

Water Resources Research®



RESEARCH ARTICLE

10.1029/2023WR035036

Groundwater Sensitivity to Climate Variations Across Australia

Xinyang Fan^{1,2,3} , Tim J. Peterson^{1,4} , Benjamin J. Henley^{1,5,6,7} , and Meenakshi Arora¹ 

Key Points:

- Climate-dominated groundwater bores throughout Australia are statistically identified ($n = 1,143$, 26%)
- Median head and recharge sensitivity to precipitation are 42 and 0.43 mm mm⁻¹
- Head and recharge sensitivity are primarily controlled by climate type and hydrogeology

¹Department of Infrastructure Engineering, University of Melbourne, Melbourne, VIC, Australia, ²Melbourne Climate Futures Academy, University of Melbourne, Melbourne, VIC, Australia, ³Division of Hydrogeology, Institute of Applied Geosciences, Karlsruhe Institute of Technology, Karlsruhe, Germany, ⁴Department of Civil Engineering, Monash University, Melbourne, VIC, Australia, ⁵School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, NSW, Australia, ⁶Securing Antarctica's Environmental Future, University of Wollongong, Wollongong, NSW, Australia, ⁷ARC Centre of Excellence for Climate Extremes, University of Melbourne, Melbourne, VIC, Australia

Supporting Information:

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

Correspondence to:

X. Fan,
fan.x1@unimelb.edu.au

Citation:

Fan, X., Peterson, T. J., Henley, B. J., & Arora, M. (2023). Groundwater sensitivity to climate variations across Australia. *Water Resources Research*, 59, e2023WR035036. <https://doi.org/10.1029/2023WR035036>

Received 4 APR 2023
Accepted 27 OCT 2023

Abstract Groundwater response to climate variations is often pivotal to managing groundwater sustainably. However, this relationship is rarely explicitly examined because of the complexity of surface to subsurface processes and the diverse impacts of multiple drivers, such as groundwater pumping and land use changes. In this paper, we address this challenge by proposing methods to quantify the sensitivity of groundwater level and recharge to temporal climate variability across Australia. Using the *HydroSight* groundwater hydrograph toolbox we first identify 1,143 out of a total of 4,350 bores as climate-driven, where historically, head was primarily driven by climate variations. Streamflow elasticity measures are then adapted to groundwater to quantify the long-term head and recharge sensitivity. We find that the national median sensitivity of head and recharge to precipitation change are 42 and 0.43 mm mm⁻¹, respectively (interquartiles: 20–77 and 0.30–0.55 mm mm⁻¹); both of which are ~8 times that of potential evapotranspiration. Nationally, the results are spatially correlated, suggestive of large-scale effects. The responses of head and recharge appear to be primarily related to climate type and hydrogeology. The more arid the climate, the higher the head sensitivity but the lower the recharge sensitivity. Porous media generally show higher head sensitivity than fractured media due to smaller aquifer specific yield, and again contrarily for that of recharge. These findings contribute to understanding the long-term impact of climate change on groundwater and thus provide valuable insights for sustainable groundwater management.

Plain Language Summary In this study, we assess the response of groundwater to meteorological variations by using long-term groundwater level records across Australia. We first identify the sites where the groundwater level has been primarily impacted by climate variations alone. The changes in groundwater level and replenishment rate (also called “recharge”) at the natural sites are simulated under a range of precipitation and evapotranspiration shifts and their relationships are statistically quantified. Results show that the national median groundwater level changes by 42 mm and the recharge changes by 0.43 mm per 1 mm change in precipitation. The response of groundwater level and recharge are found to be primarily governed by the inherent properties of the sites, such as climate type and hydrogeology. This study provides valuable insights into climate change impacts on groundwater availability and contributes to developing adaptive water management strategies.

1. Introduction

The fifth and sixth Assessment Reports of the Intergovernmental Panel on Climate Change (IPCC) state that a human-induced global warming trend is clear, but that changes to the hydrological cycle are difficult to quantify and project due primarily to model biases and intermodel spread (IPCC, 2014, 2021). As changes in temperature and precipitation significantly alter the water cycle, understanding the impact of climate change on water resources remains challenging but necessary for developing adaptive strategies (Green, 2016; Green et al., 2011; Henley et al., 2019; Taylor et al., 2013). Many factors can impact groundwater availability, such as pumping for agricultural irrigation, land use and land cover change, natural climate variability, and anthropogenic climate change (Amanambu et al., 2020; Atawneh et al., 2021; Jiménez-Cisneros et al., 2014). Due to the confounding influence of multiple drivers and limited data availability, “detection of changes in groundwater systems and attribution of those changes to climatic changes are rare” (IPCC, 2014) and “no confident assessment of groundwater

© 2023. The Authors.

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.

projections is made” to date (IPCC, 2021). Hence, the impact of climatic change on groundwater remains an open question.

Numerous studies have attempted to project groundwater level changes by calibrating a physically based model (e.g., MODFLOW) forced with meteorological data of future climate change scenarios (Allen et al., 2004; Kumar, 2016; Toews & Allen, 2009b). This approach provides the opportunity to estimate how the future climate may influence groundwater level changes, but such projections are highly uncertain because (a) the calibrated models generally lack cross-validation (Isaaks & Srivastava, 1990) and differential split sample tests (Refsgaard et al., 2014); (b) many assumptions of stationarity are made, such as land cover and land use, which have recently been challenged (Peterson et al., 2021); (c) limited observational monitoring and deficient knowledge of the subsurface (Oreskes et al., 1994); and (d) uncertainty in the predicted climate derived from both General Circulation Models and Regional Climate Models (Toews & Allen, 2009a).

Predicting groundwater level changes is challenging (Oreskes et al., 1994), and anthropogenic climate change makes this task even more onerous. As a precursor to exploring future projections, quantifying the historic impact can potentially inform understanding of the future response of groundwater to climate change. Furthermore, given that groundwater is the only widely monitored hydrological *state* variable (Peterson & Western, 2014) and often has a long memory (Fowler et al., 2020), better understanding the historic response of groundwater to climate can potentially provide insights into impacts on streamflow. Analyzing groundwater hydrographs therefore presents an opportunity to examine the cumulative impacts of climatic changes on groundwater and broader hydrology.

There are however a few obstacles to using groundwater hydrographs for sensitivity quantification including: (a) the fact that hydrographs are often short and irregularly sampled; (b) the complexity of surface to subsurface processes; (c) the heterogeneity of subsurface properties; and (d) the impacts of non-climatic drivers on the head, such as pumping for agricultural irrigation and land use change, most of which are rarely quantified and time-variant. In this study, we address the following challenges:

1. *Identifying climate-driven sites*: Many bores have been impacted by pumping at some point in the past, so identifying sites which are free from direct anthropogenic impacts is key, that is, the sites with groundwater hydrographs which have been driven primarily by climate variations rather than irrigation/land-use changes.
2. *Assessing a large number of sites*: Previous studies selected climate-driven sites based on site environment, groundwater metadata (e.g., land use and pumping records), and researchers' experience (Bloomfield et al., 2019; Cuthbert et al., 2019). However, a methodology does not exist to identify climate-driven sites among a large number of unfamiliar sites with limited prior knowledge.
3. *Accounting for groundwater memory*: There is a time delay of the head response to meteorological change (Cuthbert et al., 2019; Opie et al., 2020). That is, the observed change in the head is a cumulative response, which relates not only to the rainfall on the observed day but also to those events months or even years ago (Domenico & Schwartz, 1998; Von Asmuth et al., 2002).
4. *Quantifying sensitivity*: No known study and methodology quantifies groundwater level and recharge sensitivity to climate variations.

In light of the above challenges, here we aim to answer the two research questions: (a) in which regions of Australia is groundwater most sensitive to climate variability? We then generalize the findings to understand (b) in which types of aquifers and environments is groundwater most and least sensitive to climate variations? To achieve this aim, we develop an approach to identify climate-driven groundwater sites across the country and to quantify the sensitivity of groundwater level and recharge to climate variables and their interactions.

To expand, we aim to introduce methods to derive a first-order estimate of the sensitivity of groundwater level and recharge to climate variations in Australia. Both groundwater level and recharge sensitivity are quantified as they are impacted by different subsurface properties. The recharge is a partitioning of the infiltrated precipitation and the generation of recharge depends on the properties of the unsaturated zone (e.g., soil and vegetation roots). The groundwater level change however also depends on aquifer hydraulic properties, such as specific yield, which represents the effective porosity of the saturated zone where the water is able to flow (Kotchoni et al., 2019). Australia is chosen as study site because of its extensive groundwater borehole observational networks (>230,000 bores), the availability of continent-wide multi-decadal to centennial length meteorological observations, the variety of climate types (tropical, arid, and temperate) across the continent, and the incidence of recent droughts which have caused severe groundwater depletion (Australian Government Bureau of Meteorology, 2012; Van Dijk et al., 2013).

Below, we first describe the methods developed to identify the climate-driven sites and to quantify their groundwater level and recharge sensitivity. The spatial distributions of climate-driven sites, the sensitivity results, and the likely governing factors (e.g., climate type, hydrogeology, land use, depth to water table, aquifer property) of groundwater sensitivity are analyzed and presented in Section 3. In Section 4, we discuss the adequacy of the methods, the sensitivity results, and the implications for groundwater management under climate change. The key findings of the study are summarized in Section 5.

2. Data and Methods

2.1. Data

2.1.1. Groundwater Level and Borehole Data

Groundwater level time-series data, bore coordinates, elevations, and construction depths are obtained from the Australian Government Bureau of Meteorology (available at <http://www.bom.gov.au/water/groundwater/explorer/map.shtml>) for 232,141 bores in Australia. Under the assumptions that (a) unconfined aquifers are most likely to respond to climate variability and (b) long-term records are required to produce reliable results, a subset of bores that fulfill the following criteria are selected for use: (a) at least 200 observations per hydrograph; (b) observations covering the entire period of the Millennium Drought (Van Dijk et al., 2013) in Australia between 1997 and 2010; and (c) the bore screening minimum depth of ≤ 50 m (or where the screen depth is unavailable, a bore depth ≤ 50 m). Application of these criteria results in 5,076 sites identified for use.

2.1.2. Climate Data

The daily precipitation (P) and potential evapotranspiration (PET) from 1960 to 2018 are required to simulate past meteorological variations. The daily precipitation, maximum and minimum air temperature, relative humidity, vapor pressure, and solar radiation at each bore location are estimated using bilinear interpolation of the $0.05^\circ \times 0.05^\circ$ gridded national meteorological data (Jones et al., 2009) with an R-package *AWAPer* (Peterson et al., 2020). The interpolated climate variables are then adopted to calculate Morton (1983)'s wet environment areal PET using the method outlined by Guo et al. (2016).

Australia has a variety of climates (12 subcategories) ranging from tropical, arid, and temperate climates. The climate type at each bore location is obtained from the Köppen-Geiger climate map updated by Peel et al. (2007).

2.1.3. Land Use Data

The land use at groundwater sites is sourced from the Australian Department of Agriculture, Water and the Environment (Australian Bureau of Agriculture and Resources Economics and Sciences (ABARES), 2021). The national map has a spatial resolution of 50×50 m and contains six primary categories: conservation and natural environments (shortened to “natural”), relatively natural environments (shortened to “relatively natural”), dryland agriculture and plantations (shortened to “dryland”), irrigated agriculture and plantations (shortened to “irrigation”), intensive uses, and water bodies (e.g., rivers and lakes).

2.1.4. Hydrogeology Data

The hydrogeology of the sites is obtained from the Australian Bureau of Mineral Resources (Jacobson & Lau, 1987; Lau et al., 1987). The aquifers in Australia are divided into five classes according to the media (porous or fractured), distribution (extensive or local), and productivity (high or low-moderate). High productivity denotes those producing good quality of water at a high yield and from a shallow depth. The aquifer types include porous extensive highly productive aquifers, porous extensive aquifers with low-moderate productivity, fractured or fissured extensive highly productive aquifers, fractured or fissured extensive aquifers with low-moderate productivity, and local aquifers with generally low productivity.

2.2. Time-Series Groundwater Hydrograph Modeling

The groundwater modeling is undertaken using *HydroSight* (Peterson & Fulton, 2019; Peterson & Western, 2014) which is a nonlinear transfer function noise time-series model (available at <https://github.com/peterson-tim-j/HydroSight>). Before building the time-series models, each of the 5,076 hydrographs is analyzed for erroneous observations and outliers using *HydroSight* (Peterson & Western, 2018). The following quality checks are

undertaken: (a) observations are within the bore construction and termination dates; (b) there are no duplicate observations; (c) daily maximum head change ≤ 10 m; (d) head cannot be constant for a duration of > 120 days; and (e) observations must be less than 4 times the noise standard deviation of the model simulation. After removing any erroneous and outlier observations, only the sites that still fulfill the criteria listed in Section 2.1.1 are used for modeling.

The daily P and PET are input forcings of the groundwater models. The meteorological variables are first non-linearly transformed through a two-layer soil moisture storage (SMS) module in *HydroSight*. Compared with the one-layer module, the two-layer module restricts solutions to those giving a plausible estimate of actual evapotranspiration given the long-term site aridity, the benefit of which is improved estimates of average gross recharge (Peterson & Fulton, 2019). The climate forcings are first transformed through the upper soil layer with Equation 1 and then the deep soil layer with Equation 2 (Peterson & Fulton, 2019):

$$\frac{dS}{dt} = P_{\text{inf}} - D_{\text{shallow}} - E_{\text{shallow}} \quad (1)$$

$$\frac{dS_{\text{deep}}}{dt} = D_{\text{shallow}} - D_{\text{deep}} - E_{\text{deep}} \quad (2)$$

$$D_{\text{shallow}} = k_{\text{sat}} \left(\frac{S}{\text{SMSC}} \right)^{\beta} \quad (3)$$

$$D_{\text{deep}} = k_{\text{sat}} \left(\frac{S_{\text{deep}}}{\text{SMSC}_{\text{deep}}} \right)^{\beta} \quad (4)$$

$$E_{\text{shallow}} = \text{PET} \left(\frac{S}{\text{SMSC}} \right)^{\gamma} \quad (5)$$

$$E_{\text{deep}} = (\text{PET} - E_{\text{shallow}}) \left(\frac{S_{\text{deep}}}{\text{SMSC}_{\text{deep}}} \right)^{\gamma} \quad (6)$$

where P_{inf} and PET are daily infiltrated precipitation and potential evapotranspiration at time t ; D_{shallow} and D_{deep} are the drainage of the upper and deep layers of the two-layer SMS module at t ; E_{shallow} and E_{deep} are the evapotranspiration of both layers at t ; S and S_{deep} are the soil moisture of both layers at t ; SMSC and $\text{SMSC}_{\text{deep}}$ are the SMS capacity of both layers; k_{sat} is the vertical saturated hydraulic conductivity; β and γ are the parameters of the non-linear infiltration equations. To increase the calibration efficiency, the k_{sat} , β , and γ of the deep layer are set to the same as the upper layer. The applied assumptions are that the soil is vertically homogeneous except for the SMSC, and all precipitation infiltrates when the upper soil layer is not full (Peterson & Fulton, 2019).

The drainage of the deep soil layer (D_{deep}) is conceptualized as the gross groundwater recharge (a likely slight over-estimation due to omission of possible phreatic evapotranspiration and unsaturated zone lateral flow) (Peterson & Fulton, 2019). The recharge is then weighted with a Pearson type III distribution function (Equation 7) and integrated to give the groundwater level over time relative to a datum (Peterson & Western, 2014; Von Asmuth & Bierkens, 2005). The groundwater memory is accounted for by the weighting function which represents the historic time-series climate impact on the observed head.

$$\theta(t) = A \frac{b^n t^{n-1} \exp(-bt)}{\Gamma(n)} \quad (7)$$

where A , b , and n are the calibration parameters that define the shape of the weighting function and Γ represents a gamma function (Peterson & Western, 2014).

2.2.1. Model Calibration

Each model has nine parameters to calibrate, with five parameters (SMSC , $\text{SMSC}_{\text{deep}}$, k_{sat} , β , γ) for the two-layer SMS module and four parameters for the weighting function (A , b , n) and the exponential noise function (α). More details of the model structure and parameters are described in Peterson and Fulton (2019) and Peterson and Western (2014). Each model is calibrated to the observed hydrograph with a global calibration scheme shuffled complex evolution with principal components analysis—University of California at Irvine (SP-UCI) (Chu

et al., 2011) with 36 complexes (4 complexes per parameter). The percentage change allowed in the objective function before convergence is set to 1×10^{-5} and the calibration stops when at least 10 evolution loops meet the convergence criteria. The calibrations are performed on a high-performance computer at the University of Melbourne (Lafayette et al., 2016).

2.3. Climate-Driven Sites Identification

After modeling each site, we identify which bores are likely to be primarily influenced by climate variability and not by anthropogenic factors. A good fit of a *HydroSight* model built only with meteorological data is potentially helpful to indicate the sites where groundwater level fluctuations can be well explained by climate variations alone. Using a high threshold of model performance thus provides an opportunity to identify the sites where the water level changes have been primarily driven by climate variations. Here we use the Nash-Sutcliffe efficiency (NSE) (Nash & Sutcliffe, 1970) as the measure of goodness-of-fit:

$$\text{NSE} = 1 - \frac{\sum_{t=1}^N (H_{\text{obs},t} - H_{\text{sim},t})^2}{\sum_{t=1}^N (H_{\text{obs},t} - \bar{H}_{\text{obs}})^2} \quad (8)$$

where N is the length of the simulation period; $H_{\text{obs},t}$ and $H_{\text{sim},t}$ are the observed and simulated heads at time t ; \bar{H}_{obs} is the mean observed head over the length N . The NSE ranges from $-\infty$ to 1; 1 denotes a perfect performance and 0 denotes an estimate no better than the mean head.

We adopt an $\text{NSE} \geq 0.80$ as our criteria for identifying climate-driven sites, which equates to more than 80% of the groundwater level variance being explained by climate forcings alone (impacts of other factors are slightly possible). Different NSE thresholds (e.g., 0.85, 0.90) are explored to understand the impact of the threshold on the sensitivity results. The spatial consistency of the modeling results is assessed and compared to the maps of land use and hydrogeology.

2.4. Groundwater Level and Recharge Sensitivity Quantification

Local sensitivity is a well-established hydrological method to examine the response of streamflow to climate and the behavior of distributed groundwater models (Chiew, 2006; Hill & Tiedeman, 2006; Saltelli et al., 2019). As no known study quantifies groundwater sensitivity, we propose to adapt existing methods for streamflow sensitivity, which are often referred to as streamflow *elasticity* metrics, to quantify groundwater sensitivity. Streamflow elasticity quantifies the change in streamflow discharge for a unit change in a climate variable (Andréassian et al., 2016; Chiew, 2006). A regression measure is often used and built between the modeled discharge changes (with a rainfall-runoff model) and a range of plausible climate shifts (e.g., precipitation), and the slope is taken as the elasticity (Vano et al., 2012). To account for the joint impact of multiple climate variables (e.g., precipitation and evapotranspiration) on streamflow, a multiple linear regression (MLR) approach is often adopted (Andréassian et al., 2016; Chiew, 2006). The streamflow elasticity measures however cannot be used directly for groundwater as (a) the groundwater head responds to climatic changes with time lags (Peterson & Western, 2014), and (b) the portion of the precipitation which ultimately impacts the head is time-variant depending on the precondition of the soil.

Here we adapt the MLR method to estimate groundwater level and recharge sensitivity (a) by using the long-term mean changes (instead of instant changes due to the delayed response of the head) in the head, recharge, and climates at each climate-driven site, and (b) by using *HydroSight* to simulate the head changes since it can account for the delayed response of the head to climate variations and transform the infiltrated precipitation into the head based on the soil conditions (Equations 1–4). Both are absolute sensitivities which quantify the change (mm) in the head or recharge per 1 mm change in P or PET:

$$\Delta \bar{H}_{i,j} = \epsilon_{\text{P,H}} \cdot \Delta \left(\sum P_i \right) + \epsilon_{\text{PET,H}} \cdot \Delta \left(\sum \text{PET}_j \right) + \omega_{\text{H}} \quad (9)$$

$$\Delta \left(\sum R_{i,j} \right) = \epsilon_{\text{P,R}} \cdot \Delta \left(\sum P_i \right) + \epsilon_{\text{PET,R}} \cdot \Delta \left(\sum \text{PET}_j \right) + \omega_{\text{R}} \quad (10)$$

where

$\Delta(\sum P_i)$ is the average of the annual total precipitation changes (mm) over the record length for the daily P shifting by $i\%$ ($i \leq \pm 15\%$);

$\Delta(\sum PET_j)$ is the average of the annual total PET changes (mm) for the daily PET shifting by $j\%$ ($j \leq \pm 10\%$);

$\Delta\bar{H}_{i,j}$ is the average of the annual mean head changes (mm) between the simulated daily hydrographs with the observed climate and with P and PET shifts;

$\Delta(\sum R_{i,j})$ is the average of the annual total recharge changes (mm) between the simulated daily recharge with the observed climate and with P and PET shifts;

$\epsilon_{P,H}$ and $\epsilon_{PET,H}$ are MLR coefficients that represent the head sensitivity to P and PET in a unit of mm mm^{-1} ;

$\epsilon_{P,R}$ and $\epsilon_{PET,R}$ are MLR coefficients that represent the recharge sensitivity to P and PET in a unit of mm mm^{-1} ;

ω_H and ω_R are regression residuals and all regressions are ordinary least-square solutions ($\omega_H \sim N(0, \sigma_H)$,

$\omega_R \sim N(0, \sigma_R)$, σ_H and σ_R are the standard deviations)).

In the model simulation, the daily P shifts between $\pm 15\%$ ($n = 31$) and the daily PET between $\pm 10\%$ ($n = 21$), based on their historic range of changes (Liu et al., 2021; Ukkola et al., 2019), with an increment size of 1%. In total, 651 (i.e., 31×21) daily groundwater hydrographs are simulated for each climate-driven site. The recharge as an intermediate output of the model (Equation 4) is also simulated. The simulations are the same lengths as the observed hydrographs and are on a daily scale (not only on the irregularly observed days of the head) to provide an uninterrupted estimate of the head and recharge changes. The MLRs with a coefficient of determination $R^2 \geq 0.80$ are included in the sensitivity analysis.

While we do acknowledge value in exploring more complex and realistic climate scenarios, and *HydroSight* can explore such scenarios, we contest that examining the groundwater head and recharge response to finer-scale meteorological shifts (e.g., increased rainfall intensity) is a second-order issue behind our examination of long-term climate sensitivity. The sensitivity of head to a single precipitation event could be analytically derived but it would require prior knowledge of the soil moisture. Furthermore, even if possible, this would not provide insight into the long-term sensitivity of groundwater to climate variability. Also, we further argue that the regression step is required because the rate of head changes with P and PET has some nonlinearity, which could have been ignored by exploring a single small climate increment or undertaking a finer scaled increment within the broader range of changes. By exploring a large range of increments and using the MLR to represent the overall picture, we strike a balance between breadth and practicality in the analysis.

2.5. Specific Yield Estimation

The heterogeneous aquifer specific yields could be a main cause of the spatial discrepancy of the groundwater level and recharge sensitivity because of its role in controlling the response of the head to recharge. Specific yield is a fundamental aquifer hydraulic property that has been highly desirable for estimating the rate of drawdown and recovery from groundwater pumping. Since not all water contained in the saturated zone is freely movable and available for the wells, specific yield defines the ratio of the volume of water that can be drained by gravity to the total porosity of the saturated zone (Johnson, 1967). It is frequently used with the head data in the Water-Table Fluctuation method to estimate groundwater recharge (Fu et al., 2019; Healy & Cook, 2002; Kotchoni et al., 2019; Meinzer, 1923):

$$R_t = S_y \frac{\Delta H}{\Delta t} \quad (11)$$

where R_t is the recharge (mm) over time t , ΔH is the head change (mm) during the defined period of time Δt , and S_y is the specific yield (-) and usually expressed in percentage.

Estimating specific yield in practice, however, involves costly aquifer pumping tests (Healy & Cook, 2002), highly precise groundwater level observations, and subsequent earth-tide analysis (Chowdhury et al., 2022). Consequently, estimates are rare and often have an order of magnitude uncertainty (Crosbie et al., 2019; Healy & Cook, 2002). To overcome this challenge, we propose that the long-term mean sensitivity of head and recharge be used to estimate aquifer specific yield:

$$S_y = \frac{R_t}{\Delta H_t} \Rightarrow \frac{\Delta R_t / \Delta P}{\Delta H_t / \Delta P} \Rightarrow \frac{\epsilon_{P,R}}{\epsilon_{P,H}} \quad (12)$$

where ΔR_t is the recharge change (mm) over time t ; ΔH_t is the corresponding head change (mm), and ΔP is the precipitation change (mm) over t ; $\epsilon_{P,R}$ and $\epsilon_{P,H}$ are the long-term mean sensitivities of recharge and head to precipitation (mm mm^{-1}) in Equations 9 and 10. Note that this estimated specific yield is only a measure over the

range of depths for which the water table fluctuates, and is not necessarily an estimate of the entire (esp. deeper portion of the) aquifer.

2.6. Examination of Groundwater Sensitivity and Specific Yield

To examine the plausibility of the sensitivity results and the aquifer specific yield, the spatial correlations of these properties of all sites are analyzed. The thinking here is that the sensitivity of aquifer (recharge, level) and specific yield should be spatially correlated because physically (a) the aquifer hydrogeological properties have spatial coherence and (b) the climate variables influencing recharge and level are highly spatially correlated. Given that each site is modeled separately from other sites, a lack of spatial correlation in any of the results would warrant their rejection. The empirical variogram fitted with an isotropic spherical model (Cressie, 2015) is adopted to examine the spatial correlations: the range of the variogram represents the maximum distance within which the sites are spatially correlated; the nugget-effect ratio (=nugget/sill) of the variogram indicates the randomness or noise in the data, that is, a lower nugget-effect ratio suggests that the majority of the variance arises from the separation distance between sites rather than randomness.

To understand how aquifer types and environments influence groundwater sensitivity, we examine the head and recharge sensitivity distribution across the following categorical variables: climate type, hydrogeology, land use, depth to water table, and land surface elevation. The sensitivity distributions are compared between subgroups within each category (e.g., arid vs. temperate climates) by using the non-parametric Wilcoxon-Mann-Whitney test (Mann & Whitney, 1947; Wilcoxon, 1945) (which identifies differences in the median) at the 5% significance level. To further quantify the role of climate type, we analyze the head and recharge sensitivity versus the aridity of the climate (also called humidity index) which is calculated as the mean annual total P divided by PET at each site (Cooper et al., 2018). A $P/PET < 1$ represents a water-limited condition and > 1 indicates an energy-limited condition. The impacts of specific yield on the sensitivity are analyzed using ordinary least squares linear regression which is built between the log-transformed specific yield and the log-transformed head and recharge sensitivity individually (the logarithms use base 10).

3. Results

3.1. Groundwater Hydrograph Simulation

Figures 1a–1c show groundwater hydrographs modeled with only meteorological forcing data and the corresponding observed hydrographs at three sites in Australia. Each site is known not to be impacted by pumping, with water level changes primarily driven by climate. All three modeled hydrographs closely fit the observations before, during and after the Millennium Drought (Van Dijk et al., 2013) with an NSE > 0.90 . That is, over 90% of the observed groundwater level variance is explained by only climate variations. The long-term trends of the observed hydrographs are well simulated at all sites, despite slight underestimation of intra-annual variability (Figure 1a) which is acceptable, given that we focus on the long-term mean sensitivity of groundwater level.

In contrast, Figure 1d shows the observed and modeled hydrographs at a site known to be influenced by groundwater pumping (from 1996 onward). Initially there is a rapid drawdown (~ 2 m) in head which subsequently remains 1–2 m below the pre-pumping level, though continues to display apparent seasonal fluctuations. Modeling results show that without including pumping data, the model performance decreases by 23% (NSE: 0.87 cf. 0.67) and the uncertainty increases by ~ 1.5 m as shown by the 95% confidence interval. More significantly, modeling the site using only climate data produces an NSE below our threshold of 0.8 for the identification of a climate-driven site, and thus the correct exclusion of a site influenced by pumping. Further confirmation of this is provided by the climate-only model having a worse Akaike Information Criteria than that using both climate and pumping data (AIC: 75.5 cf. -33.7); thus suggesting the latter is a more parsimonious model. Overall these results indicate that this modeling approach and NSE criterion can successfully identify sites where the groundwater level is primarily driven by climate forcing and exclude those where other drivers have a significant influence.

3.2. Climate-Driven Groundwater Bores in Australia

Application of the selection criteria (Section 2.1.1) and outlier removal (Section 2.2) results in 4,350 bores being selected for groundwater modeling. Their locations and model performance (evaluated with NSE) are shown in Figure 2. Around 64% ($n = 2,782$) of the sites have an NSE ≥ 0.50 . More importantly, 26% ($n = 1,143$) have an

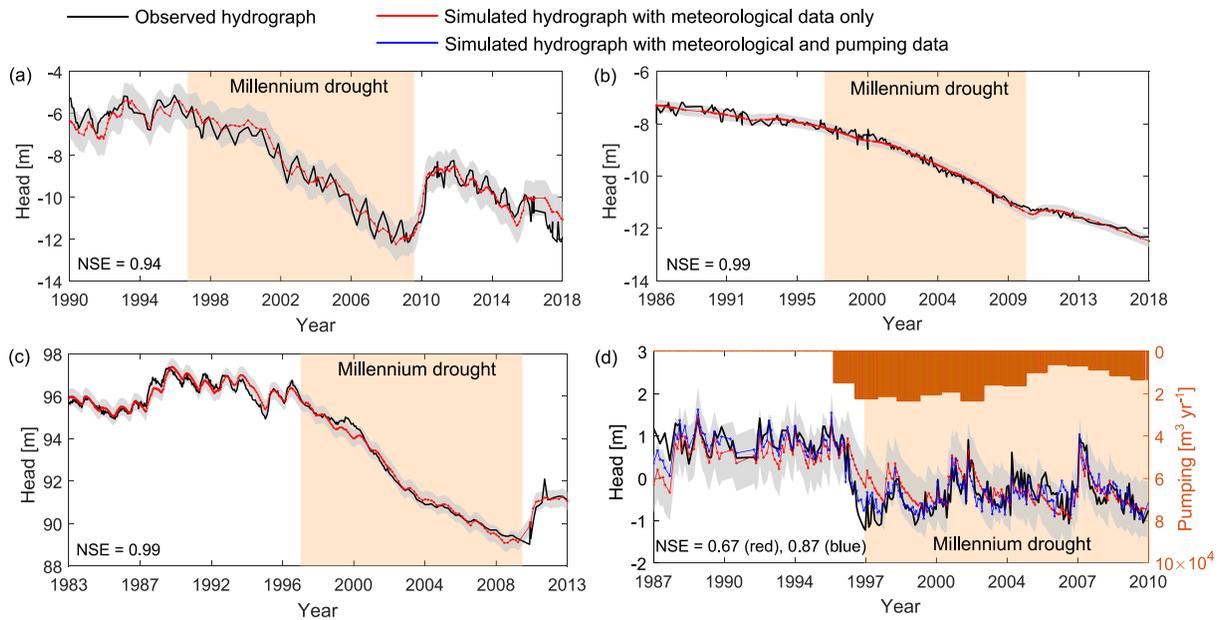


Figure 1. Long-term observed and modeled groundwater hydrographs in Victoria, Australia. (a–c) The hydrographs of three sites known to have no pumping and modeled with meteorological data alone: (a) bore ID 20054348 (Nash-Sutcliffe efficiency [NSE] = 0.94), (b) bore ID 20156664 (NSE = 0.99), and (c) bore ID 6416 (NSE = 0.99). (d) The hydrograph of a site with known pumping (bore ID 86656) and modeled with meteorological data alone (NSE = 0.67) as well as with both meteorological and pumping data (NSE = 0.87). The brown bars represent the annual pumping rate at the site; the black lines denote the observed hydrographs; the red lines denote the hydrographs modeled with meteorological data alone; the blue line is the hydrograph modeled with both meteorological and pumping data; the gray shading denotes the 95% confidence interval of the modeled hydrographs; and the orange shading denotes the Millennium Drought (~1997–2010).

NSE equal to or greater than the threshold for being primarily climate-driven (i.e., ≥ 0.80), and 11% ($n = 479$) have an $NSE \geq 0.90$. That is, climate forcing alone is able to explain at least 80% of the groundwater level variance at 26% of bores.

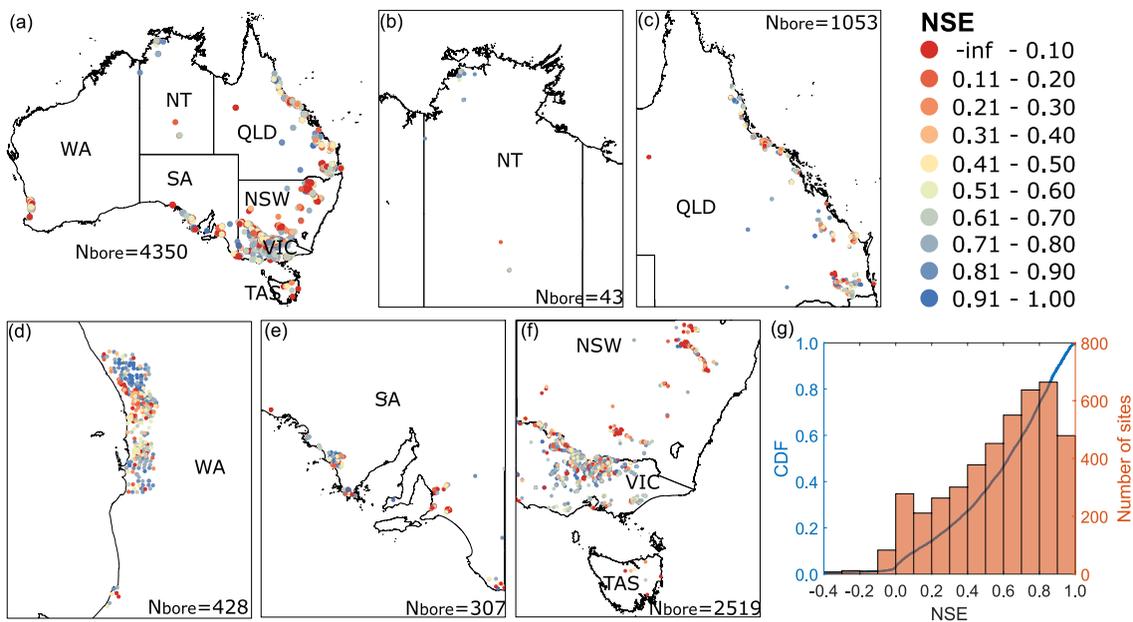


Figure 2. The Nash-Sutcliffe efficiency (NSE) distribution of 4,350 modeled groundwater hydrographs in Australia and each state. The sites with an $NSE \geq 0.80$ are selected as climate-driven sites. (a) Australia. (b) Northern Territory. (c) Queensland. (d) Western Australia. (e) South Australia. (f) New South Wales, Victoria, and Tasmania. (g) Cumulative distribution function and histogram of the NSE of all modeled sites in the country. Each dot on the maps denotes an observed borehole and the dot color denotes the NSE value.

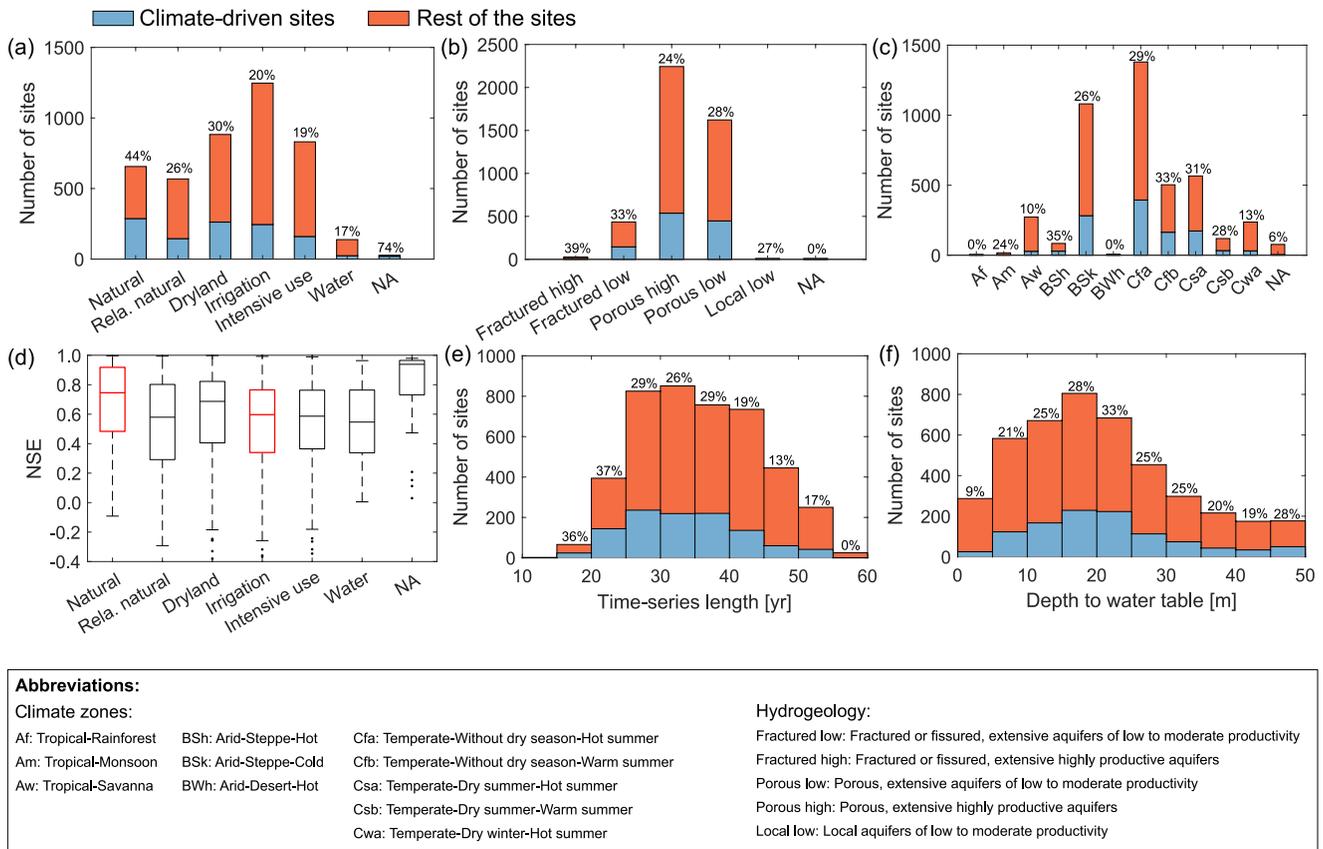


Figure 3. The distribution and statistics of all modeled groundwater sites in Australia. (a–c) The number of all modeled sites in each (a) land use, (b) hydrogeology, and (c) climate type. (d) The Nash-Sutcliffe efficiency (NSE) distribution of all modeled sites in each land use. (e, f) The time-series length and the depth to water table of all modeled hydrographs. The blue bars represent the climate-driven sites ($NSE \geq 0.80$); the orange bars represent the rest of the sites ($NSE < 0.80$). The percentage on top of the bar represents the proportion of the climate-driven sites relative to the total number of modeled sites of that bar. Note, rela. natural means “relatively natural” and NA means unavailable information.

Of these 1,143 climate-driven sites, 75% ($n = 861$) are in southeast Australia (SEA) including the states of Victoria (VIC, $n = 747$), New South Wales (NSW, $n = 28$), and South Australia (SA, $n = 86$); 11% ($n = 121$) are in southwest Australia (SWA) including the state of Western Australia (WA) only; and 14% ($n = 161$) are in northeast Australia (NEA) including the states of Queensland (QLD, $n = 143$) and Northern Territory (NT, $n = 18$). Furthermore, the spatial consistency of the NSE results (e.g., WA and NSW in Figures 2d and 2f), despite each bore being modeled independently, supports the validity of the results.

The distributions of climate-driven sites ($n = 1,143$) in each land use, hydrogeology, and climate are shown in Figures 3a–3c. Figure 3a shows that the number of climate-driven bores in the less human-impacted (natural, relatively natural, and dryland areas) versus human-impacted areas (irrigation, intensive use) is 3:2 ($n = 694:449$). Additionally 44% ($n = 287/656$) of the bores in the natural area are identified as climate-driven sites, whereas only 20% ($n = 245/1,247$) of those in the irrigation area are identified as climate-driven. Within the natural area the median NSE of all modeled sites is 0.75, which is 25% higher than that (0.60) for the irrigation area (Figure 3d). The Wilcoxon-Mann-Whitney test also finds that the null hypothesis of the natural and irrigation areas having the same NSE distribution can be rejected at the 5% significance level (p -value < 0.001). Overall, these results show that the adoption of an NSE criterion for the identification of climate-driven sites meets the reasonable expectation that such sites are more likely in natural and dryland agricultural regions and less likely in intensively used and irrigated regions.

Figure 3b shows that the number of climate-driven sites identified at porous versus fractured aquifers is 6:1 ($n = 984:156$). Around 25% ($n = 984/3,865$) of porous aquifers and 34% ($n = 156/463$) of fractured aquifers are identified as climate-driven. The number of climate-driven sites identified in tropical, arid, and temperate climates

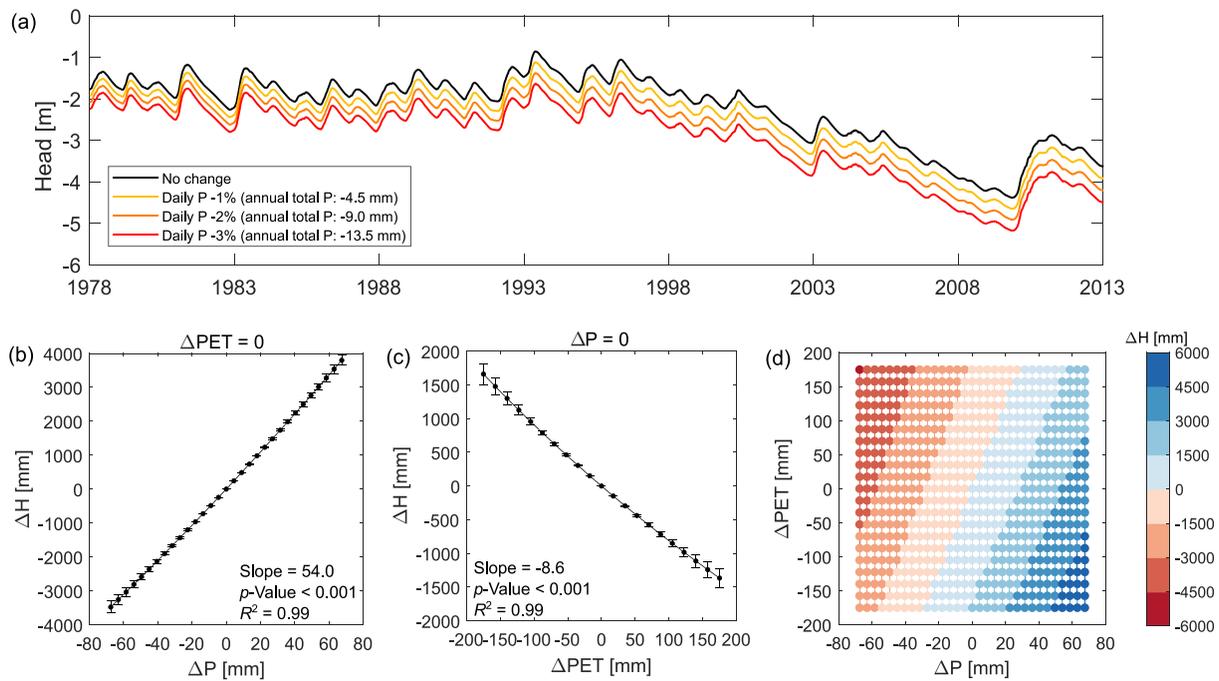


Figure 4. Groundwater head (H) changes with P and potential evapotranspiration (PET) variations at a climate-driven site (ID 20120288, Nash-Sutcliffe efficiency = 0.92) in Victoria, Australia. (a) Simulated daily groundwater head for a decline in daily P of 1%, 2%, and 3% (i.e., the average of the annual total P decreases by 4.5, 9.0, and 13.5 mm). (b, c) The average of the annual mean head changes (ΔH) against the average of the annual total P changes (ΔP) for constant ΔPET , and that against the average of the annual total PET changes (ΔPET) for constant ΔP . The solid line represents an ordinary least squares linear regression fitted between the variables. (d) The average of the annual mean head changes against the averages of the annual total P and PET changes. The grid line represents a multiple linear regression fitted between them with a $R^2 = 0.99$.

is 1:10:25 ($n = 32:311:795$) (Figure 3c). Around 11% ($n = 32/296$) of the sites in tropical climates are identified as climate-driven, whereas 27%–28% of the sites in arid ($n = 311/1,172$) and temperate ($n = 795/2,805$) climates are identified. Overall, these results show that the selection of climate-driven sites across Australia is relatively unbiased with respect to the dominant aquifer type and likewise for climate, despite a modest under-representation of sites within tropical regions, the cause of which is unknown and may arise from the deficient performance of *HydroSight* within energy-limited environments.

The length of all modeled hydrographs is between 13 and 58 years (1960–2018). The climate-driven sites have a length of 16–53 years with a median of 33 years (Figure 3e). Around 26% of the sites are identified as climate-driven at each time length on average: a slightly higher rate for those sites with ≤ 25 years and slightly lower for those with > 40 years. All modeled sites have a depth to water table (DTW) between 1.2 and 50 m. The DTW at climate-driven sites ranges 2.1–50 m with a median depth of 20 m. The ratio of the climate-driven sites identified at each DTW is relatively similar between 20% and 30%, except the top 5 m where 9% of the climate-driven sites are identified. These results show that the climate-driven sites are successfully identified at a variety of time-series lengths and depths to water table, despite a modest bias toward shorter record lengths, possibly because those with longer records are influenced by historic land clearing post-colonial settlement; and a modest omission of sites with a shallow water table, most likely because of the more complex vadose zone processes and the omission of such from *HydroSight*.

3.3. Groundwater Level and Recharge Sensitivity and Specific Yield Estimation

To illustrate the estimation of groundwater level sensitivity, Figure 4a shows the groundwater head simulated using three increments of reduced precipitation, as well as with no change. The modeled internal fluxes (e.g., SMS and deep drainage) of this site are provided in Figure S1 in Supporting Information S1. With a 1% decrease in the daily P (i.e., the average of the annual total P reduced by 4.5 mm), the daily head decreases on average by 0.24 m over the record length. When PET is constant, the mean head change is linearly related to the P change

($R^2 = 0.99$), with only minor curvature apparent at large P changes (Figure 4b). The slope provides an estimate of the long-term mean head sensitivity to P, which is a head change of 54 mm per 1 mm change in P. However, if the slope is quantified for the positive and negative P change separately, the head is found to be slightly more sensitive to a positive P change than a negative change (56.3 cf. 51.5 mm mm⁻¹).

Similarly, when PET is varied and P is constant, the head changes by -8.6 mm per 1 mm PET change (Figure 4c). Again, the head is slightly more sensitive to a negative PET change than a positive change (-9.0 cf. -7.4 mm mm⁻¹). To jointly quantify the head changes with both P and PET, we fit Equation 9 to the plane in Figure 4d. The final result for the site shows that the groundwater head is ~6 times more sensitive to P than to PET (54.1 cf. -8.6 mm mm⁻¹ respectively).

We find that 95% (1,081/1,143) of the climate-driven sites in Australia have an MLR model fit (R^2) more than 0.80. The head and recharge sensitivity of these sites are shown in Figure 5. All sites are found, as expected, to have a positive sensitivity to P and negative sensitivity to PET. The groundwater level sensitivity to P varies from ≤40 mm of head change per 1 mm of P change at 48% ($n = 518$) of sites, 40–80 mm mm⁻¹ at 28% ($n = 306$) of sites, and >80 mm mm⁻¹ at 24% ($n = 257$) of sites (Figure 5a). For PET, the absolute value of the groundwater level sensitivity varies from ≤5 mm of head change per 1 mm of PET change at 42% ($n = 452$) of sites, 5–10 mm mm⁻¹ at 30% ($n = 323$) of sites, and >10 mm mm⁻¹ at 28% ($n = 306$) of sites (Figure 5b).

Spatially, the groundwater level sensitivity to precipitation is the highest in SEA, followed by SWA and NEA in a ratio of 3:2:1 (48.5:38.5:17.2 mm mm⁻¹) when comparing the regional median (Figure 6a). The groundwater level sensitivity to PET in SEA and SWA is slightly higher than that in NEA in a ratio of 1.8:1.7:1 (-6.6:-6.2:-3.7 mm mm⁻¹ in median) (Figure 6b). The national median of the groundwater level sensitivity to P is ~42 mm mm⁻¹ and that to PET is around -6.4 mm mm⁻¹.

Similarly, the groundwater recharge also responds positively to P and negatively to PET changes (Figures 5c and 5d). The recharge sensitivity to P varies from ≤0.3 mm of recharge change per 1 mm of P change at 25% ($n = 271$) of sites, 0.3–0.5 mm mm⁻¹ at 41% ($n = 440$) of sites, and >0.5 mm mm⁻¹ at 34% ($n = 370$) of sites. For PET, the absolute value of recharge sensitivity varies from ≤0.05 mm of recharge change per 1 mm of PET change at 38% ($n = 407$) of sites, 0.05–0.08 mm mm⁻¹ at 33% ($n = 359$) of sites, and >0.08 mm mm⁻¹ at 29% ($n = 315$) of sites. The groundwater recharge sensitivity to P is the highest in SWA, followed by SEA and NEA in a ratio of 1.4:1.1:1 (0.55:0.43:0.40 mm mm⁻¹ in median) (Figure 6c). The groundwater recharge sensitivity to PET is slightly higher in SWA and NEA compared with SEA in a ratio of 2:2:1 (0.10:0.10:0.05 mm mm⁻¹ in median) (Figure 6d). The national median sensitivity of groundwater recharge to P is 0.43 mm mm⁻¹ and that to PET is -0.06 mm mm⁻¹.

Overall, the head and recharge sensitivity results are physically plausible and justify their use in estimating aquifer specific yield ([%]) shown in Figure 5e. Results show that around 55% ($n = 595$) of sites have a specific yield less than 1%; 32% ($n = 347$) of sites have a specific yield between 1% and 4%; and 13% ($n = 139$) of sites have a specific yield more than 4%. The 95th percentile of the specific yield is 10% with a maximum of 42% (Figure 6j). The national median specific yield is 0.9% and that in SWA, SEA, and NEA are 1.6%, 0.8%, and 2.1% (Figure 6e).

Figure 7a shows that both groundwater level and recharge sensitivity to P are around 8 times that of PET, which supports the assumption that precipitation plays a dominant role in driving water level and recharge changes. The groundwater level and recharge sensitivity however show a very low correlation ($R^2 < 0.1$), probably due to the extra influential factors of the head such as aquifer hydraulic properties. The groundwater level sensitivity is strongly negatively correlated with the specific yield with a slope of -0.6 to -0.7 in their log-correlations ($R^2 > 0.7$, Figure 7b), which is as expected given that a lower specific yield physically equates to a larger head reduced per unit of water extracted from an aquifer. The specific yield however also shows a slight positive correlation with recharge sensitivity ($R^2 \approx 0.4$), possibly because the recharge sensitivity is used to estimate specific yield in the derivation method (Equation 12).

Additionally, in understanding the uncertainty in the sensitivity results and specific yield due to the adopted NSE threshold (≥ 0.80) in climate-driven sites selection, our results may slightly underestimate the head sensitivity by 8%–11%, overestimate the recharge sensitivity by 2%–4% and specific yield by ~10% compared with adopting a higher NSE (e.g., ≥ 0.85 , Figure S2 in Supporting Information S1). However, a slight compromise in model performance, from an NSE of 0.85 to 0.80, increases the spatial coverage of the selected climate-driven sites by ~40% (1,143 cf. 818 sites). We therefore judge that the application of an NSE ≥ 0.80 is sensible for balancing the trade-off between the results accuracy and spatial coverage (i.e., to avoid spurious spatial correlations).

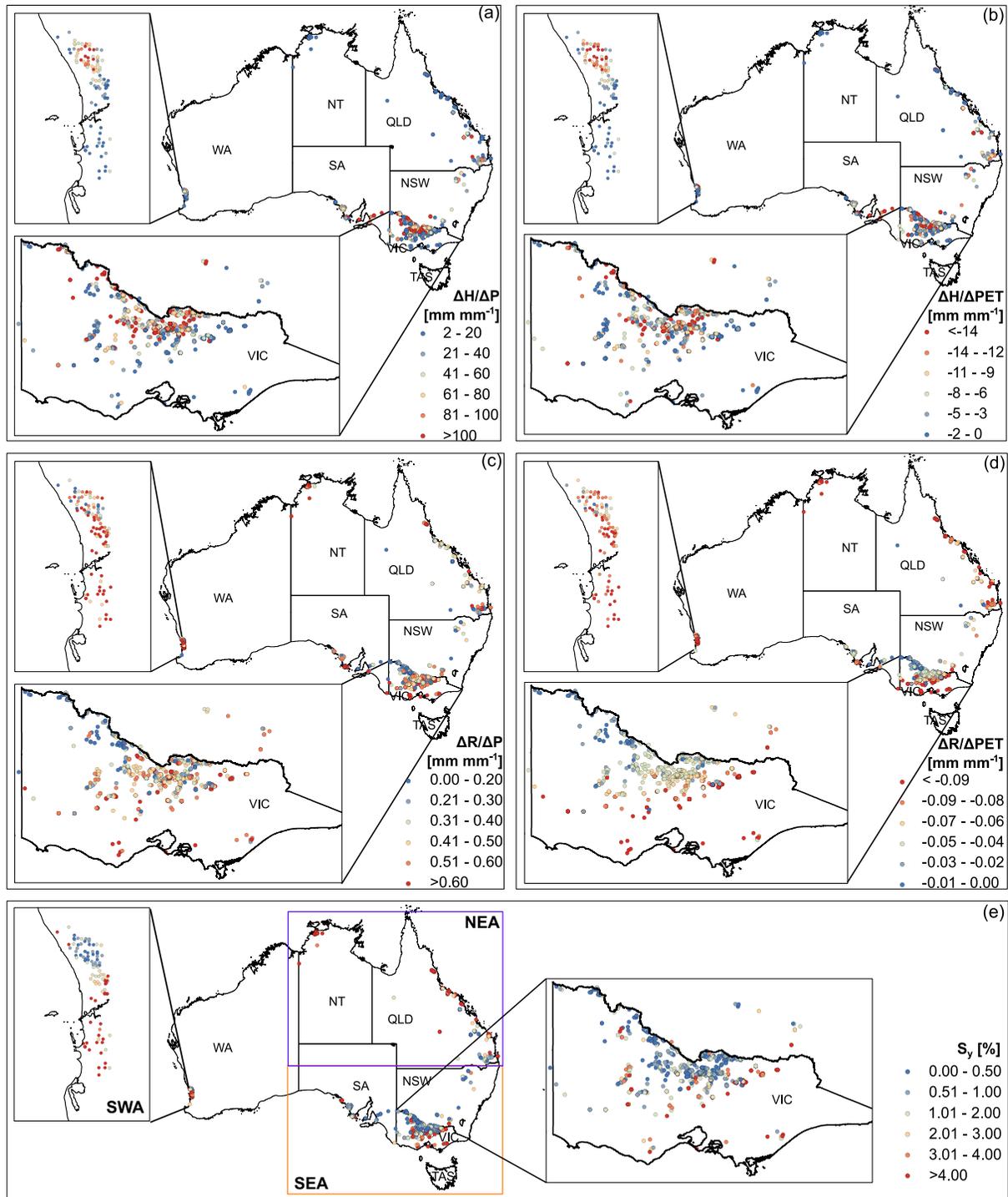


Figure 5. The sensitivity of groundwater level and recharge to P and potential evapotranspiration (PET) and the estimated specific yield in Australia (AUS). (a, b) The groundwater level (H) sensitivity to P and PET. (c, d) The recharge (R) sensitivity to P and PET. (e) The estimated aquifer specific yield (S_y) for all sites. The zoomed-in maps show the clusters of the sites in southwest Australia and Victoria (VIC). Note, southeast Australia includes VIC, South Australia, New South Wales, and Tasmania; northeast Australia includes Queensland and Northern Territory.

To examine the plausibility of the sensitivity results and specific yield, empirical and model isotropic variograms of each result are shown in Figure 8. The sensitivity results show that around half of the spatial variability is explained by the distance between bores; whereas the remaining variability is random and independent of location (i.e., nugget). More importantly, the variogram range is from tens to hundreds of kilometers. The recharge

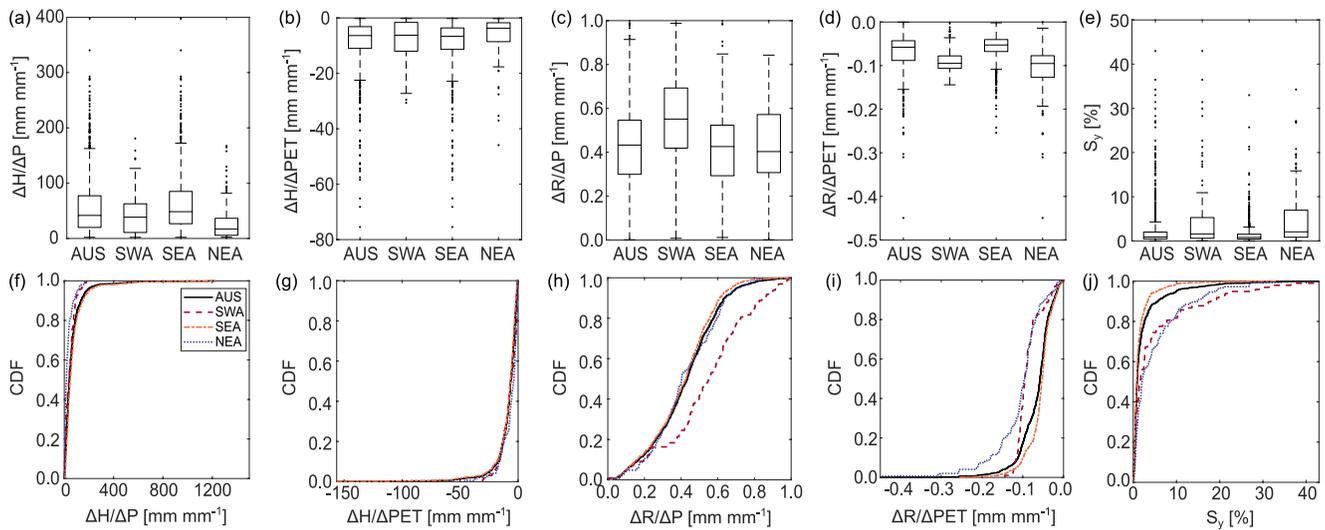


Figure 6. Boxplots and cumulative distribution functions of (a, b, f, and g) head sensitivity and (c, d, h, and i) recharge sensitivity to P and potential evapotranspiration, and (e, j) the specific yield of all sites in Australia and each region. The regional divisions are shown in Figure 5e.

sensitivity to PET has the largest range (651 km) followed by that to P (268 km), which is as expected given precipitation is more spatially variable than PET and often a stronger driver within the model. The ranges of the variograms of head sensitivity however contract to about 30 km. This is lower than expected given that the head sensitivity is a function of the aquifer hydraulic properties, but may be explained by the head sensitivity incorporating many factors (in addition to soil and climate) such as bore screening depth and proximity to other omitted drivers such as water bodies. Finally, the specific yield estimates have a nugget close to 0, suggesting very low random noise in the data, and a range of 204 km. Given that its derivation is a function of both recharge and head sensitivity, a range between the two is sensible. Overall, given each site is analyzed independently of others, the

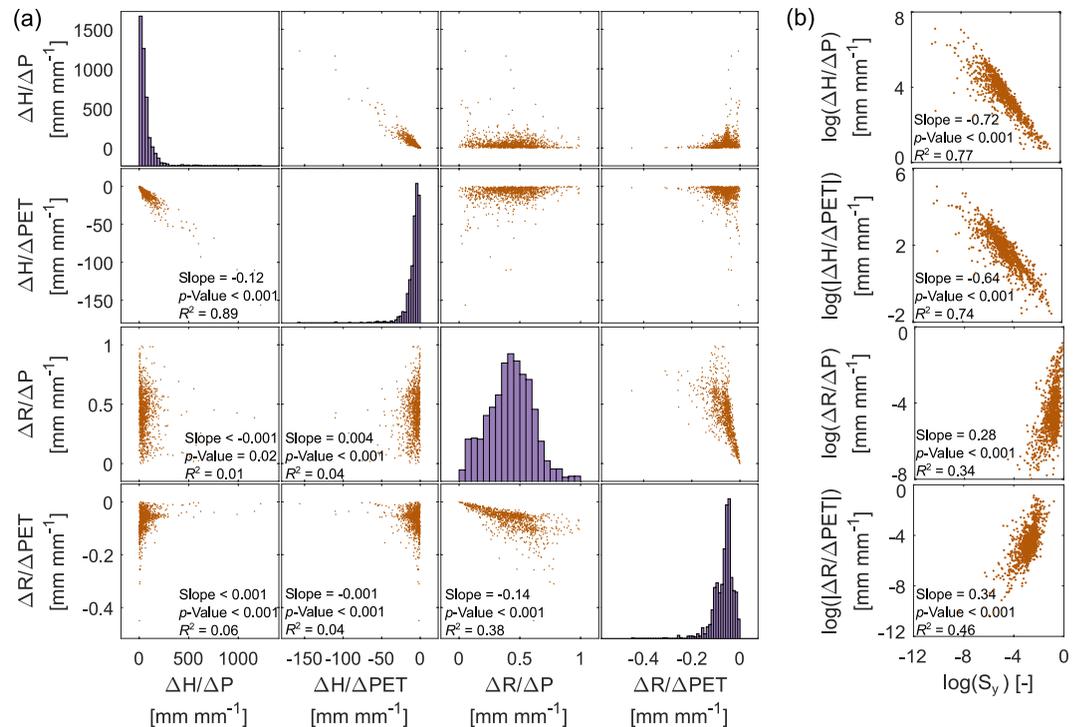


Figure 7. (a) Relationships of groundwater level and recharge sensitivity to P and potential evapotranspiration and (b) the log-correlated relationships of the sensitivity with specific yield of all sites in Australia. The logarithms use base 10.

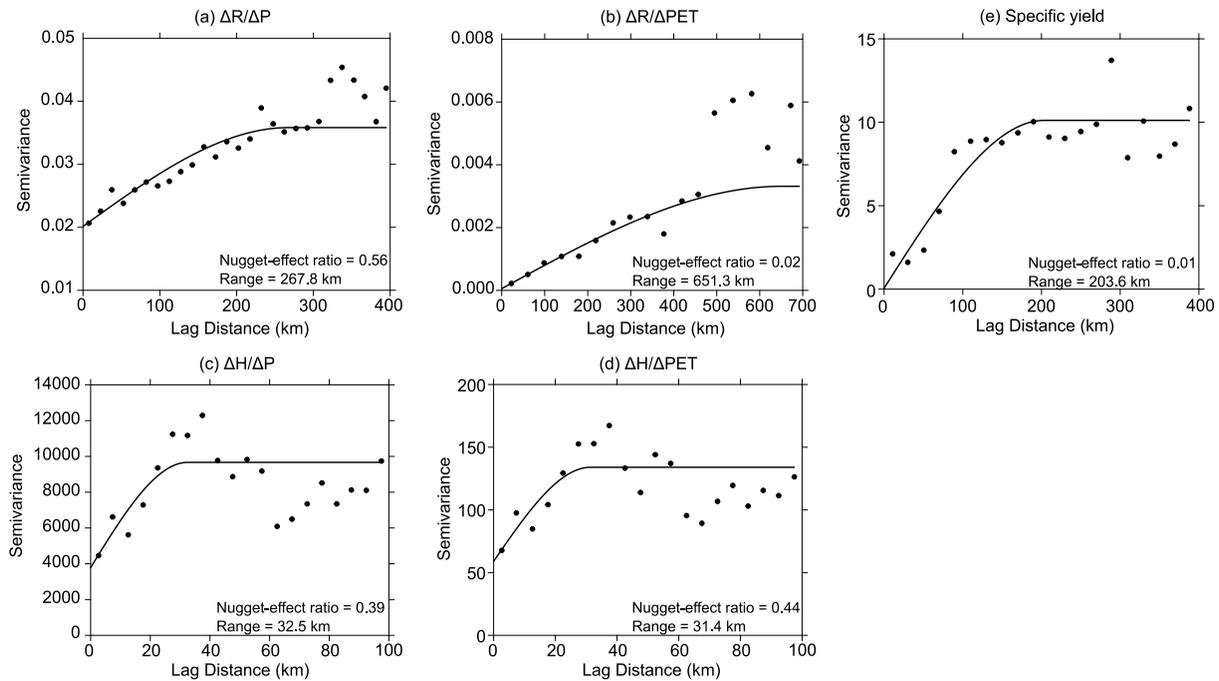


Figure 8. Empirical and model isotropic variograms of (a, b) recharge sensitivity, (c, d) head sensitivity to P and potential evapotranspiration, and (e) aquifer specific yield in Australia. The nugget-effect ratio = nugget/sill.

emergence of spatial correlations consistent with the scale of relevant climate and aquifer factors suggest that the results are reasonable.

3.4. Explanatory Factors for Groundwater Sensitivity

The magnitudes of groundwater level and recharge sensitivity are found to vary across climate types (Figures 9a–9d). The head sensitivity to P is highest in arid climates (58 mm mm⁻¹ in median), followed by temperate and tropical climates (40 and 5 mm mm⁻¹ median respectively, see Figure 9a); and similarly so for that to PET (Figure 9b). In contrast, the recharge sensitivity to P shows an opposite order with the highest sensitivity in the tropical climates (0.51 mm mm⁻¹ in median), followed by temperate and arid climates (0.45 and 0.36 mm mm⁻¹ median, see Figure 9c); and again similarly so for that to PET (Figure 9d).

Looking at the catchment properties, the median head sensitivity at porous aquifers is twice that at fractured aquifers (46 cf. 23 mm mm⁻¹ for sensitivity to P, and -7 cf. -4 mm mm⁻¹ for that to PET, see Figures 9e–9h). In contrast, the median recharge sensitivity of fractured aquifers is slightly greater than that of porous aquifers (0.47 cf. 0.42 mm mm⁻¹ for sensitivity to P, and -0.09 cf. -0.05 mm mm⁻¹ for that to PET). With regards to land use, at the irrigation areas the median head sensitivity to P and PET are 62 and -8 mm mm⁻¹, which are slightly higher than those at other categories (Figures 9i and 9j). The recharge sensitivity to P and PET however show no clear relationship with different land uses (Figures 9k and 9l). Similarly, the head and recharge sensitivity show no clear relationship with the depth to water table or land surface elevation (Figure S3 in Supporting Information S1).

To examine the spread within each climate and catchment property category, Figure 10 shows the results from the non-parametric Wilcoxon-Mann-Whitney tests comparing the distribution from a combination of each climate type, hydrogeology and land use. It shows that the head sensitivity to P and PET is significantly different between climate types; however, only porous against fractured aquifer type appears to be significant; for land use, irrigation again appears to be significantly different from other uses while natural and relatively natural uses often differ to the agricultural categories (dryland, irrigation, intensive). With regards to recharge sensitivity, again the climate type is very often a significant factor; however, unlike for head, all combinations of aquifer type differ from one another. Land use, though, does not appear to control recharge sensitivity.

Overall, these results suggest that climate and hydrogeology are the major controllers of head and recharge sensitivity, with recharge being the most sensitive to hydrogeology. Practically this suggests that the head is generally

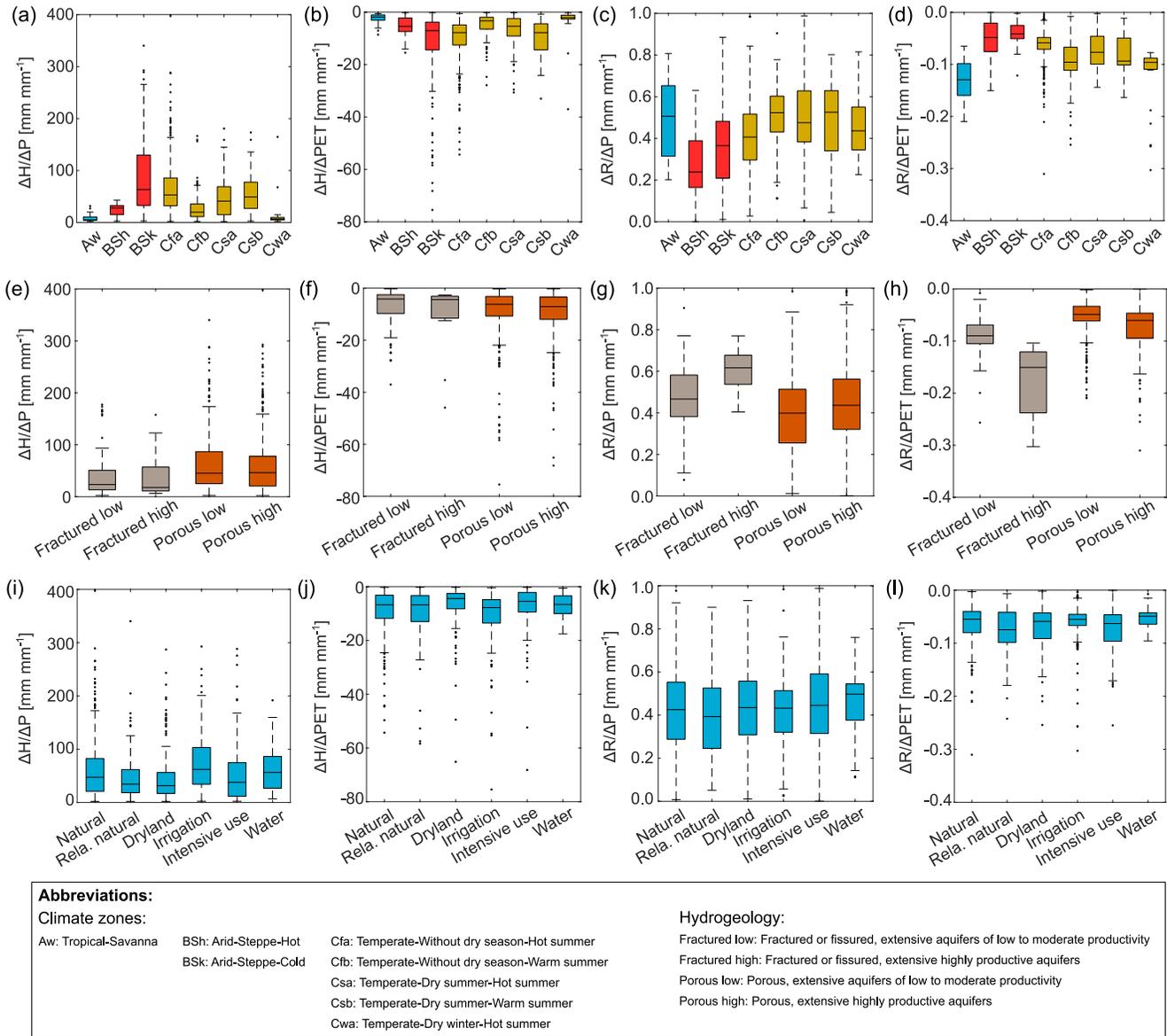


Figure 9. The distributions of groundwater level and recharge sensitivity to P and potential evapotranspiration in each (a–d) climate type, (e–h) hydrogeology, and (i–l) land use category in Australia. Note, rela. natural means “relatively natural.”

most sensitive to precipitation within more water-limited climates (arid and some temperate climates such as Csa and Csb) and within sedimentary aquifers. However, recharge appears most sensitive to precipitation within tropical and all but one (Cwa) temperate climate and within fractured aquifers. Looking at the major climate and aquifer types across Australia, this suggests that the sedimentary aquifers within the arid zones of northern VIC are most sensitive to rainfall changes, along with those at temperate sedimentary sites in southern SA and WA (Figure 5). For recharge, the tropical regions in northern Australia and temperate regions in southern VIC, WA, and SA are most sensitive to rainfall changes, particularly so within fractured aquifers. Land use however plays a minor role in recharge sensitivity, especially to P, probably because all sites are climate-dominated, where the recharge is primarily driven by P. The head however shows a different sensitivity between the agricultural uses and natural areas likely because of the anthropogenic activities that were active at some points in the past have altered the subsurface structures and properties.

To further understand how the head and recharge sensitivity changes with the aridity of the climate, Figure 11 shows that nearly all sites in Australia are in water-limited conditions ($P/PET < 1$). With the increasing aridity (i.e., decreasing P/PET), the head sensitivity increases significantly and there appear to be P/PET step changes at

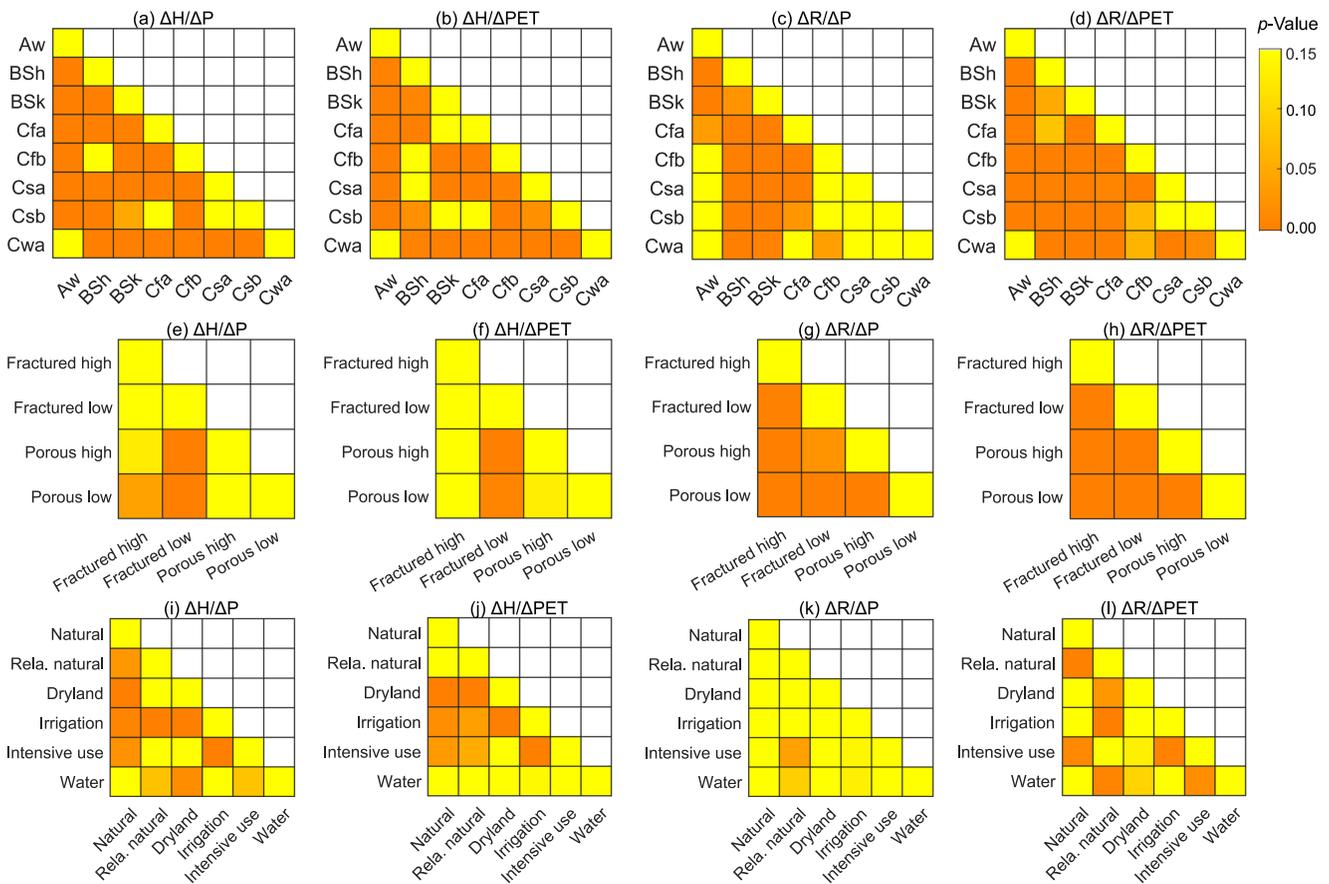


Figure 10. Differences of groundwater level and recharge sensitivity to P and potential evapotranspiration between each (a–d) climate type, (e–h) hydrogeology, and (i–l) land use category assessed with the non-parametric Wilcoxon-Mann-Whitney tests in Australia. A p -value < 0.05 means that the sensitivity is significantly different between two categories. Note, the abbreviations are the same as those in Figure 9.

thresholds of around 0.3 and 0.5 (Figures 11a and 11b). The median head sensitivity to P between the thresholds are 57, 25, and 8 mm mm^{-1} (P/PET between 0 and 0.3, 0.3 and 0.5, and 0.5 and 1.0, respectively); similarly, that to PET are -7 , -5 , and -2 mm mm^{-1} (Figures 11e and 11f). The recharge sensitivity however shows a gradual decrease with increasing climate aridity. Using the same P/PET thresholds, the median recharge sensitivity to P are 0.4, 0.5, and 0.6 mm mm^{-1} for each interval, and that to PET are -0.05 , -0.10 , and -0.16 mm mm^{-1} , respectively. Overall, these results show that the higher the aridity of the climate, the higher the head sensitivity (large step changes) whereas the lower the recharge sensitivity (more gradual changes).

4. Discussion

4.1. Climate-Driven Sites Selection

The insights into the groundwater head and recharge sensitivity across Australia, and the generalization to hydroclimatic regions, are contingent on (a) the adequacy of the *HydroSight* groundwater time-series modeling and (b) the set of climate-driven sites being unbiased with respect to site attributes and groundwater level monitoring duration and frequency.

In examining the adequacy of the modeling, the sites identified as climate-driven are found to be relatively unbiased with respect to aquifer type, record length, depth to water table, and climate; though tropical zones appear under-represented and possibly because of deficiencies in modeling of such energy-limited environments. Most climate-driven sites are within relatively natural regions or those with non-irrigated agriculture, though $\sim 20\%$ are within irrigation areas, possibly because pumping at those sites is less intense (Peterson & Fulton, 2019)) or only active for a short period in the decadal long data (≥ 16 years) and their impact turns out to be minimal. Conversely, 56%–70% of sites within natural and dryland agricultural regions are rejected as climate-driven. This

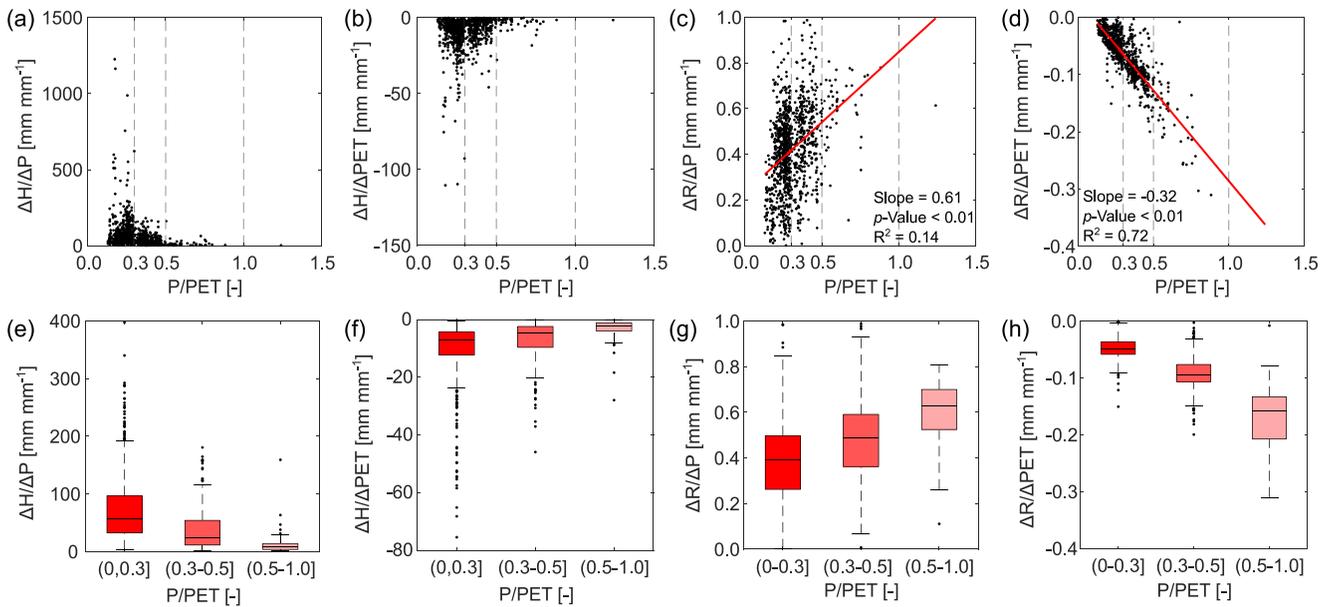


Figure 11. The sensitivity of (a, b, e, and f) groundwater level and (c, d, g, and h) recharge to P and potential evapotranspiration (PET) versus the aridity index (P/PET). The dashed lines denote the P/PET thresholds where the head sensitivity shows step changes.

might be a result of a deficient model structure for the vadose zone processes (e.g., the non-linear complexity of water infiltration and interflow) (Collenteur et al., 2021); omission of a transfer function for phreatic evapotranspiration, which could be useful for sites with long-term trends (Peterson & Western, 2014); or additional drivers omitted from the model such as land cover change, most notably plantation forestry which may decrease model performance at the sites with a shallow water table (Figure S4 in Supporting Information S1), or interaction with surface waters. Despite these challenges, the recharge estimated with *HydroSight* are found to be comparable with the estimates with independent methods or a slight overestimation in small number of bores (Crosbie et al., 2019; Kong et al., 2021; Peterson & Fulton, 2019). The estimated recharge and recharge sensitivity could hence be an upper limit. Overall, given the relatively unbiased identification of climate-driven sites, a large number of such sites ($n = 1,143$), the spatial consistency of the regions of high NSE, the relationships of the groundwater sensitivity with various climate-catchment attributes, and the propensity for non-climate driven sites to be located within intensively irrigated regions, we conclude that the approach for identification of climate-driven sites is reasonable and that the selected sites are a representative sample of bores across Australia.

4.2. Groundwater Level and Recharge Sensitivity and Specific Yield Estimation

The national median head sensitivity to P is 42 mm head change per 1 mm P, with an interquartile range of 20–77 mm mm^{-1} . Only $\sim 1\%$ ($n = 11$) sites show an extremely high head sensitivity of $>500 \text{ mm mm}^{-1}$, suggesting a low specific yield. At all sites the head has a positive sensitivity with P, which is as expected given the physics. Given that an aquifer having a specific yield of 1%, 10% of precipitation going to recharge would have a head sensitivity to P of 10 mm mm^{-1} (see Equation 12). The results of head sensitivity to P are plausible though slightly higher than expected. With regards to head sensitivity to PET, the median is -6 mm mm^{-1} , with an interquartile range of -3 to -11 mm mm^{-1} . As expected, all sites have a negative sensitivity with PET. Looking at the spatial consistency of the head sensitivity (P and PET), sites within 30 km of each other show a clear spatial correlation and $\sim 60\%$ of the variability is explained by this separation distance; with the remaining $\sim 40\%$ arising from very local scale soil and aquifer heterogeneity and variability introduced by the modeling. Overall, given that the head sensitivity are within a plausible range, have the correct sign and are spatially correlated despite being individually derived, the results are considered reasonable.

The national median recharge sensitivity to P is 0.43 mm recharge change per 1 mm P (interquartile: 0.30–0.55 mm mm^{-1}). Again, only $\sim 1\%$ ($n = 10$) sites show an extremely high recharge sensitivity ($>0.90 \text{ mm mm}^{-1}$). All sites have a positive sensitivity with P and more importantly less than 1 mm mm^{-1} , which

is again as expected given the physics. Given that rainfall partitioning is often dominated by PET and runoff, rather than recharge, and that one additional millimeter of P should produce significantly less than one additional millimeter of recharge, this further supports the physical validity of the results. Looking at the spatial consistence of the recharge sensitivity to P and PET, sites within 268 and 651 km respectively of each other show a clear spatial correlation. Nearly all (98%) of the variability is explained by the separation distance for the sensitivity to PET, in spite of $\sim 40\%$ for that to P because of the higher spatial variability of precipitation than evapotranspiration. Despite that recharge is inferred from the modeling and is not measured like head, the physical validity of the results and their high spatial correlation suggest that the results are reasonable.

The ratio of the head sensitivity to recharge sensitivity to P (Equation 12) enables estimation of the aquifer specific yield (S_y). The national median is 0.9% with an interquartile range of 0.5%–2.0%. Encouragingly the S_y is between the physical limits of 0% and 100% at all sites. Further evaluation of the validity of the results is, however, difficult given that very few independent estimates exist nationally, most likely because of the high costs of pumping tests and those that exist have a high uncertainty (Chowdhury et al., 2022; Crosbie et al., 2015). That said, for northern Australia it is estimated at 1.7%–3.6%, while in SA it is estimated at 7.5%–30%, and globally at 0%–43% (Crosbie et al., 2015, 2019; Lv et al., 2021). While the median estimate here is below this range for Australia, the distribution of results shown in Figure 6j is well within this range. Importantly however, the results presented here and the national estimates are unlikely to be directly comparable because the former is an estimate of S_y at the water table-unsaturated zone intersection while the latter is generally a vertically integrated average estimate of the entire aquifer; a difference which needs to be considered if applying these results.

Unsurprisingly the head sensitivity (P and PET) is positively highly correlated with S_y (Figure 7b); that is, lower S_y should result in a greater head change per millimeter of recharge. However, the recharge sensitivity is positively, but weakly, correlated with the S_y . This has no obvious physical explanation other a possible correlation between vadose zone vertical conductivity and S_y (i.e., more vertically conductive soils have a higher aquifer porosity) (Chen et al., 1999, 2010) or a spurious correlation arising from recharge sensitivity being used to estimate S_y .

Looking at the spatial consistence of the S_y estimates, sites within 204 km of each other show a clear spatial correlation and 99% of the variability is explained by this separation distance. The spatial correlation is broadly consistent with the extent of aquifers, which supports the finding that the estimate is representative of the physical properties of the aquifers. This is highly encouraging given that each estimate of S_y is derived independently of other estimates. Overall, despite the difficulties in independently evaluating individual S_y estimates, there is no evidence to invalidate them and considerable regional scale supporting evidence.

4.3. Governing Factors of Groundwater Level and Recharge Sensitivity

The groundwater level and recharge sensitivity to P and PET are found to significantly differ with climate types and hydrogeology (Figures 9 and 10). The inherent properties of the sites thus play a crucial role in governing groundwater sensitivity.

The head is most sensitive in arid climates and porous media and least sensitive in tropical climates and fractured media. The porous aquifers show higher head sensitivity than the fractured ones, probably because porous media have a lower specific yield which causes a larger head change per unit change of recharge (Johnson, 1967). In contrast, fractured aquifers show higher recharge sensitivity likely because of the impact of hydrogeology on the generation of recharge (e.g., matrix-flow only in porous media compared with dual-process of matrix- and fracture-flow in fractured media) (Manna et al., 2017). For the impact of climate type, the head sensitivity to P and PET is found to significantly increase with climate aridity (Figures 11e and 11f). Given that the aridity is projected to expand globally and become more severe in many countries and regions under climate change, such as Australia and Mediterranean regions (Feng & Fu, 2013; Lickley & Solomon, 2018), the head sensitivity may increase over the coming decades, especially in the arid zone with porous media.

Additionally, land use also has a modest impact on the head sensitivity (Figures 9 and 10). For example, areas with agricultural activities show a higher head sensitivity than the natural sites. Interestingly, the depth to water table however does not show a clear relationship with the head and recharge sensitivity (Figure S3 in Supporting Information S1), indicating that the declining water level will likely not significantly change the sensitivity. Together, these findings suggest that agricultural activities may increase the head sensitivity, not because of a decreasing groundwater level (e.g., water extraction) but possibly because of disturbance to the subsurface

properties and processes; that said, the differences may arise simply because of the locations chosen for agriculture being hydrologically different from nature regions.

4.4. Implications for Groundwater Management

Quantifying the groundwater level and recharge sensitivity is practically useful for protecting groundwater-dependent ecosystems (GDEs) and managing groundwater resources under the impact of climate change. For example, quantifying and overlaying the groundwater level sensitivity map with that of the GDEs allows the vulnerable GDEs to be identified. That is, the GDEs where the head sensitivity to precipitation is high may likely respond quickly to droughts, such as the ones located in central Victoria (Figure S5 in Supporting Information S1) and will need more attention and protection. Second, understanding recharge sensitivity to climate variations is valuable as water managers need this information to regulate groundwater extraction rates, which are often set as a fraction of the recharge (times the capture area) (White et al., 2016). Besides that, regions with a high recharge sensitivity to precipitation could be more vulnerable to groundwater pumping, especially during droughts due to the significantly declined recharge with precipitation.

Here we also introduce an approach to estimate the specific yield, which is a key aquifer hydraulic property and highly desirable for groundwater management. Given how sparse such estimates often are and their high utility in estimating the aquifer response time to pumping (Chowdhury et al., 2022; Crosbie et al., 2015), our results offer a considerable opportunity to the management of groundwater usage. The introduced approach enables to estimate specific yield at any site using long-term head and climate records, therefore avoiding the logistical and economic challenges of obtaining specific yield estimates through in situ measurement. Compared with the specific yield reported in other studies (Crosbie et al., 2015, 2019; Lv et al., 2021), this method provides a plausible estimate. A caveat of the method is that the estimated specific yield is only for the top part of the aquifer where the head fluctuates, and not an estimate of the entire (especially the deeper portion of the) aquifer.

5. Conclusions

Climate change is projected to impact water resources in many countries, but the predictions are highly uncertain because of the significant uncertainty in future rainfall projections and that they are heavily dependent upon model structures, climate change scenarios, and assumptions about the stationarity of hydrological processes (Oreskes et al., 1994; Peterson et al., 2021; Toews & Allen, 2009a). Given that groundwater has a long memory of historic climate variability and trends, analyzing groundwater hydrographs at climate-dominated sites potentially allows for estimating its long-term mean sensitivity to climate variations, which is an essential first step toward understanding climate change impact.

In this study, we aim to investigate (a) groundwater level and recharge sensitivity to climate variations across Australia and (b) in which environments is groundwater most sensitive to climate variations. In doing this, we first identify 1,143 climate-driven sites using a time-series groundwater toolbox *HydroSight* (Peterson & Fulton, 2019; Peterson & Western, 2014). We then adopt a MLR approach, adapted from streamflow elasticity studies, to quantify groundwater sensitivity. Results show that the head and recharge sensitivity to P is around 8 times that to PET. The national median changes in head and recharge per 1 mm P change are 42 and 0.43 mm respectively. Southeast Australia shows the highest head sensitivity, whereas SWA shows the highest recharge sensitivity. Overall, the sensitivity and specific yield estimates are found to be highly consistent with available independent evidence. Results also show that climate type and hydrogeology are the primary controlling factors, with the highest head sensitivity in arid climates and porous media and the highest recharge sensitivity appearing in tropical climates and fractured media. Moreover, land use change has a modest influence. Practically, the results contribute to identification of GDEs and groundwater engagement areas that are vulnerable to climate variability.

Additionally, the sensitivity results enable estimation of the possible impacts of climate change on groundwater. For example, head sensitivity is found to increase with climate aridity and given that arid climate zones are projected to expand and intensify in many regions and countries globally, such as Australia and Mediterranean regions (Feng & Fu, 2013; Lickley & Solomon, 2018), the groundwater head in such regions is likely to become more temporally variable. More generally, multiplication of a head sensitivity estimate by a predicted change in rainfall enables a first order estimate of the possible impact of climate change on head; and similarly for recharge

sensitivity. Climate change is, however, likely to produce more complex and nonlinear changes to head and recharge sensitivity than studied here and others are encouraged to develop and apply techniques to quantify the change in sensitivity over time.

Data Availability Statement

The time-series groundwater level data, bore coordinates, and bore depths across Australia are available on the Groundwater Explorer of the Australian Government Bureau of Meteorology, available at <http://www.bom.gov.au/water/groundwater/explorer/map.shtml>, under the Creative Commons Attribution 4.0 Australia Licence, retrieved in March 2019. The climate data at any given bore coordinates in Australia are extracted using a R-package *AWAPer* version 0.1.46 (Peterson et al., 2020) from the gridded Australian national meteorological data repository (Jones et al., 2009), available at <http://www.bom.gov.au/climate/>. The Morton CRAE potential evapotranspiration (Morton, 1983) is calculated with the R package *Evapotranspiration* version 1.16 (Guo et al., 2016). The Australian land use data are obtained from the Australian Bureau of Agriculture and Resources Economics and Sciences (ABARES) (2021), available at <https://www.agriculture.gov.au/abares/aclump/land-use/data-download>. The Australian hydrogeology data are sourced from the Australian Bureau of Mineral Resources (Jacobson & Lau, 1987; Lau et al., 1987), available at pid.geoscience.gov.au/dataset/ga/32368. The groundwater hydrograph modeling software *HydroSight* version 1.40.2.1 (Peterson & Fulton, 2019; Peterson & Western, 2014) is used and deposited at Peterson (2022) and actively maintained at <https://github.com/peterson-tim-j/HydroSight>.

References

- Allen, D. M., Mackie, D. C., & Wei, M. (2004). Groundwater and climate change: A sensitivity analysis for the Grand Forks aquifer, southern British Columbia, Canada. *Hydrogeology Journal*, 12(3), 270–290. <https://doi.org/10.1007/s10040-003-0261-9>
- Amanambu, A. C., Obarein, O. A., Mossa, J., Li, L., Ayeni, S. S., Balogun, O., et al. (2020). Groundwater system and climate change: Present status and future considerations. *Journal of Hydrology*, 589, 125163. <https://doi.org/10.1016/j.jhydrol.2020.125163>
- Andréassian, V., Coron, L., Lerat, J., & Le Moine, N. (2016). Climate elasticity of streamflow revisited – An elasticity index based on long-term hydrometeorological records. *Hydrology and Earth System Sciences*, 20(11), 4503–4524. <https://doi.org/10.5194/hess-20-4503-2016>
- Atawneh, D. A., Cartwright, N., & Bertone, E. (2021). Climate change and its impact on the projected values of groundwater recharge: A review. *Journal of Hydrology*, 601, 126602. <https://doi.org/10.1016/j.jhydrol.2021.126602>
- Australian Bureau of Agriculture and Resources Economics and Sciences (ABARES). (2021). Catchment scale land use of Australia - Update December 2020 [Dataset]. <https://doi.org/10.25814/aqjw-rq15>
- Australian Government Bureau of Meteorology. (2012). Australian water resources assessment 2012 summary report. Retrieved from www.bom.gov.au/water/awra/2012
- Bloomfield, J. P., Marchant, B. P., & McKenzie, A. A. (2019). Changes in groundwater drought associated with anthropogenic warming. *Hydrology and Earth System Sciences*, 23(3), 1393–1408. <https://doi.org/10.5194/hess-23-1393-2019>
- Chen, X., Goeke, J., & Summerside, S. (1999). Hydraulic properties and uncertainty analysis for an unconfined alluvial aquifer. *Groundwater*, 37(6), 845–854. <https://doi.org/10.1111/j.1745-6584.1999.tb01183.x>
- Chen, X., Song, J., & Wang, W. (2010). Spatial variability of specific yield and vertical hydraulic conductivity in a highly permeable alluvial aquifer. *Journal of Hydrology*, 388(3–4), 379–388. <https://doi.org/10.1016/j.jhydrol.2010.05.017>
- Chiew, F. H. S. (2006). Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Sciences Journal*, 51(4), 613–625. <https://doi.org/10.1623/hysj.51.4.613>
- Chowdhury, F., Gong, J., Rau, G. C., & Timms, W. A. (2022). Multifactor analysis of specific storage estimates and implications for transient groundwater modelling. *Hydrogeology Journal*, 30(7), 2183–2204. <https://doi.org/10.1007/s10040-022-02535-z>
- Chu, W., Gao, X., & Sorooshian, S. (2011). A new evolutionary search strategy for global optimization of high-dimensional problems. *Information Sciences*, 181(22), 4909–4927. <https://doi.org/10.1016/j.ins.2011.06.024>
- Collenteur, R. A., Bakker, M., Klammler, G., & Birk, S. (2021). Estimation of groundwater recharge from groundwater levels using nonlinear transfer function noise models and comparison to lysimeter data. *Hydrology and Earth System Sciences*, 25(5), 2931–2949. <https://doi.org/10.5194/hess-25-2931-2021>
- Cooper, M., Schaperow, J., Cooley, S., Alam, S., Smith, L., & Lettenmaier, D. (2018). Climate elasticity of low flows in the maritime western us mountains. *Water Resources Research*, 54(8), 5602–5619. <https://doi.org/10.1029/2018WR022816>
- Cressie, N. (2015). *Statistics for spatial data*. John Wiley & Sons.
- Crosbie, R. S., Davies, P., Harrington, N., & Lamontagne, S. (2015). Ground truthing groundwater-recharge estimates derived from remotely sensed evapotranspiration: A case in South Australia. *Hydrogeology Journal*, 23(2), 335–350. <https://doi.org/10.1007/s10040-014-1200-7>
- Crosbie, R. S., Doble, R. C., Turnadge, C., & Taylor, A. R. (2019). Constraining the magnitude and uncertainty of specific yield for use in the water table fluctuation method of estimating recharge. *Water Resources Research*, 55(8), 7343–7361. <https://doi.org/10.1029/2019WR025285>
- Cuthbert, M. O., Taylor, R. G., Favreau, G., Todd, M. C., Shamsudduha, M., Villholth, K. G., et al. (2019). Observed controls on resilience of groundwater to climate variability in sub-Saharan Africa. *Nature*, 572(7768), 230–234. <https://doi.org/10.1038/s41586-019-1441-7>
- Domenico, P. A., & Schwartz, F. W. (1998). *Physical and chemical hydrogeology*. Wiley.
- Feng, S., & Fu, Q. (2013). Expansion of global drylands under a warming climate. *Atmospheric Chemistry and Physics*, 13(19), 10081–10094. <https://doi.org/10.5194/acp-13-10081-2013>
- Fowler, K., Knoben, W., Peel, M., Peterson, T., Ryu, D., Saft, M., et al. (2020). Many commonly used rainfall-runoff models lack long, slow dynamics: Implications for runoff projections. *Water Resources Research*, 56(5), 1–27. <https://doi.org/10.1029/2019WR025286>
- Fu, G., Crosbie, R. S., Barron, O., Charles, S. P., Dawes, W., Shi, X., et al. (2019). Attributing variations of temporal and spatial groundwater recharge: A statistical analysis of climatic and non-climatic factors. *Journal of Hydrology*, 568, 816–834. <https://doi.org/10.1016/j.jhydrol.2018.11.022>

- Green, T. R. (2016). *Linking climate change and groundwater*. Springer International Publishing. https://doi.org/10.1007/978-3-319-23576-9_5
- Green, T. R., Taniguchi, M., Kooi, H., Gurdak, J. J., Allen, D. M., Hiscock, K. M., et al. (2011). Beneath the surface of global change: Impacts of climate change on groundwater. *Journal of Hydrology*, *405*(3–4), 532–560. <https://doi.org/10.1016/j.jhydrol.2011.05.002>
- Guo, D., Westra, S., & Maier, H. R. (2016). An R package for modelling actual, potential and reference evapotranspiration [Software]. *Environmental Modelling & Software*, *78*, 216–224. <https://doi.org/10.1016/j.envsoft.2015.12.019>
- Healy, R. W., & Cook, P. G. (2002). Using groundwater levels to estimate recharge. *Hydrogeology Journal*, *10*(1), 91–109. <https://doi.org/10.1007/s10040-001-0178-0>
- Henley, B. J., Peel, M. C., Nathan, R., King, A. D., Ukkola, A. M., Karoly, D. J., & Tan, K. S. (2019). Amplification of risks to water supply at 1.5°C and 2°C in drying climates: A case study for Melbourne, Australia. *Environmental Research Letters*, *14*(8), 084028. <https://doi.org/10.1088/1748-9326/ab26ef>
- Hill, M. C., & Tiedeman, C. R. (2006). *Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty*. John Wiley and Sons.
- IPCC. (2014). *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change*. IPCC.
- IPCC. (2021). *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.
- Isaaks, E. H., & Srivastava, R. M. (1990). *An introduction to applied geostatistics*. Oxford University Press.
- Jacobson, G., & Lau, J. C. (1987). Hydrogeology of Australia (1:5000000 scale map) [Dataset]. Retrieved from pid.geoscience.gov.au/dataset/ga/32368
- Jiménez-Cisneros, B. E., Oki, T., Arnell, N. W., Benito, G., Cogley, J. G., Döll, P., et al. (2014). Freshwater resources. In *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.
- Johnson, A. I. (1967). *Specific yield: Compilation of specific yields for various materials (No. 1662)*. US Government Printing Office.
- Jones, D. A., Wang, W., & Fawcett, R. (2009). High-quality spatial climate data-sets for Australia [Dataset]. Australian Meteorological and Oceanographic Journal, *58*(4), 233–248. <https://doi.org/10.22499/2.5804.003>
- Kong, F., Song, J., Crosbie, R. S., Barron, O., Schafer, D., & Pigois, J.-P. (2021). Groundwater hydrograph decomposition with the HydroSight model. *Frontiers in Environmental Science*, *9*, 1–19. <https://doi.org/10.3389/fenvs.2021.736400>
- Kotchoni, D. O. V., Vouillamoz, J.-M., Lawson, F., Adjomayi, P., Boukari, M., & Taylor, R. G. (2019). Relationships between rainfall and groundwater recharge in seasonally humid Benin: A comparative analysis of long-term hydrographs in sedimentary and crystalline aquifers. *Hydrogeology Journal*, *27*(2), 447–457. <https://doi.org/10.1007/s10040-018-1806-2>
- Kumar, C. P. (2016). Assessing the impact of climate change on groundwater resources. *IWRA (India) Journal*, *5*(1), 3–11. Retrieved from <https://www.indianjournals.com/ijor.aspx?target=ijor:iwra&volume=5&issue=1&article=001>
- Lafayette, L., Sauter, G., Vu, L., & Meade, B. (2016). *Spartan performance and flexibility: An HPC-cloud chimera*. OpenStack Summit. <https://doi.org/10.4225/49/58ead90dceaaa>
- Lau, J., Commander, D., & Jacobson, G. (1987). Hydrogeology of Australia [Dataset]. Bureau of Mineral Resources. Retrieved from pid.geoscience.gov.au/dataset/ga/32368
- Lickley, M., & Solomon, S. (2018). Drivers, timing and some impacts of global aridity change. *Environmental Research Letters*, *13*(10), 104010. <https://doi.org/10.1088/1748-9326/aae013>
- Liu, D. L., Teng, J., Ji, F., Anwar, M. R., Feng, P., Wang, B., et al. (2021). Characterizing spatiotemporal rainfall changes in 1960–2019 for continental Australia. *International Journal of Climatology*, *41*(S1), E2420–E2444. <https://doi.org/10.1002/joc.6855>
- Lv, M., Xu, Z., Yang, Z.-L., Lu, H., & Lv, M. (2021). A comprehensive review of specific yield in land surface and groundwater studies. *Journal of Advances in Modeling Earth Systems*, *13*(2), e2020MS002270. <https://doi.org/10.1029/2020MS002270>
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. <https://doi.org/10.1214/aoms/1177730491>
- Manna, F., Walton, K. M., Cherry, J. A., & Parker, B. L. (2017). Mechanisms of recharge in a fractured porous rock aquifer in a semi-arid region. *Journal of Hydrology*, *555*, 869–880. <https://doi.org/10.1016/j.jhydrol.2017.10.060>
- Meinzer, O. E. (1923). *The occurrence of ground water in the United States with a discussion of principles* (Vol. 50, No. (846)). University of Chicago.
- Morton, F. I. (1983). Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, *66*(1–4), 1–76. [https://doi.org/10.1016/0022-1694\(83\)90177-4](https://doi.org/10.1016/0022-1694(83)90177-4)
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Opie, S., Taylor, R. G., Brierley, C. M., Shamsudduha, M., & Cuthbert, M. O. (2020). Climate–groundwater dynamics inferred from GRACE and the role of hydraulic memory. *Earth System Dynamics*, *11*(3), 775–791. <https://doi.org/10.5194/esd-11-775-2020>
- Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the Earth sciences. *Science*, *263*(5147), 641–646. <https://doi.org/10.1126/science.263.5147.641>
- Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. <https://doi.org/10.5194/hess-11-1633-2007>
- Peterson, T. J. (2022). HydroSight (v1.40.2.1) [Software]. Zenodo. <https://doi.org/10.5281/zenodo.10023312>
- Peterson, T. J., & Fulton, S. (2019). Joint estimation of gross recharge, groundwater usage, and hydraulic properties within HydroSight [Software]. *Groundwater*, *57*(6), 860–876. <https://doi.org/10.1111/gwat.12946>
- Peterson, T. J., Saft, M., Peel, M. C., & John, A. (2021). Watersheds may not recover from drought. *Science*, *372*(6543), 745–749. <https://doi.org/10.1126/science.abd5085>
- Peterson, T. J., Wasko, C., Saft, M., & Peel, M. (2020). AWAPer: An R package for area weighted catchment daily meteorological data anywhere within Australia [Software]. *Hydrological Processes*, *34*(5), 1301–1306. <https://doi.org/10.1002/hyp.13637>
- Peterson, T. J., & Western, A. W. (2014). Nonlinear time-series modeling of unconfined groundwater head [Software]. *Water Resources Research*, *50*(10), 8330–8355. <https://doi.org/10.1002/2013WR014800>
- Peterson, T. J., & Western, A. W. (2018). Statistical interpolation of groundwater hydrographs. *Water Resources Research*, *54*(7), 4663–4680. <https://doi.org/10.1029/2017WR021838>
- Refsgaard, J. C., Madsen, H., Andréassian, V., Arnbjerg-Nielsen, K., Davidson, T. A., Drews, M., et al. (2014). A framework for testing the ability of models to project climate change and its impacts. *Climatic Change*, *122*(1–2), 271–282. <https://doi.org/10.1007/s10584-013-0990-2>

- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., et al. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, *114*, 29–39. <https://doi.org/10.1016/j.envsoft.2019.01.012>
- Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., et al. (2013). Ground water and climate change. *Nature Climate Change*, *3*(4), 322–329. <https://doi.org/10.1038/nclimate1744>
- Toews, M. W., & Allen, D. M. (2009a). Evaluating different GCMs for predicting spatial recharge in an irrigated arid region. *Journal of Hydrology*, *374*(3–4), 265–281. <https://doi.org/10.1016/j.jhydrol.2009.06.022>
- Toews, M. W., & Allen, D. M. (2009b). Simulated response of groundwater to predicted recharge in a semi-arid region using a scenario of modelled climate change. *Environmental Research Letters*, *4*(3), 035003. <https://doi.org/10.1088/1748-9326/4/3/035003>
- Ukkola, A. M., Roderick, M. L., Barker, A., & Pitman, A. J. (2019). Exploring the stationarity of Australian temperature, precipitation and pan evaporation records over the last century. *Environmental Research Letters*, *14*(12), 124035. <https://doi.org/10.1088/1748-9326/ab545c>
- Van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., et al. (2013). The Millennium drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. <https://doi.org/10.1002/wrcr.20123>
- Vano, J. A., Das, T., & Lettenmaier, D. P. (2012). Hydrologic sensitivities of Colorado river runoff to changes in precipitation and temperature. *Journal of Hydrometeorology*, *13*(3), 932–949. <https://doi.org/10.1175/jhm-d-11-069.1>
- Von Asmuth, J. R., & Bierkens, M. F. (2005). Modeling irregularly spaced residual series as a continuous stochastic process. *Water Resources Research*, *41*(12), W12404. <https://doi.org/10.1029/2004WR003726>
- Von Asmuth, J. R., Bierkens, M. F. P., & Maas, K. (2002). Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research*, *38*(12), 23-1–23-12. <https://doi.org/10.1029/2001WR001136>
- White, E. K., Peterson, T. J., Costelloe, J., Western, A. W., & Carrara, E. (2016). Can we manage groundwater? A method to determine the quantitative testability of groundwater management plans. *Water Resources Research*, *52*(6), 4863–4882. <https://doi.org/10.1002/2015WR018474>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83. <https://doi.org/10.2307/3001968>