

UC Davis

UC Davis Previously Published Works

Title

Vertical Hydraulic Gradients as Key Calibration Constraints in Regional Groundwater Models

Permalink

<https://escholarship.org/uc/item/9hw0b5zt>

Journal

Water Resources Research, 62(8)

ISSN

0043-1397

Authors

Pasner, Yara M
Fogg, Graham E
Harter, Thomas
[et al.](#)

Publication Date

2026-08-01

DOI

10.1029/2026wr043771

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Water Resources Research®

RESEARCH ARTICLE

10.1029/2026WR043771

Key Points:

- Regional groundwater models reproduce heads but diverge in vertical flow and budgets, revealing key model contrasts
- Calibrating to vertical head gradients increased vertical anisotropy and yielded a more defensible vertical flow field
- Observed vertical head gradients are increasing across California's Sacramento Valley, which signals emerging groundwater risk

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. M. Pasner,
ypasner@gmail.com

Citation:

Pasner, Y. M., Fogg, G. E., Harter, T., & Dahlke, H. E. (2026). Vertical hydraulic gradients as key calibration constraints in regional groundwater models. *Water Resources Research*, 62, e2026WR043771. <https://doi.org/10.1029/2026WR043771>

Received 6 FEB 2026
Accepted 29 JUL 2026

Author Contributions:

Conceptualization: Yara M. Pasner, Graham E. Fogg, Thomas Harter
Formal analysis: Yara M. Pasner
Funding acquisition: Helen E. Dahlke
Methodology: Yara M. Pasner, Graham E. Fogg
Supervision: Helen E. Dahlke
Visualization: Yara M. Pasner
Writing – original draft: Yara M. Pasner, Graham E. Fogg
Writing – review & editing: Yara M. Pasner, Graham E. Fogg, Thomas Harter, Helen E. Dahlke

© 2026 The Author(s). *Water Resources Research* published by Wiley Periodicals LLC on behalf of American Geophysical Union.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Vertical Hydraulic Gradients as Key Calibration Constraints in Regional Groundwater Models

Yara M. Pasner¹ , Graham E. Fogg² , Thomas Harter² , and Helen E. Dahlke² 

¹Hydrologic Sciences Graduate Group, University of California, Davis, Davis, CA, USA, ²Department of Land, Air and Water Resources, University of California, Davis, Davis, CA, USA

Abstract In typical clastic sedimentary aquifer systems, interbedded coarse- and fine-grained deposits impede vertical flow and promote horizontal flow. This effect, known as vertical anisotropy, is often underestimated in regional groundwater models, leading to overestimated vertical aquifer communication. The issue stems partly from dependence on vertically insensitive calibration targets, such as laterally distributed observations of groundwater levels. As a result, models may appear well-calibrated while differing in their representation of key processes, including surface water–groundwater exchange, recharge, and solute transport pathways. Yet few studies have demonstrated how poorly constrained vertical flow parameters can bias model outcomes. California's Sacramento Valley (19,800 km²) illustrates the value of observed vertical hydraulic gradients (VGs—representing head change with depth) for constraining aquifer parameters during model calibration. Of three publicly available models developed for this region, only one uses VGs from 81 continuously monitored well nests to calibrate aquifer parameters. Although all three models reproduce observed heads comparably well, only this model also reproduced observed VGs, exhibited stronger vertical anisotropy than the other two, and simulated markedly different water budgets. This stronger anisotropy corresponds with a vertical flow field that better matches expected recharge and discharge patterns. In addition, observed VGs appear to be increasing at most well nests (56 of 81), reflecting strengthening downward (or weakening upward) flow potential. Continued increases could intensify surface water losses and expose deeper aquifers to contamination. Our findings highlight the importance of VG-informed monitoring and model development for detecting and managing emerging risks in sedimentary aquifer systems.

1. Introduction

The acceleration of groundwater depletion in aquifer systems globally, driven by pumping and climate warming, underscores the need for accurate regional groundwater models to inform management decisions (Jasechko et al., 2024; Wada, 2016). Yet in sedimentary basins, which are among the world's most common and productive aquifer systems (Condon et al., 2021; Richts et al., 2011), models frequently misrepresent a key control on three-dimensional flow: the tendency of interbedded coarse- and fine-grained deposits to impede vertical flow and promote horizontal movement (Shepley, 2024). This bias is highlighted by a recent meta-analysis comparing field and laboratory data against values from 113 published groundwater models (Shepley, 2024), which shows that regional-scale models often simulate vertical saturated hydraulic conductivity (K_v) values that are too large relative to horizontal saturated hydraulic conductivity (K_h). That is, many models underestimate vertical anisotropy— K_h/K_v , the ratio describing how much more groundwater flows horizontally than vertically through a given medium at comparable gradients. While numerical models commonly use K_h/K_v values ≤ 10 (Shepley, 2024), empirical data from sedimentary aquifers suggest vertical anisotropy often exceeds 10^3 and can surpass 10^5 in some systems (see Section 2.2), revealing a profound mismatch between common model assumptions and geologic reality.

This underestimation can result in model underperformance and significantly impact the validity of model-based groundwater management recommendations. For example, errors in simulated K_h/K_v can distort vertical and lateral flow pathways, with important implications for both water supply and contaminant transport (Fogg, 1986). One reason vertical anisotropy is frequently misrepresented in regional models of sedimentary aquifer systems is a widespread reliance on calibration metrics that are largely insensitive to vertical flow parameters (Shepley, 2024). In particular, fitting to laterally distributed groundwater level observations provides limited constraint on vertical flow, even when those observations are distributed vertically. This is due to differences in head between distant wells being typically much larger than vertical head differences across depth intervals at any given landscape location.

A promising yet widely underexplored strategy to better calibrate vertical anisotropy in regional groundwater models is to explicitly evaluate a model's ability to reproduce observed vertical hydraulic gradients (VGs) (Fogg, 1986; Tanachaichoksirikun et al., 2020; see below). The relationship between VGs and vertical anisotropy is supported by classical groundwater flow theory: in a natural, ambient flow aquifer system with at least some amount of vertical hydraulic stress (e.g., from recharge), larger VGs occur in areas with greater vertical resistance (e.g., clays with low K_v) since larger gradients are needed to drive groundwater through less conductive materials, and vice versa (see discussions in Fogg, 1986; Meyer et al., 2016; Taylor et al., 2003).

An early example of using VGs to calibrate vertical anisotropy in a regional scale groundwater model was conducted in the Wilcox-Carrizo aquifer system of Texas, USA, where VGs helped identify regions of increased vertical connectivity that were otherwise undetectable (Fogg, 1986). Fogg (1986) demonstrated, using VGs estimated from regional head data (a method discussed in Section 5.2), that the domain generally showed large observed VGs ($VG > 0.1$, more than 1 m head drop for every 10 m increase in depth) except in vertically connected areas which showed much smaller magnitude VGs. Large values of vertical anisotropy (10^3 – 10^4) were necessary for the model to reproduce observed VGs; model scenarios with smaller anisotropy simulated small-magnitude VGs everywhere except along boundaries where measured VGs were imposed. Similarly, in Thailand's Lower Chao Phraya basin, the vertical anisotropy value used in the model had to be increased to 10^3 or 10^4 for simulated VGs to match observed behavior (Tanachaichoksirikun et al., 2020). Despite these and other similar studies (e.g., Arizona Department of Water Resources, 2023; Hanson et al., 2014, 2018; Ma et al., 2011), VGs are rarely incorporated into model development (Shepley, 2024).

In this study, we seek to illustrate how incorporating VGs during model calibration influences vertical anisotropy and the groundwater budget in a typical, clastic sedimentary aquifer system. California's Sacramento Valley (19,800 km²) presents a rare opportunity to evaluate these dynamics in an unusually data-rich, regionally significant aquifer system. Three publicly available numerical models simulate this region: the Sacramento Valley Groundwater-Surface Water Simulation Model version 1.0 (SVSim; California Department of Water Resources [DWR], 2022), the California Central Valley Groundwater-Surface Water Simulation Model fine-grid version 1.5 (C2VSim; DWR, 2025), and the Central Valley Hydrologic Model version 2 (CVHM2; Faunt et al., 2024). C2VSim and CVHM2 each rely mostly on traditional observations for aquifer parameterization (e.g., laterally distributed head), while SVSim additionally incorporates vertically differentiated heads (closely related to VGs, but not normalized by the vertical separation between screens) from a network of 81 continuously monitored multilevel well nests to calibrate K_h and K_v . By comparing these three models, we demonstrate how VG-informed calibration can reshape simulated anisotropy and vertical flow fields, and that the three models simulate substantially different groundwater budgets despite achieving similar performance against head observations. These results highlight a general limitation of head-based calibration relevant to sedimentary aquifer systems worldwide.

2. Background

2.1. Definition and Computation of the Vertical Hydraulic Gradient

The VG is the vertical component of the 3D hydraulic gradient vector. The direction of the VG, represented by the sign (positive or negative), is inversely related to the direction of vertical groundwater flow (recharge or discharge) by the vertical expression of Darcy's Law:

$$q_z = -K_v \text{ VG} \quad (1)$$

where q_z is vertical specific discharge and K_v is saturated vertical hydraulic conductivity (both have units of length per time). A positive VG (increasing head with elevation) corresponds to downward flow potential, whereas a negative VG corresponds to upward flow potential.

When evaluating VGs between two groundwater elevation data points, the VG is computed as:

$$\text{VG} = \frac{\Delta h}{\Delta z} \equiv \frac{h_1 - h_2}{z_1 - z_2} \quad (2)$$

where h_1 and h_2 are head values measured at elevations z_1 and z_2 , respectively, with $\Delta z > 0$. Multilevel groundwater monitoring systems (e.g., Cherry et al., 2015), referred to here as well nests, directly measure VGs between vertically adjacent screened intervals, where Δh is the difference in observed water levels and Δz is the distance between screen midpoints. In 3D numerical groundwater models, VGs are computed the same way, using the difference in simulated head between vertically adjacent nodes divided by the vertical distance separating those nodes.

2.2. Scale Dependency of Vertical Anisotropy

The cause of vertical anisotropy is highly scale dependent. At the local scale (vertical length scales ≤ 1 m; Shepley, 2024), vertical anisotropy generally arises from differences in matrix geometry or grain alignment within a given deposit. For example, clay is composed of laterally elongated platy materials that facilitate greater lateral than vertical flow (Fetter, 2000). The expected range of vertical anisotropy at the local (or hydrofacies) scale varies from 1 to 10 (Freeze & Cherry, 1979). In regional numerical modeling of groundwater flow (vertical length scales > 10 m; Shepley, 2024), model layers generally represent a single hydrogeologic unit (HGU), defined as a contiguous geologic unit composed of similar hydraulic properties (Anderson et al., 2015). At the HGU scale, vertical anisotropy in sedimentary aquifer systems has been proven highly variable, with values ranging from < 1 to $> 10^5$ (Anderson et al., 2015; Runkel & Meyer, 2025; Shepley, 2024). In elastic, sedimentary environments, where HGUs typically contain a mix of interbedded materials like gravel, sand, silt, and clay, large values of upscaled vertical anisotropy arise from the vertical resistance to flow of ubiquitous fine-grained materials, while laterally elongated coarse-grained deposits can facilitate significant horizontal flux within HGUs (Fogg, 1986; Fogg et al., 2000; Maples et al., 2019; Michael & Voss, 2009; Peffer, 1991; Tanachaichoksirikun et al., 2020; Weissmann et al., 2002).

2.3. Power-Law Averaging Framework to Upscale Aquifer Parameters

The power-law averaging framework is widely used to estimate K_h and K_v values at the model layer scale based on the composition and connectivity of HGUs (Journel et al., 1986; Sanchez-Vila et al., 2006; Scantlebury et al., 2025; Wen & Gómez-Hernández, 1996; Zanon et al., 2002). For two-material systems (coarse and fine), the power-law average is given by:

$$X = \left[P_c X_c^p + (1 - P_c) X_f^p \right]^{\frac{1}{p}} \quad (3)$$

where X is the upscaled aquifer parameter (e.g., K_h , K_v , or storage), P_c is the fraction of the domain that is coarse-grained material, and X_c and X_f are the endmember property values for coarse and fine materials, respectively (Faunt, 2009). The exponent p controls the averaging scheme: $p = 1$ yields the arithmetic mean, $p \rightarrow 0$ the geometric mean, and $p = -1$ the harmonic mean. For horizontally layered media under uniform horizontal hydraulic gradients, $p = 1$ gives the exact upscaled value of K_h . For uniform VGs through such horizontal layers, $p = -1$ yields the exact K_v (Freeze & Cherry, 1979). In practice, sedimentary HGUs often use $0 < p \leq 1$ to estimate K_h , which emphasizes coarse materials, while K_v is typically estimated using $-1 \leq p < 0$ which emphasizes fine-grained materials (Fetter, 2000). All models analyzed in this study use a calibrated power-law averaging approach to parameterize K_h and K_v .

3. Methods

In this study, we compare observed and simulated VGs in California's Sacramento Valley using data from 81 well nests and three public-domain regional groundwater models. This section outlines the study area, well nest data set, model structures, and the procedures used to evaluate model performance. Additional details are provided in Text S1–S2 (Supporting Information S1). Our approach builds on a previous analysis of four well nests and their associated model performance in the Sacramento Valley (Moss et al., 2013) by assessing how explicit calibration to observed VGs alters regional model behavior.

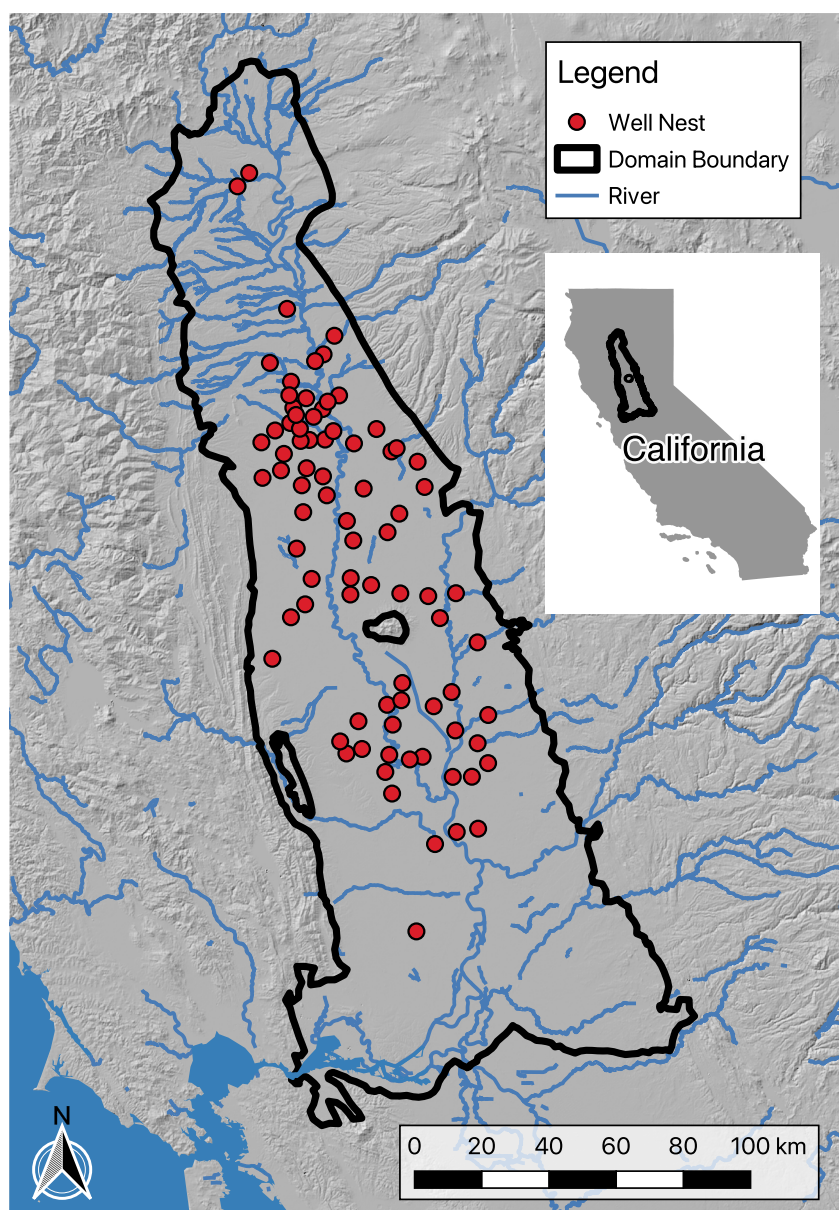


Figure 1. Map of the study domain, well nest locations, and major rivers in California's Sacramento Valley. The domain boundary corresponds to the SVSim model extent (DWR, 2022).

3.1. Study Area and Geologic Setting

The Sacramento Valley covers approximately 19,800 km² (Figure 1). It spans ~290 km from the Cascade Range southward to the Sacramento-San Joaquin River Delta, bounded by the Sierra Nevada on the east and the Coast Range and Klamath Mountains on the west (width of 40–105 km). It represents the northern third of the Central Valley, a large, asymmetric sedimentary basin underlain by Sierra Nevada batholiths and marine sedimentary rocks. The Sacramento Valley is rich in surface water and supplies a significant portion of the state's agricultural and urban demand via exports to the southern Central Valley (Hatch, 2020).

The aquifer system consists of horizontally stratified alluvial and volcanic sediments that overlie marine materials (Hatch, 2020). These deposits form interbedded, discontinuous lenses of coarse- and fine-grained materials. Unlike the San Joaquin Valley, the Sacramento Valley lacks a laterally extensive confining unit such as the Corcoran Clay; instead, increased confinement with depth arises from the cumulative vertical resistance of these

ubiquitous discontinuous fine-grained interbeds. East-side sediments from the Sierra Nevada and Cascade Range tend to be coarser than those from the west. Coarse-grained materials are also common along river channels and alluvial fans, potentially enhancing vertical connectivity in those areas (Faunt, 2009; Hatch, 2020; Maples et al., 2019; Weissmann et al., 2002). Additional details on the water use and geologic formations of the Sacramento Valley are available in Faunt (2009) and Hatch (2020).

3.2. Data Resources and Techniques

Monthly groundwater level data were downloaded from the California Natural Resources Agency (CNRA, 2025). Only data for wells of type “nested/multi-completion well” were used in this analysis, resulting in 374 unique timeseries. These data, downloaded in December 2025, spanned as far back as 1992 and were truncated at 30 September 2024 to avoid inclusion of provisional observations. We grouped timeseries by well nest using station metadata, yielding 114 well nests. We then excluded well nests outside the Sacramento Valley or without temporal overlap with model simulations, yielding 81 well nests for analysis (locations depicted in Figure 1). These well nests had three to five perforation intervals each, with screen lengths typically 4–12 m. Only three well nests had records spanning fewer than 10 years. Although nests are monitored continuously, they contain monthly water level measurements in an average of 9.7 of the 12 months each year. Well construction details and figures of the monthly timeseries data are provided in Text S4 in Supporting Information S1.

To develop a consistent terminology, we refer to groundwater level timeseries associated with the shallowest perforation as the “observed shallow head,” and that from the deepest perforation as the “observed deep head” at each well nest. Using Equation 2, we computed two types of monthly VG timeseries at each nest: (a) “top-to-bottom VG” between the shallowest and deepest perforations, and (b) “layer-pair VGs” between each vertically adjacent screen pair. To assess long-term trends in shallow head, deep head, and top-to-bottom VGs, we applied the Mann-Kendall test for monotonic trend at each well nest with at least 3 years between the first and last observation (all but one nest met this criterion), using a statistical significance level of 0.05 (Helsel et al., 2020). A trend of increasing VGs is interpreted to reflect strengthening downward (or weakening upward) flow potential over time, and vice versa.

To evaluate spatial patterns, we calculated the following metrics at each well nest: (a) average shallow head (as a proxy for water table elevation); (b) average top-to-bottom VG (to assess long-term vertical flow direction); (c) 10th and 90th percentile layer-pair VGs; and (d) variance of shallow and deep head timeseries. Percentile metrics were calculated from pooled layer-pair VG values across all observation times to represent typical low and high bounds of VG variability. Higher head variance is interpreted as a potential signal of hydraulic stress and/or greater aquifer confinement.

3.3. Models

3.3.1. CVHM2

CVHM2 (Faunt et al., 2024; Traum et al., 2024) is a groundwater model developed by the United States Geological Survey (USGS) using the finite-difference MODFLOW-One-Water Hydrologic Flow Model code (Boyce & Ferguson, 2024). CVHM2 uses a 1.6 km × 1.6 km grid with 13 vertical aquifer layers, of which 10 are active in the Sacramento Valley (layers 6–8 are inactive there). Within the Sacramento Valley (SVSim model extent; Figure 1), the active layers range from 3 to 122 m thick, with a mean thickness of 55 m. CVHM2 simulates transient water flow through the Central Valley (~53,700 km²) from April 1961 to September 2019 using monthly stress periods. The Sacramento Valley comprises the northern third of CVHM2. Although CVHM2 spans the entire valley, its development emphasized land subsidence and overdraft in the southern San Joaquin Valley, resulting in less detailed parameterization and calibration in the Sacramento Valley.

Hydraulic conductivity and storage values were parameterized in CVHM2 using a texture-based approach. Percent coarse material was estimated from a binary (coarse/fine) interpretation of well driller log data and other lithologic observations, averaged to 15 m vertical intervals and interpolated to model nodes using 3D kriging with zonally varying variograms (Faunt, 2009; Faunt et al., 2024). K_h and K_v were estimated using a calibrated power-law average (Section 2.3) with a power $p = 1$ (Equation 3) for horizontal and $p = 0.07$ for vertical averaging in the Sacramento Valley. Calibrated endmember values are $5.38 \times 10^{-3} \text{ m d}^{-1}$ (fine) and 38.3 m d^{-1} (coarse). Specific

yield (S_y) and specific storage (S_s) were assigned using zonally defined constants that vary by material type (fine vs. coarse), depth interval, and for S_s , by elastic or inelastic response.

Aquifer properties were calibrated using unitless scale factors and pilot-point multipliers. Additionally, depth-decay factors were estimated to simulate reduced conductivity (lower K_h and K_v) with depth. Calibration targets primarily relied on head, but also included drawdown trends, compaction, streamflow, and drain flows using both manual and PEST-based automated calibration (Doherty & Hunt, 2010; Faunt, 2009; Faunt et al., 2024). Across these types, 362,254 observations were used in the Central Valley-wide calibration, weighted so that each observation type contributed roughly equally to the objective function (Faunt et al., 2024).

3.3.2. C2VSim

C2VSim fine grid version 1.5 was developed by the California Department of Water Resources to support regional water management (DWR, 2025). C2VSim simulates monthly water flow for water years 1974–2021 across the entire Central Valley, using the finite-element Integrated Water Flow Model (IWFM) code (Dogrul & Kadir, 2024). C2VSim's nodes are positioned at element vertices and are spaced more closely near surface water features. Within the Sacramento Valley (SVSim model extent), it has 14,444 irregularly shaped elements with a mean area of 1.4 km² (range 0.05–7.2 km²) and four vertical layers ranging 15–884 m thick (mean 128 m). Among the three models evaluated, C2VSim incorporated the most comprehensively updated land- and water-use inputs, reconstructed for the full model period using statewide crop mapping, remote sensing products, and data submitted by local agencies under California's Sustainable Groundwater Management Act (SGMA).

C2VSim used a texture-based approach similar to CVHM2 for aquifer parameterization. Well driller logs were interpreted as binary coarse/fine material and mapped vertically into model layers. Percent coarse values were then kriged laterally to model nodes within HGUs using Texture2Par (Scantlebury et al., 2025). A power-law averaging framework was used to compute K_h , K_v , and storage parameters with exponent p values of 0.93, –0.62, and 1.0, respectively (Equation 3). Endmembers of the power-law averaging framework (X_f and X_c), hydrofacies-scale anisotropy, and depth-dependent decay parameters were calibrated at pilot points and interpolated to model nodes using Texture2Par and PEST. C2VSim calibration relied primarily on head targets (900 wells, 64,802 head data points within the Sacramento Valley) but also considered simulated water budgets and stream hydrographs, similar to CVHM2.

3.3.3. SVSim

SVSim version 1.0 (DWR, 2022) is an IWFM-based finite-element model developed by DWR to improve simulation of surface water-groundwater (SW-GW) interactions in the Sacramento Valley to support SGMA (Durbin et al., 2015). While C2VSim incorporated more recent land-use and surface-water inputs, SVSim features a finer spatial discretization, with nine vertical layers (3–430 m thick, mean 52 m) and 23,767 elements per layer, averaging around 0.83 km² in area per element.

SVSim shares its aquifer parameterization framework with C2VSim, using the same texture data, power-law coefficients, calibration targets, and software tools. However, SVSim uniquely incorporated vertical head difference timeseries from continuously monitored well nests as additional calibration targets, comprising 10,859 observations from 257 well pairs alongside 57,998 groundwater head observations from 941 wells (Bedekar, Cayar, et al., 2021). Other differences among the three models are summarized in Section 3.3.4 and Table 1, with further detail in the model's supporting documentation (Bedekar, Cayar, et al., 2021; Durbin et al., 2015).

3.3.4. Model Comparison

The three models share a common lithology field and power-law parameterization framework but differ in spatial discretization, calibration targets, and several structural features that could in principle influence vertical flow, including the vertical averaging exponents, depth-decay treatment, interborehole flow, and tile drains (Table 1). We note one asymmetry in the available documentation: observation-type weighting in the objective function is reported only for CVHM2, where types were weighted to contribute roughly equally (Faunt et al., 2024); comparable per-group weights were not available in the published documentation for the IWFM-based models (C2VSim and SVSim). Table 1 therefore compares the types of calibration targets used, not their relative weights.

Table 1
Key Similarities and Differences Among the Three Models, Reported for the Sacramento Valley Study Domain (SVSim Extent)

Feature	CVHM2	C2VSim	SVSim
Code	MODFLOW-OWHM (finite difference)	IWFM (finite element)	IWFM (finite element)
Lithology field	Shared coarse/fine texture model across all three	Shared	Shared
Horizontal grid	Regular 1.6×1.6 km square cells (2.56 km^2)	Irregular elements, mean 1.4 km^2	Irregular elements, mean 0.83 km^2
Vertical layers (in domain)	10 active, mean thickness 55 m	4, mean thickness 128 m	9, mean thickness 52 m
Hydraulic conductivity parameterization	Power-law average of coarse/fine endmembers	Same approach	Same approach
Horizontal exponent (p)	1	0.93	0.93
Vertical exponent (p)	0.07	−0.62	−0.62
Power-law endmember calibration	Single calibrated endmember set across Sacramento Valley	Endmembers + depth-decay calibrated at pilot points	Endmembers + depth-decay calibrated at pilot points
Depth decay of K	Discrete, layer-specific dimensionless multiplier; differs for K_h versus K_v	Continuous exponential, 0.007 ft^{-1} (0.023 m^{-1}), applied to endmembers before averaging; identical for K_h and K_v	Same as C2VSim
Depth decay of S_s	No	No	No
Land/water use inputs	DWR land use data set	DWR data set, most updated	DWR land use data set
Calibration targets	Head, drawdown trends, streamflow, drain flows	Head, simulated water budgets, stream hydrographs	Same as C2VSim, plus VG targets
Tile drains	Present in model, but none active in the Sacramento Valley	Simulated	Not simulated
Interborehole flow	Simulated (multi-aquifer wellbore flow)	Not simulated	Not simulated

3.4. Model Output Processing and Evaluation

To evaluate simulated heads at well nests, hydraulic head values were retrieved for all model layers at each nest location using model-specific methods (see Text S1.1 in Supporting Information S1 for details). Model layers were matched to observed screen intervals by selecting the layer whose midpoint elevation is closest to each screen's midpoint. This approach allows us to compute simulated layer-pair VGs and top-to-bottom VGs (described in Section 3.2) that are depth-matched to the observed data.

Model performance for head and VG timeseries is evaluated using scatter plots, histograms, mean error, and the Pearson correlation coefficient (r). Mean error is calculated as simulated minus observed values, so positive values indicate model overestimation. For VGs specifically, performance statistics are computed on the top-to-bottom VG only, as this metric spans depth intervals comparable to multiple model layers and is therefore the most scale-consistent basis for comparison across models with different vertical discretization. Layer-pair VGs are shown for context but are not used in performance evaluation. Simulated streamflow was also compared against observed monthly discharge from the USGS Sacramento River gauge at Freeport (gauge 11447650), located at the downstream end of the Sacramento Valley where it integrates basin-wide surface-water fluxes, over water years 1974–2015, representing the period of overlap among all three model simulations. To assess long-term trends in simulated head and VGs for each model, the Mann-Kendall test was applied at each well nest location. To ensure comparability with observed data, trend analysis was restricted to the period of temporal overlap between observed and simulated timeseries at each well nest. Trends were only estimated when the overlapping record spanned at least 3 years.

In addition to well nest and streamflow comparisons, we generate spatial maps of simulated water table elevation and VGs at all model nodes (see Text S1.2 in Supporting Information S1 for details). The simulated water table was defined as head in the shallowest saturated layer, and a top-to-bottom VG was computed at each node as the difference in head between the water table and the deepest model layer, divided by their vertical separation. September 2015 was selected for comparison as a dry, high-stress period when VGs are typically most pronounced in the Sacramento Valley. Key simulated groundwater budget components (recharge, pumping, SW-GW

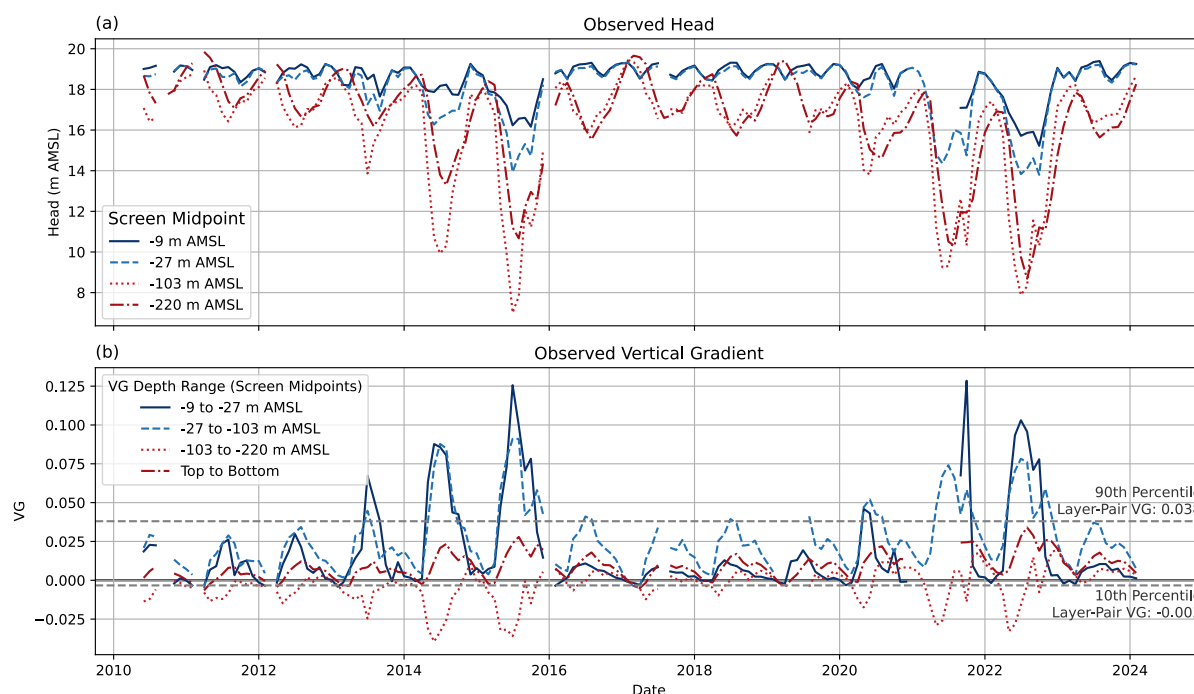


Figure 2. (a) Observed hydraulic head and (b) observed VG timeseries at the COL 10A well nest (ground surface elevation: 19.6 m AMSL). The 90th and 10th percentile of layer-pair VG values are shown as dashed horizontal lines in (b), representing the typical high and low range of the layer-pair VG distribution.

exchanges, and groundwater storage change), averaged from October 1995 to September 2015, are evaluated using model-specific methods (see Text S1.3 in Supporting Information S1; Harbaugh, 1990) to compare broader hydrologic behavior between each model. Finally, we compare calibrated aquifer parameters, including storage properties (S_y and S_s), hydraulic conductivity (K_h and K_v), and vertical anisotropy (K_h/K_v), at every saturated model node and layer in the Sacramento Valley (Text S2 in Supporting Information S1).

4. Results

Observed VGs and head dynamics vary widely across well nests in the Sacramento Valley, with many nests exhibiting significant long-term trends and strong seasonal patterns (Section 4.1). The models vary in their ability to reproduce these behaviors (Sections 4.2 and 4.3), and they differ in aquifer parameterization and simulated budget components (Sections 4.4 and 4.5). Supporting figures and summary metrics are available in Supporting Information S1.

4.1. Observed Head and VG Dynamics

4.1.1. Observed Timeseries Trends

To illustrate typical well nest behavior, we begin by exploring timeseries trends at the COL 10A well nest in Butte Basin (Figure 2). The COL 10A well nest includes four screens spanning 28–239 m below ground surface, with an average screen length of ~ 7 m. Head timeseries at COL 10A exhibit clear seasonal cycles reflecting California's annual wet and dry periods, while also rising in wet years (e.g., 2011, 2017) and declining in droughts (e.g., 2013–2015 and 2021–2023) (Figure 2a). Fluctuations in head are more pronounced at depth where annual changes can exceed 10 m. The deep head timeseries has a variance of 5.3 m^2 at COL 10A, nearly eight times that of the shallow head (0.70 m^2). Most well nests ($>80\%$) share this depth-dependent increase in head variance (Figure S3d in Supporting Information S1). This greater variability at depth likely reflects stronger pumping stresses combined with lower storativity. Near the water table, storage change is governed by specific yield ($S_y \approx 10^{-1}$), while more confined layers at depth exhibit much lower storativity ($S_s \approx 10^{-4}$ – 10^{-6} m^{-1} ; storativity $\approx 10^{-2}$ – 10^{-5} depending on effective layer thickness), which amplifies deep head responses to storage change (Section 4.4; Moss et al., 2013).

Vertical gradients at COL 10A also exhibit clear seasonal and annual cycles (Figure 2b). VGs at COL 10A (and across the basin) are generally largest in magnitude during dry seasons and drought years when layer-pair VG values can exceed 0.1. These large magnitude VGs are due to increased hydraulic stress from seasonal pumping combined with ubiquitous fine-grained interbeds that limit vertical communication, which causes shallow and deep head timeseries to separate under dry-season stress. Notably, VGs are consistently $\gg 10^{-3}$, the typical magnitude of horizontal gradients, which reflects the persistent influence of vertical anisotropy.

At shallow aquifer zones, layer-pair VG timeseries are positive during dry seasons, indicating water levels drop with increasing depth (downward flow potential). But at an elevation of approximately -103 m AMSL, dry-season VGs switch to negative values (upward flow potential). This behavior suggests this depth zone consistently draws flux from above and below during dry seasons, releasing this stress during wet seasons, likely driven by seasonal pumping or strong lateral flux around this depth. More than one-third of the nests studied (>30) show a similar trend (full list in Text S4 in Supporting Information S1).

The observed top-to-bottom VG timeseries (from the shallowest to deepest head measurements) at COL 10A generally shows small positive values, peaking around $\text{VG} = 0.03$, indicating that head generally declines with depth at this location (Figure 2b). However, the top-to-bottom VG drops below zero every wet season prior to 2021, reflecting annual recovery of these semiconfined deep aquifer zones to levels that exceed shallow heads. Such recovery suggests deeper confined zones at this location may be connected to shallow recharge zones potentially located considerable distances away (e.g., at the basin margins, where higher topographic elevation and less confining material may permit greater vertical flow). After 2021, the top-to-bottom VG remains consistently positive, even during wet seasons, indicating a possible shift in system dynamics where deep aquifer layers no longer fully recover, likely reflecting the 2020–2022 drought (the driest 3-year period in California's instrumental record) and the associated increase in groundwater pumping (Medellín-Azuara et al., 2022). Mann-Kendall trend testing confirms a slight but statistically significant increasing trend in observed top-to-bottom VGs at this site (slope: $8.5 \times 10^{-4} \text{ yr}^{-1}$). A similar trend is apparent across the basin: of 81 well nests, 56 show increasing trends, 8 decreasing trends, 16 insignificant or flat trends, and one lacks sufficient data (Table S1 in Supporting Information S1). These results indicate a widespread increase in top-to-bottom VGs, consistent with deeper groundwater levels declining more rapidly than shallow levels.

4.1.2. Observed Spatial Trends

Average shallow head observations, interpreted as a proxy for the water table, generally follow topography: values decline from >120 m AMSL in the north to approximately sea level in the southern domain, consistent with regional groundwater flow from north to south (Figure 3a). Average top-to-bottom VGs are predominantly positive across the Sacramento Valley, indicating downward flow potential through most of the domain (Figure 3d). Zones of upward flow (negative VGs) occur primarily along the north-central valley, with clusters near Corning and between Willows and Colusa. These patterns are consistent with the conceptual regional flow system described by Tóth (1963) and Freeze and Cherry (1979), characterized by widespread downward flow with localized discharge near topographic lows. Steep bedrock slopes shown by Hatch (2020, p. 5) in this region may also contribute to observed upwelling patterns (Alberti et al., 2016). Natural upwelling can be either accentuated, dampened, or reversed by pumping, depending on the relative depths of screened intervals. For example, if a shallow monitoring well screen lies within the main pumping zone, extraction lowers shallow heads which can drive negative VGs (e.g., Figure S24 in Supporting Information S1). Conversely, pumping that draws primarily from deeper zones tends to lower deep heads faster than shallow ones, which can dampen or reverse natural upwelling through time (e.g., Figure S52 in Supporting Information S1).

Large-magnitude VGs (relative to $\sim 10^{-3}$ horizontal gradients) occur throughout the Sacramento Valley. Well nests commonly exhibit 90th percentile layer-pair VGs ≥ 0.1 and deep head variance $\geq 10 \text{ m}^2$ (Figures 3c and 3f). Shallow head variance can exceed 10 m^2 in some areas, particularly near Orland, CA (Figure 3b), but is generally smaller than at depth (Figure S3d in Supporting Information S1). These observations collectively indicate limited vertical hydraulic connectivity, consistent with the hydrogeologic framework of discontinuous coarse-grained lenses embedded in fine-grained materials that restrict vertical flow (Hatch, 2020). Notably, several well nests with large VGs occur in zones where estimated percent coarse exceeds 50% (Faunt et al., 2024, p. 7), especially near Orland. This suggests that large VGs may still develop in relatively coarse-textured zones due to interbedded fine-grained materials that restrict vertical flow.

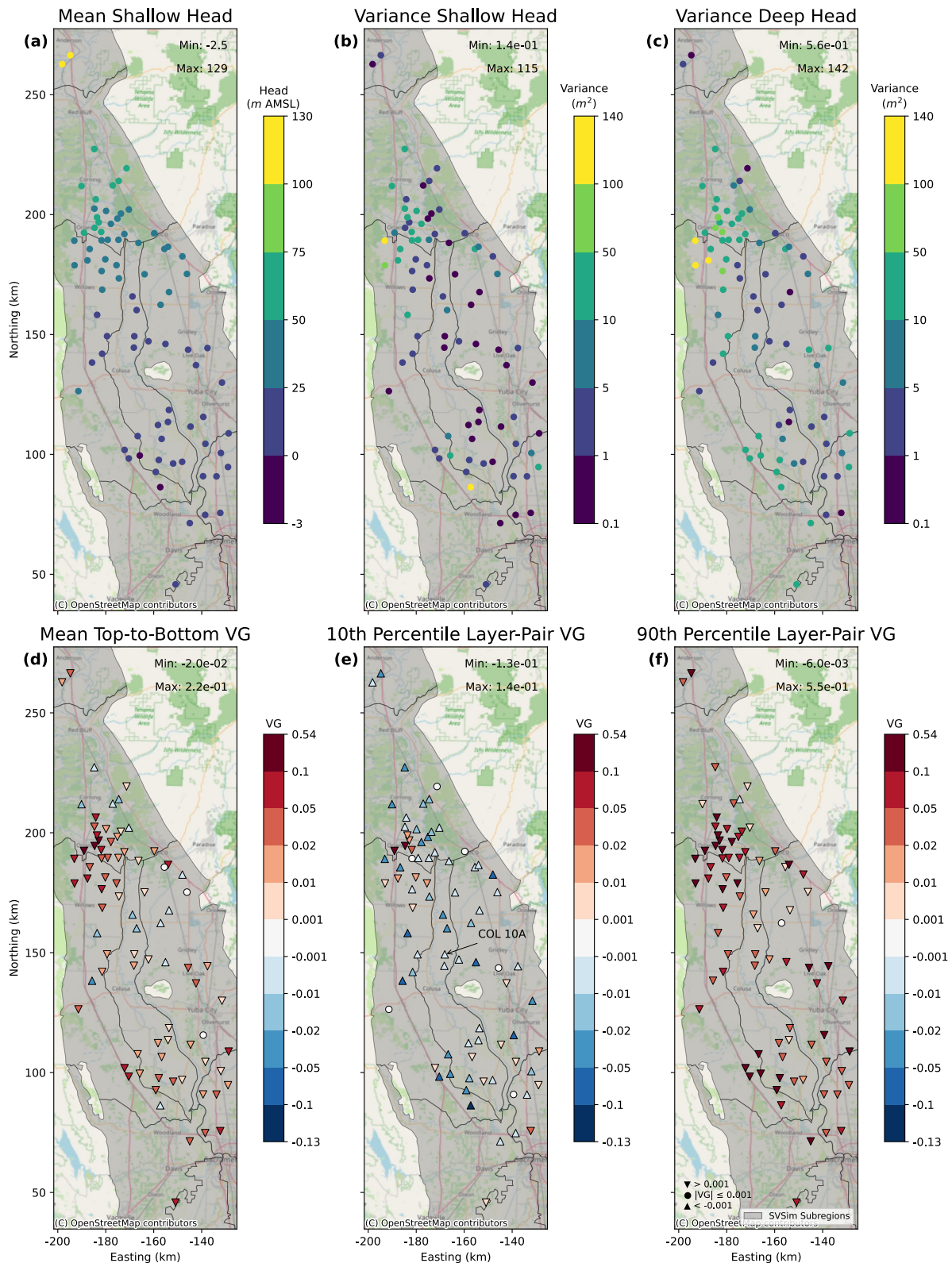


Figure 3. Maps of observed head and VG metrics calculated at each well nest location. (a) Average shallow head values, used to represent mean water table elevation; (b and c) variance of shallow and deep head timeseries, respectively; (d) average top-to-bottom VG value indicating the dominant vertical direction of flow potential through time (negative: upward flow, positive: downward flow); (e and f) the 10th and 90th percentile of layer-pair VGs, respectively, indicating the typical low and high ends of the VG distribution observed at each well nest, calculated from all layer-pair VG values across time (percentile metrics are illustrated in Figure 2).

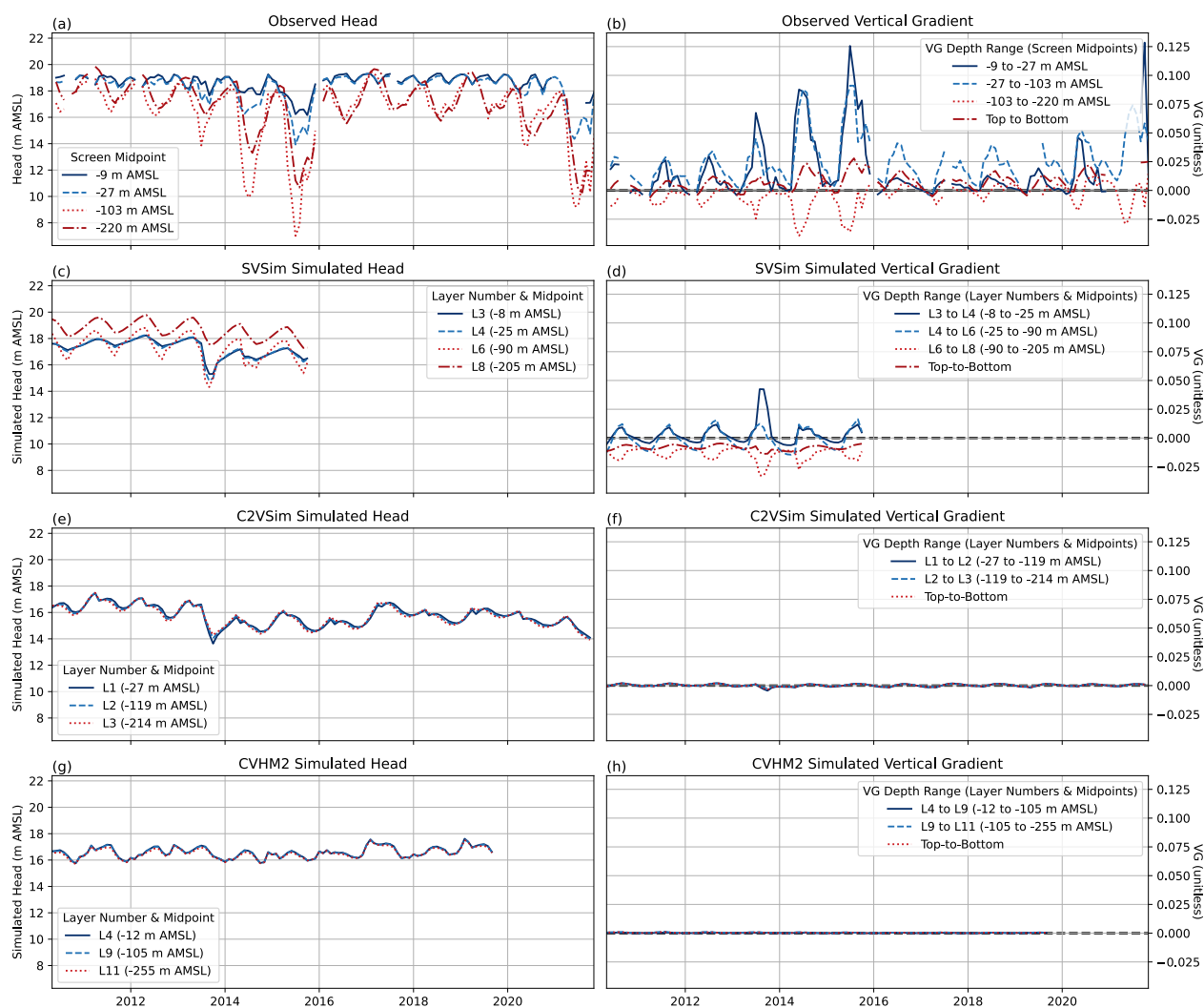


Figure 4. Head (left) and VG (right) timeseries at the COL 10A well nest location from June 2010 to September 2021. From top to bottom: observed data (a and b); SVSim (c and d); C2VSim (e and f); and CVHM2 (g and h).

Small-magnitude VGs (typically within ± 0.02) and low head variance are more common along the valley's central axis, particularly north of Woodland (Figures 3b–3f). While this region has modest percent coarse values (Faunt et al., 2024, p. 7), coarse stream deposits identified in the surficial geologic framework (Scantlebury et al., 2022, p. 8) may contribute to vertically connected pathways that are difficult to resolve in the texture model. Alternatively, reduced VG magnitudes and dampened head variability in this region may reflect the moderating effects of Sacramento River surface water availability, which supports extensive rice farming and reduces seasonal pumping demands (Faunt, 2009; Hatch, 2020).

4.2. Simulated Head and VG Dynamics

4.2.1. Simulated Timeseries Trends

To illustrate differences between each numerical simulation, we begin by comparing head and VG timeseries simulated at the COL 10A well nest location. The SVSim model effectively reproduces overall seasonal and annual head trends observed at COL 10A, though with dampened amplitude (Figures 4a–4d). Simulated annual head fluctuations are about 1 m, with a drop of ~ 3 m in 2013, and simulated deep head variance is much smaller than observed (0.98 vs. 5.3 m²). Simulated layer-pair VGs peak around ± 0.03 during drought, aligning with observed timing but underestimating peak VG magnitudes of -0.03 and 0.1 . Like the observations, VGs are positive at shallow depths and negative at deeper depths during dry seasons. However, SVSim consistently

overestimates deep heads at COL 10A, resulting in negative top-to-bottom VGs. This pattern is uncommon throughout the domain, occurring at only 13 other well nests (listed in Text S4 in Supporting Information S1).

C2VSim simulates similar seasonal drawdown as SVSim at COL 10A, but head variation across layers is minimal: values remain within 1 m of each other through time which produces very small-magnitude VGs (Figures 4e and 4f). This suggests the model behaves mostly as a single, unconfined unit at this location. These trends are largely consistent across observations.

CVHM2 shows the least variation in head through both time and depth at COL 10A, resulting in small-magnitude VGs throughout the record (Figures 4g and 4h). These trends mostly are consistent across the basin, particularly in the north and eastern portions of the valley. However, about a quarter of well nests (~23), mostly on the western side of the basin, show substantial variance in simulated head through time and/or depth (e.g., Figures S58–S61, S63, S69; full list in Text S4 in Supporting Information S1).

While observed top-to-bottom VGs show an increasing trend at COL 10A, only CVHM2 simulates a statistically significant trend at this site, and in the opposite direction (decreasing, slope: $-2.5 \times 10^{-5} \text{ yr}^{-1}$). This local mismatch mirrors broader regional patterns. When observed VG trends are evaluated over the overlapping time intervals as each model's simulation period, around 59% of valid well nests observe increasing trends. In contrast, simulations at only 23%, 24%, and 37% of well nest locations show increasing trends in CVHM2, C2VSim, and SVSim, respectively (Table S1 in Supporting Information S1). While none of the models fully replicate observed behavior, SVSim best captures the regional signal of intensifying VGs.

4.2.2. Simulated Spatial Trends

For all three models, simulated water table elevations in September 2015 exceed 100 m AMSL in the northern portion of the valley and approach sea level in the south, consistent with the general north-to-south flow direction (Figures 5a–5c). In contrast to simulated water table trends, VG patterns vary significantly between models. During September 2015, SVSim shows VGs ≥ 0.01 at ~64% of model nodes, with negative VGs (upward flow) near northern streams (Figure 5f). Large positive VGs (≥ 0.1) occur along valley margins and in the north, with smaller magnitudes along the central axis. These patterns broadly resemble observed VGs (Figures 3d–3f). By contrast, C2VSim simulates mostly small magnitude or negative VGs, with only 6% of nodes showing top-to-bottom VG ≥ 0.01 (Figure 5e). While both C2VSim and SVSim simulate negative VGs near streams, this trend is more widespread in C2VSim (32% vs. 17% of nodes, respectively).

The CVHM2 model also simulates small-magnitude VGs over a significant portion of the study domain: VG magnitudes < 0.001 are simulated over ~44% of the region in September 2015 (Figure 5d). Approximately 14% of the model area shows VGs ≥ 0.01 , primarily along the western half of the valley. Small VG magnitudes < 0.001 are primarily simulated along the north and eastern portions of the valley (Figure 5d), corresponding with large simulated K_v values (often $> 0.1 \text{ m d}^{-1}$; Figure S2 in Supporting Information S1) which likely represent coarse fluvial deposits from the Sierra Nevada Mountain range. Negative VGs cover ~10% of the valley, mostly in the southwest where the simulated water table lies below sea level (Figures 5a and 5d).

These contrasting VG patterns indicate markedly different spatial distributions of recharge and discharge among the three models, reflecting fundamental differences in how each represents vertical connectivity and SW-GW exchange. Of the three, SVSim's VG distribution most closely matches the Tóth (1963) and Freeze and Cherry (1979) conceptual flow system inferred from the observations (widespread downward flow with localized discharge near topographic lows; Section 4.1), whereas the predominantly small-magnitude or negative VGs in CVHM2 and C2VSim depart from this conceptual model. Because regional models are often used to delineate broad recharge and discharge zones, these discrepancies highlight how parameterization of vertical flow can strongly influence conceptual interpretations of basin-scale groundwater circulation.

4.3. Model Performance

Statistically, each model performed similarly in simulating both shallow and deep heads at well nests. For all models, the linear regression between simulated and observed heads yielded Pearson $r \geq 0.93$ with slope = 1.0 ± 0.1 (Figures 6a–6c). SVSim had the least overall mean error in both the shallow and deep head timeseries (−1.8 and 0.28 m respectively), while CVHM2 had the greatest overall mean error in head (2.7 and 6.2 m respectively) (Figures 6a–6c). Each model tends to overestimate head, evidenced by positive mean errors, with the exception of

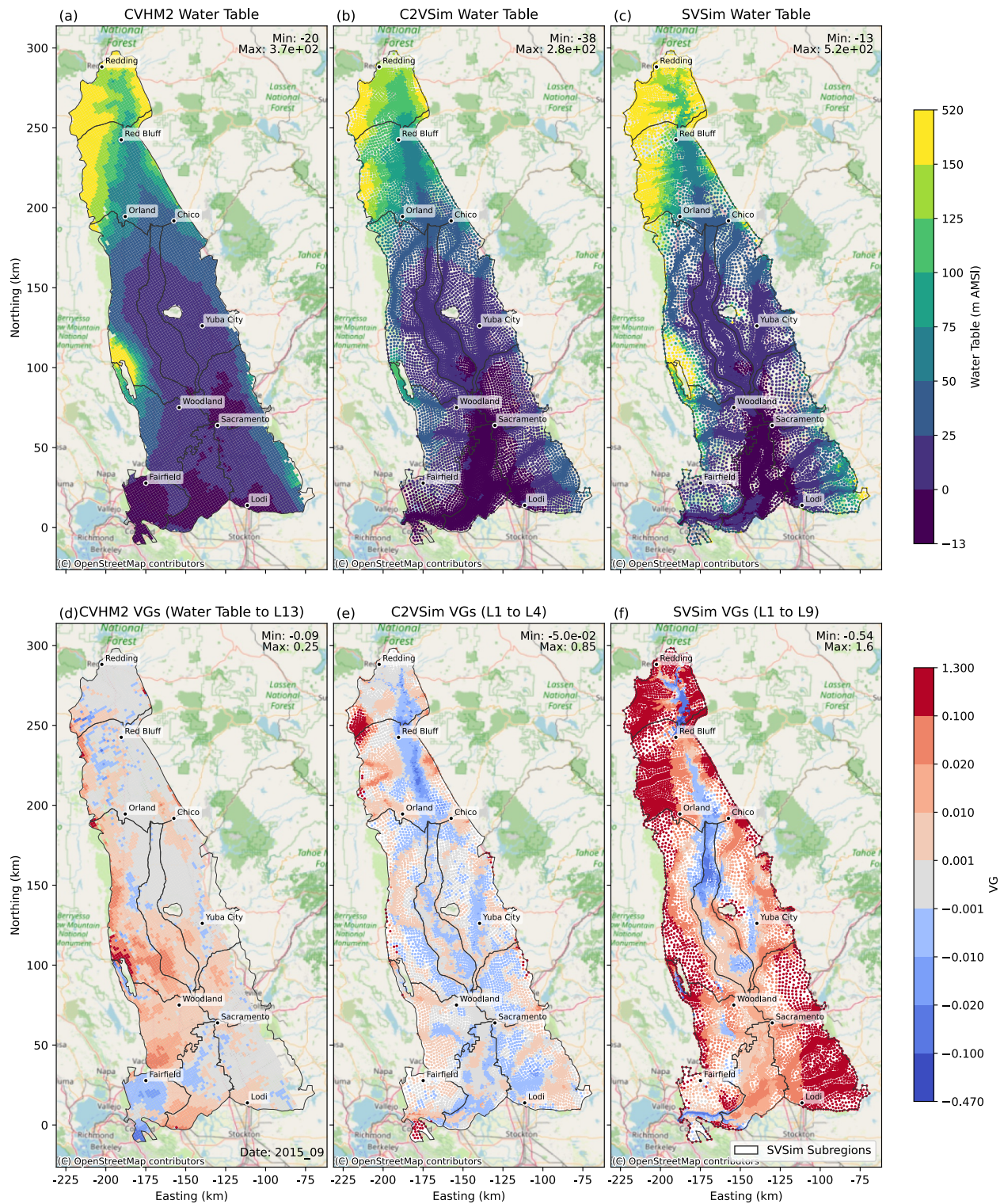


Figure 5. Simulated water table (top row) and top-to-bottom VGs (bottom row) in September 2015 at each node in CVHM2 (left), C2VSim (center), and SVSim (right). Stipple pattern in C2VSim and SVSim maps occurs because simulated head and VG values are plotted at model nodes, which are spaced more densely around surface water bodies. CVHM2 results are visualized over model cells.

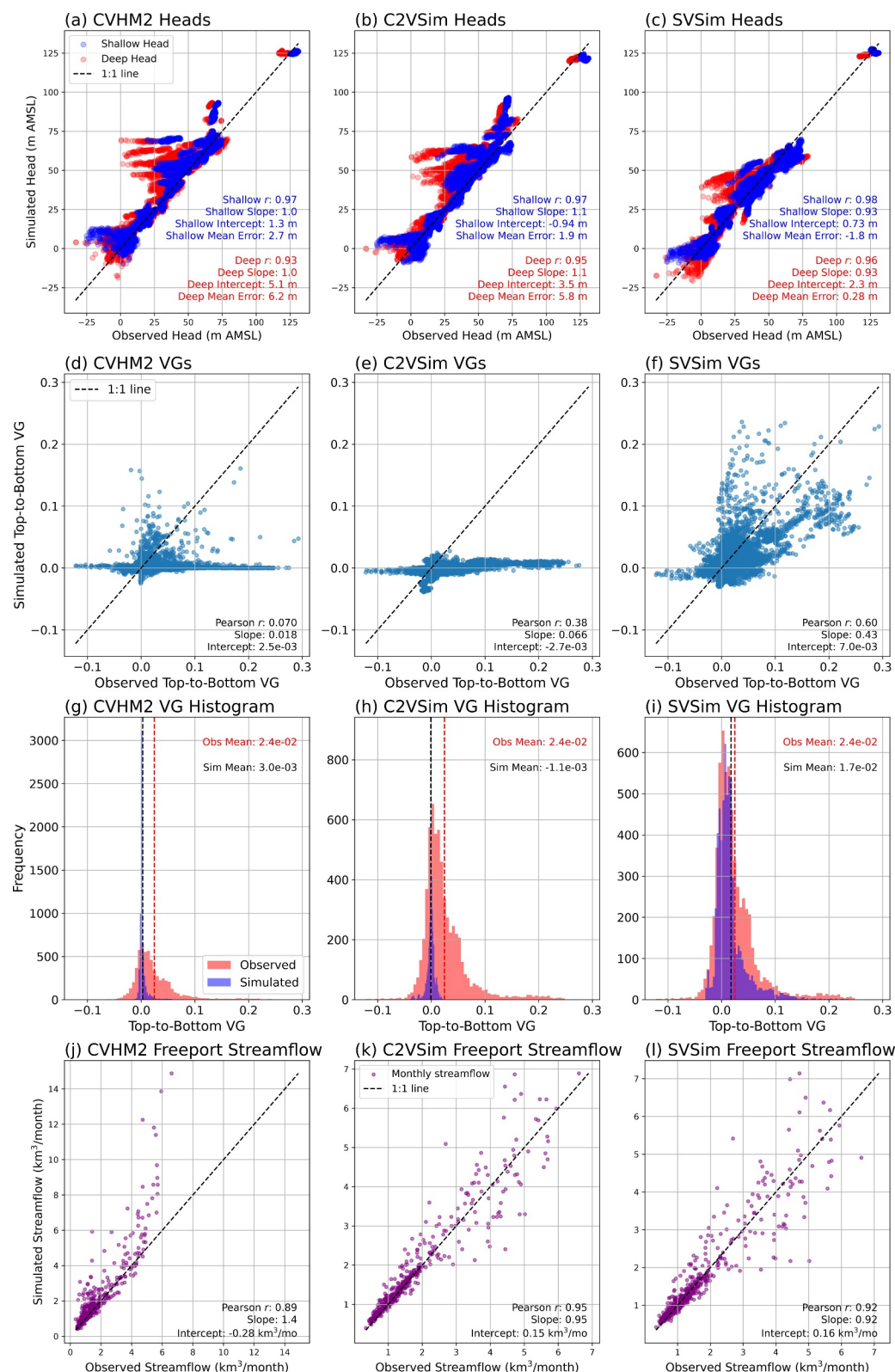


Figure 6. Scatter plots comparing observed versus simulated shallow and deep head timeseries (a–c; top row) and top-to-bottom VGs (d–f; second row). Histograms (g–i; third row) compare the distributions of observed versus simulated top-to-bottom VGs. The bottom row (j–l) compares observed versus simulated monthly streamflow in the Sacramento River at Freeport. Columns correspond to CVHM2 (left), C2VSim (middle), and SVSim (right).

shallow head timeseries in the SVSim model (Figures 6a–6c). In particular, CVHM2 and C2VSim produced significant overestimates of deep heads in the 40–75 m AMSL range, likely due to excessive connection between the shallow and deep portions of those models.

Of the models analyzed, SVSim shows the strongest relationship between simulated and observed monthly top-to-bottom VGs ($r = 0.60$, slope = 0.43; Figure 6f). It simulates a wide range of VGs, changing significantly through time, space, and with depth, consistent with spatial and temporal trends in SVSim (Figures 4d and 5f). Yet SVSim often fails to reproduce extreme VG values (Figure 6i), likely due to vertical discretization limits and unmodeled delayed drainage in fine-grained sediments (Bedekar, Cayar, et al., 2021).

By contrast, simulated top-to-bottom VGs in CVHM2 show no relationship with observed values ($r = 0.07$, slope = 0.018; Figure 6d), and are tightly clustered around zero (Figure 6g). This tight clustering is consistent with simulated timeseries at the COL 10A well nest, where minimal head variation with depth results in small magnitude VGs (Figure 4g). Furthermore, the spatial analysis for September 2015 confirms that a significant portion of the domain simulates small-magnitude VG values (Figure 5d) reinforcing the conclusion that CVHM2 underrepresents the natural variability in VGs.

The C2VSim model has a similar bias toward underestimating VG magnitudes relative to observations. C2VSim shows a weak relationship between simulated and observed VGs ($r = 0.38$, slope = 0.066; Figure 6e). The distribution of simulated top-to-bottom monthly VGs is significantly narrower than observed (Figure 6h). C2VSim has a slightly negative mean top-to-bottom VG value of -1.1×10^{-3} , compared to the observed mean of 2.4×10^{-2} (Figure 6e), indicating a bias toward simulating negative VGs (upward flow). Spatial analysis of VGs in C2VSim also support the existence of a negative-VG bias, where negative VGs were simulated in a significant portion of the domain during September 2015 (Figure 5e). This pattern reflects a tendency for deeper aquifer layers to have higher simulated heads than shallow ones.

All three models reproduce observed monthly streamflow at Freeport on the Sacramento River comparably well (Figures 6j–6l; Pearson $r = 0.89$ – 0.95 , slopes 0.92–1.4), with C2VSim tightest and CVHM2 most scattered, the latter tending to overestimate higher flows (slope = 1.4). This is a comparison to total streamflow at the basin outlet, not the SW-GW exchange term itself, and this basin-integrated signal does not by itself constrain the vertical flow distribution the VG comparison reveals. Direct validation of the exchange term would require reach-scale stream gain/loss estimates, which are not available basin-wide (Section 5.3).

4.4. Model Aquifer Parameters

Across all three models, aquifer storage parameter distributions are broadly similar. Median nodal S_y values are about 10%–11% for all models. Median S_s values are roughly an order of magnitude smaller in CVHM2 (2.0×10^{-6} m⁻¹) than in SVSim and C2VSim (8.3×10^{-5} and 6.3×10^{-5} m⁻¹ respectively; Figure S1 in Supporting Information S1). By contrast, hydraulic conductivity and vertical anisotropy vary substantially among models.

CVHM2 has the smallest median K_h (1.6 m d⁻¹) and the largest median K_v (0.21 m d⁻¹) when evaluated across all active nodes and layers (Figure S1 in Supporting Information S1). As a result, CVHM2 simulates the smallest anisotropy values, on the order of 1–10 over most of the domain (nodal median: 9.7). Model nodes with small vertical anisotropy values in CVHM2 are concentrated in the northern portion of the valley and along the eastern side (Figure S2 in Supporting Information S1). These same areas simulated small-magnitude top-to-bottom VGs during September 2015 (Figure 5d).

SVSim has the highest median K_h values (10 m d⁻¹), lowest median K_v values (0.028 m d⁻¹), and greatest vertical anisotropy (median: 330; Figure S1 in Supporting Information S1). Anisotropy is high (>100) across most of the basin (Figure S2 in Supporting Information S1), rather than dropping to the low values seen over large parts of CVHM2 and C2VSim, and SVSim correspondingly produces large-magnitude VGs across most of the domain (Figure 5f). Bedekar, Cayar, et al. (2021) report that all SVSim aquifer parameters had high sensitivities and narrow confidence intervals, indicating relatively low uncertainty in the estimated values.

C2VSim has intermediate K_h , K_v , and vertical anisotropy values (medians of 6.8 m d⁻¹, 0.081 m d⁻¹, and 67, respectively; Figure S1 in Supporting Information S1). Large K_v values (often ≥ 0.1 m d⁻¹) in C2VSim are concentrated along the central axis of the valley, roughly coincident with the Sacramento River, where simulated

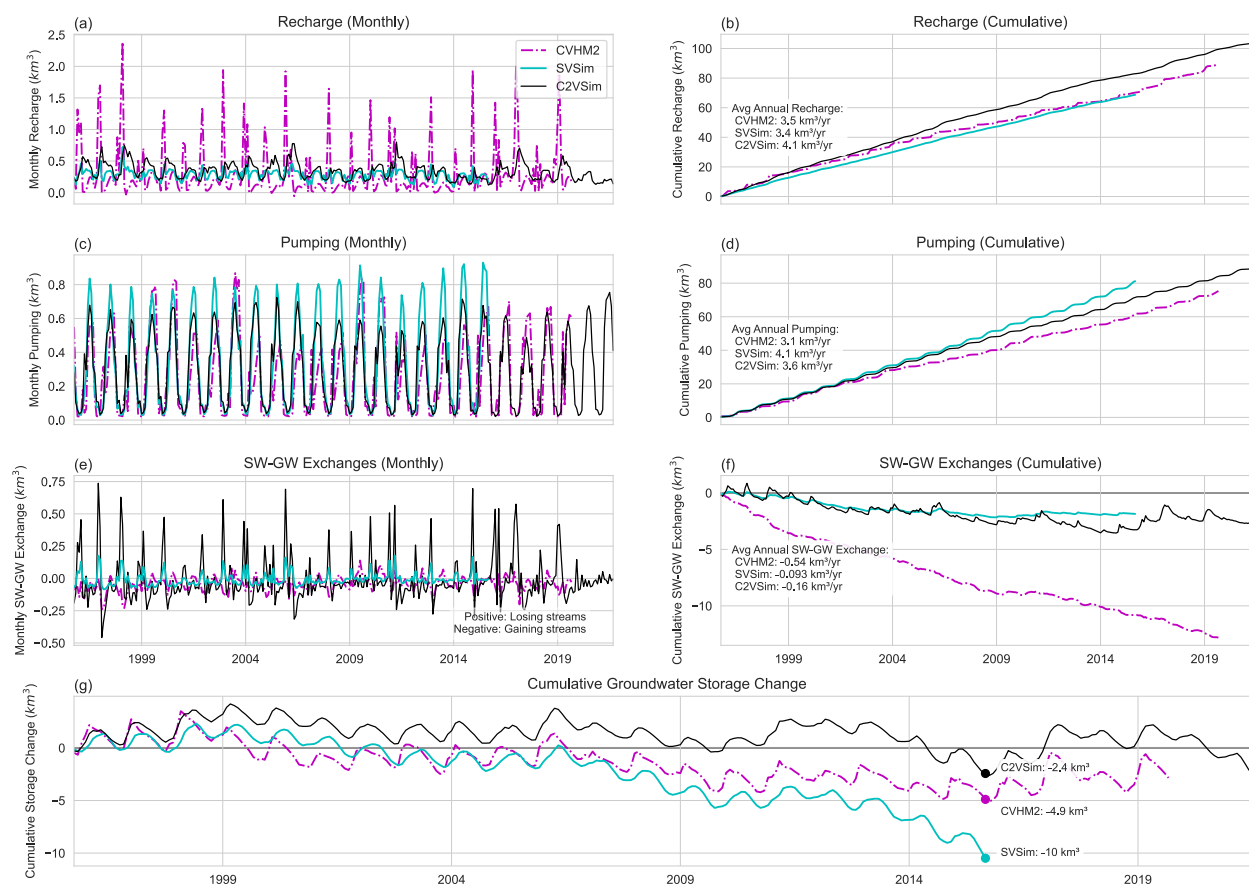


Figure 7. Monthly (a, c, e; left) and cumulative (b, d, f; right) groundwater recharge, pumping, surface water–groundwater exchanges, and (g; bottom) cumulative groundwater storage change (October 1995 to September 2015) compared between each model.

top-to-bottom VGs with small magnitudes or upward flow potential are typical during September 2015 (Figure 5e; Figure S2 in Supporting Information S1).

4.5. Simulated Groundwater Budgets

We evaluate key components from each model's simulated groundwater budget over a shared 20-year period from October 1995 to September 2015 (summarized in Figure 7). A primary difference among models is in their simulation of recharge and SW-GW exchange. CVHM2 simulates a substantially larger contribution of groundwater to surface water than the other two models, with an annual average of $-0.54 \text{ km}^3 \text{ yr}^{-1}$, compared to $-0.16 \text{ km}^3 \text{ yr}^{-1}$ in C2VSim and $-0.093 \text{ km}^3 \text{ yr}^{-1}$ in SVSim (negative values denote groundwater discharge to streams; Figure 7f). While C2VSim and SVSim show comparable cumulative SW-GW exchange, C2VSim exhibits monthly exchange fluxes that are typically two to three times larger in magnitude than the other models (Figure 7e). Similarly, C2VSim simulates about 20% more cumulative recharge (averaging $4.1 \text{ km}^3 \text{ yr}^{-1}$) than CVHM2 and SVSim (averaging 3.5 and $3.4 \text{ km}^3 \text{ yr}^{-1}$, respectively) (Figure 7b). However, CVHM2 shows peak monthly recharge fluxes that are two to three times larger than the other models (Figure 7a).

Differences in simulated SW-GW exchange and recharge, combined with differences in pumping associated with land-use representation, lead to divergent groundwater storage outcomes. SVSim shows the highest average annual pumping ($4.1 \text{ km}^3 \text{ yr}^{-1}$), followed by C2VSim ($3.6 \text{ km}^3 \text{ yr}^{-1}$) and CVHM2 ($3.1 \text{ km}^3 \text{ yr}^{-1}$) (Figure 7d). As a result, SVSim simulates the greatest cumulative groundwater depletion (-10 km^3), consistent with its higher pumping and lower recharge, while CVHM2 and C2VSim show more modest declines of -4.9 km^3 and -2.4 km^3 , respectively (Figure 7g).

5. Discussion

Our analysis shows that incorporating observed VGs during hydrogeologic parameterization of a regional-scale groundwater model in a typical clastic sedimentary aquifer system can increase simulated vertical anisotropy and improve model performance. The SVSim model, calibrated with VG data (head differentials), simulated the highest vertical anisotropy values (median K_h/K_v of 330), the smallest mean head error and strongest correlation with observed VGs. By contrast, the CVHM2 and C2VSim models show smaller vertical anisotropies (nodal medians of 9.7 and 67, respectively) and predominantly underestimated VGs. These differences in simulated VGs and vertical anisotropy correspond with differences in the simulated recharge and SW-GW exchanges, even though all three models reproduce observed shallow and deep heads comparably well. In addition to the modeling advantages of incorporating VG data, our study also highlights a concerning pattern in the Sacramento Valley: 56 out of 81 well nests observe increasing VGs over time. This trend indicates rising downward flux potential (or weakened upward flux), which could shift flow in ways that increase surface water losses and expose deeper aquifers to contamination risk (as discussed in Section 5.2). Our findings underscore the value of VG data not only for improving model performance, but also for detecting and responding to evolving system behavior.

5.1. Application of VGs to Groundwater Modeling

Vertical anisotropy is often systematically underestimated in regional-scale models of stratified sedimentary aquifers, a bias documented in a recent meta-analysis (Shepley, 2024) and attributed largely to an overreliance on vertically insensitive calibration targets, such as laterally distributed head observations that respond only weakly to changes in vertical anisotropy. Our findings support this interpretation and further demonstrate the consequences of this oversight, namely, that models calibrated without vertically sensitive data can simulate substantially different vertical flow behavior and groundwater budgets while appearing similarly well-calibrated to head. However, this consequence has only been sparsely explored in the scientific literature thus far (e.g., Fogg, 1986; Tanachaichoksirikun et al., 2020). While all three models in our analysis produce statistically acceptable fits to both shallow and deep heads and observed streamflow, they nonetheless simulate fundamentally different vertical flow behavior, resulting in strikingly different spatial patterns of recharge and discharge, and contrasting representations of recharge and SW-GW exchange. These budget differences are consequential, because regional models are among the only practical tools for estimating SW-GW exchange and recharge at the regional scale, where these fluxes cannot be measured directly. SVSim represents a substantially different conceptual flow field than the other two models, with broad recharge along the basin margins and discharge along the lower-topography central axis, consistent with the observed VG patterns and with classical regional flow concepts (Freeze & Cherry, 1979; Tóth, 1963). In addition, the SVSim documentation reports that all aquifer parameters were highly sensitive and had correspondingly narrow confidence intervals, indicating relatively low post-calibration parameter uncertainty (Bedekar, Cayar, et al., 2021). Together, the agreement with observed VG patterns, the physically expected recharge-discharge configuration, and the low post-calibration parameter uncertainty make SVSim the most defensible of the three models for representing these vertical flow terms. In addition, SVSim's stronger anisotropy may also indicate a less uncertain overall budget representation: because greater vertical anisotropy limits vertical hydraulic connectivity, substantial changes in pumping or recharge would likely distort the simulated head and VG signal more noticeably in SVSim than in the more vertically connected CVHM2 and C2VSim models. For example, additional recharge in SVSim would likely bias shallow heads high, and additional pumping would likely deepen the simulated deep head signal, whereas similar budget changes in the other two models could be absorbed with comparatively little effect on the integrated head signal. Explicitly testing which model is most accurate is identified as a priority for future research (Section 5.3).

Beyond the insensitivity of head targets, vertical anisotropy may also be underestimated in regional groundwater models because within-layer anisotropy is commonly overlooked during hydrogeologic conceptual model development, as it is difficult to estimate *a priori* (Anderson et al., 2015). For example, vertical anisotropy is scale-dependent (discussed in Section 2.2) and may be strongly influenced by fine-grained interbeds that can impose substantial vertical flow resistance and confinement, even in the absence of a laterally continuous aquitard layer. Yet these interbeds are rarely evident in regional hydrostratigraphic frameworks. As a result, parameterizing layer-scale vertical anisotropy is often deferred to calibration, where it may remain weakly constrained or effectively uncalibrated in practice (Anderson et al., 2015; Shepley, 2024). Our results illustrate that identifying vertical anisotropy early in conceptual model development is critical because doing so justifies the use of vertically sensitive calibration targets, such as VGs, to constrain anisotropic flow behavior during calibration.

Although all three models evaluated acknowledge vertically anisotropic conditions in their conceptual descriptions, only SVSim, which incorporated VGs as calibration targets, preserved this conceptual model in its final parameterization. These findings highlight the need to explicitly evaluate vertical anisotropy during initial hydrogeologic conceptual modeling and to revisit the conceptual model throughout calibration, as emphasized by Hill and Tiedeman (2007), to ensure consistency with anticipated vertical flow processes.

Our study may appear to conflict with recent findings from idealized numerical experiments (Bakker & Bot, 2025) that suggest stratified sedimentary aquifer systems can be appropriately represented using upscaled vertical anisotropy values on the order of 10. We clarify how their base assumption that hydraulic conductivity values vary by no more than two orders of magnitude within an upscaled model layer (e.g., varying from clay to fine sand or coarse sand to gravel) does not hold in clastic sedimentary aquifer systems like the Sacramento Valley, where coarse- and fine-grained materials span a far wider conductivity range and can occur within a single model layer. For example, in the American–Cosumnes fan area of the Sacramento Valley, vertical hydrofacies mean lengths are on the order of 1 m, while facies hydraulic conductivities span roughly five orders of magnitude (Maples et al., 2019), far exceeding the two-order range assumed by Bakker and Bot (2025). There is ample data and geologic evidence indicating that hydraulic conductivity in clastic sedimentary deposits ranges over four to seven orders of magnitude (Fogg & Zhang, 2016). In such environments, simulating upscaled anisotropy values $>>10$ may not only be justifiable (as shown in this study) but necessary to reproduce observed system dynamics.

The influence of vertical flow parameterization is also consequential for transport models. VGs enable vertical transport through preferential flow conduits, both anthropogenic (e.g., poorly sealed boreholes) and natural (e.g., incised valley fill deposits, stacked sand channels) (Fogg, 1986; Mukherjee et al., 2007; Nunes et al., 2021). Strong VGs can drive flow through laterally extensive low- K_v media (Chakraborty et al., 2022; Pauloo et al., 2021; von Brömssen et al., 2014). Seasonal or long-term shifts in VGs can govern which geochemically distinct sources contribute recharge at a given time, reshaping flow pathways and influencing reactive transport processes (Bexfield & Jurgens, 2014; Giambastiani et al., 2013; Z. Zhang et al., 2018). Together, these findings highlight a critical research opportunity: to develop and test transport models that explicitly represent temporally varying VGs and anisotropy, allowing direct quantification of their effects on solute and redox evolution.

The implications of misrepresenting vertical flow parameters extend beyond the scope of this study. Although our focus is on flow dynamics, many processes, including transport, subsidence, and lateral inflows, are theoretically shaped by VGs and anisotropy. Given the global prevalence of highly anisotropic conditions in stratified sedimentary aquifers (Runkel & Meyer, 2025), future work should examine how improved parameterization of vertical flow and anisotropy affect these coupled processes. Testing such approaches across diverse basin types would help clarify the broader implications of anisotropy and VG misrepresentation for regional groundwater modeling and management.

5.2. Implications for Groundwater Management

The trend of increasing VGs observed throughout the Sacramento Valley is a concern for the future management of groundwater in this region. Similar patterns have been documented in basins around the world, where anthropogenically driven downward flow gradients reflect reorganized regional flow systems, increased recharge from surface sources, and accelerated vertical transport of contaminants. For example, intensified downward flow VGs have been linked to the reversal of streamflow from gaining to losing conditions (e.g., Auterives et al., 2011; Green et al., 2024; Kwon et al., 2020; Thiros et al., 2010; Z. Zhang et al., 2018) and the contamination of deep aquifers (e.g., Currell et al., 2010; Lawrence et al., 2000; Mukherjee et al., 2007; Wang, 2018). Accordingly, the observed trend of increasing VGs in the Sacramento Valley may be an early indicator of unsustainable groundwater use, one that precedes many of the adverse outcomes regulated under California's Sustainable Groundwater Management Act, including stream depletion and groundwater quality degradation. Stream depletion is particularly critical in this context: overall, streams in the Sacramento Valley are gaining. If VGs continue to intensify, this regime could theoretically reverse, potentially reducing baseflow and impacting surface water exports from the Sacramento Valley to the agriculturally and economically vital regions of the San Joaquin Valley and Southern California.

The strong vertical anisotropy across the Sacramento Valley indicates that semiconfined conditions with low coefficients of storativity prevail, except perhaps within 20–30 m below the water table. Moreover, these conditions together with the horizontally extensive sand and gravel body network means that perturbations in

hydraulic head can travel great distances laterally, and relatively rapidly due to the high hydraulic diffusivity (e.g., Ferris et al., 1962; Theis, 1940). The evidence provided by VG thus has key management implications: actions in one part of the basin may affect conditions far away, reinforcing the need for regionally coordinated planning (Pederson, 2000). For example, recharge in one part of such a system can raise groundwater levels multiple kilometers away from the recharge site (Maples et al., 2019). Furthermore, in SVSim, Bedekar, Durbin, and Bond (2021) showed that simulated pumping can produce stream depletion in channels located tens of kilometers from the pumping wells. In addition, regionally anisotropic conditions coupled with large VGs may concentrate preferential flux through zones of high vertical connectivity. Such zones may be indicated by small-magnitude VGs surrounded by larger ones, a pattern that reflects high vertical conductivity under sustained vertical flux (Fogg, 1986), and can be especially important to recognize and manage (e.g., Guo et al., 2023; Knight et al., 2022; Weissmann et al., 2004; Y. Zhang et al., 2018).

Finally, our study highlights the value of continuously monitored multilevel head data for understanding and properly modeling groundwater system dynamics. Well nests provide direct observations of depth-dependent responses to stress, yet they remain uncommon across most basins (Shepley, 2024). Where multilevel data are unavailable, VGs can be estimated using laterally distributed single-depth wells, for example, by regressing head against position to isolate the vertical component (Fogg et al., 1991; Jasechko & Perrone, 2020). These methods depend on accurate well construction records, including perforation elevations and screen lengths (information that groundwater managers can help ensure is recorded in machine-readable formats). Expanding the use of continuous head monitoring, whether through multilevel installations or single-screen wells, offers a cost-effective path to better capture the complex, 3D flow dynamics that govern stratified aquifer systems globally (Bhalla et al., 2025).

5.3. Study Limitations and Further Research

The spatial coverage of well nest monitoring stations presents a key limitation in this study. Well nests analyzed are unevenly distributed across the domain, which may bias spatial comparisons of observed VG behavior to model outputs. For example, the southwestern Sacramento Valley, where CVHM2 simulates higher anisotropy and stronger VGs, is underrepresented, so our analysis may underestimate VG performance in that model. However, this underrepresentation is unlikely to reverse our conclusions, since CVHM2 simulates small-magnitude VGs across a significant portion of domain, departing from the observed VG pattern (Figures 3d and 5d). This limitation is more pronounced at the Central Valley scale: at this time, continuously monitored data provided by the California Department of Water Resources do not include any multilevel observations in the severely overdrafted southern San Joaquin Valley. In this region, CVHM2 and C2VSim explicitly simulate aquitard layer(s) (the Corcoran Clay) which may drive strong VGs across that layer regardless of the upscaled anisotropy simulated in the surrounding aquifers. Nevertheless, sedimentary deposits in the San Joaquin Valley other than the Corcoran Clay typically consist of 50%–80% fine-grained materials, similar to the Sacramento Valley deposits (e.g., Maples et al., 2019; Weissmann et al., 2002), indicating significant vertical anisotropy likely also characterizes much of the region. Multilevel head data provided by the USGS in the San Joaquin Valley (e.g., Everett et al., 2020) could support future VG analysis there, although no regional model of that area has been calibrated with VG data, so it does not offer the with-versus-without-VG comparison that is the focus of this study.

A simplification in our analysis is using midpoint matching to assign model layers to well screens (discussed in Text S1.1 in Supporting Information S1). A transmissivity-weighted approach could reduce head error, particularly for long screens, but could also dampen simulated VG dynamics. Similarly, previous work shows that well nest systems with shorter perforations (~1 m) may be more effective at identifying hydrogeologic unit boundaries and parameterizing vertical anisotropy than the longer-screen systems used in this study, which may yield vertically blended head signals (Meyer et al., 2014, 2023; Runkel et al., 2018). However, because the key issue identified in this study is the insufficient variability of simulated VGs compared with observations, neither refinement is likely to alter our overall conclusions.

Another concern is whether VG behavior observed at well nests is driven by local-scale pumping or recharge variability not resolved by regional-scale models. However, SVSim's reproduction of observed VG dynamics shows that these trends are not too local in scale for a regional model to resolve. Additionally, prior studies show that well nest data often reflect aquifer properties and flow patterns relevant to regional-scale flow (Meyer

et al., 2016; Runkel et al., 2018). Importantly, greater vertical anisotropy increases confinement and reduces effective storativity, which raises hydraulic diffusivity and allows head fluctuations to travel rapidly over large lateral distances (Fulton & Brodsky, 2023; Theis, 1940). As a result, head and VG measurements at a single well nest can reflect conditions across a wide area, giving them regional significance in laterally continuous sedimentary systems. Further research in the Sacramento Valley could explore the lateral continuity of patterns observed in well nests, similar to Fisher and Twining (2011) and Meyer et al. (2016).

We acknowledge that CVHM2 was primarily calibrated for the San Joaquin Valley and therefore may not accurately represent vertical flow dynamics in the Sacramento Valley (Section 3.3.1). Yet our analysis highlights that a simple modification to the calibration of CVHM2 (i.e., incorporating VGs as calibration targets) may improve the performance and could change simulated vertical fluxes and groundwater budgets.

A limitation of comparing three independently developed models is that the comparison cannot vary anisotropy in isolation; the models differ in grid, layers, simulation period, and forcing (Table 1). Yet several lines of evidence suggest that calibration choices, rather than these structural differences, are the primary driver of the divergence in simulated anisotropy. Most directly, spatial resolution does not explain the contrast: CVHM2 has vertical discretization comparable to SVSim yet produces the weakest VGs of the three. In addition, all three models are built on the same texture field and power-law upscaling framework; their K_h/K_v fields therefore differ not because of underlying geologic inputs but because of how the power-law averaging parameters were calibrated. Finally, SVSim, the only model calibrated to vertical head differences, is the only model with strong basin-wide anisotropy and good agreement with observed VGs, whereas the two models calibrated primarily to heads are weakly anisotropic and reproduce VGs poorly. Budget differences among the models may not be solely attributable to differences in vertical anisotropy, as simulated budgets also reflect divergent land- and water-use forcing. However, anisotropy constrains how much pumping and recharge a model can accommodate while still matching observed heads. Therefore, the budget from the model that best reproduces observed heads and VGs is likely better constrained than the others.

Further research could explicitly examine how including VGs in aquifer parameter calibration influences model outcomes, as in Fogg (1986) and Tanachaichoksirikun et al. (2020). A cleaner analysis would recalibrate one model with and without VG targets, holding grid, forcing, and structure fixed—isolating the effect of VG calibration on anisotropy and on the budget without the confounds that separate CVHM2, C2VSim, and SVSim (Table 1). Pairing it with independent flux constraints (e.g., seepage or baseflow estimates) would test whether stronger anisotropy yields a more faithful budget, not just better-fitting VGs. However, given the substantial modification of the Sacramento Valley's river system, and extreme spatial and temporal variability in SW-GW exchange in systems like this, existing point- or reach-scale methods are unlikely to constrain a basin-wide budget on their own; new approaches capable of resolving baseflow and SW-GW exchange variability continuously across space and time would likely be needed to make this comparison meaningful. A related open question is how heavily VG targets should be weighted relative to traditional head targets. Because vertical head differences are far smaller in magnitude than lateral head variations, VGs likely need greater weight to contribute meaningfully to the objective function rather than being overwhelmed by distributed head observations. Weighting observation types to contribute roughly equally, as documented for CVHM2 (Faunt et al., 2024), offers one approach, though the optimal balance likely depends on the aquifer system and data density. Such experiments could also clarify which modeling goals (e.g., water supply vs. transport) require detailed vertical parameterization. For example, models developed for water supply planning—rather than solute transport—in settings with thin aquifers, small-magnitude VGs, highly permeable strata, and/or well-defined aquifer–aquitard layering may be adequately represented by weakly anisotropic conceptualizations of groundwater flow (Bakker & Bot, 2025; Shepley, 2024). Lastly, while this study focuses on clastic sedimentary aquifers, VGs are increasingly used to characterize a range of heterogeneous aquifer systems that warrant further exploration (e.g., Runkel & Meyer, 2025).

6. Conclusions

We used 81 continuously monitored well nests in California's Sacramento Valley to evaluate the spatial and temporal variability of vertical hydraulic gradients and to compare their observed behavior against three publicly available regional groundwater models. Observed VGs were dynamic across space, depth, and time, often exceeding 0.1 during dry seasons and droughts. Of the three models evaluated, only SVSim, which incorporated

data comparable to VGs in calibration, sufficiently reproduced observed VG behavior. SVSim also exhibited the highest vertical anisotropy (median $K_h/K_v = 330$) and best performance in matching head trends. In contrast, CVHM2 and C2VSim, calibrated primarily to head, systematically underestimated VG magnitudes and variability. These differences in vertical flow representation correspond with differences in simulated water budgets. SVSim simulated less surface water-groundwater interaction and, consistent with its lower recharge and higher pumping, 2–4 times (-10 km^3 , vs. -4.9 and -2.4 km^3) more cumulative groundwater loss over a shared 20-year simulation period than the other models. These discrepancies highlight a critical insight: models that appear similarly well-calibrated based on head observations may simulate fundamentally different internal flow behavior, especially when vertical flow is under-constrained.

Our findings point to a clear path forward. In clastic sedimentary aquifer systems, vertically differentiated head data like VGs provide critical constraints for conceptualization and calibration workflows. Doing so enables more physically consistent representations of vertical flow and aquifer connectivity, in turn likely reducing uncertainty in simulated budgets. As 3D monitoring networks expand and continuous sensors become more accessible, leveraging VGs to enhance vertical flow representation is not just a refinement—it's a necessary step toward more physically realistic, decision-relevant models.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

All well nest station data and associated groundwater level time series are publicly available from the California Natural Resources Agency (CNRA) Open Data Portal (CNRA, 2025). Regional groundwater model outputs used in this study are also publicly available. The C2VSim (DWR, 2025) and SVSim (DWR, 2022) models are hosted on the CNRA Open Data Portal, and CVHM2 is available through the U.S. Geological Survey ScienceBase catalog (Traum et al., 2024).

Acknowledgments

This work is part of the Secure Water Future project and was supported by Agriculture and Food Research Initiative Competitive Grant 2021-69012-35916 from the USDA National Institute of Food and Agriculture. We thank lab mates in the Dahlke research group for early manuscript reviews, and we are especially grateful to Guobiao Huang, Behrooz Etebari, Vivek Bedekar, Scott Boyce, and Jon Traum for their generous feedback on model processing and interpretation.

References

- Alberti, L., Colombo, L., & Francani, V. (2016). A method for identifying and preventing the diving plumes in alluvial aquifers. *Italian Journal of Engineering Geology and Environment*, 2, 41–50. <https://doi.org/10.4408/IJEGE.2016-02.O-04>
- Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2015). *Applied groundwater modeling: Simulation of flow and advective transport* (2nd ed.). Elsevier, Academic Press.
- Arizona Department of Water Resources. (2023). Groundwater flow model of the phoenix active management area (modeling report no. 28). Retrieved from <https://azmemory.azlibrary.gov/nodes/view/326284>
- Auterives, C., Aquilina, L., Bour, O., Davranche, M., & Paquereau, V. (2011). Contribution of climatic and anthropogenic effects to the hydric deficit of peatlands. *Hydrological Processes*, 25(18), 2890–2906. <https://doi.org/10.1002/hyp.8052>
- Bakker, M., & Bot, B. (2025). The effective vertical anisotropy of layered aquifers. *Groundwater Series*, 63(1), 68–75. <https://doi.org/10.1111/gwat.13432>
- Bedekar, V., Cayar, M., Qian, F., & Durbin, T. (2021). SVSim: Sacramento Valley groundwater-surface water simulation model—SVSim TM-4: Model calibration and sensitivity analysis. *California Natural Resources Agency Open Data*. Retrieved from <https://data.cnra.ca.gov/dataset/svsim/resource/ecde4231-e721-468d-8ccf-7919a3cbd1c9>
- Bedekar, V., Durbin, T., & Bond, L. (2021). SVSim: Sacramento Valley groundwater-surface water simulation model - SVSim TM-5: Stream depletion calculation. California natural resources agency open data. Retrieved from <https://data.cnra.ca.gov/dataset/svsim/resource/8912ae8e-792f-47de-8791-06f55451b03a>
- Bexfield, L. M., & Jurgens, B. C. (2014). Effects of seasonal operation on the quality of water produced by public-supply wells. *Groundwater*, 52(S1), 10–24. <https://doi.org/10.1111/gwat.12174>
- Bhalla, S., Cherry, J. A., Konikow, L. F., Taylor, R. G., & Parker, B. L. (2025). Peak groundwater: Aquifer-scale limits to groundwater withdrawals. *Earth's Future*, 13(9), e2025EF006221. <https://doi.org/10.1029/2025EF006221>
- Boyce, S. E., & Ferguson, I. M. (2024). MODFLOW one-water hydrologic flow model (MF-OWHM) conjunctive use and integrated hydrologic flow modeling software with surface water operations, version 2.3.0: U.S. Geological Survey software release. GitLab. Retrieved from <https://code.usgs.gov/modflow/mf-owhm>
- Chakraborty, M., Mukherjee, A., & Ahmed, K. M. (2022). Regional-scale hydrogeochemical evolution across the arsenic-enriched transboundary aquifers of the Ganges River Delta system, India and Bangladesh. *Science of the Total Environment*, 823. <https://doi.org/10.1016/j.scitotenv.2022.153490>
- Cherry, J., Parker, B., Einarson, M., Chapman, S., & Meyer, J. (2015). Appendix 11: Overview of depth-discrete multilevel groundwater monitoring. In *Recommendations on model criteria for groundwater sampling, testing, and monitoring of oil and gas development in California (LLNL-TR-669645)* (pp. 199–280). Lawrence Livermore National Laboratory. Retrieved from https://water.llnl.gov/sites/water/files/2020-09/llnl_recommendations_report.pdf#page=281.12
- CNRA. (2025). Continuous groundwater level measurements (version monthly, from 1992 to present) [Dataset]. Retrieved from <https://data.cnra.ca.gov/dataset/continuous-groundwater-level-measurements>

- Condon, L. E., Kollet, S., Bierkens, M. F. P., Fogg, G. E., Maxwell, R. M., Hill, M. C., et al. (2021). Global groundwater modeling and monitoring: Opportunities and challenges. *Water Resources Research*, 57(12), e2020WR029500. <https://doi.org/10.1029/2020WR029500>
- Currell, M. J., Cartwright, I., Bradley, D. C., & Han, D. (2010). Recharge history and controls on groundwater quality in the Yuncheng Basin, North China. *Journal of Hydrology*, 385(1–4), 216–229. <https://doi.org/10.1016/j.jhydrol.2010.02.022>
- Dogrul, E. C., & Kadir, T. N. (2024). Integrated Water Flow Model (IWFModel)—IWFModel-2024.2.1594 (released October 28, 2024)—California natural resources agency open data. California natural resources agency open data. Retrieved from <https://data.cnra.ca.gov/dataset/iwfmodel-integrated-water-flow-model/resource/393a154c-0587-4790-85a7-6936504cb430>
- Doherty, J. E., & Hunt, R. J. (2010). Approaches to highly parameterized inversion: A guide to using PEST for groundwater-model calibration (Scientific Investigations Report). <https://doi.org/10.3133/sir20105169>
- Durbin, T., Cayar, M., & Amador, D. (2015). Sacramento Valley groundwater-surface water simulation model—SVSim TM-1A: Modeling approach for C2VSIM enhancement. *California Natural Resources Agency Open Data*. Retrieved from https://data.cnra.ca.gov/dataset/svsim/resource/4cf2f79f-3036-4af7-9328-8c5f2feeba7d?inner_span=True
- DWR. (2022). SVSim: Sacramento Valley groundwater-surface water simulation model. *California Natural Resources Agency Open Data*. Retrieved from <https://data.cnra.ca.gov/dataset/svsim>
- DWR. (2025). California Central Valley groundwater-surface water simulation model – Fine grid (C2VSIMFG)—C2VSIMFG V1.5 model files. *California Natural Resources Agency Open Data*. Retrieved from <https://data.cnra.ca.gov/dataset/c2vsimfg>
- Everett, R. R., Kjos, A., Brown, A. A., Gillespie, J. M., & McMahon, P. B. (2020). Multiple-well monitoring site adjacent to the Lost Hills oil field, Kern County, California. In *Open-file report (Nos. 2019–1114)*. U.S. Geological Survey. <https://doi.org/10.3133/ofr20191114>
- Faunt, C. C. (2009). *Groundwater availability of the Central Valley Aquifer, California (Professional paper)*. U.S. Geological Survey. <https://doi.org/10.3133/pp1766>
- Faunt, C. C., Traum, J. A., Boyce, S. E., Seymour, W. A., Jachens, E. R., Brandt, J. T., et al. (2024). Groundwater sustainability and land subsidence in California's Central Valley. *Water*, 16(8), 1189. <https://doi.org/10.3390/w16081189>
- Ferris, J. G., Knowles, D. B., Brown, R., & Stallman, R. H. (1962). Theory of aquifer tests (No. 1536E). <https://doi.org/10.3133/wsp1536E>
- Fetter, C. W. (2000). *Applied hydrogeology (4th ed.)*. Prentice Hall.
- Fisher, J. C., & Twining, B. V. (2011). Multilevel groundwater monitoring of hydraulic head and temperature in the eastern Snake River Plain aquifer, Idaho National Laboratory, Idaho, 2007–08: U.S. Geological Survey scientific investigations report 2010-5253 (p. 62). Retrieved from <https://pubs.usgs.gov/sir/2010/5253/>
- Fogg, G. E. (1986). Groundwater flow and sand body interconnectedness in a thick, multiple-aquifer system. *Water Resources Research*, 22(5), 679–694. <https://doi.org/10.1029/WR022i005p0679>
- Fogg, G. E., Carle, S. F., & Green, C. (2000). Connected-network paradigm for the alluvial aquifer system. In D. Zhang & C. L. Winter (Eds.), *Theory, modeling, and field investigation in hydrogeology: A special volume in honor of Shlomo P. Neumanns 60th birthday*. Geological Society of America. <https://doi.org/10.1130/0-8137-2348-5.25>
- Fogg, G. E., Kaiser, W. R., & Ambrose, M. L. (1991). *The Wilcox Group and Carrizo Sand (Paleogene) in the Sabine Uplift Area, Texas: Groundwater hydraulics and hydrochemistry*. Bureau of Economic Geology.
- Fogg, G. E., & Zhang, Y. (2016). Debates—Stochastic subsurface hydrology from theory to practice: A geologic perspective. *Water Resources Research*, 52(12), 9235–9245. <https://doi.org/10.1002/2016WR019699>
- Freeze, R. A., & Cherry, J. A. (1979). *Groundwater*. Prentice-Hall.
- Fulton, P. M., & Brodsky, E. E. (2023). Determining hydraulic diffusivity from ambient noise in subsurface flow rate and temperature data. *Geophysical Journal International*, 234(3), 2000–2006. <https://doi.org/10.1093/gji/ggad156>
- Giambastiani, B. M. S., Colombani, N., Mastrocicco, M., & Fidelibus, M. D. (2013). Characterization of the lowland coastal aquifer of Comacchio (Ferrara, Italy): Hydrology, hydrochemistry and evolution of the system. *Journal of Hydrology*, 501, 35–44. <https://doi.org/10.1016/j.jhydrol.2013.07.037>
- Green, A. J., Guardiola-Albert, C., Bravo-Utrera, M. Á., Bustamante, J., Camacho, A., Camacho, C., et al. (2024). Groundwater abstraction has caused extensive ecological damage to the Doñana World Heritage site, Spain. *Wetlands*, 44(2), 20. <https://doi.org/10.1007/s13157-023-0176-9>
- Guo, Z., Fogg, G. E., Chen, K., Pauloo, R., & Zheng, C. (2023). Sustainability of regional groundwater quality in response to managed aquifer recharge. *Water Resources Research*, 59(1), e2021WR031459. <https://doi.org/10.1029/2021WR031459>
- Hanson, R. T., Flint, L. E., Faunt, C. C., Gibbs, D. R., & Schmid, W. (2014). Hydrologic models and analysis of water availability in Cuyama Valley, California: U.S. Geological Survey scientific investigations report 2014–5150 (Scientific investigations report). (p. 150). <https://doi.org/10.3133/sir20145150>
- Hanson, R. T., Ritchie, A. B., Boyce, S. E., Galanter, A. E., Ferguson, I. A., Flint, L. E., & Henson, W. R. (2018). Rio Grande transboundary integrated hydrologic model and water-availability analysis, New Mexico and Texas, United States, and Northern Chihuahua, Mexico: U.S. Geological Survey open-file report 2018–1091 (p. 185). Retrieved from <https://pubs.usgs.gov/of/2018/1091/ofr20181091.pdf>
- Harbaugh, A. W. (1990). *A computer program for calculating subregional water budgets using results from the U.S. Geological Survey modular three-dimensional finite-difference ground-water flow model (Report nos. 90–392; open-file report)*. USGS Publications Warehouse. <https://doi.org/10.3133/ofr90392>
- Hatch, T. (2020). DWR report: California Central Valley groundwater-surface water simulation model—Fine grid (C2VSIMFG) development and calibration version 1.0. 275.
- Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E. J. (2020). Statistical methods in water resources. In *Techniques and methods (Nos. 4-A3)*. U.S. Geological Survey. <https://doi.org/10.3133/tm4A3>
- Hill, M. C., & Tiedeman, C. R. (2007). *Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty*. John Wiley & Sons.
- Jasechko, S., & Perrone, D. (2020). California's Central Valley groundwater wells run dry during recent drought. *Earth's Future*, 8(4), e2019EF001339. <https://doi.org/10.1029/2019EF001339>
- Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R. G., et al. (2024). Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature*, 625(7996), 715–721. <https://doi.org/10.1038/s41586-023-06879-8>
- Journel, A. G., Deutsch, C., & Desbarats, A. J. (1986). Power averaging for block effective permeability. SPE California Regional Meeting, SPE-15128-MS. <https://doi.org/10.2118/15128-MS>
- Knight, R., Steklova, K., Miltenberger, A., Kang, S., Goebel, M., & Fogg, G. (2022). Airborne geophysical method images fast paths for managed recharge of California's groundwater. *Environmental Research Letters*, 17(12), 124021. <https://doi.org/10.1088/1748-9326/aca344>

- Kwon, H.-I., Koh, D.-C., Jung, Y.-Y., Kim, D.-H., & Ha, K. (2020). Evaluating the impacts of intense seasonal groundwater pumping on stream-aquifer interactions in agricultural riparian zones using a multi-parameter approach. *Journal of Hydrology*, 584, 124683. <https://doi.org/10.1016/j.jhydrol.2020.124683>
- Lawrence, A. R., Gooddy, D. C., Kanatharana, P., Meesilp, W., & Ramnarong, V. (2000). Groundwater evolution beneath Hat Yai, a rapidly developing city in Thailand. *Hydrogeology Journal*, 8(5), 564–575. <https://doi.org/10.1007/s100400000098>
- Medellín-Azuara, J., Escrivá-Bou, A., Rodríguez-Flores, J. M., Cole, S. A., Abatzoglou, J., Viers, J. H., et al. (2022). Economic impacts of the 2020–22 drought on California agriculture.
- Ma, R., Zheng, C., Tonkin, M., & Zachara, J. M. (2011). Importance of considering intraborehole flow in solute transport modeling under highly dynamic flow conditions. *Journal of Contaminant Hydrology*, 123(1–2), 11–19. <https://doi.org/10.1016/j.jconhyd.2010.12.001>
- Maples, S. R., Fogg, G. E., & Maxwell, R. M. (2019). Modeling managed aquifer recharge processes in a highly heterogeneous, semi-confined aquifer system. *Hydrogeology Journal*, 27(8), 2869–2888. <https://doi.org/10.1007/s10040-019-02033-9>
- Meyer, J. R., Parker, B. L., Abbey, D. G., Shikaze, S. G., Weaver, L., Merritt, G., et al. (2023). Rock core VOC profiles diagnostic of aquitard occurrence and integrity in a multi-layered sedimentary rock aquifer flow system. *Journal of Hydrology*, 626. <https://doi.org/10.1016/j.jhydrol.2023.130347>
- Meyer, J. R., Parker, B. L., Arnaud, E., & Runkel, A. C. (2016). Combining high resolution vertical gradients and sequence stratigraphy to delineate hydrogeologic units for a contaminated sedimentary rock aquifer system. *Journal of Hydrology*, 534, 505–523. <https://doi.org/10.1016/j.jhydrol.2016.01.015>
- Meyer, J. R., Parker, B. L., & Cherry, J. A. (2014). Characteristics of high resolution hydraulic head profiles and vertical gradients in fractured sedimentary rocks. *Journal of Hydrology*, 517, 493–507. <https://doi.org/10.1016/j.jhydrol.2014.05.050>
- Michael, H. A., & Voss, C. I. (2009). Estimation of regional-scale groundwater flow properties in the Bengal Basin of India and Bangladesh. *Hydrogeology Journal*, 17(6), 1329–1346. <https://doi.org/10.1007/s10040-009-0443-1>
- Moss, D., Fogg, G. E., & Wallender, W. W. (2013). *Groundwater-surface water interaction in the Southeastern Sacramento Valley (Technical memorandum agreement number 4600007984)*. Prepared For California Department of Water Resources University of California Davis.
- Mukherjee, A., Fryar, A. E., & Howell, P. D. (2007). Regional hydrostratigraphy and groundwater flow modeling in the arsenic-affected areas of the western Bengal basin, West Bengal, India. *Hydrogeology Journal*, 15(7), 1397–1418. <https://doi.org/10.1007/s10040-007-0208-7>
- Nunes, M. A., Aravena, R., & Parker, B. L. (2021). Geochemical and isotopic evidence for pumping-induced impacts to bedrock groundwater quality in the City of Guelph, Canada. *Science of the Total Environment*, 800, Scopus. <https://doi.org/10.1016/j.scitotenv.2021.149359>
- Pauloo, R. A., Fogg, G. E., Guo, Z., & Henri, C. V. (2021). Mean flow direction modulates non-fickian transport in a heterogeneous alluvial aquifer-aquitard system. *Water Resources Research*, 57(3), e2020WR028655. <https://doi.org/10.1029/2020WR028655>
- Pederson, D. T. (2000). Effect of semi-perched ground water on monitoring of ground water levels in a developed aquifer. *Journal of the American Water Resources Association*, 36(6), 1367–1374. <https://doi.org/10.1111/j.1752-1688.2000.tb05732.x>
- Peffer, J. R. (1991). Complex aquifer-aquitard relationships at an Appalachian Plateau site. *Groundwater Series*, 29(2), 209–217. <https://doi.org/10.1111/j.1745-6584.1991.tb00512.x>
- Richts, A., Struckmeier, W. F., & Zaepke, M. (2011). WHYMAP and the groundwater resources map of the world 1:25,000,000. In J. A. A. Jones (Ed.), *Sustaining groundwater resources: A critical element in the global water crisis* (pp. 159–173). Springer. https://doi.org/10.1007/978-90-481-3426-7_10
- Runkel, A. C., & Meyer, J. R. (2025). Aquitardifer: A new hydrogeologic term for geologic materials with both aquitard and aquifer properties. *Groundwater Series*, 64(1), 70022. <https://doi.org/10.1111/gwat.70022>
- Runkel, A. C., Tipping, R. G., Steenberg, J. R., Retzler, A. J., Meyer, J. R., Parker, B. L., et al. (2018). A multidisciplinary-based conceptual model of a fractured sedimentary bedrock aquitard: Improved prediction of aquitard integrity. *Groundwater*, 26(7), 2133–2159. <https://doi.org/10.1007/s10040-018-1794-2>
- Sanchez-Vila, X., Guadagnini, A., & Carrera, J. (2006). Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics*, 44(3). <https://doi.org/10.1029/2005RG000169>
- Scantlebury, L., Bedekar, V., Karanovic, M., Tonkin, M. J., & Durbin, T. (2022). Texture2Par user's guide: A parameterization utility for IWF and MODFLOW, version 1.0.0. Retrieved from <https://data.cnra.ca.gov/dataset/svsm/resource/621ad13b-8fa5-411e-8a95-685e20c5d5bf>
- Scantlebury, L., Bedekar, V., Tonkin, M. J., Karanovic, M., & Harter, T. (2025). Texture2Par: A texture-driven tool for estimating subsurface hydraulic properties. *Environmental Modelling & Software*, 187, 106372. <https://doi.org/10.1016/j.envsoft.2025.106372>
- Shepley, M. G. (2024). Vertical hydraulic conductivity and layered heterogeneity: From measurements to models. *Hydrogeology Journal*, 32(4), 1017–1042. <https://doi.org/10.1007/s10040-024-02773-3>
- Tanachachoksirakun, P., Seeboonruang, U., & Fogg, G. E. (2020). Improving groundwater model in regional sedimentary basin using hydraulic gradients. *KSCE Journal of Civil Engineering*, 24(5), 1655–1669. <https://doi.org/10.1007/s12205-020-1781-8>
- Taylor, R. G., Cronin, A. A., Trowsdale, S. A., Baines, O. P., Barrett, M. H., & Lerner, D. N. (2003). Vertical groundwater flow in Permo-Triassic sediments underlying two cities in the Trent River Basin (UK). *Journal of Hydrology*, 284(1–4), 92–113. [https://doi.org/10.1016/S0022-1694\(03\)00276-2](https://doi.org/10.1016/S0022-1694(03)00276-2)
- Theis, C. V. (1940). The source of water derived from wells essential factors controlling the response of an aquifer to development.
- Thiros, S. A., Bexfield, L. M., Anning, D. W., & Huntington, J. M. (2010). Conceptual understanding and groundwater quality of selected basin-fill aquifers in the Southwestern United States. In *Professional paper (No. 1781)*. U.S. Geological Survey. <https://doi.org/10.3133/pp1781>
- Tóth, J. (1963). A theoretical analysis of groundwater flow in small drainage basins. *Journal of Geophysical Research*, 68(16), 4795–4812. <https://doi.org/10.1029/JZ068i016p04795>
- Traum, J. A., Faunt, C. C., & Boyce, S. E. (2024). MODFLOW-OWHM used to characterize the groundwater flow system of the Central Valley [Dataset]. U.S. Geological Survey. <https://doi.org/10.5066/P14N8VFT>
- von Brömsen, M., Markussen, L., Bhattacharya, P., Ahmed, K. M., Hossain, M., Jacks, G., et al. (2014). Hydrogeological investigation for assessment of the sustainability of low-arsenic aquifers as a safe drinking water source in regions with high-arsenic groundwater in Matlab, southeastern Bangladesh. *Journal of Hydrology*, 518, 373–392. <https://doi.org/10.1016/j.jhydrol.2014.05.055>
- Wada, Y. (2016). Modeling groundwater depletion at regional and global scales: Present state and future prospects. https://doi.org/10.1007/978-3-319-32449-4_10
- Wang, H. Y. (2018). Hydrogeological and geochemical comparison of high and low arsenic groundwaters in the Hetao Basin, Inner Mongolia. *Procedia Earth and Planetary Science*, 73–74. <https://doi.org/10.1201/9781351046633-28>
- Weissmann, G. S., Zhang, Y., Fogg, G. E., & Mount, J. F. (2004). Influence of incised-valley-fill deposits on hydrogeology of a stream-dominated alluvial fan. In J. S. Bridge & D. W. Hyndman (Eds.), *Aquifer characterization* (Vol. 80, pp. 28). SEPM Society for Sedimentary Geology. <https://doi.org/10.2110/pec.04.80.0015>

- Weissmann, G. S., Zhang, Y., LaBolle, E. M., & Fogg, G. E. (2002). Dispersion of groundwater age in an alluvial aquifer system. *Water Resources Research*, 38(10), 16-1–16-13. <https://doi.org/10.1029/2001WR000907>
- Wen, X.-H., & Gómez-Hernández, J. J. (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. *Journal of Hydrology*, 183(1), ix–xxxii. [https://doi.org/10.1016/S0022-1694\(96\)80030-8](https://doi.org/10.1016/S0022-1694(96)80030-8)
- Zanon, S., Nguyen, H., & Deutsch, C. V. (2002). *Power law averaging revisited. Center for Computational Geostatistics annual report papers; Centre for Computational Geostatistics report four*. University of Alberta. Retrieved from <https://www.ccgaberta.com/ccgresources/report04/2002-104-power-law.pdf>
- Zhang, Y., Weissmann, G. S., Fogg, G. E., Lu, B., Sun, H., & Zheng, C. (2018). Assessment of groundwater susceptibility to non-point source contaminants using three-dimensional transient indexes. *International Journal of Environmental Research and Public Health*, 15(6), Article 6. <https://doi.org/10.3390/ijerph15061177>
- Zhang, Z., Guo, H., Zhao, W., Liu, S., Cao, Y., & Jia, Y. (2018). Influences of groundwater extraction on flow dynamics and arsenic levels in the western Hetao Basin, Inner Mongolia, China. *Hydrogeology Journal*, 26(5), 1499–1512. <https://doi.org/10.1007/s10040-018-1763-9>