ABSTRACT: Megadrought in the western United States is jeopardizing water security. Groundwater regulations, such as California’s Sustainable Groundwater Management Act (SGMA), aim to preserve groundwater resources in overdrafted basins. Water agencies must establish sufficient monitoring systems to measure local groundwater abstraction and devise plans to moderate groundwater use. However, few technologies are available to monitor and regulate groundwater abstraction spatially and temporally. In this study, we deployed satellite-connected electrical current sensors on 11 agricultural groundwater pumps in Solano County, California over 2019–2022. A high correlation ($R^2 = 0.706$) was found between the in situ sensors and in-line flow meters. We then combine in situ sensor data with a land surface model to develop a multiple linear regression model of groundwater abstraction and groundwater level. Using a 10-fold cross-validation, it is found that our predictive groundwater abstraction model has approximately a 3.5% bias and a mean absolute error of 1.21 acre-feet, while our predictive groundwater level model has approximately 4.2% bias and about 5.9 acre-feet mean absolute error. Finally, we integrated these data with a blockchain-based groundwater credit trading platform to demonstrate how such a tool could be used for SGMA compliance.

KEYWORDS: groundwater, internet of things, SGMA, blockchain, drought

1. INTRODUCTION

1.1. Groundwater Use. An ongoing megadrought in the western United States is jeopardizing water security and causing drastically increased water stress for communities and agricultural producers, while incentivizing the implementation of new groundwater-focused policies. In the past decade, these unprecedented droughts have caused stresses on surface water availability and cost significant financial losses in crop and property damages. As surface waters are depleted and drought conditions continue, water managers rely more heavily on groundwater for domestic, agricultural, and industrial water.

Currently, groundwater supplies approximately 40% of the domestic and agricultural water use in California, and about 85% of the population in California depends on groundwater for some portion of their water supply. Groundwater depletion due to excessive pumping can cause land subsidence, diminished water quality, and reduced surface water supply. In times of extreme drought, sustainable groundwater management practices are critically important for the preservation of future water resources.

1.2. Groundwater Policy. Despite California’s leadership in environmental regulation and sustainability, it was the last state in the union to adopt a statewide system of groundwater regulation. Prior to the passage of the Sustainable Groundwater Management Act of 2014, regulation of California’s groundwater occurred piecemeal on a strictly local basis. This left most of the state’s groundwater resources unregulated, contributing to aquifer depletion and land subsidence. Although California’s groundwater rights doctrines are well-established, local governments have historically lacked the administrative means to prevent unsustainable practices. Prior to Sustainable Groundwater Management Act (SGMA), groundwater management was a strictly local effort (at the subbasin level or smaller) and often not a focal point of water resource management. As a result, groundwater pumping continued largely unabated for decades, with abstractions increasing during periods of drought.

A severe drought from 2011 to 2016 prompted the passage of the SGMA in 2014, which requires management of medium- and high-priority groundwater basins. SGMA authorizes pre
existing local agencies to voluntarily take on management of the resource by becoming a Groundwater Sustainability Agency (GSA) (Cal. Water Code §§ 10723, 10724). Local GSAs must develop a Groundwater Sustainability Plan (GSP) for their basin by early 2022 and achieve sustainability within 20 years. SGMA requires that GSAs determine the basin-wide rates of groundwater recharge and abstraction to equalize the water budget.

Although SGMA requires some core components in all GSPs, GSAs maintain discretion in deciding how to pursue sustainability (Cal. Water Code § 10727). GSAs are largely free to develop management strategies and practices best suited to local circumstances, so long as the GSAs provide a sufficient evidence for their choices and demonstrate measurable progress toward sustainability.

To achieve sustainable groundwater management, many areas will likely implement groundwater abstraction limits to stabilize overdraft basins.

1.3. Groundwater Monitoring and Modeling. To fulfill the requirements of SGMA, water agencies will need to actively monitor groundwater trends to establish management plans and demonstrate progress toward California’s policy requirements. Measurement of groundwater use is often considered a pre-requisite to effective management. However, by extension, such measurements are only effective if accurate, trusted, cost-effective, and actively incorporated into enforceable water conservation practices and policies.8

The United States Geological Survey (USGS) maintains monitoring wells for aquifer depth and health, but there are limited data sources for groundwater abstraction. In California, the California Statewide Groundwater Elevation Monitoring (CAGSME) Program was initiated in 2009 with the goal of establishing a locally managed network of systematic groundwater level measurements in all of California’s alluvial groundwater basins. Although this groundwater monitoring system is widespread in California’s medium- and high-priority basins, measurements often vary temporally, with some monitoring sites limited to a few groundwater level measurements over the span of decades. As a result, groundwater estimates are limited in their ability to project groundwater impacts and aid in groundwater management.9

Groundwater level models have been developed using linear regressions and machine learning techniques,10,11 and while these prediction models provide insights into trends in groundwater availability and aquifer health, they often do not distinguish between the effects of recharge and abstraction on the groundwater level.10,12−15 Although these groundwater level models provide vital information on aquifer health, they do not address spatial and temporal groundwater abstractions which tell us when and where groundwater is being used. Additionally, these models often use meteorological station data or remotely sensed satellite data as inputs but may be improved from the use of in situ groundwater abstraction data.

Implementation of electrical measurement devices to estimate groundwater abstraction has been studied extensively by the USGS, who used rate-time methods and power-consumption coefficients to calculate groundwater volume.16 This method has been implemented to estimate groundwater yield in various contexts. Chu et al.17 used electric power consumption records from pumping wells to estimate monthly groundwater yield over a 10 year period across the Choshui River Alluvial Fan in Taiwan and found that a spatio-temporal variability analysis of groundwater yield volumes is an effective method of identifying the patterns of pumped volumes in a region.

1.4. Groundwater Trading. Market-based approaches pose a viable option for tracking and trading groundwater among water users in a designated basin, as trading markets are widely accessible, scalable, and adaptive.18 Blockchain is among several technologies that may improve transparency and foster stronger market mechanisms in resource allocation.19 Resource monitoring and trading in smart communities may encourage the optimization of global consumption and save resources.20 As a pilot study, the Rosedale-Rio Bravo Water Storage District has co-developed the first open-source water accounting and trading platform in the Central Valley and found that a groundwater trading market is the most efficient and effective when developed as part of an integrated, web-based system. Blockchain-based trading platforms may foster greater transparency between practitioners, water users, and stakeholders, cultivating more effective management of water resources.18,21,22

2. METHODOLOGY

In this study, we investigate if improved groundwater management may be aided by in situ groundwater abstraction sensor networks whose data could populate a corresponding blockchain-supported groundwater trading platform.

We deployed in situ, electrical groundwater sensors in Solano County, CA (Section 2.1) to develop a data-driven model (Section 2.2) which uses geophysical inputs from both observations (Section 2.3) and simulated quantities (Section 2.4). We use 10-fold cross-validation and root mean square error, mean absolute error (MAE), coefficient of determination, and percent bias to determine the predictive skill of a multiple linear regression (MLR) groundwater abstraction model and groundwater level model. Finally, we integrate these data with a blockchain-supported groundwater trading platform, demonstrating capabilities to assist in SGMA compliance (Section 2.5).

2.1. Study Area. California has designated the Solano Subbasin, which is part of the Sacramento Valley Groundwater Basin, a medium-priority basin, and therefore is subject to SGMA compliance. Water availability in the Solano Subbasin is marked by the Sacramento-San Joaquin Delta and the Central Valley aquifer. The Solano Subbasin lies in the southwestern portion of the Sacramento Basin and the northern portion of the Sacramento-San Joaquin Delta.23

The elevation in the Solano Subbasin varies from 120 feet in the northwest corner to sea level in the south. The Solano Subbasin boundaries are defined by Putah Creek on the north, the Sacramento River on the east, the North Mokelumne River on the southeast, and the San Joaquin River on the south. The western Subbasin border lies on the hydrologic divide that separates lands draining to the San Francisco Bay from those draining to the Sacramento-San Joaquin River Delta. Annual precipitation averages in the basin range from approximately 23 inches in the western portion of the basin to 16 inches in the eastern portion of the basin.23 The historical precipitation and radiative temperature in Solano County over seven decades are shown in Figure 2.

Groundwater within the Solano Subbasin is considered to be of generally good quality and suitable for both domestic and agricultural purposes.24,25 Low permeability flood basin deposits, consisting primarily of silts and clays, occur along the eastern margin of the Subbasin. Along the Sacramento,
Figure 1. Map of California Statewide Groundwater Elevation Monitoring Program (CASGEM) monitoring wells of interest (marked with blue triangles) and electrical sensor sites on private agricultural wells (marked with red circles) in Solano County, California. All electrical sensor sites abstract groundwater from the shallow alluvial aquifer, and site locations are randomly displaced by 1 km to protect well owner privacy. Land use information sourced from the Division of Land Resource Protection, California Department of Conservation.

Figure 2. Historical precipitation and radiative temperature in Solano County from 1850 to 2022. In the latter half of the 21st century, mean radiative temperatures increase to approximately 30 degrees Celsius in this region. Data are sourced from phase 2 of the North American Land Data Assimilation System (NLDAS-2).
Mokelumne, and San Joaquin Rivers are deposits of unconsolidated silt, fine- to medium-grained sand, and gravel.26 Two primary aquifers underlay Solano County. The shallower of the two is referred to as the Alluvium aquifer and used primarily by private well owners, agricultural pumpers, and small community water systems. The deeper aquifer, known as the Tehama Formation, is the thickest water-bearing unit underlying the Solano Subbasin, ranging in thickness from 1500 to 2500 feet and provides most of the municipal water supply in the basin.27 In this study, sensors are installed on agricultural wells that abstract groundwater from the shallow alluvial aquifer. The locations of these sensor installation sites, as well as relevant groundwater well sites from the California Statewide Groundwater Elevation Monitoring Program (CASGEM), are shown in Figure 1.

Solano County is a major agricultural community that partially relies on groundwater from the Central Valley Aquifer for irrigation. Due to high connectivity between aquifer systems in Solano County and the Sacramento-San Joaquin Delta system, this area is under intense scrutiny because it supplies water to major cities downstream. Therefore, Solano County is an ideal study area to investigate the impacts of in situ groundwater abstraction monitoring under the context of regional and state-wide groundwater regulations.

2.2. Data-Driven Model and Performance Metrics. An MLR predictive model was selected to quantify linkages between hydrologic indicators and seasonal groundwater abstraction at each sensor site. This type of model is commensurate with operational28 and research surface-water forecast techniques,29,30 such as by the National Resources Conservation Council.30 The predictive technique works by fitting coefficients to individual predictor variables (e.g., streamflow, evapotranspiration, soil moisture, temperature, precipitation, estimated yield from in situ sensors, and groundwater level) with groundwater abstraction from in-line flow meters as the dependent variable. Predicted groundwater abstraction is verified with observed measurements from the in-line flow meters.

The MLR technique replicates the relationship between two or more explanatory variables and a dependent variable by fitting a linear equation to the observed data. The generic form of an MLR model is as follows

\[
y = \beta_0 + \beta_1 x_1 + \ldots + \beta_N x_N
\]

where \( y \) is the dependent variable, \( x_1, \ldots, x_N \) are observations of each of the independent variables, and \( \beta_1, \ldots, \beta_N \) are fitted coefficients.

The skill of each prediction model is determined by the coefficient of determination (\( R^2 \)), the percent bias (PBIAS), the normalized root-mean square error (RMSE), and the mean absolute error (MAE). These metrics were chosen to capture the explained variability within the model, the overall deviation from simulated and observed values, and the average magnitude of the errors in the modeled values. The percent bias is a measurement of the average tendency of the simulated values to differ from the observed value, while the root mean square error and MAE are metrics of the average model error, with the RMSE giving a relatively high weight to large errors.

The \( R^2 \) is calculated as

\[
R^2 = 1 - \frac{RSS}{TSS}
\]

where RSS is the sum of squares of residuals and TSS is the total sum of squares. The RMSE is calculated as

\[
RMSE = \left[ \frac{\sum_{i=1}^{N} (x_i - \bar{x}_i)^2}{N} \right]^{1/2}
\]

The MAE is calculated as

\[
MAE = \frac{\sum_{i=1}^{N} |x_i - \bar{x}_i|}{N}
\]

Also, the percent bias is calculated as

\[
PBIAS = \frac{(x_i - \bar{x}_i)}{x_i} \times 100
\]

where \( x_i \) is the observed time series, \( \bar{x}_i \) is the estimated time series, and \( N \) is the number of observations.

2.3. Observed and Remotely Sensed Data Sources. Four observed and remotely sensed data sources are used in this analysis: groundwater abstraction estimates are obtained from in situ electrical current sensors measuring groundwater pump activity (Figure 3); actual groundwater abstraction measurements are obtained from in-line flow meters; depth to groundwater level is obtained from the California Statewide Groundwater Elevation Monitoring Program (CASGEM); and all forcing data including model inputs such as temperature and precipitation are obtained from phase 2 of the National Land Data Assimilation System.

The use of electricity meters to estimate groundwater withdrawals has been proposed and demonstrated with varying degrees of effectiveness. In these applications, a power-conversion coefficient is determined as the ratio of water pumped to electricity consumed. These approaches are the most effective in stable aquifers.16,31

In this study, 11 electrical current sensors were installed between April 2019 and April 2020 on groundwater wells in Solano County, CA, and run time and electrical usage for each
pump were recorded daily. Subsequently, groundwater abstraction at each site is estimated using a power-conversion coefficient obtained from on-site pump tests.

At three of the sensor sites (wells 9, 10, and 11), an in-line flow meter was installed to quantitatively compare the two methods of groundwater abstraction monitoring: the electrical current sensor and conversion coefficient method with the more traditional flow meter approach. The in-line flow meters at wells 9, 10, and 11 are considered to be the “actual” groundwater abstraction throughout this analysis. Although in situ electrical current sensors provide many benefits, the utility and accuracy of this technology are dependent on a variety of factors including aquifer medium, pump age and type, and satellite connectivity. To establish the reliability of electrical current sensors in this region, the electrical current sensor data are directly compared with in-line flow meters. Weekly aggregated groundwater abstraction from the 11 sensor sites from April 2020 to April 2021 is used in this analysis. Since irrigation in Solano County is seasonal, the cumulative yearly estimated groundwater abstraction is also considered as a predictor in the MLR model. Although the end of the irrigation season depends on the type of crop being irrigated, here we define the end of a typical irrigation season as October 1. This allows us to examine trends between cumulative groundwater abstraction throughout the irrigation season and changes in depth to groundwater.

Mean depths from the ground surface to the water surface at 18 CASGEM monitoring well sites are examined from April 2019 to October 2020. Depth to the groundwater table is measured in reference to the North American Vertical Datum of 1988. These sites were selected based on the following criteria: (1) within the confines of Solano County, (2) no more than 15 miles from a sensor site, and (3) include 10 or more observational measurements of distance to groundwater elevation from April 2019 to January 2022.

Since the temporal resolution of observational measurements at each CASGEM site varies, each site was linearly interpolated to generate weekly values of depth to groundwater elevation. Linear interpolation was considered to be an appropriate and representative method of estimating weekly groundwater levels based on the high frequency of observational measurements at each site, with the frequency of groundwater level measurements at each site ranging from once every 2 weeks to once every 6 weeks.

To create a predictive model of groundwater abstraction, the in situ sensor sites are “paired” with existing CASGEM sites based on distance and correlation. The correlation coefficient of each CASGEM site is calculated against each sensor site, and the CASGEM with the highest correlation within 5 miles is chosen to be a predictor in the MLR model. Correlation between CASGEM and sensor sites was considered as a comparative metric to distinguish between wells that pump from different media (i.e., sand vs clay) or are used for different purposes (i.e., agriculture vs municipal).

Finally, temperature and precipitation are obtained from phase 2 of the National Land Data Assimilation System (NLDAS-2). NLDAS uses sophisticated numerical models of physical processes to integrate data from multiple ground- and space-based observing systems in order to produce fields of water and energy states and fluxes that are physically consistent and spatially and temporally continuous. This forcing data set is also used in land surface models to obtain other parameters of interest, as discussed in Section 2.4.

### 2.4. Modeled Data Sources

The variable infiltration capacity (VIC) land-surface model is used to obtain weekly values of surface runoff, baseflow, soil moisture, and evapotranspiration.

As a semi-distributed macroscale hydrological model, the VIC model operates by calculating both the water and surface energy budgets on a grid cell basis, while subgrid variations are captured statistically. VIC is driven by observations and therefore is expected to produce reliable estimates or hydrologic variables. Total evapotranspiration over the grid cell is computed as the sum of the canopy, vegetation, and bare soil components, weighted by the respective surface cover area fraction. To simulate streamflow, VIC results are typically post-processed with a separate routing model based on a linear transfer function to simulate the streamflow. Total column soil moisture is representative of the dynamic response of soil to the infiltrated rainfall, with diffusion allowed from the middle layer to the upper layer when the middle layer contains more water. The bottom soil layer receives moisture from the middle layer through gravity drainage, which is regulated by a Brooks-Corey relationship for the unsaturated hydraulic conductivity.

We use the 1/8° gridded NLDAS-2 forcing data set to determine weekly surface runoff, baseflow, soil moisture, and evapotranspiration. In this analysis, we define streamflow as the sum of runoff and baseflow. Hourly VIC data are aggregated to obtain weekly sums of fluxes (precipitation, streamflow, and evapotranspiration) and weekly means of state variables (temperature and soil moisture) in Solano County from April 2019 to June 2022. For clarity, these data sets are listed in Table 1.

### 2.5. Blockchain-Supported Groundwater Trading

To illustrate the potential application of an in situ, electrical
groundwater abstraction monitoring network, we created a prototype blockchain-based groundwater trading platform that integrates the in situ monitoring data. This platform illustrates the necessity of the in situ monitoring and groundwater modeling.

The experimental groundwater trading platform was created as a web-based portal designed to host groundwater abstraction and elevation data, as well as a variety of other potential data points. The platform has been designed to be compatible with records from virtually any telemetered groundwater monitoring device with little to no adaptation. This approach was pursued because a groundwater trading market benefits from being widely accessible and capable of integrating a variety of relevant information. A web-based strategy ensures that the trading system can host telemetered data from groundwater wells across the basin to track abstractions in near-real-time.

3. RESULTS AND DISCUSSION

3.1. Groundwater Abstraction Model Development. Over a 3 year period (April 2019–June 2022), there is an inverse correlation between the monthly aggregated groundwater abstraction data and the monthly aggregated groundwater level data, as shown in Figures 4 and 5. The trend between groundwater abstraction and groundwater level is described by the mean depth to groundwater level from the 18 CASGEM sites, the mean cumulative groundwater abstraction from the 11 in situ sensor sites was taken, and both data sets were normalized by the respective minimum and maximum values. To determine any changes in the relationship between the groundwater level and estimated groundwater abstraction in time, the correlation between the normalized data sets was investigated on a yearly basis. Figures 4 and 5 show the available data from 2019 to 2022. This analysis shows that the correlation between the normalized groundwater level and yearly cumulative groundwater abstraction varies between $R^2 = 0.377$ and $R^2 = 0.914$ for each year. This difference in $R^2$ value may be explained by natural recharge during wet years and surface water–groundwater interactions, indicating that the in situ, electrical groundwater sensor may provide the most useful information about groundwater levels during dry years.

Figure 4 shows a large decrease in depth to the groundwater table in October, which corresponds to the peak in cumulative groundwater abstraction at the end of the Solano County irrigation season. The aquifer replenishes from October through February and then decreases as pumping begins in March. This strong correlation between depth to the water table and observed groundwater abstraction suggests that the local groundwater level may be a highly effective and reliable indicator and identifier of local groundwater abstraction.

In this analysis, flow meter data at three sites (wells 9, 10, and 11) are assumed to be the actual groundwater abstraction. The right-side panel of Figure 6 shows a time series of the flow meter and estimated groundwater abstraction from the electrical sensors, while the left-side panel shows the associated correlations between the groundwater yield estimate from the sensor data and the corresponding coefficient obtained from on-site pump tests for each site, compared to the “ground truth” groundwater abstraction as measured by the in-line flow meters. The flow meter data in conjunction with the electrical...
sensor data at these three sites indicate that the sensor measurements and associated conversion coefficients provide a reasonable estimate of groundwater abstraction.

The timing and magnitude of groundwater abstraction are largely retained between the flow meter and electrical sensor measurements. Sites with relatively low flow (i.e., wells 9 and 10) show less variance and consistent underestimate of groundwater yield. Well 11 shows the greatest variance among the wells, with a large portion of groundwater abstraction estimates overestimated compared to the flow meter readings. Figure 6 provides further insights into this variability; for low flows (<1.5 acre-ft/day), the in situ sensors capture the magnitude and timing of groundwater abstraction, while at high flows (≥1.5 acre-ft/day), the magnitude of groundwater abstraction is underestimated by the in situ sensors. This is potentially explained by the use of variable frequency drive pump controllers, which provide efficient energy usage at higher yields. Since well 11 typically experiences flows greater than 1.5 acre-ft/day, the overall accuracy of estimated groundwater abstraction at this site is diminished. Additionally, Figure 6 shows that in situ sensors tend to underestimate groundwater abstraction at well 10, while tending to overestimate groundwater abstraction at well 11. This could be explained by a variety of factors, including variations in pump performance, differences in the underlying aquifer medium, satellite transmission failures, inaccuracies in flow meter calibration, or inadequate power supply to the sensor due to debris blocking the solar panel.

For each groundwater abstraction site (wells 9−11), an MLR is created with environmental parameters as predictor variables (precipitation, streamflow, evapotranspiration, soil moisture, temperature, depth to groundwater table, and groundwater abstraction from in situ sensors) and groundwater yield from in-line flow meters as the dependent variable. A neighboring CASGEM site is selected as the depth to the groundwater level predictor variable based on its proximity and correlation to sensor sites. Trends between individual CASGEM and sensor sites may provide insights into the timing and location of groundwater abstraction and consider how differences in aquifer medium (i.e., sand vs loam), differences in borehole screened interval depths, and differences in borehole use (i.e., agriculture vs domestic) may be used to select a more accurate groundwater abstraction predictor variable. Additionally, the hydrologic variables at each site are calculated based on the weighted distances from the location of the well to the center of the grid cell. In this way, each well site has unique values of precipitation, evapotranspiration, streamflow, soil moisture, and temperature to more accurately capture local variations in climate.

To determine which CASGEM site will be used as a predictor in the MLR for an individual in situ site, the correlation coefficient and distance between a single in situ sensor site and all CASGEM sites are considered. The most highly correlated CASGEM site within 5 miles is selected as a predictor in the MLR. Figure 7 shows a matrix of the correlation coefficient between all sensor sites and CASGEM sites. The correlation coefficient is calculated as...
where $r$ is the correlation coefficient, $\bar{x}$ is the mean of $x$, and $\bar{y}$ is the mean of $y$. The heat map shows that in situ sensor sites have a high positive correlation with other sensor sites and have varying magnitudes of inverse correlation with CASGEM sites. This further supports the strong inverse correlation between groundwater abstraction and groundwater table elevation, as shown in Figures 4 and 5. The correlation coefficient heat map also provides insights into individual sites’ behavior; the sites are listed from the most negative to the most positive correlations and indicate that in situ sensor site 7 has the most highly positive correlations, while CASGEM site 10 has the most highly inverse correlations.

Figure 6. Left panel: actual weekly groundwater abstraction (flow meter data) compared to estimated weekly groundwater abstraction from electrical sensor data. Right panel: time series of flow meter data compared to the estimated groundwater abstraction from in situ sensor data. An on-site pump test determines the conversion coefficient for each well separately, which is used to convert the electrical current sensor readings into yield (acre-ft).
Figure 7. Matrix of correlation coefficients between each in situ sensor site (denoted with a circle) and nearby CASGEM sites (denoted with a triangle). Blue colors represent inverse correlation coefficients and red colors represent positive values, with darker colors representing greater magnitude.

Figure 8. Modeled weekly groundwater abstraction using all predictors. When all predictors are included in the linear model, the abstraction at wells 9, 10, and 11 is captured compared to the actual groundwater abstraction, with an $R^2$ value of 0.921.
increased by 30%; and when temperature was not included as a predictor, the NRMSE increased by 35% and the MAE increased by 24%. The analysis of each iteration of predictor selection shows that the skill of the model is least affected by the streamflow and soil moisture predictors, as shown in Figure 9.

To demonstrate the predictive capability of the groundwater abstraction model, a 10-fold cross-validation was completed using the original model with all predictors. For this analysis,
the original sample is partitioned into a training set to develop the model and a test set to evaluate it. In 10-fold cross-validation, the original sample is randomly partitioned into 10 equal size subsamples. Of the 10 subsamples, a single subsample is retained as the validation data for testing the model and the remaining 9 subsamples are used as training data. The cross-validation process is then repeated 10 times (the folds), with each of the 10 subsamples used exactly once as the validation data. The 10 skill scores from the folds are then averaged to produce a single estimation of model performance. The results of this procedure found that the average percent bias of predicted groundwater abstraction is 3.50%, the average MAE is 1.21 acre-ft, and the normalized RMSE is 1.43. This approach supports the concept that a groundwater abstraction model can be developed and applied in a specific region to estimate the timing and quantity of groundwater abstraction.

3.2. Groundwater Level Model Development. A groundwater level model was created using measured depth from the land surface to the water table from CASGEM as the dependent variable and weekly precipitation, evapotranspiration, streamflow, temperature, soil moisture, and estimated groundwater abstraction from in situ sensors as the predictor variables. The high correlation between groundwater levels (CASGEM) and the estimated groundwater abstraction from in situ electrical current sensors, shown in Figures 4 and 5, suggests that estimated groundwater abstraction may be a useful predictor in modeling and predicting groundwater levels. To determine the relative importance of each predictor variable, seven iterations of an MLR were developed: (1) all predictions, (2) drop in situ sensor as a predictor, (3) drop precipitation, (4) drop evapotranspiration, (5) drop streamflow, (6) drop temperature, and (7) drop soil moisture. The relative $R^2$, NRMSE, and MAE were compared for each model iteration. Figure 10 shows the resulting modeled groundwater level at the paired CASGEM sites for wells 1–11. The model has an $R^2$ of 0.953, indicating that the majority of the variation in the actual groundwater levels is explained by the predictors in the model.

As before, a predictor drop analysis is completed to determine the relative skill score of each variable. Here, the original model is defined as the model iteration that includes all of the predictor variables (streamflow, evapotranspiration, soil moisture, temperature, precipitation, and estimated yield from in situ sensors). The model performance decreases when the groundwater abstraction estimate from in situ sensors, the evapotranspiration, and the temperature are not included as predictors in the MLR. When the in situ sensor data are not included as a predictor, the $R^2$ experiences a $-11.5\%$ change from the original model, and when temperature is not included as a predictor, the $R^2$ experiences a $-7.47\%$ change from the original model. While all coefficients of determination suggest that much of the variance in the model is explained, the relative decrease in $R^2$ when the in situ sensor data or the temperature data are not included as a predictor indicates that these variables are significant to the model skill. Likewise, when the in situ sensor data are not included as a predictor, the NRMSE increases by 98% and the MAE changes by 104% compared to the original model, and when temperature is not included as a predictor, the NRMSE changes from the original model by 61% and the MAE changes by 46%. For this regression model, the soil moisture, precipitation, and streamflow data have the least effect on the model skill score, as shown in Figure 11.

The predictive capability of this groundwater level model is computed from a 10-fold cross-validation. This cross-validation of the groundwater level model found that the average percent bias of predicted groundwater abstraction is 4.21%, the average MAE is 5.93 acre-ft, and the NRMSE is 1.04. This demonstrates that existing, linear regression-based groundwater level models may be improved with the use of in situ groundwater abstraction data. Additionally, the high predictive skill of a data-driven predictive model demonstrates that such a model may provide useful information to water managers and water users. Such models may be implemented in areas without existing groundwater level data to predict aquifer levels.
levels, which may reduce the lack of groundwater data in drought-prone areas.

3.3. Groundwater Trading Platform Demonstration. The experimental blockchain-supported groundwater trading platform integrates pumping and depth data from over 11 in situ monitoring devices in the Solano Subbasin. The platform also hosts data from several different types of groundwater monitoring devices. This variety was chosen as it reflects real-world conditions, ensuring that the experimental platform will function with nearly any existing monitoring device so long as it produces data that are digitized and telemetered.

Once a monitoring device is integrated into the experimental platform, the well associated with that device is assigned to the individual owner. After being authenticated, the owner can access and view their abstraction information in near-real-time. Importantly, the experimental platform’s ability to connect individual owners with specific wells provides a practical means of assigning groundwater allocations to individuals. GSAs are responsible for the creation of abstraction allocations and may opt to develop and assign them based on individual wells, landowners, irrigated acreage, or properties. In any case, a functional groundwater trading system needs to be able to reconcile individual allocations with an individual owner’s wells and properties and the associated in situ monitoring data. This reconciliation provides the owner with an accurate understanding of their allowable groundwater pumping while also enabling instantaneous tracking of individual usage against the associated allocation. Thus, pairing the monitoring data, allocations, and ownership provides the base understanding and insights necessary for any transactional system.

The experimental platform integrates real groundwater monitoring data from wells across the Solano Subbasin, assigns those wells to users, and gives each well fictional allocations (actual allocations do not yet exist). Using this approach, the experimental platform successfully hosted a ledger of allocations and individually reconciled them with actual, ongoing groundwater abstraction monitoring. This provided nearly instantaneous insights into the remaining allocation and notice when an allocation has been exceeded and by how much.

To ensure the security and privacy of the information, the experimental platform has been built using a managed blockchain. Blockchain is valuable for a wide range of applications where security, privacy, and traceability are critical. The use of blockchain can foster trust among prospective users and stakeholders by safeguarding individual allocation data. 

Figure 12. Experimental digital portal includes a marketplace component where individual groundwater users can buy and sell allocations and receive incentives for groundwater augmentation activities.
privacy, creating an immutable ledger that minimizes corruption and prevents the data from being compromised. The immutable nature of the blockchain ledger ensures that groundwater usage and allocations are accurately and unalterably tracked, avoiding even the perception of impropriety in the monitoring of groundwater usage. Although still a prototype, the experimental platform represents a valuable tool for monitoring groundwater abstractions and tracking the status of groundwater across the basin. Once fully developed, the platform could provide a means to give individual groundwater users the flexibility to buy and sell unused allocations, potentially realizing a new revenue stream while ensuring that overall groundwater use is reduced to sustainable yields. The dashboard display of this trading feature is demonstrated in Figure 12. Marketplace functions will help to direct limited water supplies to the highest and best use. Moreover, this trading system will provide the GSAs with additional management options by allowing the agencies to impose rules on the trading market that drive toward desired outcomes.

Beyond advancing planning efforts and providing flexibility to landowners, the experimental platform also incentivizes groundwater augmentation activities. The platform currently includes a system whereby individual users could request credit for implementing activities that boost groundwater recharge. Once independently verified and approved, the landowner would then receive credit in the form of additional allocations that they could use or sell in the marketplace. This recharge component is supported by the groundwater elevation modeling, which can help to direct the recharge activities to locales where they will have the most benefit and evidence the results of such actions. Even in the absence of a groundwater trading system, the ability to document recharge enables a GSA to establish an incentive system, offering a path to sustainability that focuses on supply augmentation.

4. CONCLUSIONS
The lack of reliable, widespread groundwater abstraction data is a significant impediment for achieving long-term groundwater sustainability in drought-prone areas. In this study, we examined how the in situ groundwater monitoring networks can be used to inform a statistically driven MLR model to predict groundwater abstraction and groundwater levels and enable a blockchain-supported groundwater management and trading platform in Solano County, CA.

The cost of groundwater monitoring, including the use of flow meters, water level sensors, and telemetry, can be a significant barrier to improved management. In this study, we used satellite-connected sensors that offer competitive installation costs while demonstrating reliable and usable data collection. A full analysis of costs and benefits across a variety of solutions will likely emerge over the coming years of SGMA compliance planning.

Deploying noninvasive technologies such as these to record data from the existing groundwater infrastructure could remedy the current data gaps. Moreover, the required data management system could be developed as a web-based digital platform, making these data more accessible and, as a result, more useful. Importantly, the confluence of the required groundwater monitoring network, data management system, and elevation modeling potentially creates an opportunity for GSAs to take the first step toward a market-based groundwater trading program by encouraging them to create a centralized and accessible platform that tracks groundwater use nearly instantaneously and models the implications of that use. Once established, this type of platform could be refined and expanded to facilitate groundwater trading and ensure sustainable groundwater use.

■ AUTHOR INFORMATION
Corresponding Author
Evan Thomas — Mortenson Center in Global Engineering and Resilience, University of Colorado Boulder, Boulder, Colorado 80302, United States; orcid.org/0000-0003-3095-8407; Email: ethomas@colorado.edu

Authors
Melanie Holland — Mortenson Center in Global Engineering and Resilience, University of Colorado Boulder, Boulder, Colorado 80302, United States
Chris Thomas — The Freshwater Trust, Portland, Oregon 97205, United States
Ben Livneh — Mortenson Center in Global Engineering and Resilience, University of Colorado Boulder, Boulder, Colorado 80302, United States; orcid.org/0000-0001-5445-2473
Stephanie Tatge — The Freshwater Trust, Portland, Oregon 97205, United States
Alex Johnson — The Freshwater Trust, Portland, Oregon 97205, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acsestwater.2c00214

Notes
The authors declare the following competing financial interest(s): The supplier of the sensors used in this study, SweetSense Inc., was founded and led by corresponding author Thomas.

■ ACKNOWLEDGMENTS
This work was supported by The Freshwater Trust, The Gordon and Betty Moore Foundation, the Water Foundation, and the National Science Foundation under the terms of Award no. 1738321. The authors thank Skot Croshere, Danny Wilson, Taylor Sharpe, Matthew Tolbirt, Matthew Falcone, Becky Rittenburg, David Primozich, and Erik Ringelberg and the landowners in Solano County, California who generously allowed the monitoring of their groundwater pumps.

■ REFERENCES
(1) Williams, A.; Cook, E. R.; Smerdon, J. E.; Cook, I. J. T. A.; Abatzoglou, K.; Bolles, S. H.; Baek, A. M.; Badger, B.; Livneh, B. Large contribution from anthropogenic warming to an emerging North American megadrought. Science 2020, 368, 314.
(2) Wilhite, D. A.; Svoboda, M. D.; Hayes, M. J. Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness. Water Resour. Manag. 2007, 21, 763–774.
(3) Cook, B. I.; Ault, T. R.; Smerdon, J. E. Unprecedented 21st century drought risk in the American Southwest and Central Plains. Sci. Adv. 2015, 1, No. e1400082.
(4) Chappelle, C.; Hanak, E.; Harter, T. Groundwater in California, 2017. URL. https://www.ppic.org/publication/groundwater-in-california/.
(5) Faunt, C.; Sneed, M.; Traum, J.; Brandt, J. T. Water availability and land subsidence in the Central Valley, California, USA. Hydrogeol. J. 2016, 24, 675–684.
(6) Tina Cannon Leahy. Desperate Times Call for Sensible Measures: The Making of the California Sustainable Groundwater Management Act. Environmental Law Journal, 2016; Vol. 9. URL: http://digitalcommons.law.ggu.edu/ggueljhtmlp://digitalcommons.law.ggu.edu/ggueljvol9iss1/4.

(7) Water Education Foundation. The 2014 Sustainable Groundwater Management Act: A Handbook to Understanding and Implementing the Law, 2015. URL: https://www.watereducation.org/publication/2014-sustainable-groundwater-management-act.

(8) Molle, F.; Closs, A. Groundwater metering: revisiting a ubiquitous “best practice”. Hydrogeol. J. 2021, 29, 1857.

(9) NIDIS. Technical report National Integrated Drought Information System (NIDIS) in Partnership with the California-Nevada Climate Applications Program (CNAP) and Western Regional Climate Center (WRCC), 2017.

(10) Hodgson, F. D. I. The Use of Multiple Linear Regression in Simulating Ground-Water Level Responses. Groundwater 1978, 16, 249–253.

(11) Daliakopoulos, I. N.; Coulibaly, P.; Tsanis, I. K. Groundwater level forecasting using artificial neural networks. J. Hydrol. 2005, 309, 229–240.

(12) Jasmin, J.; Murali, T.; Mallikarjuna, P.; Ismail, J.; Murali, T.; Mallikarjuna, P. Statistical Analysis of Groundwater Table Depths in Upper Swarnamukhi River Basin. J. Water Resour. Protect. 2010, 02, 577–584.

(13) Sahoo, S.; Jha, M. K. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. Hydrogeol. J. 2013, 21, 1865–1887.

(14) Maiti, S.; Tiwari, R. K. A comparative study of artificial neural networks, Bayesian neural networks and adaptive neuro-fuzzy inference system in groundwater level prediction. Environ. Earth Sci. 2014, 71, 3147–3160.

(15) Adamowski, A.; Chan, H. F. A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. 2011, 407, 28–40.

(16) Hurri, R. T.; Litke, D. W. Estimating Pumping Time and Ground-Water Withdrawals Using Energy-Consumption Data Technical Report, 1989. URL: http://pubs.er.usgs.gov/publication/wri894107.

(17) Chu, H.-J.; Lin, C.-W.; Burbey, T. J.; Ali, M. Z. Spatiotemporal Analysis of Extracted Groundwater Volumes Estimated from Electricity Consumption. Groundwater 2020, 58, 962–972.

(18) Heard, S.; Remson, E. J.; Fienup, M.; King, S. S. Groundwater Markets: An Early Case Study from Fox Canyon; The Nature Conservancy, 2019. URL: https://groundwaterresourcehub.org/public/uploads/pdfs/TNC_FoxCanyon_GroundwaterMarketCaseStudy.pdf.

(19) Chohan, W. Blockchain and Environmental Sustainability: Case of IBM’s Blockchain Water Management; SSRN Electronic Journal, 2019; pp 1556–5068.

(20) Alcarria, R.; Bordel, B.; Robles, T.; Martin, D.; Manso-Callejo, M.-A. A Blockchain-Based Authorization System for Trustworthy Resource Monitoring and Trading in Smart Communities. Sensors 2018, 18, 3561.

(21) Babitt, C.; Rosedale—Rio Bravo Water Accounting and Trading Platform: A Pilot Project Advancing Sustainable Groundwater Management, 2018. URL: https://edimaps.maps.arcgis.com/apps/Cascade/index.html?appid=7a13444157741a2a090cf4262b4206.

(22) Johnson, A.; Ringelberg, E.; Rittenburg, B.; Dumont, D. Groundwater Measurement & Trading: California Pilots Groundwater Measurement and Trading Program with IoT Sensors and Blockchain, The Water Report, 2019. URL: www.TheWaterReport.com.

(23) Thomasson, H. G., Jr.; Olmsted, P. H.; LeRoux, E. F. Geology, Water Resources and Useable Ground-Water Storage Capacity of Part of Solano County, California; United States Department of the Interior, Geological Survey: Washington, D.C., 1960.

(24) Marguerite, A. S. Reports for California by the U.S. Geological Survey, Water Resources Division. Technical Report, 1989.