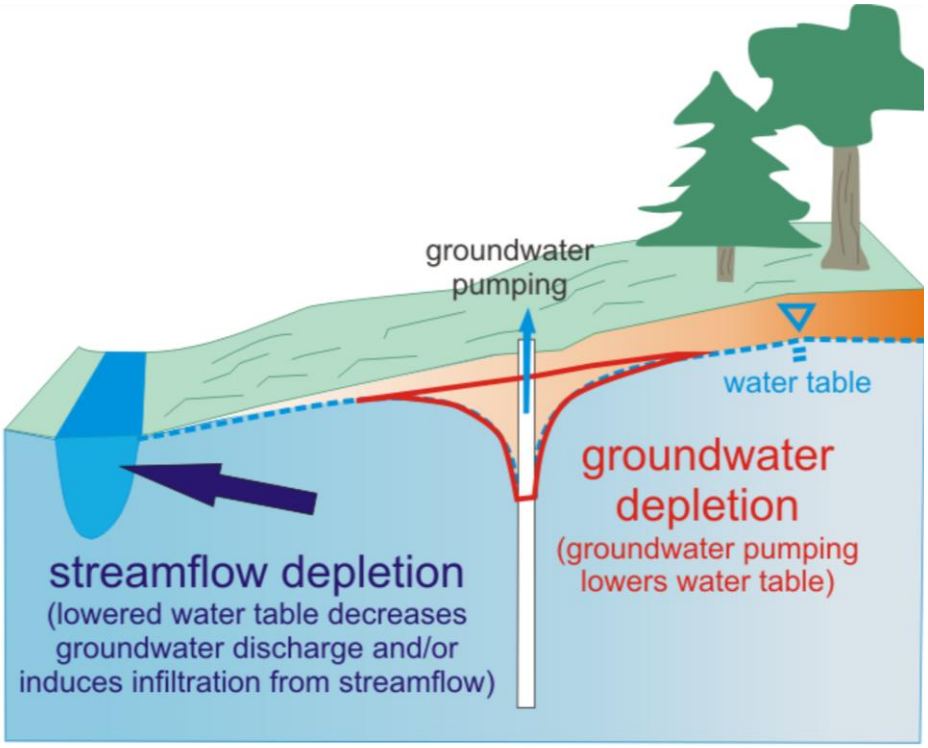
Analytical Depletion Functions and Response Time of Groundwater Pumping Impacts on Environmental Flows

Qiang Li, Samuel C. Zipper, Tom Gleeson



June 2020

Department of Civil Engineering,
University of Victoria, B.C.

University of Victoria

Author's Contact Information:

Qiang Li, Ph.D., (Email: liqiang1205@gmail.com)
Department of Civil Engineering, University of Victoria, 3800 Finnerty Road, Victoria, V8P 4C2

Tom Gleeson, Ph.D., (Email: tgleeson@uvic.ca)
Department of Civil Engineering, University of Victoria, 3800 Finnerty Road, Victoria, V8P 4C2

Samuel C. Zipper, Ph.D., (Email: samzipper@ku.edu)
Kansas Geological Survey, University of Kansas, Lawrence KS, United States of America

Cover Photographs:
Dr. Tom Gleeson from Gleeson and Richter, (2018).

Citation:
Li, Q., Zipper, S.C., and Gleeson, T. 2020. Analytical depletion functions and response times of groundwater pumping impacts on environmental flow. A report prepared for the B.C. Ministry of Environment and Climate Change Strategy. Department of Civil Engineering, University of Victoria, Victoria, B.C.

Acknowledgements

The authors would like to acknowledge the funding provided by the British Columbia Ministry of Environment and Climate Change Strategy. The authors gratefully appreciate the reviews and comments from Klaus Rathfelder (ENV), Julie-Ann Ishikawa (ENV), Shirley Wang (FLNRORD), Diana Allen (Simon Fraser University), Mike Simpson (City of Vancouver), Robin Pike (ENV), Ben Kerr (Foundry Spatial), Harm Demon (FLNRORD), and Amy Sloma (ENV).

EXECUTIVE SUMMARY

Groundwater is a critical resource for society and aquatic ecosystems. Groundwater pumping can negatively affect rivers, lakes, wetlands, and other surface water bodies by reducing groundwater discharge into surface water bodies or inducing infiltration through the streambed. Combined, these impacts are known as *streamflow depletion*. Understanding streamflow depletion can guide water management for issuing sustainable groundwater licenses while maintaining environmental flow needs.

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. As a result, streamflow depletion is typically estimated using numerical and/or analytical models. Numerical models are generally developed for site-specific investigations and require large amounts of time, effort, data, and professional experience to calibrate, validate, and implement. In contrast, analytical models are often adopted for preliminary assessment and/or in settings with limited resources or data due to their relatively low data requirement and ease of implementation. 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 impacts of pumping on streamflow depletion (e.g. Rathfelder, 2016). However, analytical models include several potentially problematic assumptions, e.g., representing streams as linear features bounded by an infinite horizontal aquifer, making it unclear whether they are appropriate tools for decision-making.

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, 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. (2019A) evaluated 50 analytical depletion functions by comparing them against an archetypal groundwater model, which is a simplified uncalibrated groundwater flow model. While these results indicated that analytical depletion functions could provide accurate predictions of streamflow depletion, they were only tested in a single watershed in California. It is not clear whether analytical depletion functions will be similarly accurate for other regions with different stream network geometries, hydrogeological characteristics, and bioclimatic conditions, such as British Columbia (BC). Another notable negative impact of groundwater pumping on streamflow is the time taken for streamflow depletion to reach the environmental flow threshold, i.e. response time to environmental flow threshold. In this report, we adopted the presumptive standard of an environmental flow threshold, which is defined as 90% of monthly natural baseflow and is recommended to meet a high level of ecological protection (Gleeson and Richter, 2018). This presumptive standard provides a universal metric that can be compared between different regions without defining the watershed specific environmental flow need (EFN) and has been previously applied in BC (Forstner et al., 2018). Therefore, the presumptive standard is used in this study rather than applying the BC EFN policy which does not differentiate the baseflow component of streamflow. Thus far, analytical depletion functions have only been evaluated for predictions of the magnitude and timing of streamflow depletion, and it is unclear whether they are suitable to estimate the response time between pumping and environmental flow thresholds being surpassed. Therefore, the goals of this study are to:

1) evaluate the accuracy of analytical depletion functions for estimating impacts of groundwater pumping on streamflow in two different regions in BC;

2) examine whether analytical depletion functions can be used to evaluate the response time to environmental flow thresholds in different watersheds in BC; and
3) understand the applicability of analytical depletion functions in evaluating streamflow depletion and environmental flows.

To accomplish these goals, we compared output from analytical depletion functions to existing calibrated numerical models in BC, which include more detailed representation of subsurface flow processes and surface water-groundwater interactions. We initially obtained four calibrated numerical models from regions across BC: BX Creek in the semi-arid interior plateau and highlands, Peace region in the boreal plain, Abbotsford-Sumas in the coastal basin and lowlands, and Bevan Wellfield in the coastal basin characterized with semi-confined aquifers. However, we found that numerical model design has a significant impact on the ability to simulate reasonable streamflow depletion estimates.

Our evaluation of analytical depletion functions, therefore, focused on the following two domains, BX Creek and Peace region. For these two domains, the existing numerical groundwater models were built in MODFLOW and calibrated for simulating steady-state conditions. We converted the models to transient conditions at a weekly time step in order to simulate the effects of seasonal groundwater pumping for irrigation on surface water features within each domain. In each domain, we implemented >90 synthetic pumping wells with seasonal pumping schedules for a period of 30 years and compared the response of each stream segment to pumping between the MODFLOW models and the analytical depletion functions.

We used three metrics to assess the performance of the analytical depletion functions relative to the MODFLOW models: (i) correct identification of the stream segment most affected by a well, quantified as a percent accuracy; (ii) correct prediction of depletion rate in the stream segment most affected by a well, quantified using normalized mean absolute error (MAE), which is the MAE divided by the highest pumping rate applied to the model; and (iii) correct prediction of depletion in all stream segments affected by a well, also quantified using normalized MAE. Agreement between the numerical model and analytical depletion functions estimates would indicate that both approaches provide similar predictions of streamflow depletion. It should be noted that the two approaches are both modelled outcomes; our study design is not capable of evaluating whether analytical depletion functions are accurately reproducing the response to pumping using real-world observations.

The MODFLOW results showed that realistic rates of groundwater pumping could have significant impacts on streamflow in both domains. However, the two domains had different responses to pumping wells due to differences in hydrogeological settings, stream networks, and model structure. For instance, groundwater pumping had limited impacts on dry or ephemeral streams in the BX Creek model and thus need to be accounted for when assessing streamflow depletion.

Two analytical depletion functions were applied in the BX Creek and Peace region domains to estimate the seasonal pumping on streamflow depletion. We found that analytical depletion functions built using two different analytical models, Glover and Balmer (1954) and Hunt (1999), produced similar streamflow depletion estimates. In the BX Creek domain, analytical depletion functions correctly identified the most affected stream segments for all wells over the entire 30-year simulation with the average normalized MAE of 14.4% (MAE is normalized by the highest pumping rate). For all affected stream segments, the average normalized MAE was 5.0%. In the Peace region, which has a larger stream network than BX Creek, the analytical depletion functions correctly identified the most affected stream segment for fewer wells than the BX Creek domain, but can reach up to 83% of the time. The average normalized MAE was 7.6% for the most affected segment and 2.3% for all stream segments. In addition, analytical depletion functions have better prediction in pumping season than those in the non-pumping season in both domains. Overall, these comparisons indicated that analytical depletion functions can be an accurate tool for estimating streamflow depletion within the pumping season at over annual to decadal timescales, with worse performance for sub-annual depletion.

The performance of analytical depletion functions varied in response to the hydrogeologic settings. Analytical depletion functions had smaller errors for wells in higher hydraulic conductivity materials, screened in shallower aquifers, and affecting segments with lower streambed conductance in the BX Creek, a finding which is consistent with other studies (e.g. Zipper et al., 2018). Conversely, in the Peace region, analytical depletion functions had smaller errors in lower hydraulic conductivity materials and deeper aquifers, while the performance over the range of streambed conductance conditions tested was similar. In both domains, there was consistently greater difference between analytical depletion functions and MODFLOW for wells within a few kilometers of a stream, which correspond to wells with the highest predicted depletion. The contrasting responses of analytical depletion functions performance to the hydrogeological setting between the BX Creek and Peace region models indicates that different drivers influence the accuracy of streamflow depletion prediction in different hydrogeological settings and thus it is important to conduct additional testing of analytical depletion functions in other areas to identify the drivers of performance for a particular hydrogeological area.

Moreover, our results suggested that analytical depletion functions can be used to assess the response time over which groundwater pumping causes streamflow depletion to exceed environmental flow thresholds. Realistic groundwater pumping in the two domains leads to depletion exceeding a presumptive standard for protecting environmental flow needs (i.e., 10% of reduction of baseflow), indicating that potential groundwater pumping in study domains should be considered in water management. In the BX Creek, the difference in response time to environmental flow threshold between analytical depletion functions and MODFLOW in the BX Creek was smaller (~ 2 years) than the difference between models in the Peace region, while analytical depletion function showed both shorter and longer response time than MODFLOW. While, in the Peace region, the analytical depletion functions showed shorter response times than MODFLOW. These results indicate that analytical depletion functions can be used to estimate the response time of environmental flows, but potential errors resulting from the local hydrogeological setting should be considered.

Finally, we found that calibrated numerical models, often considered the 'gold standard' for streamflow depletion prediction, may not always be appropriate for quantifying streamflow depletion if they were previously developed for a different purpose. Each numerical model is developed and calibrated for a specific purpose and targets chosen to represent processes relevant to that purpose. When a model that is not designed for quantifying streamflow depletion is then used for simulating the response of a stream to pumping, significantly large errors are possible. For example, the numerical model developed for the Abbotsford-Sumas domain was designed to assess regional groundwater sustainability under future climate change. In this model, mass balance changes in the aquifer are much higher than the pumping rate for a single well due to the large domain and highly permeable aquifer, and therefore the impacts of that well cannot be reliably separated from potential model error.

Recommendations:

1. Analytical depletion functions are useful for assessing streamflow depletion for perennial streams during the pumping season in the absence of numerical models, although analytical depletion functions have significant errors and uncertainties that should be considered when using the analytical depletion functions for decision making.
2. Analytical depletion functions performance and response to hydrogeological setting vary between the two domains tested in this study, which highlights the uncertainties decision-makers will face

when using analytical depletion functions. For simple stream network and aquifer types as in BX Creek, analytical depletion functions performed better in the regions with highly conductive alluvial aquifers and shallower well depth. For complex hydrogeological settings as in the Peace region, analytical depletion functions performed better for wells with lower hydraulic conductivity and deeper well depth. The findings of this study also demonstrated that the analytical depletion functions conservatively estimate the time it takes for streamflow to drop below a presumptive environmental flow threshold as a result of groundwater pumping from a single well. However, due to the different performance of analytical depletion functions over the two domains in different hydrogeological landscapes, our preliminary assessment is not sufficient to make generalized conclusions regarding the applicability of analytical depletion functions across BC or specific hydrogeological settings. Therefore, we recommend that analytical depletion functions should be used with caution for broader application and that a good understanding of the local hydrogeological conditions outside of the tested domains are needed.

3. Analytical depletion functions can be applied to estimate the response time of streamflow to pumping. In both domains, analytical depletion functions calculated a shorter response time to environmental flow threshold resulting from groundwater pumping than predicted by numerical models particularly in the Peace region. Differences in response time between two domains suggest that the lagged impacts of pumping on streamflow will vary across different hydrogeological settings. We recommend that the response time to environmental flow threshold is likely region-specific and, therefore, decisions related to the response time to environmental flow threshold should be designed for an individual watershed.
4. Like the analytical depletion functions, numerical models are mathematical simplifications of reality which are designed for a specific purpose. Even calibrated numerical models may be inaccurate for predicting streamflow depletion if they were designed for a different purpose. Therefore, we recommend that decision-makers assess the performance of any previously developed numerical models specifically for streamflow depletion prior to use, rather than assuming that a calibrated model will provide accurate predictions of streamflow depletion. Accordingly, we recommend that only numerical models that have been evaluated and demonstrated to generate accurate stream-aquifer interactions be used for streamflow depletion-related decision making.

CONTENTS

ACRONYMS

BC	British Columbia
CHD	Constant Head (in MODFLOW)
DRN	Drain Boundary Condition (in MODFLOW)
d	Well Stream Distance
EFN	Environmental Flow Needs
ENV	Ministry of Environment and Climate Change Strategy
FLNRORD	Ministry of Forests, Lands, Natural Resource Operations and Rural Development
GHB	General Head Boundary (in MODFLOW)
MAE	Mean Absolute Error
T	Transmissivity
RIV	River Boundary Condition (in MODFLOW)
S	Storativity or storage coefficient
STR	Stream Boundary Condition (in MODFLOW)
λ	Streambed Conductance

1. MOTIVATION

Groundwater pumping can reduce groundwater discharge into surface water bodies or, in severe cases, induce infiltration through the streambed (Figure 1a; Barlow and Leake, 2012). Combined, these impacts are known as *streamflow depletion* and can have significant impacts on environmental flow needs (Gleeson and Richter, 2018). Shortly after the start of pumping, most water comes from storage, but as the well continues to pump, more and more of the water comes from capture rather than storage (Figure 1b) which has profound implications for surface water and aquatic habitat. Assuming groundwater recharge or non-stream discharge (e.g., phreatophytic evapotranspiration) does not change, all abstraction eventually leads to a reduction in streamflow (Figure 1c). Streamflow depletion can be insignificant relative to environmental flows or may occur over a long time (hundreds to thousands of years), and the critical information needed for government staff making water allocation decisions is the timing and severity of streamflow depletion, relative to environmental flows. In addition, streamflow quality and quantity have been significantly disturbed by climate and land cover and land-use change (e.g. forest logging and fires) which threaten salmonid populations, for example, as well as the health of aquatic habitat (Kiffney et al., 2002; Wei et al., 2018). Moreover, groundwater demand for agricultural irrigation is likely to increase under the warming climate (St. Jacques et al., 2010), which may further impact streamflow, particularly during the dry season. Therefore, understanding the impacts of streamflow depletion quantity and timing relative to environmental flow thresholds is critically important for sound water management.

Numerical and analytical models are two typical approaches to assess the impacts of groundwater pumping on streamflow. Numerical models are typically used for site-specific investigations and require extensive effort and knowledge to calibrate and validate. Sometimes, water managers cautiously transfer the knowledge gained from one location to another, but differences in hydrogeological and climate conditions between areas can cause significant uncertainty and hinder 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). However, analytical models have many simplifying assumptions, for instance, an infinite horizontal aquifer bounded by a single linear stream, which can limit their application in the real-world hydrogeological settings with complex stream networks of multiple and meandering stream segments.

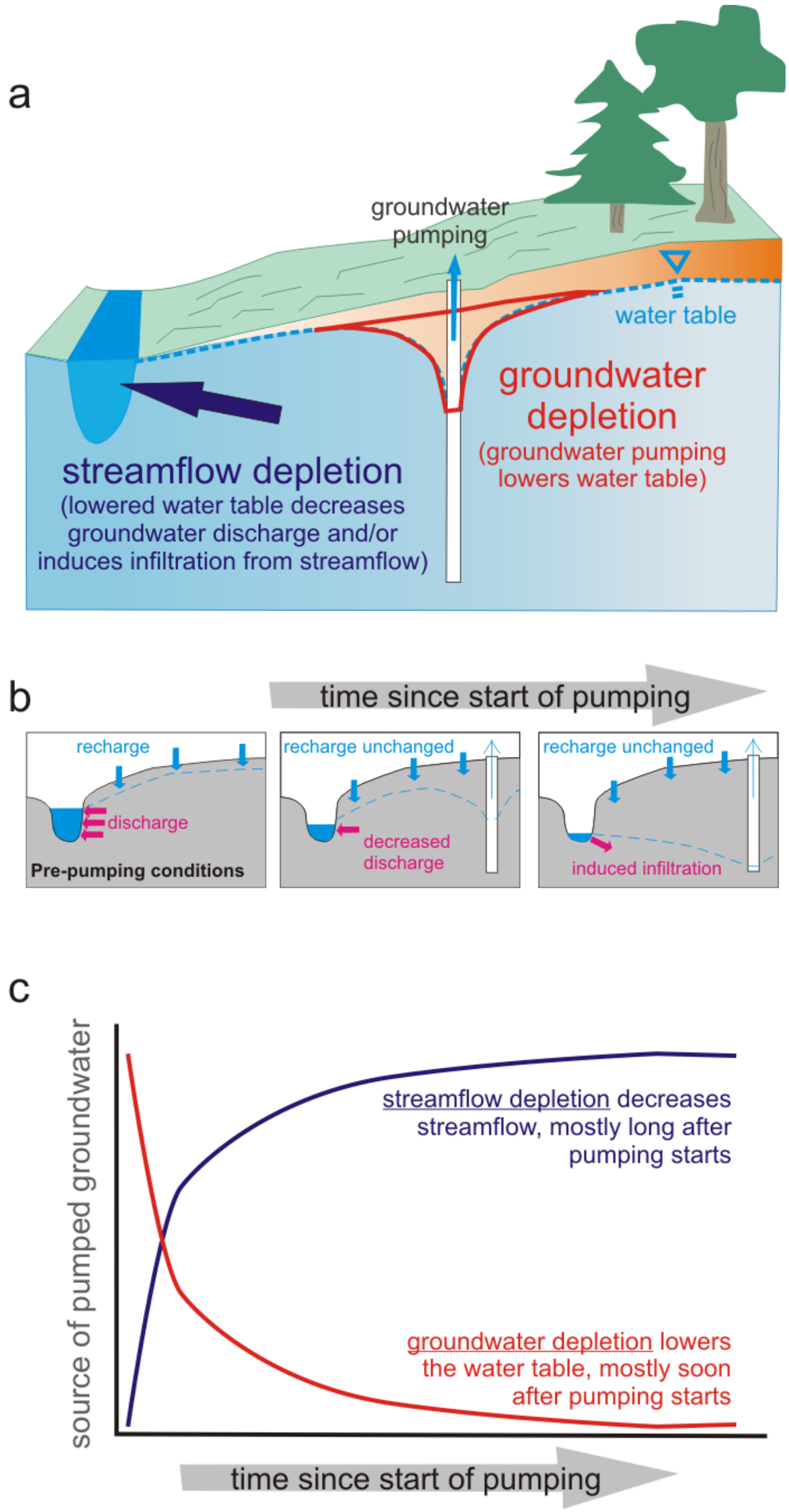


Figure 1: a) Streamflow depletion and groundwater depletion by groundwater pumping. (b) Streamflow depletion change with pumping duration: decreased groundwater discharge or induced infiltration of water from the river leads to streamflow depletion. (c) Source of pumping groundwater from streamflow depletion and groundwater depletion through time since the start of pumping. Graphics from Gleeson and Richter, 2018.

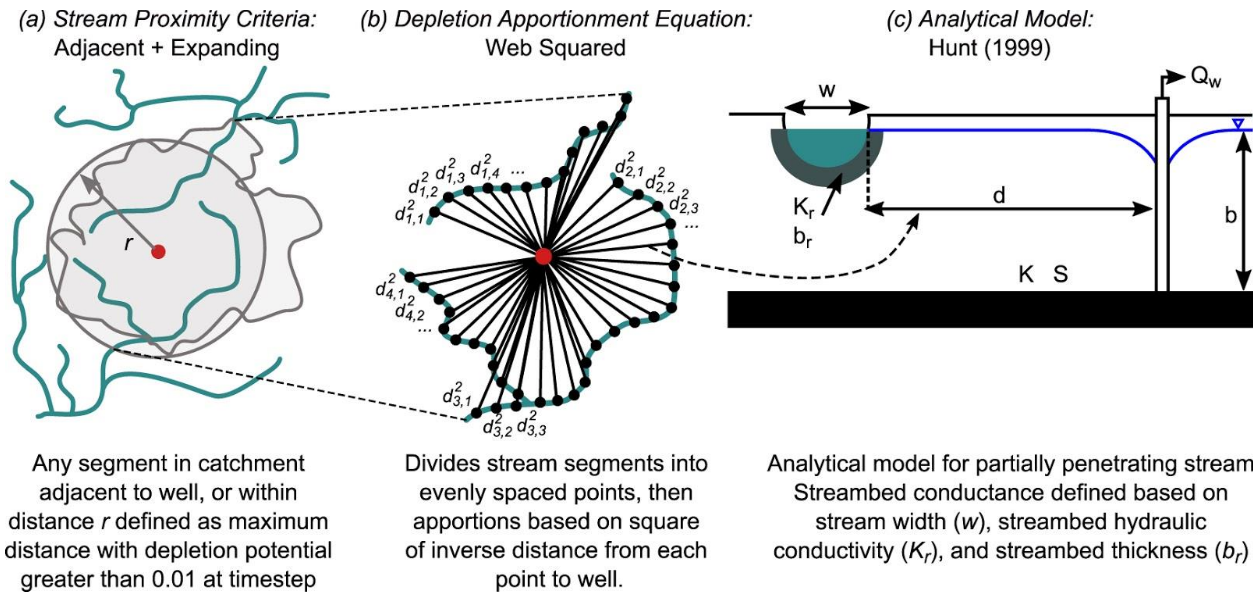


Figure 2: The best performing analytical depletion function developed by Zipper et al. (2019 A), including: (a) stream proximity criteria; (b) depletion apportionment equation; and (c) analytical model. Figure from Zipper et al. (2019B) used under Creative Commons BY 3.0 license.

To advance the application of analytical models in real-world settings with multiple stream segments and sinuous stream networks, Zipper et al. (2019A) developed analytical depletion functions (Figure 2), which consist of (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. (2019A) evaluated 50 analytical depletion functions by comparing them against an archetypal groundwater model, which is an un-calibrated and simplified numerical groundwater model. Although the analytical depletion functions provided accurate predictions of streamflow depletion in this test, several knowledge gaps remain for their real-world application. First, analytical depletion functions have only been tested in limited domains, for instance, Nanaimo, British Columbia (Zipper et al., 2018) and Navarro River watershed, California, USA (Zipper et al., 2019A). It is unclear whether analytical depletion functions would be similarly accurate in other domains with different stream networks and hydrogeological conditions. Second, both of these prior studies assessed analytical depletion function performance by comparing the results to un-calibrated archetypal numerical models, which simplify hydrogeological conditions and hence may not be representative of real-world settings with complex hydrogeological settings and stream networks. It is, therefore, necessary to test analytical depletion functions against calibrated numerical models to enhance the confidence in applying analytical depletion functions in the real-world. Finally, groundwater pumping can negatively affect the EFN of a stream, which is defined as *"the volume and timing of water flow required for proper functioning of the aquatic ecosystem"* (Province of British Columbia, 2016). This definition emphasizes the importance of both flow quantities and timing. In addition, the *Water Sustainability Act* (2016) mandates that groundwater pumping must not reduce streamflow in hydraulically connected surface waters below the EFN.

Thus far, there are many methods existing to assess the environmental flows for the specific watershed. For instance, BC has developed the Modified-Tennant method (Ptolemy and Lewis, 2002), which incorporates the biological and geomorphological information for a specific region. In this method, biological information considers fish species and their life stages, such as spawning, incubation, migration, active rearing, as well as ecological needs (Ptolemy and Lewis, 2002). The flow duration curve is another commonly used method to quantify EFN and multiple flow duration indices are used worldwide, e.g. 7Q2 is the lowest flow for seven consecutive days within a 2-year return period and this index has been applied in Quebec (Caissie et al. 2007). In BC, hydrologists are promoting 30Q2 as a new metric for the environmental flow needs (Personal communication, 2020). Clearly, these indices are watershed-specific, which require large efforts to quantify and make it challenging to compare EFN thresholds between watersheds. Moreover, they are often developed at the watershed scale due to the availability of observed hydrology data. As such, the spatial EFN of each individual stream segment within a watershed is often unknown. Recently, Gleeson and Richter (2018) proposed a presumptive standard of an environmental flow threshold, which is defined as 90% of monthly natural baseflow that should be maintained in order to meet the high level of ecological protection. A presumptive standard provides a universal metric that can be compared between different regions and within a watershed. In addition, the presumptive standard has been validated in BC (Forstner et al., 2018). Although previous literature has shown that analytical depletion functions can accurately assess streamflow depletion in comparison to un-calibrated archetypal numerical models (e.g., Zipper et al., 2018), there are knowledge gaps regarding the utility of these tools to support management decisions related to groundwater pumping and response time to environmental flow thresholds, which is defined as the duration of pumping time for pumping a well to cause streamflow to drop below an environmental flow threshold. In response to these knowledge gaps, this study has three goals:

1) evaluate the accuracy of analytical depletion functions for estimating impacts of groundwater pumping on streamflow through comparisons to calibrated numerical models;
2) examine whether analytical depletion functions can be used to evaluate the response time to environmental flow threshold using the presumptive standard; and
3) understand the applicability of analytical depletion functions in evaluating streamflow depletion and environmental flow threshold.

2. METHODS

To assess the performance of analytical depletion functions across hydrogeological and climate settings, we selected two study domains within BC that have contrasting hydrogeology, climate, topography, and ecology (Figure 1). For these two domains, we modified existing pre-calibrated numerical models, built in MODFLOW, to simulate the pumping impacts on streamflow depletion for comparison with analytical depletion functions. We treated streamflow depletion simulated by the numerical models as reference or "observed" values for comparison with analytical depletion functions output since numerical models include more detailed process-based representation of subsurface flow, and both models were calibrated, though 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 (Figure 3). In this study, two of these domains were investigated. The BX Creek model represents interior plateaus and highlands and the Peace region model represents boreal plains. In addition, these domains offer contrasting stream network complexities, as BX Creek represents a small aquifer with simple stream networks while the Peace region model represents a large regional aquifer with a complex stream network.

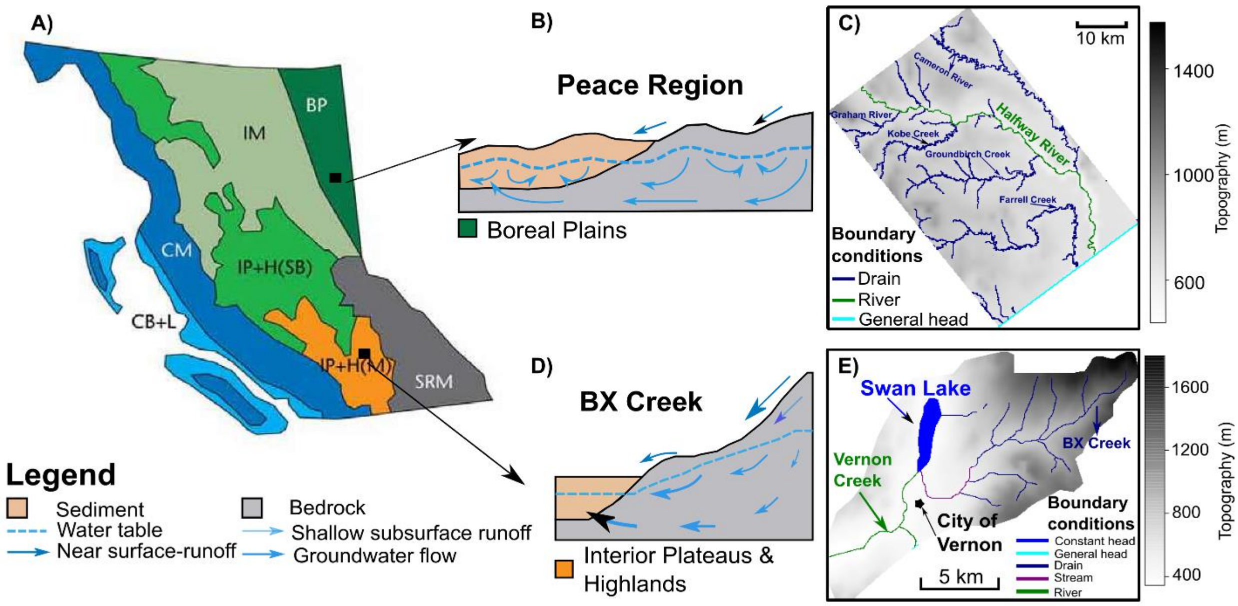


Figure 3: A) Hydrogeological landscapes in British Columbia, Canada. CB+L: Coastal Basins and Lowlands; CM: Coastal Mountains; IP+H(M): Interior Plateaus & Highlands (Montane); IP+H(SB): Interior Plateaus & Highlands (Sub-Boreal); IM: Interior Mountains; SRM: Southern Rocky Mountains; and BP: Boreal 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., 2009A used Attribution-NonCommercial-NoDerivs 2.5 Canada (CC BY-NC-ND 2.5 CA).

2.1.1 Numerical model in Interior Plateaus and Highlands (BX Creek)

The BX Creek is located in the southern interior of British Columbia (Figure 3). The numerical model in the BX Creek was initially developed to validate a long-term recharge rate estimation method for mountainous regions, and the numerical model revealed that recharge rate estimation method was accurate (Smerdon et al., 2009B). BX Creek is characterized by snowmelt-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 the valley bottom is recharged by surface runoff as well as receiving local and regional groundwater flows (Figure 3). Model development, conceptualization, and parameterization are described in detail in Smerdon et al. (2009B). Here we briefly review the model settings. The model has a uniform grid resolution of 50 x 50 m with 327 rows, 400 columns, and 8 layers aligned with lateral boundaries between layers following the land surface. The modelled area is about 164.8 km^2. Based on the borehole logs from 611 water wells, the subsurface is parameterized as four hydrostratigraphic units (Table 1, Figure 4 and Figure 5).

Table 1: Hydraulic conductivity of four hydrostratigraphic units in the BX Creek model.

Materials	**K_x (m/s)**	**K_y (m/s)**	**K_z (m/s)**	**Specific Yield (S_y)**	**Specific Storage (S_s) (m^{-1})**
Alluvial sediment and aquifer	1 x 10^{-5}	1 x 10^{-5}	1 x 10^{-6}	0.15	1 x 10^{-5}
Glaciolacustrine sediments	7 x 10^{-7}	7 x 10^{-7}	1 x 10^{-8}	0.15	1 x 10^{-5}
Mixed sediments	1 x 10^{-6}	1 x 10^{-6}	1 x 10^{-7}	0.15	1 x 10^{-5}
Bedrock	4 x 10^{-8}	4 x 10^{-8}	4 x 10^{-8}	0.02	1 x 10^{-4}

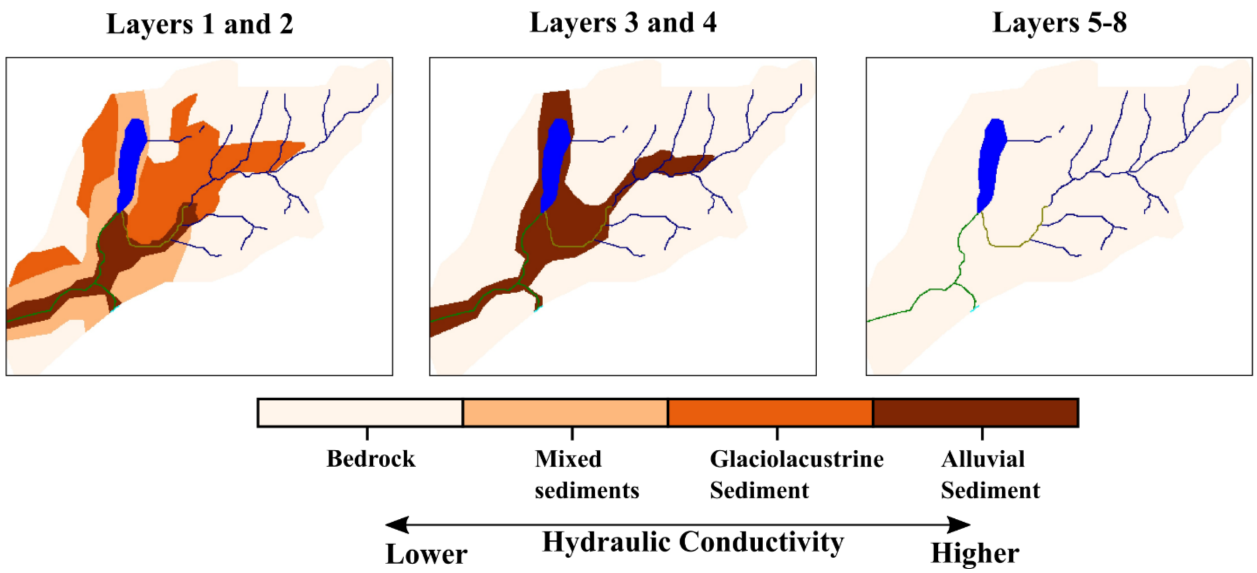


Figure 4: Spatial distribution of hydrostratigraphic units in the MODFLOW for the BX Creek. Deep model layers have large numbers. The hydraulic conductivity of the materials is listed in the Table 1.

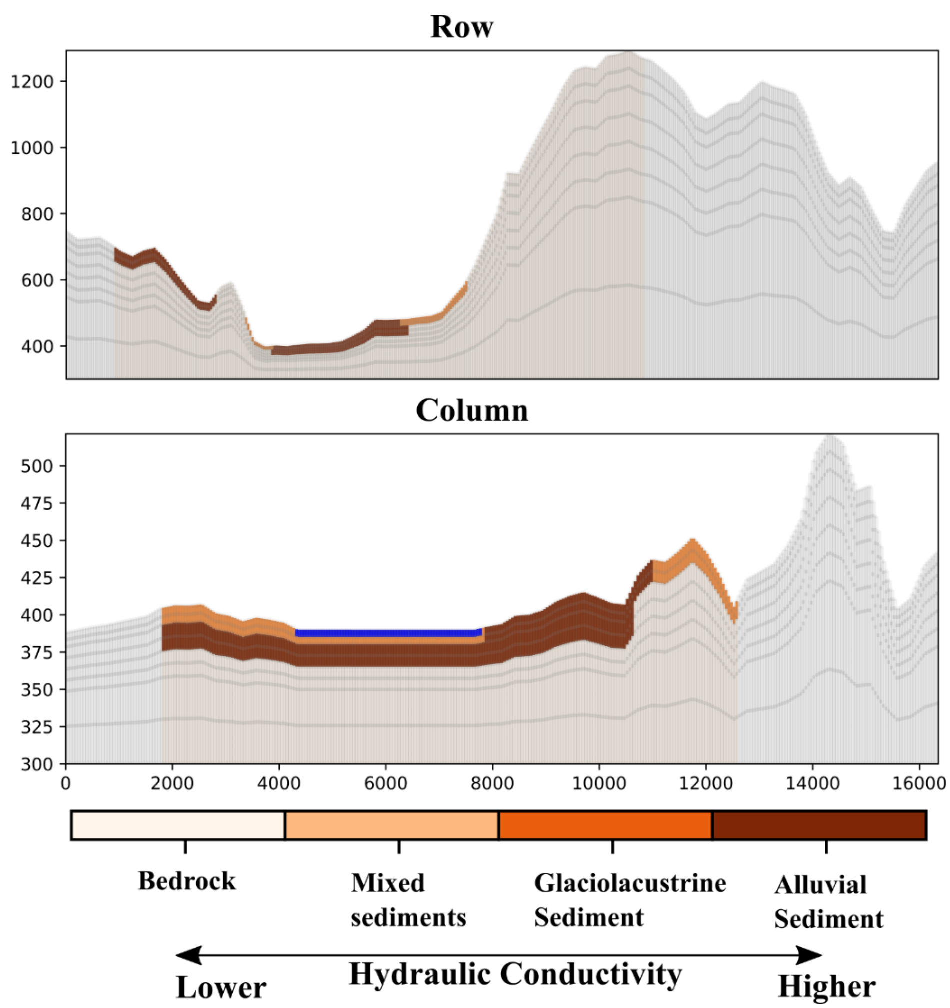


Figure 5: Cross-section view of hydrostratigraphic units in the MODFLOW for the BX Creek.

The original BX Creek model includes five boundary conditions: constant head (CHD package in MODFLOW), general head (GHB), drain (DRN), river (RIV), and stream (STR). As shown in Figure 6A, 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 depletion functions are not designed to represent seepage faces. We found that removing these drain blocks had minor impacts on the model mass balance change, so they were removed from the numerical model (Figure 6B) while keeping the other boundary conditions.

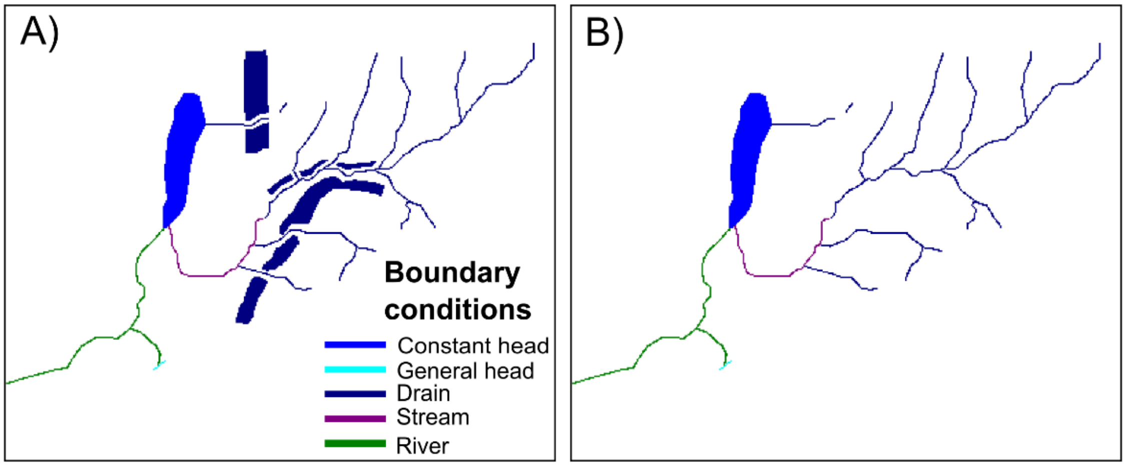


Figure 6: Flow boundary conditions in the BX Creek model in the MODFLOW Settings. A) BX Creek model with drain blocks representing the groundwater seepage at the mid-elevations. B) Drain blocks were removed from the BX Creek for streamflow depletion assessment.

2.1.2 Numerical model in the boreal plain (Peace region)

Unlike the conceptual model in the BX Creek, the boreal plains have a gentle undulating topography and water tends to be retained on the surface, forming wetlands and ponds (Smerdon et al., 2019A). Such landscapes also form nested groundwater flow systems, and hence complex surface and groundwater interactions are found in this region (Figure 3). The Peace region is located in the northeast of British Columbia (Figure 3). The objective of the Peace region MODFLOW model was to understand the connection between local and regional groundwater flow in the buried valley aquifers, and simulation results suggested that buried valleys are not regionally connected throughout the whole network (Morgan, 2018; Morgan et al., 2019).

The MODFLOW model in the Peace region has 327 rows and 308 columns with a uniform cell size of 200 m x 200 m and 20 layers aligned with the watershed topography. The modelled area is about 1952 km^2. In this model, six hydrostratigraphic units were assigned and are shown in Figure 7-8 and Table 2. To simulate the surface water features, three boundary conditions were assigned, including drain (DRN), river (RIV), and general head (GHB) in MODFLOW (Figure 11A). The river boundary represents the Halfway River and the drain boundary represents its tributaries. The general head, which is in the south boundary, is to account for the influence of the Peace River, the mainstream of Halfway River (Figure 3).

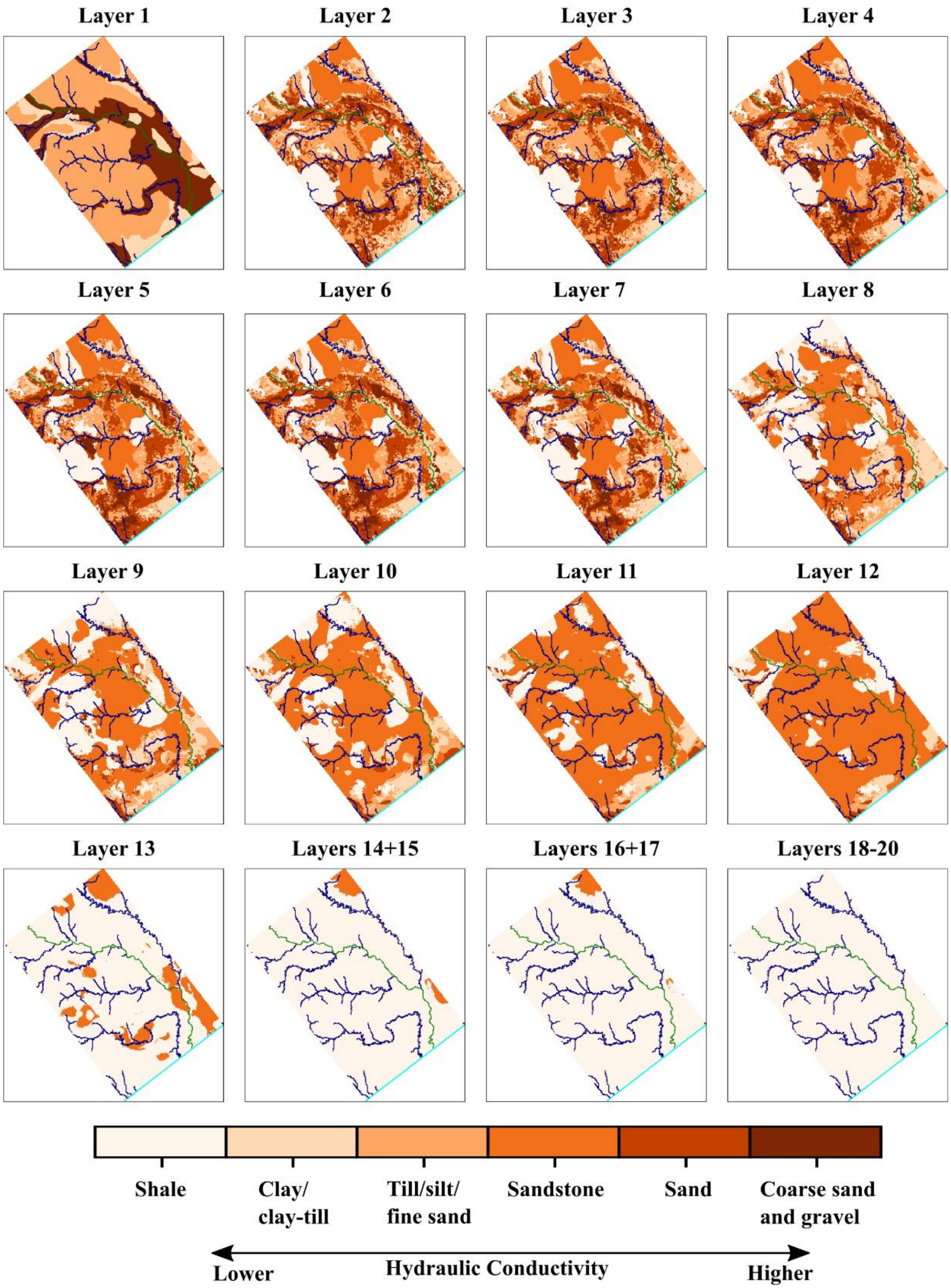


Figure 7: Spatial distribution of hydrostratigraphic units in the MODFLOW for the Peace region. Deeper layers have larger layer numbers. The hydraulic conductivity of the material is listed in Table 2.

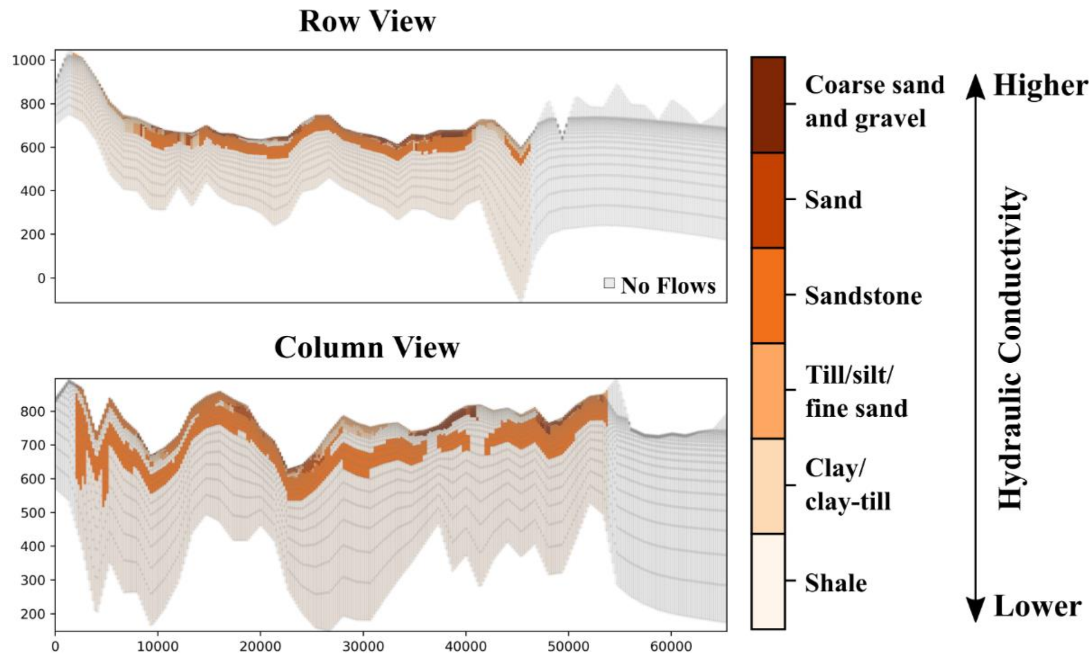


Figure 8: Cross-section view of hydrostratigraphic units in the MODFLOW model for the Peace region. The hydraulic conductivity of the material is listed in Table 2.

Table 2: Hydraulic conductivity of six hydrostratigraphic units in the Peace region model

Materials	K_x (m/s)	K_y (m/s)	K_z (m/s)	Specific Yield (S_y)	Specific Storage (S_s) (m^{-1})
Coarse sand and gravel	3×10^{-3}	3×10^{-3}	1×10^{-3}	0.15	2×10^{-4}
Sand	5×10^{-5}	5×10^{-5}	1×10^{-5}	0.15	2×10^{-4}
Till/silt/fine sand	4×10^{-7}	4×10^{-7}	2×10^{-8}	0.02	1×10^{-5}
Clay/Clay-till	1×10^{-8}	1×10^{-8}	1×10^{-10}	0.02	1×10^{-5}
Sandstone	1×10^{-5}	1×10^{-5}	1×10^{-6}	0.02	1×10^{-5}
Shale	7.7×10^{-10}	7.7×10^{-10}	7.7×10^{-12}	0.02	1×10^{-5}

2.2 The impacts of pumping on streamflow depletion using numerical models

2.2.1 Pumping schedule

To test the impacts of pumping on streamflow, we developed pumping schedules typical of agricultural irrigation in the two domains using the BC Agriculture Water Calculator (http://www.bcagriculturewatercalculator.ca/), which estimates the monthly water demand for various crops based on the soil types and climate conditions. We used the same pumping schedule based on BX Creek irrigation demands for both domains because we wanted to have a consistent stress on groundwater system for direct comparison of streamflow depletion in the two domains, and the average demands in BX Creek were higher than that in the Peace Region. We selected four dominant agricultural crops in the BX Creek (apple, cheery, forage and grape) and assumed the irrigated area from each groundwater well was considered to be equal to the size of the average land parcels within the domain. 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 (Figure 9).

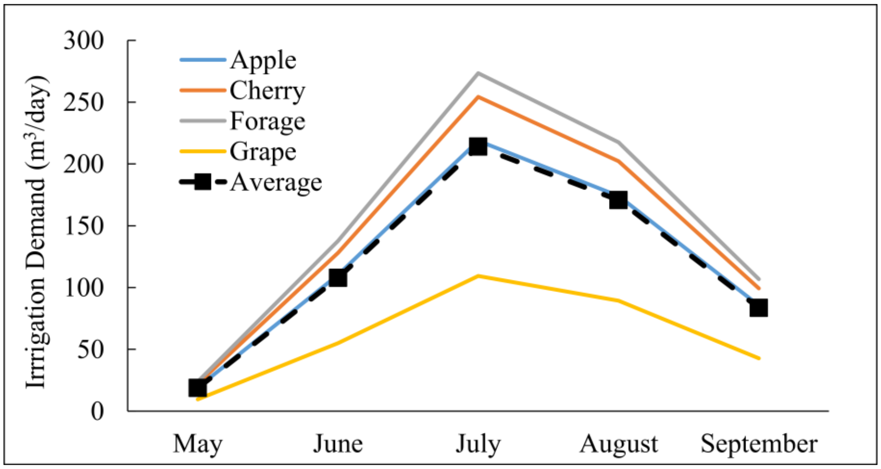


Figure 9: Monthly water demand for major crops in the interior plateau and highland were estimated using the BC Agriculture Water Calculator (http://www.bcagriculturewatercalculator.ca/).

2.2.2 Streamflow depletion assessment using numerical models

The numerical models for BX Creek and the Peace region were originally calibrated for steady-state conditions. Because our goal was to simulate time-varying streamflow depletion by groundwater pumping and associated response time to environmental flow threshold, we converted both models to transient conditions using a weekly stress period. To do so, the boundary conditions in the steady-state MODFLOW model were then used for each stress period to reduce the uncertainties in the model simulations. We converted the annual recharge rate to weekly for each stress period (i.e., no seasonal variation). We used specific yield and specific storage that were initially estimated based on the hydrostratigraphic units used in the steady-state calibrated models and adjusted these values within a reasonable range (Table 1 and Table 2) to minimize mass balance errors in numerical models. To stabilize numerical model performance, 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 the transient pumping test simulations. Finally, a total simulation period was set for 35 years with the first five years as a spin-up to stabilize the numerical models and the pumping test for the final 30 years with a total number of 1820 weeks.

To systematically compare the analytical and numerical models in a number of settings, we introduced synthetic groundwater wells at regular spacing throughout the two domains to stress the aquifer. Synthetic wells were screened at a depth of 15 m below the water table in the BX Creek to ensure pumping from the shallow aquifer, 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. As a result, a total of 99 and 96 synthetic pumping wells were created in the BX Creek and Peace region, respectively, corresponding to a well density of 0.6 wells/km^2 and 0.05 wells/km^2 for the BX Creek and Peace region, respectively. Each well has the same seasonal pumping schedule for each of the 30 simulation years (Figure 9).

To calculate streamflow depletion caused by each synthetic well, the model was 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 pumped simulation and the baseline simulation.

To quantify the streamflow depletion in MODFLOW, we calculated the annual streamflow depletion factor, which is the ratio of the sum of streamflow depletion across all affected streams in the domain 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.

2.3 Analytical depletion functions

In this study, we used the highest performing stream proximity criteria and depletion apportionment equation identified by Zipper et al. (2019A) (Figure 2). Stream proximity criteria define the stream segments, which could potentially be impacted by a well, and 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 apportioned to each stream segment. We used the *Web Squared* depletion apportionment equation, which splits each stream segment into a finite number of points (e.g. space between each point is 5 meters) and apportions based on the square of the inverse distance of each stream segment to the well as shown in Eqn. (1).

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

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 in 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 from the depletion apportionment equation. As such, the streamflow depletion by a pumping well is estimated. 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 Eqn. (2).

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

where S is the aquifer storage coefficient (e.g., specific yield in an unconfined aquifer), T is the aquifer transmissivity, t is the time since the start of pumping, d is the well-stream distance, and Q_w is the pumping rate.

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

$$Q_a = Q_w * (\mathrm{erfc}(\sqrt{\frac{Sd^2}{4Tt}} - \exp(\frac{\lambda^2 t}{4ST} + \frac{\lambda d}{2T})\mathrm{erfc}(\frac{\lambda^2 t}{4ST} + \frac{\lambda d}{4Tt})) \quad (3)$$

where λ is the streambed conductance. The streambed conductance is defined as $\lambda = w_r * \frac{K_r}{b_r}$, where w_r is the width of the stream segments.

The analytical depletion functions 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 to minimize differences in parameters between the MODFLOW and analytical depletion functions. For the analytical models, 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 analytical depletion function input parameter is calculated as follows:

1) Transmissivity (*T*). MODFLOW uses hydraulic conductivity as input while the analytical models use transmissivity, which is equal to hydraulic conductivity multiplied by the aquifer thickness. To obtain the transmissivity for the analytical models, the average hydraulic conductivity between the surface and well depth in the MODFLOW model was used. The aquifer thickness at the well location in the steady-state condition was used.
2) Specific yield (*S*). Average specific yield between the surface and well depth in the MODFLOW model at the well locations was 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 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 and the conductance is not required. We used the highest conductance values of other boundary conditions for the constant head, while λ of other boundary conditions was provided 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. The inclusion of these boundaries in analytical depletion functions could lead to a biased streamflow depletion assessment. Therefore, these flow features were excluded for the analytical depletion functions in the BX Creek domain. In contrast, both gaining and ephemeral stream types are detected within the drain boundary in the Peace region (Figure 11C) and thus all features were included in the analytical depletion functions.

2.4 Metrics for comparing analytical depletion functions and numerical model results

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

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

2) *Magnitude of primary impact* quantifies how big the difference is between analytical depletion functions and MODFLOW in predicting the streamflow depletion. We quantify this as the normalized MAE (mean absolute error), which is normalized by the highest pumping rate between the streamflow depletion predicted by the analytical depletion functions and MODFLOW.
3) *Magnitude of overall impact* quantifies the difference in all affected stream segments by analytical depletion functions and MODFLOW. This metric is expressed by the MAE of streamflow depletion caused by the analytical depletion functions and MODFLOW.

2.5 Application of analytical depletion functions in assessing response time of environmental flow needs

Although many methods have been used to define the EFN threshold (see Section 1), in this study, we use the presumptive standard of an environmental flow threshold, which states the decrease in monthly natural baseflow by groundwater pumping should be less than 10% (Gleeson and Richter, 2018). As a proxy for baseflow, we used weekly stream-aquifer exchange estimated by MODFLOW for each stream segment. Stream-aquifer exchange in the no-pumping baseline scenario was treated as a proxy for natural baseflow. The response time to the environmental flow threshold is defined as the pumping duration at which streamflow depletion is equal to 10% of this natural baseflow. In this report, we initially evaluated the response time to the environmental flow threshold for each weekly stress period and for the most affected stream segments identified by the MODFLOW models and analytical depletion functions. We first selected stream segments that reached 1% of baseflow reduction due to pumping (based on estimates from analytical depletion functions) and then examined response time of selected segments to reach 5% and 10% of baseflow reduction. When the response time of some segments was longer than the simulation time of 30 years, the response time of these segments was set to 30 years so that response time simulated by each method has a value and thus they were comparable. The difference in response time between the analytical depletion functions and MODFLOW models was calculated to examine the performance of analytical depletion functions in simulating the response time. Positive differences indicated that MODFLOW simulated a shorter response times, while negative differences indicated analytical depletion functions estimated a shorter response time (e.g. Figure 16).

3. <u>RESULTS AND DISCUSSION</u>

3.1 Streamflow depletion simulation by MODFLOW

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 (Figure 10C). Our numerical 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 Figure 10 D-E, 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 duartion. 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% of wells in the fifth year. For each specific year of pumping, streamflow depletion factor variation ranged from 0 to 100% across the study domain.

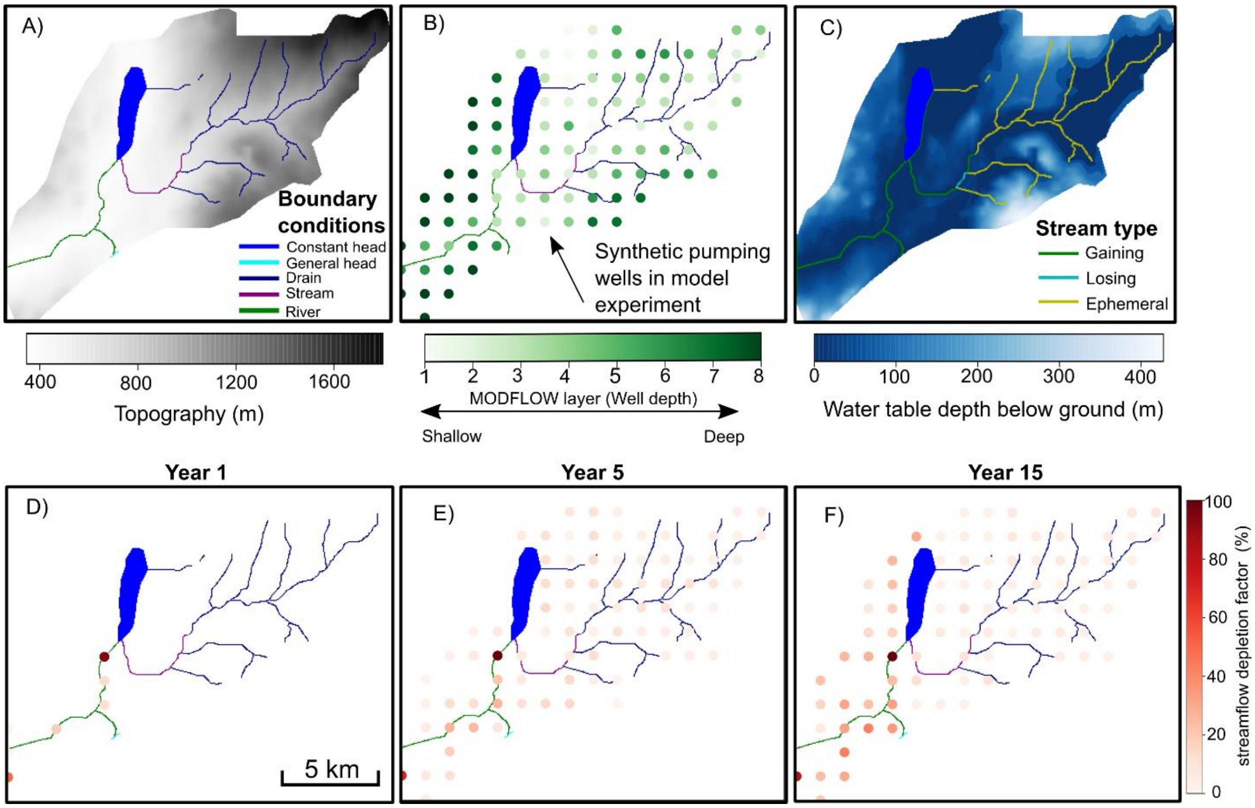


Figure 10: A) Boundary conditions and topography in the interior plateau and highlands (BX Creek); B) spatial distribution of synthetic pumping wells and their depths in the numerical model; C) streamflow types, including gaining, losing, and ephemeral stream segments. D), E), and F) are cumulative streamflow 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).

Spatially, the streamflow depletion in the first year was primarily caused by wells in the lowland portions of the domain and expanded to wells in upland areas in later years. In this domain, we found that there was no depletion from stream segments or ephemeral streams in the upper domain where the MODFLOW model uses drain flow boundary to represent surface water (Figure 10A). Some wells located in upland areas caused depletion in later years, which was sourced from nearby down-gradient features such as the constant head (lake) and rivers (Figure 10C). Similarly, there was no depletion detected from the general head boundary. In the BX Creek MODFLOW, the general head boundary was designed to ensure that the downstream segments are not dry over the simulation. In MODFLOW, general head boundaries simulate fluxes as proportional to the difference in hydraulic heads between the stream and the aquifer. As shown in Figure 10C depicting model results over the pumping period, the flux exchange between the aquifer and general head boundary are small or even nearly zero because the difference of hydraulic head between the general head boundary and the surrounding areas are minor (Figure 10C). Overall, simulation in the BX domain reveals that the boundary conditions in numerical models can affect the assessment of streamflow depletion and should be carefully designed when investigating the impacts of flow responses to pumping. Therefore, future studies are needed to examine the best performed boundary conditions for streamflow depletion assessment.

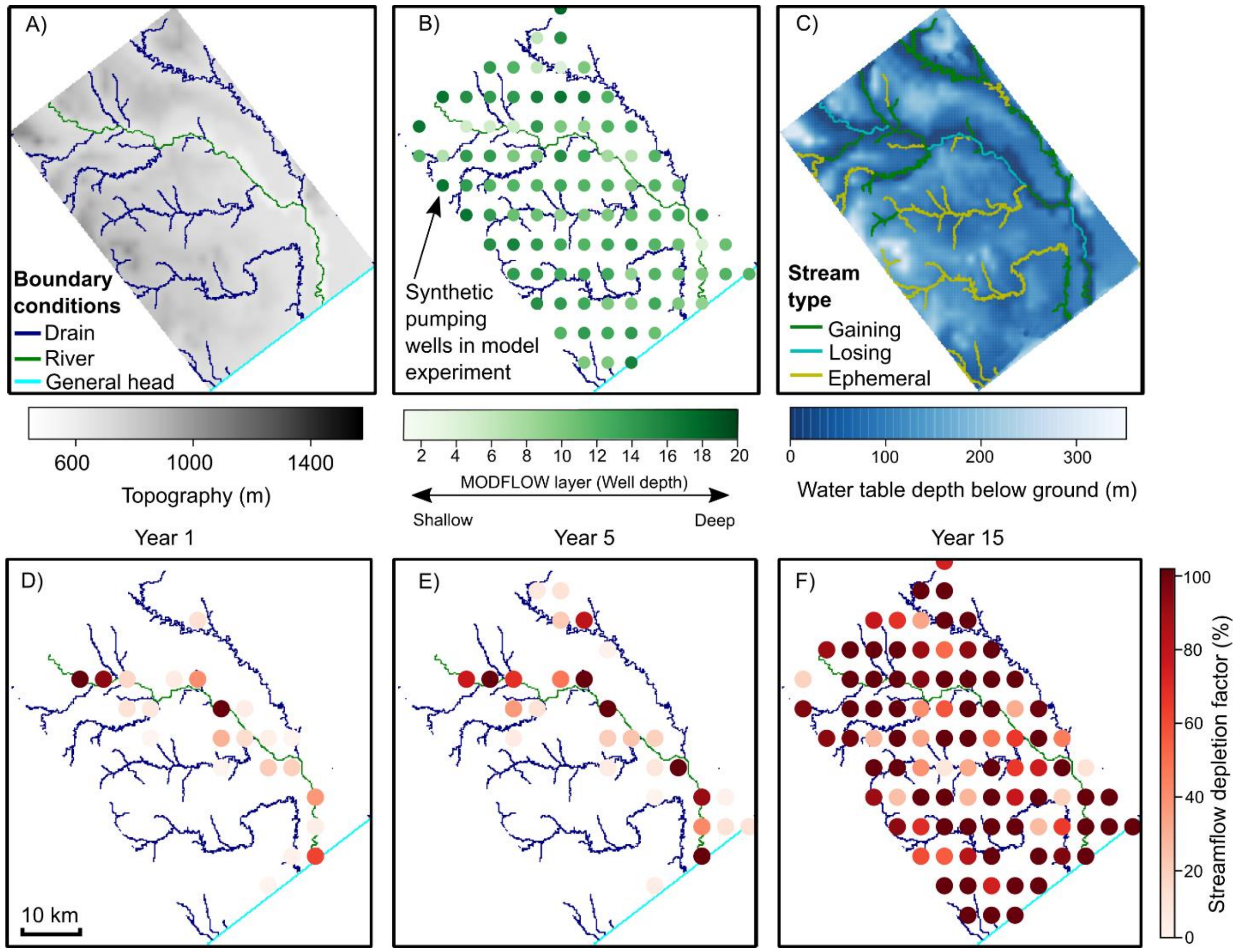


Figure 11: A) Boundary conditions and topography in the boreal plain (Peace region); B) spatial distribution of synthetic pumping wells and their depths in the MODFLOW; C) streamflow types, including gaining, losing, and ephemeral stream segments. D), E), and F) are annual streamflow depletion factor (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).

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, 52% of synthetic groundwater wells caused detectable streamflow depletion within the 30-year pumping experiment (Figure 11 and Figure B2). For instance, in the first eight years of pumping, less than 32% of groundwater wells led to streamflow depletion, but beginning in the ninth year more than 85% of wells caused detectable streamflow depletion, and by the final year all pumping wells affected at least one stream in the domain. Similar to the BX Creek model, the streamflow depletion factor ranged from 0 to 100% and the number of wells causing streamflow depletion increased with longer pumping time. Conversely, the groundwater pumping had larger impacts on the streams in the Peace region (larger streamflow depletion factor in Figure 11F compared to Figure 10F), indicating that the same pumping schedule in two different regions with different hydrogeological settings stream networks can have differing impacts on streams, which further highlights that streamflow depletion assessment should consider site-specific factors.

In summary, the two domains consistently showed that groundwater pumping can have significant impacts on streamflow under realistic pumping scenarios for the region. Therefore, the potential negative impacts of pumping on streamflow in these regions should be considered when making water management decisions. However, the two domains have different responses to pumping wells due to differences in the hydrogeological settings and stream networks. Also, we found that boundary conditions in numerical models should be carefully selected when investigating the impacts of flow responses to pumping as the response to pumping varies across boundary condition types. Since dry stream channels disconnected from the water table do not experience streamflow depletion, they should be identified and addressed prior to streamflow depletion estimation.

3.2 Streamflow depletion calculated by analytical depletion functions

We calculated depletion using two different analytical depletion functions in the two domains for comparison against the MODFLOW results (Figures A1-A2 and B1-B2). Over the entire 30-year simulation period, the average normalized MAE for analytical depletion functions including the Glover model and Hunt model were 5.0% and 5.1% for the BX Creek and 2.3% and 2.2% for the Peace region, respectively, compared to MODFLOW (Figure 12). We thus conclude that the performance of the analytical depletion functions including the Glover model and the Hunt model was similar, which is consistent with results from Zipper et al. (2019A). The similarity between analytical depletion functions incorporating different analytical models indicates that streambed conductance is not an important driver of streamflow depletion dynamics in the BX Creek and Peace region numerical models. In BX Creek, analytical depletion functions correctly identified the most affected stream segments for all wells over the entire 30-year simulation. The highest MAE of the most affected segments was 55 % of the highest pumping rate with an average of 14.4% for all time steps. For all affected stream segments, the largest MAE was 35.8% with an average of 5.0% of the highest pumping rate. In the Peace region, which has a larger stream network than BX Creek, the analytical depletion functions correctly identified the most affected stream segment for >40% of wells in the first year of pumping, and accuracy increased up to 83% by the twelfth year of pumping as the number of wells causing detectable streamflow depletion increased. The MAE for the most affected segment ranged from 0.04% to 14.8% with an average of 7.6% of the highest pumping rate throughout the pumping period. In addition, the average MAE of all affected streamflow segments was around 2.3% of the highest pumping rate (Figure 12). Overall, the difference between analytical depletion functions and MODFLOW results is relatively small.

Across both domains, we found that the normalized MAE for all affected stream segments was smaller than that for the most affected stream. This is because the most affected stream segments typically had short well-stream distances and accordingly large predicted depletions. Therefore, small errors relative to the depletion rate could still lead to a large magnitude of differences between the analytical depletion functions and MODFLOW. In contrast, for all affected stream segments, many stream segments with relatively little depletion were included, and these also tended to have smaller differences between analytical depletion functions and MODFLOW. The MAE also decreased throughout the 30-year simulation, indicating predictions of analytical depletion functions may be more accurate for longer (annual and decadal) than shorter (seasonal) impacts. In addition, the two domains consistently showed that the MAE of the most- and all-affected stream segments was significantly higher in the non-pumping season than pumping season. For instance, the average normalized MAE of the most affected stream segments for the pumping season and non-pumping seasons were 8.3% and 22.6%, respectively in the BX Creek, and 7.1% and 11.1%, respectively, in the Peace region. Similarly, the average normalized MAE of all affected stream segments for the pumping season and non-pumping seasons are 4.0% and 4.8%, respectively in the BX Creek, and 2.2% and 2.5% respectively in the Peace region. This implies that analytical depletion functions are a useful tool to estimate streamflow depletion for intermittently pumped wells since analytical depletion functions have a better performance during the pumping season when impacts tend to be larger.

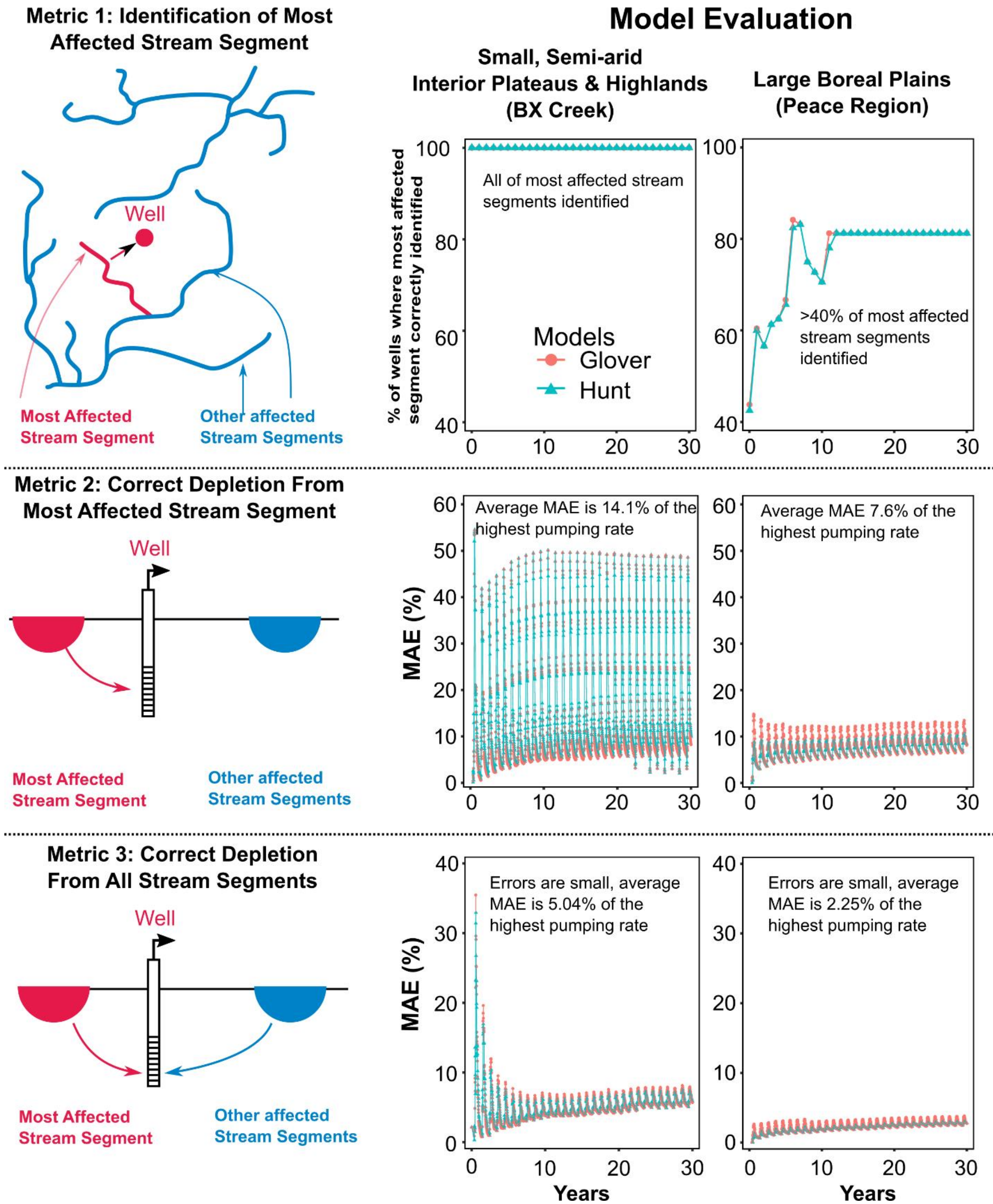


Figure 12: Performance of the analytical depletion functions in BX Creek and Peace region for metric 1 (identification of most affected stream segments), metric 2 (correct depletion from most affected stream segment), and metric 3 (correct depletion from all stream segments).

Our results showed that analytical depletion functions can correctly identify the most affected stream segment for 100% of wells in a simple stream network (BX Creek) and as much as 83% of the time in a more complex stream network (Peace region). Similarly, Zipper et al. (2019A) showed the analytical depletion functions can correctly identify the most affected segment ~85% for a mountainous domain in California, USA. Therefore, comparing across the three domains, we conclude that analytical depletion functions can accurately identify the most affected stream segments in many real-world settings. In addition, the MAE between the analytical depletion functions and MODFLOW was small, the highest MAE in the two domains was <15% of the highest pumping rate. Similarly, in California Zipper et al. (2019A) reported that MAE was <20% of the range in observed streamflow depletion for the most affected stream segments. In summary, these results suggest the analytical depletion functions are most effective at estimating streamflow depletion over the annual to decadal timescales but are less well-suited for sub-annual impacts. However, more case studies are needed to evaluate the application of analytical depletion functions in additional real-world settings.

3.3 Analytical depletion functions sensitivity to hydrogeological setting

To evaluate where the analytical depletion functions are most likely to estimate streamflow depletion accurately in each domain, we explored the response of analytical depletion functions to four model parameters: hydraulic conductivity, streambed conductance, well-stream distance, and pumping depth in both domains. In BX Creek, we found that analytical depletion functions performed better in materials with higher hydraulic conductivity, lower streambed conductance, and wells within the top five layers (Figure 13). Interestingly, we found opposite responses of analytical depletion functions to several characteristics in the Peace region (Figure 14). Specifically, analytical depletion functions performed better in lower conductivity sediments and wells in the deeper layers, while no clear response to streambed conductance was observed. The possible reasons are that 1) the parameters (*T* and *S*) used in the analytical depletion functions 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, 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 (Figure 7 and Figure 8). The parameters at the well location thus do not represent the average condition between well and streams, and this mismatch between the hydrostratigraphy at the well and the hydrostratigraphy between well and stream may have caused a different response between two domains. 2) Although we implemented a similar number of wells in the two domains, the BX Creek has larger well density than that of Peace region. Given larger domain size, well density, and cell size, groundwater pumping from a large number of wells was from groundwater storage rather than streamflow depletion. Therefore, these may have caused a different response between the two domains.

In addition, one of the key input parameters in analytical depletion functions is the well-stream distance, which is a critically important variable for a water allocation officer. In both domains, there were consistently greater differences between the analytical depletion functions and MODFLOW for wells within kilometers of a stream, which correspond to wells with the highest predicted depletion. As well-stream distances increased, the predicted depletion decreases and thus leads to a smaller MAE. The analytical depletion functions' sensitivity in the BX Creek are consistent with Zipper et al. (2019A), who found that analytical depletion functions performed better in places which are relatively flat (where alluvial aquifers are most likely to exist), near-surface water table (shallower well depth) and within a few kilometers of the downgradient perennial streams. While we found similar results in BX Creek, our comparison in the Peace region shows that drivers of analytical depletion functions performance variability differ across two domains and streamflow depletion response to hydrogeological characteristics is most likely to be region-specific. As such, conclusions related to specific parameters observed in one domain cannot be assumed to apply in other regions.

Across the two domains, we found that analytical depletion functions including the Glover and Hunt models produced similar streamflow depletion estimates (Figures A1-A2 and B1-B2) and responded similarly to hydrogeological characteristics (Figure 13 and Figure 14). The MAE for the most affected stream for analytical depletion functions including the Glover model and Hunt model were 13.6% and 14.4% respectively, in the BX Creek; and 7.9% and 7.3%, 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., 1995; Spalding and Khaleel, 1991). Theoretically, as the streambed conductance decreases, greater differences between analytical depletion functions including the Glover and Hunt models would be expected. In the two domains, we found that the difference between analytical depletion functions including the Glover and Hunt 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 and Hunt.

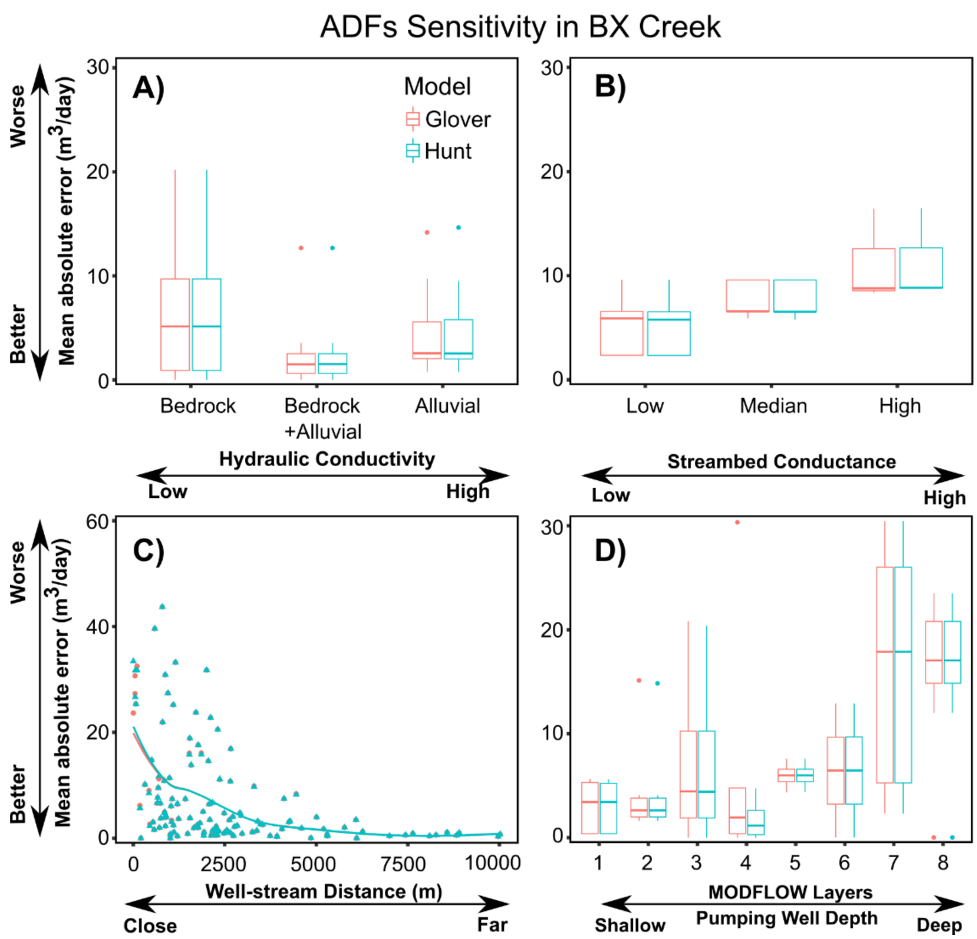


Figure 13: Analytical depletion functions sensitivity for BX Creek (MAE of streamflow depletion between the analytical depletion functions and MODFLOW) over hydraulic conductivity, streambed conductance, well-stream distance and pumping well depth in the small-arid interior plateau and highlands.

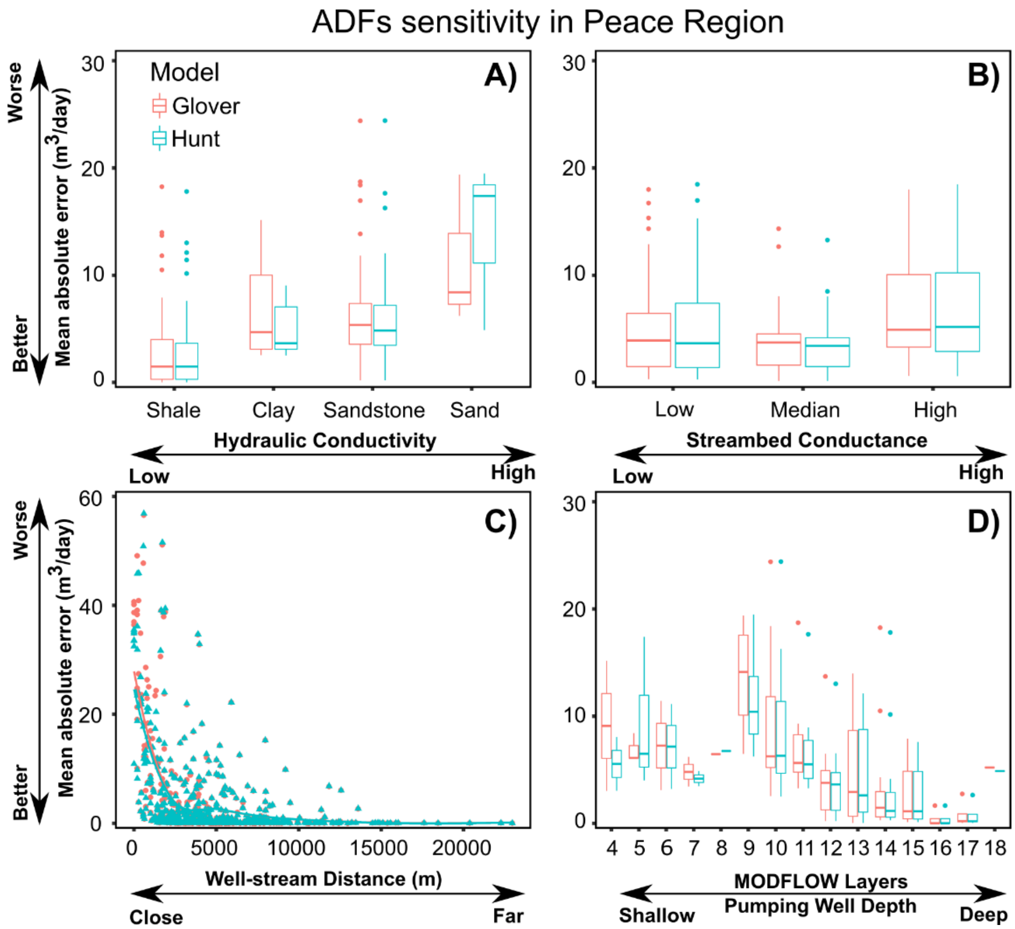


Figure 14: Analytical depletion functions sensitivity for the Peace Region (MAE of streamflow depletion between the analytical depletion functions and MODFLOW) over hydraulic conductivity, streambed conductance, well-stream distance and pumping well depth in the boreal plain.

3.4 Analytical depletion functions performance in estimating response time to environmental flow thresholds

In the two domains, analytical depletion functions and MODFLOW calculated the response time for streamflow depletion to equal 1%, 5%, and 10% of natural baseflow, respectively, for the most affected streams identified by analytical depletion functions; longer pumping times were required to reach higher thresholds (Figure 15). We found that the realistic pumping rates used in our study domains can lead to streamflow depletion beyond the presumptive threshold (10% reduction of baseflow). Both study domains also showed that response times of affected streams were variable as the magnitudes of streamflow depletion and baseflow of affected streams were different. Equivalent volumes of streamflow depletion can have different response times because natural baseflow rates vary among stream segments . Therefore, our study suggests that groundwater pumping in study domains should be managed to minimize the potential negative impacts on environmental flows.

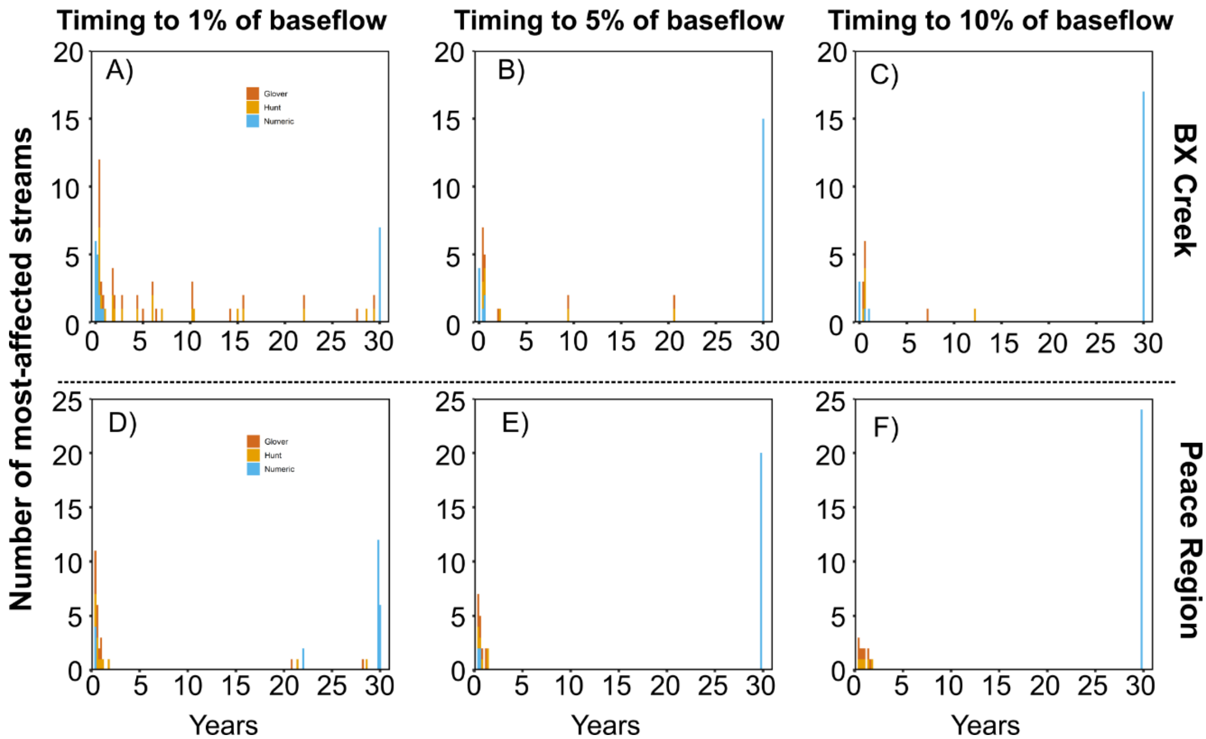


Figure 15: Stacked histograms of response time (years) of environmental flow thresholds of 1%, 5%, and 10% decrease in baseflow in the BX Creek (A, B, and C) and Peace region (D, E, and F), respectively. The bins of each bar is for 0.2 years.

The analytical depletion functions using the Glover model and Hunt model had similar response times in two domains. We, therefore, averaged response times predicted by analytical depletion functions with Glover model and Hunt model as a whole to show overall performance of analytical depeletion functions. We then compared the difference between the analytical depletion functions and MODFLOW. As mentioned in Section 2.5, the positive values of difference imply that MODFLOW predicts a shorter response time, while negative values show that analytical depletion functions predict a shorter response time. As a result, in the BX Creek, response time simulated by analytical depletion functions were both shorter and longer than MODFLOW for the most affected streams of each threshold (Figure 16A). The mean differences in response time of streamflow depletion at 1%, 5%, and 10% of baseflow were -2.1, -4.06, and +1.78 years, respectively. In the Peace region, the mean differences in response times of streamflow depletion for 1%, 5%, and 10% of baseflow were -19.5, +9.6, and -14.5 years, respectively. The inconsistent trends between the thresholds were mainly due to the setting of the longest response time being 30 years. The different performance of the response time in the two domains may be due to the difference in hydrogeological settings and stream networks. Overall, analytical depletion functions in the BX Creek had smaller differences in response time, and generally negative differences in the Peace region indicated that analytical response time estimates are conservative relative to MODFLOW, despite the fact that the differences were big. Our results were consistent with Rathfelder (2016), who compared the several analytical depletion models against a groundwater model built for Grand Forks, BC, suggesting that analytical models predicted a quicker streamflow depletion response to pumping than an existing MODFLOW model. In summary, comparison across the two domains showed that analytical depletion functions can be applied to estimate the response time, but the potential errors of analytical depletion functions and local hydrogeological conditions should be considered.

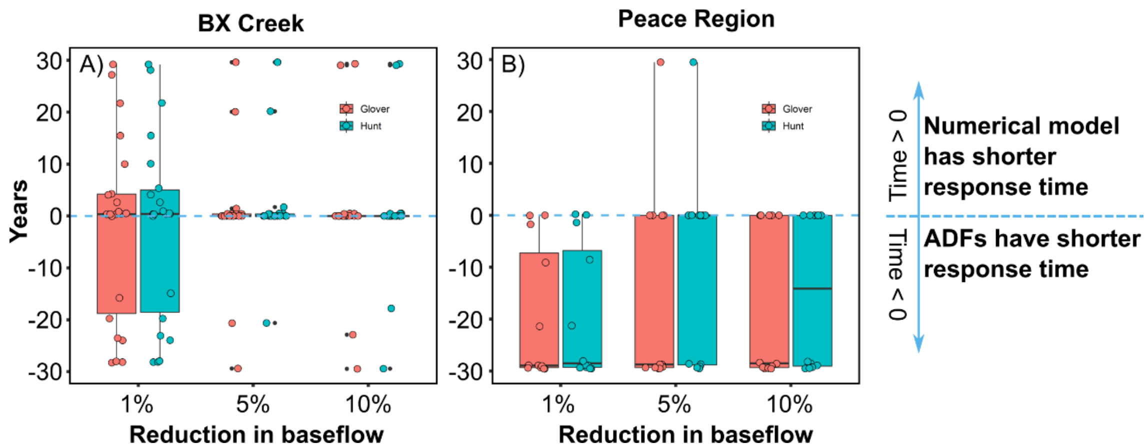


Figure 16: Difference between analytical depletion functions with Glover model and Hunt model and numerical model (i.e., response time of ADF - Numeric) of environmental flow thresholds of 1%, 5%, and 10% decrease in baseflow in the BX Creek (A) and Peace region (B). The average time difference of 1%, 5%, and 10% baseflow reduction were -2.1, -4.06, and +1.78 years in BX Creek, and -19.5, +9.6, and -14.5 years in the Peace region, respectively.

Previous studies showed that analytical depletion functions are a rapid and accurate method in quantifying magnitudes of streamflow depletion caused by groundwater pumping wells (Zipper et al., 2018; 2019A). Yet, the application of analytical depletion functions for estimating the response time to environmental flow thresholds has not been examined previously. In addition, the EFN policy highlighted the importance of both quantity and timing of streamflow needed for the proper functioning of the aquatic ecosystems. In this study, we used a presumptive 10% standard (Gleeson and Richter, 2018), which not only considers baseflow, but also can be used to assess the response time spatially for stream segments to groundwater pumping. We found that groundwater pumping had significant impacts on the presumptive environmental flow threshold in two domains. Moreover, the analytical depletion functions predicted a smaller difference in response time in BX Creek and shorter response time in the Peace region, which implies analytical depletion functions can be used to predict response time. However, given the inconsistent response time in two domains, more case studies using analytical depletion functions to estimate environmental flow needs are needed. This also highlights that response times vary with different hydrological settings and differing environmental flow needs both within and across domains, which suggest that management of the response time needs to be region- or watershed-specific.

3.5 Applicability across BC: uncertainty, limitations, and future needs for analytical depletion functions and numerical models

The results of this study support previous work indicating that analytical depletion functions are an accurate tool for estimating streamflow depletion (Huggins et al., 2018; Zipper et al., 2018; 2019A). However, the results also showed that streamflow depletion responds differently to hydrogeological conditions and physiographic settings of the two domains studied (Peace region and BX Creek). Due to the different performance of analytical depletion functions in these two hydrogeological settings, our results are not sufficient to make generalized conclusions regarding the applicability of analytical depletion functions across BC or within specific hydrogeological settings. Therefore, we recommend that analytical depletion functions should be used with caution for broader application, and that further research to understand the impacts of local hydrogeological conditions (outside of the tested domains) is necessary. Our findings also demonstrated that analytical depletion functions conservatively estimate the time it takes for streamflow to drop below a presumptive environmental flow threshold as a result of groundwater pumping from a single well.

While our analysis primarily focused on the evaluation of analytical depletion functions, we also found substantial uncertainty associated with using previously calibrated numerical models for streamflow depletion assessment if the models were developed for a different purpose. When a model that was not designed for quantifying streamflow depletion and is then used for simulating the response of a stream to pumping, significant errors and misinterpretations are possible. In this study, we initially selected four pre-calibrated numerical models in four different hydrostratigraphic regions in BC. Besides the two domains studied in this report, we also tested the performance of analytical depletion functions in the Abbotsford-Sumas area (coastal basin and lowlands; Allen et al., 2008) and the Bevan Wellfield area (coastal basin characterized with semi-confined aquifers; Piteau Associates, 2010). Unfortunately, the results for these areas are not included in this report because these models were not suitable for this type of study. For example, the Abbotsford-Sumas model was designed to assess regional groundwater sustainability under future climate change. In this model, mass balance changes in the aquifer are much higher than the designed pumping rate for a single well in our experiments due to the large domain and highly permeable aquifer, and therefore the impacts of a single well cannot be reliably separated from potential model error. The Bevan wellfield model was developed to assess the potential impacts of the Bevan well field operation on groundwater and nearby surface water features. However, boundary conditions in the model are not appropriate for simulating streamflow depletion. Even in the BX Creek and Peace region models, some of the boundary conditions were challenging to compare to the analytical depletion functions. Notably, the drain boundary can represent different stream types, i.e., ephemeral streams for BX Creek and losing and ephemeral streams for the Peace region. Groundwater pumping has limited impacts on ephemeral streams in the numerical models since they are typically disconnected from the water table. Therefore, the selection of boundary conditions could potentially affect the stream depletion results in the numerical models and thus influence our comparison between the analytical depletion functions and numerical model. In summary, in order to provide reliable streamflow depletion assessment, numerical models should be specifically developed and calibrated for surface and groundwater interactions.

Based on these results, additional study is needed to advance understanding of streamflow depletion and guide the application of analytical depletion functions before the results of this study can be generalized across BC or a specific hydrogeological condition. Here, we highlight some possible directions for future research studies.

1) **Cumulative impacts of multiple pumping wells:** We compared the performance of analytical depletion functions to numerical models by turning one pumping well on at a time. However, multiple groundwater wells typically exist in a watershed. Future studies are needed to examine whether analytical depletion functions are an appropriate tool for estimating streamflow 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) **Streamflow recovery after pumping:** We focused on the response time to environmental flow thresholds due to groundwater pumping. However, the recovery time to environmental flow threshold after cessation of pumping (Gleeson and Richter, 2018) is also of interest to water managers. It is unclear, thus far, whether analytical depletion functions are suitable to examine post-pumping recovery.
3) **Clarifying appropriate numerical models for comparison**: We assessed the performance of analytical depletion functions by comparing them to previously calibrated numerical models, so

the degree to which analytical depletion functions results can be considered representative of real-world conditions depends on the representative of the numerical models. The selection of boundary conditions affects numerical model results and thus increase uncertainty in the degree to which our analytical depletion predictions would match real-world conditions.

4. CONCLUSION

In this study, we evaluated the performance of analytical depletion functions, which include depletion apportionment equations, stream proximity criteria, and analytical depletion models to understand the utility of analytical depletion functions in two different hydrogeological settings within British Columbia: interior plateau and highlands (using BX Creek as a case study) and boreal plains (using Peace region). BX Creek has a simpler hydrogeological setting and stream network, while the Peace region is larger and more complex. Such comparison leads to a more comprehensive understanding of the utility of analytical depletion functions in varying hydrogeologic settings.

Using MODFLOW simulations, we found that groundwater pumping can have a significant impact on streamflow in both regions. However, the two domains have different responses to pumping wells due to differences in the 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 estimates from the analytical depletion functions was compared to the numerical models. Across the two domains, we found that analytical depletion functions correctly identified the most affected stream segment by all wells in BX Creek domain and 83% of the time in the Peace region over the entire 30-year simulation. The mean absolute errors (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 14.1 % and 17.6% of the highest pumping rate (214 m^3/day) in the BX Creek and Peace region, respectively over the 30-year simulation period. For all affected stream segments, the average MAE was 5.0% of the highest pumping rate in BX Creek and 2.3% in the Peace region. Also, the MAE decreased throughout the 30-year simulation, indicating analytical depletion functions predictions may be more accurate over longer (decadal) than shorter (seasonal) timeframes. In addition, we found that analytical depletion functions predictions were more accurate during the pumping season compared to the non-pumping season. Overall, we conclude that analytical depletion functions provide reasonable predictions to estimate streamflow depletion in the pumping season for perennial streams over yearly or longer timescales.

Moreover, analytical depletion functions can be used to assess the response time to environmental flow thresholds to pumping. The streamflow depletion based by realistic groundwater pumping surpassed the presumptive standard of environmental flow needs in both domains, indicating that groundwater pumping in study domains should be addressed in water management strategies. In BX Creek, the difference in response time of threshold in BX Creek was smaller than that in the Peace region, but results from both regions showed that the analytical depletion functions predicted shorter response times than MODFLOW. Combined, analytical depletion functions can be applied to estimate the response time and potential error resulted from the local hydrogeological setting should be considered.

We found variable drivers of the performance of analytical depletion functions across these two hydrogeologic settings. In the BX Creek model, analytical depletion functions have smaller errors for wells in higher hydraulic conductivity materials, shallower aquifers, and in areas with lower streambed conductance. Conversely, in the Peace region, analytical depletion functions have smaller errors in lower hydraulic conductivity materials and deeper aquifers, while the performance over the streambed conductance is similar. In both domains, the performance of analytical depletion functions 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 analytical depletion functions performance to hydrogeological setting between the BX Creek and Peace region stresses the importance of additional testing of analytical depletion functions in different regions in British Columbia to better identify the drivers of performance.

REFERENCES

Ahlfeld, D. P., Schneider, J. C., & Spalding, C. P. (2016). Effects of nonlinear model response on allocation of streamflow depletion: Exemplified by the case of Beaver Creek, USA. Hydrogeology Journal, 24(7), 1835 ‑ 1845. https://doi.org/10.1007/s10040-016-1438-3

Allen, D. M., Schuurman, N., Deshpande, A., & Scibek, J. (2008). Data integration and standardization in cross-border hydrogeological studies: a novel approach to hydrostratigraphic model development. Environmental geology, 53(7), 1441-1453. https://doi.org/10.1007/s00254-007-0753-3

Barlow, P. M., & Leake, S. A. (2012). Streamflow depletion by wells—Understanding and managing the effects of groundwater pumping on streamflow (No. Circular 1376). Reston VA: U.S. Geological Survey. https://pubs.usgs.gov/circ/1376/.

Caissie, D., N. El-Jabi, and C. Hebert. (2007). Comparison of hydrologically based instream flow methods using a resampling technique. *Canadian Journal of Civil Engineering*. 34(1): 66-74. https://doi.org/10.1139/l06-095.

Forstner, T., Gleeson, T., Borrett, L., Allen, D.M., Wei, M., and Baye, A. (2018). Mapping aquifer stress, groundwater recharge, groundwater use, and the contribution of groundwater to environmental flows for unconfined aquifers across British Columbia. Water Science Series, WSS2018-04. Prov. B.C., Victoria B.C

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.* 34(1): 83-92. 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, Transactions American Geophysical Union, 35*(3), 468–470. https://doi.org/10.1029/TR035i003p00468

Hatfield, T., A. Lewis, D. Ohlson, and M. Bradford. (2003). Development of instream flow thresholds as guidelines for reviewing proposed water uses. Consultants report for British Columbia Ministry of Sustainable Resource Management and British Columbia Ministry of Water, Land and Air Protection. Victoria, BC. 88p.

Huang, C. ‑ S., Yang, T., & Yeh, H. ‑ D. (2018). Review of analytical models to stream depletion induced by pumping: Guide to model selection. *Journal of Hydrology*, 561, 277 ‑ 285. https://doi.org/10.1016/j.jhydrol.2018.04.015

Huggins, X., Gleeson, T., Eckstrand, H. and Kerr, B., (2018). Streamflow Depletion Modeling: Methods for an Adaptable and Conjunctive Water Management Decision Support Tool. *JAWRA Journal of the American Water Resources Association*, 54(5), 1024-1038. https://doi.org/10.1111/1752-1688.12659

Hunt, B. (1999). Unsteady Stream Depletion from Ground Water Pumping. *Ground Water*, 37(1), 98–102. https://doi.org/10.1111/j.1745-6584.1999.tb00962.x

Kiffney, P., J. P. Bull, and M. C. Feller (2002), Climatic and hydrologic variability in a coastal watershed of southwestern British Columbia, *JAWRA Journal of the American Water Resources Association*., 38, 1437–1451, http://doi.org/10.1111/j.1752-1688.2002.tb04357.x.

Lackey, G., Neupauer, R. M., & Pitlick, J. (2015). Effects of streambed conductance on stream depletion. *Water*, 7(12), 271 - 287. https://doi.org/10.3390/w7010271

Morgan, S.E. (2018). Investigating the role of buried valley aquifer systems in the regional hydrogeology of the central Peace region in northeast British Columbia, Master's thesis. Simon Fraser University.

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

Personal Communication. (2020). Comments from Robin Pike, BC ENV.

Piteau Associates. (2010). Hydrogeological investigation for environmental assessment - Bevan road well field, Abbotsford, B.C. Project 2914, May 2010.

Province of British Columbia (2016) Environmental Flow Needs policy. https://www2.gov.bc.ca/assets/gov/environment/air-land-water/water/water-rights/efn_policy_mar-2016_signed.pdf (accessed on Jan. 30, 2020)

Ptolemy, R. and A. Lewis. 2002. Rationale for Multiple British Columbia Instream Flow Standards to Maintain Ecosystem Function and Biodiversity. Draft for Agency Review. prepared for Ministry of Water, Land and Air Protection and Ministry of Sustainable Resource Management

Rathfelder, K. M. (2016). Modelling tools for estimating effects of groundwater pumping on surface waters*. Water Science Series* No. WSS2016- 09. Province of B.C.: Ministry of Environment.

Schneider, J. C., Ahlfeld, D. P., & Spalding, C. P. (2017). Allocation of streamflow depletion impacts under nonlinear conditions. JAWRA Journal of the American Water Resources Association, 53(3), 697 - 706. https://doi.org/10.1111/1752 - 1688.12525

Smerdon, B.D., Allen, D.M., Grasby, S.E. and Berg, M.A., (2009A). An approach for predicting groundwater recharge in mountainous watersheds. *Journal of Hydrology*, *365*(3-4), 156-172. https://doi.org/10.1016/j.jhydrol.2008.11.023

Smerdon, B.D., Redding, T. and Beckers, J., (2009B). An overview of the effects of forest management on groundwater hydrology. *Journal of Ecosystems and Management*, *10*(1). Access on March 20, 2019 http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.533.4354&rep=rep1&type=pdf

Sophocleous, M., Koussis, A., Martin, J. L., & Perkins, S. P. (1995). Evaluation of simplified stream-aquifer depletion models for water rights administration. *Ground Water*, 33(4), 579-588. 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 Resources Research*, 27(4), 597–609. https://doi.org/10.1029/91WR00001

St. Jacques, J.-M., D. J. Sauchyn, and Y. Zhao (2010), Northern Rocky Mountain streamflow records: Global warming trends, human impacts or natural variability? *Geophysical Research Letters.* 37, L06407, http://doi.org/10.1029/2009GL042045

Water Sustainability Act. (2016) http://www.bclaws.ca/civix/document/id/complete/statreg/14015. Accessed on Jan. 30, 2020.

Wei, X., Li, Q., Zhang, M., Giles-Hansen, K., Liu, W., Fan, H., Wang, Y., Zhou, G., Piao, S. and Liu, S. (2018). Vegetation cover—another dominant factor in determining global water resources in forested regions. *Global Change Biology*, 24(2), pp.786-795. https://doi.org/10.1111/gcb.13983

Zipper, S.C., Dallemagne, T., Gleeson, T., Boerman, T.C. and Hartmann, A., (2018). Groundwater pumping impacts on real stream networks: Testing the performance of simple management tools. *Water Resources Research*, *54*(8), 5471-5486. https://doi.org/10.1029/2018WR022707

Zipper, S. C. (2019). streamDepletr: Estimate streamflow depletion due to groundwater pumping (Version R package version 0.1.0). Retrieved from https://CRAN.R - project.org/package=streamDepletr

Zipper, S.C., Gleeson, T., Kerr, B., Howard, J.K., Rohde, M.M., Carah, J. and Zimmerman, J., (2019A). Rapid and accurate estimates of streamflow depletion caused by groundwater pumping using analytical depletion functions. *Water Resources Research*, 55(7), pp.5807-5829. https://doi.org/10.1029/2018WR024403

Zipper, S.C., Carah, J.K., Dillis, C., Gleeson, T., Kerr, B., Rohde, M.M., Howard, J.K. and Zimmerman, J.K. (2019B). Cannabis and residential groundwater pumping impacts on streamflow and ecosystems in Northern California. *Environmental Research Communications*, 1(12), p.125005. https://doi.org/10.1088/2515-7620/ab534d

APPENDIX A: FIGURES FOR THE BX CREEK MODEL

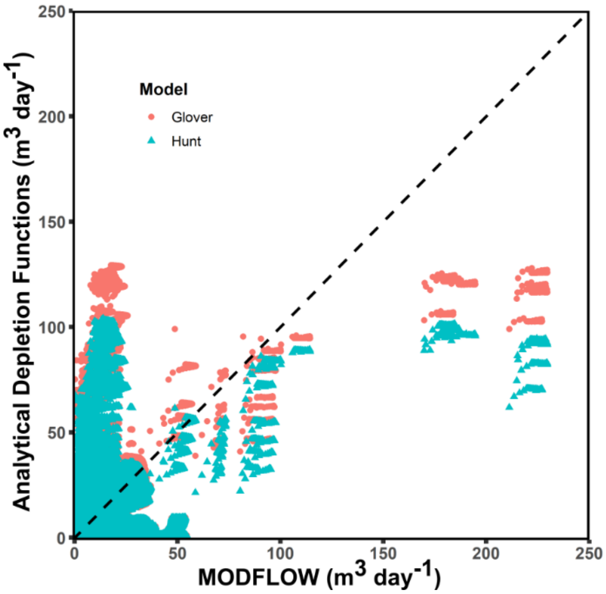


Figure A1: Streamflow depletion comparison between analytical depletion functions simulated by two analytical models and MODFLOW in the BX Creek for the simulation period of 30 years.

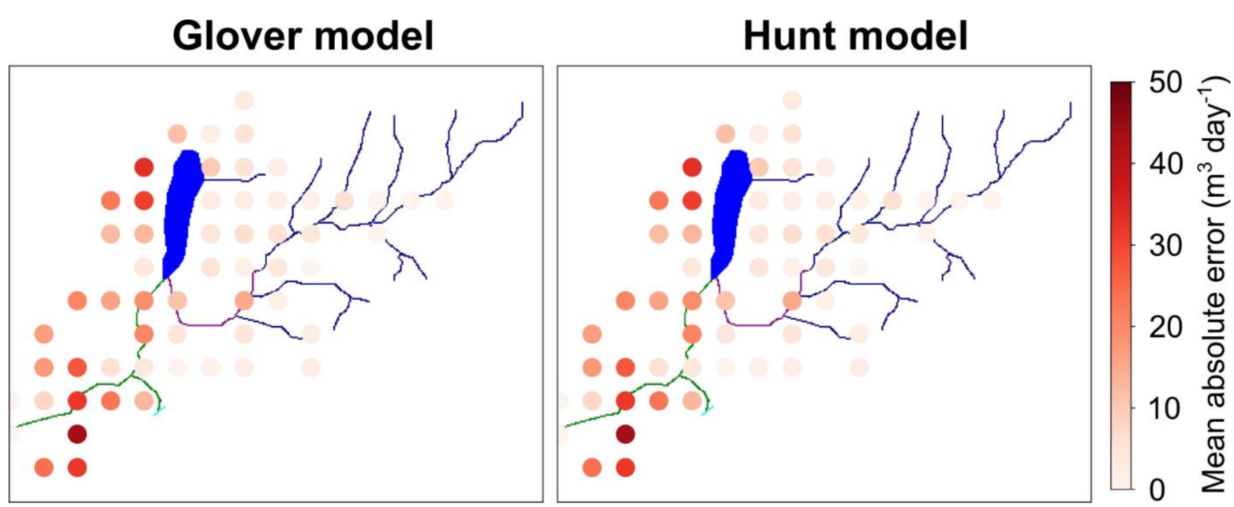


Figure A2: Spatial distribution of mean absolute error between analytical depletion functions and numerical models in the BX Creek over the simulation period of 30 years.

APPENDIX B: FIGURES FOR THE PEACE REGION

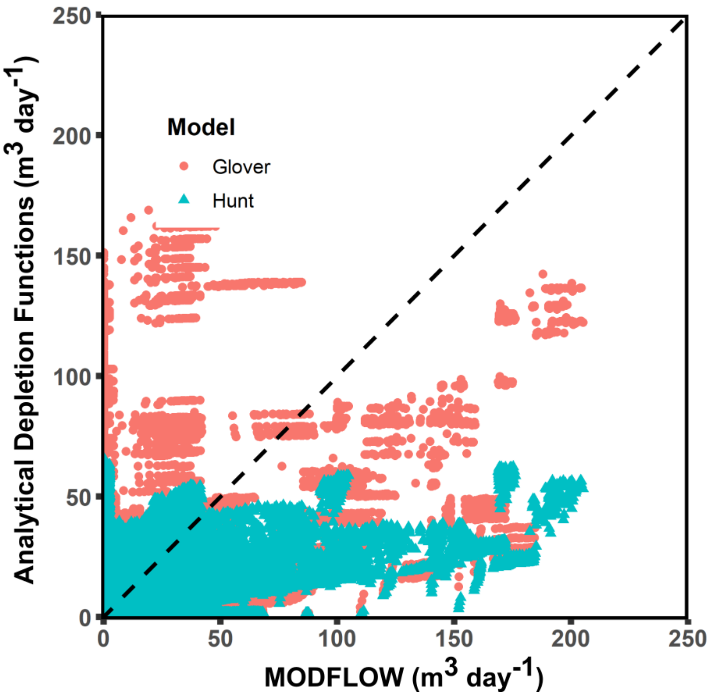


Figure B1: Streamflow depletion comparison between analytical depletion functions simulated by two analytical models and MODFLOW in the Peace region for the simulation period of 30 years.

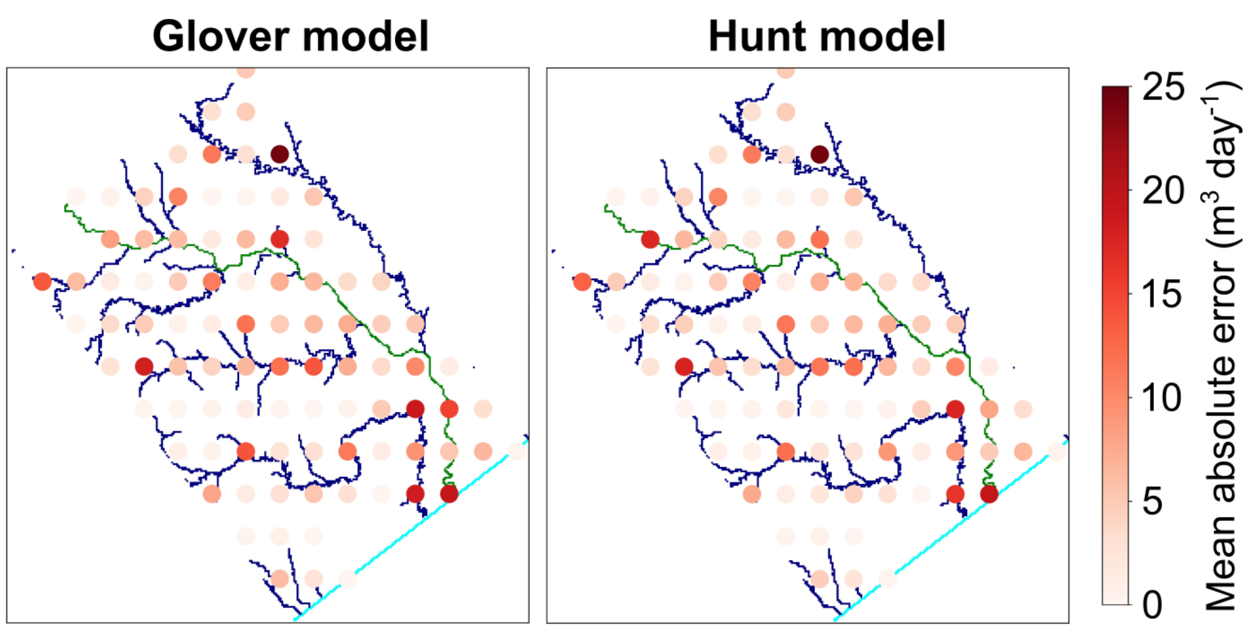


Figure B2: Spatial distribution of mean absolute error between analytical depletion functions and numerical models in the Peace region over the simulation period of 30 years.