Title

Optimizing Managed Aquifer Recharge Locations in California's Central Valley Using an Evolutionary Multi-Objective Genetic Algorithm Coupled With a Hydrological Simulation Model

Journal

Water Resources Research, 59(5)

ISSN

0043-1397

Authors

Kourakos, Georgios

Brunetti, Giuseppe

Bigelow, Daniel P

et al.

Publication Date

2023-05-01

DOI

10.1029/2022wr034129

Optimizing Managed Aquifer Recharge Locations in California's Central Valley using an Evolutionary Multi-Objective Genetic Algorithm coupled with a Hydrological Simulation Model

Georgios Kourakos[a], Giuseppe Brunetti[b], Daniel P. Bigelow[c], Steven Wallander[d], Helen E Dahlke[a]

a) University of California, Davis; Department of Land, Air, and Water Resources One Shields Ave., Davis, CA 95616-8628, United States

b) University of Natural Resources and Life Sciences Vienna, Austria

c) Oregon State University, Department of Applied Economics, Corvallis, OR, United States

d) Economic Research Service, United States Department of Agriculture, Washington D.C., USA

Key Points:

- Combining a simulation-optimization framework with a groundwater model can identify suitable managed aquifer recharge locations
- Recharging water from local rivers within the same basin is the most cost-efficient MAR approach per unit of groundwater storage gain
- Diverting more water may not yield the most efficient groundwater storage gain due to higher transportation cost

Abstract

Managed aquifer recharge (MAR) can provide long-term storage of excess surface water for later use. While decades of research have focused on the physical processes of MAR and identifying suitable MAR locations, very little research has been done on how to consider competing factors

and tradeoffs in siting MAR facilities. This study proposes the use of a simulation-optimization (SO) framework to map out a cost-effectiveness frontier for MAR by combining an evolutionary algorithm with two objective functions that seek to maximize groundwater storage gains while minimizing MAR cost. We present the theoretical framework along with a real-world application to California's Central Valley. The result of the SO framework is a Pareto front that allows identifying suitable MAR locations for different levels of groundwater storage gain and associated MAR project costs, so stakeholders can evaluate different choices based on cost, benefits, and tradeoffs of MAR sites. Application of the SO framework to the Central Valley shows groundwater can be recharged from high-magnitude (95^{th} percentile) flows at a marginal cost of $57 to $110 million per km^3. If the 10 percent largest flows are recharged the total groundwater storage gain would double and the marginal costs would drop to between $30 and $50 million per km^3. If recharge water is sourced from outside local basins (e.g. the Sacramento-San Joaquin Delta), groundwater storage gain is approximately 25% to 80% greater than can be achieved by recharging local flows, but the total cost is about 10-15% higher because of additional lift cost.

Keywords: managed aquifer recharge, multi-objective optimization, hydro-economic modeling, California

1 Introduction

About 25% of the global population depends on groundwater for drinking water supplies (Grönwall & Danert 2020). Groundwater also provides about 40% of the irrigation water that supports global food production (Siebert et al., 2010; Aeschbach-Hertig and Gleeson 2012; Döll et al., 2012). Our ability to continue to rely on groundwater sources for future needs, however, is threatened by growing consumptive use and climate change. Many regions that are heavily dependent on groundwater have experienced severe groundwater storage depletion (Wada et al., 2010; Bierkens and Wada 2019; Jia et al., 2019; Ajaz et al., 2020; Dangar et al 2021). To address these water scarcity concerns, a number of groundwater management actions have been adopted

around the globe, such as the Water Framework Directive (WFD, 2000/60/EC) and the Groundwater Directive (GWD, 2006/118/EC) in the European Union, the National Groundwater Action Plan in Australia, and other national legislation in China (Shen 2015; Yu et al., 2018; Wang et al., 2020), India (Jha and Sinha 2010), and South Africa (Pietersen et al., 2012). Many of these policy frameworks have the goal of achieving sustainable use of groundwater resources by protecting aquifers from deterioration and chemical pollution (Scanlon et al., 2012).

Managed Aquifer Recharge (MAR) is increasingly being recognized as a promising management option (Dillon et al., 2020). MAR is defined as the intentional replenishment of an aquifer with water that can be used at a later time (e.g., during a drought) or in different places after water transport. In MAR the aquifer is treated as a water bank (Maliva, 2015), where water can be stored during periods when supply exceeds demand. MAR is widely recognized as one of the cheapest forms of water storage (Dillon 2005) and has been studied for several decades. Dillon et al. (2019) and Zhang et al. (2020) provide a thorough review of MAR practices and challenges observed over the last six decades. Depending on local needs and conditions, different MAR technologies are used to replenish groundwater (Dillon 2005; Gale 2005; Tzoraki et al., 2018; Standen and Monteiro 2020).

Increasing MAR worldwide is an important policy challenge and often starts with identifying suitable locations for MAR facilities. A recent review of MAR site suitability studies (Sallwey et al., 2019) observed that the majority of studies identify suitable sites by combining multiple biophysical factors such as hydrogeology, geomorphology or soil suitability (Russo et al., 2014; O'Geen et al., 2015), land use (Marwaha et al. 2021), but also groundwater quality (Waterhouse et al, 2020), water availability (Kocis and Dahlke, 2017; Dahlke & Kocis, 2018), and economic feasibility (Tran et al., 2019, 2020a,b) in a multi-criteria decision analysis (MCDA). Marwaha et al. (2021), for example, identified MAR sites near rural communities in the southern Central Valley, California by combining soil maps, land use maps, existing conveyance infrastructure information, and general groundwater flow directions in a MCDA. Likewise, Zaidi et al., (2015) identified suitable MAR locations in Saudi Arabia by overlaying maps of soil texture, vadose zone thickness, land use, and slope. A common alternative approach to these MCDA analyses conducted in Geographic Information Systems is the identification of MAR locations and their

benefits and risks through numerical modeling (Tzoraki et al., 2018; Kacimov et al., 2016; Zlotnik et al., 2017; Wurl and Imaz-Lamadrid, 2018; Ganot and Dahlke, 2021; Bachtouli and Comte, 2019). Scherberg et al. (2014), for example, used the Integrated Water Flow Model (IWFM) (Dogrul, 2013) to quantify the impact of selected MAR scenarios in the Walla Walla River basin, USA. Similarly, Niswonger et al., (2017) used a numerical model (MODFLOW) to examine the impact of six agricultural MAR scenarios on groundwater storage, groundwater stream interaction, and groundwater levels. Ghasemizade et al., (2019) and Kourakos et al., (2019) also used IWFM to study the effect of flooding different farm field MAR sites with excess surface water (Ag-MAR) in a sub-basin of the Central Valley (CV).

Many of the above modeling studies investigated a fixed and generally limited but pre-determined number of MAR locations and management scenarios within relatively small study areas, which provides the advantage that the number of decision variables and computational time are relatively small. Only a few studies to date have used a simulation–optimization (SO) framework together with a groundwater model to determine optimal MAR locations based on different sets of model parameters that are dependent on several decision variables in the optimization problem). Most of these studies have focused on groundwater remediation (Zheng and Wang 1999; Cunha 2002; Yeh 2015), groundwater abstractions (Datta and Kourakos 2015; Danapour et al., 2021), and coastal aquifer management (Kourakos and Mantoglou, 2015) and only a few have been applied to MAR. The few studies that have used SO frameworks for MAR have solely focused on aquifer storage and recovery (Hernandez et al., 2014; Marchi et al., 2016; Forghani and Peralta 2018; Al-Maktoumi et al., 2020). For example, Fatkhutdinov and Stefan (2019) used a multi-objective SO to identify optimal well locations and recharge rates in a hypothetical test case. Ebrahim et al. (2016) developed a MODFLOW groundwater model for a catchment in Oman and combined it with an optimization algorithm to increase aquifer recharge by optimizing extraction and recovery rates subject to water level constraints.

The above referenced studies have primarily looked at physical factors that influence MAR locations. In addition, in MAR design the economic cost is a significant parameter to consider. Marques et al., 2010 optimized the water availability and the economic benefits of irrigated agriculture, yet their approach did not consider the simulation of groundwater storage. On the

other hand, Gailey et al., (2019a), optimized reservoir releases to maximize the available volume subject to available funds by coupling the optimization with a groundwater simulation model. Because of their small study area, they were able to consider all available land for MAR siting. In this study we propose a simulation–optimization framework to identify suitable MAR locations that maximize groundwater storage benefits while minimizing MAR cost. Our method is suitable for large scale studies where considering a full list of all available land is prohibitive. Developing and testing a simulation-optimization framework to site MAR locations improves our understanding on how to deploy MAR as a water resources management strategy while considering multiple competing environmental and economic benefits and constraints. The balance between groundwater benefits and economic cost are represented by a Pareto frontier, which allows decision makers to evaluate multiple water source, recharge volume, and recharge location trade-offs. The framework was tested in the San Joaquin–Tulare Lake Basin in California, one of the most productive agricultural areas in the world.

2 Methodology

In this study, we propose a simulation-optimization (SO) framework to identify optimal recharge facility locations for MAR that maximize groundwater storage recovery. Generally, the groundwater storage is likely to increase as the number of MAR sites increase. However, implementing MAR facilities is costly, hence one can expect the total MAR cost to increase as the number of MAR sites increases. To balance the environmental benefits of MAR with the cost of each MAR facility we use a multi-objective optimization framework with two competing objectives: i) maximization of groundwater storage and ii) minimization of economic cost. This formulation combines an output frontier (for a single output) and a cost frontier, such that any point on the frontier is optimal in the sense of being cost-effective (i.e.: being the least cost option for achieving a given quantity of water recharged) (Førsund et al., 1980). Mathematically the optimization formulation is expressed as:

$$\left.\begin{aligned} \max_{D} O_{env}(D) &= f(S^G) \\ \min_{D} O_{cost}(D) &= g(C) \end{aligned}\right\} \quad (1)$$

subject to:

$$D_i^V \le R_i^V, i \in [1, \dots, Nr] \quad (2)$$

where $O_{env}(D)$ and $O_{cost}(D)$ represent the environmental and economic objective function respectively. S^G is the groundwater storage gain and C represents cost. Both depend on $D=(D_1, D_2, \dots, D_{Nr})$ which are the decision variables that express the active diversion routes and Nr is the total number of the active diversion routes for each scenario. A diversion route is characterized by a river point where the water is diverted and a MAR facility that receives the diverted water. The optimization is constraint by the surface water availability. For every diversion route the requested diversion volume D_i^V must be less than the available amount of water in the river R_i^V. In the following paragraphs we expand on the formulation of the two objective functions and constraints

2.1 Environmental Objective Function

The environmental objective O_{env} of MAR is a function of the groundwater storage gain S^G. The groundwater storage (S) is not a field measurable value and it can only be estimated via numerical simulations. The majority of groundwater simulation models such as MODFLOW, IWFM, FEFLOW etc. are able to calculate groundwater storage per discretization unit (e.g. grid cell or mesh element). In transient-state models the storage value at each discretization unit is a time series In our study we are interested in the long-term storage, and we assume that the storage gain of a single time step can be representative of the overall effect of a MAR scenario as long as the MAR scenario has been active for a sufficiently long period. Therefore, the storage gain can be calculated as:

$$S^G = \sum_{i=1}^{N_{el}} S(t^{SA})_i^{MAR} - S(t^{SA})_i^{BAU} \quad (3)$$

where $S(t^{SA})_i^{MAR}$ is the calculated storage for element i under a given scenario, $S(t^{SA})_i^{BAU}$ is the groundwater storage for the business as usual (BAU) scenario where no additional MAR

operations are considered, both calculated at time t^{SA} which is selected as representative time for storage assessment and N_{el} the number of discretization units considered in the calculation.

2.2 Economic Objective Function

In this study the economic objective function O_{cost} is formulated considering the cost of establishing and operating a MAR facility. The total MAR cost is the sum of four components that form the major components of the construction and operation costs of MAR facilities (Ross and Hasnain, 2018):

$$C = C_{land} + C_{capital} + C_{lift} + C_{conv} \tag{4}$$

where land C_{Land} and capital $C_{capital}$ costs are fixed ("capital") and depend upon the maximum amount of water diverted in any time period to a MAR facility. The pumping lift C_{lift} and conveyance C_{conv} costs are variable ("marginal") and depend upon the total amount of water diverted to a MAR site over the full period and, due to discounting, the timing of those diversions. Since the optimization seeks out lower cost portfolios of MAR sites, this objective function essentially penalizes sites that have high fixed costs, which is related to the size of the MAR facility, and sites that have high variable costs (e.g. distance between diversion points and MAR sites, changes in elevation).

Land cost

We calculate the land cost as:

$$C_{land} = \sum_{i=1}^{Nr} P_{land,i} c_a \frac{Q_{max,i}}{c_b} \tag{5}$$

where $P_{land,i}$ is the land price($\$/L^2$), of the acquired land i, $Q_{max,i}$ (L^3) is the maximum diversion volume of water to i MAR site within a specified time period; c_a(-) is a coefficient that is used to convert the maximum diversion volume available for a time period to a maximum daily value; and $c_b(L)$ reflects the maximum depth of water that can be recharged through MAR. The land cost depends on the land price and on the size of the MAR facility, which is determined by the

maximum amount of water that can be recharged at any time over the full simulation period. Typically, transient-state hydrologic models operate on time steps larger than daily. A typical time step in regional model is a month. Therefore, an additional assumption is needed to calculate a daily maximum diversion volume, $Q_{max,i}$ based on the frequency at which source water for recharge is available. If wastewater is used for MAR, one can expect a constant volume of water to be available for recharge every day. However, if high-magnitude flow is used for recharge, water availability can highly vary in space and time depending on the frequency at which rain events and storm flows occur. For example, Kocis and Dahlke (2017) estimated for the California Central Valley that high-magnitude flows (e.g. flood flows) are only available for a few days each month. Hence, we introduce coefficient c_a to account for the fact that the total diversion volume over a time step would not occur in one day $c_a = 1$ and neither be spread evenly over the time step $c_a = 1/N_{days}^{step}$ where N_{days}^{step} is the number of days within a time step. To estimate the corresponding land area needed to recharge the maximum daily volume we introduce coefficient $c_b(L)$ that corresponds to the maximum water depth that can be accommodated at a MAR site.

Capital cost

We construct capital cost in a similar manner since the total size of MAR basin depends on the maximum volume of water stored in the basin at any point over the run of the model:

$$C_{capital} = \sum_{i=1}^{Nr} P_{basin,i} c_a c_b Q_{max,i} \tag{6}$$

where $P_{basin,i}$ is the per-unit-area construction cost for a basin i.

Pumping lift cost

We construct the pumping cost based on $Q_{t,i}(L^3/T)$, the total amount of water diverted to i MAR site in a given year (t), as follows:

$$C_{lift} = \sum_{i=1}^{Nr} \sum_{t=1}^{T} \delta_t (P_{el} E_{lift} X_{lift,i} Q_{t,i}) \tag{7}$$

where P_{el} is the price for electricity ($\$/kWh$), E_{lift} is the energy lift (Wh), $X_{lift,i}(L)$ is the vertical lift from the diversion point to the MAR site, and δ_t is the discount factor for each year.

Conveyance cost

We estimate the conveyance cost as:

$$C_{conv} = \sum_{i=1}^{Nr} \sum_{t=1}^{T} \delta_t \left(P_{conv} X_{dist,i} Q_{t,i} \right) \tag{8}$$

where $X_{dist,i}(L)$ is the distance from the diversion point to the MAR site i; P_{conv} is the per volume per distance conveyance cost; and $Q_{t,i}$ and δ_t are defined as above. In reality and depending on data availability the conveyance cost for MAR could consider several additional factors such as additional pumping costs, canal maintenance costs, regulatory costs, and conveyance losses. Since the pumping and lift cost is already accounted for in our study, any additional pumping costs would involve efforts to overcome friction within the conveyance system to transfer water at junctions within the system.

2.3 Constraints

The main constraint in our optimization formulation is the surface water availability. In this paper we explore the possibility of utilizing the excess water that is available from high magnitude flows (HMF). HMF is defined (Kocis and Dahlke 2017) as the water volume that can be diverted after a certain streamflow threshold is reached. Here we set the threshold V^p as a specified percentile of the cumulative probability function of the daily or monthly flows and assume that the water volume that is higher than this threshold is available without adverse downstream effects.

In addition, we impose a limit on the water volume that can be diverted. At times streamflow can be much higher than the threshold V^p, effectively exceeding the capacity of the conveyance infrastructure. To utilize the entire HMF volume available for MAR would require increasing the capacity of the existing conveyance infrastructure, which is quite expensive considering that these flows are available once every 5 to 10 years (Kocis and Dahlke, 2017). To avoid expanding

existing infrastructure, we introduce a *cap* parameter to enforce a maximum diversion amount. Therefore, at any given time the available water for diversion is:

$$D_t^V = \begin{cases} min\left\{R_t^V - V^p, cap\right\} & if\ R_t^V > V^p \\ 0 & R_t^V \le V^p \end{cases} \quad (9)$$

where R_t^V is the volume of water that flows at time t.

2.4 Multi-objective optimization

The decision variables of the multi-objective formulation (eq. 1) are all the possible diversion routes from rivers to MAR facilities. The decision variables represent whether water is allowed to flow from a river to a MAR facility or not, encoded as binary values. This results in an optimization formulation where the objective and constrain function gradients cannot be calculated. A typical method to tackle such optimization problems is to use genetic algorithms (GAs). GAs solve optimization problems by imitating the natural evolution of species over several generations, where an initial population is evolved by genetic processes to improve the species and no gradient information is required. In each generation, a number of individuals (the fittest) are selected and recombined to produce the population of the next generation using genetic processes such as *selection* (select the fittest individuals of a population), *mating* (crossover of genes between parents), and *mutation* (random changes in genes). Since the fittest members have a higher likelihood to survive, the overall population will improve to a near global optimum after a few generations. The details of the GA are provided in the SI (TEXT S1). GAs are known for their ability to solve highly complex and non-linear optimization problems. However, they often require a significant number of objective function evaluations. One of the reasons is that during the genetic evolution, the population may become saturated, i.e. the same individual appears at increasing frequency either within the same population or different generations. When the calculation of the objective function is not time consuming (e.g. it is a simple algebraic expression) it may be acceptable to allow evaluating the same solution several times until the genetic operators find a better offspring. However, if the objective function involves running complex hydrologic models as in our MAR formulation, then evaluating the same candidate solution more than once makes the GA optimization inefficient. In this work we

use a hybrid genetic algorithm which combines the well-known GA NSGAII (Non Dominating Sorting Algorithm II, Deb et. al., 2002) and components of tabu search algorithms (Glover 1989).

In particular, we introduce a tabu list which keeps the encoding of every solution that has been evaluated. All the new individuals are first compared against the tabu list. If they are identical with any member of the tabu list, the individuals undergo further mutation until they are unique. Using this approach has two distinct advantages. First it does not allow the evaluation of the same individual. Secondly, the mutation rate becomes self-adaptive. In general, the mutation operator is a mechanism to allow the GA to escape the local optimum. The mutation rate is usually a very small percent. Using large percentages makes the GA optimization unstable to converge to the global optimum. In our approach we are still setting a constant mutation rate which is however increased adaptively when the individual members of the population become identical, allowing the algorithm to converge to the optimum solutions while making it easier to search new areas of the decision variable space.

3 Application

3.1 Study Area

The simulation-optimization framework was applied to the San Joaquin–Tulare Lake basin (33,000 km^2, 35°-38°N, 118°-120°W) located in the southern part of the Central Valley (CV) in California, USA (Fig. 1). The region is one of the most productive agricultural areas in the world with a total farm output of over $41.5 billion in 2019 (CDFA, 2020). Situated in a Mediterranean climate, about 40% of water used in California comes from groundwater. As such, the region is representative for many semi-arid agricultural regions around the world that rely on groundwater for irrigation. Crop production in the CV uses a complex system of surface reservoirs and canals that move water from the wetter north to the drier south. Surface water in the state is managed through two major water infrastructure projects. The State Water Project (SWP) diverts water from the Feather River and Oroville reservoir to southern California via the California Aqueduct

and the A.D. Edmonton Pumping Plant over the Tehachapi Mountains to the Los Angeles metropolitan area. The Central Valley Project (CVP) operated by the federal Bureau of Reclamation consists of multiple reservoirs, canals, hydroelectric power plants and other facilities including the Friant-Kern Canal (FKC), a 245 km aqueduct located on the east side of the San Joaquin-Tulare Lake basin, which delivers water to Fresno, Tulare and Kern counties (Fig. 1). Total annual precipitation averages 582 mm a^{-1} in California but can be as high as 3000 mm in the north and <150 mm a^{-1} in the south. Average annual evapotranspiration (1961–2003) in the CV ranges from 1100 to 1450 mm a^{-1}, with the highest rates observed in the San Joaquin-Tulare Lake basins (Faunt 2009). In the past two decades, prolonged droughts and below-average precipitation have much reduced surface water reliability (Famiglietti et al., 2011; Xiao et al., 2017) leading to increased groundwater use and groundwater storage depletion in the Central Valley alluvial aquifer. Groundwater depletion has caused a number of other issues such as subsidence (Faunt et al., 2016), water quality degradation (Gailey 2017; Levy et al., 2021), loss of groundwater dependent ecosystems (Duffy & Kahara 2011) and increasing pumping costs due to groundwater level decline (Vasco et al., 2019; Gailey et al., 2019b; Pauloo et al., 2020).

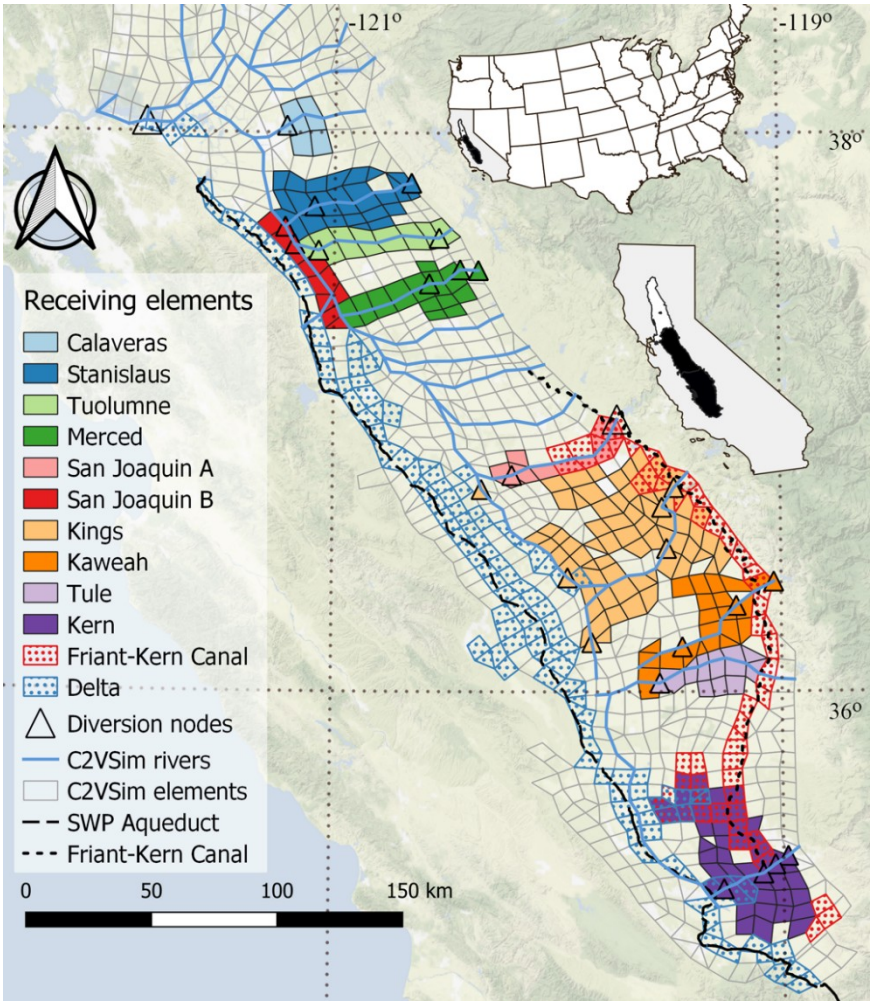


Fig. 1 Description of study area. The figure shows the diversion nodes and their corresponding receiving elements in matching colors.

3.2 Groundwater Storage computation

To estimate the groundwater storage efficiency from MAR operations we use an integrated hydrologic model that simulates the surface water – groundwater balance in the region. Specifically, we use the California Central Valley Groundwater-Surface Water Simulation Model (C2Vsim), developed by the California Department of Water Resources (CDWR) (Brush et al. 2013), to identify suitable MAR locations considering economic cost of MAR and hydrologic benefits. A retrospective modeling approach, similar to that of Kourakos et al. (2019) and Gailey et al. (2019a), is used to identify promising MAR locations in the San Joaquin – Tulare Lake basin over the period 1969 to 2009.

C2Vsim uses the Integrated Water Flow Model (IWFM) code, which is a finite element-based model code that simulates the majority of the hydrologic processes that take place in agricultural basins including groundwater and stream flow, stream-groundwater interaction, soil water balance, lake storage, and land surface flows (Brush et al. 2013). In addition, IWFM is capable of calculating agricultural and urban water demands considering a number of factors such as soils, land use/land cover, evapotranspiration, domestic water use, and population. The IWFM code was specifically developed by CDWR to support water resources management and planning efforts of water districts and Groundwater Sustainability Agencies (GSA) in California. More detailed information on the model structure, water balance components, calibration and validation can be found in Brush et al. (2013).

Multiple versions of the C2Vsim model have been developed over the last 15 years. In this study we use the C2Vsim Coarse Grid v3.02 version, which has a run time of a few minutes. The model simulates the historic surface water – groundwater system from 1921-2009 on a monthly time step. However, for this study we consider only the second half of the total simulation period from 1963-2009. Fig. 1 shows the model domain for the study area, the stream network, and the finite element mesh for the model. The coarse mesh C2Vsim model has an average finite element size of 37.19 km^2 ± 13.54 km^2 while vertically the model is discretized into 3 layers. At that scale the model cannot capture fine details of groundwater flow such as flow around wells or represent recharge basins at field accuracy.

Within the C2Vsim framework, MAR is simulated as diversions of high-magnitude streamflow from stream nodes to finite elements that serve as spreading grounds. Most surface water in California's CV is legally allocated, according to Kocis & Dahlke (2017) winter flood flows and runoff from high-magnitude storm events (e.g., flow above the 90^{th} percentile) that exceed environmental flow requirements provide the only viable and physically available surface water source left in California to expand MAR. For the San Joaquin-Tulare Basins, Kocis & Dahlke (2017) estimated that high-magnitude flows occur on average 4.7 out of 10 years from just a few storm events (5–7 1-day peak events) lasting on average for 10–30 days between November and April. In C2Vsim, surface water diversions are defined as time series, however, during the simulation the model adjusts the diversion amount to the streamflow available at each node. If

the requested diversion amount is greater than the available streamflow, the diversion amount is adjusted to the available streamflow and the difference is reported as diversion shortage. The water that is diverted from a stream node is applied as groundwater recharge (i.e., influx of water from the unsaturated zone into the first groundwater layer) to selected finite elements. The diverted water does not infiltrate and percolate through the root and unsaturated zone and is rather assumed to instantly reach the groundwater table, which is justified based on the fact that most recharge occurs in winter when evapotranspiration losses are small. The authors acknowledge that this approach does not consider the potential time lag that flow through the unsaturated zone will create. The time lag depends on factors that vary in space such as percolation rates, depth, soil etc. For example, in the northern part of the CV the unsaturated depth is generally less than 30 m (Gailey et al., 2019a) while in the southern part it can be as high as 200 m. Therefore, the influence of the time lag is different across the CV even within our study area. However, this time lag is negligible when assessing the long-term water budget and response of the groundwater – surface water system to MAR (e.g., see Maples et al. 2019 for a discussion).

3.3 Economic Objective Function for the test application

Land transaction data provided by CoreLogic (https://www.corelogic.com/) was used to estimate the land cost, P_{land}, in each finite element. P_{land} varies across elements from about \$1.48/m^2 (\$6,000/acre) to \$4.94/m^2 (\$20,000/acre) (Fig. S3). The source water for MAR in our application is high magnitude flow (HMF), defined as flow either above the 90th or 95th percentile of the full record (Kocis & Dahlke, 2017). Since C2Vsim operates on a monthly time step we assume that the recharge water is only available for five days $c_a = 1/5$ each month based on statistical analyses done by Kocis and Dahlke (2017). To estimate the corresponding land area needed to recharge the maximum daily volume we assume, based on field experiments and water availability (Dahlke & Kocis, 2018; Ma et al., 2022), that a maximum water depth of 4.52 m (15ft) ($c_b = 15\,ft$) can be recharged within the winter rainy season (Nov-Apr). As the hydrologic model is too coarse to capture any spatial variability at a field level, the diverted water is spread

proportionally to the element area so that the elements that are receiving diversions from the same diversion node have the same rate. This reflects the fact that in larger elements it is likely to acquire more land.

To calculate the capital cost, we set a fixed per-acre construction cost P_{basin} at \$1.24 m^2 (\$5,000/acre) for the entire CV (personal communication Jonathan Parker, Kern Water Bank Authority).

To calculate the pumping cost, we estimate the vertical lift $X_{lift}(\text{ft})$ from a diversion point to a MAR site from the simulation model input data (Fig. S4). Based on the assumption that a perfectly efficient pump would take 2.7 Wh to lift one cubic meter of water one meter in elevation (1.02 kWh to lift one acre-foot of water one foot) we assume it would require $E_{lift} = 3.857\,\text{Wh}$ per meter ($1.45\,kWh$ per foot) of lift for a pump with 70% efficiency. We also assume that irrigation districts would pay the average commercial price for electricity of $P_{el} = \$\,0.17/kWh$. To construct the discount factor for each year, we use a discount rate of 3 percent. This is a real discount rate that incorporates the assumption of an annual 2 percent inflation in energy prices.

For the conveyance cost we estimate the distance X_{dist} in miles from the diversion point to the barycenter of the recharge sites (Fig. S5) and estimate the overall conveyance cost P_{conv} as \$0.02 per acre-foot per mile (\$0.026 per m^3 per m).

3.4 Recharge volume and MAR land scenarios

In this application of the SO framework, we explore different recharge volume and MAR land scenarios. The main source of water for all MAR scenarios evaluated in this study is high-magnitude flow as defined above (Kocis & Dahlke, 2017). For the optimization, we assume that conveyance infrastructure is mostly used during the summer months for transporting irrigation water and that conveyance is available at full capacity during the winter months to transport HMF to recharge areas. In addition, we only use existing diversion nodes and agricultural lands defined in the model to receive surface water from the diversion nodes. No new diversion nodes or agricultural areas are explored in the model for MAR to avoid the associated cost of building

new water infrastructure, which allows identifying optimal MAR locations that maximize groundwater storage under the current infrastructure.

For the MAR land scenarios, we consider three possible sources of HMF in the optimization: (i) Local scenario: diversion of HMF from ten major rivers exiting the Sierra Nevada mountains on the east side of the San Joaquin-Tulare Lake basin onto nearby elements located within the native river basin. This scenario does not transfer water across basins; (ii) Friant-Kern Canal scenario: diversion of HMF from rivers and the Friant-Kern Canal onto associated elements within nearby groundwater basin; and (iii) Delta scenario: diversion of HMF from local rivers, the Friant-Kern Canal and the Sacramento-San Joaquin delta onto associated elements within the San Joaquin-Tulare Lake basin (Fig. 1). Colored elements and triangles in Fig. 1 show the ten rivers used in all three scenarios, surface water diversion nodes along each river, and associated elements receiving water from each river. Fig. 1 also highlights that for some rivers there exist multiple diversion nodes. More detailed descriptions of the diversion scenarios and their set up in the optimization framework are given in section S2 and Table S1 in the Supplemental Materials.

~~*Table 1: Surface water diversion scenarios used in this study and their set up in the optimization framework.*~~

~~**Diversion scenario**~~	~~**Decision variables**~~	~~**Populations**~~	~~**Generations**~~
~~**Local**~~	~~258~~	~~256~~	~~400~~
~~**Friant-Kern Canal**~~	~~320~~	~~384~~	~~300~~
~~**Delta**~~	~~459~~	~~512~~	~~400~~

In our study, two HMF thresholds (90th and 95th percentile) are evaluated. Fig. 2a shows the 90th and 95th HMF percentiles for the Tuolumne River, which correspond to 113 m^3/s and 150 m^3/s, respectively. When HMF flows occur, the volume of water that is available for MAR is quite large (Fig. 2a). Conveyance of HMF likely requires an increase in existing conveyance infrastructure capacity, which is quite expensive considering that these flows are only available once every decade. To avoid this, we enforce a maximum diversion amount (cap). Two caps are considered in this study: 1) 100,000 acre-feet (100 TAF) per month corresponding to 0.123 BCM

(billion cubic meter) and 2) 200 TAF corresponding to 0.247 BCM/month. Therefore, the total amount that is diverted in the model for MAR ranges between the percentile and the cap. For the Tuolumne River the total amount of water that can be diverted during the simulation period (10/1965 – 9/2009) ranges between 2 BCM (95th percentile 100 TAF cap scenario) and 6 BCM (90th percentile – 200 TAF cap scenario) (Fig. 2b). Since our simulation assumes that HMF occurs for a few days every month ($c_a = 1/5$), the maximum available HMF diverted in a given month is cap/5. Likewise, because we assume that a maximum water depth of 4.52 m (15ft) ($c_b = 15\,ft$) can be recharged within the winter rainy season (Nov-Apr), the recharge volume requires a maximum MAR basin area of 5.4 km^2 (1.34 acres) for the 100 TAF cap and 10.8 km^2 (2.67 acres) for the 200 TAF cap.

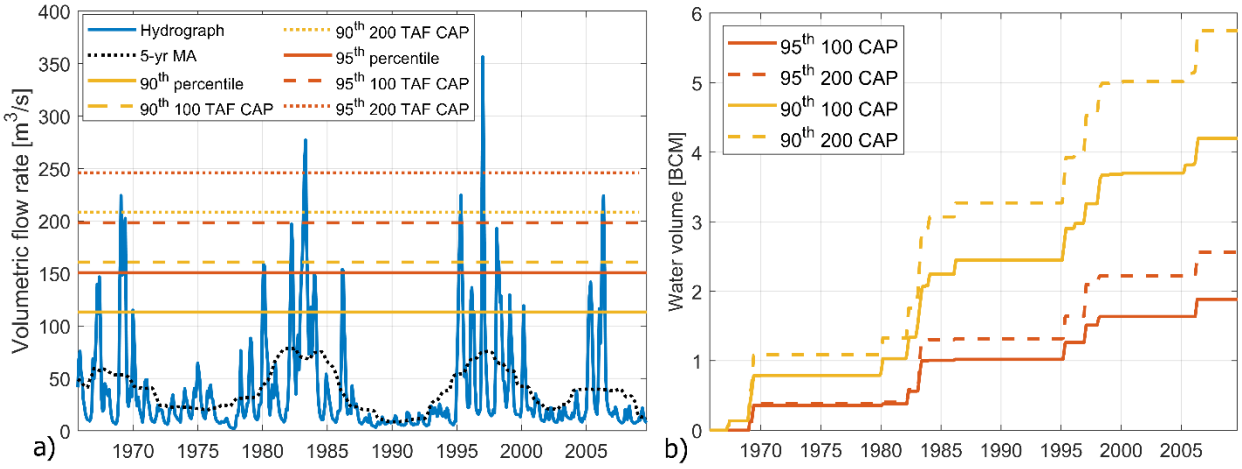


Fig. 2 a) Streamflow hydrograph extracted from C2Vsim for the Tuolumne River. B) Cumulative HMF amount available for MAR for the 95th and 90th percentiles and the annual diversion limits of 100 TAF and 200 TAF, respectively. MA in the legend means moving average.

Depending on the HMF percentile and the cap threshold, the total cumulative amount of water available for recharge over the 46-year, 1963-2009 modeling period is quite variable among the scenarios. For the local scenario, the 95th percentile provides significant excess water only during four wet years (1969, 1983, 1995-1998, 2006), while for the 90th percentile there are more than ten years where excess water is available. Imposing a cap on the diversion amounts mainly influences the total amount available from larger water sources (e.g. Delta, FKC, larger river) but has limited effect on the amounts available from smaller systems. Overall, cumulative (1969-2009) diversion amounts vary between 0.9 and 11 BCM for the ten rivers depending on the HMF

threshold and cap. In comparison, the cumulative diversion amounts available from the FKC and Delta scenarios vary between 2.5 km^3 and 9.5 km^3 and 3.2 and 12.3 km^3, respectively, with the 90th percentile and 200 TAF cap scenario providing the most water (Fig. 3). In reality, HMF volumes available from the Delta are considerably larger than 200 TAF per month, but the majority of this water is fully allocated to downstream users.

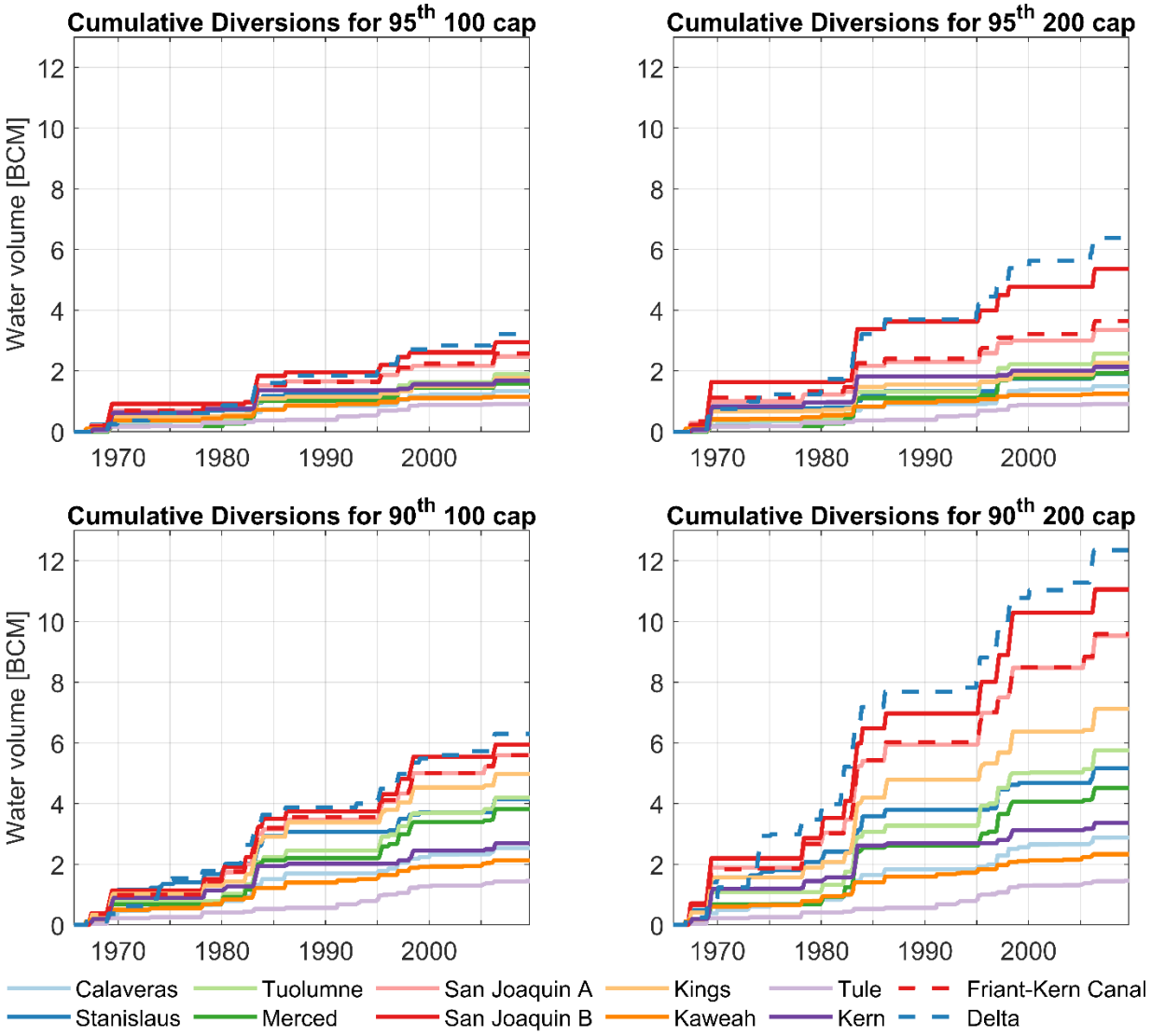


Fig. 3 Cumulative surface water diversion amounts available from the ten rivers within the San Joaquin-Tulare Lake basin, the Friant-Kern canal, and the Delta.

4 Results and Discussion

4.1 Optimization

For each scenario we solve four optimizations considering the different HMF thresholds (90th, 95th) and cap limits (100 TAF, 200 TAF), resulting in 12 total optimizations. In the following paragraph we describe in detail the optimization results of the SO framework application using the example of the local scenario with 95th percentile flows and a cap of 100 TAF. The goal of this exercise is to examine the evolution of the optimization and test whether it converged to a near optimum solution, which is done by calculating the hypervolume (i.e., area between the Pareto front solutions and a reference solution that is dominated by every possible solution) for each generation. During the first 100 generations the hypervolume expands rather rapidly (Fig. 4a) as the optimization finds solutions that push the Pareto front away from the reference point (e.g. the worst-case solution for both objectives). After 100 generations the hypervolume increases at a slower rate until it reaches a plateau after about 300 generations. Fig. 4a (orange line) shows that the number of solutions is increasing over the first 150 generations following the same pattern as the hypervolume metric. After 150 generations the number of solutions starts to oscillate between 80 and 110. At the same time, we observe that the change in hypervolume for these generations is rather negligible. During the first stages of the optimization the entire Pareto front moves towards the optimal Pareto front while in the later generations the Pareto front undergoes minor shifts, and the optimization focuses on the density of the Pareto front. When comparing the optimizations across different scenarios, we find that the shape of the hypervolume evolution is very similar in all optimizations conducted in this study (Fig. ~~S5~~S6), indicating that the evolution of the number of Pareto solutions depends highly on the random genetic operations, which is very different for each optimization but always follows a similar pattern – a rapid increase in the number of Pareto solutions followed by an oscillation pattern.

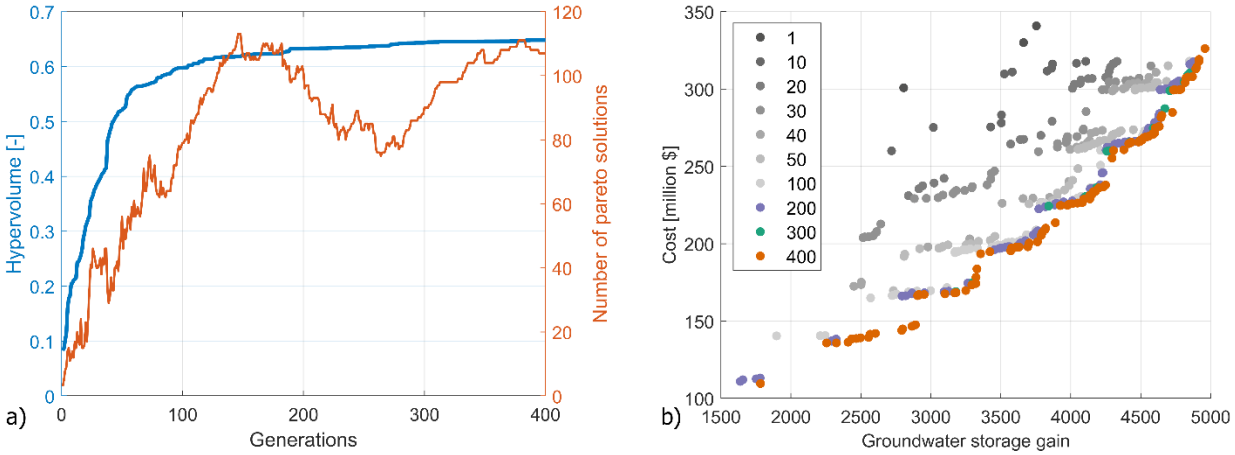


Fig. 4: a) Normalized hypervolume (left axis) and number of solutions (right axis) for the optimization for the local scenario with 95th percentile HMF and a 100TAF cap b) Evolution of the Pareto fronts for selected generations.

Fig. 4b shows that within 50-100 generations the Pareto front is shifted very close to the final solution. However, the individual solutions of the final (400) Pareto front are spread more evenly and are also expanded on both sides of the front. Similar trends were observed in all other optimization runs, which gives confidence that the final solutions are very close to the near global optimum Pareto front.

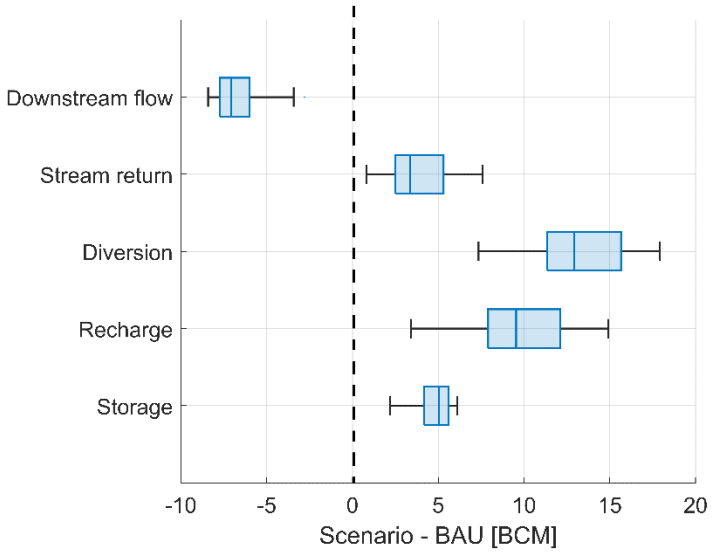


Fig. 5 Difference in water budget components shown for the Pareto solutions for the local, 95th percentile and 100 TAF cap scenario in comparison to the BAU scenario. Diversion is the amount of surface water diverted from rivers for recharge and Recharge is the amount of water applied on selected finite elements for recharge minus transportation losses.

To understand the impact of the Pareto solutions on the long-term water budget, we examine the water budget components that are impacted by the MAR scenarios. These are the downstream flows Q_{down}, stream return flows Q_{return}, diversion flows $Q_{¿}$, groundwater recharge Q_{rch}, and groundwater storage Q_{strg}. The upstream flows Q_{up} that are used in the following equations are identical between the BAU and MAR scenarios. The water budget at each stream node dictates that $Q_{down}=Q_{up}+Q_{return}-Q_{diversion}$, while at each element the groundwater storage is $Q_{strg}=\sum Q_i+Q_{rch}-Q_{return}$ where $\sum Q_i$ represent all of the remaining flows in or out of an element. Fig. 5 shows the water budget components for each Pareto solution of the local, 95th percentile, 100 TAF cap scenario in comparison to the business-as-usual (BAU) scenario using the C2Vsim model (Fig. 5). The diversion of streamflow for MAR is negatively impacting downstream flows below the diversion point, as indicated by the lower streamflows in the MAR scenarios compared to the BAU simulation. However, the downstream flow reduction is approximately half of the diverted amount, indicating that about half of the recharged water is returned back to the stream at a later time thus increasing the flow downstream. As expected, the groundwater storage at the end of the simulation period is greater than the BAU scenario. The increase in storage is approximately 30% of the diversion amount. The remaining 70% is returned to the stream. When comparing the water budget components for all other scenarios, we observe that this pattern is consistent across all optimizations, however the exact water budget amounts differ considerably among the individual Pareto solutions.

4.2 Pareto solutions of different MAR scenarios

4.2.1The role of cost in Pareto solutions

Fig. 6 shows a step-like pattern for the Pareto front for the local, 95th percentile, 100 TAF scenario, which is mainly due to solutions being clustered based on the number of rivers considered in each solution. When examining the cost function of the Pareto front, we observe that the cost between clusters either increases stepwise (i.e., from $N_{¿}$ 5 to 6 and 7 to 8, 9 to 10) or sharply (i.e. $N_{¿}$: 6, 8). The cost jumps are mainly due to the fixed capital cost of a MAR basin, which is increasing for each river that is added to the solution.

As shown in Fig. 6, for the smallest number of rivers considered (N=4), total MAR cost is mainly made up of land cost (70% of total cost; blue filled area) and capital cost (30%; orange filled area). In contrast, the lift cost exhibits a non-linear increase, because for every diversion node there are elements with zero lift (elevation of diversion point is higher than the receiving element) and elements with non-zero lift cost (elevation of diversion point is lower than the receiving element). Based on the optimization formulation (eq.1), elements with zero lift cost are generally more attractive. However, elements with zero lift cost can also be limited in their storage gain (e.g. due to storage availability, proximity to streams etc.). To overcome this limit, the algorithm has to either include non-zero lift cost elements or add an extra diversion from another river which expands the candidate list of elements for MAR including additional zero and non-zero lift cost elements. Fig. 6 suggests that it is cheaper to overcome this limit with non-zero cost element due to the high capital cost. However, there is a point where the lift cost is greater than the capital cost. For these cases, MAR locations that increase groundwater storage are located at higher elevations than the diversion node, hence requiring a more energy to lift the water to the recharge area. When a new river is added, the lift cost decreases sharply since optimal recharge locations can be found at locations with zero lift cost, thereby reducing the energy costs required to transport the diverted water. However, adding a new river increases capital and land cost, since the total number of MAR locations increases compared to a scenario that diverts water from fewer rivers. The overall conveyance costs are negligible in this scenario.

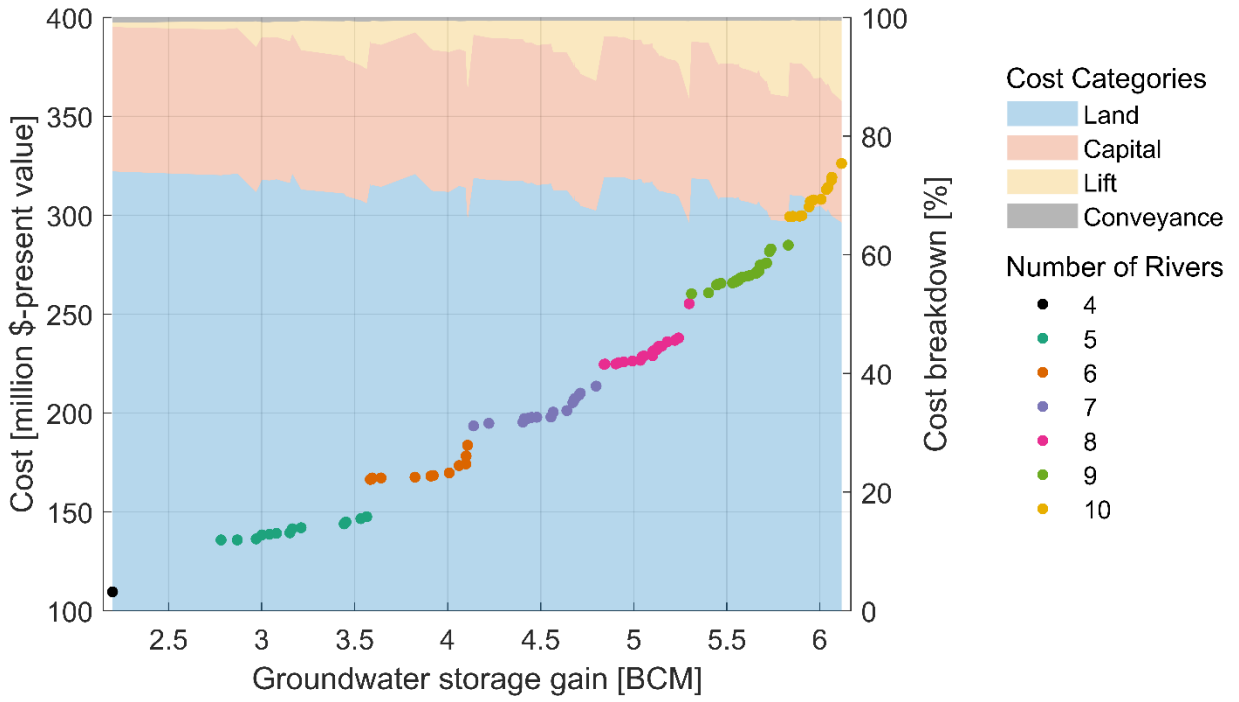


Fig. 6 Pareto solutions for the local, 95th percentile, 100 TAF scenario. Solutions are clustered according to the number of diversion points. The background colors indicate the cost breakdown as a percent fraction of total cost (right axis).

4.2.2Efficiency of Pareto solutions

A key consideration for the hydrologic objective function is the selection of MAR sites that maximize storage and reduce return flows of groundwater to streams. The trade-off between these two competing objectives can be evaluated by assessing the efficiency of a particular solution in terms of the share of total diverted water used for recharge that remains in the groundwater aquifer. As shown in Fig. 7a, for the lowest cost, lowest storage solution about 33% (1 BCM) of the total diverted water recharged (3.2 BCM) returned to streams and 67% (2.2 BCM) remained in aquifer storage. In contrast, for the most expensive solution with the largest groundwater storage gain, 55% (7.2 BCM) returned to streams and only 45% (6 BCM) remained in groundwater storage. However, as clearly shown in Fig. 7b, as the cost of the solution increases, which also increases the groundwater storage gain, the overall efficiency of MAR decreases at a rate of 14% per $100 million. Note that the variability between the groups decreases as the number of rivers increases. When five rivers are considered, the MAR efficiency varies between 65 and 80%, while for 10 rivers it varies between 44 and 48%. This is because

there are multiple combinations of 5-river sets and the number of combinations decreases as the river number increases.

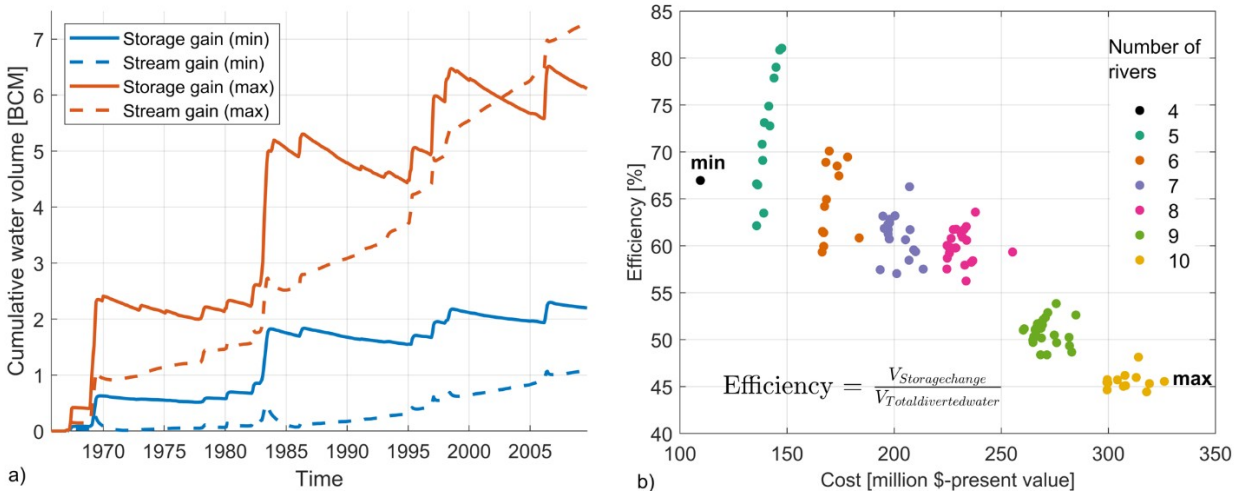


Fig. 7 a) Groundwater storage and streamflow gain for the two most extreme solutions of the optimization for the local, 95th percentile, 100 TAF cap scenario b) Efficiency of Pareto solutions for the local, 95th percentile, 100 TAF cap plotted as a function of cost and grouped according to the number of activated rivers.

4.2.3Local diversions scenarios

Fig. 8a shows the Pareto fronts for the different HMF thresholds (90th and 95th percentile) and cap (100 and 200 TAF) scenarios. As mentioned before, both the HMF threshold and the cap limit affect the total amount of water diverted for MAR. For both HMF thresholds, there is a range of solutions where the same amount of storage gain can be achieved at a lower cost if the cap is lowered. For example, using the 90th percentile (yellow and purple dots), there is a range between 5.5 and 12.5 BCM of storage gain where the Pareto solutions of the 100 TAF cap achieve the same storage gain as the 200 TAF cap, but at a much lower cost. However, because higher cap values allow more water to be diverted, the overall storage gain that can be achieved under the 200 TAF cap scenario is higher. Under the 200 TAF scenario the diverted amount is doubled, yet we see that the maximum storage gain is less than double for both the 95th and 90th percentiles.

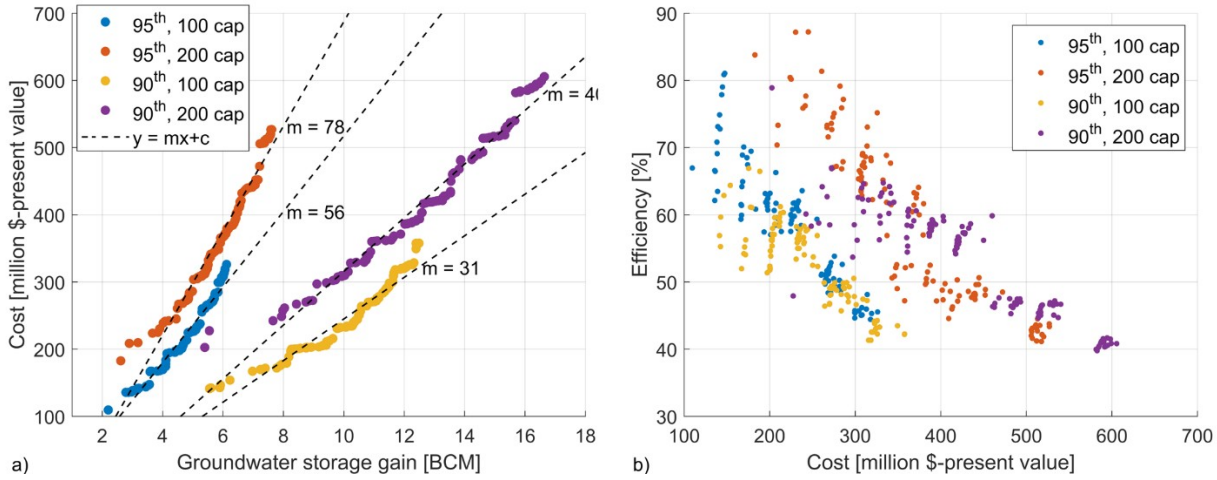


Fig. 8 a) Pareto fronts for all local scenario optimizations b) Efficiency of pareto solutions for the 4 local scenarios.

Fig. 8a also highlights the importance of the frequency of diversion events. The Pareto sets that use the 95th percentile exhibit an average diversion cost of $56 and $78 million per BCM of storage gain for the 100 and 200 TAF cap scenarios, respectively. The diversion cost of both 90th percentile scenarios are approximately halved at $31 and $40 million per BCM for the 100 TAF and 200 TAF scenarios, respectively. By comparing the Pareto fronts of both 100 TAF cap scenarios (blue and yellow fronts) we see that for the price of $300 million for example not only is a higher volume of water diverted but also the storage gain is doubled for the same cost.

When comparing all optimizations, we observe that the efficiency of the Pareto solutions follows the same pattern (Fig. 8b). In general, lower cost solutions, which divert less water compared to the high-cost solutions, have a higher efficiency because these solutions minimize the return flow to streams therefore yielding higher efficiency. Note also that the optimizations for the 95th percentile return several (10-20) solutions with efficiencies greater than 70%. These are low-cost solutions which divert water from a few rivers (5-6), therefore there are many more combinations of MAR sites that minimize baseflow compared to scenarios where all rivers and potential MAR sites are used. On the other hand, optimizations that use the 200 TAF cap are generally more efficient for a given cost compared to the 100 TAF cap mainly due to the frequency of water availability. For $300 million, the efficiencies of the 100 TAF cap Pareto

607 solutions vary between 45-50%, while for the 200 TAF cap the efficiencies vary between 55 –
608 75%.

609 The Pareto front of the local, 95th percentile, 100 TAF scenario consists of 107 individual
610 solutions. Although each solution consists of a unique combination of diversion nodes and finite
611 elements receiving water for MAR, there are several elements that tend to get selected more
612 often than others by the SO framework. The elements selected under the two HMF thresholds
613 and two cap scenarios follow a very similar probability selection pattern (Fig. 9). Elements
614 located south of the Kings and Kern rivers and close to the San Joaquin River are selected more
615 frequently (yellow colors). This is mainly due to the overall falling price gradient from east to
616 west (Fig. S3). For the diversions from the San Joaquin, Merced, and Stanislaus rivers we see
617 that there is one element for each diversion node, that is most frequently selected in the Pareto
618 solutions while the other elements are either never selected or only a few times. It is interesting
619 to note that although the Tuolumne River provides higher water volumes compared to the other
620 rivers the optimizations do not choose elements that correspond to the Tuolumne River. This is
621 likely due to the fact that land prices of the receiving elements within the Tuolumne River basin
622 are the highest in the study area (Fig. S3).

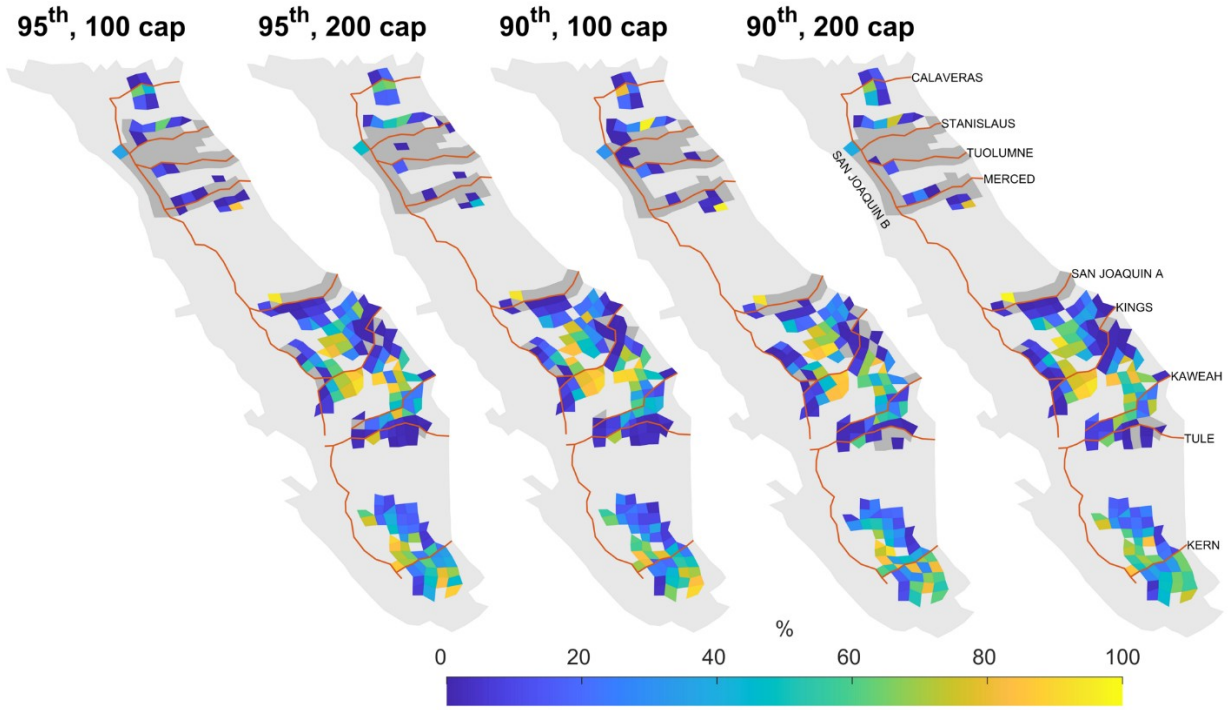


Fig. 9 Probability (in percent) of finite elements being selected for recharge by a Pareto solution for the local scenario.

4.2.4Friant Kern Canal and Delta diversions scenarios

In the Friant-Kern-Canal (FKC) and the Delta scenarios (Fig. 10), where finite elements can receive water from 2 or 3 diversion nodes at a time, the resulting Pareto fronts are very similar to those for the local scenario. However, the FKC and Delta scenarios show a different behavior in terms of cost breakdown. Land cost makes up about 60% and 50% of the total cost for the FKC and Delta scenarios, respectively. Since the capital cost is fixed across these scenarios the change in total cost is mainly caused by the higher energy or lift cost. For the Delta scenario the energy cost is almost 40% of the total cost since water needs to be lifted from mean sea level to the receiving elements (Fig. 10). The FKC and Delta scenario also show a higher conveyance cost compared to the local scenario, which is caused by the longer distances over which water is being transported.

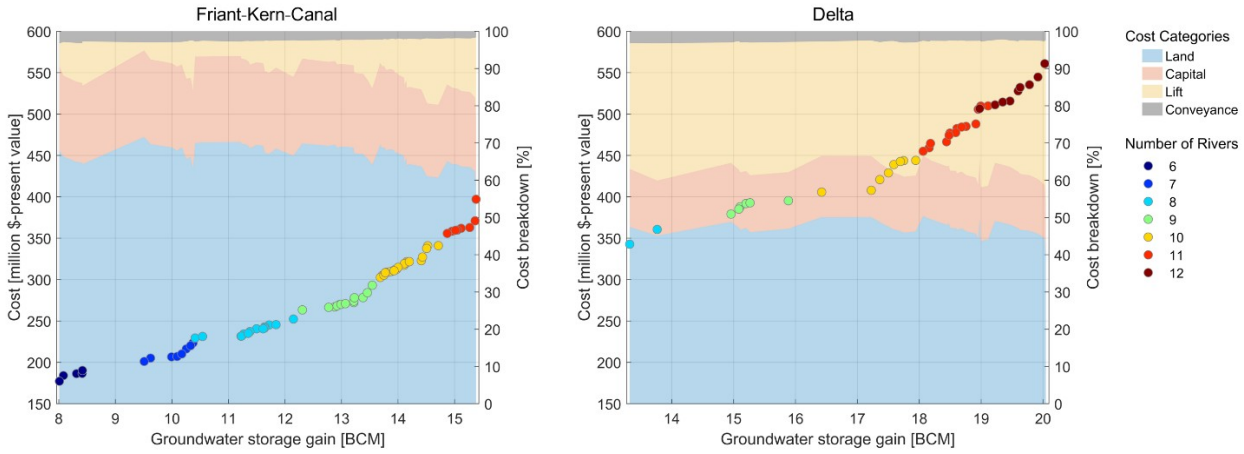


Fig. 10 Pareto fronts and cost breakdown for the 90th percentile and 100 TAF cap scenario for the Friant-Kern-Canal and Delta diversion scenarios.

Because of the higher diversion amounts possible under the FKC and Delta scenarios, the groundwater storage gain is larger for both scenarios than for the local scenario. However, the increase in surface water used for MAR in both scenarios does not translate into a proportional increase in groundwater storage. Large increases in storage are only achieved for the 200 TAF cap scenarios (Fig. 11). The FKC optimizations show an approximately 25% greater storage gain than the local scenario, while for the Delta optimizations the additional storage gain is 50-80%. However, the increased storage comes at a higher total cost, which is about 10-15% higher than observed in the local scenario. On the low storage gain, low-cost end of the Pareto solutions, the FKC and Delta scenarios are quite similar (e.g., for the same price, the storage gain is of the same order), but the overall cost per unit storage gain (dashed lines) is slightly lower in the FKC and Delta scenarios compared to the local scenario. One exception to this pattern are the 90th percentile runs, which due to the higher amount of diverted water resulted in considerably higher storage gains compared to the 95th percentile scenarios.

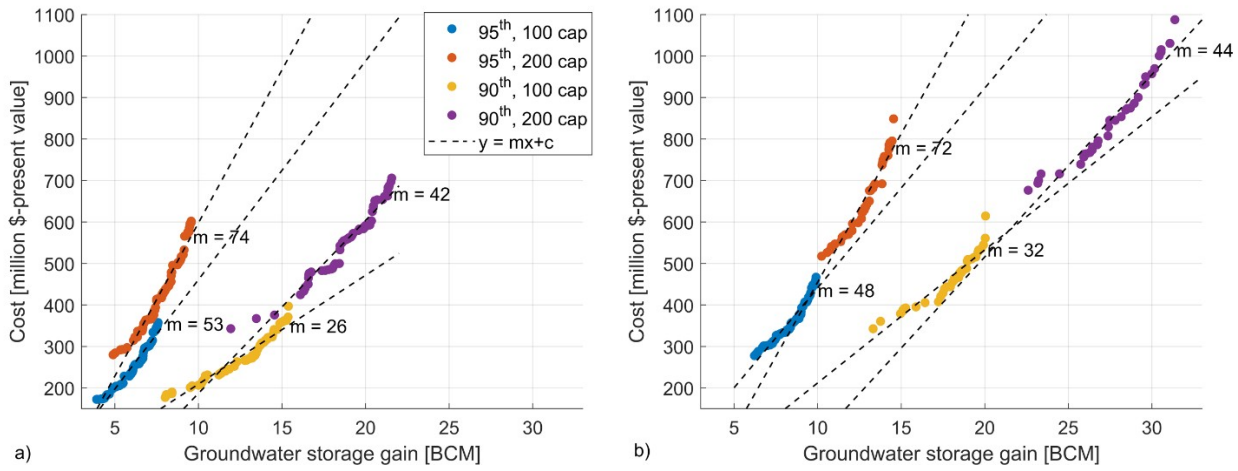
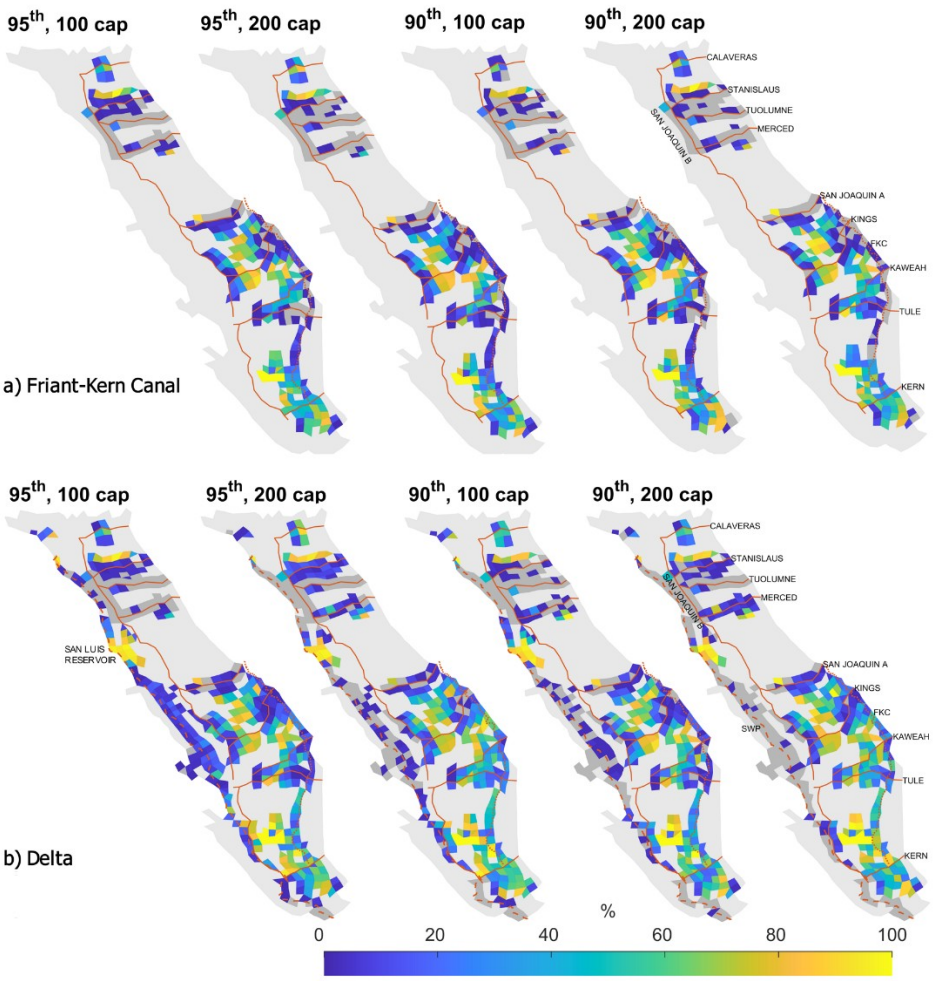


Fig. 11 Pareto fronts for all HMF scenario optimizations for the Friant-Kern Canal scenario (a) and the Delta scenario (b).

The selection probability of individual finite elements in the FKC scenario is similar to the local diversion scenario. The majority of selected elements that can receive water from the FKC are located in the northwest of Kern County (e.g. region north of the Kern River shown in Fig. 12). In particular, there are three elements (Fig. 12a) that have nearly 100% selection probability, which are selected because of their low land and lift cost (east to west topographic gradient), and as they are not very close to streams and can increase storage gain without losing significant amounts of water to baseflow. Another group of elements with high selection probability can be observed north of the Stanislaus River where land prices are lower compared to surrounding areas (element prices in the Stanislaus and Tuolumne River basin are among the highest in the study area; Fig. S3). However, we observe that the selection probability of these elements is higher in the local diversion scenarios than the FKC scenario. The low-cost solutions of the FKC scenario correspond to the mid-cost solutions of the local diversion scenario (Fig. 8a), with the cost difference mainly coming from the increased lift cost in the FKC scenario. Lastly, we observe that elements near the eastern boundary of the study area have very low selection probability in the FKC scenarios. This is partly due to the higher land prices in this area as well as low storage gain capacity that these elements have (Fig. S3).



676

677 *Fig. 12 Probability of element selection for the Friant Kern Canal (a) and Delta diversions (b). Dark grey*
678 *elements correspond to elements available for recharge that have not been selected by any solution.*

679

680 For the Delta diversion scenario, selection of elements is again mainly influenced by land price
681 and energy cost. Elements with the highest selection probability are located near the San Luis
682 Reservoir (Fig. 12b), however, as more water becomes available for MAR, fewer elements are
683 selected by the optimization to receive water from the Delta since more recharge can be
684 accommodated by the Sierra tributaries and the FKC (e.g., Fig. 12a versus Fig. 12b). In Fig. 12,
685 dark grey elements indicate elements available for recharge that have not been selected by any
686 solution. In the 95th percentile – 100 TAF cap Delta scenario at least 25 elements are selected to
687 receive water for MAR, while only 7 elements are selected under the 90th percentile – 200 TAF

cap scenario. This is because the local diversions are always cheaper in terms of energy and conveyance cost.

5 Conclusions

Managed Aquifer Recharge (MAR) is increasingly being recognized as a promising management option to replenish overdrafted aquifers in groundwater-dependent agricultural regions. However, identifying suitable locations for MAR remains a challenge since it often involves a combination of multiple biophysical factors and rarely economic factors (e.g., capital cost, conveyance cost, land cost) that are often perceived as constraints on the selection of a MAR site. We propose the use of a simulation-optimization framework that allows identifying trade-offs between environmental benefits of MAR and the cost of implementing managed aquifer recharge facilities in multiple locations. The proposed simulation-optimization framework utilizes a NSGAII multi-objective genetic algorithm with an integrated surface water-groundwater model capable of evaluating different water balance components and MAR scenarios. As such it provides a solution front (i.e. Pareto front) that allows stakeholders to evaluate tradeoffs between competing criteria such as MAR project cost and environmental benefits.

The SO framework was tested with the California Central Valley Groundwater-Surface Water Simulation Model (C2Vsim) in the San Joaquin–Tulare Lake Basin in California with the goal to identify suitable MAR locations that maximize groundwater storage and minimize MAR cost over a 46-year simulation period (1963-2009). Several diversion scenarios (90^{th} and 95^{th} percentile flows, water sources (local river flow vs imported water), and diversion caps [100 and 200 thousand acre-feet]) were evaluated with the SO framework, which provided some specific insights into the benefits of using the SO framework in hydro-economic modeling.

1. Using the SO framework with a sophisticated numerical groundwater model allowed evaluating a wide variety of possible scenarios stakeholders might want to consider when

discussing the implementation of MAR projects. Stakeholders can choose between MAR sites that maximize the groundwater storage gain, sites that minimize MAR project cost, or any level of compromise in-between.

2. Parallelizing the code allowed simulation of several scenarios with several hundreds of different decision variables and solutions in a matter of a few days.

3. Diverting and recharging excess water from local rivers within the same basin is the most cost-efficient MAR approach in terms of cost per unit of groundwater storage gain due to the low water transportation cost.

4. Diverting more water may not yield the most efficient groundwater storage gain (e.g. cost per unit of groundwater storage gained in the aquifer). Total MAR cost for imported water is higher (10-15%) due to longer distances, elevation changes, and potential seepage losses encountered during transport.

5. Depending on physiographic characteristics (e.g. elevation and distance of recharge area to diversion points) and land prices, some recharge locations are preferably selected over other locations. This is mainly due to the enormous storage gain that these areas can provide.

Currently the model results are based on a hydrologic model that has a relatively low spatial resolution (~10 km^2) and distinguishes only three groundwater layers. It therefore lacks detailed hydrogeologic characterization of the subsurface. Future studies may consider using more refined models such as the fine-mesh C2Vsim model (~ 2.7 km^2) as well as incorporating more detailed subsurface information from airborne electromagnetic surveys currently acquired over California's CV to improve subsurface characterization. Lastly, this study does not take into account groundwater quality, which can be an important factor in MAR design (Guo et al., 2023).

Overall, the results indicate that different combinations of high-magnitude flow diversion thresholds and receiving MAR locations maximize groundwater storage at minimum cost, which provides water resources managers with different options depending on source water availability for recharge. Water resource managers can explore the trade-offs between

different scenarios in a webtool of the results, which is available at http://subsurface.gr/joomla/MAR/OptimResults.html.

Data Availability Statement

The data and results of this paper can be found at https://www.hydroshare.org/resource/f3a3e4e3697e484e8d71852b92ffbc90/

Acknowledgements

This work was supported by the USDA Economic Research Service through a cooperative agreement 58-6000-7-0090 and by National Science Foundation Award #-1716130. The findings and conclusions in this publication are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy.

6 References

Aeschbach-Hertig, W., & Gleeson, T. (2012). Regional strategies for the accelerating global problem of groundwater depletion. In Nature Geoscience (Vol. 5, Issue 12, pp. 853–861). https://doi.org/10.1038/ngeo1617

Ajaz, A., Datta, S., & Stoodley, S. (2020). High plains aquifer-state of affairs of irrigated agriculture and role of irrigation in the sustainability paradigm. In Sustainability (Switzerland) (Vol. 12, Issue 9). https://doi.org/10.3390/su12093714

Al-Maktoumi, A., Zekri, S., El-Rawy, M., Abdalla, O., Al-Abri, R., Triki, C., & Bazargan-Lari, M. R. (2020). Aquifer storage and recovery, and managed aquifer recharge of reclaimed water for management of coastal aquifers. Desalination and Water Treatment, 176, 67–77. https://doi.org/10.5004/dwt.2020.25499

Bachtouli, S., & Comte, J. C. (2019). Regional-Scale Analysis of the Effect of Managed Aquifer Recharge on Saltwater Intrusion in Irrigated Coastal Aquifers: Long-Term Groundwater Observations and Model Simulations in NE Tunisia. Journal of Coastal Research, 35(1), 91–109. https://doi.org/10.2112/JCOASTRES-D-17-00174.1

Bierkens, M. F. P., & Wada, Y. (2019). Non-renewable groundwater use and groundwater depletion: A review. In Environmental Research Letters (Vol. 14, Issue 6). https://doi.org/10.1088/1748-9326/ab1a5f

Brush, C. F., Dogrul, E. C., & Kadir, T. N. (2013). Development and calibration of the California Central Valley Groundwater-Surface Water Simulation Model (C2Vsim), version 3.02-CG. Sacramento, CA: Bay-Delta Office, California Department of Water Resources.

CDFA – California Department of Food and Agriculture. (2020). California Agricultural Statistics Review 2019–2020. State of California.

Cunha, M. Da Conceição (2002). Groundwater cleanup: The optimization perspective (a literature review). Engineering Optimization, 34(6), 689-702.

Dahlke, H.E. and Kocis, T.N. (2018) Streamflow availability rating identifies high magnitude flows for groundwater recharge in the Central Valley. California Agriculture Journal, 72(3): 162-169. https://doi.org/10.3733/ca.2018a0032.

Danapour, M., Fienen, M. N., Højberg, A. L., Jensen, K. H., & Stisen, S. (2021). Multi-Constrained Catchment Scale Optimization of Groundwater Abstraction Using Linear Programming. Groundwater, 59(4), 503–516. https://doi.org/10.1111/gwat.13083

Dangar, S., Asoka, A., & Mishra, V. (2021). Causes and implications of groundwater depletion in India: A review. In Journal of Hydrology (Vol. 596). https://doi.org/10.1016/j.jhydrol.2021.126103

Datta, B., & Kourakos, G. (2015). Preface: Optimization for groundwater characterization and management. Hydrogeology Journal, 23(6), 1043–1049. https://doi.org/10.1007/s10040-015-1297-3

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197. https://doi.org/10.1109/4235.996017

Dillon, P. (2005). Future management of aquifer recharge. Hydrogeology Journal, 13(1), 313–316. https://doi.org/10.1007/s10040-004-0413-6

Dillon, P., Stuyfzand, P., Grischek, T., Lluria, M., Pyne, R. D. G., Jain, R. C., Bear, J., Schwarz, J., Wang, W., Fernandez, E., Stefan, C., Pettenati, M., van der Gun, J., Sprenger, C., Massmann, G., Scanlon, B. R., Xanke, J., Jokela, P., Zheng, Y., … Sapiano, M. (2019). Sixty years of global progress in managed aquifer recharge. Hydrogeology Journal, 27(1), 1–30. https://doi.org/10.1007/s10040-018-1841-z

Dillon, P., Page, D., Vanderzalm, J., Toze, S., Simmons, C., Hose, G., Martin, R., Johnston, K., Higginson, S., & Morris, R. (2020). Lessons from 10 years of experience with Australia's risk-based guidelines for managed aquifer recharge. Water, 12(2). https://doi.org/10.3390/w12020537

Dogrul, E.C. (2013) Integrated Water Flow Model (IWFM v3.02): Theoretical documentation. Sacramento(CA): Central Valley Modeling Unit, Modeling Support Branch, Bay-Delta Office, California Department of Water Resources

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G., & Scanlon, B. R. (2012). Impact of water withdrawals from groundwater and surface water on continental water storage variations. Journal of Geodynamics, 59–60, 143–156. https://doi.org/10.1016/j.jog.2011.05.001

Duffy, W. G., & Kahara, S. N. (2011). Wetland ecosystem services in California's Central Valley and implications for the Wetland Reserve Program. Ecological Applications, 21(3 SUPPL.), S18–S30. https://doi.org/10.1890/09-1338.1

Ebrahim, G. Y., Jonoski, A., Al-Maktoumi, A., Ahmed, M., & Mynett, A. (2016). Simulation-Optimization Approach for Evaluating the Feasibility of Managed Aquifer Recharge in the Samail Lower Catchment, Oman. Journal of Water Resources Planning and Management, 142(2). https://doi.org/10.1061/(asce)wr.1943-5452.0000588

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., Swenson, S. C., De Linage, C. R., & Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. Geophysical Research Letters, 38(3), L03403. https://doi.org/10.1029/2010GL046442

Fatkhutdinov, A., & Stefan, C. (2019). Multi-Objective Optimization of Managed Aquifer Recharge. Groundwater, 57(2), 238–244. https://doi.org/10.1111/gwat.12793

Faunt, C.C., ed., (2009), Groundwater Availability of the Central Valley Aquifer, California: U.S. Geological Survey Professional Paper 1766, 225 p

Faunt, C. C., Sneed, M., Traum, J., & Brandt, J. T. (2016). Water availability and land subsidence in the Central Valley, California, USA. Hydrogeology Journal, 24(3), 675–684. https://doi.org/10.1007/s10040-015-1339-x

Forghani, A., & Peralta, R. C. (2018). Intelligent performance evaluation of aquifer storage and recovery systems in freshwater aquifers. Journal of Hydrology, 563, 599–608. https://doi.org/10.1016/j.jhydrol.2018.06.042

Førsund, F. R., Lovell, C. K., & Schmidt, P. (1980). A survey of frontier production functions and of their relationship to efficiency measurement. Journal of econometrics, 13(1), 5-25.

Gailey, R. M., Fogg, G. E., Lund, J. R., & Medellín-Azuara, J. (2019a). Maximizing on-farm groundwater recharge with surface reservoir releases: a planning approach and case study in California, USA. Hydrogeology Journal, 27(4), 1183–1206. https://doi.org/10.1007/s10040-019-01936-x

Gailey, R. M. (2017). Inactive supply wells as conduits for flow and contaminant migration: conditions of occurrence and suggestions for management. Hydrogeology Journal, 25(7), 2163–2183. https://doi.org/10.1007/s10040-017-1588-y

Gailey, R. M., Lund, J. R., & Medellín-Azuara, J. (2019b). Domestic well reliability: evaluating supply interruptions from groundwater overdraft, estimating costs and managing economic externalities. Hydrogeology Journal, 27(4), 1159–1182. https://doi.org/10.1007/s10040-019-01929-w

Gale, I.N (2005). Strategies for Managed Aquifer Recharge (MAR) in Semi-Arid Areas; United Nations Educational, Scientific and Cultural Organisation (UNESCO): Paris, France, 2005

Ganot, Y., & Dahlke, H. E. (2021). A model for estimating Ag-MAR flooding duration based on crop tolerance, root depth, and soil texture data. Agricultural Water Management, 255. https://doi.org/10.1016/j.agwat.2021.107031

Ghasemizade, M., Asante, K. O., Petersen, C., Kocis, T., Dahlke, H. E., & Harter, T. (2019). An integrated approach toward sustainability via groundwater banking in the southern Central Valley, California. Water Resources Research, 55(4), 2742-2759.

Glover, F. (1989). Tabu Search—Part I. ORSA Journal on Computing, 1(3), 190–206. https://doi.org/10.1287/ijoc.1.3.190

Grönwall, J., & Danert, K. (2020). Regarding groundwater and drinking water access through a human rights lens: Self-Supply as a norm. Water (Switzerland), 12(2). https://doi.org/10.3390/w12020419

Guo, Z., Fogg, G. E., Chen, K., Pauloo, R., & Zheng, C. (2023). Sustainability of regional groundwater quality in response to managed aquifer recharge. Water Resour. Res., 59, e2021WR031459. https://doi.org/10.1029/2021WR031459

GWD [Groundwater Directive] 2006/118/EC: European Community. Directive 2006/118/EC of the European Parliament and of the Council of 12 December 2006 on the protection of groundwater against pollution and deterioration. Off. J. Eur. Commun. 2006, 372, 19–31.

Hernandez, E. A., Uddameri, V., & Arreola, M. A. (2014). A multi-period optimization model for conjunctive surface water-ground water use via aquifer storage and recovery in Corpus Christi, Texas. Environmental Earth Sciences, 71(6), 2589–2604. https://doi.org/10.1007/s12665-013-2900-3

Jha, B. M., & Sinha, S. K. (2010). Towards better management of ground water resources in India. Water and Energy International, 67(1).

Jia, X., O'Connor, D., Hou, D., Jin, Y., Li, G., Zheng, C., Ok, Y. S., Tsang, D. C. W., & Luo, J. (2019). Groundwater depletion and contamination: Spatial distribution of groundwater resources sustainability in China. Science of the Total Environment, 672, 551–562. https://doi.org/10.1016/j.scitotenv.2019.03.457

Kacimov, A. R., Zlotnik, V., Al-Maktoumi, A., & Al-Abri, R. (2016). Modeling of transient water table response to managed aquifer recharge: A lagoon in Muscat, Oman. Environmental Earth Sciences, 75(4). https://doi.org/10.1007/s12665-015-5137-5

Kocis, T. N., & Dahlke, H. E. (2017). Availability of high-magnitude streamflow for groundwater banking in the Central Valley, California. Environmental Research Letters, 12(8). https://doi.org/10.1088/1748-9326/aa7b1b

Kourakos, G., & Mantoglou, A. (2015). An efficient simulation-optimization coupling for management of coastal aquifers. Hydrogeology Journal, 23(6), 1167–1179. https://doi.org/10.1007/s10040-015-1293-7

Kourakos, G., Dahlke, H. E., & Harter, T. (2019). Increasing Groundwater Availability and Seasonal Base Flow Through Agricultural Managed Aquifer Recharge in an Irrigated Basin. Water Resources Research, 55(9), 7464–7492. https://doi.org/10.1029/2018WR024019

Levy, Z. F., Jurgens, B. C., Burow, K. R., Voss, S. A., Faulkner, K. E., Arroyo-Lopez, J. A., & Fram, M. S. (2021). Critical Aquifer Overdraft Accelerates Degradation of Groundwater Quality

in California's Central Valley During Drought. Geophysical Research Letters, 48(17). https://doi.org/10.1029/2021GL094398

Ma, X., Dahlke, H.E., Duncan, R., Doll, D., Martinez, P., Lampinen, B., and A. Volder. (2020). Winter flooding recharges groundwater in almond orchards with limited effects on tree root dynamics, growth and yield. California Agriculture Journal, 76(2): 7. https://doi.org/10.3733/ca.2022a0008

Maliva, R. G. (2015). Managed aquifer recharge: State-of-the-art and opportunities. Water Science and Technology: Water Supply, 15(3), 578–588. https://doi.org/10.2166/ws.2015.009

Maples, S.R., Fogg, G.E. & Maxwell, R.M. (2019). Modeling managed aquifer recharge processes in a highly heterogeneous, semi-confined aquifer system. Hydrogeol J (27), 2869–2888 https://doi.org/10.1007/s10040-019-02033-9

Marchi, A., Dandy, G. C., & Maier, H. R. (2016). Integrated Approach for Optimizing the Design of Aquifer Storage and Recovery Stormwater Harvesting Schemes Accounting for Externalities and Climate Change. Journal of Water Resources Planning and Management, 142(4). https://doi.org/10.1061/(asce)wr.1943-5452.0000628

Marques, G. F., Lund, J. R., & Howitt, R. E. (2010). Modeling Conjunctive Use Operations and Farm Decisions with Two-Stage Stochastic Quadratic Programming. Journal of Water Resources Planning and Management, 136(3), 386–394. https://doi.org/10.1061/(asce)wr.1943-5452.0000045

Marwaha, N., Kourakos, G., Levintal, E., & Dahlke, H. E. (2021). Identifying Agricultural Managed Aquifer Recharge Locations to Benefit Drinking Water Supply in Rural Communities. Water Resources Research, 57(3). https://doi.org/10.1029/2020WR028811

Niswonger, R. G., Morway, E. D., Triana, E., & Huntington, J. L. (2017). Managed aquifer recharge through off-season irrigation in agricultural regions. Water Resources Research, 53(8), 6970–6992. https://doi.org/10.1002/2017WR020458

O'Geen, A., Saal, M., Dahlke, H., Doll, D., Elkins, R., Fulton, A., Fogg, G., Harter, T., Hopmans, J., Ingels, C. & Niederholzer, F. (2015). Soil suitability index identifies potential areas for groundwater banking on agricultural lands. California Agriculture 69, no. 2: 75-84.

Pauloo, R.A., Escriva-Bou, A., Dahlke, H., Fencl, A., Guillon, H. and Fogg, G.E., 2020. Domestic well vulnerability to drought duration and unsustainable groundwater management in California's Central Valley. *Environmental Research Letters*, *15*(4), p.044010.

Pietersen, K., Beekman, H. E., Holland, M., & Adams, S. (2012). Groundwater governance in South Africa: A status assessment. Water SA, 38(3). https://doi.org/10.4314/wsa.v38i3.11

Ross, A., & Hasnain, S. (2018). Factors affecting the cost of managed aquifer recharge (MAR) schemes. Sustainable Water Resources Management, 4, 179-190.

Russo, T. A., Fisher, A. T., Lockwood, B. S., Assessment of managed aquifer recharge potential and impacts using a geographical information system and numerical modeling, Groundwater, doi: 10.1111/gwat.12213, 2014.

Sallwey, J., Bonilla Valverde, J.P., Vásquez López, F., Junghanns, R. and Stefan, C., 2019. Suitability maps for managed aquifer recharge: a review of multi-criteria decision analysis studies. Environmental Reviews, 27(2), pp.138-150.

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Proceedings of the National Academy of Sciences of the United States of America, 109(24), 9320–9325. https://doi.org/10.1073/pnas.1200311109

Scherberg, J., Baker, T., Selker, J. S., & Henry, R. (2014). Design of Managed Aquifer Recharge for Agricultural and Ecological Water Supply Assessed Through Numerical Modeling. Water Resources Management, 28(14), 4971–4984. https://doi.org/10.1007/s11269-014-0780-2

Shen, D. (2015). Groundwater management in China. Water Policy, 17(1), 61–82. https://doi.org/10.2166/wp.2014.135

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation - A global inventory. Hydrology and Earth System Sciences, 14(10), 1863–1880. https://doi.org/10.5194/hess-14-1863-2010

Standen, K., & Monteiro, J. P. (2020). In-channel managed aquifer recharge: A review of current development worldwide and future potential in Europe. Water (Switzerland), 12(11). https://doi.org/10.3390/w12113099

Tran, D. Q., K. F. Kovacs, and S. Wallander. 2019. "Long run optimization of landscape level irrigation through managed aquifer recharge or expanded surface reservoirs." Journal of Hydrology, 579, 124220.

Tran, D. Q., K. F. Kovacs, and S. Wallander. 2020a. "Water conservation with managed aquifer recharge under increased drought risk." Environmental Management, 66(4), 664-682.

Tran, D.Q., Kovacs, K.F. and West, G.H., 2020b. Spatial economic predictions of managed aquifer recharge for an agricultural landscape. Agricultural Water Management, 241, p.106337.

Tzoraki, O., Dokou, Z., Christodoulou, G., Gaganis, P., & Karatzas, G. (2018). Assessing the efficiency of a coastal Managed Aquifer Recharge (MAR) system in Cyprus. Science of the Total Environment, 626, 875–886. https://doi.org/10.1016/j.scitotenv.2018.01.160

Vasco, D.W., Farr, T.G., Jeanne, P., Doughty, C. and Nico, P., 2019. Satellite-based monitoring of groundwater depletion in California's Central Valley. *Scientific reports*, *9*(1), pp.1-14.

Wada, Y., Van Beek, L. P. H., Van Kempen, C. M., Reckman, J. W. T. M., Vasak, S., & Bierkens, M. F. P. (2010). Global depletion of groundwater resources. Geophysical Research Letters, 37(20). https://doi.org/10.1029/2010GL044571

Wang, J., Jiang, Y., Wang, H., Huang, Q., & Deng, H. (2020). Groundwater irrigation and management in northern China: status, trends, and challenges. International Journal of Water Resources Development, 36(4), 670–696. https://doi.org/10.1080/07900627.2019.1584094

Waterhouse, H., Bachand, S., Mountjoy, D., Choperena, J., Bachand, P. A. M., Dahlke, H. E., & Horwath, W. R. (2020). Agricultural managed aquifer recharge - water quality factors to consider. California Agriculture, 74(3), 144–154. https://doi.org/10.3733/CA.2020A0020

WFD (Water Framework Directive), 2000/60/EC. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. OJ L 327, 22.12.2000, pp 1- 51.

Wurl, J., & Imaz-Lamadrid, M. A. (2018). Coupled surface water and groundwater model to design managed aquifer recharge for the valley of Santo Domingo, B.C.S., Mexico. Sustainable Water Resources Management, 4(2), 361–369. https://doi.org/10.1007/s40899-017-0211-7

Xiao, M., Koppa, A., Mekonnen, Z., Pagán, B.R., Zhan, S., Cao, Q., Aierken, A., Lee, H. and Lettenmaier, D.P., 2017. How much groundwater did California's Central Valley lose during the 2012–2016 drought?. *Geophysical Research Letters*, *44*(10), pp.4872-4879.

Yeh, W. W.-G. (2015). Review: Optimization methods for groundwater modeling and management. Hydrogeology Journal, 23(6), 1051–1065. https://doi.org/10.1007/s10040-015-1260-3

Yu, L., Ding, Y., Chen, F., Hou, J., Liu, G., Tang, S., Ling, M., Liu, Y., Yan, Y., & An, N. (2018). Groundwater resources protection and management in China. Water Policy, 20(3), 447–460. https://doi.org/10.2166/wp.2017.035

Zaidi, F. K., Nazzal, Y., Ahmed, I., Naeem, M., & Jafri, M. K. (2015). Identification of potential artificial groundwater recharge zones in Northwestern Saudi Arabia using GIS and Boolean logic. Journal of African Earth Sciences, 111, 156–169. https://doi.org/10.1016/j.jafrearsci.2015.07.008.

Zhang, H., Xu, Y., & Kanyerere, T. (2020). A review of the managed aquifer recharge: Historical development, current situation and perspectives. In Physics and Chemistry of the Earth (Vol. 118). https://doi.org/10.1016/j.pce.2020.102887

Zheng, C., and Wang, P. P. (1999), An integrated global and local optimization approach for remediation system design, Water Resour. Res., 35(1), 137-148, doi:10.1029/1998WR900032

Zlotnik, V. A., Kacimov, A., & Al-Maktoumi, A. (2017). Estimating Groundwater Mounding in Sloping Aquifers for Managed Aquifer Recharge. Groundwater, 55(6), 797–810. https://doi.org/10.1111/gwat.12530