Check for updates

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR027591

Special Section:
The Quest for Sustainability of Heavily Stressed Aquifers at Regional to Global Scales

Comparing Streamflow Depletion Estimation Approaches in a Heavily Stressed, Conjunctively Managed Aquifer

Samuel C. Zipper[1], Tom Gleeson[2,3], Qiang Li[2,3,4], and Ben Kerr[3]

[1]Kansas Geological Survey, University of Kansas, Lawrence, KS, USA, [2]Department of Civil Engineering, University of Victoria, Victoria, BC, Canada, [3]Foundry Spatial Ltd., Victoria, BC, Canada, [4]Center for Ecological Forecasting and Global Change, College of Forestry, Northwest A&F University, Yangling, China

Key Points:

- We compared streamflow depletion estimated by analytical depletion functions to a calibrated numerical model in a heavily stressed aquifer
- Analytical depletion functions had similar estimates of streamflow depletion with lower data and computational costs than numerical models
- Analytical depletion functions are a potential tool for decision making in settings where numerical models are not available

Supporting Information:

- Supporting Information S1

Correspondence to:
S. C. Zipper,
samzipper@ku.edu

Citation:
Zipper, S. C., Gleeson, T., Li, Q., & Kerr, B. (2021). Comparing streamflow depletion estimation approaches in a heavily stressed, conjunctively managed aquifer. *Water Resources Research*, *57*, e2020WR027591. https://doi.org/10.1029/2020WR027591

Received 26 MAR 2020
Accepted 25 NOV 2020

Abstract Estimating reductions in streamflow caused by groundwater pumping ("streamflow depletion") is critical for conjunctive groundwater-surface water management. Streamflow depletion can be quantified using analytical models, which have low data requirements but many simplifying assumptions, or numerical models, which represent physical processes more realistically but have high data, effort, and expertise requirements. Analytical depletion functions are a new tool that address some of the limitations of analytical models, but to date have only been evaluated in limited hydrogeological settings. Here, we compare eight different analytical depletion functions to streamflow depletion estimates from a calibrated MODFLOW numerical model used for conjunctive water management in the heavily stressed Republican River region of the High Plains Aquifer (USA). We find mostly strong agreement between the analytical depletion functions and the numerical model, though analytical depletion function estimates of depletion are lower for wells close to surface water features in high transmissivity settings. Compared to previous work, there is little variability among the eight analytical depletion functions, indicating that function formulation plays a minor role in this domain. Agreement between the modeling approaches is strongly influenced by hydrostratigraphic parameters (i.e., aquifer storage and transmissivity), suggesting accurate subsurface data are essential to estimating streamflow depletion regardless of modeling approach. Additionally, agreement between the two approaches is insensitive to pumping rate, confirming a key assumption of analytical models. Overall, analytical depletion functions provide comparable estimates of streamflow depletion to numerical models at a fraction of the time and data requirements.

Plain Language Summary Estimating the impacts of groundwater pumping on streamflow ("streamflow depletion") is challenging but essential for effectively managing water resources. In this study, we test a low-cost, low-effort approach (called an "analytical depletion function") to estimate streamflow depletion by comparing it to a more complex tool that is currently used for water management in a heavily irrigated setting in the central US. We find that there is general agreement between the analytical depletion function and the more complex approach. We also test whether analytical depletion function performance is better or worse for different conditions and find that performance is similar regardless of pumping rate but very sensitive to properties of the subsurface. Overall, our results indicate that analytical depletion functions could be useful tools for estimating streamflow depletion where more complex approaches are unavailable, but having accurate data about the subsurface is essential.

1. Introduction

Groundwater is an essential contributor to streamflow around the world (Beck et al., 2013), providing a relatively cool and stable supply of water particularly during dry periods. Groundwater inflow to streams ("baseflow") is essential for a number of aquatic and groundwater-dependent ecosystems (Larsen & Woelfle-Erskine, 2018; Rohde et al., 2017). However, groundwater abstractions can lead to reductions in streamflow ("streamflow depletion") via the capture of discharge, which includes interception of water which otherwise would have discharged into a stream or, in extreme cases, induced infiltration from a previously gaining stream (Barlow et al., 2018; Bredehoeft, 2002; Bredehoeft et al., 1982).

Streamflow depletion cannot be directly observed and is often difficult to estimate due to complex groundwater flow systems, lag times between pumping and streamflow reductions, and natural variability in

© 2020. American Geophysical Union. All Rights Reserved.

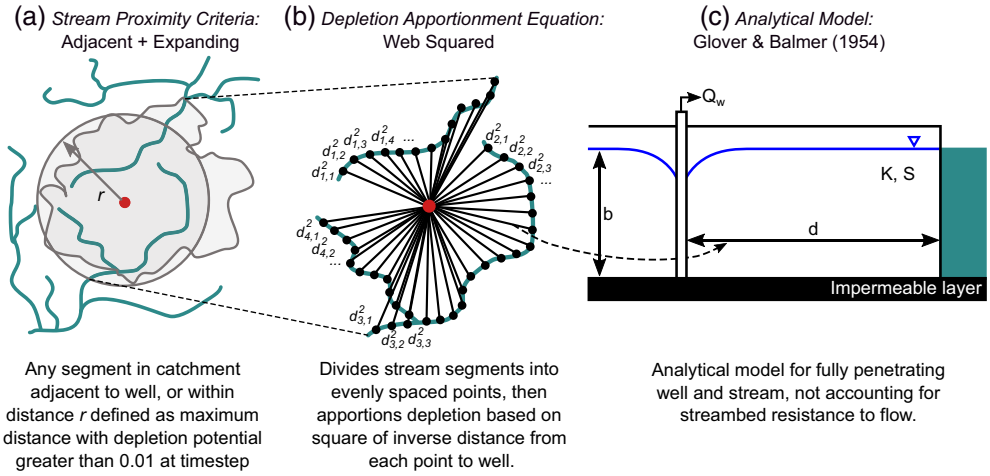

Figure 1. Components of an analytical depletion function: (a) stream proximity criteria, which identify stream segments that may be affected by a well; (b) a depletion apportionment equation, which divides streamflow depletion caused by the well among affected segments; (c) an analytical model, which calculates the depletion in each individual stream segment. Modified from Zipper, Carah, et al. (2019) under CC-BY 3.0 license.

streamflow resulting from other processes such as weather, water control structures such as dams, and surface water withdrawals (Barlow & Leake, 2012). For this reason, numerical models (e.g., MODFLOW; McDonald & Harbaugh, 1988), which are physics-based simulations of groundwater flow processes, are the typical approach for streamflow depletion and conjunctive groundwater-surface water management (Barlow & Leake, 2012; Fienen et al., 2018). However, there are large time, effort, and computational costs associated with numerical models (Fienen et al., 2015, 2016). Numerical models are, therefore, not available in most settings.

In locations where numerical models are not available, analytical models are often used instead (Hamilton & Seelbach, 2011; Huang et al., 2018). Analytical models are simpler representations of stream-aquifer interactions, but contain many limiting assumptions such as one (or occasionally two) streams, homogeneous subsurface conditions, and simplified stream and aquifer geometry. While analytical models have been proposed that account for some of these assumptions (Butler et al., 2007; S. K. Singh, 2009; Yeh et al., 2008; Zlotnik & Tartakovsky, 2008), most real-world environments violate the core assumptions of analytical models including settings where there are multiple and/or sinuous streams.

Recently, analytical depletion functions (Figure 1) have been proposed as a potential extension of existing analytical models which empirically address some of these limitations for use in complex, real-world settings (Zipper, Gleeson, et al., 2019). An analytical depletion function consists of (i) *stream proximity criteria*, which identify the streams that may be affected by a well; (ii) a *depletion apportionment equation*, which calculates how depletion from a single well should be allocated to multiple stream segments meeting the stream proximity criteria; and (iii) an *analytical model*, which estimates depletion in each stream segment meeting the proximity criteria.

To date, analytical depletion functions have only been subjected to limited testing in a small number of study domains. All of these evaluations have compared analytical depletion functions to numerical model results because it is not possible to estimate regional-scale, segment-resolution streamflow depletion using observational data (Barlow & Leake, 2012). During the development of the State of Michigan's Water Withdrawal Assessment Tool, Reeves et al. (2009) compared nine different depletion apportionment equations for a small watershed in Michigan and found that an inverse distance-based approach best matched output from a MODFLOW numerical model. Subsequently, Zipper, Dallemagne, et al. (2018) evaluated five depletion apportionment equations for several real-world stream networks in British Columbia, finding that a new inverse distance-based approach which considers stream geometry (called web squared; Figure 1b) performed the best under steady-state conditions. Zipper, Gleeson, et al. (2019) introduced the concept of stream proximity criteria (Figure 1a) and performed a sensitivity analysis to 50 combinations of

stream proximity criteria, depletion apportionment equation, and analytical model, finding that analytical depletion functions were able to accurately estimate the distribution and magnitude of streams affected by a well. However, Zipper, Gleeson, et al. (2019) compared analytical depletion functions to an archetypal numerical model with several simplifications including a homogeneous subsurface and stream properties. Q. Li et al. (2020) conducted the first comparison of analytical depletion functions to a calibrated numerical model for two regions in British Columbia, but this was in unstressed conditions with only a single well pumping at any given time and did not evaluate different analytical depletion function formulations.

As a result, it remains unknown whether analytical depletion functions are suitable tools for heavily stressed aquifers with significant ongoing pumping, particularly since evidence indicates that cumulative impacts of multiple wells may not be linearly additive (Ahlfeld et al., 2016). Furthermore, the degree to which well and hydrostratigraphic characteristics influence the performance of analytical depletion functions has not been previously evaluated. Thus, the ability of analytical depletion functions to predict streamflow depletion in complex, heterogeneous, and highly stressed real-world settings where cumulative impacts of multiple wells are occurring simultaneously remains unknown. In this study, we address this knowledge gap by comparing a suite of analytical depletion functions to a calibrated groundwater model of the Republican River Basin (USA) which is currently used for conjunctive water management (RRCA, 2003). Specifically, we ask:

(1) Do analytical depletion functions estimates of streamflow depletion in a complex, highly stressed, unconfined aquifer agree with an existing calibrated numerical model?
(2) How does agreement between analytical depletion functions and the numerical model vary as a function of analytical depletion function formulation, hydrogeological properties, stream properties, landscape attributes, and time of year?

Answering these questions will enable further use and development of analytical depletion functions as a real-world water management tool.

2. Methods

This study focused on a comparison between the analytical depletion functions and a calibrated numerical model because it is not possible to obtain regional-scale, segment-resolution streamflow depletion estimates from observational data (Barlow & Leake, 2012). At the scale of an individual stream segment, analytical approaches can be evaluated via controlled field experiments (Hunt et al., 2001; Kollet & Zlotnik, 2003). However, at large spatial scales, streamflow depletion is obscured by variability in weather, lag times between pumping and depletion, and other management activities such as surface water withdrawals (Gleeson & Richter, 2018), and therefore calibrated numerical models are the preferred approach to quantify streamflow depletion (Barlow & Leake, 2012).

2.1. Republican River Compact Administration Groundwater Model

The Republican River Watershed drains approximately 64,500 km^2 (24,900 mi^2) of the US High Plains, flowing through Colorado, Nebraska, and Kansas (RRCA, 2003). There is significant irrigated agriculture within the watershed (Deines et al., 2017, 2019) and, as a result, the surface and groundwater resources in the Republican River Watershed and surrounding High Plains Aquifer are heavily stressed (Butler et al., 2018; Haacker et al., 2016). To allocate and manage the water of the Republican River, the three states entered into the Republican River Compact in 1942 (Khan & Brown, 2019). Following a US Supreme Court decision in 2002, representatives from Kansas, Nebraska, Colorado, the US Bureau of Reclamation, and the US Geological Survey collaboratively constructed the Republican River Compact Administration (RRCA) groundwater model as a tool for quantifying streamflow depletion caused by groundwater pumping. The model is updated annually to guide water allocations among the three states. The model is described in detail in RRCA (2003), and all model files are available on the RRCA website (https://www.republicanrivercompact.org/). We selected this numerical model for use in our study because it is an existing calibrated decision support tool for a complex conjunctive water management decision-making in a heavily irrigated setting and has been deemed to produce acceptable estimates of streamflow depletion by a group of state

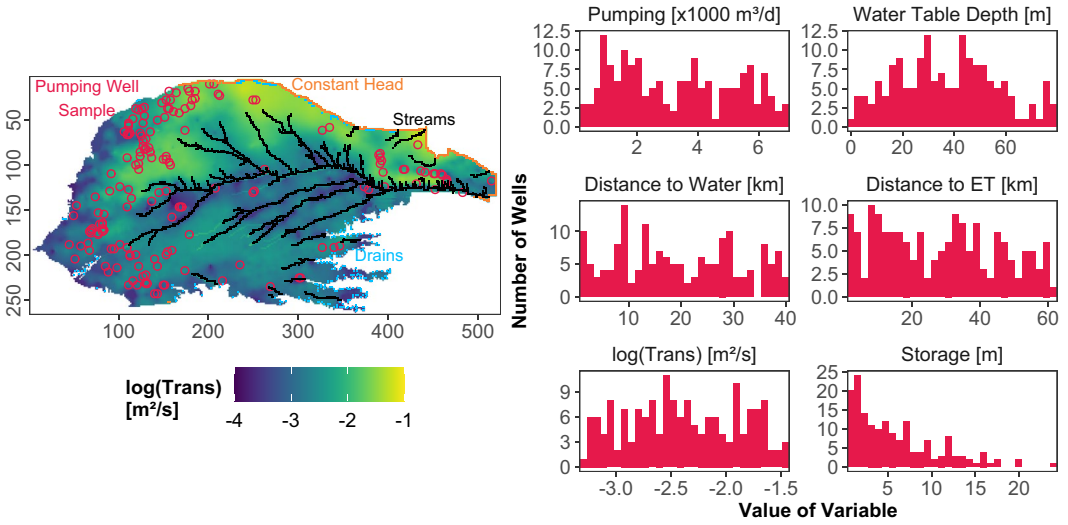

Figure 2. (Left) Map showing Republican River Compact Administration MODFLOW model domain, hydrostratigraphic properties, and selected pumping wells for pumping experiment. log(Trans) is log-10 transformed transmissivity. (Right) Properties of wells sampled for pumping experiment, which are described in detail in Section 2.3.1 of the text.

and federal scientists, engineers, and water managers (RRCA, 2003). Since the numerical model we use was calibrated to reproduce historical changes in baseflow by regional experts (described below), it represents the best available regional estimates of streamflow depletion, and has been widely used for previous scientific studies (de Graaf et al., 2019; Hrozencik et al., 2017; Hu et al., 2015, 2017; Mulligan et al., 2014; Ou et al., 2018).

The RRCA model spans an area larger than the Republican River watershed and is bounded by the Platte River in the north and outcrops of the High Plains Aquifer on the east, west, and south. The model is constructed using MODFLOW-2000 (Harbaugh et al., 2000) covering a total active domain of 77,868 km^2 (30,065 mi^2), which is discretized into 2.6 km^2 (1 mi^2) grid cells. The model is updated annually with each year's estimated pumping data submitted by each state. Here, we use version 12p, which is the version originally released that spans the period 1918–2000 at a monthly timestep. While the High Plains Aquifer in this region is unconfined, the RRCA model assumes constant transmissivity to improve model stability. To appropriately represent unconfined aquifer storage properties with a confined parameterization, the RRCA calculates specific storage for each grid cell as the estimated specific yield divided by saturated thickness in each grid cell (RRCA, 2005).

The model represents surface water features using a combination of three MODFLOW packages (Figure 2; Figure S1): the stream package (STR), which is used for the Republican River and tributaries and allows stream cells to dry in response to pumping; constant head boundaries (CHBs), which are used for the Platte River at the north edge and the eastern edge of the model; and the drain package (DRN), which represent springs and are primarily along the southeastern portion of the domain. Direct uptake of groundwater by phreatophytes is represented using the evapotranspiration (EVT) package and primarily occurs in cells along stream channels with shallow groundwater. There are a total of 9,372 pumping wells in the domain, each of which has a monthly pumping schedule, and pumping primarily occurs during the June–September growing season (Figures S2 and S3).

The model was calibrated via comparison to groundwater levels (350,233 records from 10,835 locations, with more focus on wells with a longer period of record) and baseflow (65 analyses in which baseflow was manually separated from daily streamflow using the pilot point method by the model development team) for the historical pumped period. Hydraulic conductivity and precipitation recharge rates were the primary calibration parameters. Since the calibration period included the expansion of pumping across the watershed, calibrating to baseflow across the historic period provides confidence that the model is able

to simulate the response of groundwater-surface water interactions to groundwater pumping. Complete calibration results are available at the RRCA website for baseflow (https://www.republicanrivercompact.org/v12p/html/ch07.html) and groundwater levels (https://www.republicanrivercompact.org/v12p/html/ch08.html). Since baseflow is the more relevant calibration target for our investigation, we included several calibration scatterplots in the supplemental information (Figures S4–S6). There tends to be more variability in performance in smaller tributaries compared to the Republican River main stem. Fit between observations and predictions was evaluated by experts from the model development team, which was made up of representatives from each of the affected states and the federal government, and deemed to be acceptable for water allocation decisions related to streamflow depletion based on professional judgment, though no fit statistics were presented in model documentation due to the complexity of the model domain (RRCA, 2003).

2.2. Analytical Depletion Functions

Analytical depletion functions are described in detail in Zipper, Gleeson, et al. (2019), so only a brief overview is presented here. Analytical depletion functions consist of three components (Figure 1):

(i) Stream proximity criteria, which identify the stream segments that may be depleted by a well as a subset of the total stream network
(ii) A depletion apportionment equation, which calculates the fraction (f_i) of the well's total depletion that is sourced from each stream segment (i) meeting the stream proximity criteria. For each well, f_i of all stream segments must sum to 1.0
(iii) An analytical model, which calculates the streamflow depletion for each stream segment without considering other stream segments (Qa_i). The calculation of Qa_i follows the typical use of analytical models which assume infinite stream length

For that well, the estimated volumetric streamflow depletion in each segment, Qs_i, is then calculated as, $Qs_i = f_i \times Qa_i$.

The inclusion of stream proximity criteria and depletion apportionment equations in analytical depletion functions are intended to empirically account for two major limitations of analytical models: analytical models typically only include one or a limited number of streams and do not address stream sinuosity. Depletion apportionment equations that subdivide stream segments into multiple points, such as the web and web squared approaches developed by Zipper, Dallemagne, et al. (2018), approximate the integral of streamflow depletion along a stream segment of finite length, which addresses the problematic analytical assumption of infinite stream length (Kollet et al., 2002). However, analytical depletion functions still include many of the assumptions of analytical models, one of which is that pumping does not cause any changes in capture (groundwater recharge or discharge) apart from streamflow depletion (Barlow & Leake, 2012; Bredehoeft, 2002; Bredehoeft et al., 1982).

Numerous options exist for stream proximity criteria, depletion apportionment equations, and analytical models. These components were systematically evaluated in Zipper, Gleeson, et al. (2019) via comparison to an uncalibrated numerical model of the Navarro River Watershed (California, USA). Zipper, Gleeson, et al. (2019) found that analytical depletion function performance was most sensitive to the choice of depletion apportionment equation, secondarily sensitive to the choice of stream proximity criteria, and relatively insensitive to the choice of analytical model.

In this study, we will compare a subset of 8 of the 50 analytical depletion functions evaluated by Zipper, Gleeson, et al. (2019). Since Zipper, Gleeson, et al. (2019) found the greatest sensitivity of model performance to the choice of depletion apportionment equation, we compared four unique but related depletion apportionment equations here: web and web squared, which were the two best-performing equations for the Navarro River Watershed, CA (Zipper, Gleeson, et al., 2019); and inverse distance and inverse distance squared, the former of which was the best-performing equation for the Kalamazoo Valley, MI (Reeves et al., 2009). All four depletion apportionment equations can be expressed as:

$$f_i = \frac{\sum_{p=1,P_i} \frac{1}{d_{i,p}^w}}{\sum_{j=1,n}\left(\sum_{p=1,P_j} \frac{1}{d_{j,p}^w}\right)} \quad (1)$$

The result, f_i, is the fraction of total depletion occurring in stream segment i. Required inputs include d, the distance from the well to the stream segment; w, a weighting factor that is equal to 1 for inverse distance and web and equal to 2 for inverse distance squared and web squared; n is the total number of affected stream segments identified by the stream proximity criteria (turquoise lines in Figure 1b); and P is the number of points each stream segment is divided into (black dots in Figure 1b). The inverse distance and inverse distance squared methods are simplified formulations of Equation 1 where only the closest point to the well on each stream segment is used and therefore $P = 1$.

Since stream proximity criteria were of secondary importance, we compared two stream proximity criteria: adjacent catchments only, which was used by Reeves et al. (2009), and adjacent + expanding, which Zipper, Gleeson, et al. (2019) found worked best in the Navarro River Watershed. The adjacent + expanding methods includes both adjacent catchments to the well and any catchments in which the estimated streamflow depletion would be greater than or equal to 1% of the pumping rate at a given timestep. As a result, the number of stream segments meeting the stream proximity criteria for a given pumping well increases with the time since the onset of pumping.

We tested a single, relatively simple Glover and Balmer (1954) analytical model (herein referred to as the Glover model; Equation 2):

$$Qa = Qw \times \mathrm{erfc}\left(\sqrt{\frac{Sd^2}{4Tt}}\right) \quad (2)$$

In Equation 2, Qw is the pumping rate of the well; S is the storativity, typically specific yield for unconfined aquifers; T is the transmissivity; and t is the time since pumping began. To account for monthly variation in Qw, we used superposition techniques to turn the wells on and off (Jenkins, 1968). The Glover model contains numerous simplifying assumptions which are violated in the RRCA domain, including a stream fully penetrating to the bottom of the aquifer, a single linear stream, an aquifer of infinite extent extending away from the stream, and no changes in transmissivity in response to pumping. Since Zipper, Gleeson, et al. (2019) found minimal sensitivity to the choice of analytical model and the RRCA MODFLOW model represents surface water using a mixture of packages, some of which do not include streambed conductance as an input which is required for more complex analytical models such as Hunt (1999), we did not test multiple analytical models in this study. However, many analytical models exist with different formulations (reviewed in Huang et al., 2018) and in other hydrogeological settings, different analytical models may be appropriate.

2.3. Analytical Depletion Function Performance Evaluation

2.3.1. Selecting Pumping Well Sample

The goal of our study is to examine the performance of analytical depletion functions relative to a MODFLOW model over a range of hydrogeological and physiographic characteristics. Therefore, we selected a subset of 166 wells (of 9,372 total pumping wells in the domain) based on the following characteristics: (i) the mean pumping rate, (ii) predevelopment water table depth from MODFLOW steady-state output, (iii) the transmissivity of the MODFLOW cell containing the well (log-transformed), (iv) the storativity of the MODFLOW cell containing the well, (v) the distance from the well to the closest surface water feature (active STR, DRN, or CHB grid cell), and (vi) the distance from the well to the closest cell with potential phreatophytic ET (active EVT cell). The distribution of these properties among our experimental sample is shown in Figure 2 and the distribution among all wells is shown in Figure S7. To span each of these characteristics

as uniformly as possible, we used the Latin Hypercube Sampling method (McKay et al., 1979) to randomly sample 250 parameter combinations spanning these six characteristics (Zipper, Lamontagne-Hallé, et al., 2018). We then selected the pumping well that had the shortest euclidean distance in parameter space to each parameter sample, resulting in a total of 166 unique pumping wells in our final evaluation because some wells were closest to multiple samples. Several of the selected wells are spatially close to each other because they are in locations that have relatively rare parameter conditions within the multidimensional parameter space, such as high distance to water (Figure S7), and therefore, multiple nearby wells were selected to effectively sample a wide range of parameter combinations.

2.3.2. Calculating Streamflow Depletion in MODFLOW Model

To calculate the streamflow depletion associated with each of these pumping wells in the MODFLOW model, we first ran a baseline RRCA simulation to calculate the stream-aquifer flux under baseline conditions for each cell containing a surface water feature. The cell-resolution stream-aquifer fluxes were then aggregated to net stream-aquifer fluxes for each stream segment, which were defined (for the STR package) or manually delineated based on stream network geometry (for the DRN and CHB packages). Since many of the DRN cells were discontinuous (Figure 2) and represent springs or seeps rather than stream channels, each DRN segment was delineated as a cluster of nearby DRN cells. To evaluate the sensitivity of the calculated performance metrics to the inclusion of these features, we performed analytical depletion function calculations both including and excluding DRN cells. We then turned off each of the selected wells one-at-a-time for 166 unique numerical experiments (which we refer to as "pumping simulations" herein) and calculated the net stream-aquifer flux for each stream segments in each pumping simulation. Turning off each well one-at-a-time and comparing to the baseline simulation isolates the amount of streamflow depletion caused by that specific well. Quantitatively, the decrease in stream-aquifer flux in the baseline simulation relative to a pumping simulation is equal to the streamflow depletion caused by that pumping well (and potential numerical model error), and therefore, we can calculate streamflow depletion in each segment and at each stress period for each of our 166 pumping wells. We automated MODFLOW runs using the FloPy package for Python (Bakker et al., 2016, 2018).

Since changing groundwater model boundary conditions (such as pumping rates) can impact model convergence and stability, we screened MODFLOW output for anomalous results prior to comparison. We used two approaches to screen for anomalous results in all MODFLOW simulations, each of which corresponds to a single pumping well. First, we identified additive outliers in the timeseries of the difference in overall mass balance error between the pumping simulation and the baseline simulation (Chen & Liu, 1993; López-de-Lacalle, 2019). Second, we identified any stress period where MODFLOW estimated negative streamflow depletion either in an individual segment or summed across all segments exceeding 1% of the maximum pumping rate for that well. For both of these comparisons, we identified the stress periods at which anomalous MODFLOW results occurred and limited our comparison to the time period between the onset of pumping and the stress period prior to the first anomalous MODFLOW mass balance change (Figure S8). Mass balance outliers indicating convergence issues were identified for 31 of the 166 pumping experiments tested, occurring as early as the 499th stress period and as late as the final (996th) stress period. Stress periods with anomalous results from these 31 pumping simulations were removed from analysis as described above. For all other pumping experiments, our comparison was conducted between the onset of pumping and the end of the model simulation period (December 2000).

2.3.3. Calculating Streamflow Depletion With Analytical Depletion Functions

Input for the analytical depletion functions was extracted directly from the RRCA MODFLOW model. Using consistent parameters between the two approaches was intended to focus our comparison on the impact of the simplifications of analytical depletion functions relative to the numerical model on predicted streamflow depletion, since we do not do a direct comparison to field observations of streamflow depletion. The pumping rate (Qw) was extracted from the MODFLOW WEL package (Figure S3), with each grid cell representing a unique pumping well. The well-stream geometry (d) was extracted as the distance from each well to each grid cell with a STR, CHB, and DRN cell, which were grouped into the same segments used by the MODFLOW model which are described in Section 2.3.2. Due to the discretization of the MODFLOW model, the point spacing for the discretization of segments in the web and web squared depletion apportionment equations was 2.6 km^2 (1 mi^2). Effective transmissivity (T) and storativity (S) were calculated as

the average T and S for all MODFLOW cells intersected by a straight line between the well and each stream segment. For S, we used specific yield, rather than specific storage, as input to the analytical depletion functions because the use of specific storage in the RRCA model is an artifact of model design, as described in Section 2.1. The output from the analytical depletion functions is the estimated streamflow depletion in each stream segment caused by each well throughout the entire RRCA simulation period.

2.3.4. Comparison Between Analytical Depletion Functions and MODFLOW

Due to the lack of segment-resolution, regional-scale streamflow depletion estimates, the calibrated MODFLOW numerical model is the best available source of streamflow depletion estimates in the Republican River basin. Accordingly, we interpret similar estimates by the two approaches as evidence that the analytical depletion functions are also able to accurately quantify streamflow depletion. However, we acknowledge the potential for error in both the RRCA MODFLOW model and analytical depletion functions, and therefore, our study focuses on agreement and differences between the two approaches.

To systematically assess different aspects of analytical depletion function performance, we calculated four fit metrics which are described below. To assess the drivers of analytical depletion function performance, we conducted a regional sensitivity analysis for each of the fit metrics in response to well and landscape characteristics (Figure 2; Section 2.3.1). The regional sensitivity analysis is meant to identify conditions under which the performance of analytical depletion functions exceeds a defined performance threshold (Pianosi et al., 2016; Spear & Hornberger, 1980; Wagener et al., 2001), which we set separately for each of the four fit metrics:

1. *Spatial distribution of primary impact:* The percentage of pumping simulations in which the stream segment most affected by groundwater pumping matched between the analytical depletion functions and MODFLOW model. For regional sensitivity analysis, a threshold value of 50% was used to separate good (match > 50%) from poor (match < 50%) performance
2. *Magnitude of primary impact:* The mean absolute difference (MAD) between the volumetric depletion (Qs) in each stream segment predicted by the analytical depletion function and MODFLOW model, normalized to the range of predicted Qs values from MODFLOW. The normalization allows for comparison across a range of depletion conditions—for instance, a difference in predicted Qs of 100 $m^3\ d^{-1}$ is more problematic when actual Qs is 200 $m^3\ d^{-1}$ than when actual Qs is 5,000 $m^3\ d^{-1}$. For regional sensitivity analysis, a threshold value of 0.25 was used to separate good (normalized MAD ≤ 0.25) from poor (normalized MAD > 0.25) performance. Note that we use the term MAD, rather than Mean Absolute Error because differences between the MODFLOW model and analytical depletion functions may be caused by errors in either of the two approaches
3. *Spatial distribution of overall impacts:* The Kling-Gupta Efficiency (KGE) between the volumetric depletion (Qs) in each stream segment predicted by the analytical depletion function and MODFLOW model. KGE is a hydrological performance indicator which accounts for differences in correlation, variability, and bias (Gupta et al., 2009; Kling et al., 2012). Similar to the Nash-Sutcliffe Efficiency, a KGE value of 1.0 indicates perfect agreement and lower values indicate worse agreement. As a benchmark, KGE > −0.41 indicates better agreement than simply using the mean (Knoben et al., 2019). For regional sensitivity analysis, a threshold value of −0.41 was used to separate good (KGE ≥ −0.41) from poor (KGE < −0.41) performance
4. *Magnitude of overall impacts:* The bias between the total streamflow depletion across all segments for an individual well simulated by the analytical depletion function and the MODFLOW model, normalized to the range of predicted total streamflow depletion from MODFLOW. For regional sensitivity analysis, a threshold value of 75% was used to separate good (absolute bias ≤ 75%) from poor (absolute bias > 75%) performance

3. Results and Discussion

3.1. Overall Performance

Overall, we find a strong agreement between the analytical depletion functions and MODFLOW predictions of streamflow depletion. Variability among analytical depletion function formulations is explored in the following section. The best-performing analytical depletion function combined the adjacent + expanding

Table 1
Analytical Depletion Function Performance for Different Fit Metrics, Calculated for Each Month of Simulations and Averaged Over the Final 20 Years of the Simulations.

Stream proximity criteria	Depletion apportionment equation	Spatial distribution of primary impact (% most-affected correct)	Magnitude of primary impact (MAD segment streamflow depletion, normalized)	Spatial distribution of overall impacts (KGE, segment streamflow depletion)	Magnitude of overall impacts (% bias, total streamflow depletion)
Adjacent	Inverse distance	50.7	0.056	0.701	−8.0
Adjacent	Inverse distance squared	50.8	0.054	0.695	15.1
Adjacent	Web	**52.4**	**0.045**	0.699	−20.9
Adjacent	Web squared	**53.9**	0.048	**0.767**	4.2
Adjacent + expanding	Inverse distance	50.0	0.057	0.697	−18.0
Adjacent + expanding	Inverse distance squared	50.1	0.053	0.737	9.2
Adjacent + expanding	Web	**52.5**	0.047	0.671	−26.5
Adjacent + expanding	Web squared	**53.9**	0.048	**0.779**	**0.4**
Analytical only	Analytical only	**53.9**	0.06	0.555	32.5

Note. All of the analytical depletion functions shown in this table included DRN features.

Abbreviations: KGE, Kling-Gupta Efficiency; MAD, mean absolute difference.

Bold values in each column indicate analytical depletion functions which are not significantly different from the best-performing function for that metric ($p > 0.05$ using Tukey's honest significant difference test).

stream proximity criteria, the web squared depletion apportionment equation, and the Glover analytical model, which agrees with the results from a previous comparison in the Navarro River Watershed (Zipper, Gleeson, et al., 2019).

There is strong agreement between this analytical depletion function and the MODFLOW model across our four performance criteria (Table 1). Over the final 20 years of the simulation period, the most-affected stream segment is identified correctly for 53.9% of pumping wells, the MAD of predicted depletion in the most-affected segment is 0.048 of the range in predicted depletion, the KGE of predicted depletion across all segments is 0.779, and the bias for predicted total streamflow depletion is 0.4%. Analytical depletion function performance is significantly better than using a standalone analytical model (without stream proximity criteria or depletion apportionment equations) for all metrics except the identification of the most affected stream segment (Table 1), highlighting the ability of analytical depletion functions to improve predictions for real-world settings compared to a standalone analytical model.

Identification of the most-affected segment is substantially lower than previous work, which was generally >70% correct in the Navarro River Watershed (Zipper, Gleeson, et al., 2019). Results from the standalone analytical model, which always used the stream segment closest to each well, had the same skill in identifying the most-affected segment (53.9%) which indicates that for 46.1% of wells the most-affected stream segment was not the closest stream segment to the well. This may be driven by the fact that well-stream distances and numerical model discretization are an order of magnitude larger in this domain compared to the Navarro River watershed, and therefore subsurface controls on flow such as spatial heterogeneity in T and S exert a stronger control over the distribution of pumping impacts.

Despite a negligible overall bias, this analytical depletion function tends to have a higher estimate than MODFLOW for both segment-resolution and total streamflow depletion at low magnitudes and a lower estimate at high magnitudes (Figure 3). The largest differences between MODFLOW and analytical depletion function estimates tend to be driven by a relatively small number of wells which are located near the edge of the domain, which is consistent for both segment-resolution and total streamflow depletion. The cluster of points in which the analytical depletion function produces substantially higher estimates of depletion compared to MODFLOW are associated with two wells on the southeastern margin of the domain that have several extreme characteristics relative to the overall well sample. These wells are the minimum possible distance from surface water cells given the model discretization (one grid cell, or 1.6 km) yet relatively far from EVT cells (>45 km, or 85th percentile among well sample), and also have very low transmissivity

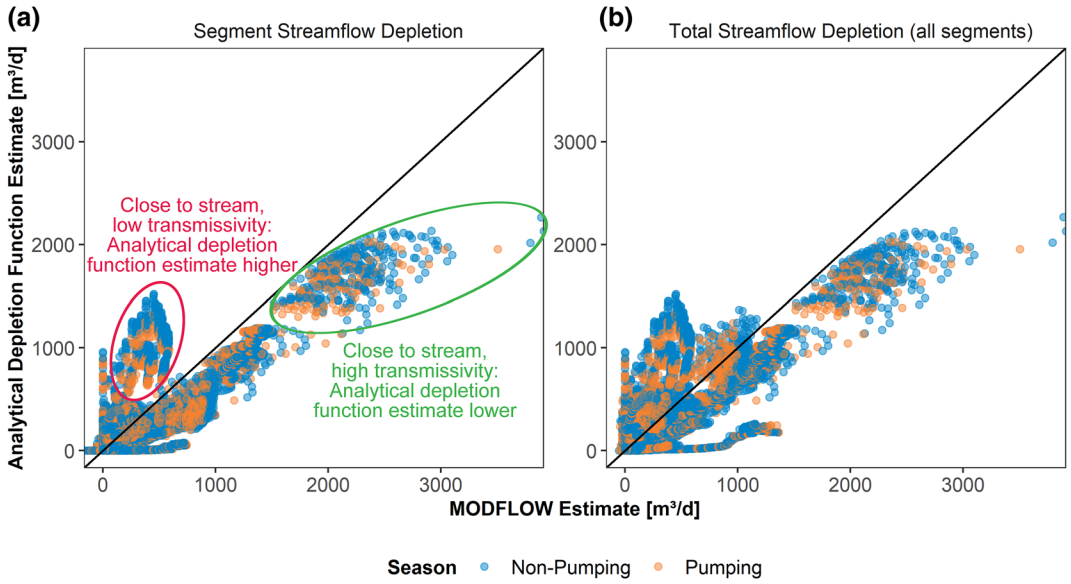

Figure 3. Comparison between analytical depletion function and MODFLOW predictions of (a) segment-scale streamflow depletion, (b) total streamflow depletion summed across all segments for a given well. Circles on depletion plot indicate "higher estimate" and "lower estimate" points discussed in Section 3.1 text.

(~50 m^2 d^{-1}, or 6th percentile). In contrast, the points in which analytical depletion functions have the largest underestimate relative to MODFLOW are associated with two wells on the northern edge of the domain that are also extremely close to a stream (two grid cells, or 3.2 km) but have a very high transmissivity (~1900 m^2 d^{-1}, or 95th percentile).

The primary differences between segment-resolution streamflow depletion (Figure 3a) and total streamflow depletion caused by a well (Figure 3b) occur at low magnitudes of depletion. As discussed above, all points with high magnitudes of depletion tend to be very close to a surface water feature of some sort and therefore depletion is dominated by a single segment. In contrast, wells causing low levels of depletion tend to be far from stream segments and therefore depletion is distributed throughout more segments, but remains low even when the depletion is added together across all segments.

While there is variability in bias among analytical depletion functions (Table 1), none of the analytical depletion function formulations predict substantially higher depletion than the best-performing analytical depletion function (Figures S9 and S10), indicating that the underestimate in depletion at high magnitudes (Figure 3) may be a persistent problem for analytical depletion functions in this setting. This finding is in contrast to the typical assumption that analytical approaches provide conservative estimates of streamflow depletion (Rathfelder, 2016; Sophocleous et al., 1995) and may be problematic from a water management perspective because underestimating the depletion from the wells with the largest impacts could lead to an overallocation of water resources (Zipper, Dallemagne, et al., 2018). As a result, analytical depletion functions should not be considered a "worst-case" estimate of depletion, but rather a minimally biased estimate which may overestimate or underestimate depletion relative to the MODFLOW model.

3.2. Performance Response to Analytical Depletion Function Formulation and Input Data

Analytical depletion formulation had relatively little impact on most of the model performance metrics. Comparing among stream proximity criteria, there is little difference between the Adjacent and Adjacent + Expanding stream proximity criteria (Table 1), likely due to the large size of the domain and relatively sparse stream network compared with Zipper, Gleeson, et al. (2019). Comparing among depletion apportionment equations, the inverse distance and inverse distance squared depletion apportionment equations do not perform the best for any of the fit metrics evaluated (Table 1) indicating that considering stream network geometry with the web and web squared improves performance, though differences

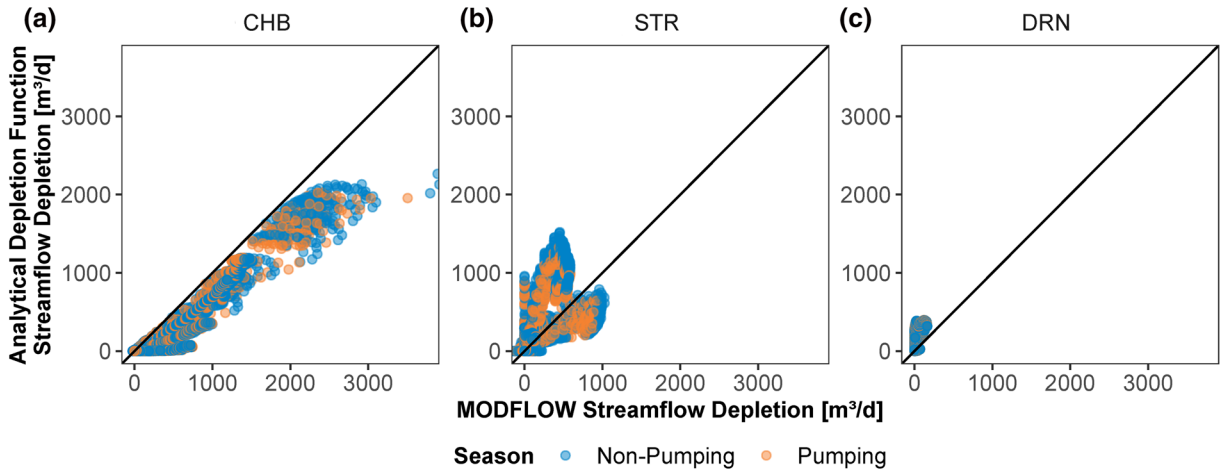

Figure 4. Comparison between analytical depletion function and MODFLOW predictions of segment-scale streamflow depletion for (a) CHB package, (b) STR, and (c) DRN in MODFLOW. CHB, constant head boundary; DRN, drain package; STR, stream package.

between approaches are only a few percentage points. Notably, the web squared depletion apportionment equation performed the best at the spatial distribution of the overall impacts (regardless of stream proximity criteria used), demonstrating its effectiveness at identifying streamflow depletion across a stream network. The superior overall performance of the web squared depletion apportionment equation relative to the web equation is due to a negative bias for wells with high amounts of total streamflow depletion (Table 1; Figure S10). This is caused by the increased weight given to near-well stream segments in the web squared approach (Zipper, Dallemagne, et al., 2018).

Both the magnitude of predicted depletion and the relationship between MODFLOW and analytical depletion functions varied as a function of the boundary condition used in the MODFLOW model. In general, the highest levels of depletion tended to be predicted for the CHB, which runs along the north side of the model domain (Figure 2a), and analytical depletion function estimates of depletion were consistently lower than MODFLOW (Figure 4a). In contrast, predicted depletion from the STR tended to be more evenly distributed with a mixture of overestimates and underestimates relative to the MODFLOW model (Figure 4b). Depletion from drain features (which are diffuse boundaries representing springs in this model; Figure 2a) was small, but analytical depletion functions had consistently higher estimates than MODFLOW (Figure 4c).

The differences in predicted depletion among these boundary conditions are likely driven by a combination of hydrostratigraphy and MODFLOW model design. First, the CHBs are found along the north side of the model domain and this region has the highest estimated transmissivity (Figure 2a) due to more conductive sediments and a greater saturated thickness (RRCA, 2003). These higher conductivity materials, along with the close proximity of some wells to the stream (discussed in Section 3.1), explains why the largest depletion estimates are found for the northern part of the domain along the CHB. Second, MODFLOW uses a streambed conductance term to simulate potential low-conductivity streambed materials for stream and drain features but not for CHBs. This streambed layer is not represented in the Glover analytical model we use in our analytical depletion functions. Use of an analytical model that includes streambed conductance, such as Hunt (1999), may improve agreement for stream and drain boundary conditions, but would cause further underestimation of depletion in CHBs.

Since analytical models are not traditionally designed for use in diffuse discharge features such as the springs represented by the DRN, we also compare analytical depletion function performance with and without drain features (Figure S11). Removing drains had relatively small impacts on model performance, particularly at high levels of depletion. Overall, removing DRNs meant that estimates of depletion in DRN segments went to 0, and as a result estimates of depletion in some other boundary conditions increased (Figure S11). Removing DRN features from the analytical depletion functions had mixed impacts on our four performance metrics. When DRNs were excluded from analytical depletion function calculations, the identification of the most affected segments degraded (42.2%, compared to 53.9% when DRNs were

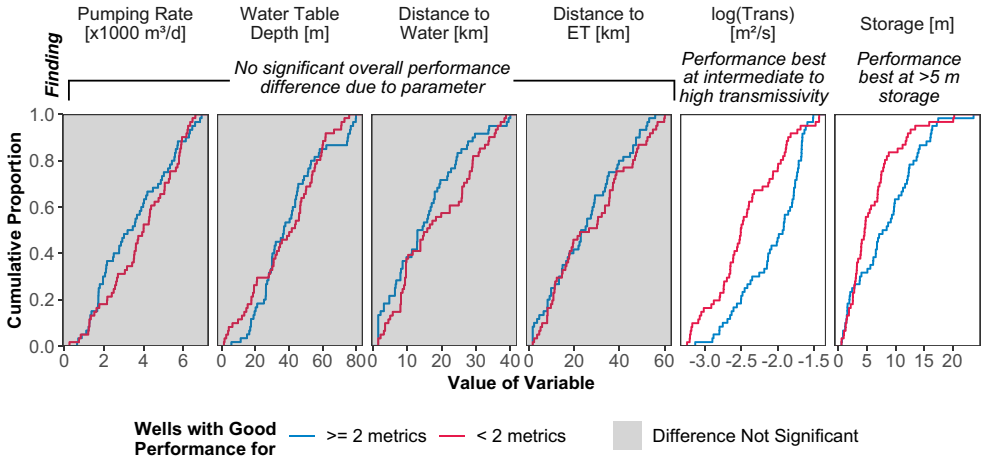

Figure 5. Regional sensitivity analysis results, expressed as empirical cumulative distribution functions for wells with good performance in at least two fit metrics and wells with good performance in less than two fit metrics. In shaded panels, distributions of the two samples are expected to be drawn from the same distribution (two-sample Kolmogorov-Smirnov test $p > 0.05$).

included) and total streamflow depletion had a negative bias of −8.4% (compared to 0.4% with DRNs). However, normalized MAD of depletion for the most affected segment improved to 0.043 without DRNs (compared to 0.048 with DRNs) and KGE for depletion of all stream segments rose to 0.811 without DRNs (compared to 0.779 with DRNs).

3.3. Well and Landscape Drivers of Performance Variability

Using the threshold defined in Section 2.3.4 for each performance metric, we found that 13 wells had good performance for all four metrics, 13 wells had good performance for three metrics, 33 wells had good performance for two metrics, 38 wells had good performance for a single metric, and 23 wells did not have good performance for any metric. To investigate the relative importance of each parameter with comparable sample sizes, we compared wells with good performance in at least two fit metrics to wells with good performance in less than two fit metrics (Figure 5). Differences in the empirical cumulative distribution functions between the two samples for a given parameter indicates a potentially significant impact of that parameter on overall analytical depletion function performance (Pianosi et al., 2016; Wagener et al., 2001). Comparing across all fit metrics concurrently, the only two parameters that led to significant differences between these two samples were the hydrostratigraphic properties of transmissivity and aquifer storage (the product of specific yield and saturated thickness). Wells tended to have better performance at intermediate to high values of transmissivity, agreeing with the observed drivers of over- and underestimated depletion (Figure 3). Similarly, agreement between the analytical depletion functions and the numerical model was greater at intermediate to high (>5 m) values of aquifer storage (Figure 5).

Investigating individual fit metrics, the regional sensitivity analysis found a significant impact of all of the well and landscape properties except pumping rate on at least one of the fit metrics, but no well or landscape property had a significant impact on all fit metrics (Figure 6). For pumping rate, none of the fit metrics differed significantly between wells with a good and poor fit, which supports the long-held assumption that streamflow depletion can be estimated as a fraction of pumping, independent of abstraction rate (Glover & Balmer, 1954; Hunt, 1999; Theis, 1941). Transmissivity significantly affected the most fit metrics, with intermediate transmissivity values associated with improved prediction of depletion (MAD in the most-affected segment and KGE in all segments) and bias. The difference between the "good" and "poor" wells at intermediate values of transmissivity further supports the observation that performance degrades in extremely high and low transmissivity settings near streams (Figure 3), and that variability in hydraulic conductivity is an important control over analytical depletion function accuracy (B.-D. Li et al., 2016; Sophocleous et al., 1995). Interestingly, aquifer storage had different effects depending on whether the primary impacts

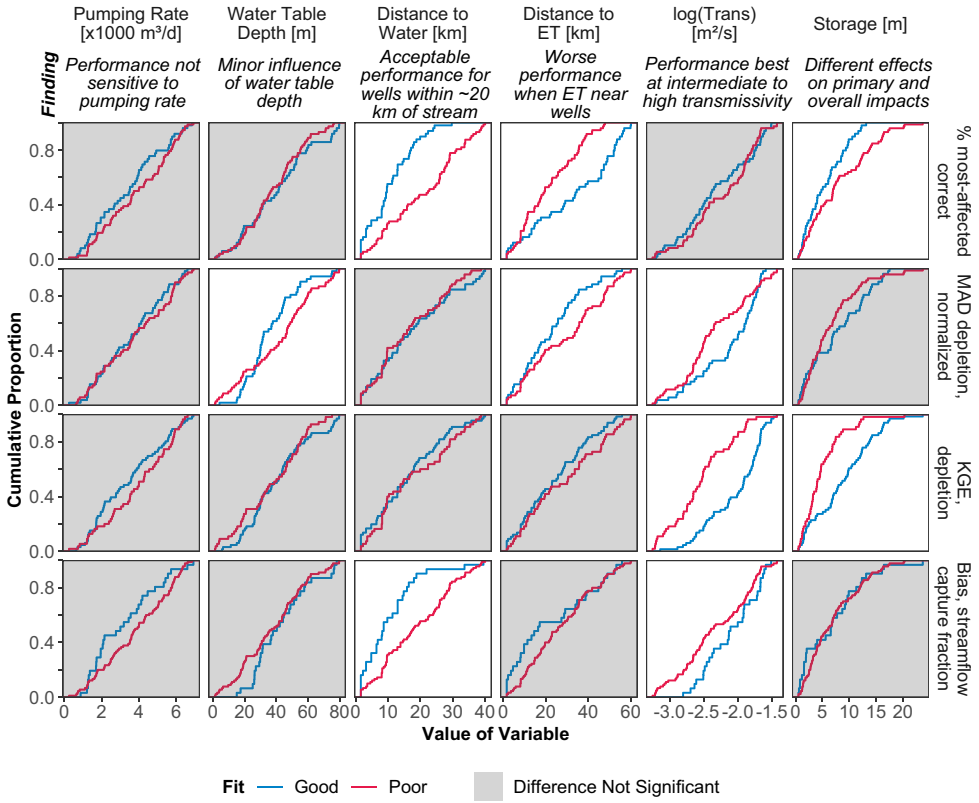

Figure 6. Regional sensitivity analysis results, expressed as empirical cumulative distribution functions for well characteristics for each performance criteria. In shaded panels, distributions of "good" and "poor" groups are expected to be drawn from the same distribution (two-sample Kolmogorov-Smirnov test $p > 0.05$).

or the overall impacts were being considered. Lower storage values tended to be associated with better agreement between the MODFLOW model and the analytical depletion functions for identification of the most affected segment, whereas higher storage was associated with better performance for the KGE of all stream segments (Figure 6). The differences in the performance-storage relationship indicate that assessing the relative impact and importance of a given parameter depends on the aspect of model performance being considered.

There was an apparent threshold-type response as distance to the closest surface water feature increased; the identification of the most-affected segment and the total streamflow depletion bias both degraded significantly at distances greater than ~20 km. For distance to cells with ET, which is an alternate source of capture in the MODFLOW model that is not considered by the analytical depletion functions, wells that performed poorly for identifying the most-affected segment and MAD of depletion were primarily concentrated at shorter distances, while the MAD of depletion estimates improved at further distances to ET. Since phreatophytic ET is primarily concentrated along stream channels in the MODFLOW model (RRCA, 2003), this indicates a spatial interplay between streamflow and ET capture sources which merits future investigation (Condon & Maxwell, 2019). The water table depth had only a minor influence on the fit metrics, with better depletion predictions at intermediate water table depths (~20–50 m). It is important to note that we were assessing fit between the analytical depletion functions and the MODFLOW model, not agreement with field observations (which are not available). As a result, the division of fit into "good" and "poor" categories may be driven by errors in the MODFLOW model instead of or in addition to errors in the analytical depletion functions, and potential errors in the MODFLOW model may also vary in response to parameters evaluated here such as well-stream distance.

While we did not use the distance to the edge of the model domain as one of the variables guiding our well sample selection, we conducted a *post hoc* analysis to evaluate whether it had a significant impact on

performance. We found that analytical depletion functions more successfully identify the most-affected segment and have an acceptable bias for wells that were closer to no-flow boundaries (Figure S12). This response is very similar to the observed influence of distance to the closest surface water feature (Figure 5) and we were not able to isolate the impacts of these no-flow boundaries because they are often colocated with or near surface water features (Figure 2). In aquifers of limited lateral extent, analytical models for bounded aquifers (Huang et al., 2018) may be useful methods to integrate into analytical depletion functions, but would need additional testing.

Combined, our analysis indicates that data collection efforts should prioritize high-accuracy estimates of transmissivity and storativity to improve accuracy of both streamflow depletion and groundwater depletion predictions, since hydraulic diffusivity (the ratio of storativity to transmissivity) is fundamental to the timing and magnitude of streamflow depletion (Barlow & Leake, 2012). In areas with high-quality water use and water level data, emerging data-driven approaches may be a valuable tool for improving storativity estimates (Butler et al., 2016, 2020; Whittemore et al., 2016).

3.4. Synthesis With Previous Analytical Depletion Function Evaluations

This work extends previous evaluations of analytical depletion functions by comparing their output to a calibrated numerical model in a highly stressed aquifer, a setting where analytical depletion functions have not previously been tested, and by systematically assessing the influence of well and hydrostratigraphic characteristics on results. Synthesizing across studies, we find general agreement that the adjacent + expanding stream proximity criteria and the web squared depletion apportionment equation produce the best agreement with numerical model output (Zipper, Dallemagne, et al., 2018; Zipper, Gleeson, et al., 2019). We also found performance was best when wells are close to streams, a finding that is consistent with previous work in coastal California (Zipper, Gleeson, et al., 2019) but in contrast to a study in British Columbia that found better agreement for wells further from streams (Q. Li et al., 2020). We also extend this previous work by testing performance across 166 wells with a variety of pumping rates and demonstrated that performance is insensitive to pumping rate, indicating that analytical depletion functions are likely to be equally useful regardless of the magnitude of groundwater abstractions.

This study and previous work raise several key questions for further evaluation. First, we identify a potential spatial performance-related interaction between the distance from the well to the closest stream and closest ET cell (Figure 6). Additional testing is necessary to determine the conditions in which phreatophytic ET confounds analytical depletion function estimates of streamflow depletion (Condon & Maxwell, 2019). Second, this and previous evaluations have focused on evaluation of a single well in isolation. While the current study investigates performance in the context of a heavily stressed aquifer with many pumping wells, we isolated effects of an individual well by turning wells on/off one-at-a-time. While it is widely assumed that the output from analytical models is additive, work in the Republican River Watershed has shown this may not be the case (Schneider et al., 2017). For application of analytical depletion functions in heavily stressed aquifers, systematic testing of cumulative impacts by evaluating the impacts of multiple wells concurrently is critical. Third, recent field investigations found that analytical model performance in an urban setting varied as a function of stream stage (Flores et al., 2020) and will be important to test analytical depletion functions in a variety of stream stage conditions. Finally, we evaluated analytical depletion functions via comparison to a calibrated numerical model, but not via direct comparison to field-based estimates of streamflow depletion. This is because there are not currently techniques that are able to quantify segment-resolution, regional-scale estimates of streamflow depletion from observational data, and therefore, numerical models are considered the most accurate approach (Barlow & Leake, 2012). Further work to develop improved large-scale high-resolution streamflow depletion benchmarking datasets would benefit both the evaluation of analytical depletion functions as well as numerical models and water managers.

Given the low computational and data requirements of analytical depletion functions relative to numerical models, they may be a particularly valuable tool for applications requiring many streamflow depletion estimates under different conditions such as decision support tools, time series analysis, and simulation-optimization management modeling. Huggins et al. (2018) developed a workflow for integrating depletion apportionment equations and analytical models into existing web-based water decision support tools. Furthermore, analytical depletion functions could be used to improve parameterization of pumping impacts

in time series analysis approaches, which typically require a head response function that is often based on analytical methods (Bakker & Schaars, 2019; Obergfell et al., 2019; Shapoori et al., 2015). Finally, simulation-optimization models require the ability to test many different management approaches (A. Singh, 2014; Wagner, 1995). While this can be accomplished in relatively small domains using numerical models (Fienen et al., 2018), analytical depletion functions may complement other approaches such as metamodeling (Fienen et al., 2015) to provide estimates of streamflow depletion in large model domains under diverse scenarios and identify optimal management solutions. For all of these potential applications, however, care should be taken to ensure that uncertainty and limitations of analytical approaches are appropriately considered, quantified, and shared with relevant stakeholders, so that decision-makers can determine whether the accuracy is sufficient for their needs.

4. Conclusions

Reliable estimates of streamflow depletion are critical for effective conjunctive management of groundwater and surface water resources. This study is the first systematic evaluation of analytical depletion functions for use in a heavily stressed unconfined aquifer and assesses how agreement between the numerical and analytical model varies as a function of well and hydrostratigraphic characteristics. We found that analytical depletion functions can produce similar estimates of streamflow depletion to an existing calibrated numerical model during both the pumping and nonpumping seasons, though they tend to over- or underestimate depletion relative to MODFLOW for wells very close to surface water features. Comparing among eight different analytical depletion functions, we found relatively little sensitivity to analytical depletion function formulation, but a strong response to the hydrostratigraphic properties of transmissivity and aquifer storage, indicating the critical importance of reliable parameter estimates. Among the analytical depletion functions, the one that performs most similarly to the numerical model included time-varying stream proximity criteria and a depletion apportionment equation that accounted for stream network geometry, which is consistent with previous studies. The analytical depletion function and numerical model are most similar for wells within ~20 km of a stream and intermediate values of transmissivity, with no sensitivity to pumping rate. These results do not suggest that numerical models should be replaced or superseded by analytical depletion functions, but rather that analytical depletion functions are a useful low-cost, low-effort approach to obtain comparable estimates of streamflow depletion in settings where calibrated numerical models are not available.

Data Availability Statement

The Republican River Compact Administration groundwater model is available at http://www.republicanrivercompact.org/. Code and data developed during this study are available via the Open Science Framework at https://doi.org/10.17605/OSF.IO/CPV94 .

Acknowledgments
We gratefully acknowledge the many people involved in the development of the Republican River Compact Administration MODFLOW model. This work was funded by a Natural Sciences and Engineering Research Council Collaborative Research and Development Grant (NSERC CRD) to the University of Victoria and Foundry Spatial. We also appreciate helpful feedback from the editor, associate editor, Paul Barlow, and two anonymous reviewers.

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*, 1835–1845. https://doi.org/10.1007/s10040-016-1438-3

Bakker, M., Post, V., Langevin, C. D., et al. (2016). Scripting MODFLOW model development using Python and FloPy. *Ground Water*, *54*, 733–739. https://doi.org/10.1111/gwat.12413

Bakker, M., Post, V., Langevin, C. D., et al. (2018). *FloPy* (Version 3.29). http://dx.doi.org/10.5066/F7BK19FH

Bakker, M., & Schaars, F. (2019). Solving groundwater flow problems with time series analysis: You may not even need another model. *Ground Water*, *57*(6), 826–833. https://doi.org/10.1111/gwat.12927

Barlow, P. M., & Leake, S. A. (2012). *Streamflow depletion by wells—Understanding and managing the effects of groundwater pumping on streamflow*. Reston, VA: U.S. Geological Survey.

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

Beck, H. E., van Dijk, A. I. J. M., Miralles, D. G, et al. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, *49*, 7843–7863. https://doi.org/10.1002/2013WR013918

Bredehoeft, J. D. (2002). The water budget myth revisited: Why hydrogeologists model. *Ground Water*, *40*, 340–345. https://doi.org/10.1111/j.1745-6584.2002.tb02511.x

Bredehoeft, J. D., Papadopulos, S. S., & Cooper, H. H. (1982). Groundwater: The water budget myth. In *Scientific basis of water resource management* (pp. 51–57). Washington, DC: National Research Council. Retrieved from https://www.nap.edu/catalog/19530/scientific-basis-of-water-resource-management

Butler, J. J., Bohling, G. C., Whittemore, D. O., & Wilson, B. B. (2020). A roadblock on the path to aquifer sustainability: Underestimating the impact of pumping reductions. *Environmental Research Letters*, *15*, 014003. https://doi.org/10.1088/1748-9326/ab6002

Butler, J. J., Whittemore, D. O., Wilson, B. B., & Bohling, G. C. (2016). A new approach for assessing the future of aquifers supporting irrigated agriculture. *Geophysical Research Letters*, *43*, 2004–2010. https://doi.org/10.1002/2016GL067879

Butler, J. J., Whittemore, D. O., Wilson, B. B., & Bohling, G. C. (2018). Sustainability of aquifers supporting irrigated agriculture: A case study of the High Plains Aquifer in Kansas. *Water International*, *43*, 815–828. https://doi.org/10.1080/02508060.2018.1515566

Butler, J. J., Zhan, X., & Zlotnik, V. A. (2007). Pumping-induced drawdown and stream depletion in a leaky aquifer system. *Ground Water*, *45*, 178–186. https://doi.org/10.1111/j.1745-6584.2006.00272.x

Chen, C., & Liu, L.-M. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, *88*, 284–297. https://doi.org/10.2307/2290724

Condon, L. E., & Maxwell, R. M. (2019). Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion. *Science Advances*, *5*, eaav4574. https://doi.org/10.1126/sciadv.aav4574

de Graaf, I. E. M., Gleeson, T., van Beek, L. P. H., Sutanudjaja, E. H., & Bierkens, M. F. P. (2019). Environmental flow limits to global groundwater pumping. *Nature*, *574*(7776), 90–94. https://doi.org/10.1038/s41586-019-1594-4

Deines, J. M., Kendall, A. D., Crowley, M. A., et al. (2019). Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. *Remote Sensing of Environment*, *233*, 111400. https://doi.org/10.1016/j.rse.2019.111400

Deines, J. M., Kendall, A. D., & Hyndman, D. W. (2017). Annual irrigation dynamics in the U.S. northern High Plains derived from Landsat satellite data. *Geophysical Research Letters*, *44*, 9350–9360. https://doi.org/10.1002/2017GL074071

Fienen, M. N., Bradbury, K. R., Kniffin, M., & Barlow, P. M. (2018). Depletion mapping and constrained optimization to support managing groundwater extraction. *Ground Water*, *56*, 18–31. https://doi.org/10.1111/gwat.12536

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. *Environmental Modelling & Software*, *77*, 95–107. https://doi.org/10.1016/j.envsoft.2015.11.023

Fienen, M. N., Nolan, B. T., Feinstein, D. T., & Starn, J. J. (2015). Metamodels to bridge the gap between modeling and decision support. *Ground Water*, *53*, 511–512. https://doi.org/10.1111/gwat.12339

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

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*, 468–470. https://doi.org/10.1029/TR035i003p00468

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 modeling. *Journal of Hydrology*, *377*, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Haacker, E. M. K., Kendall, A. D., & Hyndman, D. W. (2016). Water level declines in the High Plains Aquifer: Predevelopment to resource senescence. *Ground Water*, *54*, 231–242. https://doi.org/10.1111/gwat.12350

Hamilton, D. A., & Seelbach, P. W. (2011). *Michigan's water withdrawal assessment process and internet screening tool* (Fisheries Division Special Report 55).

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000). *MODFLOW-2000, the U.S. Geological Survey modular ground-water model—User guide to modularization concepts and the ground-water flow process*. U.S. Geological Survey.

Hrozencik, R. A., Manning, D. T., Suter, J. F., Goemans, C., & Bailey, R. T. (2017). The heterogeneous impacts of groundwater management policies in the Republican River Basin of Colorado. *Water Resources Research*, *53*(12), 10757–10778. https://doi.org/10.1002/2017WR020927

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

Hu, Y., Cai, X., & DuPont, B. (2015). Design of a web-based application of the coupled multi-agent system model and environmental model for watershed management analysis using Hadoop. *Environmental Modelling & Software*, *70*, 149–162. https://doi.org/10.1016/j.envsoft.2015.04.011

Huggins, X., Gleeson, T., Eckstrand, H., & 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*, 98–102. 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

Hu, Y., Quinn, C. J., Cai, X., & Garfinkle, N. W. (2017). Combining human and machine intelligence to derive agents' behavioral rules for groundwater irrigation. *Advances in Water Resources*, *109*, 29–40. https://doi.org/10.1016/j.advwatres.2017.08.009

Jenkins, C. T. (1968). Techniques for computing rate and volume of stream depletion by wells. *Ground Water*, *6*, 37–46. https://doi.org/10.1111/j.1745-6584.1968.tb01641.x

Khan, H. F., & Brown, C. M. (2019). Effect of hydrogeologic and Climatic variability on performance of a groundwater market. *Water Resources Research*, *55*, 4304–4321. https://doi.org/10.1029/2018WR024180

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424*(425), 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011

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

Kollet, S. J., & Zlotnik, V. A. (2003). Stream depletion predictions using pumping test data from a heterogeneous stream–aquifer system (a case study from the Great Plains, USA). *Journal of Hydrology*, *281*(1), 96–114. https://doi.org/10.1016/S0022-1694(03)00203-8

Kollet, S. J., Zlotnik, V. A., & Ledder, G. (2002). "A stream depletion field experiment," by Bruce Hunt, Julian Weir, and Bente Clausen. *Ground Water*, *40*(4), 448–449. https://doi.org/10.1111/j.1745-6584.2002.tb02523.x

Larsen, L. G., & Woelfle-Erskine, C. (2018). Groundwater is key to salmonid persistence and recruitment in intermittent Mediterranean-climate streams. *Water Resources Research*, *54*, 8909–8930. https://doi.org/10.1029/2018WR023324

Li, B.-D., Sampath, P. V., & Liao, H. (2016). A study on the hydro-geologic factors that affect streamflow depletion. *Hydrological Sciences Journal*, *61*, 2133–2143. https://doi.org/10.1080/02626667.2015.1083649

Li, Q., Zipper, S. C., & Gleeson, T. (2020). Streamflow depletion from groundwater pumping in contrasting hydrogeological landscapes: Evaluation and sensitivity of a new management tool. *Journal of Hydrology*. *590*, 125568https://doi.org/10.1016/j.jhydrol.2020.125568

López-de-Lacalle, J. (2019). *tsoutliers: Detection of outliers in time series* (Version R package version 0.6-8). Retrieved from https://CRAN.R-project.org/package=tsoutliers

McDonald, M. G., & Harbaugh, A. W. (1988). *A modular three-dimensional finite-difference ground-water flow model*. U.S. Geological Survey.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*, 239–245. https://doi.org/10.2307/1268522

Mulligan, K. B., Brown, C., Yang, Y.-C. E., & Ahlfeld, D. P. (2014). Assessing groundwater policy with coupled economic-groundwater hydrologic modeling. *Water Resources Research*, *50*(3), 2257–2275. https://doi.org/10.1002/2013WR013666

Obergfell, C., Bakker, M., & Maas, K. (2019). Identification and Explanation of a change in the groundwater regime using time series analysis. *Ground Water*, *57*(6), 886–894. https://doi.org/10.1111/gwat.12891

Ou, G., Munoz-Arriola, F., Uden, D. R., Martin, D., Allen, C. R., & Shank, N. (2018). Climate change implications for irrigation and groundwater in the Republican River Basin, USA. *Climatic Change*, *151*(2), 303–316. https://doi.org/10.1007/s10584-018-2278-z

Pianosi, F., Beven, K., Freer, J, et al. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, *79*, 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008

Rathfelder, K. M. (2016). *Modeling tools for estimating effects of groundwater pumping on surface waters*. Victoria, BC: Ministry of Environment.

Reeves, H. W., Hamilton, D. A., Seelbach, P. W., & Asher, A. J. (2009). *Ground-water-withdrawal component of the Michigan water-withdrawal screening tool*. Reston, VA: U.S. Geological Survey.

Rohde, M. M., Froend, R., & Howard, J. (2017). A global synthesis of managing groundwater dependent ecosystems under sustainable groundwater policy. *Ground Water*, *55*(3), 293–301. https://doi.org/10.1111/gwat.12511

RRCA. (2003). *Republican river compact administration ground water model*. Republican River Compact Administration. Retrieved from https://www.republicanrivercompact.org

RRCA (2005). *Republican river compact administration groundwater model user manual (DRAFT) V12s 1.0*, Colorado, Nebraska, and Kansas: Republican River Compact Administration.

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*, 697–706. https://doi.org/10.1111/1752-1688.12525

Shapoori, V., Peterson, T. J., Western, A. W., & Costelloe, J. F. (2015). Top-down groundwater hydrograph time-series modeling for climate-pumping decomposition. *Hydrogeology Journal*, *23*(4), 819–836. https://doi.org/10.1007/s10040-014-1223-0

Singh, S. K. (2009). Flow depletion induced by pumping well from stream perpendicularly intersecting impermeable/recharge boundary. *Journal of Irrigation and Drainage Engineering*, *135*, 499–504. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000095

Singh, A. (2014). Simulation–optimization modeling for conjunctive water use management. *Agricultural Water Management*, *141*, 23–29. https://doi.org/10.1016/j.agwat.2014.04.003

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*, 579–588. https://doi.org/10.1111/j.1745-6584.1995.tb00313.x

Spear, R. C., & Hornberger, G. M. (1980). Eutrophication in Peel Inlet—II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, *14*, 43–49. https://doi.org/10.1016/0043-1354(80)90040-8

Theis, C. V. (1941). The effect of a well on the flow of a nearby stream. *Eos, Transactions, American Geophysical Union*, *22*, 734–738. https://doi.org/10.1029/TR022i003p00734

Wagner, B. J. (1995). Recent advances in simulation-optimization groundwater management modeling. *Reviews of Geophysics*, *33*(S2), 1021–1028. https://doi.org/10.1029/95rg00394

Wagener, T., Boyle, D. P., Lees, M. J, et al. (2001). A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, *5*, 13–26. https://doi.org/10.5194/hess-5-13-2001

Whittemore, D. O., Butler, J. J., & Wilson, B. B. (2016). Assessing the major drivers of water-level declines: New insights into the future of heavily stressed aquifers. *Hydrological Sciences Journal*, *61*, 134–145. https://doi.org/10.1080/02626667.2014.959958

Yeh, H.-D., Chang, Y.-C., & Zlotnik, V. A. (2008). Stream depletion rate and volume from groundwater pumping in wedge-shape aquifers. *Journal of Hydrology*, *349*, 501–511. https://doi.org/10.1016/j.jhydrol.2007.11.025

Zipper, S. C., Carah, J. K., Dillis, C., et al. (2019). Cannabis and residential groundwater pumping impacts on streamflow and ecosystems in Northern California. *Environmental Research Communications*, *1*, 125005. https://doi.org/10.1088/2515-7620/ab534d

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

Zipper, S. C., Gleeson, T., Kerr, B., et al. (2019). Rapid and accurate estimates of streamflow depletion caused by groundwater pumping using analytical depletion functions. *Water Resources Research*, *55*, 5807–5829. https://doi.org/10.1029/2018WR024403

Zipper, S. C., Lamontagne-Hallé, P., McKenzie, J. M., & Rocha, A. V. (2018). Groundwater controls on post-fire permafrost thaw: Water and energy balance effects. *Journal of Geophysical Research: Earth Surface*, *123*, 2677–2694. https://doi.org/10.1029/2018JF004611

Zlotnik, V. A., & Tartakovsky, D. M. (2008). Stream depletion by groundwater pumping in leaky aquifers. *Journal of Hydrologic Engineering*, *13*, 43–50. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:2(43)