ADFs across hydrogeological settings, we selected two study domains within British Columbia that have contrasting hydrogeology, climate, topography, and ecology (Fig. 1). For these two domains, we modified existing calibrated numerical models, built in MODFLOW, to simulate the pumping impacts on streamflow depletion for comparison with ADFs. We treated streamflow depletion simulated by the numerical models as reference or "observed" values for comparison with ADF output since numerical models include more detailed process-based representation of subsurface flow. While both models were calibrated to field observations, we acknowledge that

even numerical models are an imprecise mathematical representation of reality.

2.1. Study domains

Seven hydrogeological landscapes have been classified based on the physiographic, groundwater regions, and biogeoclimatic zones in British Columbia, Canada (Fig. 1). In this study, we selected two domains in contrasting settings based on numerical model availability and consistency: BX Creek model represents interior plateaus and highlands and the Peace region model represents boreal plains. In addition, these domains have contrasting aquifer and stream network characteristics, as BX Creek represents a small aquifer (165 km^2) with a relatively simple stream network while the Peace region model represents a large regional aquifer (1952 km^2) with a complex stream network.

2.1.1. BX Creek numerical model

The numerical model in the BX Creek was initially developed to validate a long-term recharge rate estimation method for mountainous regions (Smerdon et al., 2009A). BX Creek is characterized by snow-melt-dominated uplands and a dry valley bottom. In the uplands, surface runoff and high flows occur during the snow-melt seasons and groundwater recharge is minimal. Springs and groundwater seepage occur at mid-elevations. The aquifer in valley bottom is recharged by surface runoff as well as receiving local and regional groundwater flows (Fig. 1; (Smerdon et al., 2009B). Model development, conceptualization, and parameterization are described in detail in Smerdon et al. (2009A). The model configuration including spatial distribution of hydraulic conductivity and model layers can be found in Supplementary Information Figs. S1–S3. Here we briefly review the model settings. The model has a uniform grid resolution of 50 × 50 m with 327 rows, 400 columns, and 8 layers with lateral boundaries between layers following the land surface (Fig. S3). The total model domain has an area of 165 km^2. Based on the borehole logs from 611 water wells, the subsurface is parameterized as four hydrostratigraphic units (Table 1, Fig. S2).

The original BX Creek model includes five boundary conditions: constant head (BAS package in MODFLOW), general head (GHB), drain (DRN), river (RIV), and stream (STR) (Fig. S1). As shown in Fig. S1, there are four blocks of drains in the domain to represent seepage faces in the mid-elevation regions, which tend to be found in areas with steep slopes near the valley floor. However, analytical approaches are not designed to represent seepage faces (Huang et al., 2018). We found that removing these drain blocks had minor impacts on the model mass balance, so they were removed from the numerical model (Fig. S1) while keeping the other boundary conditions.

The steady-state model was originally calibrated against hydraulic heads data from 196 groundwater wells, of which 22 were located closely and in a single model cell. The root-mean-square-error of the calibration was 44.6 m, which is 3.6% of the observed hydraulic head range in the watershed. The simulated groundwater discharge from the upland creek (drain boundary) was 90% of average observed low flow. Additionally, this model's solute transport processes were also calibrated. Simulated hydrochemical data also showed strong agreement with the observed data within the domain, providing the further confidence in model performance (Smerdon et al., 2009A). Overall, this calibration suggests that the numerical model represents essential groundwater flow processes in the watershed.

2.1.2. Peace region numerical model

In the boreal plains, surface runoff is minor due to gentle undulating topography which retains water on the surface in wetlands and ponds. Such landscapes also form nested groundwater flow systems with complex surface and groundwater interactions (Fig. 1; (Smerdon et al., 2009B). The objective of the Peace region numerical model was to understand the connection between local and regional groundwater
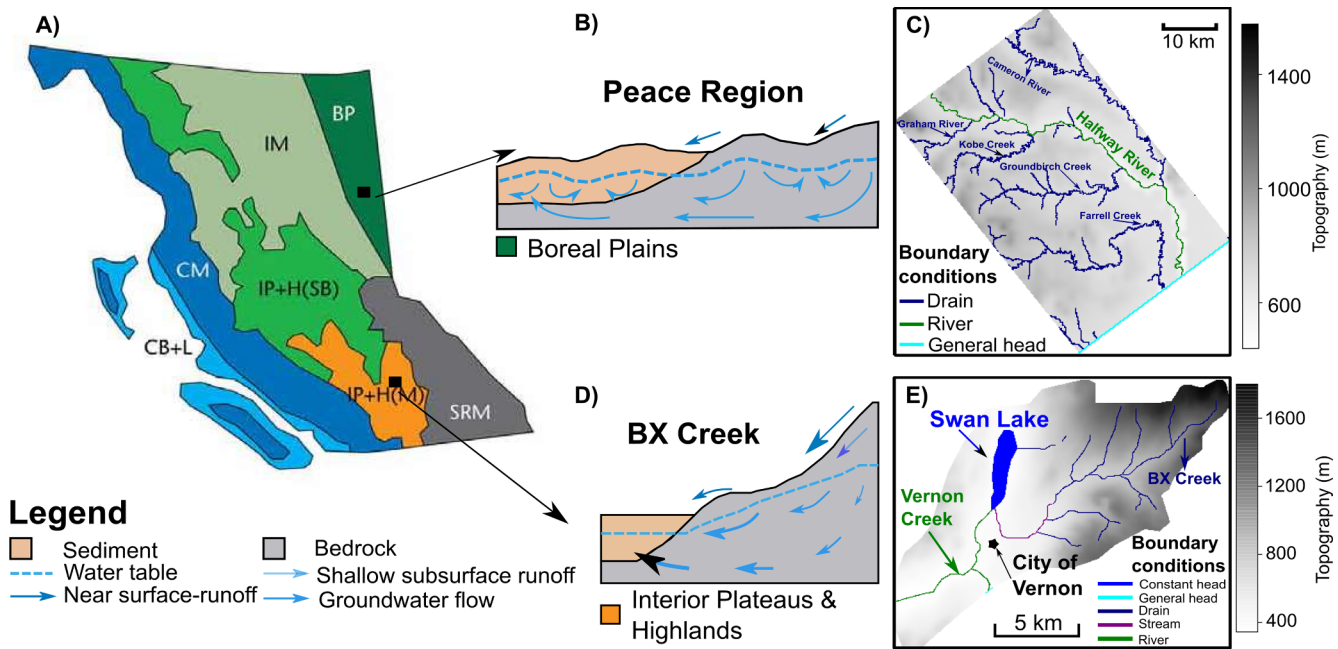
Fig. 1. A) Hydrogeological landscapes in British Columbia, Canada. CB + L: Coastal Basins and Lowlands; CM: Coastal Mountains; IP + H(M): Interior Plateaus & Highlands (Montane); IP + H(SB): Interior Plateaus & Highlands (Sub-Boreal); IM: Interior Mountains; SRM: Southern Rocky Mountains; and BP: Boreal Plains. Conceptual models in BP and IP + H for the Peace region (B) and BX Creek (D), respectively. C) and E) are watershed locations and river names in each domain. Figures A), B), and D) were modified from Smerdon et al., 2009B used Attribution-NonCommercial-NoDerivs 2.5 Canada (CC BY-NC-ND 2.5 CA).

flow in the buried valley aquifers, and simulation results suggested that buried valleys are not regionally connected throughout the whole network (Morgan et al., 2019).

The MODFLOW model in the Peace region has 327 rows and 308 columns with the uniform cell size of 200 m × 200 m and 20 layers following watershed topography. The model area is about 1952 km^2. The numerical model has six hydrostratigraphic units (Table 1; Figs. S6 and S7). To simulate surface water features, three boundary conditions were assigned including drain (DRN), river (RIV), and general head (GHB) in MODFLOW (Fig. 4A). The river boundary represents the Halfway River and the drain boundary represents its tributaries. The general head boundary condition, which is in the south boundary, is to account for the influence of the Peace River, the mainstream of Halfway River (Fig. 1). The details of the numerical model settings can be found in Morgan et al. (2019).

In the Peace region, there were no long-term groundwater levels for calibration. During the period of 1980–2017, only 20 wells had intermittent groundwater level measurements which were in different seasons and different years. As a result, hydraulic head data were not used to calibrate the steady-state model as they do not represent the long-term average hydrogeological conditions in this domain. Instead, the MODFLOW model was calibrated using estimated baseflow (i.e.,

stream-aquifer exchange in MODFLOW) derived from streamflow data. Overall, MODFLOW estimates of baseflow were within ± 3.5% of mean annual discharge for the study domain (Morgan et al., 2019).

2.2. The impacts of pumping on streamflow depletion using numerical models

2.2.1. Converting numerical models from steady-state to transient

The numerical models for BX Creek and the Peace region were calibrated for steady-state conditions. Because our goal was to simulate time-varying streamflow depletion by groundwater pumping, we converted both models to transient conditions using a weekly stress period with the same boundary conditions as in the steady-state MODFLOW model. For recharge, we converted the annual recharge rate to a constant weekly rate. We estimated specific yield and specific storage based on the hydrostratigraphic units used in the steady-state calibrated models, and adjusted these values within a reasonable range (Table 1) to minimize mass balance error in numerical models (Leake, 2011). To obtain appropriate initial conditions for the transient numerical models, each numerical model was run for 40 years without pumping until a dynamic equilibrium was reached, and the output of hydraulic heads was used as initial conditions for subsequent transient pumping test

Table 1
Hydraulic conductivity of hydrostratigraphic units in the BX Creek model and Peace region model.

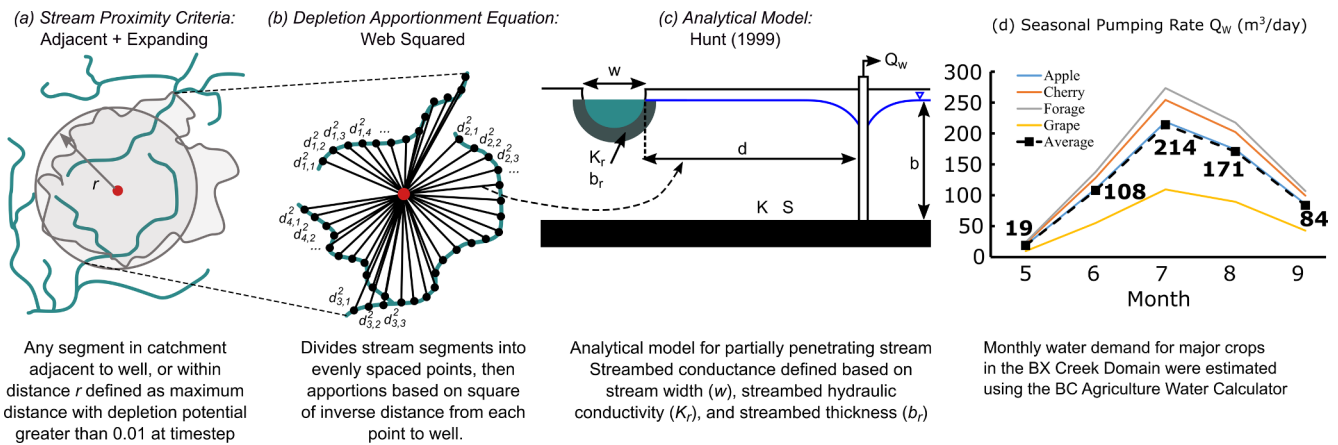| Domains	Materials	Horizontal hydraulic conductivity (m s^{-1})	Vertical hydraulic conductivity (m s^{-1})	Specific Yield (S$_y$)	Specific Storage (S$_s$ m^{-1})
BX Creek Model (Interior Plateaus & Highlands)	Alluvial sediment and aquifer	1×10^{-5}	1×10^{-6}	0.15	1×10^{-5}
	Glaciolacustrine sediments	7×10^{-7}	1×10^{-8}	0.15	1×10^{-5}
	Mixed sediments	1×10^{-6}	1×10^{-7}	0.15	1×10^{-5}
	Bedrock	4×10^{-8}	4×10^{-8}	0.02	1×10^{-4}
Peace region (Boreal Plains)	Coarse sand and gravel	3×10^{-3}	1×10^{-3}	0.15	2×10^{-4}
	Sand	5×10^{-5}	1×10^{-5}	0.15	2×10^{-4}
	Sandstone	1×10^{-5}	1×10^{-6}	0.02	1×10^{-5}
	Till/silt/fine sand	4×10^{-7}	2×10^{-8}	0.02	1×10^{-5}
	Clay/Clay-till	1×10^{-8}	1×10^{-10}	0.02	1×10^{-5}
	Shale	7.7×10^{-10}	7.7×10^{-12}	0.02	1×10^{-5}

Fig. 2. The best performing analytical depletion function developed by Zipper et al. (2019A), including: (a) stream proximity criteria; (b) depletion apportionment equation; (c) analytical model; and (d) Seasonal pumping rate used for synthetic wells. Figure modified from Zipper et al. (2019B) used under Creative Commons BY 3.0 license.

simulations. The transient pumping test simulations were 35 years in length, with the first five years as an additional spin-up period without pumping, followed by a 30-year (1820 weeks) pumping test.

2.2.2. Pumping schedule

To test the impacts of pumping on streamflow, we developed pumping schedules typical of agricultural irrigation in the BX Creek region using the British Columbia Agriculture Water Calculator (http://www.bcagriculturewatercalculator.ca/), which estimates the monthly water demand for various crops based on soil type and climate conditions. We selected four dominant agricultural crops in BX Creek (apple, cherry, forage and grape) and assumed the irrigated area from each groundwater well was equal to the size of the average land parcels within the domain. Then, water demand was calculated for each type of crop for the irrigated area over the growing season. Finally, the water demand was averaged across the dominant crops. Specifically, monthly pumping rates for May, June, July, August, and September are 19, 108, 214, 171, and 84 m^3 day^{-1}, respectively (Fig. 2). We used the same pumping schedule for BX Creek and the Peace region to simulate a consistent stress on the groundwater system for direct comparison of streamflow depletion in the two domains, and the average irrigation demand in BX Creek was higher than that in the Peace region. Previous literature showed that streamflow depletion factor (i.e., streamflow depletion/pumping rate) derived by analytical models and ADFs are not sensitive to pumping rates (Barlow and Leake, 2012; Zipper et al., 2018). Therefore, pumping schedule designed in this study allows us to examine a more realistic groundwater pumping impacts on streamflow.

2.2.3. Streamflow depletion assessment using numerical models

To systematically compare the ADFs and numerical models, we introduced synthetic groundwater wells at regular spacing throughout two domains to stress the aquifer (Fienen et al., 2016). A total of 99 and 96 synthetic pumping wells were created in the BX Creek and Peace region, respectively. The well density (total number of wells over model domain area) are 0.6 and 0.05 wells/km^2 for the BX Creek and Peace region, respectively. Synthetic wells were screened at a depth of 15 m below the water table in the BX Creek to ensure pumping from the shallow aquifers, and 35 m below the water table in the Peace region to ensure wells did not dry up in response to pumping during the simulation period. The water table from the steady-state model simulations was used to define the pumping well depths. Each well has the same annual pumping schedule for the entire 30-year pumping simulations (Fig. 2d). To calculate streamflow depletion caused by each synthetic well, the model was first run with no groundwater pumping (all wells

turned off) as a baseline simulation, and the stream-aquifer flux was calculated for each stream segment for each stress period. Then, a new simulation was conducted for each synthetic well turned on one-at-a-time and the stream-aquifer flux was calculated for each stream segment and stress period in this pumped scenario. Streamflow depletion caused by each individual well in each stream segment for a stress period was calculated as the difference between the pumping and the baseline simulation.

We extracted stream segments from the MODFLOW simulation results and then we calculated the annual streamflow depletion factor, which is the ratio of the annual sum of streamflow depletion that occurred in a specific year over the annual pumping rate to represent temporal changes in streamflow depletion (Barlow and Leake, 2012). In this way, streamflow depletion occurring during the pumping season and time-lagged depletion occurring during non-pumping periods were both included for assessment. In addition, two MODFLOW models reached dynamic equilibrium stressed by constant weekly recharge rates and hence led to a stable surface water and groundwater interactions. As such, the impacts of pumping for different streamflow regimes (e.g., high and low flow periods) cannot be assessed in this study. Nevertheless, Flores et al. (2020) showed that analytical depletion models have a better performance over constant flow period based on field experiments. Therefore, our research design augments the possibility of comparison between two domains with different hydrological landscape and stream networks.

2.3. Streamflow depletion assessment using analytical depletion functions

In this study, we used the highest performing stream proximity criteria and depletion apportionment equation identified by Zipper et al. (2019A); Zipper et al. (2020) and compared two different analytical models (Fig. 2). Stream proximity criteria identify the stream segments which could potentially be impacted by a well. In this study, we used the *Adjacent + Expanding* stream proximity criteria, which include any stream segment that is in a catchment adjacent to the well or is within the maximum radial distance where depletion would be at least 1% of the pumping rate at a given time step. Depletion apportionment equations estimate the fraction of total depletion allocated to each stream segment. The *Web Squared* depletion apportionment equation splits each stream segment into a finite number of points (e.g. space between each point is 5 m) and apportions based on the square of the inverse distance of each stream segment to the well as shown in Eq. (1).

$$f_i = \frac{\sum_{P=1}^{P_i} \frac{1}{d_{i,p}^2}}{\sum_{j=1}^{n} \sum_{P=1}^{P_i} \frac{1}{d_{j,p}^2}}$$

(1)

where f_i is the depletion fraction of total streamflow depletion from a well apportioned to a stream segment, P is the total number of points which a stream segment is divided into the web squared equation, d is the distance from a well to a stream segment, and n is the total number of stream segments meeting the stream proximity criteria.

The final step to estimate segment-resolution streamflow depletion is to use traditional analytical models to calculate the depletion for each stream segment, which is then adjusted based on the fraction of total depletion calculated using Eq. (1). In this study, we compared the Glover (Glover and Balmer, 1954) and Hunt analytical models (Hunt, 1999), both of which are commonly used to calculate the streamflow depletion due to their simplicity of implementation. The Glover method assumes that streams fully penetrate the aquifer, and there is no resistance to flow through the streambed. The volumetric streamflow depletion rate, Q_a of a stream segment can be calculated by Eq. (2).

$$Q_a = Q_w * \mathrm{erfc}\left(\sqrt{\frac{Sd^2}{4Tt}} \right)$$

(2)

Where, $erfc()$ is complementary error function, S is the aquifer storage coefficient (e.g., specific yield in an unconfined aquifer, unitless), T is the aquifer transmissivity (L^2/T, L is for length and T is for time), t is the time since the start of pumping (T), d is the well-stream distance (L), and Q_w is the pumping rate (L^3/T).

The Hunt model assumes the streams partially penetrate the aquifer and there is a streambed clogging layer of a finite thickness (b_r, L) and hydraulic conductivity (K_r, L/T) impeding water exchange between the aquifer and the stream. The Hunt model defines volumetric streamflow depletion as

$$Q_a = Q_w * \left(\mathrm{erfc}\left(\sqrt{\frac{Sd^2}{4Tt}} - \exp\left(\frac{\lambda^2 t}{4ST} + \frac{\lambda d}{2T} \right) \mathrm{erfc}\left(\frac{\lambda^2 t}{4ST} + \frac{\lambda d}{4Tt} \right) \right) \right)$$

(3)

where λ is the streambed conductance. The streambed conductance is defined as where w_r (L) is the width of the stream segments. Unlike the conductance in MODFLOW has a unit of L/T, λ in Hunt model has a unit of L^2/T and interpreted as the conductance per unit length of stream.

The ADFs were implemented in the R package "streamDepletr" (Zipper, 2019). The input parameters of the analytical depletion functions, including, *T, S, d,* and λ, were extracted from MODFLOW so that differences in parameters between the MODFLOW and ADFs are minimized. For the analytical models' assumptions, hydrostratigraphic input parameters (*T, S*) should ideally be averaged between the well locations and affected stream segments. However, since detailed subsurface information is typically unavailable in watershed management settings, we used *T* and *S* parameters based on the values at the well locations where aquifer testing (e.g. pump tests, borehole logs) would typically be available. Specifically, each ADF input parameter was calculated as follows:

1) Transmissivity (*T*). MODFLOW uses hydraulic conductivity as input while analytical models use transmissivity, which is equal to hydraulic conductivity multiplied by the aquifer thickness. To obtain the transmissivity for the ADFs, the average hydraulic conductivity between the surface and well depth in the MODFLOW model was used and multiplied by the thickness from the water table to the well depth.
2) Specific yield (*S*). Average specific yield between the surface and well depth in the MODFLOW at the well locations was used as input for analytical depletion functions.
3) Well-stream distance (*d*). The horizontal distance between a well and affected stream was calculated based on the model domain.

4) For the Hunt model, streambed conductance (λ) is needed. For the Peace region, λ was available for all flow boundary conditions used in the MODFLOW model. However, in the BX Creek model, the constant head boundary, which does not require conductance, was used to simulate a lake. We used the highest conductance values of other boundary conditions for the constant head, while λ of other flow boundary conditions was extracted from the MODFLOW model.

In BX Creek, the MODFLOW simulations revealed that no depletion occurred for the drain and general head boundary conditions, which are consistently located in the upper streams of the domain and represent intermittent headwater streams (Section 3.1). The inclusion of these boundaries in ADFs could lead to a biased streamflow depletion assessment since depletion cannot occur from dry streams. Therefore, these flow features were excluded for the ADFs in the BX Creek domain. In contrast, both gaining and ephemeral stream types are detected within the drain boundary in the Peace region (Fig. 4C) and thus all features were included in the ADFs. The different response of the drain boundary conditions in the two domains is described below in Section 3.1.

2.4. Evaluation metrics for comparing between analytical depletion functions and numerical models

To evaluate the performance of ADFs, we used three metrics proposed by Zipper et al. (2019A):

Metric 1: *Spatial distribution of primary impact* evaluates whether the ADFs can correctly identify the most affected stream segments by a pumping well. It is quantified as the percentage of wells in which ADFs and MODFLOW identify the same stream segment with the greatest predicted depletion.

Metric 2: *Magnitude of primary impact* quantifies how big the difference is between ADFs and MODFLOW in predicting the streamflow depletion. We quantify this as the normalized MAE (mean absolute error), which is the streamflow depletion MAE normalized by the highest pumping rate simulated ($214 \, \mathrm{m^3 \, day^{-1}}$). In addition, we used Kling-Gupta Efficiency (Gupta et al., 2009), or KGE to examine the fit between ADFs and MODFLOW. The KGE is a hydrological fit metric, which integrates correlation, bias, and variability between two datasets. The KGE of 1 is a perfect fit and lower values implying worse performance, with KGE < -0.41 indicating poor performance (Knoben et al, 2019).

Metric 3: *Magnitude of overall impact* quantifies the difference in all affected stream segments by ADFs and MODFLOW. For this metric, we used a) MAE to quantify the difference in streamflow depletion predictions from the ADFs and MODFLOW; b) KGE to quantify the fit between the two predictions; and c) accuracy of predicted streamflow capture fraction, which is the cumulative depletion summed across all stream segments from a given well at a given time step (Barlow et al., 2018; Zipper et al., 2019A). This is quantified as the MAE between the capture fraction estimated by ADFs and those by MODFLOW, normalized by the range in capture fraction among all wells from MODFLOW. The three metrics can comprehensively assess the performance of analytical depletion functions.

2.5. Sensitivity analysis of analytical depletion functions and their performance in different hydrogeological settings

To understand the drivers of ADFs performance and guide the application of ADFs in real-world settings, we conducted a one-at-a-time (OAT) local sensitivity analysis of ADF input parameters (i.e., *T, S,* and λ). To do so, *T* and λ inputs to the ADFs were increased/decreased by 10 and 100 times their original values, and *S* was increased/decreased by 10% and 20%. In practice, the variations of tested parameters are within variability ranges when hydrostratigraphic unit is identified in the study region and thus the outcomes can enhance our understanding
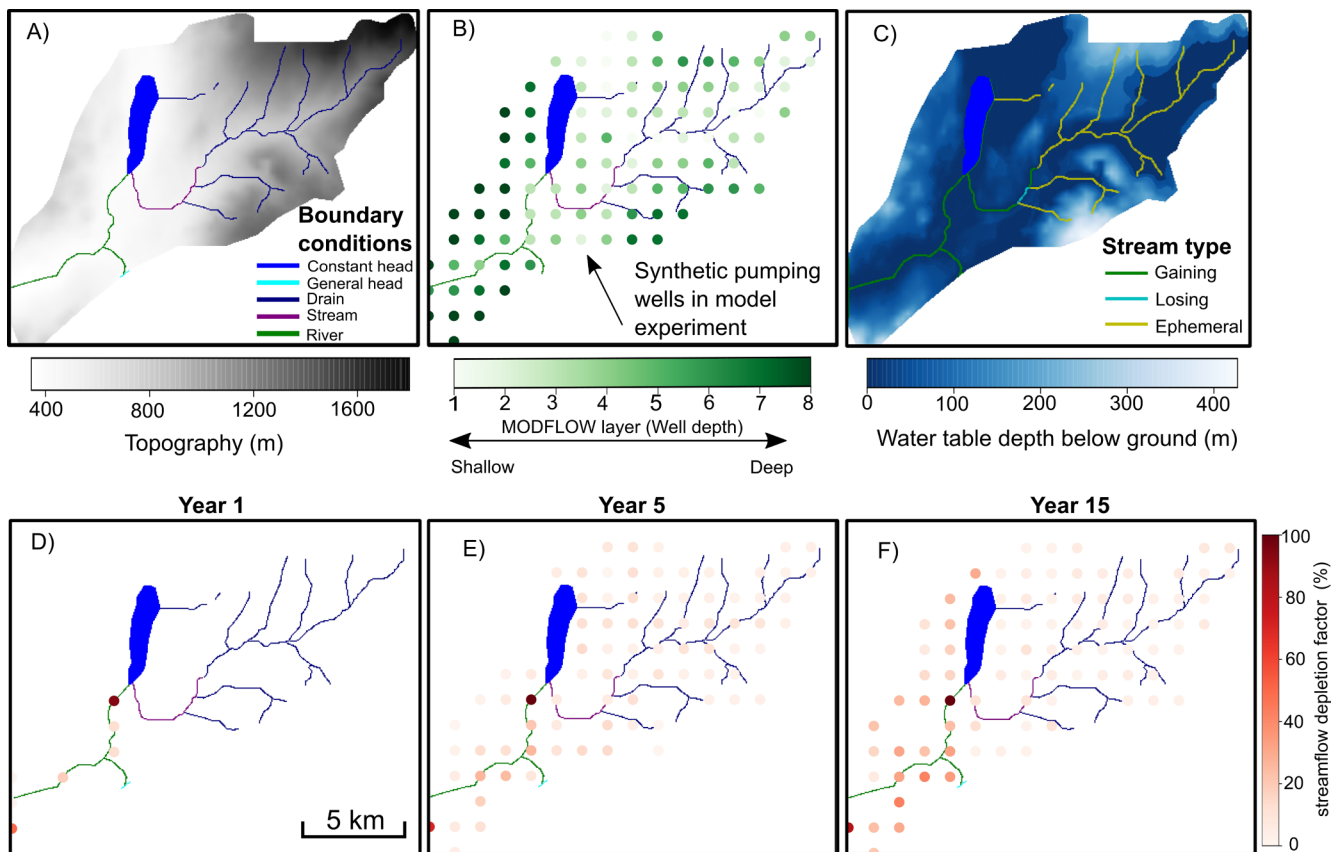
Fig. 3. A) Boundary conditions and topography in the interior plateau and highlands (BX Creek); B) spatial distribution of synthetic pumping wells and their depths in the numerical model; C) streamflow types, including gaining, losing, and ephemeral stream segments. D), E), and F) are cumulative streamflow depletion factors (a ratio of the sum of streamflow depletion that occurred in the year by annual pumping rate) at the 1st, 5th, and 20th years of pumping, respectively. Red colour bar corresponds to the annual streamflow depletion factor in panel D), E), and F).

of the uncertainties in ADFs for the real-world application. The distribution and mean streamflow depletion difference among sensitivity analyses and the baseline were compared. We also took advantage of existing hydrostratigraphic variability in our two domains to examine the impacts of pumping on the performance of the ADFs. We accomplished this by calculating MAE between ADFs and MODFLOW for each well, and investigating variability in MAE in response to hydrological conductivity, streambed conductance, well-stream distance, and well depth in both domains.

3. Results and discussion

3.1. Streamflow depletion assessment using numerical models

In the BX Creek domain, which is relatively small with a simple stream network compared to the Peace region, there are a mixture of gaining, losing, and ephemeral streams (Fig. 3C). Our pumping experiments found that 88% of synthetic groundwater wells caused detectable streamflow depletion over the simulation period. The other 12% of pumping wells tested affected only groundwater storage and thus caused no depletion of surface water features. As shown in Fig. 3D–F, both the total number of pumping wells causing detectable streamflow depletion and the stream depletion factor (annual depletion divided by annual pumping rate) in a given well increased with pumping time getting longer. Specifically, only 7% of wells caused detectable streamflow depletion in the first year, which increased to 65% in the second year and finally stabilized around 85% in the fifth year. For each specific year of pumping, streamflow depletion factor ranged from 0 to 100% across the study domain.

Spatially, the streamflow depletion in the first year was primarily

caused by wells close to surface water features in the lowland portions of the domain and expanded to wells in upland areas in the domain in later years, which were simulated by RIV (Fig. 3A, rivers) and STR (Fig. 3A, streams) flow boundary conditions in MODFLOW. In this domain, we found that there was no depletion from ephemeral streams in the upper domain where the MODFLOW model uses the drain flow boundary condition (Fig. 3A, drains) or DRN in MODFLOW to represent surface water. This is because in MODFLOW, aquifer-drain fluxes drop to zero when hydraulic head in drain cells falls below a threshold level representing the bottom of the stream channel. In our simulations, head was below the drain channel elevations for the entire simulation, which is consistent with regional understanding that these ephemeral streams are disconnected from the groundwater system and primarily transport surface water (e.g., snowmelt and overland flow). Despite the fact that the drain boundary condition represents headwater drainages with no flow during the simulations (Fig. 3C), some wells located in upland areas still caused depletion, which was sourced from nearby, down-gradient features such as the constant head (lake) and rivers.

Similarly, there was no depletion detected from the general head boundary condition (Fig. 3A, general head), partially because it only covers four grid cells in the MODFLOW model. The flux simulated between the general head boundary condition and surrounding model areas is always proportional to the difference in hydraulic heads in MODFLOW (Harbaugh, 2005). In the BX Creek model, the general head boundary condition was designed to ensure that the downstream segments do not dry. Fig. 3C depicting model results over the pumping period and the detailed examination of model output showed that the fluxes exchange between the aquifer and general head boundary condition are small or even nearly zero because the difference of hydraulic head between the general head boundary condition and the

Fig. 4. A) Boundary conditions and topography in the boreal plain (Peace region); B) spatial distribution of synthetic pumping wells and their depths in the MODFLOW; C) streamflow types, including gaining, losing, and ephemeral stream segments. D), E), and F) are annual streamflow depletion factor (a ratio of the sum of streamflow depletion occurred in the year by the annual pumping rate) at the 1st, 5th, and 20th years of pumping, respectively. Red colour bar corresponds to the annual streamflow depletion factor in panel D), E), and F).

surrounding areas are minor, leading to small flux exchanges. Overall, simulations in the BX Creek domain revealed that aquifer-stream interactions differed among boundary conditions and setting in the landscape, indicating that boundary condition choice may affect streamflow response to pumping and therefore should be carefully designed when investigating the impacts of flow responses to pumping.

The Peace region is a large-scale regional domain with a more complex stream network than BX Creek and also contains gaining, losing, and ephemeral streams. In this model, all of synthetic groundwater wells caused detectable streamflow depletion within the 30-year pumping experiments (Figs. 4 and S4). Like the BX Creek model, gaining and losing streams near pumping well responded quickly over to pumping. Simulated streamflow depletion factors ranged from 0 to 100% across wells and the number of wells causing streamflow depletion increased with longer pumping time. For instance, in the first eight years of pumping, < 32% of groundwater wells led to detectable streamflow depletion, while more than 85% of wells caused streamflow depletion from ninth year after onset of pumping onwards. By the final year of pumping, all pumping wells have affected at least one streams in the domain. In contrast to BX Creek, gaining and ephemeral stream segments were found for drain flow boundary condition and we further revealed that pumping had no impacts on ephemeral streams (Fig. 4A and C), indicating that stream types should be considered in streamflow depletion assessment.

Overall, groundwater pumping had larger impacts on streams in the Peace region with the same pumping schedule (larger streamflow depletion factor in Fig. 4F compared to Fig. 3F). This is likely because 1) different hydrogeological settings and stream networks can have differing impacts on streams, which further highlights that streamflow depletion assessment should consider site-specific factors; 2) the MODFLOW setting can play a role in different responses of streamflow depletion. In this study, BX Creek model is a small watershed domain and has 8 vertical layers with a horizontal grid size of 50 m, while the Peace region domain is a regional model has 20 vertical layers with a horizontal grid size of 200 m (Fig. 1). The literature showed that discretization (i.e., grid size of a horizontal cell and depth of vertical layer) can produce differences in stream leakages and thus can influence streamflow depletion prediction (Brunner et al., 2010; Mehl and Hill, 2010). MODFLOW is a finite-difference model and the distribution of hydraulic head and flux exchange does not vary within a grid cell. However, they are not dependent on the size of grid cell in reality. Studies showed that a finer cell could have a detailed representation of hydraulic heads and fluxes exchanges and thus may lead to a better simulation of groundwater flow processes (Morel-Seytoux, 2009). Similarly, when vertical groundwater flow is significant, adopting a coarse vertical layer depth can lead to errors in hydraulic heads and flow exchanges between layers (Brunner et al., 2010). Such differences in discretization between two models could contribute to different
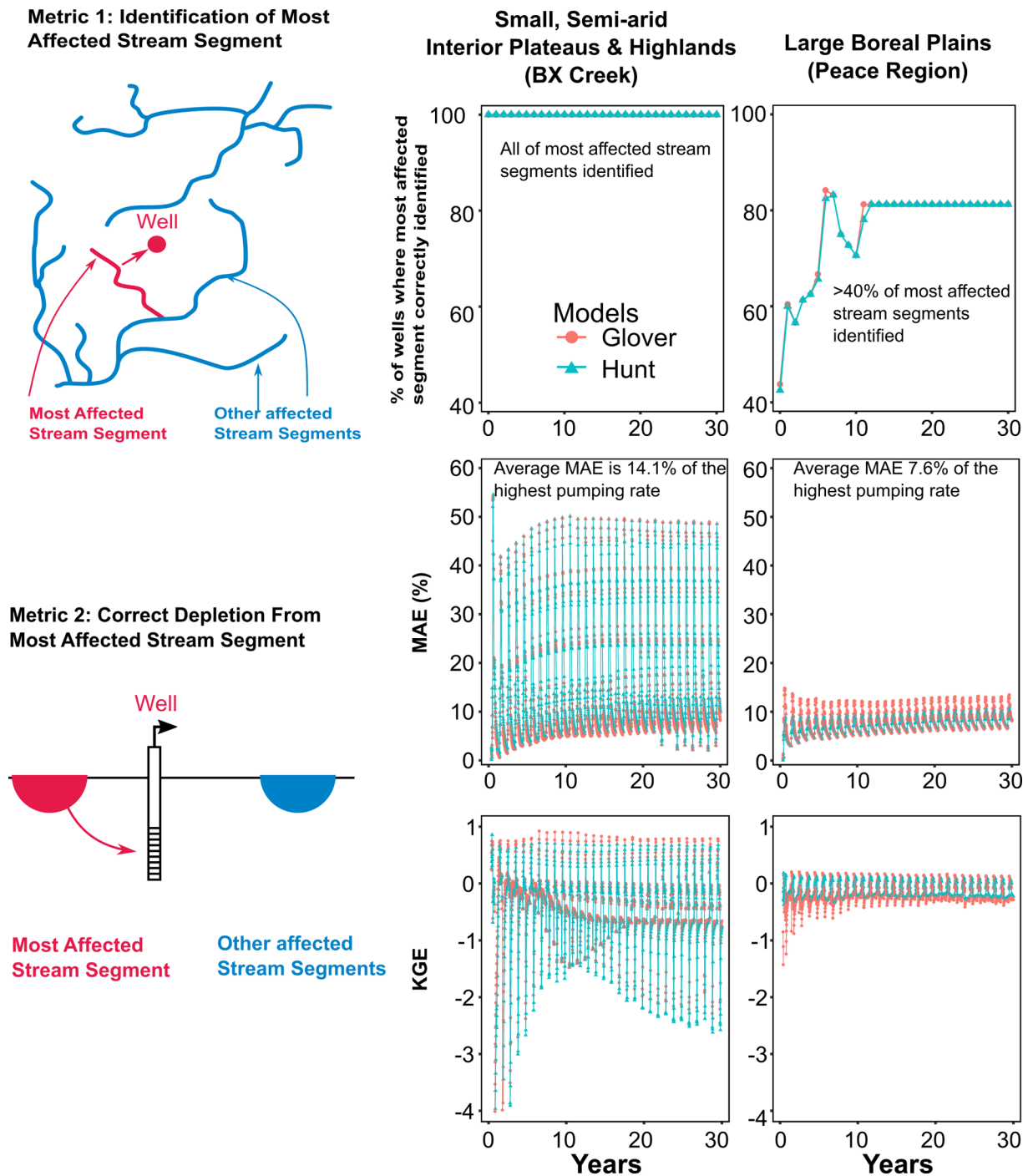
segment ranged

Fig. 5. Performance of the analytical depletion functions in BX Creek and Peace region for performance metric 1 (identification of most-affected stream segment) and performance metric 2 (correct depletion from most-affected segment, quantified using MAE and KGE).

Hunt model had similar responses to changes in T, S, and λ in ADFs in the two domains (Figs. 7, S10, and Table S1), which further demonstrate that the choice of analytical models does not significantly affect streamflow depletion estimates in these study domains. We, therefore, focus our discussion of the sensitivity analysis results on the Hunt model, which includes streambed conductance. We found that changes in T and λ lead to large difference in streamflow depletion as the density distributions are more widespread for T and λ than that of S (Fig. 7). Additionally, the average difference relative to the original values of streamflow depletion (Table S1) was larger for T than of λ or S, indicating that T has the largest effect on results while S is the least sensitive parameters to ADFs in two domains under sensitivity tests.

However, it should be noted that we only changed the S by 20% of its baseline value while T and λ varied by two orders of magnitude as we were intended to keep parameter values realistic with real-world hydrostratigraphic materials in these regions. In general, streamflow depletion will be underestimated when T and λ are underestimated, or overestimated when S is overestimated (and vice versa). As a result, we suggest that obtaining accurate transmissivity estimates should be given the first priority when applying the ADFs.

To evaluate the impact of hydrostratigraphic conditions on ADF performance, we explored the response of ADFs to four model parameters: hydraulic conductivity, streambed conductance, well-stream distance, and pumping depth in both domains. In BX Creek, we found
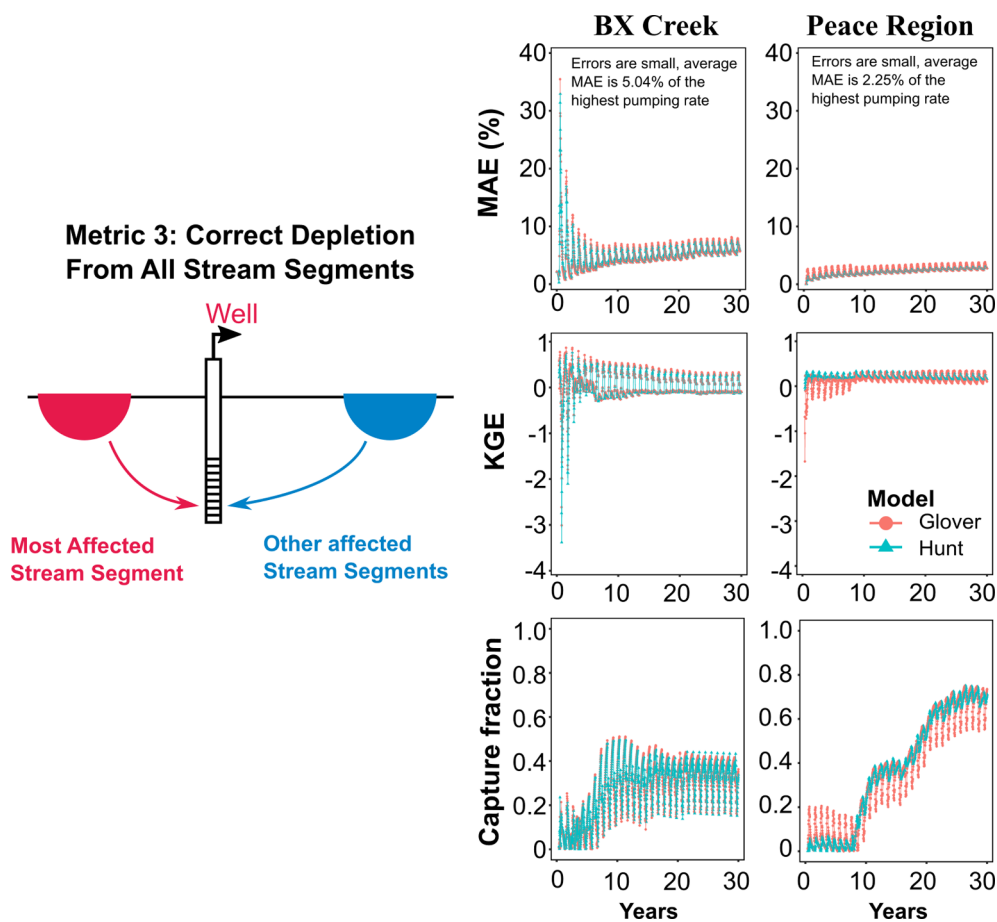
Fig. 6. Performance of the analytical depletion functions in BX Creek and Peace region for performance metric 3 (correct depletion from all affected stream segments), expressed as mean absolute error normalized by highest pumping rate, KGE, accuracy of predicted streamflow capture fraction.

that ADFs performed better in materials with higher hydraulic conductivity (compared to bedrock materials), lower streambed conductance, and wells within the top five model layers (Fig. 8). Interestingly, we found opposite responses of ADFs to several characteristics in the Peace region (Fig. 9). Specifically, ADFs performed better in lower conductivity sediments and for wells in deeper layers to affect streams with lower streambed conductance. One possible reason is that the parameters (T and S) used in the ADFs were obtained from well locations. In theory, they should be the average values between the wells and affected streams. In the BX Creek model, due to the simple hydrostratigraphic units and smaller domain, the selected parameters were more representative of the average values between well and streams. In contrast, the Peace region has a large spatial variation in hydraulic conductivity and longer well-stream distances (Figs. S6 and S7). The parameters at the well location, therefore, are less likely to represent the average condition between well and streams, and this mismatch between the hydrostratigraphy at the well and the hydrostratigraphy between the well and stream may have caused a different response between the two domains. Moreover, the ADFs had higher MAE in higher hydraulic conductivity zones with higher streambed conductance than smaller ones, potentially leading to larger estimates of depletion and accordingly larger MAE. The hydraulic conductivity range in the BX Creek domain is three orders of magnitude compared to seven orders of magnitude in the Peace region, and apparently there is a much wider range of variability in Peace region which may contribute to different responses in two domains.

In addition, one of the key input parameters in ADFs is the well-stream distance. In both domains, there was consistently greater differences between the ADFs and MODFLOW for wells within ∼2 km of a

stream, which correspond to wells with the highest predicted depletion (Figs. 8 and 9). As well-stream distance increased, the predicted depletion decreases and thus leads to a smaller MAE. The analytical depletion functions' sensitivity in the BX Creek is similar to that observed by Zipper et al. (2019A), who found that ADFs performed better in places which are relatively flat (where alluvial aquifers are most likely to exist), with a near-surface water table (shallower well depth), and within a few kilometers of the down-gradient perennial streams. While we found similar results in BX Creek, our comparison in the Peace region showed that the drivers of ADF performance variability differed across the two domains and that streamflow depletion response to hydrogeological characteristics is most likely to be region-specific. As such, our results indicate that conclusions related to specific parameters observed in one domain cannot be assumed to apply in other regions.

Across the two domains, we found that ADFs including the Glover and Hunt models produced similar streamflow depletion estimates (Figs. S4-S5 and S8-S9) and responded similarly to hydrogeological characteristics (Figs. 8 and 9). The MAE for the most affected stream for ADFs including the Glover and Hunt models were 13.6% and 14.4%, respectively, in the BX Creek; and 5.0% and 5.1%, respectively, in the Peace region. Comparison across the two domains revealed that the Hunt model had a consistently better match with MODFLOW, suggesting that considering the streambed conductance can lead to smaller errors, though the differences were slight. Previous work also found that streambed conductance can influence the performance of analytical models (Lackey et al., 2015; Sophocleous et al., 1988, 1995; Spalding and Khaleel, 1991). For instance, Hunt et al., (2001) conducted a field experiment in a small domain in New Zealand and found that streambed conductance determined the accuracy of streamflow
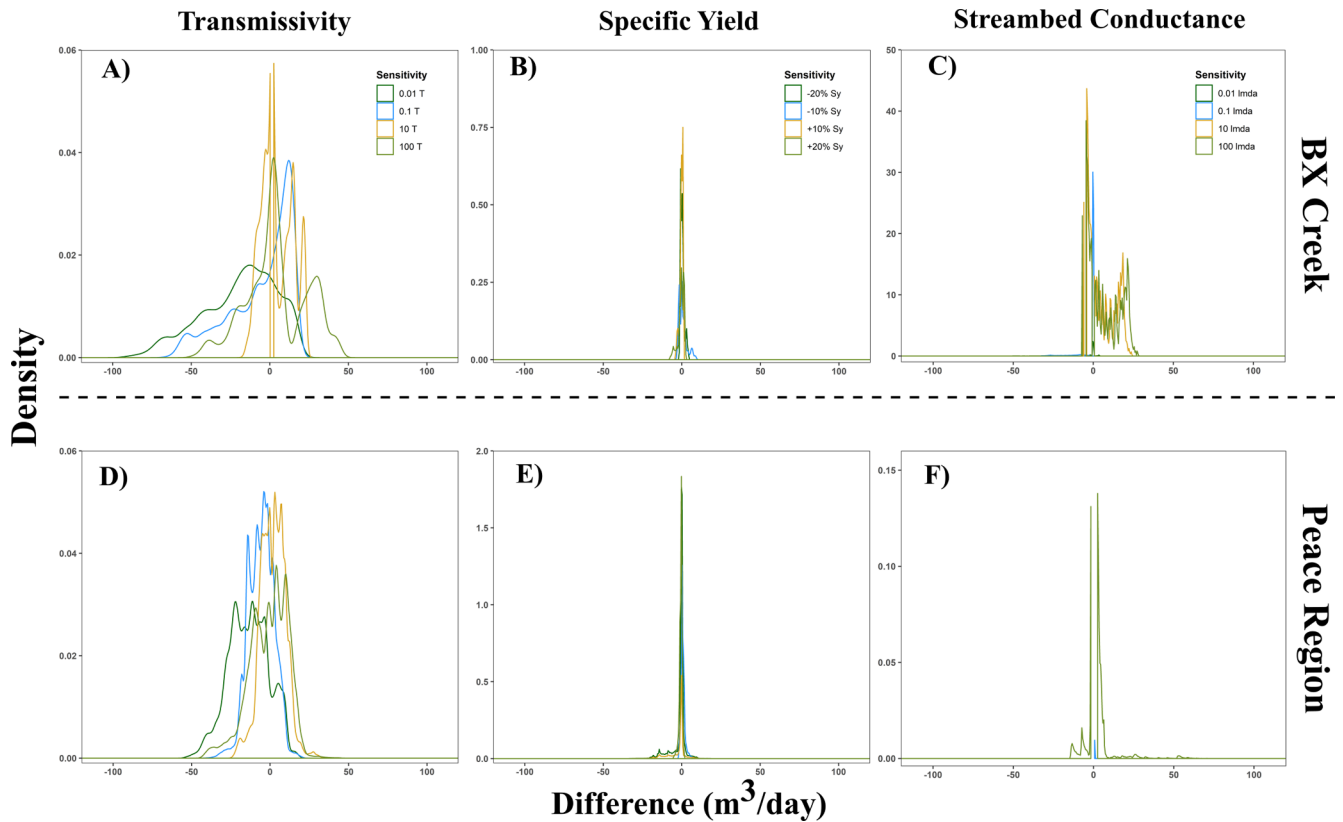
Fig. 7. Sensitivity analysis of analytical depletion function in BX Creek (A, B, and C) and Peace region (D, E and F). The horizontal axis is the streamflow depletion difference between the tested parameters and original parameter.

depletion using the Hunt model. In contrast, streambed conductance may play a minor role in streamflow depletion (Fox et al., 2011). Theoretically, as the streambed conductance decreases, greater differences between ADFs including the Glover and Hunt models would be expected. In the two domains, we found that the difference between ADFs including the Glover and Hunt models was minor over a range of streambed conductance conditions, suggesting that in these two model domains, the streambed conductance may not be a significant factor leading to the differences between Glover model and Hunt model. Our results are consistent with Fox et al. (2011), which conducted a direct measurement of streambed conductance along the North Canadian River in Oklahoma, the USA and concluded that Glover model and Hunt model derived similar estimates of streamflow depletion, indicating that streambed conductance did not add values to streamflow depletion estimation.

3.4. Applicability of analytical depletion function: Uncertainty limitation and future needs for analytical functions and numerical models

The results of this study confirm that ADFs are an accurate tool for estimating streamflow depletion (Huggins et al., 2018; Zipper et al., 2018, 2020; Zipper et al., 2019A). However, the results also show that streamflow depletion responds differently to hydrogeological and physiographic characteristics of the Peace region and BX Creek (Figs. 8 and 9). Due to the different performance of ADFs in these two different hydrogeological landscapes, our results are not sufficient to make generalized conclusions regarding the applicability of ADFs across British Columbia or within specific hydrogeological settings. Therefore, we recommend that ADFs should be used with caution for broader application, that good understanding of local hydrogeological

conditions outside of the tested domains are needed, and that additional testing of ADF performance prior to use in decision-making should be a priority.

While our analysis primarily focused on the evaluation of ADFs, we also found substantial uncertainty associated with using previously calibrated numerical models for streamflow depletion assessment if those models were developed for a different purpose. In the BX Creek and Peace region models some of the flow boundary conditions were challenging to compare to the ADFs. Notably, the drain boundary represented different stream types across our two domains, i.e., ephemeral streams in BX Creek and losing and ephemeral streams in the Peace region. Groundwater pumping has limited impacts on ephemeral streams in these numerical models since they are typically disconnected from the water table. In addition, it should be noted that flow processes in unsaturated zone are not simulated by MODFLOW, which result in the river leakage being directly added to aquifers and hence leading to a biased calculation of streamflow depletion, particularly for disconnected systems between streams and the aquifer. The potential uncertainty from such MODFLOW assumptions should be addressed. Therefore, the selection of boundary conditions could potentially affect the stream depletion results in numerical models and thus influence our comparison between the ADFs and the numerical model. In summary, in order to provide reliable streamflow depletion assessments, numerical models should be specifically selected, developed, and calibrated for surface and groundwater interactions.

Based on these results, more case studies are needed to advance understanding of streamflow depletion and guide application of ADFs for real-world settings with complex stream networks. Here, we highlight some possible directions for future research studies.
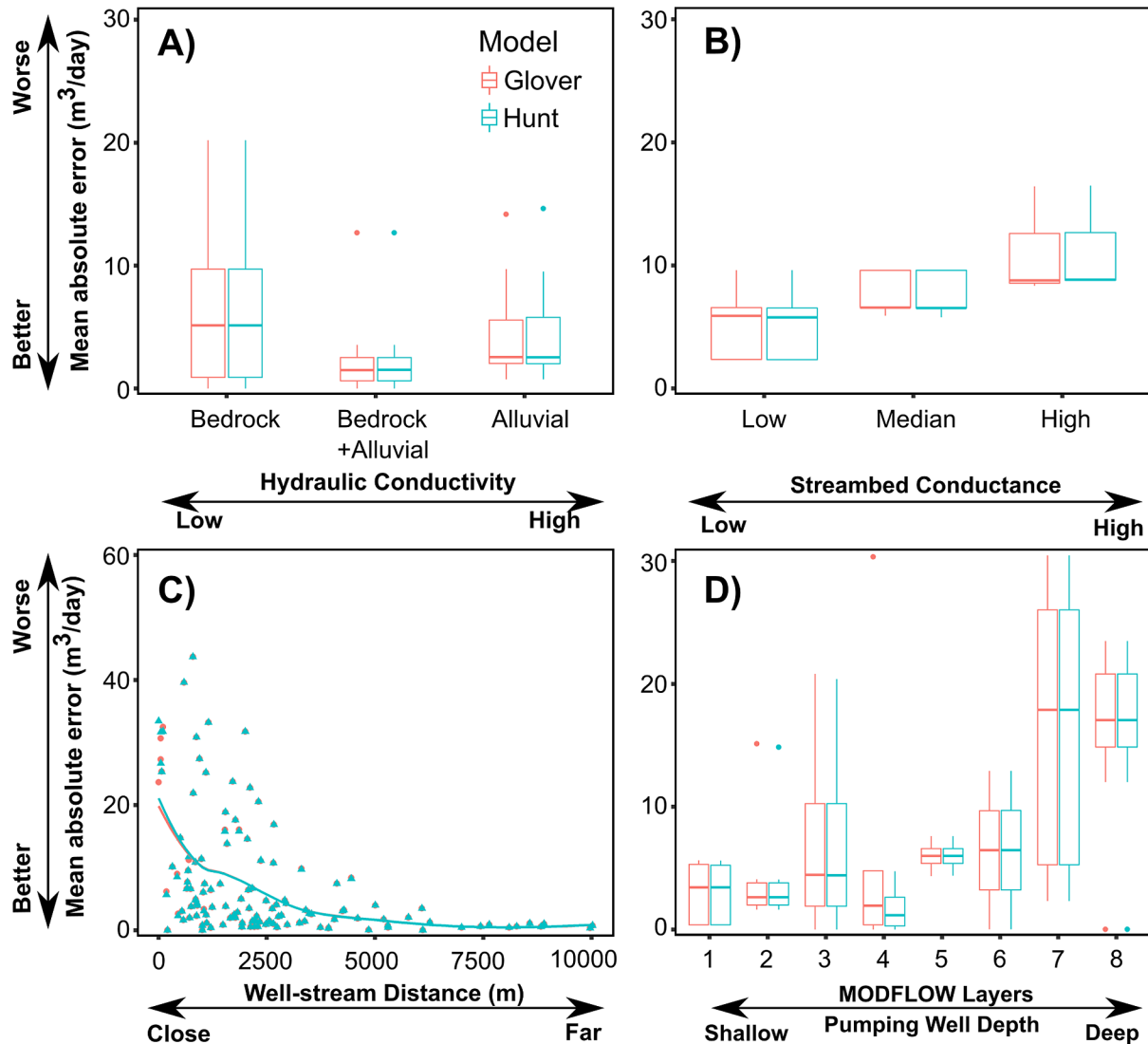
Fig. 8. Analytical depletion function sensitivity to different hydrostratigraphic and physiographic characteristics for the BX Creek model. Sensitivity expressed as MAE of streamflow depletion estimated by the analytical depletion functions compared to MODFLOW and compared to hydraulic conductivity (A), streambed conductance (B), well-stream distance (C) and pumping well depth (D) in the small-arid interior plateau and highlands.

1) **Cumulative impacts of multiple pumping wells:** We compared the performance of ADFs to numerical models by turning one pumping well on at a time. However, multiple groundwater wells typically exist in an aquifer. Future studies are needed to examine whether ADFs are an appropriate tool for estimating stream flow depletion when considering the cumulative effects of multiple wells. Existing literature indicates that the total impacts from multiple groundwater wells may not be equal to the sum of the effects of individual wells (Ahlfeld et al., 2016; Schneider et al., 2017).

2) **Clarifying appropriate numerical models for comparison**: We assessed the performance of ADF by comparing to the previously calibrated numerical models, so the degree to which ADF results can be considered representative of real-world conditions depends on the representativeness of the numerical models. The selection of flow boundary conditions affects numerical model results and thus increases uncertainty in the degree to which our ADF predictions would match real-world conditions. The MODFLOW used in this study did not include flow process in unsaturated zone. Fully integrated models such as HydroGeoSphere (Brunner and Simmons, 2012) can improve the simulation of the partially saturated zone and groundwater and surface water interactions, but in practice are rarely available in calibrated models. Indeed, they are not available in the regions where we conducted our study. Future studies are recommended to adopt unsaturated zone flow (UZF) package in MODFLOW (Niswonger et al., 2006) or to use other models [e.g., HydroGeoSphere and ParFlow (Maxwell and Condon, 2016)] that have capability to simulate both unsaturated and saturated flow processes to minimize uncertainties.

4. Synthesis and conclusions

In this study, we evaluated the performance of analytical depletion functions, which include depletion apportionment equations, stream proximity criteria, and analytical models, to understand the utility of
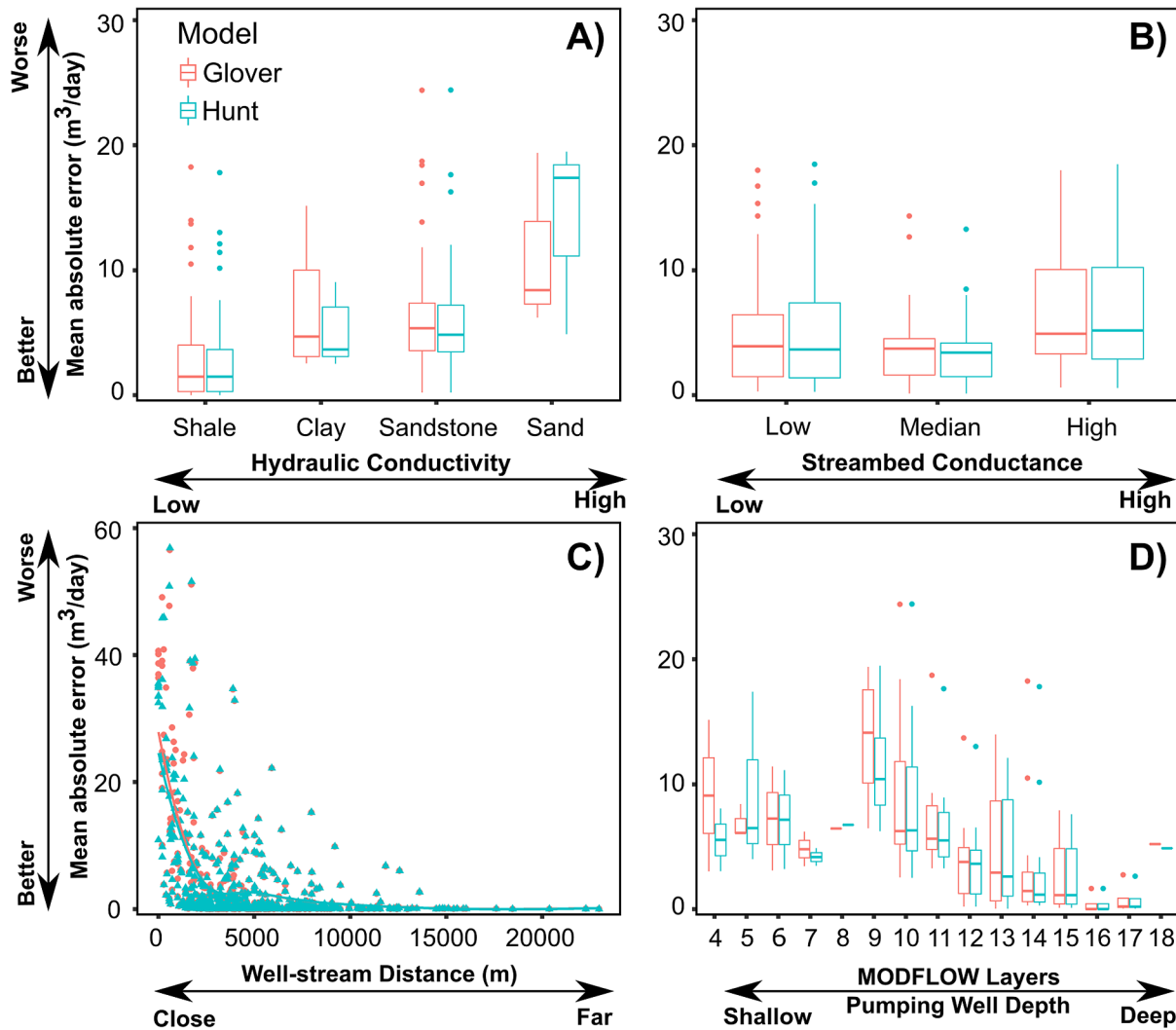
Fig. 9. Analytical depletion function sensitivity to different hydrostratigraphic and physiographic characteristics for the Peace region model. Sensitivity expressed as MAE of streamflow depletion estimated by the analytical depletion functions compared to MODFLOW and compared to hydraulic conductivity (A), streambed conductance (B), well-stream distance (C), and pumping well depth (D) in the boreal plains.

ADFs in two different hydrogeological settings. Specifically, BX Creek has a simpler hydrogeological setting and stream network, while the Peace region is larger and more complex. Using MODFLOW simulations, we found that groundwater pumping can have significant impacts on streamflow in both regions. However, the two domains have different responses to pumping wells due to differences in hydrogeological settings and stream networks. Further, we found that streamflow depletion varies as a result of numerical model structure, with limited impacts on dry or ephemeral streams represented using the drain package in MODFLOW.

Streamflow depletion derived from the ADFs was compared to the numerical models. Across the two domains, we found that ADFs correctly identified all wells for most affected stream segments in the BX Creek and as much as 81% of the time in the Peace region over the entire 30-year simulation. The mean absolute error (MAE) of the most affected stream segment were relatively small compared to the pumping rate. Specifically, the average MAE of the most affected stream segments was 14.4% and 7.6% of the highest pumping rate (214 m³/ day) in the BX Creek and Peace region, respectively. For all affected

stream segments, the average MAE was 5.0% of the highest pumping rate in the BX Creek and 2.3% in the Peace region. In addition, we also found that ADFs predictions were more accurate during the pumping season compared to the non-pumping season. Overall, ADFs provide reasonable predictions to estimate streamflow depletion in the pumping season for perennial streams.

We found variable factors of ADFs performance across these two the hydrogeologic settings. In BX Creek, ADFs have smaller errors for wells in higher hydraulic conductivity materials, shallower aquifer, and lower streambed conductance. Conversely, in the Peace region, ADFs have smaller errors in lower hydraulic conductivity materials, deeper aquifers, and lower streambed conductance. In both regions, the performance of ADFs is most variable closest to streams, with increasing MAE in the first couple of kilometers, corresponding to areas where predicted depletion is the highest. The contrasting responses of ADFs performance to hydrogeological setting between the BX Creek and Peace region stresses the importance of additional testing of ADFs in different regions to better identify the drivers of performance.

In summary, we conclude that ADFs are useful and accurate tools for streamflow depletion prediction for conjunctive water management. We, therefore, recommend this tool can be applied for assessing the impacts of groundwater pumping on streamflow and its associated aquatic functioning. However, a good understanding of local hydro-geological conditions is required to address the potential uncertainty of ADFs and ensure the prediction accuracy.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2020.125568.

References

Aeschbach-Hertig, W., Gleeson, T., 2012. Regional strategies for the accelerating global problem of groundwater depletion. Nat. Geosci. https://doi.org/10.1038/ngeo1617.

Ahlfeld, D.P., Schneider, J.C., Spalding, C.P., 2016. Effets de la réponse non-linéaire d'un modèle sur l'attribution des responsabilités de l'affaiblissement du débit d'un cours d'eau: illustration par le cas de Beaver Creek aux États-Unis d'Amérique. Hydrogeol. J. 24 (7), 1835–1845. https://doi.org/10.1007/s10040-016-1438-3.

Bakker, M., Post, V., Langevin, C.D., Hughes, J.D., White, J.T., Starn, J.J., Fienen, M.N., 2016. Scripting MODFLOW model development using python and FloPy. Groundwater 54 (5), 733–739. https://doi.org/10.1111/gwat.12413.

Barlow, P.M., Leake, S.A., 2012. Streamflow Depletion by Wells-Understanding and Managing the Effects of Groundwater Pumping on Streamflow. U.S. Geological Survey Circular, pp. 1376.

Barlow, P.M., Leake, S.A., Fienen, M.N., 2018. Capture versus capture zones: Clarifying terminology related to sources of water to wells. Groundwater 56 (5), 694–704. https://doi.org/10.1111/gwat.12661.

Booth, J.M., 2007. The book of inkscape. Digital Art. https://doi.org/10.1017/S0003598X00017725.

Brunner, P., Simmons, C.T., Cook, P.G., Therrien, R., 2010. Modeling surface water-groundwater interaction with MODFLOW: Some considerations. Groundwater 48 (2), 174–180. https://doi.org/10.1111/j.1745-6584.2009.00644.x.

Brunner, P., Simmons, C.T., 2012. HydroGeoSphere: A fully integrated, physically based hydrological model. Ground Water 50 (2), 170–176. https://doi.org/10.1111/j.1745-6584.2011.00882.x.

Chen, X.H., 2000. Streamflow depletion due to groundwater irrigation: Analysis of reduced baseflow and induced stream infiltration. Proceedings of the International Symposium on Hydrogeology and the Environment.

Fox, G.A., Heeren, D.M., Kizer, M.A., 2011. Evaluation of a stream-aquifer analysis test for deriving reach-scale streambed conductance. Trans. ASABE 54 (2), 473–479. https://doi.org/10.13031/2013.36450.

Gleeson, T., & Richter, B. (2018). How much groundwater can we pump and protect environmental flows through time? Presumptive standards for conjunctive management of aquifers and rivers. River Research and Applications. https://doi.org/10.1002/rra.3185.

Glover, R.E., Balmer, G.G., 1954. River depletion resulting from pumping a well near a river. Eos Trans. Am. Geophys. Union. https://doi.org/10.1029/TR035i003p00468.

de Graaf, I.E.M., Gleeson (Rens), T., van Beek, L.P.H., Sutanudjaja, E.H., Bierkens, M.F.P., 2019. Environmental flow limits to global groundwater pumping. Nature. https://doi.org/10.1038/s41586-019-1594-4.

Fienen, M.N., Nolan, B.T., Feinstein, D.T., 2016. Evaluating the sources of water to wells: Three techniques for metamodeling of a groundwater flow model. Environ. Modell. Software 77, 95–107. https://doi.org/10.1111/gwat.12389.

Flores, L., Bailey, R.T., Kraeger-Rovey, C., 2020. Analyzing the effects of groundwater pumping on an urban stream-aquifer system. JAWRA J. Am. Water Resourc. Assoc. 56 (2), 310–322. https://doi.org/10.1111/1752-1688.12827.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Harbaugh, A.W. (2005). MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process (pp. 6–A16). Reston, VA: US Department of the Interior, US Geological Survey.

Huang, C.S., Yang, T., Yeh, H. Der, 2018. Review of analytical models to stream depletion induced by pumping: Guide to model selection. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2018.04.015.

Huggins, X., Gleeson, T., Eckstrand, H., Kerr, B., 2018. Streamflow depletion modeling: Methods for an adaptable and conjunctive water management decision support tool. J. Am. Water Resour. Assoc. https://doi.org/10.1111/1752-1688.12659.

Hunt, B., 1999. Unsteady stream depletion from ground water pumping. Ground Water. https://doi.org/10.1111/j.1745-6584.1999.tb00962.x.

Hunt, B., Weir, J., Clausen, B., 2001. A stream depletion field experiment. Ground Water 39 (2), 283–289. https://doi.org/10.1111/j.1745-6584.2001.tb02310.x.

Kiffney, P.M., Bull, J.P., Feller, M.C., 2002. Climatic and hydrologic variability in a coastal watershed of southwestern British Columbia. J. Am. Water Resour. Assoc. https://doi.org/10.1111/j.1752-1688.2002.tb04357.x.

Knoben, W.J.M., Freer, J.E., Woods, R.A., 2019. (2019) Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. Hydrol. Earth Syst. Sci. 23, 4323–4331. https://doi.org/10.5194/hess-23-4323-2019.

Kurylyk, B. L., Macquarrie, K. T. B., Linnansaari, T., Cunjak, R. A., & Curry, R. A. (2015). Preserving, augmenting, and creating cold-water thermal refugia in rivers: Concepts derived from research on the Miramichi River, New Brunswick (Canada). Ecohydrology. https://doi.org/10.1002/eco.1566.

Lackey, G., Neupauer, R.M., Pitlick, J., 2015. Effects of streambed conductance on stream depletion. Water (Switzerland). https://doi.org/10.3390/w7010271.

Leake, S.A., 2011. Capture-Rates and directions of groundwater flow don't matter!. Groundwater 49 (4), 456–458. https://doi.org/10.1111/j.1745-6584.2010.00797.x.

Maxwell, R.M., Condon, L.E., 2016. Connections between groundwater flow and transpiration partitioning. Science. https://doi.org/10.1126/science.aaf7891.

Mehl, S., Hill, M.C., 2010. Grid-size dependence of Cauchy boundary conditions used to simulate stream–aquifer interactions. Adv. Water Resour. 33 (4), 430–442. https://doi.org/10.1016/j.advwatres.2010.01.008.

Morel-Seytoux, H.J., 2009. The turning factor in the estimation of stream-aquifer seepage. Groundwater 47 (2), 205–212. https://doi.org/10.1111/j.1745-6584.2008.00512.x.

Morgan, S.E., Allen, D.M., Kirste, D., Salas, C.J., 2019. Investigating the hydraulic role of a large buried valley network on regional groundwater flow. Hydrogeol. J. 27 (7), 2377–2397. https://doi.org/10.1007/s10040-019-01995-0.

Niswonger, R. G., Prudic, D. E., and Regan, R. S. (2006). Documentation of the unsaturated-zone flow (UZF1) package for modeling unsaturated flow between the land surface and the water table with MODFLOW-2005 (No. 6-A19).

Power, G., Brown, R.S., Imhof, J.G., 1999. Groundwater and fish – Insights from northern North America. Hydrol. Process. 13 (3), 401–422. https://doi.org/10.1002/(Sici)1099-1085(19990228)13:3<401:Aid-Hyp746>3.3.Co;2-1.

Rathfelder, K. M., (2016). Modelling tools for estimating effects of groundwater pumping on surface waters. Province of British Columbia, Water Science Series. WSS2016-09. http://a100.gov.bc.ca/appsdata/acat/documents/r51878/tools4streamdepletion_1484093475019_4092907088.pdf (access on Feb. 17, 2020).

Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature. https://doi.org/10.1038/nature08238.

Schneider, J.C., Ahlfeld, D.P., Spalding, C.P., 2017. Allocation of streamflow depletion impacts under nonlinear conditions. J. Am. Water Resour. Assoc. https://doi.org/10.1111/1752-1688.12525.

Smerdon, B.D., Allen, D.M., Grasby, S.E., Berg, M.A., 2009A. An approach for predicting groundwater recharge in mountainous watersheds. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2008.11.023.

Smerdon, Brian, D., Redding, T., Beckers, J., 2009B. An overview of the effects of forest management on groundwater hydrology. J. Ecosyst. Manage. 10 (1).

Sophocleous, M., Townsend, M.A., Vogler, L.D., McClain, T.J., Marks, E.T., Coble, G.R., 1988. Experimental studies in stream-aquifer interaction along the Arkansas River in central Kansas—Field testing and analysis. J. Hydrol. 98 (3–4), 249–273. https://doi.org/10.1016/0022-1694(88)90017-0.

Sophocleous, M., Koussis, A., Martin, J.L., Perkins, S.P., 1995. Evaluation of simplified stream-aquifer depletion models for water rights administration. Groundwater.

https://doi.org/10.1111/j.1745-6584.1995.tb00313.x.

Spalding, C.P., Khaleel, R., 1991. An evaluation of analytical solutions to estimate drawdowns and stream depletions by wells. Water Resour. Res. https://doi.org/10.1029/91WR00001.

Team, R. C, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Wada, Y., Wisser, D., Bierkens, M.F.P., 2014. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. Earth Syst. Dyn. https://doi.org/10.5194/esd-5-15-2014.

Wei, X.H., Li, Q., Zhang, M.F., Giles-Hansen, K., Liu, W.F., Fan, H.B., et al., 2018. Vegetation cover-another dominant factor in determining global water resources in forested regions. Glob. Change Biol. 24 (2), 786–795. https://doi.org/10.1111/gcb.13983.

Zipper, S.C., Dallemagne, T., Gleeson, T., Boerman, T.C., Hartmann, A., 2018. Groundwater pumping impacts on real stream networks: testing the performance of simple management tools. Water Resour. Res. https://doi.org/10.1029/2018WR022707.

Zipper, S.C., Gleeson, T., Kerr, B., Howard, J.K., Rohde, M.M., Carah, J., Zimmerman, J., 2019A. Rapid and accurate estimates of streamflow depletion caused by groundwater pumping using analytical depletion functions. Water Resour. Res. https://doi.org/10.1029/2018WR024403.

Zipper, S.C., Carah, J.K., Dillis, C., Gleeson, T., Kerr, B., Rohde, M.M., Howard, J.K., Zimmerman, J.K., 2019B. Cannabis and residential groundwater pumping impacts on streamflow and ecosystems in Northern California. Environ. Res. Commun. 1 (12). https://doi.org/10.1088/2515-7620/ab534d.

Zipper, S.C. (2019) streamDepletr: Estimate Streamflow Depletion Due to Groundwater Pumping Online: (https://CRAN.R-project.org/package=streamDepletr).

Zipper, S.C., Gleeson, T., Li, Q., Kerr, B., 2020. Streamflow depletion estimation for conjunctive water management in a heavily-stressed aquifer using analytical depletion functions. EarthArXiv. https://doi.org/10.31223/osf.io/hqrmu.