Supplementary Information

Supplementary Information for "A hydrogeologic tipping point controls climate-driven saltwater intrusion"

Chuyang Liu[1,*], Steven B. Yabusaki[2], Christopher G. Patterson[3], Erika H. Beyer[4], Daniel M. Tartakovsky[5], Dipankar Dwivedi[1]

[1]Geochemistry Department, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.
[2]Earth Systems Science, Pacific Northwest National Laboratory, Richland, Washington 99352, USA.
[3]Department of Shore Security, Naval Facilities Engineering and Expeditionary Warfare Center, Port Hueneme, California 93043, USA.
[4]Water Resources and Planning, Santa Ana Watershed Project Authority, Riverside, California 92503, USA.
[5]Department of Energy Science and Engineering, Stanford University, Stanford, California 94035, USA.

Supplementary Section 1: Model accuracy

Among 200 historical simulations, we selected the best-fitting realization as the base model. Field measurements of groundwater elevation and salinity were obtained from the Naval Installation Restoration Information Solution (NIRIS) database [36]. These data were collected from monitoring wells across multiple sites as part of various environmental restoration programs [43]. However, we noticed that differences in monitoring objectives, remediation strategies, and the site histories resulted in substantial variability in the spatial and temporal resolution of the available data. In addition, vertical datums and casing–top elevations are not consistently reported in the database. To fill these data gaps, we used a 1-m resolution digital elevation model (DEM) model from the U.S. Geological Survey with the NAVD88 datum [44]. To minimize uncertainty, we identified and removed potential outliers and artifacts by cross-checking measurements with adjacent wells (within 200 m).

We evaluated the model performance for groundwater flow by comparing simulated water levels with 5,207 measurements from seven monitoring well clusters (Fig. S6). We defined the water level residual as the difference between the observed and simulated water level and calculated it for each measurement. Our results show that the median and mean residual in most clusters are close to zero ($\leq$ 0.47 m), indicating a good agreement between model and observations (Fig. S7, Table S2). Spatially, average residuals are typically moderate (light green to light yellow color in Fig. S8), particularly in wells near the shoreline.

We then evaluated the model performance for salinity transport by a similar approach across four monitoring well clusters (Fig. S9). Since the measurement depths were not consistently reported in the database, we heuristically compared observations to the best-matching simulated salinity along each well's vertical profile. Our results show the median and mean salinity residual are typically less than 1.2 PSU, suggesting that the model can effectively capture salinity transport dynamics (Fig. S10, Table S3). Spatially, monitoring wells near the shoreline also typically show moderate average residuals ($<$ 1 PSU; light orange color in Fig. S11).

Supplementary Section 2: Water flow and salinity transport models

We simulated water flow and salinity transport in a coastal aquifer by a refined 3D unstructured mesh and a conditional seepage boundary condition to capture dynamic interactions with the sea. When the tidal water level exceeds the local ground surface elevation, a Dirichlet pressure boundary condition was applied to simulate coastal inundation. Conversely, groundwater can exfiltrate the ground surface when subsurface water pressure is above atmospheric pressure.

To capture complex boundary geometry and topographic variations, we generated a terrain-following 3D unstructured mesh by Watershed Workflow (WW)[45], with surface elevations derived from the 1-m resolution DEM. This mesh was densely refined in coastal regions (minimum mesh cell edge length of 0.1 m) to better resolve complex boundary geometry and saltwater-freshwater mixing zones, while a coarser mesh was

kept in the interior domain to account for recharge and regional subsurface freshwater flow efficiently.

To improve the estimation of vertical infiltration, we assigned a uniform cell thickness of 0.1 m to the top two soil layers. Remaining cells within the top 4 m of soil were assigned a uniform thickness of 0.2 m. Below this depth, subsequent cells had varying cell thicknesses, gradually increasing from 0.5 m to 1 m, and then to 2 m, until reaching a flat domain bottom at -14 m.

Since WW-generated mesh is not directly compatible with PFLOTRAN, we used the Mesh Toolkit [46] and exodus2pflotran tool [47] to convert the mesh format. The latter was modified to conform to PFLOTRAN's requirements for vertex ordering and data structure.

To get a reasonable initial state for various parameter sets, we spun up each model until it reached a quasi-steady state. The hydrostatic pressure was initially assigned based on measured tidal elevations from January 1, 2017. A zero salinity concentration was applied throughout the domain, representing a freshwater system. We used a cyclic spin-up method, repeatedly simulating the first two years of the entire simulation period.

Supplementary Figures

Fig. S1 Study area in Norfolk, Virginia, USA. The white outline delineates the modeled shoreline, which is bounded by Willoughby Bay, Elizabeth River, and Lafayette River. The red point indicates the National Oceanic and Atmospheric Administration (NOAA) tidal station location. Mason Creek is highlighted in blue, and two inland segments are shown in green. The inset map shows the location of the study area within Virginia. Maps were created by QGIS 3.4[48].

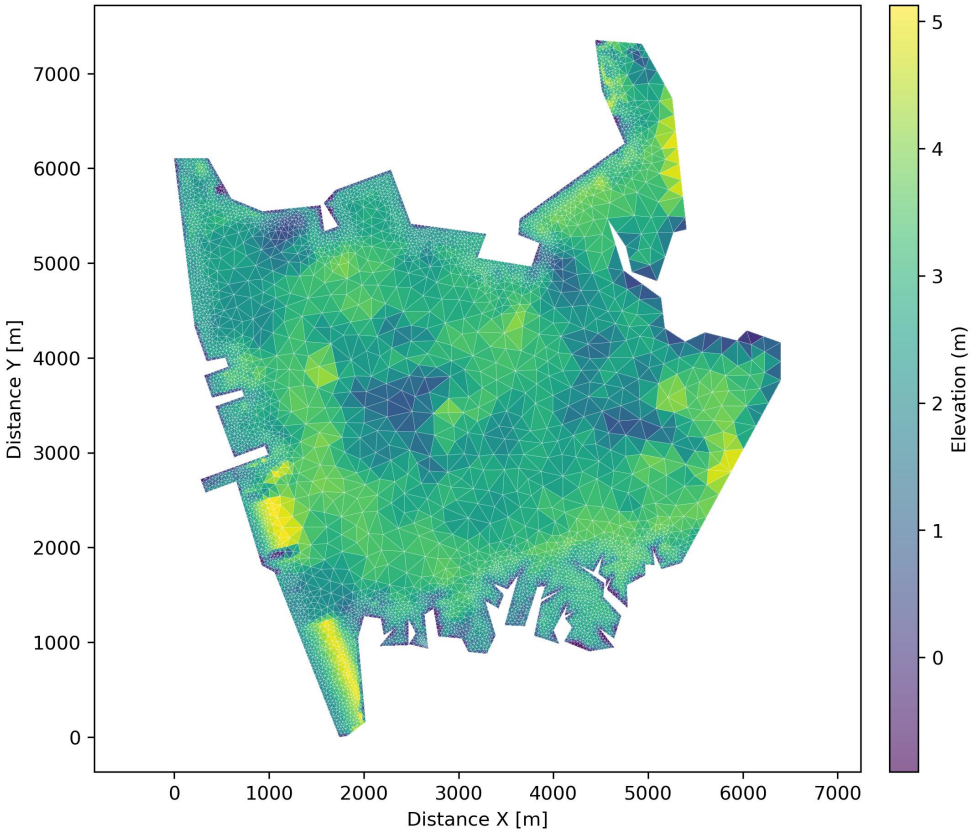

Fig. S2 Surface elevation of the study area.

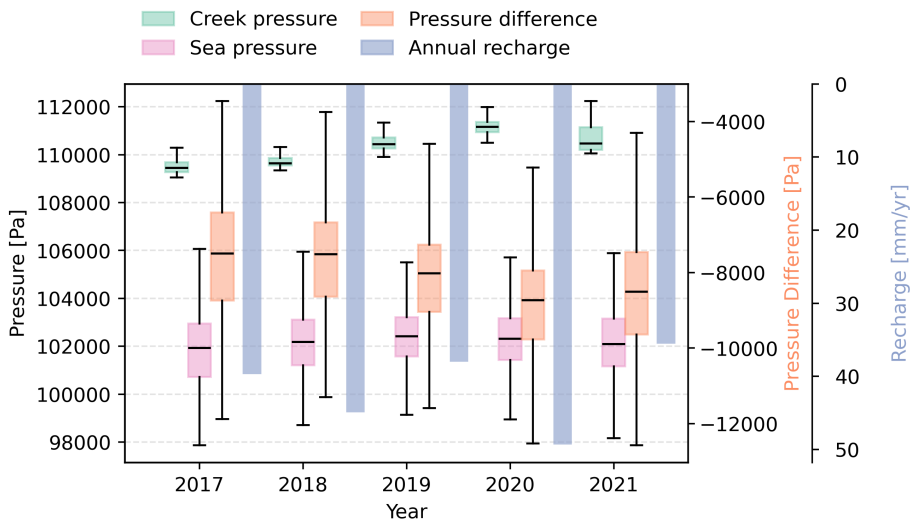

Fig. S3 Interannual variability of key drivers of salinity dynamics. The cyan and pink boxes show the pressure of the creek and sea, respectively. The orange box represents the pressure difference between the sea and the creek. The blue bar indicated the annual total groundwater recharge.

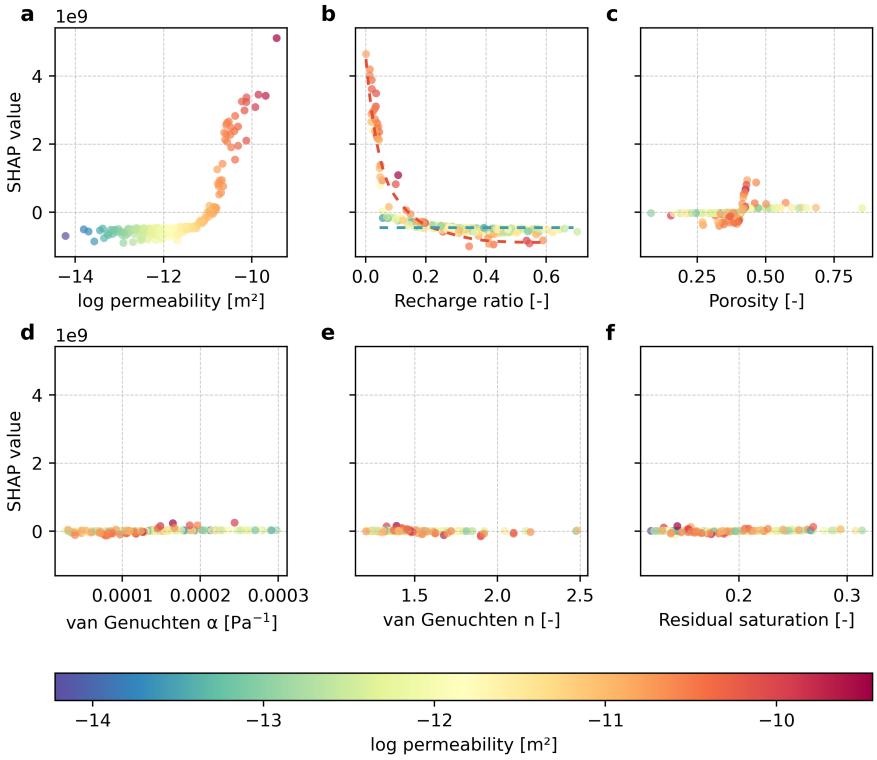

Fig. S4 Permeability and recharge are two dominant drivers in salinization. (a-f) Relative contribution of (a) permeability, (b) recharge ratio, (c) porosity, (d) van Genuchten α, (e) van Genuchten n, and (f) residual saturation on extreme salinization mass (ESM). The order of subplots is ranked by the average absolute SHAP values in descending order, representing the average impact of each variable on ESM. Orange and blue dashed lines in subplot b show two contribution patterns under different permeability ranges.

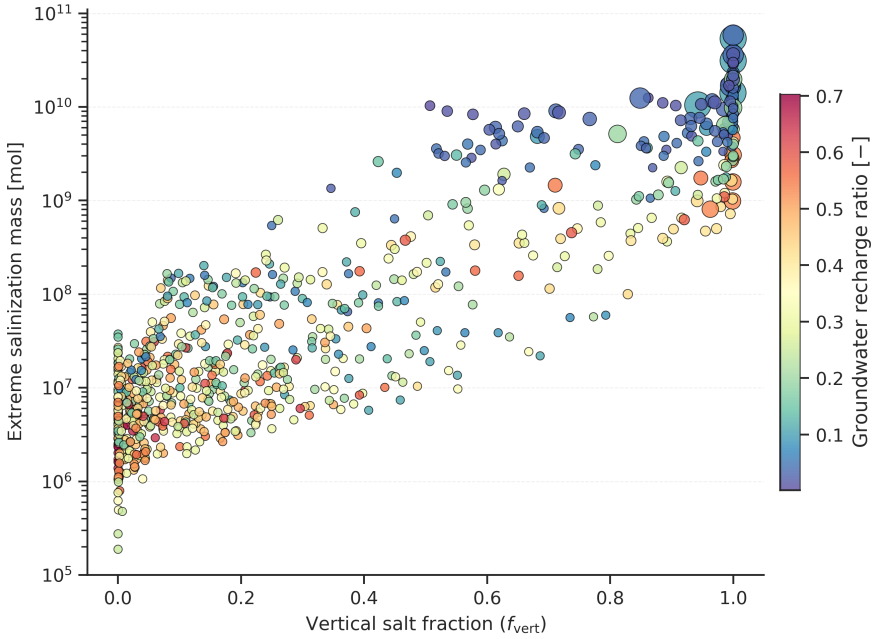

Fig. S5 Pathway dominance controls salinization severity. Extreme salinization mass over vertical salt fraction.

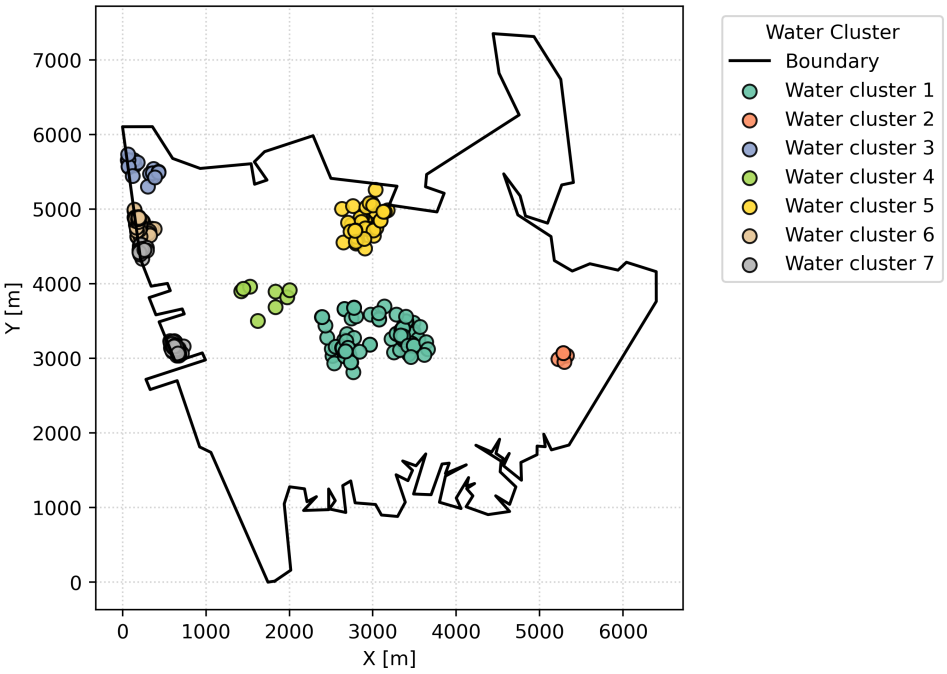

Fig. S6 Spatial distribution of monitoring well clusters for groundwater level measurements. The black solid line represents the boundary of the simulation domain. Point scatters highlight the location of monitoring wells. The color of the point indicates the classification of the monitoring well cluster.

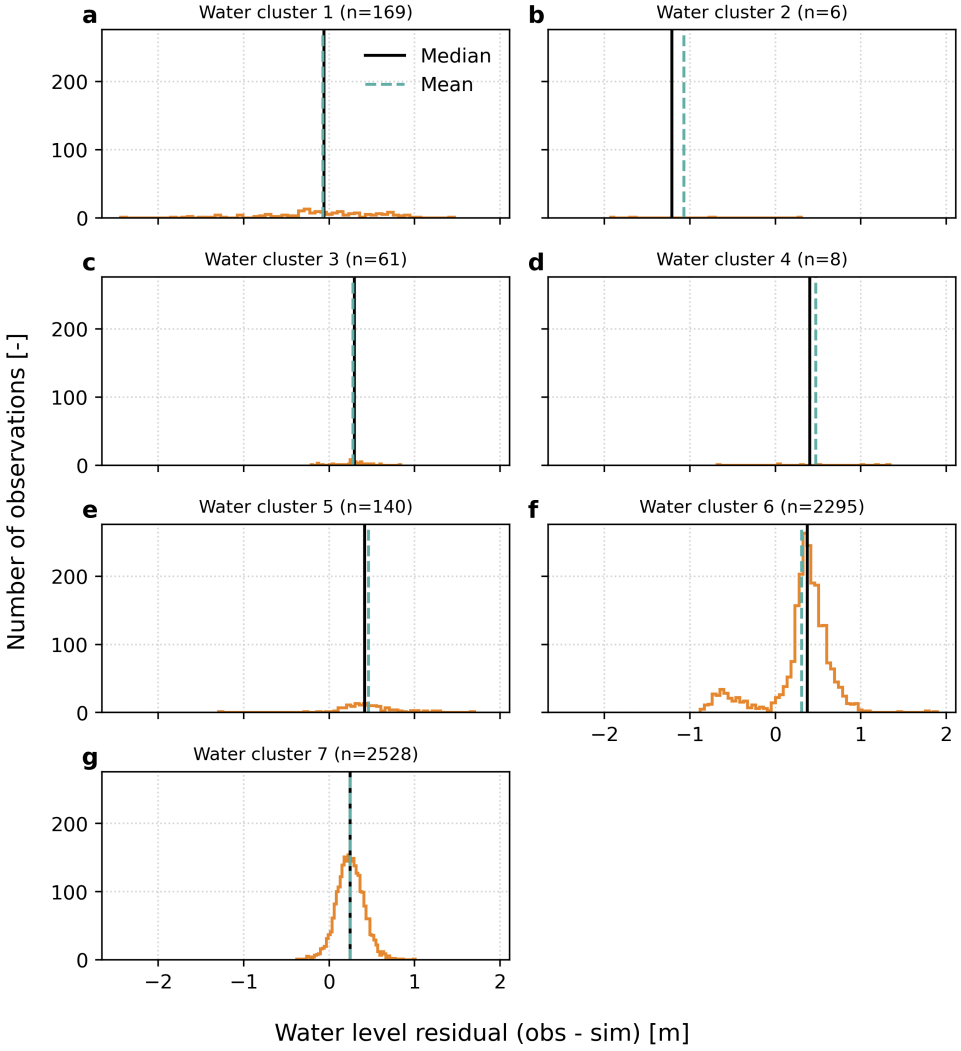

Fig. S7 The histograms of water level residual for seven water clusters. Water level residual is defined as the difference between observed and simulated water levels. The cyan dashed line represents the mean water level residuals for each water cluster. The black solid line indicates the median of water level residuals for each water cluster. The number in the bracelet shows the total measurements for each cluster.

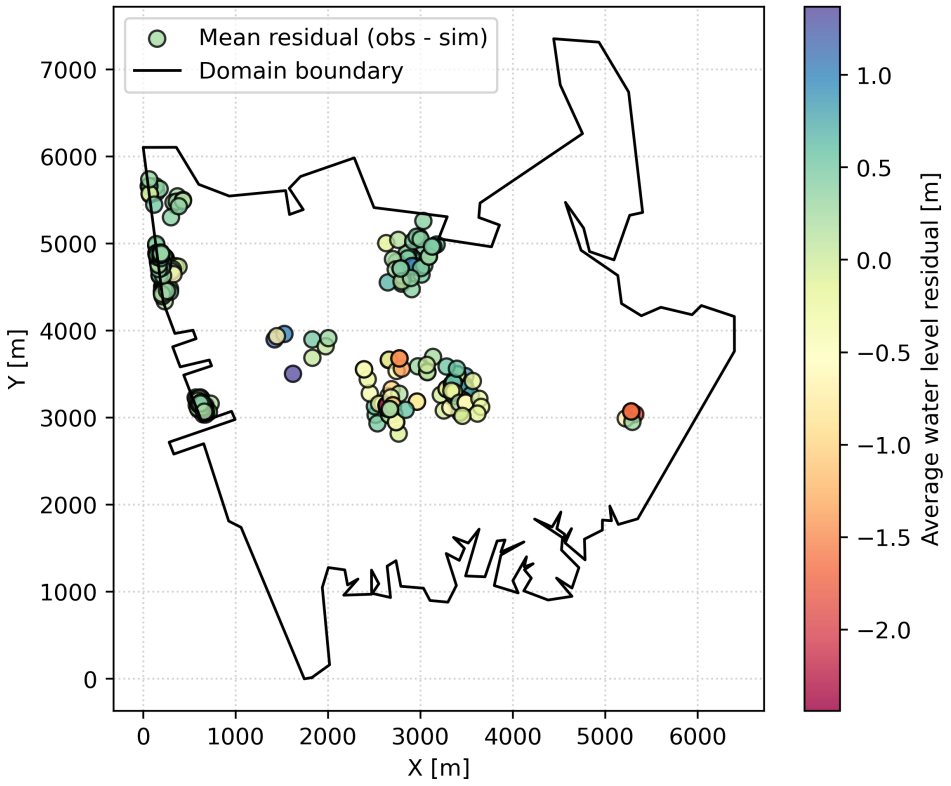

Fig. S8 The spatial distribution of the average water level residual. Water level residual is defined as the difference between observed and simulated water levels. Water level residual is averaged for each monitoring well. The black solid line represents the boundary of the model domain.

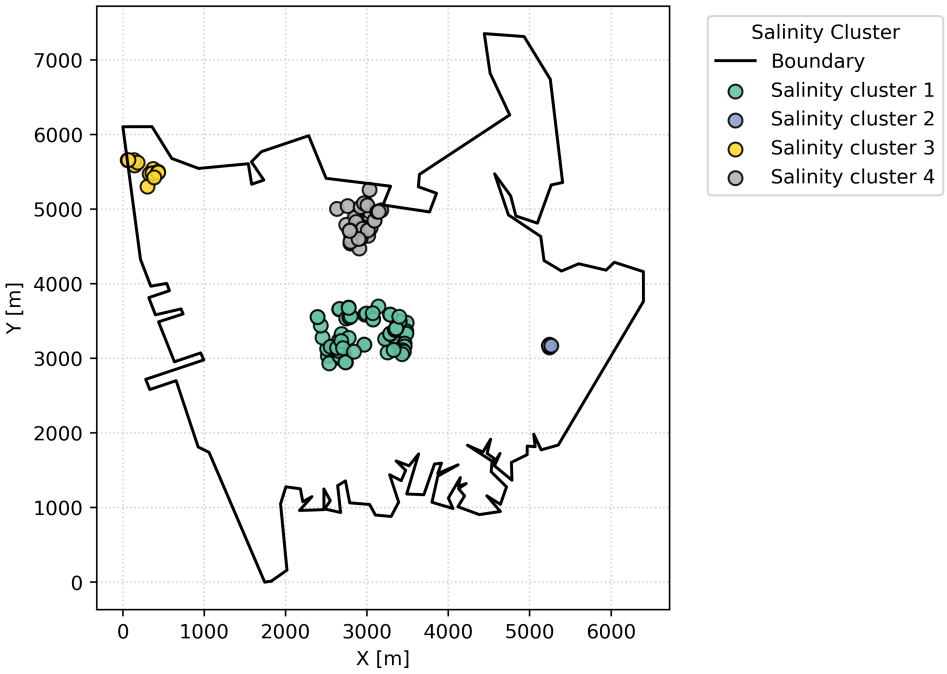

Fig. S9 Spatial distribution of monitoring well clusters for salinity measurements. The black dashed lines represent the boundary of the simulation domain. Point scatters highlight the location of monitoring wells. The color of the point indicates the classification of the monitoring well cluster.

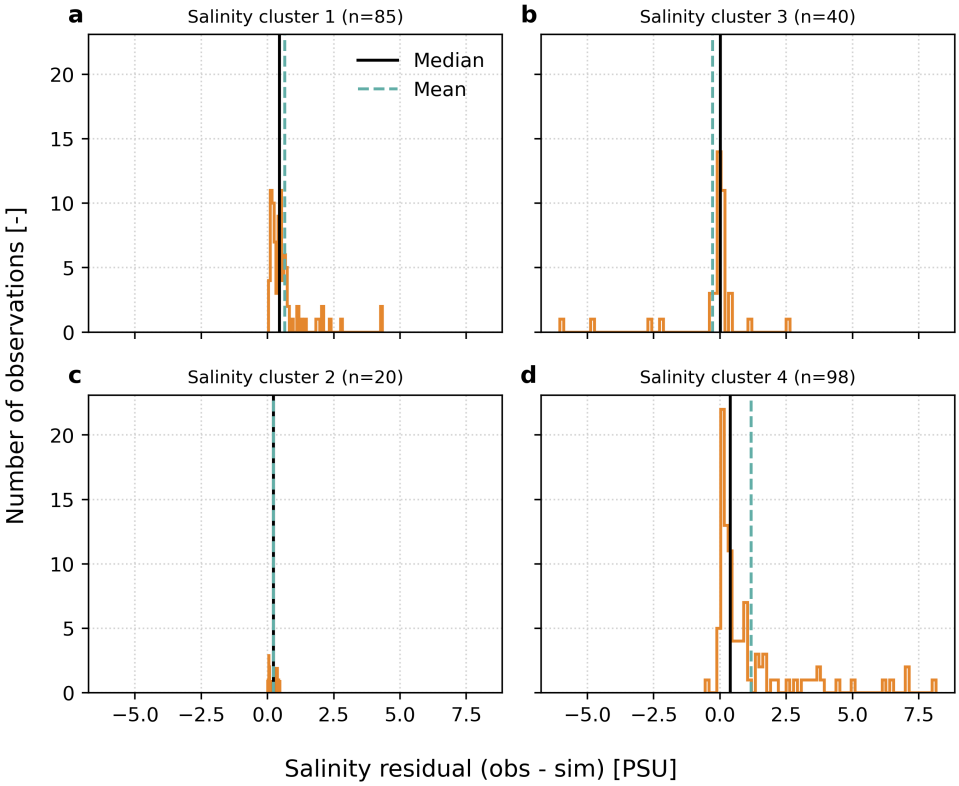

Fig. S10 The histograms of salinity residual for four salinity clusters. Salinity residual is defined as the difference between observed and simulated salinity. The cyan dashed line represents the mean residuals for each salinity cluster. The black solid line indicates the median residuals for each salinity cluster. The number in the bracelet shows the total measurements for each cluster.

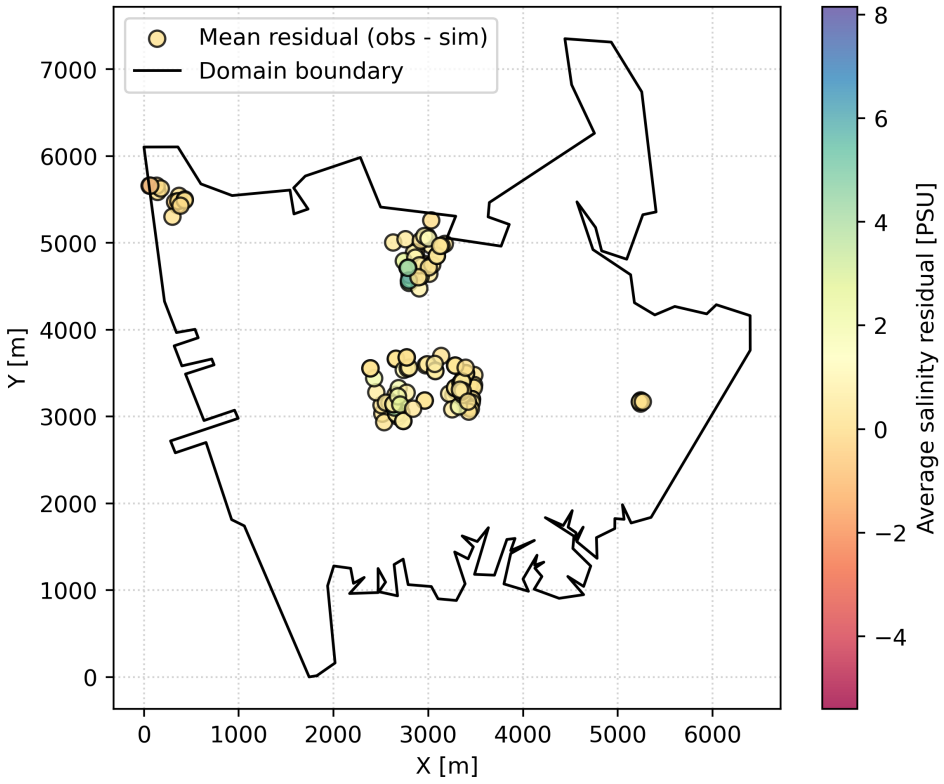

Fig. S11 The spatial distribution of the average salinity residual. Salinity residual is defined as the difference between observed and simulated salinity. Salinity residual is averaged for each monitoring well. The black solid line represents the boundary of the model domain.

Supplementary Tables

Table S1 Estimated input parameter probability distributions.

Parameter [units]	Count	Distribution shape	Statistical variables
Recharge ratio [–]	167,865	Exponential power	1.8 (shape) –0.017 (location) 0.49 (scale)
Residual saturation [–]	23,870	Gamma	2.5 (shape) 0.12 (location) 0.029 (scale)
van Genuchten α [Pa^{-1}]	23,870	Chi-square	3.3 (degrees of freedom) 2.4×10^{-5} (location) 4.2×10^{-5} (scale)
van Genuchten n [–]	23,870	Cauchy	1.4 (location) 0.098 (scale)
Permeability [m^2]	24,477	Log-normal	2.0 (shape) 1.0×10^{-16} (location) 1.2×10^{-12} (scale)
Porosity [–]	22,659	Cauchy	0.39 (location) 0.037 (scale)

Table S2 Model residual statistics for groundwater-level measurements (m) across monitoring well clusters.

Water cluster	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Arithmetic mean	-0.07	-1.1	0.28	0.47	0.46	0.31	0.25
Standard deviation	0.65	0.77	0.18	0.64	0.33	0.36	0.16
Percentiles							
10th	-0.91	-1.8	0.035	-0.17	0.16	-0.33	0.05
25th	-0.35	-1.7	0.23	0.05	0.28	0.25	0.15
50th (median)	-0.06	-1.2	0.29	0.40	0.42	0.37	0.24
75th	0.42	-0.72	0.38	1.10	0.60	0.50	0.35
90th	0.73	-0.20	0.47	1.20	0.85	0.64	0.45

Table S3 Model residual statistics for salinity measurements (PSU) across monitoring well clusters.

Salinity cluster	**1**	**2**	**3**	**4**
Arithmetic mean	0.65	0.23	-0.28	1.2
Standard deviation	0.78	0.16	1.4	1.7
Percentiles				
10th	0.14	0.05	-0.51	0.04
25th	0.22	0.07	-0.11	0.15
50th (median)	0.45	0.22	-0.01	0.39
75th	0.66	0.38	0.09	1.4
90th	1.4	0.42	0.38	3.6