HAL
open science

Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico

G. Schoups, C. Lee Addams, S. M. Gorelick

To cite this version:

G. Schoups, C. Lee Addams, S. M. Gorelick. Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico. Hydrology and Earth System Sciences Discussions, 2005, 2 (5), pp.2061-2109. hal-00298720

HAL Id: hal-00298720

https://hal.science/hal-00298720v1

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hydrol. Earth Sys. Sci. Discuss., 2, 2061–2109, 2005
www.copernicus.org/EGU/hess/hessd/2/2061/
SRef-ID: 1812-2116/hessd/2005-2-2061
European Geosciences Union

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico

G. Schoups[1], C. Lee Addams[1,2], and S. M. Gorelick[1]

[1]Department of Geological and Environmental Sciences, Stanford University, Stanford, California, USA
[2]The Earth Institute at Columbia University, International Research Institute for Climate Prediction, Columbia University, Palisades, New York, USA

Received: 27 July 2005 – Accepted: 13 September 2005 – Published: 26 September 2005

Correspondence to: G. Schoups (gerrit@stanford.edu)

© 2005 Author(s). This work is licensed under a Creative Commons License.

Abstract

Multi-objective optimization was used to calibrate a regional surface water-groundwater model of the Yaqui Valley, a 6800 km^2 irrigated agricultural region located along the Sea of Cortez in Sonora, Mexico. The main advantage of the method is that it accounts for both parameter and model structural uncertainty. In this case, results show that the effect of including the process of bare soil evaporation is significantly greater than the effects of parameter uncertainty. Furthermore, by treating the different objectives independently, a better identification of the model parameters is achieved compared to a single-objective approach, since the various objectives are sensitive to different parameters. The simulated water balance shows that 15–20% of the water that enters the irrigation canals is lost by seepage to groundwater. The main discharge mechanisms in the Valley are crop evapotranspiration (53%), non-agricultural evapotranspiration and bare soil evaporation (19%), surface drainage to the Sea of Cortez (15%), and groundwater pumping (9%). In comparison, groundwater discharge to the estuary was relatively insignificant (less than 1%). The model was further refined by identifying zonal Kv and Kh values based on a spatial analysis of the model residuals. This is the first comprehensive flow model of this important surface water-aquifer system, and it will be used in future work to identify optimal groundwater management strategies.

1. Introduction

Calibration of hydrologic models consists of an iterative process during which the model parameters are adjusted such that the model better mimics the observed dynamics of the system under consideration. Due to the time-consuming nature of this process, automated calibration methods have been developed for finding optimal parameter values that best fit the data. The match between simulated and observed variables is usually quantified by a single objective function, such as the root mean square error (RMSE). The calibration is then cast as an optimization problem in which parameter

values are sought that minimize the objective function. These optimization algorithms typically perform a local search of the parameter space, starting from an initial estimate and leading to a local optimum (e.g. Poeter and Hill, 1998; Doherty, 2000). In addition, global algorithms have been developed that search the entire parameter space and hence result in globally optimal parameter values (e.g. Duan et al., 1992; Vrugt et al., 2003a).

In many applications it is desirable that the model reproduces different types of observations (e.g. water levels and drainage rates). In that situation, calibration of the model to one data type does not guarantee successful simulation of the other data types. Calibration to multiple targets is typically done by assigning weights to the different types of data and combining them into a single objective function for parameter optimization (Hill, 1998). However, the subjective prior selection of weights typically affects the calibrated parameter values. In that case, the single-objective calibration can be repeated using different values for the weights in each optimization run (Madsen, 2003), although this approach is computationally very intensive. Alternatively, a full multi-objective optimization is conducted that identifies the entire optimal set of non-dominated or Pareto solutions within a single optimization run (Gupta et al., 2003). These Pareto solutions explicitly represent trade-offs between the various objectives. The advantage is that no prior weights need to be assigned to the different objectives, since they are treated independently (Schoups et al., 2005). In addition, the multi-objective approach includes more information about the hydrologic system in the parameter identification process, thereby potentially leading to better estimates of the parameters (Boyle et al., 2000). On the other hand, significant trade-offs in fitting two or more objectives may indicate an error in the model structure (Refsgaard and Henriksen, 2004), for example a relevant physical process may not be accounted for or it may be wrongly parameterized.

This paper discusses the application of a multi-objective global optimization approach to the calibration of a regional surface water-groundwater flow model. The study area is the 6800 km^2 Yaqui Valley in the state of Sonora, one of the most important agri-

cultural regions in Mexico. Irrigated agriculture in the Yaqui Valley has since 1942 relied on the supply of water from surface reservoirs. A recent prolonged eight year drought (1996–2004) however has drawn down these reservoir levels below sustainable levels, resulting in severe cuts in water supply and widespread fallowing (Addams, 2004). For the first time in 40 years, due to the effects of the drought on reservoir depletion, wheat was not grown in Yaqui Valley, which is the center for the "Green Revolution" for wheat in Mexico. Due to uncertainties in the future supply of surface water for irrigation, farmers in the Yaqui Valley will depend more and more on groundwater as an additional or even primary source of irrigation water.

The integrated surface water-groundwater model presented here revises the original model of Addams (2004) and serves as a first step to developing a comprehensive water management plan for the region. Unlike previous hydrogeologic research in the Yaqui Valley (Diaz, 1995; Islas, 1998; Steinich and Chavarria, 2000), the flow model presented here incorporates all known spatially distributed stresses on the system, including pumping, drainage, irrigation canal seepage, and field irrigation losses (Addams, 2004). Using hydraulic heads, canal seepage rates, and drainage volumes, a multi-objective calibration problem is formulated and solved using the recently developed Multi-Objective Shuffled Complex Evolution Metropolis (MOSCEM-UA) global optimization algorithm of Vrugt et al. (2003b).

The paper is organized as follows. First, the study area, hydrology, and hydrogeology of the Yaqui Valley are briefly described. The conceptual model for groundwater flow is introduced, and the methods used for its initial parameterization are discussed. This is followed by an introduction to the calibration algorithm and the available measurements used to calibrate the model. Next, the results of applying the multi-objective optimization algorithm are discussed in detail, highlighting the advantages of this approach compared to more traditional methods and summarizing insights into the regional flow system of the Yaqui Valley.

2. Methods

2.1. Study area: water resources and hydrogeology

Figure 1 shows the location of the Yaqui Valley, which lies between the coastal plain of the Sea of Cortez to the southwest and the Sierra Madre Mountains to the northeast. The climate is semi-arid with an average annual precipitation of ~300 mm, most of it falling in the summer from June to September. Annual potential evapotranspiration averages 2000 mm. Most of the farmland in the Valley is part of the Yaqui Irrigation District (Fig. 1). The dominant crop is winter wheat, which is grown from November to April, and is irrigated using a combination of surface water and groundwater. The surface water system consists of three reservoirs in series on the Yaqui River, the largest and furthest downstream being the Oviachic reservoir (Fig. 1). Surface water releases from Oviachic reservoir are conveyed to the Yaqui Irrigation District by means of open unlined canals. Along the way, water is diverted by the various agricultural water management units, known as modules, that make up the Irrigation District. Water is further distributed to individual fields within each module by means of a network of secondary irrigation canals. Almost 600 wells have also been installed (Fig. 1) to provide additional water for crop production, although these are never all active in the same year. Some of the wells are privately operated whereas others are managed by the District. Throughout the Irrigation District, a drainage network (Fig. 4) has been installed to drain surplus irrigation water from fields out to the Sea of Cortez. These drains are primarily open drainage ditches, with a small percentage of subsurface drainage pipes, at a depth of 1 to 2 m below the land surface. Most of the soils in the valley are clayey vertisols with organic matter contents less than 1% (Lobell et al., 2002).

The Yaqui Valley coastal aquifer system is primarily composed of alluvial fill consisting of Quaternary deposits that include consolidated and unconsolidated gravels, sands, clays, and evaporate deposits (Gonzales and Marin, 2000). The groundwater system can be conceptually divided into three layers, each of which has a particular hydraulic function (Fig. 2). The first 10–20 m below the surface constitute the shallow

aquifer, which receives irrigation-related recharge from canal seepage and field-level irrigation losses. Part of the irrigation return and water from the seasonally rising local water table is intercepted by the network of 1 to 2 m deep agricultural drains. Another part is removed by the native vegetation around the canals (phreatophytes) or it evaporates from the shallow water table. The remaining water either flows laterally toward the coast through the 10 to 2 m thick water table aquifer, where it discharges beneath the estuary, or it flows downward. Vertical flow is governed by the underlying discontinuous confining layer with a thickness ranging from 5 to 80 m. In some locations it consists of clay and in other regions it is rich in sand and gravel, so that overall it exhibits a high degree of variability in texture and hydraulic conductivity. Finally, the deep aquifer represents the zone screened by most production wells in the region. Generally, it lies 30 to 100 m below the surface, ranging in thickness from 30 to 170 m. The aquifer consists of Quaternary alluvium and underlying older unconsolidated Tertiary deposits. Depending on the magnitude of contrast in hydraulic conductivity with the overlying confining layer, the deep aquifer behaves either as a confined, semi-confined, or unconfined system. Addams (2004) provides more details on how layer geometry was determined for modeling purposes.

An estimate of the available storage in the deep aquifer under the Irrigation District is made by summing the available storage under confined conditions (above the top of the screened wells), assuming a specific storage of $10^{-4}\,\mathrm{m}^{-1}$, plus the remaining storage under unconfined conditions, assuming a specific yield of 0.2 and counting 2/3 of the aquifer thickness. This results in a value of approximately 100 000 MCM (Million Cubic Meters), which is about 16 times the available storage in the Yaqui reservoir system (6000 MCM). In addition, the supply of surface water is variable and uncertain as was evident during the recent drought. Hence, available water in the deep aquifer will likely play a central role in a sustainable water management plan in the Yaqui Valley. Historically, groundwater use has been limited due to the availability of cheap surface water and the relatively high cost of pumping.

2.2. Hydrologic model

In this section, we discuss the concepts and methods used to simulate flow in the integrated surface water-groundwater system of the Yaqui Valley. For further details we refer to Addams (2004). Special attention is paid to representation of the near-surface hydrologic processes, such as recharge, evaporation, and near-surface drainage, and to the coupling between the surface water and groundwater systems. To reduce uncertainty, we independently estimate as many parameters as possible (e.g. recharge), and then find optimal values for the remaining parameters.

The groundwater component of the flow model was represented using the transient 3D groundwater flow equation,

$$\frac{\partial}{\partial x}\left(K_x\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial h}{\partial z}\right)-W=S\frac{\partial h}{\partial t}, \quad (1)$$

where K_x, K_y, and K_z are hydraulic conductivity values in the x, y, and z direction (L/T), h is hydraulic head (L), t is time (T), W is a source-sink term (1/T) representing recharge, pumping, evaporation, or drainage, and S is a storage parameter (1/L) related to either specific storage S_s (1/L) when confined, or specific yield S_y (–) when unconfined. Groundwater flow was simulated using Modflow-2000 (Harbaugh et al., 2000). The surface water component was simulated using a routing model for flow in the main irrigation canals, including a sink-source term representing water exchange between the canals and the groundwater system (Prudic et al., 2004).

2.2.1. Spatial and temporal discretization

The lateral boundaries of the model domain are defined by the contact between the alluvial deposits and bedrock in the northwest, the Sea of Cortez near the shoreline in the southeast, and nearly stationary groundwater flow lines in the northeast and southwest. The model domain was discretized into a regular finite difference grid of three layers, each containing 60 rows and 70 columns, resulting in 12 600 cells with an

area of $2 \times 2\,km^2$ each. This level of horizontal discretization is deemed sufficient, since the interest is in characterizing the regional flow system rather than accounting for flow near individual wells. Cell thicknesses range from 5 to 170 m. The grid was oriented at a 25° azimuth angle to align it with the principal direction of groundwater flow from northwest to southeast. There were 5064 active cells within the model boundaries shown in Fig. 1.

In the vertical direction, the three layers represented the shallow aquifer, confining layer, and deep aquifer as discussed earlier. The top of the first layer coincides with the land surface elevation, which was determined from a Digital Elevation Model (DEM) and a paper elevation map obtained from the National Water Commission (CNA) in Mexico. The bottom of the first layer, which is also the top of the second layer, was estimated by interpolating elevations from selected well logs (Fig. 3) each of which exhibits a change from coarse-grained to fine-grained material. Similarly, the top of the third layer corresponds to a transition from fine-grained to coarse-grained material observed in well logs. This transition also often coincided with the top of the well screen. Finally, the thickness of the third layer was defined as twice the well screen lengths, giving thicknesses of approximately 30 to 170 m. As such, the third layer represents the approximate aquifer thickness of the productive aquifer material, rather than the total permeable thickness, which extends deeper as evidenced by geophysical measurements (Steinich and Chavarria, 2000).

The calibration period extends over 24 years, starting in October 1973 and ending in September 1997, corresponding to crop years 1974–1997. This period was selected for calibration because it covers a long record of observed aquifer head data, including periods of increased groundwater pumping, as well as a period of measured rates of agricultural drainage (1988–1997). Since our primary interest is in the long-term response of the groundwater system, the entire simulation period was discretized into 24 annual stress periods for which average annual boundary conditions were specified. Each stress period was divided into 10 time steps, using a fixed time step multiplier equal to 1.2. This degree of temporal discretization was tested to determine that it

was sufficiently accurate and yet led to reasonable computational times for the 3000 calibration simulations.

2.2.2. Initial conditions

Initial values for hydraulic heads were estimated by running the model at steady-state using time-averaged boundary conditions for the period 1972–1974 when data show that heads in the deep aquifer, pumping rates, and reservoir releases were all relatively constant. The head distribution from the steady-state model was then used as initial condition for the transient run from 1974 to 1997.

2.2.3. Irrigation, crop *ET*, and recharge

Recharge from field-scale irrigations is independently estimated using a mass-balance calculated each year and for each module in the Irrigation District,

$$R = SW + GW - ET_c, \tag{2}$$

where R is annual recharge (L/T), SW is surface water delivery to the module (includes conveyance losses during transport of water from the main irrigation canals to the fields) (L/T), GW is irrigation using groundwater (L/T), and ET_c is annual crop ET (L/T). Annual rates of surface-water irrigation delivered to the entire Yaqui Irrigation District were distributed over the modules assuming that the spatial distribution of surface-water allocation was the same as that during the period 1996–1999 when module specific data were available. Annual groundwater pumping rates were available for each well throughout the entire calibration period. Total groundwater irrigation within each module was estimated by calculating total pumping from all wells located within that module, excluding those wells that discharge directly into the main irrigation canals. Finally, district-wide crop ET was estimated from annual crop acreage data and local crop water demand values (Addams, 2004). The main crop is winter wheat (50% of planted acreage), but a variety of other crops are grown as well, including soybeans,

maize, cotton, safflower, and various vegetable crops. Estimates of annual consumptive use for these crops were obtained from local sources (CNA; Ortiz-Monasterio, pers. comm., 2001). The total crop *ET* was then distributed over the modules assuming the same spatial distribution as observed for the period 1996–1999. Note that the recharge values calculated with Eq. (2) include recharge from irrigation water applied at the field-scale, as well as conveyance losses through the secondary irrigation canals which distribute water from discharge points on the main irrigation canals to the fields. Average annual volumes of surface-water use, groundwater use, and crop water demand (consumptive use) for the period 1974–1997 were 2065, 271, and 1512 MCM, respectively. For further details of the spatial treatment of irrigation-related recharge and crop *ET*, see Addams (2004).

2.2.4. Canal seepage

Since the irrigation canals are unlined, significant amounts of water (up to 500–600 MCM annually) are lost by seepage to groundwater before canal water reaches the end of the canals. Canal volumetric seepage rates are simulated by discretizing the main canals into 145 reaches with inflow of surface water from the Oviachic reservoir (Fig. 1) specified annually at the head of each canal. Water is then routed through the canals using a water budget for each reach,

$$Q_{out} = Q_{in} + Q_{gw} - Q_{sw} - Q_{seep}, \tag{3}$$

where Q_{out} is outflow from the reach (L^3/T), Q_{in} is inflow into the reach (L^3/T), Q_{gw} represents pumping of groundwater into the canal (L^3/T), Q_{sw} is diversion of surface water for irrigation (L^3/T), and Q_{seep} is canal seepage into the aquifer (L^3/T). Direct evaporation from the canals was negligible and not considered. Additions to each canal reach by groundwater pumping, Q_{gw}, are estimated from annual CNA records of well-specific pumping, whereas water diversions to the agricultural modules are calculated from annual surface water use in each module, *SW* in Eq. (2), and the location of its

diversion points along the canals. The canal-groundwater interaction term, Q_{seep}, is calculated using Darcy's law,

$$Q_{seep} = \frac{K_s w_s L_s}{m_s}(h_s - h) \text{ for } h \geq h_{bot} \tag{4a}$$

v

$$Q_{seep} = \frac{K_s w_s L_s}{m_s}(h_s - h_{bot}) \quad \text{for} \quad h < h_{bot}, \tag{4b}$$

where K_s is hydraulic conductivity of the canal bed (L/T), w_s is canal width (L), L_s is canal length (L), m_s is thickness of the canal bed (L), h_s is head in the canal reach (L), h_{bot} is elevation of the canal bed bottom (L), and h is head in the underlying water table aquifer (L). For each canal reach, values for w_s and L_s were estimated from District records, and m_s was set at a uniform value of 0.1 m. The elevation of the canal bottom was defined relative to the land surface elevation. However, due to the coarse spatial resolution of the groundwater model (grid cells of 2 by 2 km), the elevation of the canal bottom was specified higher than the average land surface elevation of the grid cell in areas of steeper topography near the mountain front (Fig. 1). Uniform initial values for K_s equal to 0.014 and 0.01 m/yr were estimated from data on canal seepage rates and estimated water depths in the canals, whereas the lined yet leaky parts of the canal were assigned an initial K_s of 0.001 m/yr. Due to the uncertainty on the values of K_s, the initial estimates were multiplied by a scaling factor, f_{Ks} (–), which was subject to calibration (Table 1). Heads, h_s, or water depths, $d_s = h_s - h_{bot} - m_s$, in the canals are calculated using the following approximation to Manning's equation for a 45° trapezoidal cross-section,

$$d_s = b\sqrt{Q_s}, \tag{5}$$

where d_s is water depth (L), Q_s is flow at the midpoint of the canal reach (L^3/T), and b is a coefficient ($(L/T)^{-0.5}$) estimated as a function of roughness n, slope S, and width

w_s of each canal reach (Addams, 2004). The canal-aquifer system represented by Eqs. (1), (3), and (5) interacts through the seepage term in Eq. (4), requiring an iterative solution at each time step. The system was solved using the stream package of Prudic et al. (2004) and Modflow (Harbaugh et al., 2000).

2.2.5. Recharge outside the Irrigation District

Outside the Irrigation District there are three additional sources of recharge to groundwater: recharge from irrigation in the Yaqui Colonies, Yaqui River infiltration, and mountain-front recharge. The Yaqui Colonies, tribal lands located to the north of the Irrigation District, have a perpetual right to approximately 250 MCM/yr of surface water, which was estimated to result in a constant uniform recharge rate of 0.37 m/yr based on the assumption of recharge mechanisms similar to the District. Yaqui River infiltration upstream of the irrigation canals (Fig. 1) was calculated annually from tabulated Irrigation District data. Except in extremely wet conditions, no water flows in the Yaqui River past the intake of the irrigation canals (Addams, 2004). Precipitation is only a significant source of recharge near the mountain-front boundaries (Fig. 1). This mountain-front recharge was estimated as (Anderson et al., 1992),

$$\log_{10}(Q_{mfr}) = -1.34 + \log_{10}(P), \tag{6}$$

where Q_{mfr} is the volume of mountain-front recharge (MCM/yr), and P is the volume of precipitation on the upland watershed (MCM/yr). Although there is considerable interannual variability in precipitation, the resulting recharge is dampened by slow unsaturated zone percolation (Flint et al., 2000). Therefore, long-term average precipitation at the Oviachic Reservoir (0.2 m/year) was used for P, resulting in constant recharge rates of 0.03 m/year and 0.04 m/year for the Bacatete and Baroyeca mountain fronts, respectively. Direct precipitation in the valley was neglected as a source of recharge in this semi-arid climate with summer temperatures that reach 40°C.

2.2.6. Non-agricultural *ET* and bare soil evaporation

In addition to crop evapotranspiration, one also needs to account for evaporation from native and riparian vegetation along the irrigation canals and near the coast, as well as bare soil evaporation after of the crop growing season. Annual potential evaporation averages 2 m/yr, approximately 60% or 1.2 m of which occurs from May to October, i.e. after the main wheat growing season. Actual evaporation rates are limited by the availability of moisture near the soil surface, which mainly depends on the soil hydraulic properties and the water table depth (Gardner, 1958; Gardner and Fireman, 1958). Hypothetical analysis of evaporation for a clay soil, which is the main soil type in the area, as a function of water table depth was estimated using Hydrus (Šimůnek et al., 1998). Results suggest an approximate linear relation between relative evaporation and water table depth, as shown in Fig. 5. Maximum evaporation rates occur when the water table is within 0.2 m of the land surface and no evaporation takes place when the water table approaches 2 m depth. These values were used to calculate evaporation as a function of water table depth. Maximum evaporation rate was set at 2 m/yr outside the Irrigation District. Inside the District consumptive use by crops during the growing season is already accounted for in Eq. (2), hence the potential evaporation rate was set at 1.2 m/yr. Although this neglects any additional evaporation that may occur during the growing season, the simulations indicated that the upper limit of 1.2 m/yr was typically not reached within the District. A linear approximation was applied by fitting to the "data" points from Hydrus and then calculating the relative evaporation as a function of water table depth (Fig. 5).

2.2.7. Water table drainage through drainage network

Throughout the irrigation district, a drainage network has been installed to keep the water table from rising to the ground surface. The network drains (Fig. 4) surplus irrigation water from fields out to the Sea of Cortez. These drains consist partly of open drainage ditches, and partly of subsurface drainage pipes. In the model, groundwater

flow to drains is simulated as a linear head-dependent sink,

$$Q_d = L_d C_d \left[h - h_d\right] \tag{7}$$

where Q_d is drain flow (L^3/T), L_d is total drain length (L), C_d is drain conductance (L/T), h_d is drain elevation (L), and h is hydraulic head in the grid cell (L). The total drain length, L_d, in each grid cell was determined by overlaying the drainage network map onto the model grid. Drain elevations, h_d, for each drain segment were determined from land surface elevations and assuming a uniform drain depth of 2 m. Finally, drain conductance values, C_d, were estimated using monthly observed drainage volumes and water table elevations from January 1996 to September 1997 for 13 sub-areas. These initial C_d values were multiplied by a uniform scaling factor, f_{Cd} (–), which was subject to calibration (Table 1).

2.2.8. Groundwater pumping

Annual well pumping data were gathered from CNA sources for a total of 591 wells (Fig. 1), although only a part of these are active in any given year. All pumping was assigned to the main aquifer, i.e. layer 3.

2.2.9. Boundary conditions

The bottom boundary of the model domain approximately corresponds to the deepest extent below which groundwater flow in the deep aquifer is not influenced by agricultural pumping. Hence, a zero-gradient boundary condition was used at the bottom boundary, which prevents any water from leaving or entering the model through the bottom.

At the northwestern boundary, a flow line extends from the Bacatete mountain range to the Sea of Cortez, hence a zero-gradient boundary condition is assumed there. The southeastern boundary also coincides with a flow line. Historically, these flow lines have been fairly stable and insensitive to pumping within the Irrigation District. In addition, geophysical evidence suggests that there is high volcanic bedrock in the subsur-

face between the Yaqui and Mayo River Valley to the southeast, which further justifies adopting a no-flow condition in this location (ITC, 1979). The northeast boundary of the model corresponds to the bedrock-alluvium interface, hence no flow is assumed to enter or leave the model there as well. Finally, the southwest offshore boundary parallels the shoreline and is simulated as a constant head boundary, with heads set equal to zero (sea level) in the uppermost layer, allowing water to leave or enter the domain through that layer. However, no-flow boundaries are assumed for the second and third layers along the offshore boundary to simulate artesian conditions observed in an offshore island well (Fig. 3). Because discharge of groundwater occurs beneath the Sea of Cortez, it was important that the model reproduce the head value in this offshore island well because it constrains the degree of submarine confinement of the aquifer. Vertical hydraulic conductivity in other locations along the coastline was high enough to enable deep groundwater discharge into the Sea of Cortez through the constant-head boundary of the first layer.

2.2.10. Hydraulic properties

The spatial distribution of horizontal (K_h) and vertical (K_v) hydraulic conductivities was estimated from lithologic data and well tests. It was assumed that the aquifer is horizontally isotropic in each layer but vertically anisotropic, hence $K_x = K_y = K_h$, and $K_z = K_v$. For the deep aquifer, specific capacity values were calculated for 41 wells with pumping rates, static head levels, and dynamic pumping head levels recorded at the time of drilling (Fig. 3). These specific capacity values were converted into transmissivity values using the following empirical relationship (Razack and Huntley, 1991),

$$T = 15.3\, S_c^{0.67} \tag{8}$$

where T is transmissivity and S_c is specific capacity, both in units of m^2/day. Since the permeable formation extends for some distance beneath the bottom of the well screen,

corresponding values for K_h were estimated as,

$$K_{h3} = \mathrm{T}/2b \tag{9}$$

where b is the well screen length (L). The resulting K_{h3} values were interpolated to the model grid. Finally, initial values of vertical hydraulic conductivity K_v for were estimated to be an order of magnitude smaller than the corresponding K_{h3} values, i.e. $K_{v3}=K_{h3}/10$.

Since no well tests were available for the shallow layers, initial conductivity values for them were estimated using well log data from 55 wells (Fig. 3). The lithological categories observed in the well logs were aggregated into a low permeable class (clay and caliche) and a high permeable class (sand and gravel). At each well location, the total and fractional thickness of these materials was calculated in the upper two layers, respectively. These thicknesses were then spatially interpolated within each layer to provide an initial estimate of the spatial distribution of hydraulic conductivity. The spatial distribution of K_h in layer 1 was estimated as,

$$K_{h1} = 10^{1-x} \tag{10}$$

where x is the interpolated clay fractional thickness (0–1), resulting in values for K_{h1} between 1 and 10 m/day. A hydrogeologically reasonable estimate of 10:1 was adopted for the uncalibrated vertical anisotropy ratio in layer 1, or $K_{v1}=K_{h1}/10$. Finally, for layer 2 the vertical hydraulic conductivity was first estimated as,

$$K_{v2} = 10^{-ay} \tag{11}$$

where y is the interpolated clay thickness (0–50 m). The value of a was set to 0.1 in order to reproduce artesian conditions observed in the offshore well in Fig. 3. This results in initial estimates for K_{v2} between 10^{-5} and 1 m/day. Horizontal hydraulic conductivity in the layer was estimated using a 10:1 anisotropy ratio, hence $K_{h2}=10K_{v2}$.

Given the uncertainty of these initial estimates, the K_v and K_h values in each layer were scaled by factors, f_{Kvi} and f_{Khi}, where i is the layer index. These six scaling

factors were subject to calibration (Table 1). Finally, spatially uniform values for specific yield, S_y (–), and specific storage, S_s (1/L), were also estimated by calibration.

2.3. Parameter optimization

2.3.1. Calibration parameters and targets

A total of 10 parameters were subject to calibration, as listed in Table 1. For each parameter a reasonable physical prior range of values was specified centered around the initial estimates discussed in the previous section. The model is calibrated using data on water table elevations, aquifer heads, drainage volumes, and canal seepage volumes for the period 1974–1997. Figure 6 shows the well locations with water table and aquifer head observations during this period. Note however that every well doesn't necessarily have a measurement each year. Water table measurements in September are compared to simulated values at the end of each annual stress period. Aquifer heads are measured annually in October, when wells were turned off for 1-2 days prior to measurement and when almost no pumping occurs since the main growing season is from November to April. These measurements are compared to simulated heads at the end of each annual stress period. Since annually averaged groundwater pumping is used in the model, this implicitly assumes that the head at the end of the water year is not affected by intra-annual pumping changes. The third data set used for calibration consists of annual drainage rates in the agricultural drainage network (Fig. 4) during 1988–1997 measured at various discharge locations near the Sea of Cortez. Finally, the Irrigation District also maintained records of the total annual volume of surface water lost by seepage from the main canals during 1974–1997. Those values were compared to the corresponding simulated volumes.

Since there are four different types of data to calibrate to, we can formulate four

independent root mean square error objective functions,

$$RMSE_k = \sqrt{\sum_{i=1}^{n_k} \left[\frac{\psi_{k,OBS}(i) - \psi_{k,SIM}(i)}{n_k} \right]^2} \tag{12}$$

where k indicates the type of measurement (water table "*wt*", aquifer head "*aq*", drainage volume "*drain*", and canal seepage volume "*seep*"), $RMSE_k$ is the root mean square error for data type k, n_k is the number of measurements for data type k (891, 3506, 130, and 48 respectively for "*wt*", "*aq*", "*drain*", and "*seep*"), $\psi_{k,OBS}(i)$ is the ith observation of data type k, and $\psi_{k,SIM}(i)$ is the corresponding simulated value. Ideally, we would like to identify an optimal parameter set that minimizes all four objective functions simultaneously. However, this is typically not possible due to errors in the conceptual model. Therefore, the trade-offs between matching the different objectives were investigated by performing a multi-objective calibration and evaluating the tradeoff relations in the Pareto optimal surfaces.

2.3.2. Calibration algorithm

The goal of the multi-objective optimization is to minimize $\boldsymbol{F}(\boldsymbol{p}) = \langle f_1(\boldsymbol{p}), f_2(\boldsymbol{p}), \ldots, f_N(\boldsymbol{p}) \rangle$ with respect to $\boldsymbol{p}$, where $\boldsymbol{F}$ is the vector of objectives, $f_k(\boldsymbol{p})$ is the kth objective function in Eq. (1), and $\boldsymbol{p}$ is a vector of model parameters (Gupta et al., 2003). The solution to this problem will not be a single "best" parameter set, but will consist of a Pareto optimal set of solutions corresponding to trade-offs among the objectives. Formally, the Pareto set consists of parameter combinations p_i with the following properties: (1) for all non-members p_n there exists at least one member p_i that dominates p_n, and (2) it is not possible to find another member p_j within the Pareto set that dominates p_i. By definition, p_i dominates p_j if, for all k, $f_k(p_i) < f_k(p_j)$. Our four-objective optimization problem was solved using the Multi-Objective Shuffled Complex Evolution Metropolis algorithm, MOSCEM-UA (Vrugt et al., 2003b). Within a single optimization run, this algorithm generates a Pareto set of parameter combinations that best fit the data according to

multiple objective functions. The algorithm starts by randomly generating s parameter sets from uniform probability distributions defined by the a priori specified parameter ranges. For each parameter set, the model is run and the values of the four objective functions are calculated. Based on the performance on these objectives, each parameter set is characterized by a single point in the multi-objective function space. Using the concept of Pareto dominance defined earlier, the initial population of s parameter sets is divided into dominated and non-dominated points. The non-dominated parameter sets i are assigned a fitness value r_i,

$$r_i = \frac{n_i}{s} \tag{13}$$

where n_i is the number of parameter sets dominated by parameter set i. The dominated parameter sets j on the other hand are assigned a fitness r_j,

$$r_j = 1 + \sum_{i \leq j} r_i \tag{14}$$

where the summation is over all parameter sets i that dominate parameter set j. Parameter sets with low fitness values are retained, resulting in a preference of non-dominated parameter sets at the extremes of the Pareto front. This strategy prevents convergence in the compromise region of the objectives. The algorithm proceeds by dividing the initial parameter population s into a number of so-called complexes. Within each complex new parameter sets are generated by sampling from a multi-normal distribution estimated from all points in the complex, and new points are accepted or rejected based on their fitness value. After a prescribed number of iterations, all parameter sets are shuffled and new complexes are formed. Repeated application of these steps causes the population to converge to the Pareto set of solutions. The Pareto parameter set also contains the single objective solutions at the extremes of the Pareto solution set (end-members). Therefore, using one optimization run the MOSCEM-UA algorithm generates all information needed to evaluate best parameter values for each data type (see Vrugt et al., 2003b).

Here the MOSCEM-UA algorithm was run with an initial population s of 200 parameter sets (sets of 10 parameters), randomly selected from the prior ranges defined in Table 1. This initial sample set was then divided into 5 complexes for subsequent optimization for a total of 3000 model simulations or evaluations.

3. Results and discussion

The results section is organized as follows. First, all four calibration targets (water table elevations, aquifer heads, drainage rates, and canal seepage rates) are considered simultaneously using the multi-objective analysis algorithm. This analysis reveals trade-offs between the various objectives and provides insight into parameter sensitivities. Second, a single "best" parameter set is selected for which more detailed comparisons are shown between simulated and observed variables. Third, results from the multi-objective analysis in combination with a spatial analysis of model residuals are used to improve the model performance on aquifer heads. A second calibration using the refined model is then performed, focusing on aquifer heads alone.

3.1. Four-objective calibration

Figure 6 presents bi-criterion plots for the four-objective calibration after 3000 model evaluations with the MOSCEM-UA algorithm. Each plot is a marginal multi-objective curve showing two criteria out of four total criteria dimensions. Each dot represents one forward simulation of the groundwater model. The large cross indicates the simulation that yields the minimum distance in the four-dimensional normalized objective function space. Two observations are made based on these plots. First, from the first three plots in Fig. 6 it can be seen that there is very little variation in the RMSE_{wt} for the range of parameter values considered here (Table 1). This indicates that the simulation of water table elevations is quite insensitive to any of the parameters. This may be explained by the fact that the shallow water tables in the model are controlled by drainage in the en-

gineered agricultural drains and evaporation, which tend to control and suppress large variations in simulated water table elevations. Second, the remaining bi-criterion plots that do not involve $RMSE_{wt}$ (bottom three plots in Fig. 6), all exhibit trade-offs along a right angle. This indicates that improvements in fitting one objective can be made without deteriorating the fit to the other objective. In other words, these results suggest that there is little trade-off between fitting the different objectives, and that they can be considered independently. The absence of any significant trade-offs indicates that the model is generally well conceptualized, and that most relevant hydrologic processes are accounted for.

At this point it is useful to consider the effect of bare-soil evaporation on the results of the multi-objective optimization, particularly since the initial groundwater model conceptualization did not include bare soil evaporation as a significant physical process (Addams, 2004). Figure 7 shows the same plots as Fig. 6, but now bare soil evaporation was omitted. The main differences are that (1) without evaporation there is much more scatter in the plots involving $RMSE_{wt}$ (top 3 plots), and (2) there is a strong trade-off between the drainage and canal seepage objectives. In other words, no single parameter combination exists that yields a correct simulation of both drainage and canal seepage, suggesting that a key physical process is missing in the conceptualization of the shallow groundwater system. Correct simulation of canal seepage results in too much water discharging through the surface drains, hence pointing to an additional sink that needs to be included (evaporation). On the other hand, drainage can only be correctly simulated by reducing canal seepage (e.g. by decreasing the canal bed conductivity). It was not possible to discharge the extra water as subsurface flow to the Sea of Cortez within the realistic range of hydraulic conductivities. None of these trade-offs are present in the plots that include bare soil evaporation (Fig. 6). This illustrates that multi-objective optimization is a useful method for simultaneously dealing with parameter and model structural uncertainty.

In the following paragraphs we investigate how well the model parameters are identified by the four objective functions. Figure 8 presents so-called dotty plots (Beven

and Freer, 2001) for each of the four objectives (a–d) as a function of each of the 10 calibration parameters. Note that all parameter scales on the horizontal axes are log-transformed, except for S_y. Starting with the water table elevations, Figure 8a indicates that RMSE_{wt} is relatively more sensitive to the horizontal hydraulic conductivity of the deep aquifer, K_{h3}, and the conductivity of the canal bottom, K_s, compared to the other parameters. However, as discussed above, RMSE_{wt} is not sensitive to any of the parameters as evidenced by its small variation (around 2.5–3 m). Secondly, Fig. 8b shows that the simulated deep aquifer heads are most sensitive to the deep aquifer hydraulic conductivity, K_{h3}, and less obviously the vertical hydraulic conductivity of the confining layer, K_{v2}. Best performance on aquifer heads, as indicated by low values for RMSE_{aq}, are obtained for f_{Kh3} values greater than 0. Hence, the multi-objective analysis shows that increases in K_{h3} result in a better simulation of the aquifer heads compared to using the initially estimated values for this parameter (which corresponds to f_{Kh3}=0). The minimum value of RMSE_{aq} identified by the MOSCEM-UA algorithm is 7.2 m (Fig. 8, Table 3). Figure 8b suggests that improvements in the simulation of aquifer heads may be possible by re-examining the K_{v2} and K_{h3} parameterizations. This will be investigated in the next section. With regard to the simulation of drainage rates, Figure 8c clearly demonstrates that RMSE_{drain} mainly depends on the value of the drain conductances (f_{Cd}). On the other hand, RMSE_{seep} is primarily sensitive to the canal bed hydraulic conductivity, K_s (Fig. 8d). Overall, the results in Fig. 8 suggest that the different objectives are sensitive to different parameters, implying that each objective could be fitted independently by finding the optimum value(s) for its most sensitive parameter(s). This confirms the conclusions that were drawn from Fig. 6 about the independence of the objective functions. It also means that the interaction between the surface water and groundwater systems is limited. We further conclude that improvements in the simulation of heads in the deep aquifer are most likely by reconsidering the K_{v2} and K_{h3} parameterizations, and this could be done without significantly affecting the other three objectives.

It is important to realize that these findings most likely would have been obscured

in a single-objective analysis, as the different data types would have been lumped into a single objective function, thereby blurring the sensitivity of the parameters to certain data types. Furthermore, the sensitivity and identifiability of the parameters would then essentially depend on the weights assigned to the different data types. Therefore, the independence of the different objectives in the multi-objective analysis increases the information content of the data and results in better parameter estimates.

What the dotty plots do not show is possible correlation between the parameters, which if present could cause poor parameter sensitivity and identifiability. Table 2 shows correlation coefficients between the 10 parameters based on a subset of simulations in the region of compromise of the four-objective space, identified by the following criteria: $RMSE_{wt}<3$, $RMSE_{aq}<10$, $RMSE_{drain}<30$, and $RMSE_{seep}<50$. Note that these correlations were calculated on the estimates of the parameter values, not the log-transformed values. The strongest correlation in Table 2 equals −0.25, between parameters f_{Cd} and f_{Ks}, suggesting that, for the sake of parameter identifiability, parameter correlations are small (Hill 1998).

Table 3 summarizes the optimal parameter and RMSE values when minimizing the four objectives separately, as well as for the compromise solution which minimizes the Euclidean distance in the normalized objective function space (shown as a cross in Fig. 6). The range in optimal parameter values for the various objective functions is still quite large compared to the prior ranges specified in Table 1. In other words, the parameter uncertainty associated with the multi-objective Pareto solution is large, typically an order of magnitude. As was discussed earlier, this is due to the fact that the different objective functions are sensitive to different parameters. However, this does not mean that the parameters are necessarily not well identified or that the model is in error. For example, canal seepage is very sensitive to the canal bed hydraulic conductivity (Fig. 8), hence this parameter is known with certainty. Its Pareto uncertainty on the other hand is large, because the other objectives are not that sensitive to it. Note also that the parameter values of the compromise solution usually lie close to the parameter value obtained by minimizing the objective function that is most sensitive to

that parameter. For example, in the compromise solution, parameter f_{Ks} takes on the value of 1.3, which is the same as its optimal value when fitting to the seepage rates only. The same can be said about parameters f_{Cd} and f_{Kh3}. Furthermore, Table 3 shows that the compromise solution does only slightly worse in fitting the four objectives compared to the single objective results. Hence, little trade-off or compromise exists between fitting the four objective functions.

3.2. "Best" parameter set

Next, we select the compromise solution discussed in Table 3 as our "best" model and examine how well it mimics the observed dynamics in the system. The results are presented in Figs. 9, 10, 11 and 12. Figure 9 shows scatter plots of simulated versus observed water table elevations, aquifer heads, drainage rates, and canal seepage rates. The corresponding RMSE values are listed in Table 3. The prediction of water table elevations can be considered very good, whereas the plot of simulated vs. observed aquifer heads still shows a lot of scatter (RMSE=7.7 m). The spatial and temporal variation in drainage volumes is fairly well predicted, (RMSE=17.6 MCM/yr), especially in view of the uncertainty associated with the drainage measurements. Canal seepage rates in the two canals are predicted quite well (RMSE=26 MCM/yr), as is also apparent in Fig. 10, which presents time-series of observed and simulated canal seepage and total drainage rates.

Figure 11 shows pie charts of the simulated time-averaged (1974–1997) annual water balance. The two main sources of water inflow into the groundwater system are, first, irrigation water applied on agricultural fields and seepage from secondary irrigation canals (2445 MCM/year or 84% on average), and second, seepage from the main irrigation canals (454 MCM/year or 16%). Groundwater discharge is composed of crop evapotranspiration (1513 MCM/year or 53%), non-agricultural evapotranspiration and bare-soil evaporation (531 MCM/year or 19%), surface and subsurface agricultural drainage discharging into the Sea of Cortez through the outlets of the drainage network (448 MCM/year or 15%), and groundwater pumping (271 MCM/year or 9%).

Note that subsurface groundwater discharge to the Sea of Cortez accounts only for 12 MCM/year which is less than 1% of the total annual outflow. These simulated values compare well to observed canal seepage volumes (473 MCM/year on average) and drainage volumes (499 MCM/year observed on average during 1988–1997, compared to 477 MCM/year simulated over the same period for that part of the drainage network that is monitored).

Finally, we investigated the model performance on aquifer heads using the best parameter set, and suggest an improvement in the model structure based on a spatial analysis of the model residuals. Figure 12a shows a map of the time-averaged residual error, interpolated from errors calculated for the wells shown in the same figure. There are clear spatial patterns, with regions of consistent over-estimation (residual error greater than zero) and under-estimation (residual error smaller than zero). This suggests that some adjustment is needed in the spatial distribution of hydraulic properties, as initially estimated from well log and specific capacity data. Instead of using a single uniform scaling factor, we can use the spatial patterns in the model residual map to construct distinct zones of uniform scaling factors. This is discussed further in the next section.

3.3. Model refinement for aquifer heads

Based on a comparison of simulated and observed values using the best parameter set identified by the multi-objective analysis, we concluded that the model does reasonably well in simulating water table elevations, canal seepage and drainage rates. However, results for the aquifer heads were less satisfactory as shown by the large value of $RMSE_{aq}$ (7.7 m) and the model residual map in Fig. 12a. Unfortunately, there is only limited information on the geology in this area. The available information (well logs) was used to specify the initial spatial distribution of conductivities, but there is considerable uncertainty on these estimates. Here an attempt is made to improve the model performance by including more (indirect) information about the geology, namely in the form of hydraulic heads in the deep aquifer. First, the spatial patterns in the

model residual map were used to construct distinct zones of uniform scaling factors. This zonation is shown in Fig. 12b. Results from the multi-objective calibration also indicated that the simulation of aquifer heads is most sensitive to the scaling factors for K_{h3} and K_{v2}. Therefore, the model was refined by introducing spatially varying scaling factors for K_{h3} and K_{v2}, defined by the zonation in Fig. 12b. Based on the results of the multi-objective calibration run, changing these parameter values will have little effect on the simulation of water table elevations, drainage rates, and canal seepage rates.

A single-objective calibration was performed on aquifer heads only, using 8 calibration parameters, i.e. scaling factors for K_{h3} and K_{v2} for each of the four zones in Fig. 12b. All the other parameters that were considered before (Table 1) are now fixed at their "best" values as identified by the multi-objective calibration (Table 3). The SCEM-UA global optimization algorithm (Vrugt et al., 2003a) was used to identify the parameter values that minimize RMSE_{aq}. Optimal parameter values identified by the algorithm after 3000 model runs are listed in Table 4. Corresponding values for the various RMSE values are also shown. We can see that the zonation of the hydraulic properties resulted in a decrease of RMSE_{aq} from 7.7 (Table 3) to 5.6 m (Table 4). In addition, the zonation also caused a better fit to steady-state aquifer heads (1972–1974), as measured by a decrease in the steady-state RMSE values from 6.3 to 3.0 m. Values of the other three objectives not included in the new calibration (Table 4) are similar, even slightly better, than those before the second calibration (Table 3). Nevertheless, spatial patterns in the model residuals still exist (Fig. 12c), although the errors have decreased. The optimal parameter values in Table 4 are consistent with the spatial patterns of over- and under-prediction in Fig. 12a. The over-predictions in zones 2 and 4 result in smaller values for f_{Kv2} in these zones compared to the best parameter set in Table 3, whereas the opposite occurs for zone 3, where heads were consistently under-predicted. Note that the scaling factors are again relative to the originally estimated hydraulic conductivities using well log and specific capacity data. Although in Table 4, the optimal values of f_{Kv2} for zones 2 and 4 are at or near the prior minimum value of 0.01, the results could not be improved by letting f_{Kv2} reach even lower values.

In other words, once f_{Kv2} is small enough very little water will percolate vertically. Final calibrated maps of K_{v2} and K_{h3}, obtained by multiplying the original estimates with the optimal scaling factors in Table 4, are shown in Fig. 13. The large horizontal conductivities in the northern part of the study area agree with the occurrence of coarse gravel deposits observed in the well logs of that area.

Finally, parameter uncertainty was estimated on a set of behavioral models, defined by the criterion that $\mathrm{RMSE}_{aq}<6$ m. Minimum and maximum parameter values for this subset of 229 parameter sets are shown in Table 4, and typically span an order of magnitude. The resulting prediction uncertainty is shown in Fig. 14. This figure presents time-series of observed and simulated aquifer heads for various locations, which are identified by a number in each graph corresponding to the numbered wells in Fig. 4. The optimal model is shown as a dark solid line, whereas the grey lines represent the range in prediction uncertainty associated with all parameter sets that result in an $\mathrm{RMSE}_{aq}<6$ m. In most cases, the ranges in predicted heads bracket the observations. Several simulated and observed heads decline during the 1970s and into the early 1980s, followed by a recovery in head levels after that. This corresponds to a period of increased groundwater pumping in the early 1980s, followed by an increase in surface water supply and a decrease in groundwater pumping after 1985.

4. Conclusions

This paper presented the application of multi-objective optimization to calibrate a regional surface water-groundwater model for the Yaqui Valley in Mexico. Irrigated agriculture in this highly engineered hydrologic system with its surface reservoirs empty is in serous jeopardy. Groundwater as an increasingly valuable commodity and. a groundwater model will be an increasingly important tool to ensure that water is better managed to sustain crop production in the future. Our model integrates the surface water and groundwater systems and accounts for the spatial distribution of annual recharge from irrigation, subsurface drainage, agricultural pumping, and irrigation

canal seepage. A four-objective calibration problem was formulated to find improved estimates of the hydraulic parameters using head data from the shallow and deep aquifers, as well as data on agricultural drainage and canal seepage volumes. The main advantage of the method is that it accounts for both parameter and model structural uncertainty. In this case, results show that the effect of including bare soil evaporation is greater than the effects of parameter uncertainty, as evidenced by a strong trade-off between the drainage and seepage objectives when bare soil evaporation is not included in the model. Furthermore, by treating the different objectives independently, the method allowed for better identification of the model parameters compared to a single-objective approach, since the various objectives were sensitive to different parameters. Large parameter variation (uncertainty) of the Pareto set of solutions, which includes the best-fit end-members for each of the objective functions, does not always point to a model structural error, because such large parameter uncertainty may also be caused by the insensitivity of one of the objectives for the parameter. The shape of the trade-off curve is a much better indicator of model structural error. The simulated water balance shows that 15–20% of the water that enters the irrigation canals is lost by seepage to groundwater. The main discharge mechanisms in the Valley are crop evapotranspiration (53%), non-agricultural evapotranspiration and bare soil evaporation (19%), surface drainage to the Sea of Cortez (15%), and groundwater pumping (9%). In comparison, groundwater discharge to the estuary was relatively insignificant (less than 1%). Heads in the deep aquifer were most sensitive to the vertical conductivity of the confining layer (K_v) and the horizontal conductivity in the deep aquifer (K_h). The model was further refined by identifying zonal K_v and K_h values based on a spatial analysis of the model residuals. Subsequent calibration of K_v and K_h to aquifer head only (single-objective) resulted in further improvements in simulated heads. The final calibrated parameter values and their associated uncertainties presented here will be used in future work on identifying optimal groundwater management strategies and their reliability. Since groundwater storage in the Valley is about 16 times greater than the available surface water in the reservoirs, sustainable water resource allocation will

undoubtedly rely on water pumped from storage.

Acknowledgements. We would like to acknowledge data contributions from the Yaqui Irrigation District and from the CNA in Mexico. We would like to thank J. Vrugt for providing us with the MOSCEM-UA and SCEM-UA optimization algorithms. We are most grateful to the David and Lucile Packard Foundation for their support of this research. We are also grateful to the UPS Foundation for their support of the initial phases of this work.

References

Addams, C. L.: Water resource policy evaluation using a combined hydrological-economic-agronomic modeling framework: Yaqui Valley, Sonora, Mexico, PhD, Stanford University, USA, 2004.

Anderson, T. W., Freethey, G. W., and Tucci, P.: Geohydrology and water resources of alluvial basins in south central Arizona and parts of adjacent states, USGS Professional Paper 1406-B, 1992.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001.

Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods, Water Resour. Res., 36, 3663–3674, 2000.

Diaz, S.: Planeacion del uso conjunto de aguas superficiales y subterráneas en el Valle del Yaqui, Sonora: Modelación del acuífero del Valle del Yaqui, Instituto Technológico de Sonora, Obregon, Mexico, 1995.

Doherty, J.: PEST, model-independent parameter estimation. Watermark Numerical Computing, Townsville, 2000.

Duan, Q., Gupta, V. K., and Sorooshian, S.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resour. Res., 28, 1015–1031, 1992.

Flint, A. L., Flint, L. E., Hevesi, J. A., D'Agnese, F. D., and Faunt, C.: Estimation of regional recharge and travel time through the unsaturated zone in arid climates, in: Dynamics of fluids in fractured rock, edited by: Faybishenko, B., Witherspoon, P. A., and Benson, S. M., AGU, 115–128, 2000.

Gardner, W. R.: Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from a water table, Soil Sci., 85, 228–232, 1958.
Gardner, W. R. and Fireman, M.: Laboratory studies of evaporation from soil columns in the presence of a water table, Soil Sci., 85, 244–249, 1958.
Gonzales, R. and Marin, L. E.: Modelo Hidrogeológico conceptual del acuífero del Valle del Yaqui, Sonora en un contexto geológico regional, Instituto Technológico de Sonora, Obregon, Mexico, 2000.
Gupta, H. V., Bastidas, L., Vrugt, J. A., and Sorooshian, S.: Multiple criteria global optimization for watershed model calibration, in: Calibration of Watershed Models, edited by: Duan, Q., AGU Water Science and Applications Series, 6, 125–132, 2003.
Harbaugh, A. W., Banta, E. R., Hill, M. C., and McDonald, M. G.: MODFLOW-2000, the U.S. Geological Survey modular ground-water model: User guide to modularization concepts and the Ground-Water Flow Process, USGS Open-File Report 00-92, 2000.
Hill, M. C.: Methods and guidelines for effective model calibration, USGS Water Resources Investigations Report 98-4005, 1998.
Islas, L. A.: Determinación y análisis espacial prospectivo de las variables hidrodinámicas del acuífero del valle del Yaqui, Sonora, partiendo de información escasa, Instituto Technológico de Sonora, Obregon, Mexico, 1998.
ITC: Geofísica integral del los Valles del Yaqui y del Mayo, Mexico City, 1979.
Lobell, D. B., Ortiz-Monasterio, J. I., Addams, C. L., and Asner, G. P.: Soil, climate, and management impacts on regional agricultural productivity from remote sensing, Agric. and Forest Meteor., 114, 31–43, 2002.
Madsen, H.: Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives, Adv. Water Resour., 26, 205–216, 2003.
Poeter, E. P. and Hill, M. C.: Documentation of UCODE: A computer code for universal inverse modeling, USGS Water Resources Investigations Report 98-4080, 1998.
Prudic, D. E., Konikow, L. F., and Banta, E. R.: A new streamflow-routing package to simulate stream-aquifer interaction with MODFLOW-2000, USGS Open-File Report 2004-1042, 2004.
Razack, M. and Huntley, D.: Assessing transmissivity from specific capacity in a large and heterogeneous alluvial aquifer, Ground Water, 29, 856-861, 1991.
Refsgaard, J. C. and Henriksen, H. J.: Modeling guidelines, terminology and guiding principles, Adv. Water Resour., 27, 71-82, 2004.
Schoups, G., Hopmans, J. W., Young, C. A., Vrugt, J. A., and Wallender, W. W.: Multi-criteria

optimization of a regional spatially-distributed subsurface water flow model, J. Hydrol., in press, 2005.

Šimůnek, J., Šejna, M., and van Genuchten, M. Th.: The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media, International Ground Water Modeling Center, Colorado School of Mines, 1998.

Steinich, B. and Chavarria, J. A.: Determination of hydrogeological characteristics and mapping of the sea water intrusion of the Yaqui Valley aquifer, Sonora, Mexico, in: Aquatic ecosystems of Mexico: status and scope, edited by: Munawar, M. L., Manawar, I. F., and Malley, D. F., Backhuys Publishers, Netherlands, 2000.

Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters, Water Resour. Res., 39, 1201–1218, 2003a.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multi-objective optimization of hydrologic models, Water Resour. Res., 39, 1214–1232, 2003b.

Table 1. Calibration parameters and their prior uncertainty ranges.

Parameter	Description	Units	Minimum	Maximum	Transformation
$f_{Kv,h}(i)$	Scaling factors for hydraulic conductivity (v, vertical; h, horizontal; i = 1...3, layer index)	–	0.1	10.0	Log_{10}
S_y	Specific yield	–	0.10	0.30	None
S_s	Specific storage	1/m	10^{-5}	10^{-3}	Log_{10}
f_{Cd}	Scaling factor for drain conductance	–	0.1	10.0	Log_{10}
f_{Ks}	Scaling factor for conductivity of canal bottom	–	0.1	10.0	Log_{10}

Table 2. Parameter correlation coefficient matrix for a subset of simulations, identified by the following criteria: $RMSE_{wt}$<3, $RMSE_{aq}$<10, $RMSE_{drain}$<30, and $RMSE_{seep}$<50.

	f_{Kv1}	f_{Kh1}	f_{Kv2}	f_{Kh2}	f_{Kv3}	f_{Kh3}	S_y	S_s	f_{Cd}	f_{Ks}
f_{Kv1}	**1.00**	−0.04	0.04	0.11	−0.10	0.03	0.15	−0.20	0.03	0.07
f_{Kh1}	−0.04	**1.00**	0.06	−0.04	0.19	0.03	0.10	0.19	−0.11	0.07
f_{Kv2}	0.04	0.06	**1.00**	−0.12	0.11	−0.01	0.15	−0.21	0.02	−0.06
f_{Kh2}	0.11	−0.04	−0.12	**1.00**	−0.20	−0.02	0.01	−0.21	−0.01	−0.14
f_{Kv3}	−0.10	0.19	0.11	−0.20	**1.00**	0.01	−0.03	−0.02	−0.12	−0.17
f_{Kh3}	0.03	0.03	−0.01	−0.02	0.01	**1.00**	0.12	0.00	−0.25	0.16
S_y	0.15	0.10	0.15	0.01	−0.03	0.12	**1.00**	0.12	−0.18	−0.10
S_s	−0.20	0.19	−0.21	−0.21	−0.02	0.00	0.12	**1.00**	−0.03	0.08
f_{Cd}	0.03	−0.11	0.02	−0.01	−0.12	−0.25	−0.18	−0.03	**1.00**	−0.14
f_{Ks}	0.07	0.07	−0.06	−0.14	−0.17	0.16	−0.10	0.08	−0.14	**1.00**

Table 3. Optimal parameter and RMSE values for various objectives. "Euclid-dist" refers to the distance in the normalized four-dimensional objective function space.

	Minimize $RMSE_{wt}$	Minimize $RMSE_{aq}$	Minimize $RMSE_{drain}$	Minimize $RMSE_{seep}$	Minimize "Euclid-dist"
f_{Kv1}	2.9	3.3	1.1	4.3	0.6
f_{Kh1}	7.8	1.6	7.3	1.2	2.0
f_{Kv2}	0.4	1.1	0.4	0.5	0.2
f_{Kh2}	9.3	7.3	1.6	0.5	1.7
f_{Kv3}	0.4	8.4	1.2	0.5	0.4
f_{Kh3}	8.6	9.1	2.9	0.9	7.0
S_y	0.26	0.11	0.15	0.18	0.2
S_s	2.0E-04	1.8E-05	6.6E-05	3.3E-05	5.3E-04
f_{Cd}	0.1	5.1	1.7	2.4	1.8
f_{Ks}	0.6	0.4	0.9	1.3	1.3
$RMSE_{wt}$ (m)	**2.5**	2.6	2.8	2.9	2.8
$RMSE_{aq}$ (m)	7.4	**7.2**	9.4	15.5	7.7
$RMSE_{drain}$ (MCM)	41.4	23.6	**15.3**	23.6	17.6
$RMSE_{seep}$ (MCM)	125.2	154.9	72.8	**25.4**	26.6

Table 4. Optimal parameter and RMSE values for various objectives. The different zones are mapped in Fig. 12b. * These are minimum and maximum values of all parameter sets with an $RMSE_{aq}$<6 m.

	Zone	Minimum (prior)	Maximum (prior)	Optimum (minimize $RMSE_{aq}$)	Minimum*	Maximum*
f_{Kv2}	1	0.01	100	1.0	0.08	9.4
	2	0.01	100	0.01	0.01	0.19
	3	0.01	100	37	6	97
	4	0.01	100	0.013	0.01	0.15
f_{Kh3}	1	0.1	10	5	1.2	10
	2	0.1	10	8.7	0.8	10
	3	0.1	10	1.2	0.12	10
	4	0.1	10	0.40	0.16	10
$RMSE_{wt}$ (m)				2.7		
$RMSE_{aq}$ (m)				5.6		
$RMSE_{drain}$ (MCM)				17.4		
$RMSE_{seep}$ (MCM)				26.2		

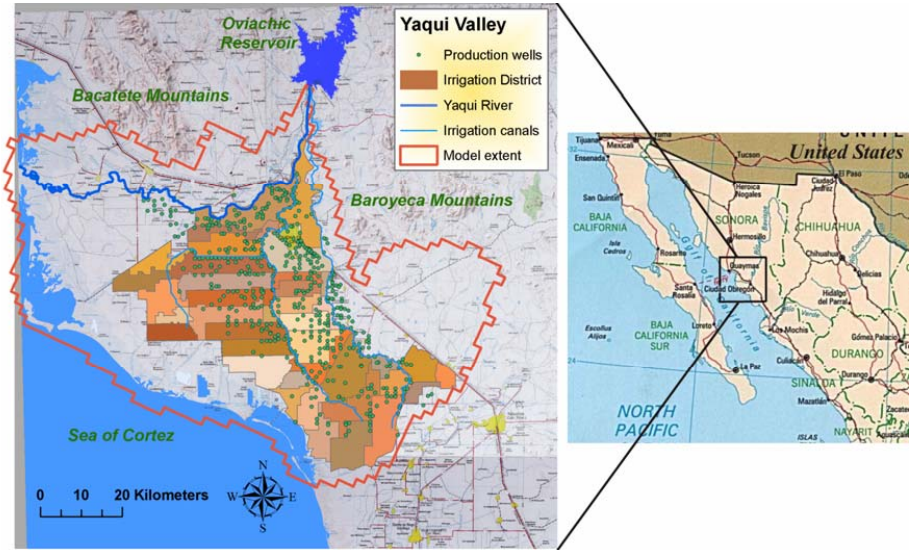

Fig. 1. Location of the Yaqui Valley study area, the Yaqui Irrigation District, and extent of the groundwater model.

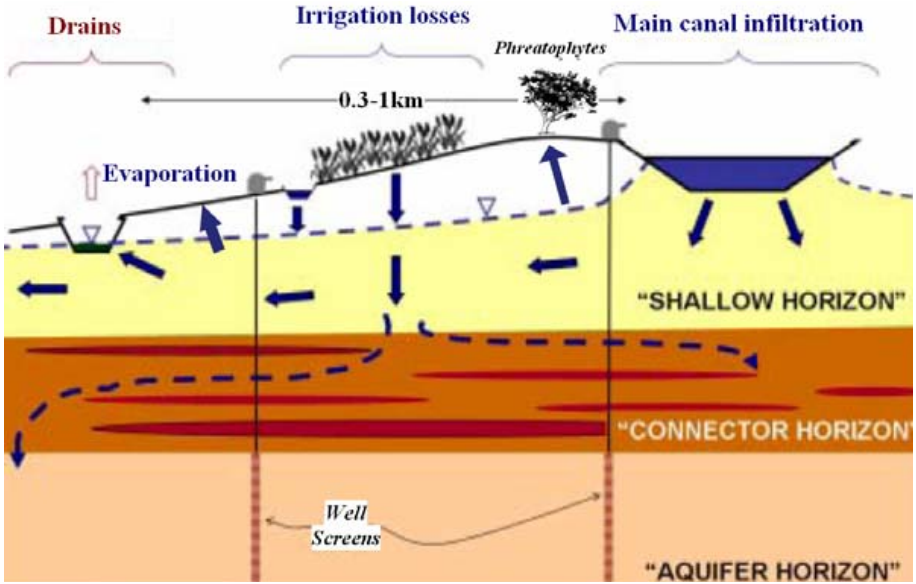

Fig. 2. Conceptual model of the Yaqui Valley surface-groundwater system.

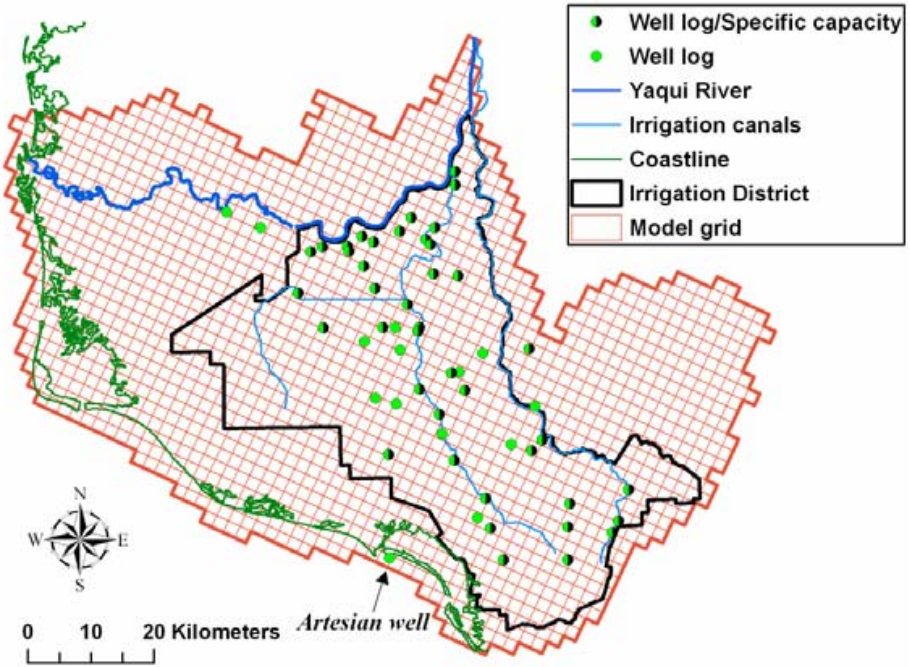

Fig. 3. Locations of wells with well log and specific capacity data used to define model layers and hydraulic properties. The model grid is also shown.

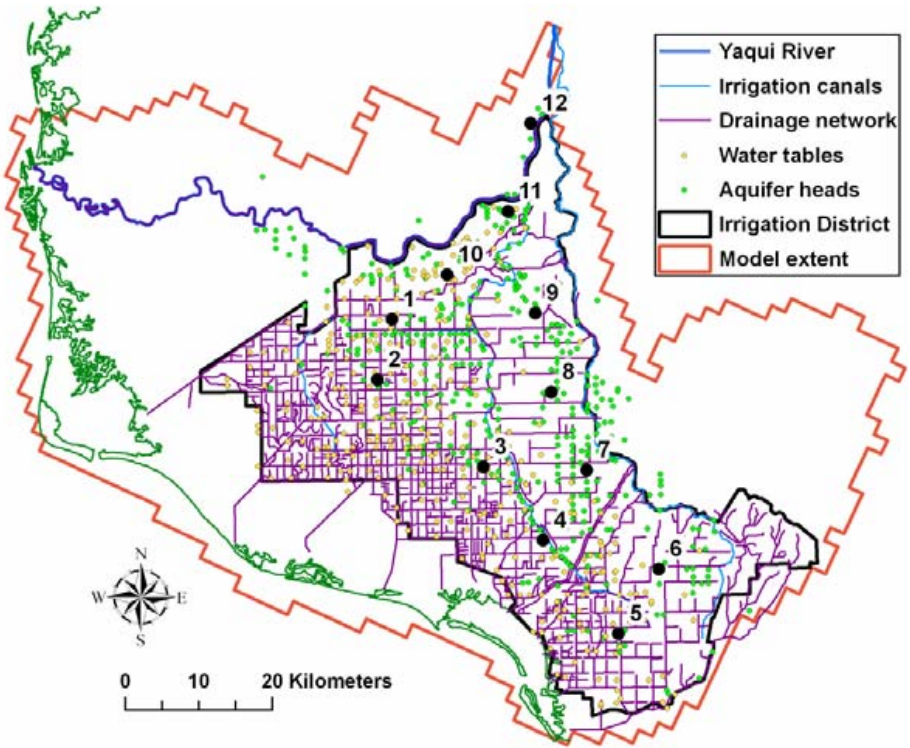

Fig. 4. Locations of wells with observed water table elevations and aquifer heads used for model calibration. The drainage networks with measured drainage volumes are also shown. Numbers correspond to wells for which time-series are shown in Fig. 14.

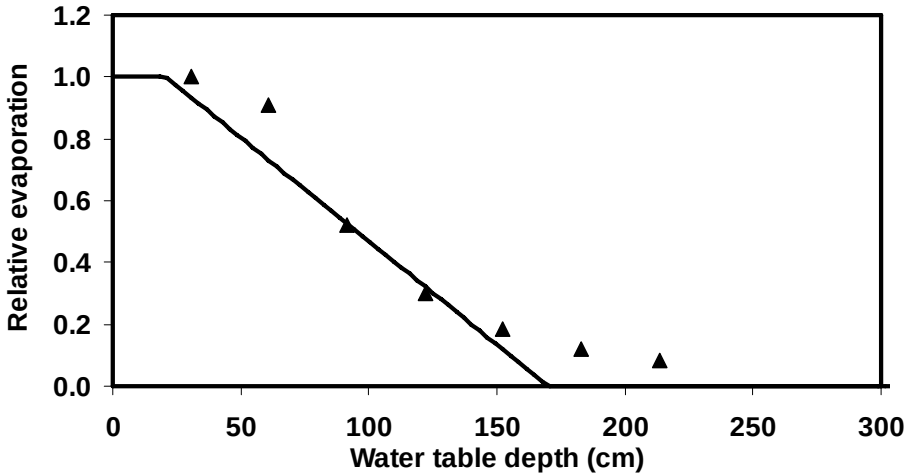

Fig. 5. Relation between relative evaporation and water table depth for a clay soil, simulated with the HYDRUS code (triangles) and approximated with a piecewise linear function.

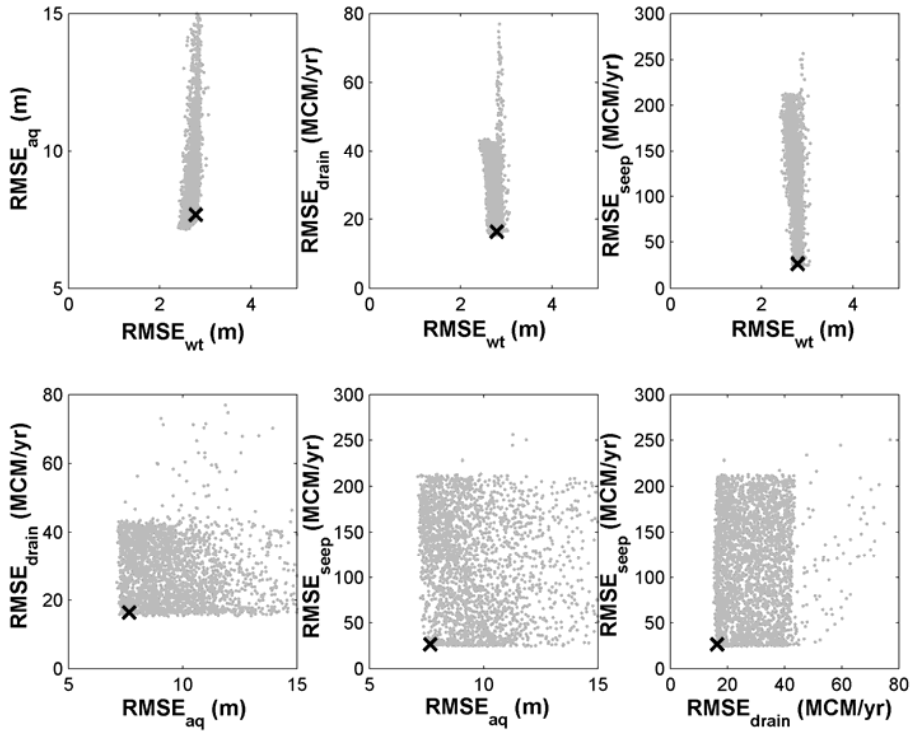

Fig. 6. Bi-criterion plots for the four-objective calibration run after 3000 model evaluations with the MOSCEM-UA algorithm.

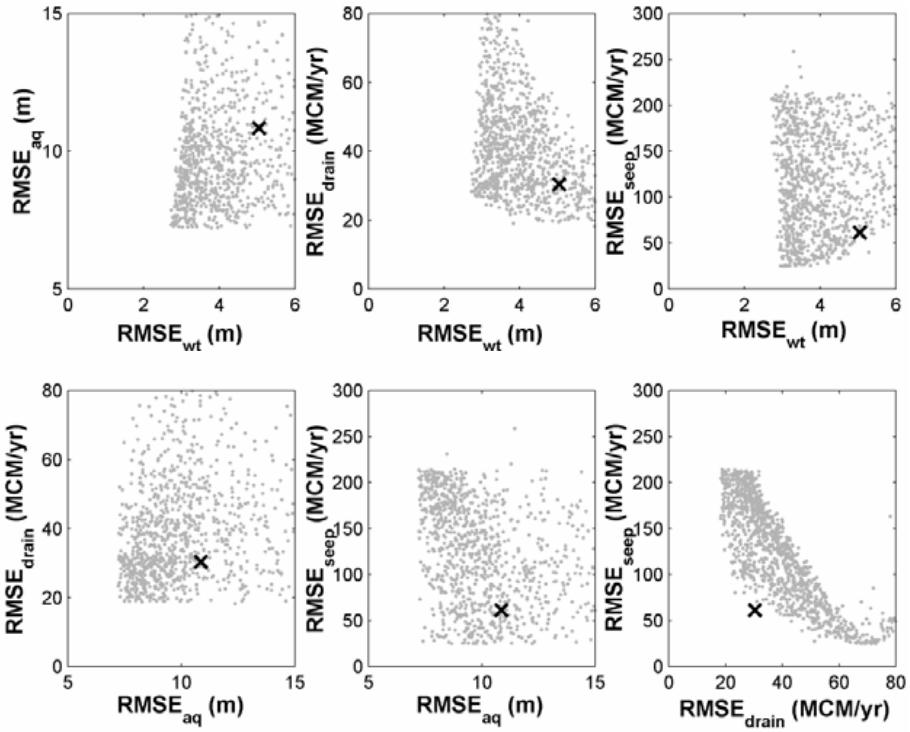

Fig. 7. Bi-criterion plots for the four-objective calibration run after 3000 model evaluations with the MOSCEM-UA algorithm, without bare soil evaporation.

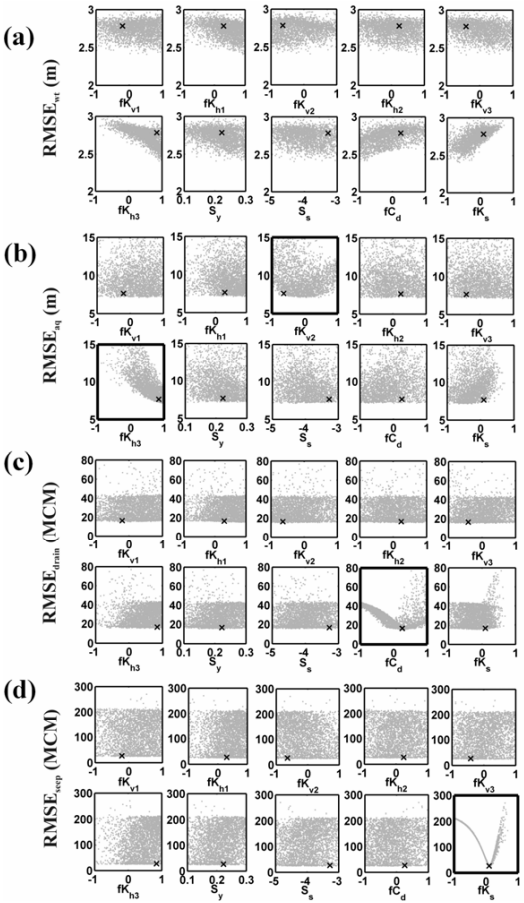

Fig. 8. Dotty plots for the four-objective calibration run, showing sensitivity of the various objective functions to the calibration parameters. **(a)** water table elevations, **(b)** aquifer heads, **(c)**drainage rates, and **(d)** canal seepage rates. All parameter scales are log-transformed, except S_y. Bold frames highlight significant effects of a parameter on an objective function.

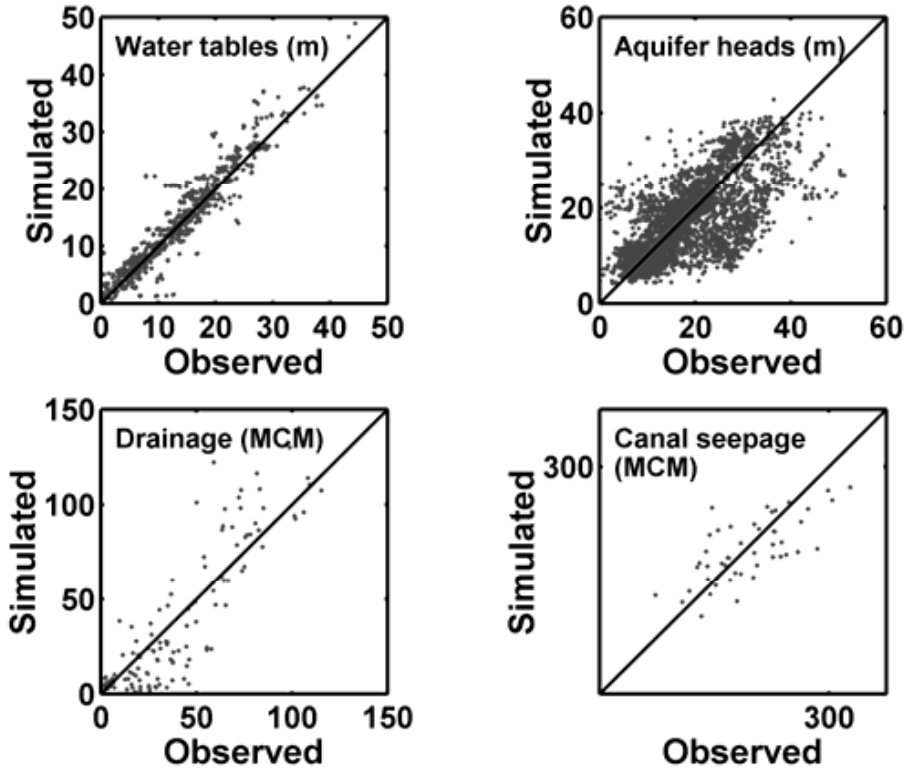

Fig. 9. Scatter plots of simulated versus observed water table elevations, aquifer heads, drainage volumes, and canal seepage volumes.

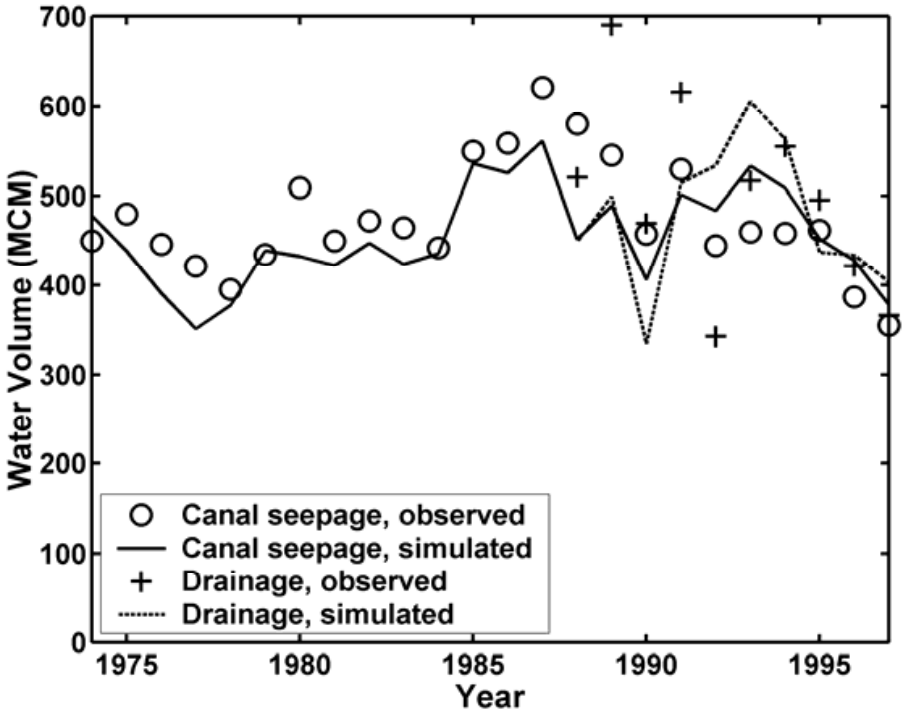

Fig. 10. Time-series of observed and simulated canal seepage and total drainage volume.

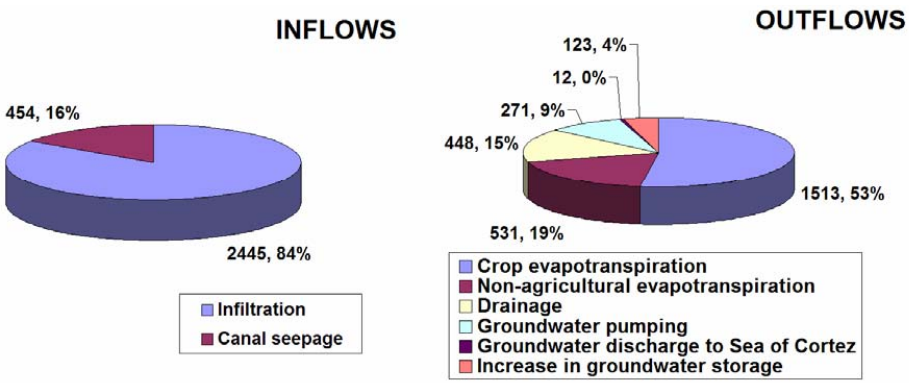

Fig. 11. Pie chart of the annual simulated water balance, averaged over the period 1974–1997 expressed as MCM and as percentages of the total inflows and outflows. Recharge below the root-zone, calculated as the difference between infiltration and crop *ET*, constitutes 65% of the total inflow into groundwater.

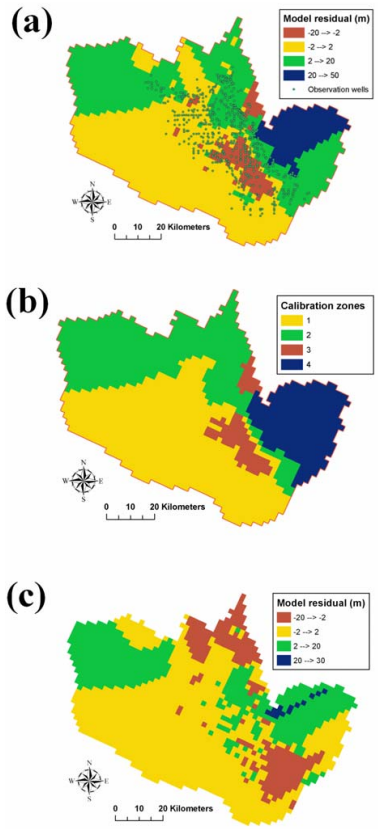

Fig. 12. **(a)** Interpolated map of time-averaged model residuals for simulated aquifer heads using the best parameter set of the multi-objective calibration, **(b)** corresponding zones of uniform K_{v2} and K_{h3} scaling factors used in the single-objective calibration, and **(c)** resulting time-averaged residuals after the second calibration.

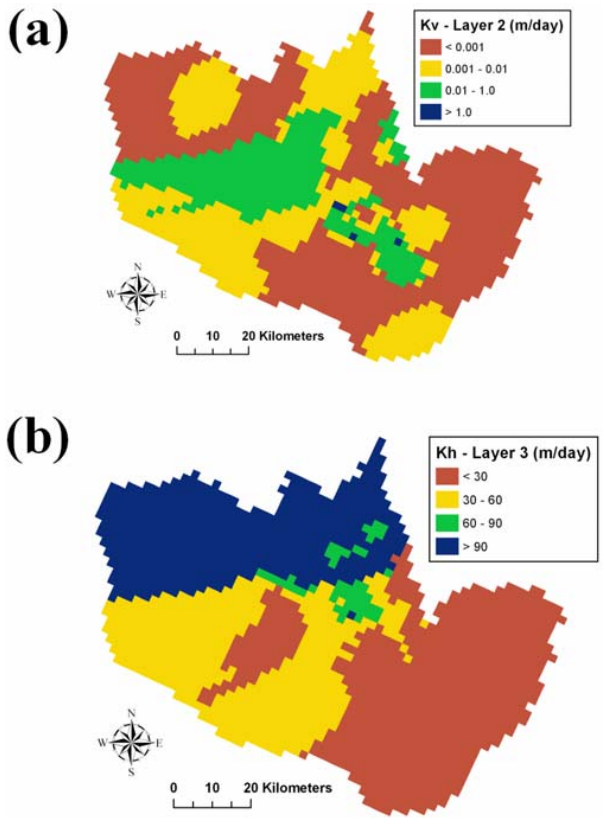

Fig. 13. Calibrated maps of K_{v2} **(a)** and K_{h3} **(b)** using the optimal parameter set in Table 4 after the second calibration.

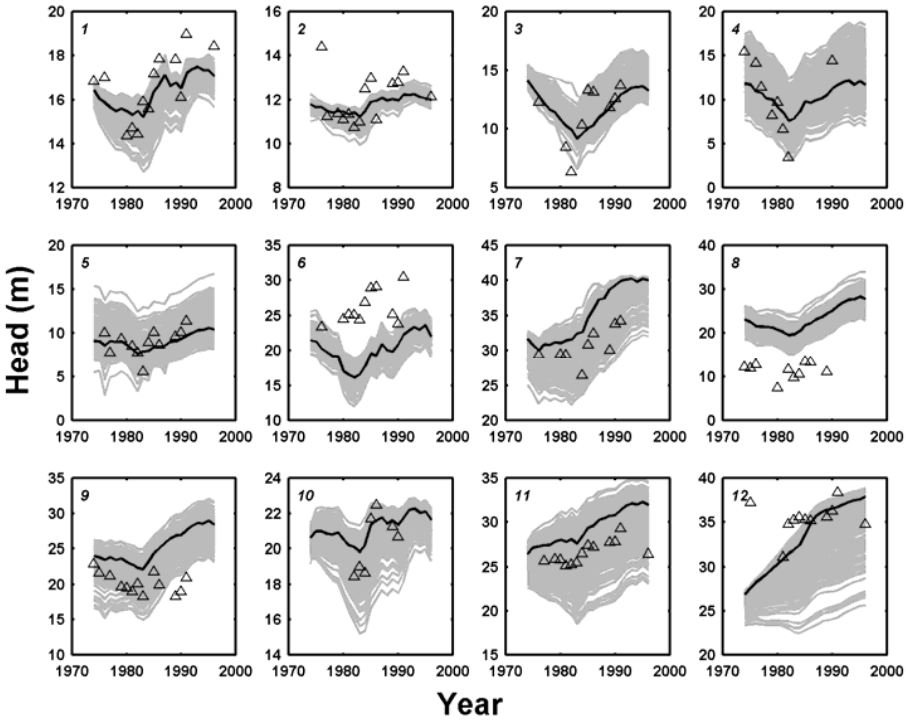

Fig. 14. Time-series of observed (triangles) and simulated (solid lines) aquifer heads for various locations, shown in Fig. 4. The dark solid line corresponds to the optimal solution of the single-objective calibration. The grey lines represent all simulations with an $RMSE_{aq}$ less than 6 m.