

ACCEPTED MANUSCRIPT • OPEN ACCESS

Quantifying uncertainty in groundwater depth from sparse well data in the California Central Valley

To cite this article before publication: Amanda Quay *et al* 2020 *Environ. Res. Lett.* in press <https://doi.org/10.1088/1748-9326/ab88fb>

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2020 The Author(s). Published by IOP Publishing Ltd.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

**Quantifying uncertainty in groundwater depth
from sparse well data in the California Central
Valley**

Environmental Research Letters

Amanda N. Quay[†], Amanda S. Hering[‡], and Meagan S. Mauter^{†*}

[†] Department of Civil and Environmental Engineering, Stanford University,
Stanford, California, 94305

[‡] Department of Statistical Science, Baylor University, Waco, Texas, 76798

* Corresponding author: email: mauter@stanford.edu; Tel. +1 (650) 725-4911

18 ABSTRACT

19 Groundwater is a critical freshwater resource for irrigation in the California Central Valley,
20 particularly in times of drought. Groundwater depth has dropped rapidly in California's
21 overdrafted basins, but irregular monitoring across space and time limits the accuracy of the
22 groundwater depth projections in the Groundwater Sustainability Plans required by the
23 California Sustainable Groundwater Management Act (SGMA). This work constructs a
24 Bayesian hierarchical model for predicting groundwater depth from sparse monitoring data in
25 three Central Valley counties. We apply this model to generate 300 m resolution monthly
26 groundwater depth estimates for drought years 2013-2015, and compare our smoothed
27 groundwater depth map to smoothed rasterized maps published by the CA Department of Water
28 Resources. Finally, we quantify uncertainty in groundwater depth predictions that are made by
29 imputing missing well data and interpolating predictions across the study domain, which is
30 helpful in directing future sampling efforts towards areas with high uncertainty. The BHM model
31 accurately captures the spatiotemporal pattern in groundwater depth, as evidenced by 94.54% of
32 withheld test samples' true depth being covered by the 95% prediction interval drawn from the
33 BHM posterior distribution. The model converged despite a very sparse dataset, demonstrating
34 broad applicability for evaluating changes in regional groundwater depth as required by SGMA.
35 Depth prediction intervals can also help prioritize future groundwater depth sampling activity
36 and increase the utility of groundwater depth maps in total storage predictions by enabling
37 sensitivity analysis.

INTRODUCTION

Groundwater is essential to the California Central Valley’s (CV) highly productive agricultural economy,¹ serving as a critical buffer against both seasonal variations in water availability and short-term droughts.^{2, 3} Groundwater withdrawals peak during late summer months when agricultural water demand is highest and surface water availability is limited.^{2, 4} Groundwater also serves as the primary water source for perennial crop growers during periods of extreme drought. For example, during the 2012-2016 drought, groundwater increased from 43% to 70% of agricultural water supply in the CV.⁵

As a result of the California CV’s extensive dependence on groundwater, the region has suffered from serious groundwater overdraft and its associated consequences.⁶ Groundwater storage in the CV aquifer system has been depleted at an average rate of 1.85 km³ per year since 1962,⁷ and during the 2012-2016 drought, groundwater depletion was roughly quadruple that rate.⁸ This level of groundwater overdraft causes a multitude of interconnected consequences, including permanent reductions in groundwater storage potential, land subsidence, seawater intrusion, and increased groundwater pumping costs for growers.^{5, 9}

In 2014, California lawmakers passed the Sustainable Groundwater Management Act (SGMA), creating the state’s first enforceable groundwater management guidelines and providing a framework for basin-level groundwater decision making.^{5, 6, 10} Under SGMA, basins deemed high- or medium-priority must achieve long-term groundwater sustainability over a 20-year period, halt overdraft, and bring groundwater basins into balanced pumping and recharge.¹¹⁻¹³ This must be accomplished without causing undesirable side effects such as significant and unreasonable groundwater storage reduction or land subsidence.¹¹⁻¹⁴ SGMA also grants authority to groundwater sustainability agencies (GSAs) and, if necessary, the State to monitor groundwater, develop and implement groundwater sustainability plans, and enforce these plans through groundwater use fees and limits on extraction.

The success of SGMA will depend on the accuracy of groundwater monitoring and modeling, but existing tools suffer from limited generalizability across basins and poor quantification of data and model uncertainty. Well-level spatial and temporal monitoring of depth-to-groundwater provides accurate readings proximal to the well, but well-level monitoring over large spatial regions or temporal periods requires costly large-scale data collection.¹⁰

California's most comprehensive groundwater monitoring effort, the California Statewide Groundwater Elevation Monitoring Program (CASGEM), provides depth-to-groundwater measurements collected by the Department of Water Resources (DWR) and local monitoring parties.^{14, 15} DWR produces interpolated maps of groundwater depth in the most sample-dense geographic regions for the months of March and October, when sampling efforts peak.¹⁵ As such, these maps provide limited geographic coverage and exclude data from less frequently sampled months that could be used to inform temporal trends. Additionally, the interpolation techniques used to generate CASGEM maps cannot differentiate between real changes in groundwater depth and fluctuations in data availability from non-uniform spatiotemporal sampling patterns.¹⁶ Finally, existing CASGEM maps do not report uncertainty in depth to groundwater estimates. This impedes policy decision making and efficient identification of data gaps to improve well sampling efficiency.¹⁷ Neither conventional interpolation models nor sophisticated simulation models estimate groundwater levels with the high spatial and temporal resolution relevant to GSA, while also providing robust uncertainty quantification.¹⁸

Bayesian hierarchical models (BHM) have gained traction for environmental science applications due to their ability to translate spatial and temporal information into conditional models on a wide variety of scales and to propagate the uncertainty of missing values into subsequent analysis.¹⁹⁻²² The popularity of BHM has coincided with rapid developments in computational technology that allow for greater model complexity.²³⁻²⁵ For example, Manago et al. (2019) developed a BHM for groundwater depth in urban Los Angeles, offering depth prediction that is more robust to varying frequency of observed well data than traditional interpolation methods.¹⁶ BHM allow for multiple imputation of missing well data and, therefore, reflect the uncertainty of the missing values in subsequent analyses such as spatial prediction and risk analyses.^{16, 26} While the BHM's hierarchical approach harnesses simple models to capture a complex posterior distribution,^{27, 28} the computational complexity of specifying a BHM can become impractical as data density decreases. A model offering unambiguous uncertainty estimation, such as a BHM, is useful in prioritizing sampling locations to improve prediction uncertainty when environmental sampling is cost-limited.

Motivated by the need for locally-relevant and near-real-time groundwater level monitoring with quantifiable uncertainty, this work develops a BHM that is subsequently used

for multiple imputation and interpolation to generate monthly groundwater depth maps in California’s Central Valley at 300 m spatial resolution. The model addresses current gaps in depth-to-groundwater predictions by incorporating the uncertainty due to both imputing missing data and making predictions at new locations. This approach maintains higher temporal resolution of publicly available depth-to-groundwater reports than existing simple interpolation models, while also quantifying uncertainty in those predictions. We expand upon the model presented by Manago et al. (2019),¹⁶ modifying their BHM formulation for a substantially larger spatial area across the Central Valley counties of San Joaquin, Stanislaus, and Merced. Finally, we describe a process for constructing a BHM with sparse data, probing computational limits of the BHM and extending the model’s applicability to large, sparse spatiotemporal datasets. Our model offers computational efficiency improvements such as task parallelization, allowing for greater uncertainty handling over a larger spatial domain. This framework can be applied to other groundwater monitoring areas under SGMA and to other water-stressed areas outside of California, offering insight into when and where additional data collection would be most useful to reducing uncertainty in groundwater depth predictions.

METHODOLOGY

To estimate monthly groundwater levels across San Joaquin, Stanislaus, and Merced in 2013-2015, we proceed through the following modeling framework. First, we collected, cleaned, and aggregated groundwater depth data from the California DWR into monthly, 300 m resolution space-time observation cells. After withholding a randomly selected 10% of observed cells for out-of-sample testing, we initialized the BHM with vague priors and iterated its Markov chain algorithm for 11,000 steps to yield a posterior distribution for each parameter in the model. Next, we used out-of-sample data to internally validate the BHM. For monthly predictions of groundwater depth across the study area, multiple imputation of missing cells was performed prior to Ordinary Kriging. We then generated groundwater maps of the predicted mean and 95% prediction interval. Finally, we compared modeling results from our BHM’s monthly interpolation products with DWR’s groundwater depth maps, where available and intersecting.

Data and Study Area. The study area is the 20 km-buffered intersection of San Joaquin, Stanislaus, and Merced counties with CA Bulletin 118 groundwater basins as defined by the

California DWR (Figure 1a). We obtain groundwater measurements for 2013-2015 from the DWR's CASGEM program. The data represent the distance from the ground surface to the water table level, so a larger distance corresponds to a lower groundwater level.

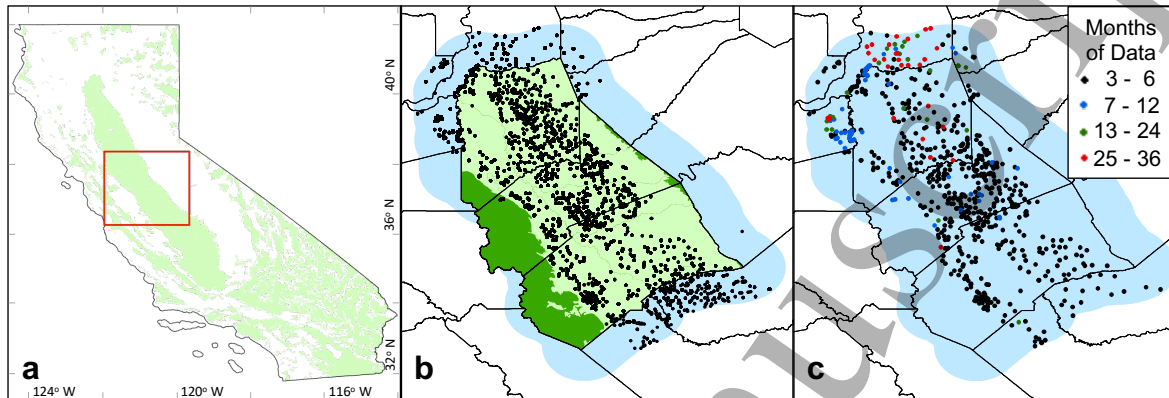


Figure 1. Spatial and temporal extent of study area. (a) CA Department of Water Resources' Bulletin 118 Groundwater Basins data product defines boundaries for California's groundwater basins, provided in light green. (b) Groundwater wells reporting data over the study time period (black) are heterogeneously dispersed. The study area is the intersection of Merced, San Joaquin, and Stanislaus counties (dark green) with the groundwater basin boundaries, plus a 20 km buffer of this intersection (light blue). (c) Groundwater well locations after cleaning and aggregation. Data are sparse, with most cells reporting less than seven months over the 36-month study time period.

Wells reporting groundwater depth over the 2013-2015 period are heterogeneously dispersed across the study area (Figure 1b). Depth-to-groundwater data are also heterogeneous across time, with both the sampling frequency and the spatial distribution of sampling varying widely from month to month. After performing quality control on the data to remove surface-water level readings, wells were spatially aggregated into 300 m resolution cells. We also explored spatial aggregation of the wells at 150 m and 600 m resolution. Cells with an in-cell, in-month standard deviation of > 9 m were assumed erroneous due to measurement or input error and were removed from the dataset. The total percentage of data removed was 2.91% across the three years.

The 300 m-resolved spatial cells were consolidated into monthly temporal bins. We averaged each monthly temporal bin's groundwater readings by applying equal weight to each well in the cell. This avoided over-weighting more frequently-sampled wells. Some cells contain well(s) reporting only a few times a year, while other cells contain groundwater depth data for as many as 33 months over the 36-month study period. In order to remove overly sparse spatial locations and facilitate model convergence, we removed cells containing less than three monthly depth

readings over the study period from the dataset. Cells with higher temporal reporting density were clustered in the northern quadrant of the study area (Figure 1c).

After data cropping, cleaning, and aggregation, there were 690 cells reporting data across 36 monthly time steps. The mean distance between monitoring wells was $\bar{D} = 83.2 \text{ km}$. The mean and median missingness in the dataset were 86.39% and 90.58%, respectively, indicating that the dataset is very sparse. The sparseness of this dataset exceeds that of prior BHM implementations for groundwater depth by nearly an order of magnitude.¹⁶ Only about 10% of our cells had data presence each month annually, with peaks in March and October near 40%. Supplementary Information (SI) Section 1 provides accessory illustrations of spatial and temporal well reporting frequency.

Bayesian Hierarchical Model Implementation. We implement a four stage BHM to impute missing groundwater data.¹⁶ The BHM is structured into the following levels:

1. Data Model: $[data \mid spatial \text{ process}, temporal \text{ process}, parameters]$
2. Spatial Process Model: $[spatial \text{ process} \mid parameters]$
3. Temporal Process Model: $[temporal \text{ process} \mid parameters]$
4. Parameter Model: $[parameters]$

The posterior distribution of the process and parameters updated with the data is:

$$[spatial \text{ process}, temporal \text{ process}, parameters \mid data] \sim [data \mid spatial \text{ process}, temporal \text{ process}, parameters] \times [spatial \text{ process} \mid parameters] \times [temporal \text{ process} \mid parameters] \times [parameters]$$

A full specification of the structure, parameters, and hyperparameters of the data model, process models, and parameter model is provided in SI Section 2. SI Section 2 also reports results validating the model assumptions for the study's large, sparse dataset and details of the sampling and imputation algorithm. Prior distributions of hyperparameters were selected to be vague at the scale of the data and conjugate when possible. The BHM iteratively updates parameter distributions for 11,000 steps. Internal validation of the BHM with out-of-sample data is available in SI Section 3, including an assessment of the model's ability to fit data with a parameter recovery experiment.

Multiple Posterior Sampling and Interpolation. The BHM generates posterior distributions of hyperparameters, which can be used to predict cells' missing groundwater levels multiple times across every time step. The BHM posterior distribution provides 10,000 predictions for each missing data point. For each time step, 100 sets of fully imputed datasets are created by randomly selecting from the 10,000 possible depth predictions for that location and time step.

We computed ordinary kriging interpolations of each of these sets of complete cell imputations on a 300 m grid overlaying the study area, producing 100 interpolated groundwater maps. Using this map manifold, we calculate each grid cell's mean, 95% prediction interval from 2.5% and 97.5% empirical quantiles, and variance. The variance includes the uncertainty due to both imputation of missing data and prediction at new locations. We also calculated depth to groundwater and variance at 150 m and 600 m resolutions.

External Comparison of Modeling Results. California's DWR produces bi-annual 'depth below ground surface' interpolated layers, produced in March and October to coincide with their heavily sampled collection periods.²⁹ The color ramp layer, hereafter termed the DWR layer, provides a smoothed approximation of the groundwater elevation "surface" based on interpolations of the measurement data.²⁹ These interpolated maps are used as a 'status quo' interpolated map for coinciding months of the BHM's interpolated maps. We use this product as a point of visual comparison for our model's final interpolated layer.

We first aggregated the DWR layer into the same 300 m grid as described in the previous section. We then take the difference between the BHM mean prediction and the DWR layer for each cell. We completed external comparison of modeling results for all 6 months that the DWR produced these layers in the 2013-2015 period.

RESULTS

Bayesian Hierarchical Model Convergence. We used Markov chain Monte Carlo (MCMC) for parameter estimation, such that the Markov chain's stationary distribution composes the parameters' posterior distribution.³⁰ Per iteration, our MCMC algorithm (i) drew most parameters by means of Gibbs steps from their full conditional distributions, (ii) drew parameters controlling the strength of spatial and temporal dependence by means of Metropolis-Hastings steps, and (iii) imputed missing groundwater depth values using parameters' updated

conditional distributions.³¹ Parameters converged within the first 1000 iterations, indicating that increased sparsity of our dataset did not inhibit parameter convergence despite longer computational times. From every numerically generated realization of the Markov chain beyond this burn-in period, we constructed posterior distributions of the parameters ($N = 10,000$). SI Section 3 includes the complete description of the MCMC algorithm and a Markov chain illustration.

The majority of the data model parameters' posterior distributions were normally distributed. One of the two seasonal parameters' distribution is centered around zero, which indicates that the majority of seasonal fluctuation in the groundwater readings were encompassed in a single parameter of the Fourier pair. The spatial dependence strength parameter (θ_s) was normally distributed with a mean value of 16.6, which indicates that cells beyond 56 km away from each other are statistically spatially uncorrelated. The temporal dependence strength parameter (α) was very high, demonstrating the utility of harnessing temporal short-term memory for groundwater depth predictions within each cell.

Monthly Multiple Posterior Sampling and Interpolation. After sampling from the posterior parameter distributions to generate multiple groundwater depth predictions at each 300 m grid cell across the study area, we calculated the mean, 95% prediction interval, and variance of each 300 m cell to generate descriptive map sets for each month. Figure 2a shows the mean predicted groundwater depth from 100 samples of fully imputed data based on the posterior distribution for October 2013. We observe a consistent spatial pattern that persists across all months. This demonstrates that the vector of spatial random effects (s) in the spatial process model effectively captured strong spatial patterns from depth to groundwater training data and propagated these spatial patterns when imputing missing data. The depth to groundwater mean maps across 150 m, 300 m, and 600 m resolutions were nearly identical, with predictions consistent within one meter across >98% of the study area (SI Section 4).

We quantify the prediction variance due to uncertainty associated with imputing missing data and uncertainty associated with spatial prediction and present the sum of six consecutive cell-by-cell variance maps (October 2013 through March 2014) in Figure 2B. Regions with high summed variance, such as the southeast corner of Merced county, are high value candidates for additional groundwater depth sampling. SI Section 4 includes cell-by-cell variance for each

month constituting Figure 2B, and provides an additional month (March 2014) of groundwater depth maps and external validation. Within-month variance distributions were particularly high for months with sparse groundwater depth sampling, such as January 2014. While monthly BHM predictions clearly offer the benefit of short-term memory in the temporal process model, it appears that well sampling within the prediction month remains critical to maintaining low prediction variance.

We observe that the variance of depth estimates increases with increasing resolution when comparing depth to groundwater variance maps across 150 m, 300 m, and 600 m resolutions. At coarser resolution, more wells are averaged together in-cell during the aggregation step pre-BHM and variance decreases. The spatial pattern of highest variance locations remains relatively constant across all three resolutions of analysis, demonstrating the BHM's consistent identification of data gaps for additional groundwater depth sampling efforts.

External Comparison of Modeling Results. We compare our modeling results with the biannual interpolated DWR layers produced under the CASGEM program. As seen in Figure 2c for October 2013, the DWR layer covers only the fraction of the study area for which uncertainty is acceptably low. Our prediction layer's bounds exceed that of the DWR layer because our model uses spatiotemporal information beyond within-month sampling. This approach is supported by strong spatial dependence persisting for at least 60 km across the area and very strong temporal dependence when verifying BHM assumptions (SI Section 2).

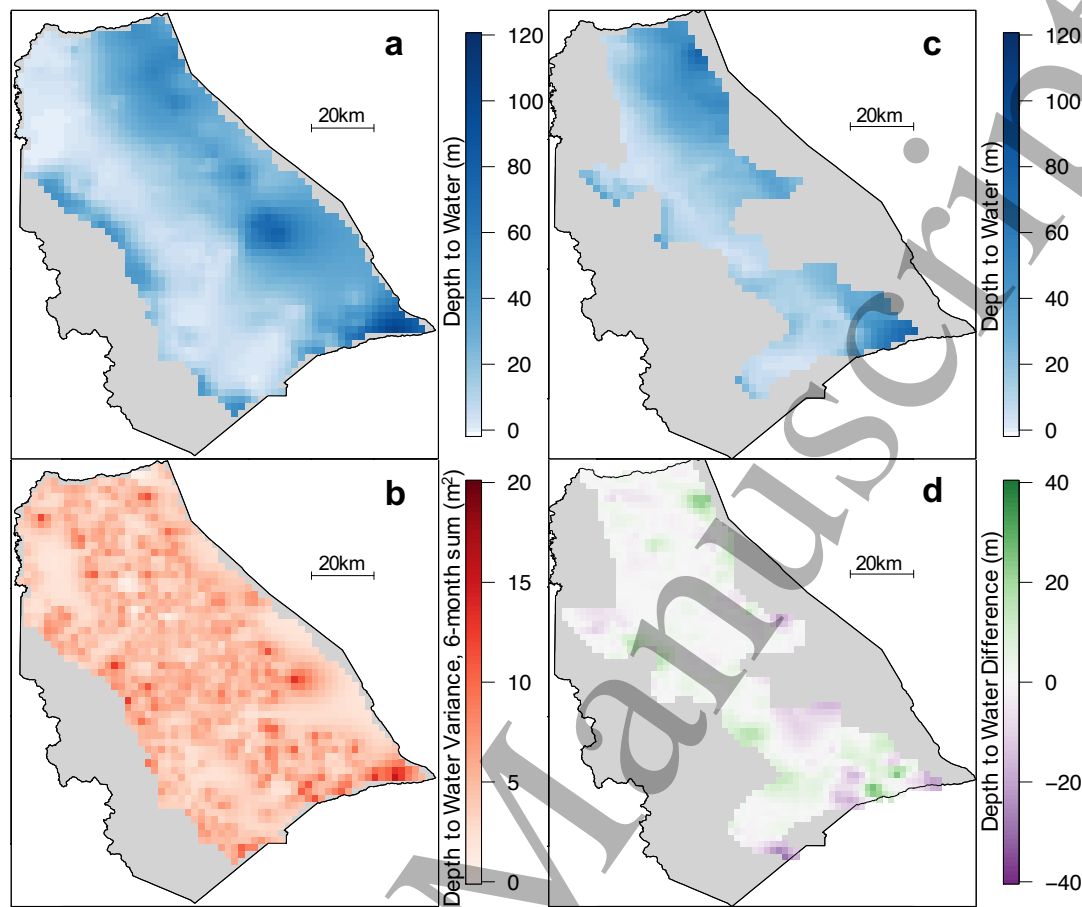


Figure 2. BHM extends spatial and temporal coverage of depth to groundwater estimates. (a) The October 2013 mean interpolated groundwater depth map after multiple posterior distribution imputation (N=100) from the BHM. (b) Sum of six months' depth-to-groundwater variance maps: October 2013 through March 2014. (c) DWR's interpolated groundwater depth map for October 2013 covers a smaller spatial area. (d) The difference between intersecting DWR and BHM mean depth, mapped in (c) and (a) respectively, shows strong agreement for the majority of the region; deviation on the edges is likely due to DWR's lack of data use beyond within-month observed spatial edge.

Figure 2d plots the per-pixel difference between the DWR layer and our model's mean prediction for the given month. We define exact match-up as when the difference between the two pixels is 0 m. For October 2013, 94.7% of 300 m-resolved differences were within 5 m of exact match-up. For March 2014, 94% of the cells were within 5 m. Thus, we can conclude that the interpolated product from DWR and our model's multiply interpolated mean product have strong consistency. We see the strongest disagreement between the two maps on the edges of the DWR layer. We hypothesize that the deviation between the models is primarily due to shortcomings in the local averaging method used to generate the DWR layers (details of DWR

layer generation available in SI Section 5). The DWR method does not account for well depth-to-groundwater readings beyond the map's extent or exploit readings beyond its subjectively specified time period. In contrast, the BHM model and subsequent ordinary kriging are performed using data that includes a data-driven selection of a 40 km buffer zone around the map area and used observed temporal patterns in the data to make predictions at cells lacking depth reporting in a month. While the DWR layer's resolution was originally 30 m and our modeling layer's resolution is 300 m, our model offers a clear improvement in terms of groundwater depth map temporal frequency. Our model also benefits from an ability to easily scale spatial resolution to 30 m, though higher resolutions require longer computational times.

DISCUSSION

We demonstrate the utility of BHM techniques for estimating depth-to-groundwater in the California CV. Previous efforts to apply BHMs to estimating GW depth only tested the validity of BHM models on high density datasets. Our model application and validation demonstrates parameter convergence even as computational times per Markov chain iteration increased. This model framework is generalizable to basin-by-basin analysis across all high- and medium-priority basins subject to the 2014 SGMA policy on a monthly basis, offering groundwater depth evaluation at a high spatial resolution in shorter monthly time updates.

This combination of model fidelity and transparency in prediction uncertainty make the Bayesian hierarchical modeling approach particularly useful for informing future sampling campaigns. Uniformly increasing spatial and temporal density of depth-to-groundwater across an entire study area can be very costly.¹⁰ However, one of CASGEM's goals is to identify and address data gaps in their groundwater monitoring program in accordance with Water Code section 10933(a).³² Uncertainty quantification using this BHM could aid in this assessment. For example, wells with highly-variable groundwater imputations would benefit the most from additional sampling wells. In addition, large variations in month-to-month predictions present high-value opportunities for increased frequency of data collection.

Depth-to-groundwater models are critical for informing policy in California, and the probabilistic and high resolution attributes of our BHM-enabled groundwater maps make them well-suited for this application. SGMA program performance evaluation may require incorporation of a variety of key metrics, including depth-to-groundwater. For instance, SGAs

1
2
3 308 may experiment with various groundwater restoration approaches, such as on-farm groundwater
4
5 309 recharge using water capture and storage services provided by natural landscape features.³³ Our
6
7 310 approach can be used to both forecast groundwater depth in a future month and update
8
9 311 groundwater depth predictions when public data for that future month becomes available. In this
10
11 312 example, comparison of the mapped groundwater depth forecast and prediction can be used to
12
13 313 assess the efficacy of basin recharge policies. The simplicity of uncertainty propagation from our
14
15 314 framework can be a valuable instrument for communicating confidence in such policy
16
17 315 evaluation.

18 316 This BHM framework enables sensitivity analysis of groundwater depth in further models or
19
20 317 analyses. For instance, depth-to-groundwater is a key uncertain variable in a variety of
21
22 318 agronomic analyses. Furthermore, groundwater flow models seeking to compare results to
23
24 319 empirical depth-to-groundwater data would benefit from uncertainty intervals around empirical
25
26 320 groundwater depth for internal validation. A variety of methods for propagating uncertainty
27
28 321 through a model or other analysis, such as stochastic simulation, would be well-suited to using
29
30 322 this depth-to-groundwater product that quantifies uncertainty.

31
32
33 323 **CONCLUSIONS**

34 324 This study employs a BHM coupled with multiple imputation and spatial prediction to
35
36 325 produce monthly groundwater depth maps in California's Central Valley at 300 m resolution.
37
38 326 This method allows for tractable uncertainty estimation of predicted groundwater depth across
39
40 327 policy-relevant space and time domains. As compared with existing approaches for monitoring
41
42 328 groundwater levels in California, our estimates offer quantitative estimates of model uncertainty
43
44 329 without sacrificing spatial and temporal resolution relevant to local policymakers and policy
45
46 330 enforcers.

47 331 Groundwater in California is inextricably tied to national food security,³ and SGMA has
48
49 332 recently created the first set of enforceable groundwater management guidelines for replenishing
50
51 333 overdrafted groundwater basins across the state. This work, when expanded to other high- and
52
53 334 medium- priority groundwater basins in California, can fill a critical gap for GSAs as they
54
55 335 complete their groundwater monitoring tasks. Our model is generalizable and constructed from
56
57 336 easily-collected data, offering recently-empowered local groundwater policy enforcers options
58
59 337 both as a planning tool and a policy evaluation tool.

SUPPORTING INFORMATION

Study area and temporal coverage illustrative figures (SI Section 1, Figures S1 and S2); details on the BHM's statistical underpinnings and associated Markov chain (SI Section 2, Figures S3 through S7); details of internal validation including a parameter recovery exercise (SI Section 3, Figures S8 through S10); BHM and DWR comparisons for March 2014 to complement main text's October 2013 maps (SI Section 4, Figure S11); month-by-month depth to water variance maps to complement main text's summed variance map (SI Section 4, Figure S12); multi-resolution depth to groundwater mean and variance maps (SI Section 4, Figure S13); details of the biannual interpolated mapping method employed by DWR (SI Section 5).

ACKNOWLEDGEMENT

This work was supported by the National Science Foundation (NSF) CBET 1604853 and the Pittsburgh chapter of the Achievement Rewards for College Scientists (ARCS) Foundation. The first author also gratefully acknowledges K. Manago for her insights pertaining to the groundwater BHM developed in Manago et al. (2019), and C. Doherty for administering computing services used for this research. All data used in modeling are available in open literature and are cited in this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available. The DWR periodic groundwater level dataset is available through the California Open Data Portal.³⁴

REFERENCES

1. California Agricultural Statistics Review: 2015-2016. (2016). California Department of Food and Agriculture.
2. Tackling the challenges of agricultural Water Use. In *OECD Agriculture*, OECD, 2016. Retrieved from <https://www.oecd.org/tad/policynotes/challenges-groundwater-use.pdf>
3. Gleick, P. H., Water strategies for the next administration. *Science* **2016**, *354*, (6312), 555.
4. Siebert, S.; Burke, J.; Faures, J. M.; Frenken, K.; Hoogeveen, J.; Döll, P.; Portmann, F. T., Groundwater use for irrigation – a global inventory. *Hydrol. Earth Syst. Sci.* **2010**, *14*, (10), 1863-1880.
5. Faunt, C. C.; Sneed, M.; Traum, J.; Brandt, J. T., Water availability and land subsidence in the Central Valley, California, USA. *Hydrogeology Journal* **2016**, *24*, (3), 675-684.
6. Famiglietti, J. S., The global groundwater crisis. *Nature Climate Change* **2014**, *4*, 945.
7. Faunt, C. C. e. *Groundwater Availability of the Central Valley Aquifer, California*; USGS Professional Paper 1766, 2009.
8. Xiao, M.; Koppa, A.; Mekonnen, Z.; Pagán, B. R.; Zhan, S.; Cao, Q.; Aierken, A.; Lee, H.; Lettenmaier, D. P., How much groundwater did California's Central Valley lose during the 2012–2016 drought? *Geophysical Research Letters* **2017**, *44*, (10), 4872-4879.
9. Chappelle, C.; Hanak, E.; Harter, T., Groundwater in California. In *Public Policy Institute of California*, 2017.
10. Newman, C.; Howitt, R. E.; MacEwan, D., How are Western water districts managing groundwater basins? *California Agriculture* **2018**, *72*, (1), 28-37.
11. Dickinson, AB-1739 Groundwater Management. In *CA GOV*, 2014.
12. Pavley, SB 1168 Groundwater Management. In *CA WAT*, 2014.
13. Pavley, SB 1319 Groundwater. In *CA WAT*, 2014.
14. Leahy, T. C., Desperate Times Call for Sensible Measures: The Making of the California Sustainable Groundwater Management Act The Waste of Water in 21st Century California. *Golden Gate University Environmental Law Journal* **2015**, *9*, 5-40.
15. California Statewide Groundwater Elevation Monitoring Program. In *CA Department of Water Resources*, 2018.
16. Manago, K. F.; Hogue, R. S.; Porter, A.; Hering, A. S., A Bayesian hierarchical model for multiple imputation of urban spatio-temporal groundwater levels. *Statistics & Probability Letters* **2019**, *144*, 44-51.
17. Sustainable Groundwater Management Act (SGMA) Program Guide. In *San Luis & Delta-Mendota Water Authority*, 2017.
18. Foglia, L.; Neumann, J.; Tolley, D. G.; Orloff, S.; Snyder, R. L.; Harter, T., Modeling guides groundwater management in a basin with river-aquifer interactions. *California Agriculture* **2018**, *72*, (1), 84-95.
19. Cressie, N.; Calder, C. A.; Clark, J. S.; Hoef, J. M. V.; Wikle, C. K., Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* **2009**, *19*, (3), 553-570.
20. Gronewold, A. D.; Borsuk, M. E., Improving Water Quality Assessments through a Hierarchical Bayesian Analysis of Variability. *Environmental Science & Technology* **2010**, *44*, (20), 7858-7864.

21. Malve, O.; Qian, S. S., Estimating Nutrients and Chlorophyll a Relationships in Finnish Lakes. *Environmental Science & Technology* **2006**, *40*, (24), 7848-7853.
22. Das, P.; Patskoski, J.; Sankarasubramanian, A., Modeling the Irrigation Withdrawals Over the Coterminous US Using a Hierarchical Modeling Approach. *Water Resources Research* **2018**, *54*, (6), 3769-3787.
23. Berliner, L. M.; Milliff, R. F.; Wikle, C. K., Bayesian hierarchical modeling of air-sea interaction. *Journal of Geophysical Research: Oceans* **2003**, *108*, (C4).
24. Obenour, D. R.; Gronewold, A. D.; Stow, C. A.; Scavia, D., Using a Bayesian hierarchical model to improve Lake Erie cyanobacteria bloom forecasts. *Water Resources Research* **2014**, *50*, (10), 7847-7860.
25. Dong, Z.; Hu, J., Development of Lead Source-specific Exposure Standards Based on Aggregate Exposure Assessment: Bayesian Inversion from Biomonitoring Information to Multipathway Exposure. *Environmental Science & Technology* **2012**, *46*, (2), 1144-1152.
26. Hopke, P. K.; Liu, C.; Rubin, D. B., Multiple Imputation for Multivariate Data with Missing and Below-Threshold Measurements: Time-Series Concentrations of Pollutants in the Arctic. *Biometrics* **2001**, *57*, (1), 22-33.
27. Goyal, A.; Small, M. J.; von Stackelberg, K.; Burmistrov, D.; Jones, N., Estimation of Fugitive Lead Emission Rates from Secondary Lead Facilities using Hierarchical Bayesian Models. *Environmental Science & Technology* **2005**, *39*, (13), 4929-4937.
28. Wainwright, H. M.; Flores Orozco, A.; Bucker, M.; Dafflon, B.; Chen, J.; Hubbard, S. S.; Williams, K. H., Hierarchical Bayesian method for mapping biogeochemical hot spots using induced polarization imaging. *Water Resources Research* **2016**, *52*, (1), 533-551.
29. Groundwater Information Center Interactive Map Application. In CA Department of Water Resources; **2018**.
30. Cauchemez, S.; Carrat, F.; Viboud, C.; Valleron, A. J.; Boëlle, P. Y., A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Statistics in Medicine* **2004**, *23*, (22), 3469-3487.
31. Gelfand, A. E.; Smith, A. F. M., Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association* **1990**, *85*, (410), 398-409.
32. *Status Report on Implementation of the California Statewide Groundwater Elevation Monitoring Program Years 2012-2015*; **2016**.
33. Kearns, F.; Parker, D., Supporting sustainable groundwater management. *California Agriculture* **2018**, *72*, (1), 6-7.
34. California Department of Water Resources, Periodic Groundwater Level Measurements. *California Water Data Library*, <https://data.cnra.ca.gov/dataset/periodic-groundwater-level-measurements>.