Multi-criteria optimization of a regional spatially-distributed subsurface water flow model

Article *in* Journal of Hydrology · September 2005

Impact Factor: 3.05 · DOI: 10.1016/j.jhydrol.2005.01.001

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 311 (2005) 20–48
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Multi-criteria optimization of a regional spatially-distributed subsurface water flow model

G. Schoups[a], J.W. Hopmans[a,*], C.A. Young[a], J.A. Vrugt[b], W.W. Wallender[a]

[a]*Hydrology Program, Department of Land, Air and Water Resources (LAWR), University of California, 123 Viehmeyer Hall, Davis, CA 95616, USA*
[b]*Faculty of Science, Institute for Biodiversity and Ecosystem Dynamics, Universiteit van Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands*

Received 9 June 2004; revised 3 December 2004; accepted 3 January 2005

Abstract

This paper presents the multi-criteria calibration of a regional distributed subsurface water flow model for a 1400 km^2 irrigated agricultural area in the western San Joaquin Valley of California. Two global optimization algorithms were used to identify model parameters using data on spatially distributed local water table depth measurements, district-average groundwater pumping and district-average subsurface drainage data. Model parameters that were subjected to calibration included irrigation efficiency, effective drain depth and conductance, crop evapotranspiration correction coefficient, saturated hydraulic conductivity and specific yield values of coarse and fine fractions, and saturated hydraulic conductivity values defining water fluxes across domain boundaries. Using the single-objective function approach, the three measurement types were weighted into a single-objective function for global optimization purposes. Additionally, a three-objective multi-criteria optimization problem was formulated in which no prior weighting of the individual objectives was specified. The single-objective optimization approach resulted in identifiable parameters with relatively small uncertainties, however, most likely values for various optimized parameter approached the outer bounds of their physical-realistic ranges. The normalized weighting of the single-objective function approach emphasized the pumping and drainage data more than the water table depth data. In the multi-objective approach, the objective function of each measurement type was treated independently, so that no subjective preferences were assigned a priori. Within a single optimization run, a Pareto set of solutions was generated, which included the optimal results for each end-member of each of the three objective functions. The results showed a moderate trade-off between pumping and water table predictions, and a slight independence of drainage predictions from the other two measurements. The estimated Pareto set exhibited large parameter uncertainty, indicating possible model structural inadequacies. We further show that the magnitude of prediction uncertainties associated with the Pareto parameter uncertainty is large for making water table predictions, but much smaller for drainage and pumping predictions. Trade-offs between fitting shallow and deep water tables were revealed by considering additional performance criteria for model evaluation, namely BIAS

* Corresponding author. Fax: +1 530 752 5262.
E-mail address: jwhopmans@ucdavis.edu (J.W. Hopmans).

doi:10.1016/j.jhydrol.2005.01.001

and RMSE values for six water table depth groups. These results point to possible model improvements by spatially distributing some of the model parameters.

Keywords: Vadose zone hydrology; Salinity; Drainage; Irrigation efficiency; Measurement error

1. Introduction

Two fundamental problems complicate large-scale (field to regional) modeling of variably-saturated subsurface flow and transport. First, no fundamental 'laws' are available that describe flow and transport at the field or regional scale. Although Darcy's law, the Richards equation (RE) and the advection–dispersion equation (ADE) are valid at the local scale, their application at the field-scale is questionable, for example, in the presence of preferential flow and transport (Šimůnek et al., 2003). Second, parameterization of large-scale models is hampered by the tremendous spatial heterogeneity of the subsurface, as well as the spatial and temporal variations in boundary conditions. Consequently, methodologies are needed that quantify the propagation of model structure and model input uncertainties in model output. Essentially, two broad approaches may be distinguished to achieve this.

The first approach consists of upscaling the small-scale physics, as represented by the RE and ADE, to the larger scale of interest using statistical averaging procedures. Model parameters and boundary and initial conditions are treated as stochastic variables with prescribed probability density functions (pdf's), which are estimated from a large number of local-scale measurements. The propagation of parameter uncertainty in the simulation results is accomplished by Monte Carlo (MC) analysis by solving the large-scale hydrological model for many realizations with each realization generated from the pdf's of parameter and boundary condition values. An example of this approach is the GLUE methodology of Beven and Freer (2001). Their methodology recognizes that often many parameter sets may describe the available data, resulting in equifinality of the parameters (Beven, 1993). Although MC analysis is conceptually straightforward, it can become computationally very intensive. Stochastic methods based on first-order perturbation (Zhang, 2002) and cumulant expansion (Wood and Kavvas, 1999) approximations have been developed that achieve similar upscaling results as MC analysis, but at a much lower computational costs. Though, the problems of estimating the parameter pdf's from local-scale measurements still remains.

The second approach to large-scale modeling is to directly apply the hydrologic model at the scale of interest using single-valued effective parameters. These effective parameters are defined such that the model results using the uniform parameter field and the heterogeneous parameter field are identical (Blöschl and Sivapalan, 1995). Effective parameters are usually estimated from inverse modeling, whereby the model parameters are adjusted by minimizing differences between observed and predicted output variables (residuals). Due to the inherent nonlinear nature of variably-saturated flow in porous media, however, no effective parameters exist in general that are a function of the porous media only (Beven, 2001). Nevertheless, effective parameter values may still produce very good results for specific situations (e.g. Wildenschild and Jensen, 1999) or for a range of flow conditions. Whereas traditional inverse modeling has focused on estimating the optimal parameter sets, recent studies have developed algorithms that also provide estimates of parameter and model structure uncertainties, especially with applications in watershed scale rainfall-runoff modeling (Gupta et al., 1998).

An example of this approach is the Shuffled Complex Evolution Metropolis global optimization algorithm, SCEM-UA, recently developed at the University of Amsterdam (Vrugt et al., 2003a). The SCEM-UA algorithm is a general purpose global optimization algorithm that provides an efficient estimate of the most likely parameter set and its underlying posterior probability distribution, defining parameter uncertainty within a single optimization run. The algorithm is an extension of the SCE-UA population evolution method developed by Duan et al. (1992). In this single-objective function approach,

one may define multiple criteria, but merge them into a single function for optimization. Alternatively, model structural uncertainty can be investigated using multi-objective optimization (Gupta et al., 1998), whereby the different performance criteria are treated independently, so that a full multi-criteria optimization is conducted that identifies the entire optimal set of Pareto solutions. By analyzing the trade-offs among the different criteria, one can better understand the limitations of the hydrologic model structure and can suggest model improvements (Refsgaard and Henriksen, 2004). In the absence of significant model structure errors, the multi-objective approach includes more information about the hydrologic system in the parameter identification process, thereby leading to a better identification of the parameters (Boyle et al., 2000). On the other hand, significant trade-offs in fitting two or more objectives may indicate model structure error. Yapo et al. (1998) and Vrugt et al. (2003b) developed algorithms capable of solving the multi-objective problem in a single optimization run.

Most research on the identification of effective model parameter values for vadose zone flow and transport has focused on laboratory-scale experiments. The capabilities and limitations of the inverse approach for the identification of soil hydraulic properties from laboratory soil cores may be considered reasonably well understood (e.g. Kool and Parker, 1988; van Dam et al., 1994; Hopmans et al., 2002). With much of the focus of inverse modeling in vadose zone hydrology and soil physics on the laboratory scale, very little has been done at regional scales. Most large-scale inverse modeling applications have been in the areas of groundwater (e.g. McLaughlin and Townley, 1996; Medina and Carrera, 1996; Weiss and Smith, 1998; Karpouzos et al., 2001) and surface water hydrology (e.g. Duan et al., 2003). However, recently, Madsen (2003) applied the SCE-UA global optimization algorithm to the calibration of the distributed physically-based MIKE-SHE watershed model that included the coupling of the vadose zone with the underlying aquifer. Vrugt et al. (accepted for publication) used the SCEM-UA algorithm to estimate spatially distributed soil hydraulic properties using a physically-based three-dimensional distributed variably-saturated water flow model, by incorporating spatially-distributed tile drainage flow measurements from the Broadview Water District (BWD), CA, in the objective function.

In this paper we present the calibration of a regional distributed variably-saturated water flow model for a 1400 km^2 agricultural area in the western San Joaquin Valley, CA. We use two global optimization algorithms to identify model parameters using data on spatially-distributed water table depth measurements and water district-average data on annual groundwater pumping and annual subsurface drainage flow measurements for a part of study area. In the first approach, these three measurement types are merged in a single-objective function for optimization with the SCEM-UA algorithm of Vrugt et al. (2003a). In the second approach, the independence of the three types of measurements is maintained by formulating a three-objective optimization problem, which is solved with the Multi-Objective SCEM-UA, or MOSCEM-UA, algorithm of Vrugt et al. (2003b). We discuss in detail the results of applying the single and multi-objective optimization algorithms to the calibration of the distributed water flow model, with optimized simulation results compared to measured water tables, groundwater pumping, and subsurface drainage for the 1984–1997 period. The focus here is on the relative merits of the single and multi-criteria approaches for model calibration. Independent evaluation of the calibrated model against data not used in the calibration will be part of a subsequent study.

2. Methods

2.1. Study area

The hydrologic domain of this study represents a 1400 km^2 irrigated agricultural region in western Fresno County on the westside of the San Joaquin Valley (Fig. 1A). The Mediterranean type climate of the area is semi-arid, with dry and hot summers, and foggy, rainy winters (Kahrl, 1979). Farms are organized into larger water districts for water distribution and drainage management (Fig. 1A). The area is topographically flat with southwest–northeast sloping deposits with slopes less than 1%.

The hydrogeology was described by Belitz and Heimes (1990). The study area is underlain by the Corcoran clay, a clay deposit that extends throughout

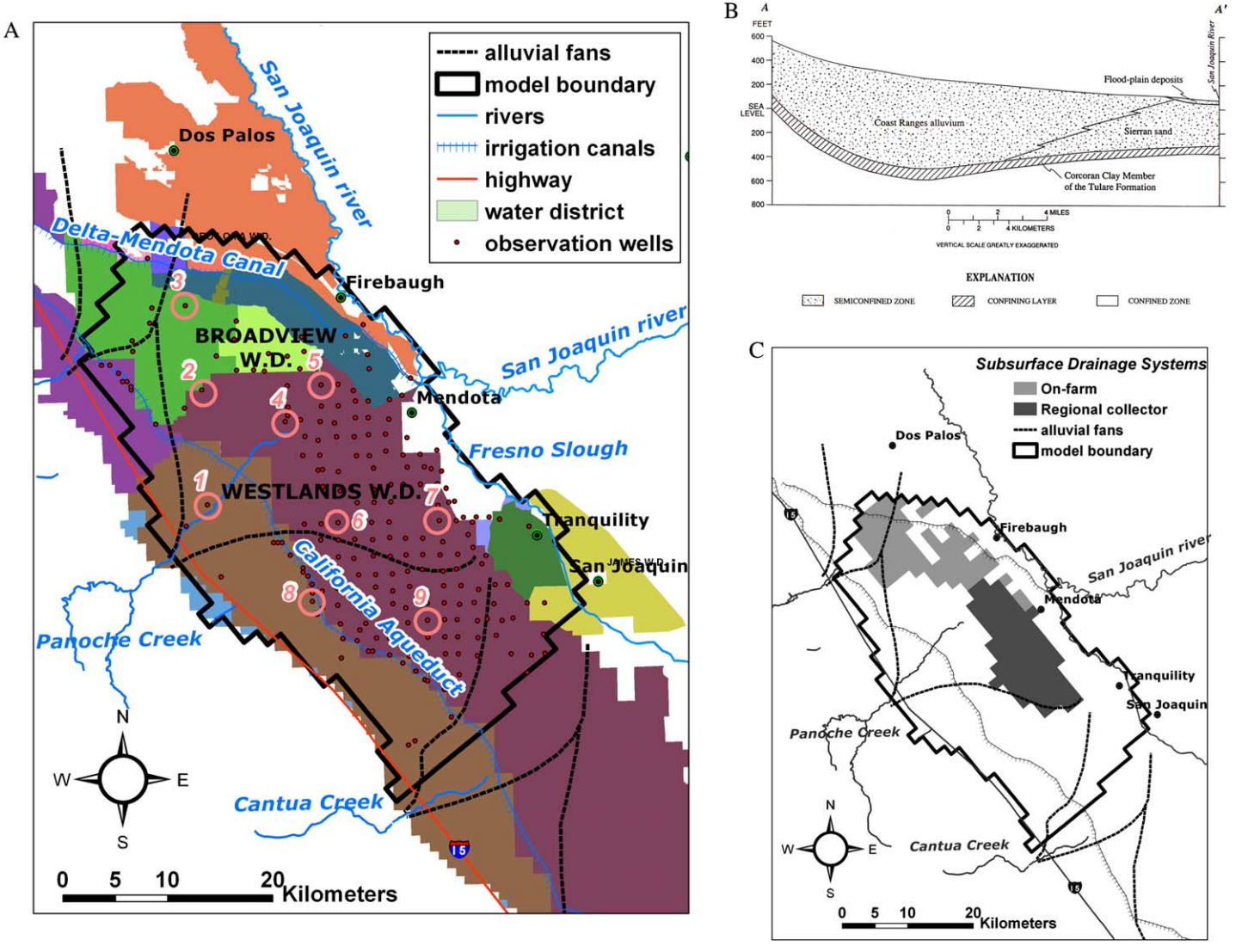

Fig. 1. (A) Model domain, water districts, and observation wells; (B) geologic cross-section along the AA′ transect shown in Fig. 1A (from Belitz et al., 1993); and (C) location of subsurface drainage systems.

the westside, with depths ranging from 244 (valley margin) to 30 m (near river) below the soil surface, as shown in Fig. 1B. The Corcoran clay divides the groundwater system into an upper semi-confined zone and a lower confined zone. The semi-confined zone consists of three hydrogeologic units: Coast Range alluvium, Sierran sand, and flood-plain deposits. The alluvial soils derived from Coast Range alluvium are generally fine-textured soils. The main soil types are clay (52% of the study area), clay loam (35%), loam (4%), and sandy loam (9%). The heaviest soils are found in the valley trough near the San Joaquin River. These soils have clay contents from 40 to 60%. Cotton is the major crop, with lesser areas planted in tomatoes, melons, vegetables, and orchards. Subsurface tile drains were installed in nearly 135,000 acres (55,000 ha) to control shallow water tables (Fig. 1C). On-farm drainage systems consist of parallel networks of perforated drain laterals, typically 1.8–2.7 m below the surface, and spaced horizontally from 30 to 180 m apart (Fio, 1994). The main source of irrigation water is surface water originating from the Sacramento Valley in northern California. In dry years, groundwater supplements surface water for irrigation. Groundwater pumping for irrigation occurs primarily from the confined aquifer below the Corcoran clay (Belitz et al., 1993; Gronberg and Belitz, 1992).

Groundwater levels were measured in 242 monitoring wells (Fig. 1A), from which annual October readings will be used here to calibrate the model (2236 measurements). Water level measurements were also available for April and July of each year. Although these data provide further information on the within-year flow dynamics, they were not used in the calibration, since the interest here is mainly in long-term prediction (e.g. can we predict the impact of a drought?). Additional measurements that are used for the model calibrations include total annual groundwater pumping data for Westlands Water District, WWD (13 measurements), and total annual subsurface drainage rates from BWD (11 measurements).

2.2. Hydrologic model

The numerical model used is MOD-HMS (Panday and Huyakorn, 2004). This model is a MODFLOW-based distributed watershed model, by which the three-dimensional variably-saturated subsurface flow is solved for the following equation (Panday and Huyakorn, 2004)

$$\frac{\partial}{\partial x}\left(K_{xx}k_{\mathrm{rw}}\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_{yy}k_{\mathrm{rw}}\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_{zz}k_{\mathrm{rw}}\frac{\partial h}{\partial z}\right)-W=n\frac{\partial S_{\mathrm{w}}}{\partial t}+S_{\mathrm{w}}S_{\mathrm{s}}\frac{\partial h}{\partial t} \quad (1a)$$

In Eq. (1a), x, y, and z are Cartesian coordinates [L]; t is time [T]; K_{xx}, K_{yy}, and K_{zz} are the principal components of saturated hydraulic conductivity along the x, y, and z axes, respectively [L/T]; k_{rw} is the relative hydraulic conductivity that is a function of degree of water saturation [-], $S_{\mathrm{w}}=\theta/n$, where θ is volumetric moisture content and n is porosity; h is total hydraulic head [L], so that $h=\psi+z$, where ψ and z denote the soil water pressure head and gravitational head [L], respectively, with z defined positive upwards; W is volumetric water flux per unit volume, representing sources and/or sinks [1/T]; and S_{s} is the specific storage of the saturated porous material [1/L]. For each grid cell, the left-hand side represents boundary fluxes due to: (1) head gradients as described by Darcy's law, and (2) sources (e.g. injection well) and sinks (e.g. evapotranspiration, drainage systems). The right-hand side describes changes in storage due to: (1) saturation/desaturation of the porous medium (first term), and (2) compressibility of the water and porous medium (second term). In unsaturated flow, the second term is usually neglected, whereas the first term is omitted in groundwater applications. Eq. (1a) is solved numerically using a mass-lumped fully implicit finite difference method with adaptive time stepping.

In this application, the sink term W consists of three terms

$$W=\frac{1}{\Delta z}(\mathrm{ET_a}+Q_{\mathrm{p}}+Q_{\mathrm{d}}) \quad (1b)$$

where Δz is thickness [L] of the grid cell, $\mathrm{ET_a}$ is evapotranspiration rate [L/T], Q_{p} is pumping rate [L/T], and Q_{d} is subsurface drainage rate [L/T]. Note that these three terms are written as fluxes, defined as volumetric flow rate per unit area of grid cell. Their calculations are discussed in Section 2.2.2.

The saturation dependence of soil water pressure head and relative permeability are typically described

by the constitutive expressions of van Genuchten (1980). The nonlinearities arising from the dependence of K and ψ on water saturation are determined with Newton–Raphson linearization, leading to a large computational burden when simulating three-dimensional variably-saturated flow for large domains. Furthermore, given the large uncertainties when simulating at the field or regional scale, the soil hydraulic parameters may not be well identified (Vrugt et al., accepted for publication). Alternatively, the soil hydraulic functions can be linearized at the grid-scale, assuming vertical equilibrium (Huyakorn et al., 1994), and are referred to as 'pseudo-soil' functions (HydroGeologic, 2001). In this approach the nonlinear water retention and conductivity functions at a point are replaced by discrete functions, with degree of saturation and relative hydraulic conductivity equal to zero when the soil water pressure is negative and equal to one when the pressure head is positive. These point values are integrated across the thickness of the grid cell that contains the water table to yield linear soil hydraulic functions. These linear grid-scale representative functions define saturation, S_w, and relative *horizontal* conductivity, k_{rw}, values that increase linearly from 0, when the water table is at or below the bottom of the grid cell, to 1 when the water table is at or above the top of the cell, or

$$\begin{aligned} &S_w = k_{rw} = 1, \ \text{for } \psi/\Delta z \geq 0.5; \\ &S_w = k_{rw} = 0.5 + \frac{\psi}{\Delta z}, \ \text{for } -0.5 < \psi/\Delta z < 0.5; \\ &S_w = k_{rw} = 0, \ \text{for } \psi/\Delta z \leq -0.5 \end{aligned} \qquad (2)$$

where Δz is thickness [L] of the grid cell with the water table and ψ is the pseudo-pressure head [L] at the node (Huyakorn et al., 1994). In the vertical direction, k_{rw} is always equal to 1. Note that Eq. (2) is identical to the vertical equilibrium assumption of the depth-integrated two-dimensional Boussinesq equation, i.e. water above the water table is assumed to be in hydrostatic equilibrium. The linearization of the water retention curve in Eq. (1a) results in a moisture-independent, depth-integrated soil-specific water capacity described by the specific yield parameter, S_y [-], with a value equal to n. The use of pseudo-soil functions constitutes a computationally attractive compromise between the rigorous variably-saturated flow modeling using the *van Genuchten* relationships, and the simplified MODFLOW approach for which cells become inactive when the water table drops below the bottom of the cell (McDonald and Harbaugh, 1988). In the approach used here, when the water table drops below the bottom of a grid cell, Eq. (1a) is still solved but with the right-hand side equal to zero, i.e. changes in storage above the water table are neglected. This procedure avoids convergence problems with (in)activation of cells encountered in MODFLOW (Doherty, 2001).

2.2.1. Spatial and temporal discretization

The horizontal boundaries of the model domain coincide with the hydrologic boundaries of the regional groundwater flow model of Belitz et al. (1993) and Belitz and Phillips (1995), defined by the San Joaquin river on the east, the Coast Range foothills in the west, and the no-flow boundaries in the north and south of the regional flow domain. The model domain was discretized into a regular finite difference grid of 2960 square cells of 804.65 m (0.5 mi) side length and 64 ha (160 acres) areas, corresponding with a typical field size, resulting in 74 rows and 40 columns. This grid size is half of the discretization used by Belitz et al. (1993). The field grid size assumes that the impact of smaller-scale soil variations within each field on subsurface water flow are small, relatively to the flow effects by differences in surface boundary conditions (Refsgaard et al., 1999).

In the vertical direction, the model domain extends from the land surface to the top of the Corcoran clay (Fig. 1B), with surface elevations and clay depths defined by Belitz et al. (1993). The vadose zone-groundwater flow system above the Corcoran clay was divided into six layers of increasing thickness from the surface downwards with the first 2-m thick layer representing the crop root zone and the second and third layer with sizes of 4.5 and 10 m. The three bottom layers (4, 5, and 6) were of variable thickness, assigned 3/16, 5/16, and 8/16 of the remaining depth interval. Consequently, the total number of model grid cells was equal to 17,760. However, part of the model domain was inactive because of cells that accommodate the irregular natural boundaries, so that the total number of active model cells was 12,720. Fig. 1A also specifies the locations of the various water districts

(WD) within this geographic area, specifically WWD and BWD along the SE–NW central axis.

The simulation period was 13 years, starting in October 1984 and ending in September 1997. This period was selected for calibration because: (1) it contains the most and best available data, and (2) it includes a drought period with increased fallowing and groundwater pumping. Hence, model performance during this period provides a good test since a range in hydrologic conditions is encountered. The focus will be on the relative merits of the single and multi-criteria approaches for model calibration. Independent evaluation of the calibrated model against data not used in the calibration will be part of a subsequent study. The entire simulation period was discretized into 13 annual time intervals for which annual-average boundary conditions were specified for each grid block. The time steps of the computations were much shorter, typically weekly to monthly, with the step controlled by the convergence of the numerical solution at each time step.

2.2.2. Initial conditions, boundary conditions, sinks and sources

2.2.2.1. Initial condition. The spatial distribution of initial head was determined from spatially-interpolated groundwater table depth measurements of October 1984, assuming hydrostatic conditions. The following boundary condition fluxes apply to the field or grid scale.

2.2.2.2. Crop evapotranspiration and soil evaporation. For each year of the simulation, crop acreage data were collected for each WD. Crops, including fallow, were randomly distributed over all cells within a district with the simulated total crop acreage equal to that reported, thus leading to the annual assignment of a single crop to each grid cell. Little fallowing generally occurs in wet years, but during droughts up to 50% of irrigable land may be fallowed. The standard crop coefficient method was used for estimating the average potential crop evapotranspiration, ET_c, (Allen et al., 1998) during the growing season for 25 different crops

$$ET_c = f_{ET} K_c ET_0 \quad (3)$$

where K_c is the crop coefficient [-] and ET_0 denotes the reference evapotranspiration for well-watered completely-covered grass [L/T]. Local information on time-varying crop-specific crop coefficients, planting dates, and other plant data were taken from Snyder et al. (1989). Daily values for reference evapotranspiration, ET_0, were obtained from the Firebaugh CIMIS weather station and applied uniformly over the entire study area. This is justified because the study area falls completely within one ET_0 zone (zone 15), as mapped by the California Department of Water Resources (http://wwwcimis.water.ca.gov/cimis/images/etomap.jpg). An additional crop correction factor, f_{ET} [-], was added, see Eq. (3), to account for uncertainties associated with estimating annual field-scale crop water demand using this approach, for example, because of water stress and non-uniform crop development. Its single value was subject to calibration. Hence, it is assumed that non-uniformity in crop development is independent of crop type.

Average soil evaporation outside the growing season, E_a [L/T], was calculated as the sum of contributions from infiltrated rain and from evaporation from a shallow water table. Both terms were estimated from a suite of HYDRUS-1D simulations (Šimůnek et al., 1998), using a wide range of fixed water table depths and the four main soil types of the study area (Table 1). These HYDRUS-1D simulations used daily boundary conditions for rainfall and potential evaporation, and soil hydraulic functions for the main soil types estimated using the ROSETTA pseudo-transfer functions (Schaap et al., 1998). A piecewise linear model as a function of water table depth, z_{wt}, was fitted to the annual evaporation values simulated with HYDRUS-1D, resulting in soil-specific values of z_{wtmin}, water table depth at which evaporation is not limited by the soil hydraulic

Table 1
Evaporation parameters for the main soil types of the study area

Soil type	Areal fraction	z_{wtmin} (cm)[a]	z_{wtmax} (cm)[a]
Clay	0.52	20.7	203.6
Clay loam	0.35	9.1	192.0
Loam	0.04	55.2	250.5
Sandy loam	0.09	27.7	195.1

[a] z_{wtmin} is the water table depth at which evaporation is not limited by the soil hydraulic conductivity; z_{wtmax} is the water table depth at which groundwater table contribution to soil evaporation stops.

conductivity, and z_{wtmax}, the water table depth at which groundwater table contribution to soil evaporation stops (Table 1). The loam soil has the largest values for z_{wtmin} and z_{wtmax}, indicating that it has the largest potential for capillary rise. The smallest values are listed for the clay loam soil. Evaporation of infiltrated rain was estimated from the HYDRUS-1D simulations with a deep water table, where capillary rise was not present.

Total annual evapotranspiration ET_a [L/T] was computed from:

$$ET_a = E_a + ET_c \tag{4}$$

Total annual evapotranspiration is applied as a sink term in Eq. (1b) for the top model layer only.

2.2.2.3. Rainfall, irrigation, and groundwater pumping. Average annual rainfall, R [L/T], for the simulation period 1984–1997, as determined from Hanford weather station data, was 203 mm/year. This is similar to the average rainfall in the Firebaugh CIMIS station. Since most rain evaporates, its value is not critical for the purpose of this study. Rainfall was applied uniformly over the entire domain, and varied annually. The main irrigation water source is surface water for which annual total values were available for each district of the study area. However, local pumped groundwater is used as a complementary water source in times of shortages. This water is pumped mainly from the confined aquifer below the Corcoran clay. Since only district-aggregated values of surface water deliveries were available and pumped groundwater amounts were unknown, we needed a method to distribute total applied water across the fields within each district. To estimate field-scale irrigation water applications (AW), we developed the following procedure. First, the annual irrigation requirement, IR [L/T], for each grid cell or crop was computed from

$$IR = \mathrm{Max}\left\{0, \frac{ET_c - R_{gs}}{IE}\right\} \tag{5}$$

where IE denotes the seasonal field-scale irrigation efficiency [-], and R_{gs} is rainfall within the crop's growing season [L/T]. Thus, IR depends on crop type and varies between grid cells and years. The regional study of Gronberg and Belitz (1992) concluded that IE was highly correlated with water table depth, because farmers in areas with shallow water tables irrigate more efficiently because capillary rise satisfies part of the crop water demand in these areas, whereas over-irrigation may lead to soil aeration problems. As pointed out by Ayars and Schrale (1991), IE values tend to be higher for fine-textured soils with low hydraulic conductivity. We assumed that IE is defined as a piecewise linear function of water table depth, or

$$\begin{aligned} &IE = IE_{deep}, \quad \text{for } z_{wt} > 7 \text{ m};\\ &IE = IE_{shallow} + \frac{z_{wt} - 1}{7 - 1}(IE_{deep} - IE_{shallow}),\\ &\quad \text{for } 1 < z_{wt} < 7 \text{ m};\\ &IE = IE_{shallow}, \quad \text{for } z_{wt} < 1 \text{ m} \end{aligned} \tag{6}$$

where $IE_{shallow}$ and IE_{deep} represent the maximum and minimum values of irrigation efficiency that were subjected to calibration. Reported values of district-scale surface water deliveries (SW_d) were disaggregated to the field scale, assuming that surface water deliveries were proportional to field-scale irrigation requirements, to estimate the field-scale surface water allocation, SW [L/T], or

$$SW = IR\frac{SW_d}{\sum_d IR} \tag{7}$$

where the summation denotes the irrigation requirement of all grid cells within water district d. Recycled drainage water in BWD, DW [L/T], was distributed similarly. The contribution of groundwater pumping GW [L/T] was computed by difference, or

$$\begin{aligned} &GW = \mathrm{Max}\{0, IR - SW - DW\} \quad \text{and}\\ &GW_d = \sum_d GW \end{aligned} \tag{8}$$

where GW_d defines the annual district groundwater pumping [L/T], computed from the summation over all grid cells within water district d. The total annual amount of water applied for each field, AW [L/T], was determined from

$$AW = R + SW + GW + DW \tag{9a}$$

where R is annual rainfall [L/T]. Gronberg and Belitz (1992) estimated the spatial distribution of the percentage of groundwater pumping from the

semi-confined and confined zones. The amount of water removed from the semi-confined zone by pumping is then

$$Q_p = f_p \mathrm{GW} \tag{9b}$$

where f_p is the fraction of pumping from the semi-confined zone. Since the confined zone is not included in the model, Q_p is applied as a sink term in Eq. (1b). Pumped groundwater is taken from the bottom model layer only. The average rate of groundwater pumping in WWD will be compared to measurements and was computed from

$$Q_{\mathrm{pump}} = \frac{1}{n_p} \sum_{j=1}^{n_p} \mathrm{GW} \tag{9c}$$

where n_p is the number of grid cells within WWD.

2.2.2.4. Infiltration, ponding, and run-off. The MODHMS model partitions the applied irrigation and rainfall water into infiltration, ponding, and run-off, using a user-specified maximum ponding depth of 10 cm. For our simulations, the amount of water removed from the system by runoff was less than 0.5% of the total water balance.

2.2.2.5. Subsurface drains. Subsurface drains were primarily installed in Westlands Water District (WWD), and in Panoche, Broadview and Firebaugh Canal water districts (Fig. 1C). Some of these drains were installed as early as 1958. Since that time, the drained area followed the areal expansion of the shallow water table (Fio, 1994). The regional-collector drains in WWD were installed in an area of about 42,000 acres, and became operative in 1980. However, the drains were closed in 1985 due to high selenium concentrations in the drainwater. North of WWD is an area underlain by on-farm drains. Drain flow was modeled as a linear head-dependent sink

$$Q_d = C_d[h - (z_0 - d_e)] \tag{10a}$$

where Q_d is the grid-cell drainage flux [L/T]; C_d is drain conductance [1/T]; h is total hydraulic head in the grid cell [L]; z_0 is the land surface elevation [L], and d_e is an effective drain depth [L]. The single-valued dram parameters, C_d and d_e were subject to calibration. Q_d is applied as a sink term in Eq. (1b). Depending on the drain depth, subsurface drainage water is taken from one of the two upper layers. The average rate of subsurface drainage in BWD will be compared to measurements and is calculated as

$$Q_{\mathrm{drain}} = \frac{1}{n_d} \sum_{j=1}^{n_d} Q_d \tag{10b}$$

where j is an index looping over all grid cells in BWD and n_d is the number of grid cells within BWD.

2.2.2.6. Bottom boundary. The bottom of the model domain coincides with the Corcoran clay, which is underlain by a confined aquifer system used for groundwater pumping. Seepage through the Corcoran clay was modeled using the following linear head-dependent flux boundary condition

$$Q_b = C_b(h_b - h) \tag{11}$$

where Q_b is water flux across the bottom boundary [L/T]; C_b is the corresponding conductance [1/T]; and h and h_b denote the total hydraulic heads in the bottom grid cell and the confined aquifer below it [L], respectively. The bottom conductance, C_b, was calculated from

$$C_b = \frac{K_{\mathrm{Corc}}}{L_{\mathrm{Corc}}} \tag{12}$$

where L_{Corc} [L] and K_{Corc} [L/T] are the thickness and saturated hydraulic conductivity of the Corcoran clay, respectively. Thus, specification of the lower boundary condition required values of L_{Corc}, K_{Corc}, and the hydraulic head in the confined aquifer, h_b. Spatially-distributed thickness values of the Corcoran clay ranged between 10 and 30 m, and were taken from Belitz et al. (1993). Values for h_b were varied in space and time based on spatially-distributed confined well data for several years (1980, 1985, 1989, 1995, and 1997). Values for K_{Corc} were reported by Belitz et al. (1993) and Johnson et al. (1968), ranging from 5.8 to 0.009–0.06 cm/year, whereas historical hydraulic head observations across the Corcoran clay suggested very low seepage rates (Davis and Coplen, 1989). In this study, K_{Corc} was subject to calibration.

2.2.2.7. Lateral boundaries. The lateral flow boundary conditions were simulated as in Belitz et al. (1993). The western boundary of the model (Fig. 1A) near the Coast Range was a no-flow boundary condition.

Since both the northern and southern boundaries are parallel with observed flow lines of Little Panoche Creek and Cantua Creek, respectively, also these were treated as no-flow boundaries. Finally, also the eastern and northeastern boundaries from the soil surface down to the 6.5 m depth were treated as no flow boundaries, because they consist of low-conductivity flood-basin clays. However, the same boundary below the 6.5 m depth with the Sierran sand was treated as a head boundary, or

$$Q_l = C_l(h_l - h) \tag{13}$$

where Q_l is lateral water flux across the model boundary [L/T]; C_l is the lateral conductance [1/T]; and h and h_l denote the total hydraulic head in the grid cell and east of the grid cell [L], respectively. Values for h_l were based on historic well data from wells located to the east of the model domain. The lateral conductance was determined from

$$C_l = \frac{K_h}{L_l} \tag{14}$$

where L_l [L] is the distance between the two boundary nodes, set to 3000 m (Belitz et al., 1993) and K_h [L/T] denotes the saturated horizontal hydraulic conductivity of the Sierran deposits, which was subjected to calibration, as discussed in Section 2.2.3.

2.2.3. Hydraulic properties

The spatial distribution of vertical, K [L/T] and horizontal saturated hydraulic conductivity, K_h, values was determined from interpolated coarse fractions, based on textural data from 446 well logs and 23 electrical geophysical logs (Laudon and Belitz, 1991; Belitz et al., 1993). Hydraulic conductivity values were calculated as power averages of the coarse/fine end-member conductivities, K_F and K_C, (Desbarats, 1998), or

$$K_{h,v} = [K_C^{\omega} f_C + K_F^{\omega}(1 - f_C)]^{1/\omega} \tag{15}$$

where f_C [-] defines the cell's coarse fraction and ω is the averaging exponent. We selected arithmetic averaging for the horizontal direction ($\omega = 1$) and geometric averaging for the vertical direction ($\omega = -1$). Different K_C values were specified for the coarse-grained deposits derived from the Coast Ranges $K_{C,cr}$, and for Sierran sand deposits, $K_{C,ss}$. The spatial distribution of the Sierran sand deposits was based on Belitz et al. (1993). Data from slug tests (Belitz et al., 1993) resulted in $K_{C,cr}$ ranges in the range of 1500–1.2×10^6 cm/year, with an average of 3.5×10^5 cm/year, whereas $K_{C,ss}$ values ranged from 8.5×10^4 to 2.8×10^6 cm/year, with an average of 1.2×10^6 cm/year. Both average K_C values were scaled by (i.e. divided by) a common scaling factor, f_{Kc}, which was calibrated. The value of K_F was also calibrated.

Using the pseudo-soil functions of Eq. (2) requires specification of two storage-related effective parameters: specific storage, S_s [1/L], and specific yield, S_y [-]. Spatially-distributed values for S_y were computed from the distribution of coarse fractions, f_C, and end-member values, $S_{y,C}$ and $S_{y,F}$, or:

$$S_y = S_{y,C} f_C + S_{y,F}(1 - f_C) \tag{16}$$

Values for both $S_{y,C}$ and $S_{y,F}$ were subject to calibration. Finally, we assigned a spatially-uniform value for specific storage, S_s, equal to 10^{-7} cm^{-1} (Belitz et al., 1993). Table 2 presents all calibration parameters, including their prior ranges.

For each realization of the parameter optimization procedure, the simulation model was run for the 13-year period using all parameters in Table 2. However, since rates of evaporation, applied water, and groundwater pumping is controlled by table depth, we implemented the following procedure. First, the groundwater model was run for a single year, and the corresponding calculated water table depth values were used to update the evaporation, applied water, and groundwater pumping rates for the next year. Subsequently, the model was run for the following year with these new rates, until all 13 years were simulated.

2.3. Parameter optimization

Two different optimization methods were used to calibrate the regional flow model. The first approach uses a single-objective function, consisting of a weighted sum of squared error terms that represent residuals of groundwater table levels, groundwater pumping and subsurface drainage rates. The second, multi-objective approach optimizes three separate

Table 2
Calibration parameters and their prior uncertainty ranges

Parameter	Description	Units	Minimum	Maximum	Sampling strategy
K_F	Hydraulic conductivity of the fine fraction	m/year	0.3	1.0	Uniform log values
$S_{y,C}$	Specific yield of the coarse fraction	–	0.05	0.35	Uniform
$S_{y,F}$	Specific yield of the fine fraction	–	0.05	0.40	Uniform
K_{Corc}	Hydraulic conductivity of the Corcoran clay	m/year	0.001	0.10	Uniform log values
f_{ET}	ET correction coefficient	–	0.8	1.2	Uniform
$IE_{shallow}$	Maximum irrigation efficiency	–	0.7	1.2	Uniform
IE_{deep}	Minimum irrigation efficiency	–	0.5	0.8	Uniform
f_{Kc}	Scaling factor for the hydraulic conductivity of the coarse fraction	–	0.5	5	Uniform
d_e	Drain depth	m	1.5	3.0	Uniform
C_d	Drain conductance	1/year	0.45	4.5	Uniform

objective functions for each of the three data types simultaneously.

The Shuffled Complex Evolution Metropolis (SCEM-UA) global optimization algorithm of Vrugt et al. (2003a) was used in the single-objective approach, as was most recently done in Vrugt et al. (accepted for publication) to calibrate vadose zone parameters of the BWD. A detailed description of the SCEM-UA algorithm is given in Vrugt et al. (2003a). The objective function was defined as

$$\mathrm{OF} = \frac{1}{n_{\mathrm{wt}}}\sum_{i=1}^{n_{\mathrm{wt}}}\left[\frac{z_{\mathrm{wt,OBS}}(i)-z_{\mathrm{wt,SIM}}(i)}{z_{\mathrm{wt,OBS}}(i)}\right]^2 + \frac{1}{n_{\mathrm{pump}}}\sum_{i=1}^{n_{\mathrm{pump}}}\left[\frac{Q_{\mathrm{pump,OBS}}(i)-Q_{\mathrm{pump,SIM}}(i)}{Q_{\mathrm{pump,OBS}}(i)}\right]^2 + \frac{1}{n_{\mathrm{drain}}}\sum_{i=1}^{n_{\mathrm{drain}}}\left[\frac{Q_{\mathrm{drain,OBS}}(i)-Q_{\mathrm{drain,SIM}}(i)}{Q_{\mathrm{drain,OBS}}(i)}\right]^2 \quad (17)$$

where Q_{pump} is the average rate [L/T] of groundwater pumping in WWD, defined in Eq. (9c), Q_{drain} is the average rate [L/T] of subsurface drainage in BWD, defined in Eq. (10b), and n_{wt}, n_{pump}, and n_{drain} are the number of observations for groundwater table depth (2236), groundwater pumping (13), and subsurface drainage (11), respectively. Thus, the SCEM-UA derived parameter estimates and uncertainties, as estimated from minimizing Eq. (17), are determined such that the various measurement types are equally weighted. We note that the subjective weighting by observation number and magnitude is not in full accordance with the classical Bayesian approach of the SCEM-UA algorithm. In the Bayesian approach the likelihood function to be minimized is normalized only by the measurement error. However, to reduce the sensitivity of the optimization procedure to the relatively larger number of water table measurements, we implemented Eq. (17) instead of the classical Bayesian likelihood function.

While in the single-objective SCEM-UA approach, weighting of the different objectives into a single aggregated scalar is necessary, multi-objective optimization allows the user to treat each objective function independently, so that no subjective weighting of the objectives needs to be defined a priori. The multi-objective optimization problem is defined as follows (Gupta et al., 2003)

$$\text{Minimize } F(\boldsymbol{p}) = \langle f_1(\boldsymbol{p}), f_2(\boldsymbol{p}), \ldots, f_N(\boldsymbol{p})\rangle \text{ with respect to } \boldsymbol{p}$$

where F is the vector of objectives, $f_k(\boldsymbol{p})$ is the kth objective function ($k=1\ldots N$), and $\boldsymbol{p}$ is a vector of model parameters. In general, the solution to this problem will not be a single 'best' parameter set, but will consist of a Pareto set, $P(\boldsymbol{p})$, of solutions corresponding to trade-offs among the objectives. Formally, the Pareto set consists of parameter combinations $\boldsymbol{p}_i$ with the following properties: (1) for all non-members $\boldsymbol{p}_j$ there exists at least one member $\boldsymbol{p}_i$ such that $P(\boldsymbol{p}_i)$ is strictly less than $P(\boldsymbol{p}_j)$, and (2) it is not possible to find a member $\boldsymbol{p}_j$ within the Pareto set such that $P(\boldsymbol{p}_j)$ is strictly less than $P(\boldsymbol{p}_i)$. By definition, $P(\boldsymbol{p}_i)$ is strictly less than $P(\boldsymbol{p}_j)$ if, for all $k=1\ldots N$, $f_k(\boldsymbol{p}_i) < f_k(\boldsymbol{p}_j)$. The multi-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. Parameter sets are assigned a ranking based on their performance of these objectives, and it is this ranking rather than the value of a single-objective function that is used to evolve the population towards the Pareto trade-off surface. 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. We simultaneously minimized three-objective functions, corresponding to the RMSE values of groundwater table depths [L], groundwater pumping [L/T], and subsurface drainage [L/T]:

$$\mathrm{RMSE}_{\mathrm{pump}} = \sqrt{\frac{1}{n_{\mathrm{pump}}}\sum_{i=1}^{n_{\mathrm{pump}}}[Q_{\mathrm{pump,OBS}}(i) - Q_{\mathrm{pump,SIM}}(i)]^2} \tag{18a}$$

$$\mathrm{RMSE}_{\mathrm{drain}} = \sqrt{\frac{1}{n_{\mathrm{drain}}}\sum_{i=1}^{n_{\mathrm{drain}}}[Q_{\mathrm{drain,OBS}}(i) - Q_{\mathrm{drain,SIM}}(i)]^2} \tag{18b}$$

$$\mathrm{RMSE}_{\mathrm{wt}} = \sqrt{\frac{1}{n_{\mathrm{wt}}}\sum_{i=1}^{n_{\mathrm{wt}}}[z_{\mathrm{wt,OBS}}(i) - z_{\mathrm{wt,SIM}}(i)]^2} \tag{18c}$$

Both the single (SCEM-UA) and multiple (MOSCEM-UA) optimization algorithms were conducted with an initial population of 500 parameter sets, randomly selected from the prior ranges defined in Table 2. This initial sample sets was divided into five complexes. Each optimization was based on a total of 4000 model evaluations, which is an adequate number of simulations to obtain an initial approximation of the Pareto set with a reasonable level of computational effort. Each optimization consisting of 4000 model evaluations required approximately 70 h of CPU time on a 1.5 GHz desktop computer.

In addition to the three objectives listed above, we also calculated RMSE and BIAS values for an additional six groundwater table classes, dividing groundwater table depths in groups of different depth intervals. For example, model bias for pumping rates was computed (%) by:

$$\mathrm{BIAS} = \frac{\sum_{i=1}^{n_{\mathrm{pump}}}[Q_{\mathrm{pump,OBS}}(i) - Q_{\mathrm{pump,SIM}}(i)]}{\sum_{i=1}^{n_{\mathrm{pump}}} Q_{\mathrm{pump,OBS}}(i)} \times 100\% \tag{19}$$

Note that according to the definition of BIAS in Eq. (19), a negative BIAS indicates that the model overpredicts, whereas an underprediction results in a positive BIAS.

3. Results and discussion

3.1. Single-objective optimization results

The optimum parameter set after 4000 simulations with the smallest OF value is presented in Table 3. The values of these optimum parameters may be compared to other studies, in particular the work of Belitz et al. (1993), as listed in Table 3. In general, the optimum parameter values are similar to the values estimated by Belitz et al. (1993), however, the parameters are not always directly comparable. The value for K_F is somewhat smaller than the one in Belitz et al. (1993), however, here geometric averaging, Eq. (15), was used for vertical conductivity instead of harmonic averaging by Belitz et al. (1993). Since geometric averaging gives more weight to the coarse fraction conductivity, a smaller value for K_F may give similar values for vertical conductivity. The values for K_C are smaller than the ones estimated by Belitz et al. (1993) by a factor of 3.26, resulting in smaller horizontal (mainly) and vertical conductivities, and in less water leaving the domain as horizontal flow through the eastern boundary of the model. Belitz et al. (1993) calibrated average values for specific yield above and below 7 m equal to 0.30 and 0.20, respectively. The corresponding values above and below 7 m obtained here are 0.31 and 0.23. Note that these values were obtained by using average coarse fractions for the sediments above and below 7 m (9.6 and 36%, respectively) in Eq. (16), together with calibrated values of 0.06 and 0.34 for 'coarse' and 'fine' specific yield (Table 3). The conductivity of

Table 3
Optimal parameter sets for several objectives

Parameter	Units	Minimal OF, Eq. (17)	Minimal $RMSE_{wt}$	Minimal $RMSE_{pump}$	Minimal $RMSE_{drain}$	MO_1	MO_2	MO_3	MO_4	B&P[a]
K_F	m/year	0.33	0.31	0.57	0.37	0.39	0.38	0.58	0.63	0.44
$S_{y,C}$	–	0.06	0.14	0.05	0.30	0.13	0.30	0.17	0.08	0.20
$S_{y,F}$	–	0.34	0.37	0.06	0.31	0.40	0.31	0.40	0.40	0.30
K_{Corc}	m/year	0.001	0.016	0.003	0.002	0.001	0.005	0.018	0.009	0.06
f_{ET}	–	0.80	0.98	0.82	0.81	0.84	0.82	0.87	0.87	1.00
$IE_{shallow}$	–	1.19	0.92	1.17	0.74	1.17	0.83	1.15	0.98	0.80
IE_{deep}	–	0.54	0.55	0.56	0.68	0.61	0.67	0.52	0.55	0.65
f_{Kc}	–	3.26	4.35	2.64	4.31	4.64	4.77	4.92	4.67	1.00
d_e	m	2.55	2.74	2.07	2.46	2.36	2.46	2.26	1.66	2.20
C_d	1/year	0.45	1.38	0.67	0.49	2.08	0.59	2.31	2.03	0.59

MO_1–MO_4 denote optimal solutions obtained by the multi-objective optimization, as discussed in Section 3.4. MO_1, minimum Euclidean distance in the normalized space of the three main objectives; MO_2, minimum $RMSE_{wt<2}$ and absolute $BIAS_{wt<2}$ less than 5%; MO_3, minimal $RMSE_{2<wt<3}$ and absolute $BIAS_{2<wt<3}$ less than 5%; MO_4, minimum Euclidean distance in the normalized two-objective space of $RMSE_{wt<2}$ and $RMSE_{2<wt<3}$.

[a] Parameter values estimated by Belitz et al. (1993).

the Corcoran clay, K_{Corc}, is an order of magnitude smaller than the one calibrated by Belitz et al. (1993), but is more in line with the measurements of Johnson et al. (1968). The optimum value of 0.8 for the crop correction factor, f_{ET}, results in an average crop ET_c value for cotton, the main crop in the study area, equal to 0.58 m, which is similar to the value of 0.66 m used by Belitz et al. (1993). The value of 0.8 for f_{ET} also agrees with calibrated values obtained by Vrugt et al. (accepted for publication) in BWD, and with field observations of non-uniform crop development by Burt and Styles (1994) in the northern part of the study area. Belitz et al. (1993) used irrigation efficiencies equal to 0.80, 0.72, and 0.65 for the water table depth intervals 0–3, 3–7, and >7 m. The corresponding optimum values obtained here averaged over the same depth intervals are, respectively, 1.10, 0.73, and 0.54. In other words, compared to Belitz et al. (1993), the optimum solution obtained here results in a shift of applied irrigation water from the shallow water table area (0–3 m) to the deep water table area (deeper than 7 m). Irrigation efficiencies larger than one in the presence of a shallow water table area are consistent with experimental studies in the study area (e.g. Grismer and Gates, 1988). Finally, values for the drainage parameters, d_e and C_d, are similar to those of Belitz et al. (1993).

In Fig. 2A we present histograms for each of the 10 optimized parameters for the last 2000 parameter sets sampled by SCEM-UA, and the corresponding normalized parameter plots, respectively. A zero value of the normalized parameter means that its optimized value is equal to the lower bound of the a priori defined parameter range. Table 4 presents standard statistics for the sample of 2000 last-sampled parameter sets, including cross-correlation, mean, median, and CV values for each parameter. First, we note that the OF value reached an approximately constant value, already after the first set of 500 simulations. The resulting spread of optimized parameter values in Fig. 2a and b is quite small for all parameters. This result indicates that parameter uncertainty has been significantly reduced, when compared to the prior ranges specified in Table 2 and that the parameter values are highly identifiable. However, Fig. 2a and b also show that three parameters approach the lower bound of the physically-realistic prior distribution; these are K_{Corc} (Eq. (12)), the saturated hydraulic conductivity of the Corcoran clay, the crop correction factor f_{ET} (Eq. (3)), and the drain conductance, C_d (Eq. (10a)). There could be various reasons of why this occurred. First, the 'physically' possible prior ranges may be specified too narrow. Though we realize that the optimization of a small number of parameters for such a large spatial domain with complex soil heterogeneities and boundary conditions may lead to effective parameters that not necessarily obey the physically-based

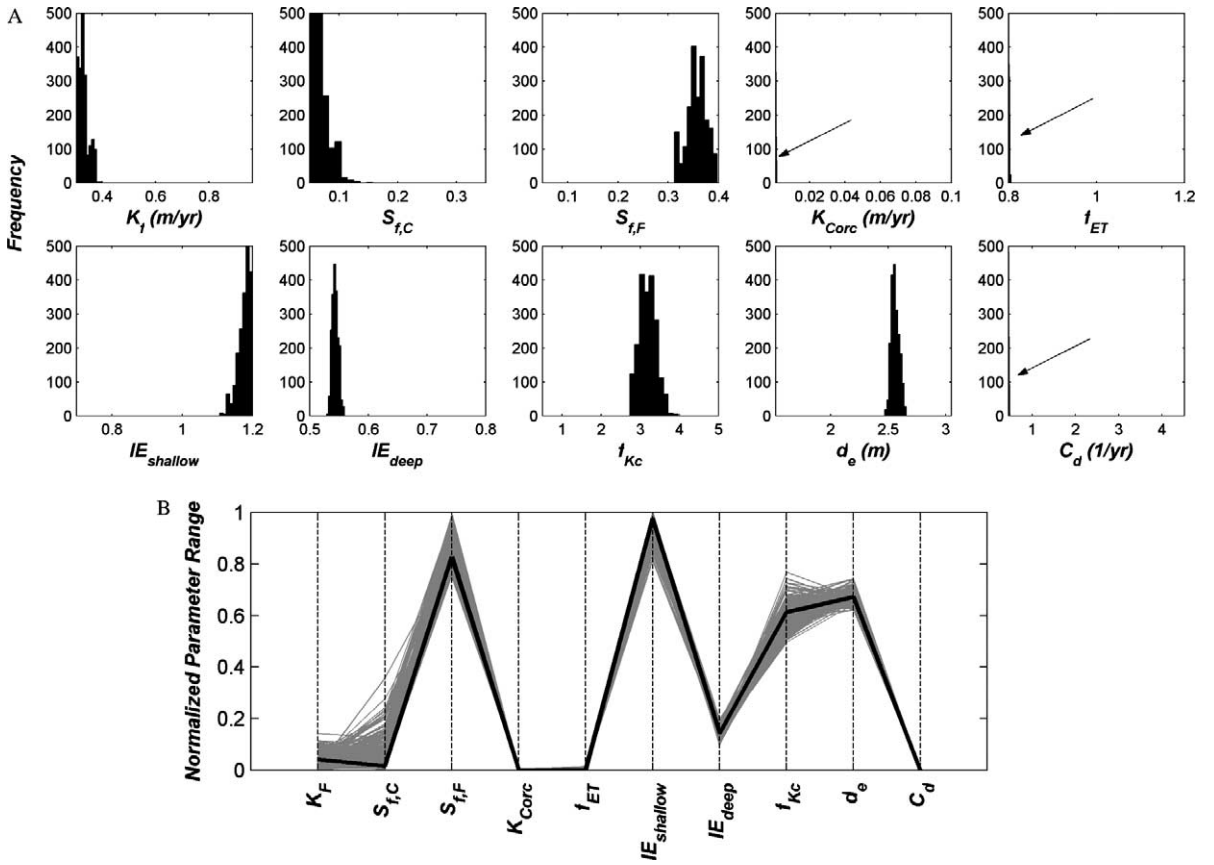

Fig. 2. (A) Histograms, and (B) normalized parameter plots for the last 2000 parameter sets sampled with the SCEM-UA algorithm; the black line indicates the optimum parameter set.

hydrologic laws, we believe that some parameters may move outside their physical range, to compensate for model structure, and input/output measurement errors. Second, parameters can attain unrealistic values if they are insensitive to the output or correlated. However, as the cross-correlation values in Table 4 show, they are generally very small as compared to other studies, such as Vrugt et al.

Table 4
Statistics of the last 2000 parameter sets sampled by the SCEM-UA algorithm

	Correlation coefficients											
	K_F	$S_{y,C}$	$S_{y,F}$	K_{Corc}	f_{ET}	$IE_{shallow}$	IE_{deep}	f_{Kc}	d_e	C_d	Median	Mean	CV (%)
K_F	1.00	−0.11	0.20	−0.02	0.12	0.04	0.08	−0.16	−0.53	0.17	0.33	0.33	5.55
$S_{y,C}$	−0.11	1.00	−0.22	0.22	−0.23	−0.28	−0.06	−0.03	0.02	−0.21	0.06	0.07	20.76
$S_{y,F}$	0.20	−0.22	1.00	−0.06	0.09	0.39	0.08	−0.14	−0.13	0.11	0.36	0.36	5.41
K_{Corc}	−0.02	0.22	−0.06	1.00	−0.14	−0.05	−0.22	0.19	0.04	−0.25	0.001	0.001	12.20
f_{ET}	0.12	−0.23	0.09	−0.14	1.00	0.11	0.20	−0.13	0.04	0.11	0.80	0.80	0.14
$IE_{shallow}$	0.04	−0.28	0.39	−0.05	0.11	1.00	0.01	0.14	0.07	−0.01	1.18	1.18	1.45
IE_{deep}	0.08	−0.06	0.08	−0.22	0.20	0.01	1.00	−0.15	−0.13	−0.06	0.54	0.54	0.99
f_{Kc}	−0.16	−0.03	−0.14	0.19	−0.13	0.14	−0.15	1.00	−0.15	−0.10	3.17	3.19	6.73
d_e	−0.53	0.02	−0.13	0.04	0.04	0.07	−0.13	−0.15	1.00	−0.12	2.56	2.56	1.32
C_d	0.17	−0.21	0.11	−0.25	0.11	−0.01	−0.06	−0.10	−0.12	1.00	0.46	0.46	1.40

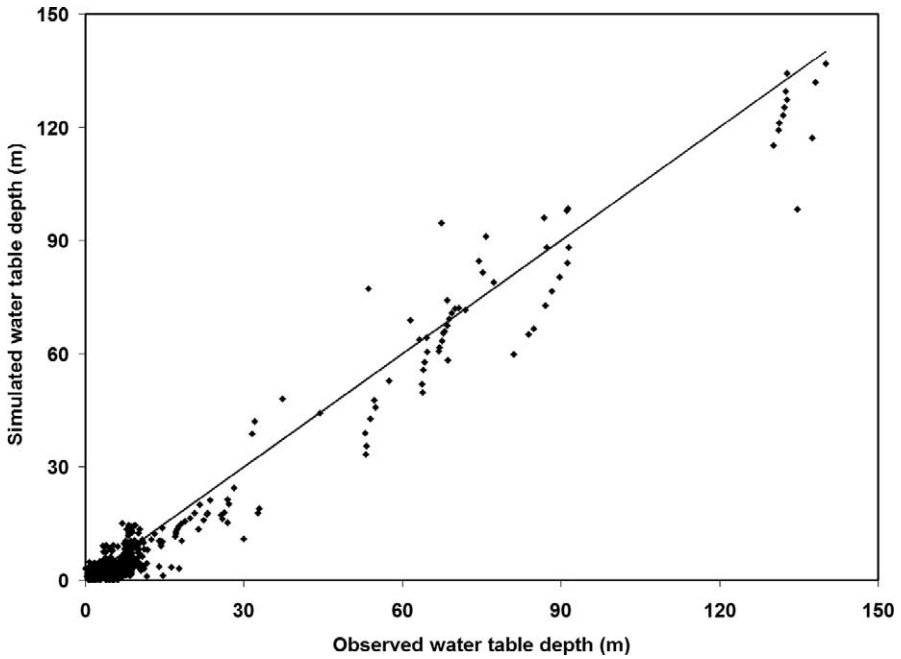

Fig. 3. Scatter plot of simulated vs. observed water table depths for the parameter set (Table 3) which minimizes OF, Eq. (17).

(accepted for publication). As pointed out by Boyle et al. (2000), the single-criterion approach can provide good parameter estimates and model fit, but parameter values may be hydrologically unrealistic.

Next, observations are compared with predicted output variables for the 13-year calibration period in Figs. 3–6. For this period, the 1980s were relatively wet, whereas the early 1990s included a 3–5 year drought, followed by wetter years in the 1990s.

The comparison between measured and simulated groundwater table depths in Fig. 3 show that the model generally predicts the regional variation in

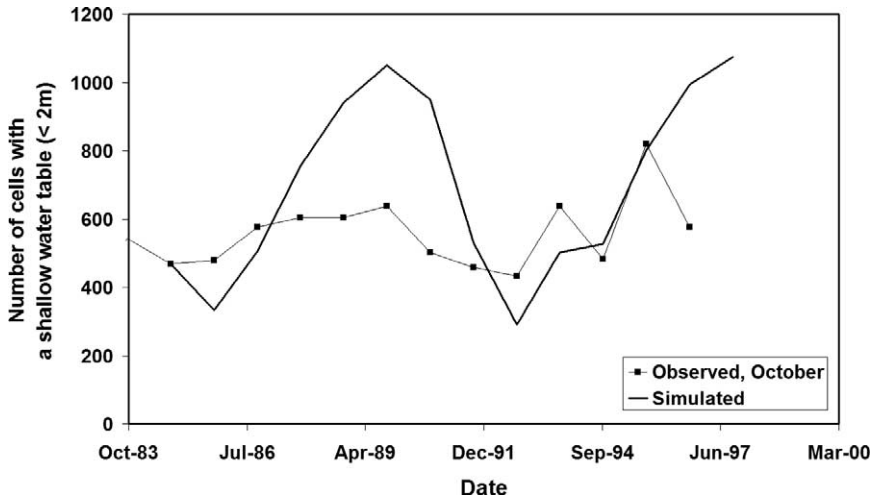

Fig. 4. Observed and simulated number of model cells with a shallow water table (<2 m) for the parameter set (Table 3) which minimizes OF, Eq. (17).

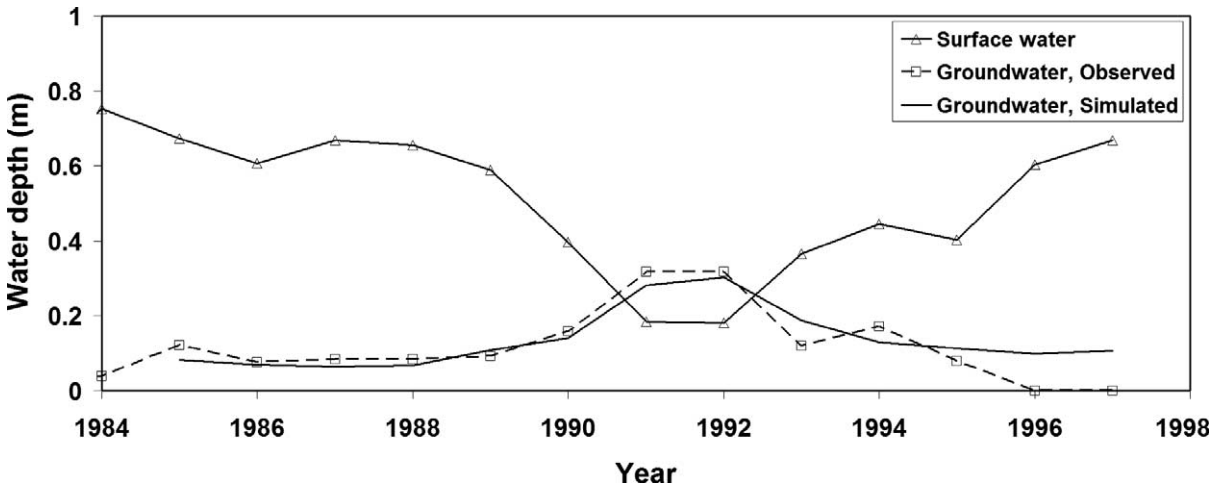

Fig. 5. Observed vs. simulated groundwater pumping in Westlands Water District (WWD) for the parameter set (Table 3) which minimizes OF, Eq. (17).

water table depths very well, with an RMSE value of 2.76 m, and a correlation coefficient of 0.98. However, these statistics are by themselves not very useful, realizing that the majority of wells are shallow with water table depths less than 15 m, and that the data scatter is relatively large for this depth range. Some of these local variations are caused by randomly distributing crops and corresponding boundary conditions within each district. A time series of the number of model cells with measured and predicted groundwater tables within 2 m from the land surface is presented in Fig. 4. For this comparison, the shallow water table observations were spatially-interpolated using kriging, to obtain 'measured' shallow water tables for each model cell for the month of October. In general, groundwater tables are closest to the soil

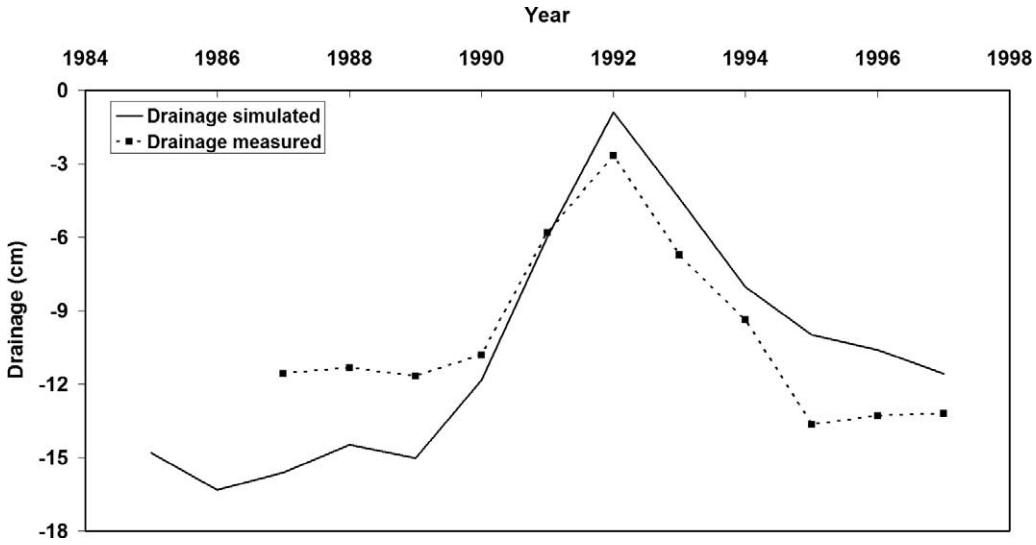

Fig. 6. Observed and simulated drainage in Broadview Water District (BWD) for the parameter set (Table 3) which minimizes OF, Eq. (17).

surface after the winter rains and during the irrigation season (April–July), and they are deepest at the end of the growing season before the winter (October). The simulations are representative of the October water levels, since initial conditions were taken in October 1984 and annual boundary conditions are used. Considering the long-term trends in Fig. 4, observations show gradual fluctuations with an increase in the area with shallow water tables before and after the early 1990s drought. Both the measured and simulated decrease in the shallow water table area during the drought years was caused by increased fallowing and groundwater pumping. The predicted trend of the areal coverage of shallow water table agrees fairly well with the observations, however, the simulated temporal dynamics is much more pronounced, resulting in an under-prediction of water table depths at the end of the 1980s before the drought. This will be further discussed in Section 3.4, where predicted water table maps are compared to observed ones.

Fig. 5 shows a comparison of simulated and observed groundwater pumping in WWD (Fig. 1A) during the 13-year calibration period, and includes the surface water delivery data. Typically, surface water deliveries vary between wet and dry years, with high deliveries in the 1980s, a pronounced decrease during the drought period, and a gradual recovery in the mid-late 1990s. The decrease in surface water allocation during the drought was partly compensated for by increased groundwater pumping. However, as Fig. 5 shows, the total amount of available irrigation water was significantly reduced during the drought years, resulting in increased acreage of fallow fields. Our calculated pumping using Eq. (8) is very close to reported groundwater pumping data.

A comparison of measured with predicted annual subsurface drainage fluxes for BWD (Fig. 1A) is presented in Fig. 6. As the other graphs, the effects of the early 1990s drought on drainage flows are evident, with reduced drainage flows coinciding with lower groundwater tables and water applications. Nevertheless, the calibrated regional flow model slightly over-predicts drainage in the 1980s and under-predicts drainage in the 1990s. To better understand the magnitude of prediction errors, we present a detailed annual water balance for BWD in Table 5, including annual values for applied water (AW), crop evapotranspiration (ET_c) and soil evaporation outside the growing season (E_a), soil water storage (dstor), lateral water fluxes across BWD boundaries (Q_x and Q_y), tile drainage flow (Q_d), and water flux across the Corcoran clay (Q_b). There is no pumping in this district. The control volume for the water balance

Table 5
Simulated annual water balance in Broadview Water District, BWD

	AW	ET_c	E_a	Q_d	dstor	Q_x	Q_y	Q_{pump}	Q_b
1985	106.03	−49.75	−24.8	−14.82	−14.05	−1.86	−0.72	0	−0.09
1986	96.96	−47.29	−27.64	−16.31	−3.21	−1.73	−0.66	0	−0.09
1987	87.36	−43.37	−25.17	−15.61	−1.05	−1.55	−0.56	0	−0.09
1988	87.54	−49.38	−22.87	−14.49	1.09	−1.35	−0.44	0	−0.09
1989	92.35	−48.84	−24.21	−15.02	−2.44	−1.3	−0.4	0	−0.07
1990	72.31	−46.15	−17.61	−11.83	5.19	−1.34	−0.46	0	−0.07
1991	48.83	−31.92	−16.87	−5.99	8.31	−1.21	−1.05	0	−0.07
1992	38.77	−26.26	−15.48	−0.9	6.17	−1.11	−1.12	0	−0.07
1993	73.35	−38.32	−21.8	−4.42	−6.82	−1.15	−0.82	0	−0.07
1994	81.41	−48.62	−16.64	−8.04	−5.67	−1.25	−0.84	0	−0.07
1995	93.52	−50.95	−26.18	−9.99	−4.29	−1.31	−0.64	0	−0.11
1996	88.46	−53.43	−21.66	−10.61	−0.73	−1.2	−0.72	0	−0.11
1997	94.73	−53.14	−25.39	−11.57	−2.69	−1.34	−0.42	0	−0.11
Average	81.66	−45.19	−22.02	−10.74	−1.55	−1.36	−0.68	0	−0.08

All units are in cm/year. AW, total applied water, Eq. (9a); ET_c crop evapotranspiration, E_a, soil evaporation; Q_d, tile drainage; dstor, change in water storage from the land surface to the top of the Corcoran clay; Q_x, net horizontal groundwater flow in x direction (along rows); Q_y, net horizontal groundwater flow in y direction (along columns); Q_{pump}: groundwater pumping from bottom layer; Q_b: seepage through the Corcoran clay. Positive/negative values correspond to sources/sinks of water, respectively.

extends from the land surface to the top of the Corcoran clay, with negative values indicating water leaving the simulated domain, except for dstor where a negative value reflects water storage gain. Table 5 shows that the main components of this district's water balance are AW (average of 82 cm) and ET_c (average of 45 cm), followed by soil evaporation (average of 22 cm), and tile drainage (average of 11 cm). Seepage through the Corcoran clay is less than 1 mm/year, whereas net horizontal groundwater flows are directed out of the district at an average rate of about, 1 cm/year, but never larger than 2 cm/year. Based on the water balance, we can conclude that drainage flow prediction errors are small and that they are largely a function of uncertainties in the local water balance (e.g. increased crop stress during the 1990s).

3.2. Three-objective optimization results

Although the single-optimization approach resulted in a very good model fit for drainage and groundwater pumping, the overall performance on the water tables was relatively poor. This is mainly a result of the applied weighting procedure in Eq. (17), emphasizing the drainage and pumping more than the regional groundwater tables. Alternatively, rather than iteratively conducting multiple single-objective optimizations with different weight factors, we conducted a multi-objective optimization with separate minimization of residuals of groundwater table, drainage, and pumping data. The advantage of the multi-objective approach is that independence of the objectives is maintained, whereas it does not require the (subjective) specification of prior weights (Gupta et al., 1998). Secondly, for a single optimization run, it generates a Pareto trade-off surface between the three objectives, and provides the best-fit parameter sets that minimize each objective individually (Vrugt et al., 2003b).

The results of the three-objective optimization with the MOSCEM-UA algorithm are presented in Figs. 7–13. Fig. 7 presents the three so-called bi-criterion plots for the three objectives containing all 4000 model evaluations. In the absence of model structural inadequacies and input and output measurement errors, no trade-offs in the fitting of the objectives will be found, and one set of parameters will minimize each of these objectives simultaneously. Not surprisingly, the MOSCEM-UA results clearly illustrates that trade-offs occur between the three different objectives. It means that optimization of each objective function separately, results in different optimized parameter sets. Interestingly, the shape of the bi-criterion plots differs among the objectives. The Pareto trade-off between groundwater table and pumping data (Fig. 7A) is along a right angle, indicating that significant improvements in water table predictions can be made without losing accuracy for the groundwater pumping predictions. In contrast, the other two bi-criterion plots (Fig. 7B and C) that include drainage are much more along a vertical straight line, indicating much more sensitivity to the drainage than the other two data types. Most likely that is so, because the drainage data apply only

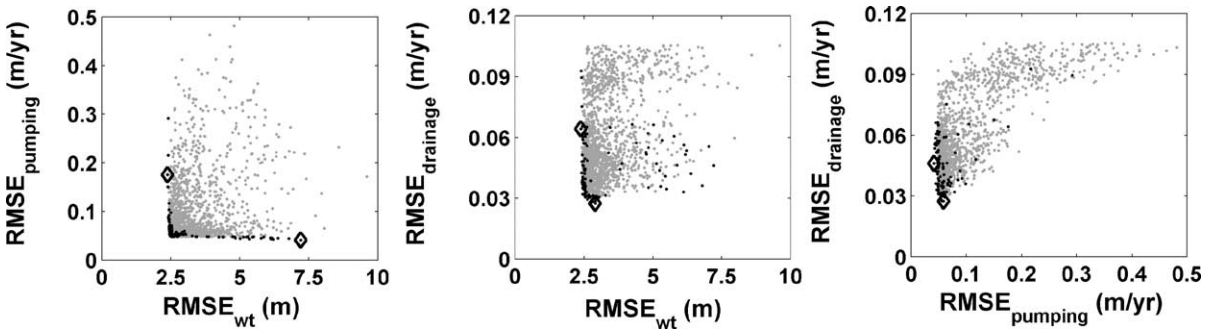

Fig. 7. Bi-criterion plots for the three-objective optimization with the MOSCEM-UA algorithm. Each dot corresponds to one of 4000 MODHMS simulations. The dark dots are two-dimensional projections of the three-objective trade-off (Pareto) surface, whereas the open diamonds correspond to the best-fit solutions for each individual objective.

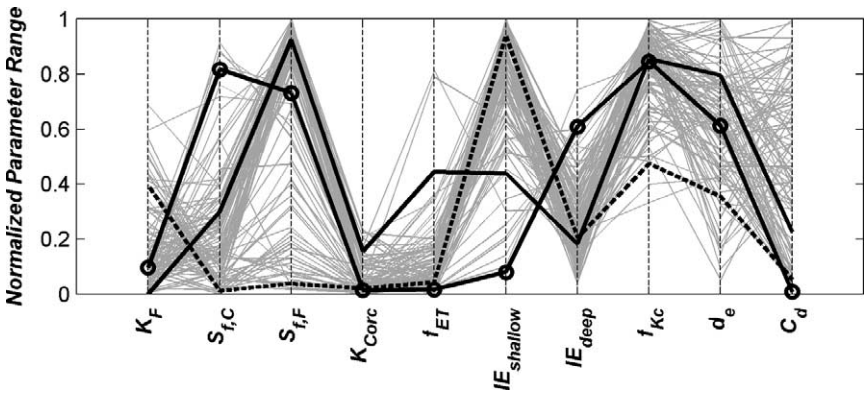

Fig. 8. Normalized parameter plots showing parameter variations over the three-criteria Pareto set. Each gray line corresponds to a member of the Pareto set. Darker lines correspond to optimal parameter sets for water table depth (solid), groundwater pumping (dashed), and drainage (dotted).

to the BWD (Fig. 1A), whereas the other two data types represent the whole region. Furthermore, the drainage predictions are affected by the two drainage parameters, d_e and C_d, but the other two data types are insensitive to these parameters. Consequently, we could have chosen to calibrate on the groundwater tables and pumping rates only in a first step, followed by optimization of the drainage parameters only in a second step, thereby reducing the number of parameters that are optimized simultaneously. Note also that some of the Pareto points in Fig. 7B and C fall outside the 'front'. These correspond to parameter sets that do well on the third objective not plotted in the bi-criterion plot.

Fig. 8 shows normalized parameter plots for each member of the three-objective Pareto set. The wide spread of most parameter values indicate their large uncertainty, and support the validity of the 'equifinality' concept of Beven (1993) that states that an infinite number of optimized parameter sets may result in equally-good model performances. The large parameter uncertainty of the Pareto set is likely due to a combination of errors related to model structure, measurements, and model input, but could also be caused by correlations between different parameters. Whereas in the absence of model structural inadequacies, and input and output measurement errors, additional types of data should improve parameter identifiability, the reverse is true when these errors are significant. In that case, the parameter values need to be changed to match different objectives by compensating for model error, which leads to large parameter uncertainty. However, parameters may still be well identified in simulating a single objective (see Fig. 2B), but different parameter sets will be needed for different objectives. Therefore, model calibration against a single objective may mask problems with the model structure. Obviously, the parameter uncertainty may also result from the low information content of the measurements. For example, the uncertainty of the drainage parameters in Fig. 8 is probably due to the insensitivity of the water table and pumping objectives to these parameters. We also note that unlike most other parameters, the crop correction factor, f_{ET}, is relatively well identified by the multi-objective optimization, with best-fit values that are between 0.8 and 0.9. These values are consistent with other studies, such as the one by Burt and Styles (1994). They concluded that crop growth and development is non-uniform for many fields, resulting in effective field-scale crop coefficients that are about 20% smaller than typical local-scale values reported in the literature (e.g. Snyder et al., 1989; Allen et al., 1998). The parameter K_{Corc} is also relatively well identified, with optimum values at the lower end of the prior range. Larger values for K_{Corc} result in too much seepage through the Corcoran clay, resulting in lower water tables, less drainage, and an increase in groundwater pumping.

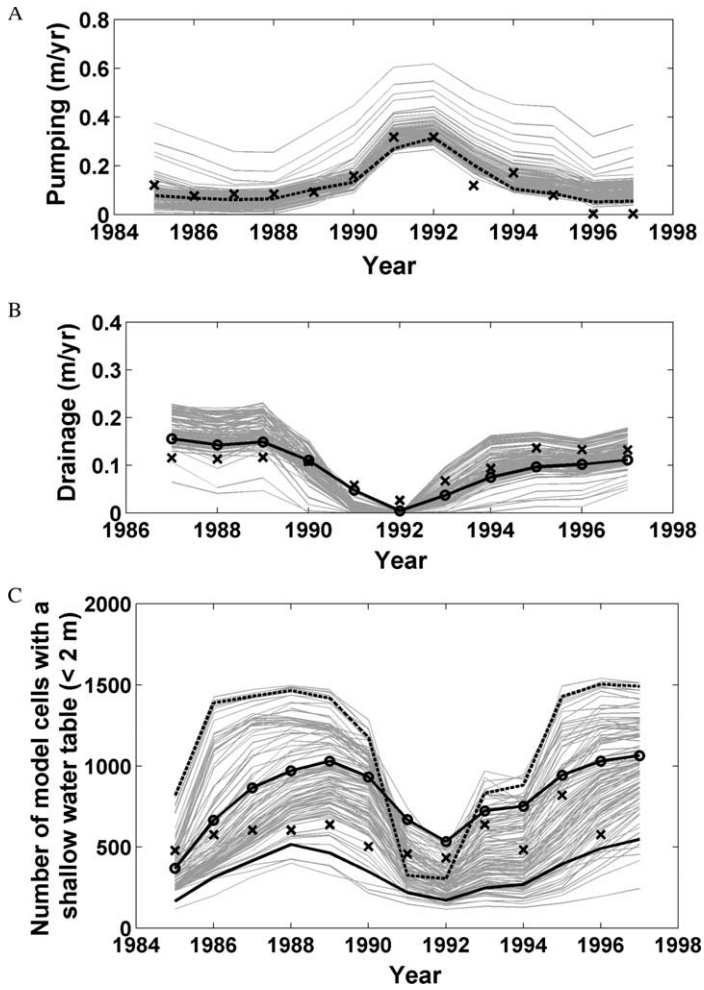

Fig. 9. Effect of Pareto parameter uncertainty on: (A) groundwater pumping predictions in Westlands Water District (WWD), (B) subsurface drainage predictions in Broadview Water District (BWD), and (C) prediction of number of model cells with a shallow water table (<2 m) in October. Observed values are shown as crosses; each member of the Pareto set is represented by a gray line; optima for water tables (solid), pumping (dashed), and drainage (dotted) are also shown. These optima correspond to the parameter sets in Table 3 that minimize one of three objectives in Eq. (18).

Low values for f_{ET} and K_{Corc} are consistent in that they both increase the amount of water storage in the model domain.

Figs. 9–11 summarize the effects of the Pareto parameter uncertainty on prediction uncertainty of groundwater pumping, subsurface drainage, and water table depths. Fig. 9A shows results for groundwater pumping. The optimal solution matches the observed data, as presented by the cross symbols, very well. We conclude that, despite the large parameter uncertainty of the Pareto set, the prediction uncertainty for groundwater pumping is small. A similar graph was plotted for drainage in Fig. 9B. Again, the optimal solution matches the data well, but the prediction uncertainty is larger than was the case for pumping. Figs. 9C and 10 demonstrate the large uncertainties of the predicted groundwater tables. As Fig. 9C shows, the prediction of the number of model cells with a shallow water table (<2 m) is associated with large uncertainty. The parameter sets minimizing

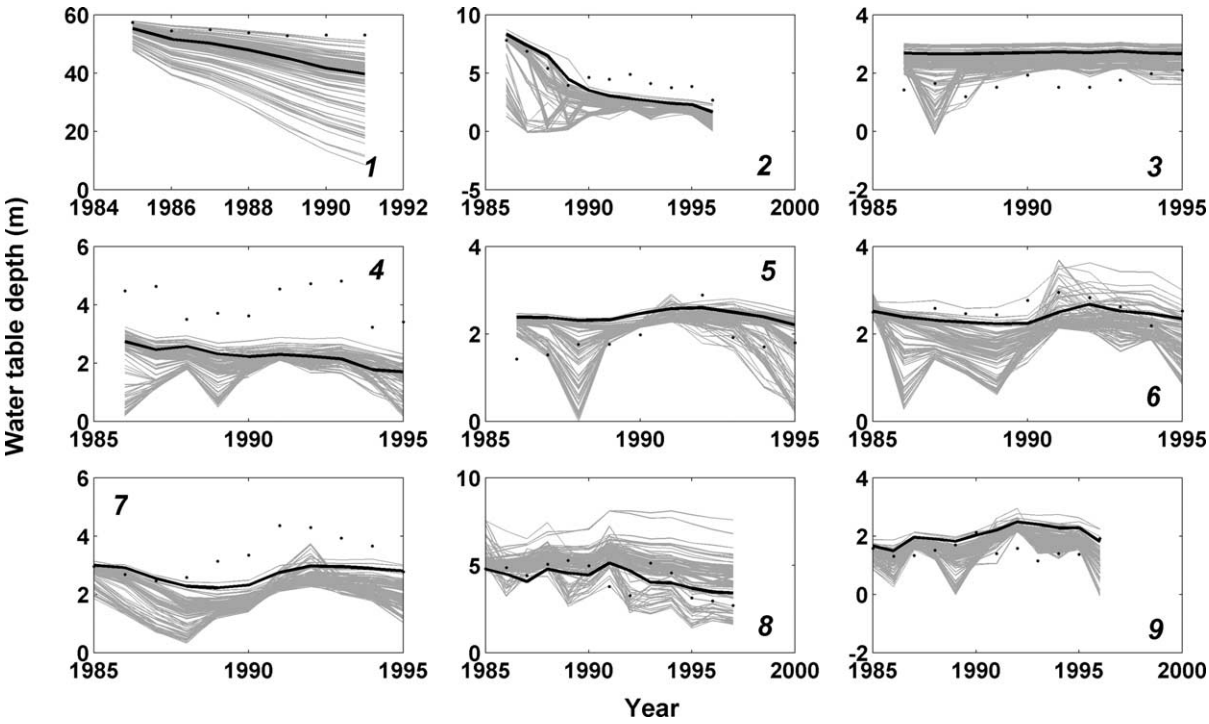

Fig. 10. Effect of Pareto parameter uncertainty on prediction of water table depth predictions: observed (crosses), each member of the Pareto set is represented by a gray line; results for best water table parameter set shown as dark lines. The observation well numbers refer to those in Fig. 1A.

the drainage and pumping RMSE values result in too much dynamics. This is consistent with the single-objective results in Fig. 4. The parameter set that minimizes the water table RMSE results in less dynamics, but gives a slight yet consistent under-prediction of the number of model cells with a shallow water table in October of each year. Fig. 10 shows results for predicting water levels in individual wells. The optimal solution, which minimizes $RMSE_{wt}$ in Eq. (18c), fits the data well for some locations (e.g. well number 2), but severely under-predicts or over-predicts in other cases. Generally, the prediction errors and uncertainties are smaller when the water table is closer to the land surface, because in that case water level fluctuations are bounded by the land surface and controlled by evaporation from the water table. The large uncertainties for water table prediction at individual well locations partly arise from the uncertainties in the field-scale boundary conditions, since crop acreages and irrigation deliveries are only known at the district scale, and are randomly distributed over all fields within the district.

Finally, in Fig. 11 we present maps of the 13-year time-averaged leaching rate through the bottom of the root-zone (A) and associated standard deviation (B), computed for the complete Pareto parameter set. The leaching rate is of interest because it is a driving force in soil salinization. Average leaching rates are generally less than 0.2 m/year for shallow water table areas, but increase to larger than 0.2 m/year in the western, upslope direction. Leaching rates are less than 0.1 m/year in the inter-fan areas, at the very eastern edge of Panoche Creek alluvial fan, and in the regional-collector drainage area (Fig. 1C) for which the drainage system was plugged in 1985. In contrast, some of the areas underlain by on-farm drains (Fig. 1C) show leaching rates of 0.1–0.2 m magnitude. Areas of negative leaching rates occur in

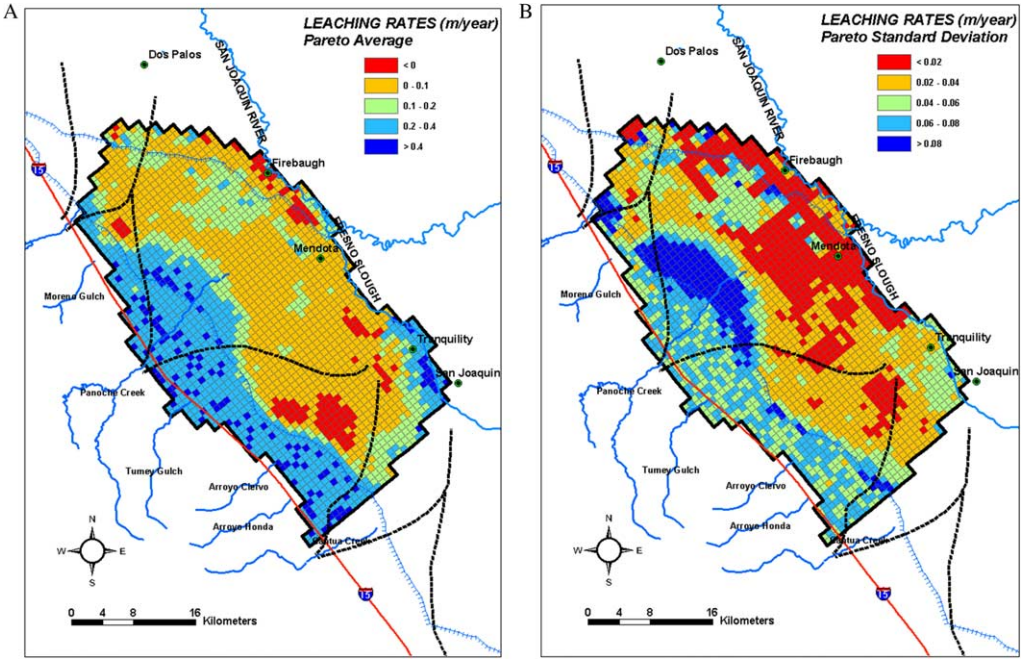

Fig. 11. Effect of Pareto parameter uncertainty on prediction of spatially distributed time-averaged (1984–1997) leaching rates: (A) Pareto average leaching rates, (B) Pareto standard deviation of leaching rates.

the inter-fan areas—these correspond to shallow water table areas dominated by upward capillary flow. Leaching rate variation over the Pareto set is shown in Fig. 11B. The largest variations were determined for the mid-fan area of the Panoche fan, for which the density of observation wells was the smallest (Fig. 1A). Also, a significantly larger average and variation in leaching rates was computed for the drained areas (Fig. 1C), as compared to similar positions in the landscape that were not drained. The subsurface tile drains control the water table and allow farmers to apply more water for leaching of salts. Furthermore, capillary rise from the water table is less in areas with subsurface drains. The large variation in leaching rates in the drained area is probably due to the Pareto uncertainty of the drain parameters (d_e and C_d), which do not influence leaching outside of the drained area.

3.3. Consideration of additional performance criteria

As part of the results in Section 3.2, it was concluded that the parameter variation and corresponding prediction uncertainty was relatively high for the Pareto parameter set of solutions. Poor parameter identifiability can be the result of a combination of causes, including model structural inadequacies, and input and output measurement errors. In addition, it might be a consequence of low parameter sensitivity to the calibration data.

In the following, we will investigate whether the parameter uncertainty can be reduced by rejecting parameter sets from the Pareto set that do not satisfy additional performance criteria. In particular, we look at two types of additional information, namely: (1) model BIAS for the model predictions of groundwater table, pumping rate, and drainage data, and (2) model RMSE and BIAS for specific water table depth classes. We include the BIAS criterion, because it would indeed be desirable for any prediction model to have a near-zero bias (Boyle et al., 2000). Second, it has been suggested that distributed models should be tested against distributed data (e.g. Refsgaard, 1997; Beven, 2001). Distributed data are bound to contain a lot of additional information on model performance. Consequently,

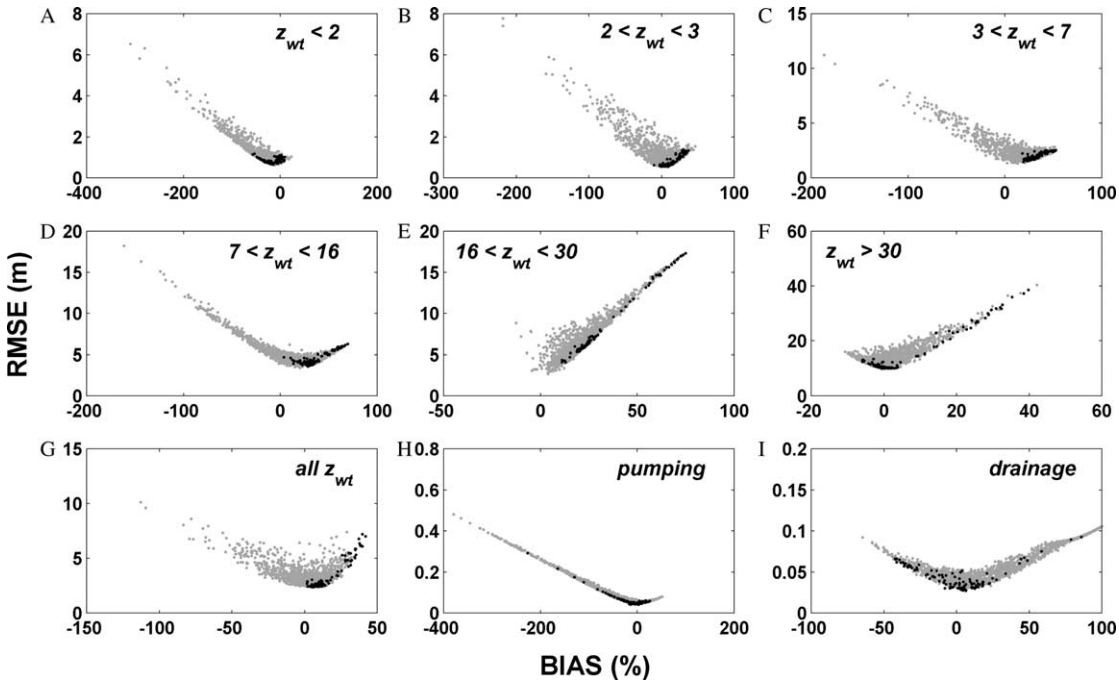

Fig. 12. Root mean square error (RMSE) as a function of %BIAS for all nine model performance criteria. (A) $z_{wt} < 2$ m, (B) $2 < z_{wt} < 3$ m, (C) $3 < z_{wt} < 7$ m, (D) $7 < z_{wt} < 16$ m, (E) $16 < z_{wt} < 30$ m, (F) $z_{wt} > 30$ m, (G) all groundwater tables, (H) all pumping rate data, and (I) all drainage rate data. The Pareto set of solutions are shown in black.

we divided the large number of wells (Fig. 1A) into six water table depth groups, ranging from very shallow (<2 m deep) to very deep (>30 m deep). Since there is a clear trend in water table depth from east to west through the study area, these water table groups also represent separate geographical regions.

Plots of RMSE vs. %BIAS for each of the three main objectives and the six additional water table depth groups are presented in Fig. 12. Each plot consists of all 4000 optimized parameter solutions. In addition, each plot includes the Pareto points (black dots). These plots show that increasing RMSE values create larger BIAS, resulting in a clear minimum. For most criteria this minimum coincides with near-zero BIAS, except for water table depths between 3 and 16 m, indicating additional trade-offs exist between low RMSE values and near-zero BIAS for those depths. However, not always are these minima clear defined. For example, the Pareto results of the drainage data, Fig. 12I, show a large range in model BIAS near the minimum RMSE. It suggests that model BIAS should be used as well to better define the best-fit prediction model and to reduce parameter uncertainty. For example, we may select to eliminate all parameter sets from the Pareto set that result in absolute BIAS values larger than 5% for the shallow water table (<2 m) predictions. This resulted in a subset of 24 parameter sets (from the 106 Pareto points). However, the parameter ranges of this subset relative to the Pareto set (Table 6) indicate that parameter uncertainty was only significantly reduced for the f_{ET} and K_{Corc} parameters. This indicates that the shallow water table BIAS is mostly sensitive to these two parameters.

The results of Fig. 12A–F also provide further insights into the spatial distribution of model performance. First, they show that the bias of the Pareto set of solutions varies between water table depth intervals. Specifically, for the most shallow groundwater tables in Fig. 12A, the Pareto set results in mostly negative BIAS values, indicating that these overestimate the water table depth. In contrast,

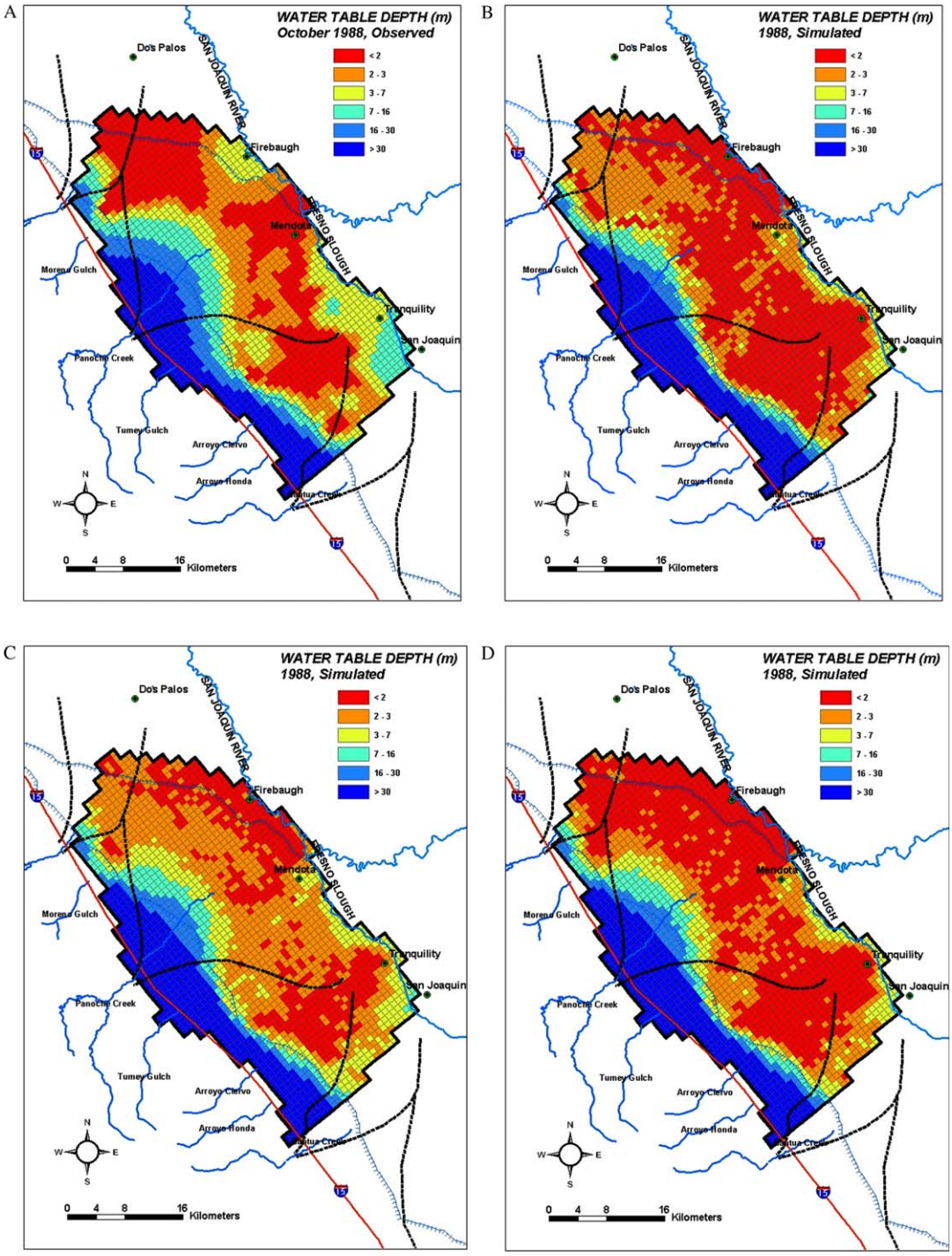

Fig. 13. (A) Measured water table map for October 1988, and (B–D) simulated depth to water table maps for October 1988 using parameter sets for cases (1), (2), and (8), respectively, as discussed in the text.

Table 6
Parameter ranges for all members of the Pareto set that result in a BIAS of 5% or less for the shallow water tables (<2 m), as a percentage of the parameter ranges of the entire Pareto set

	Relative range (%)
K_F	82
$S_{y,C}$	96
$S_{y,F}$	84
K_{Corc}	25
f_{ET}	18
$IE_{shallow}$	91
IE_{deep}	76
f_{Kc}	89
d_e	93
C_d	91

the Pareto set for most other water table depth intervals have positive BIAS, indicating an under-prediction of water table depth. Possibly, this may be related to a shortcoming of the soil evaporation model, as discussed in Section 2.2.2., with model errors largely affecting the shallow water table predictions. Another cause might be related to the calibration of single-valued parameters, assuming that these are spatially uniform and temporally constant. Since the groundwater table classes generally coincide with different geographical areas, better groundwater table predictions might have been achieved by spatially distributing calibration parameters, thereby reducing trade-offs between matching shallow (<2 m) and deeper water table depths. For example, Vrugt et al. (accepted for publication) concluded that significant improvements were made in the prediction of spatially distributed drainage rates, if calibration parameters were spatially distributed, however, with the risk of over-parameterization. As one approach for spatially distributing the f_{ET} parameter, one could explicitly simulate water stress in Eq. (3) as a function of root-zone moisture content. However, it would require defining soil hydraulic functions other than the pseudo-soil approach of Eq. (2) assumed in this study.

The model prediction error has different implications for regions with shallow and deep groundwater tables. For example a 10% BIAS for the deeper groundwater tables will likely not affect drainage, ponding and capillary rise with associated soil salinity. However, this same percentage value for the shallow water tables might radically change the near surface hydrology. Therefore, the performance criteria used and the associated best-fit parameter set will depend on the intended use of the model (Refsgaard and Henriksen, 2004). This is illustrated in Section 3.4.

3.4. Selecting a 'best' parameter set

Given the trade-offs that occur between the fitting of the different objectives of the hydrologic system, it is now clear that the success of identifying a single 'best' parameter set will depend on the selection of performance criteria and intended use of the hydrologic model. For example, minimizing the groundwater table RMSE might lead to large estimation errors of groundwater pumping (Fig. 7A). Here, we wish to illustrate how the Pareto solution set resulting from a multi-objective optimization contains all required information to select best-fit parameter sets, depending on the importance of the individual objectives as defined by the user. This is indeed the main strength of multi-objective optimization.

The 'best' parameter sets for both the single and multi-objective function optimizations, emphasizing different goals of the hydrologic model, are presented in Table 3. In addition, Table 7 includes both %BIAS and RMSE values for all optimizations, using all nine data types. The following objectives were considered: (1) minimization of the single weighted objective function of Eq. (17), (2) minimization of $RMSE_{wt}$, (3) minimization of $RMSE_{pump}$, (4) minimization of $RMSE_{drain}$, (5) selection from the Pareto set of the parameter set with the minimum Euclidean distance in the normalized space of the three main objectives (MO_1), (6) selection from the Pareto set of the parameter set which minimizes $RMSE_{wt<2}$ and with an absolute $BIAS_{wt<2}$ less than 5% (MO_2), (7) selection from the Pareto set of the parameter set which minimizes $RMSE_{2<wt<3}$ and with an absolute $BIAS_{2<wt<3}$ less than 5% (MO_3), and (8) selection from the Pareto set of the parameter set with the minimum Euclidean distance in the normalized two-objective space of $RMSE_{wt<2}$ and $RMSE_{2<wt<3}$ (MO_4). For cases (1) and (5), the user will be equally interested in any of the three main objectives. As such, a 'best' parameter set may be determined by using the single optimization

Table 7
Model performance criteria corresponding to parameter sets in Table 3

	Minimal OF, Eq. (17)	Minimal $RMSE_{wt}$	Minimal $RMSE_{pump}$	Minimal $RMSE_{drain}$	MO_1	MO_2	MO_3	MO_4
$BIAS_{wt<2}$	−14.78	−44.36	2.12	6.99	−20.57	−4.51	−41.44	−20.41
$RMSE_{wt<2}$	0.73	0.96	1.10	0.82	0.71	0.73	0.91	0.67
$BIAS_{2<wt<3}$	17.14	2.74	29.94	32.39	13.81	25.04	−0.52	14.59
$RMSE_{2<wt<3}$	0.79	0.58	1.38	1.15	0.68	0.94	0.55	0.66
$BIAS_{3<wt<7}$	33.41	26.30	42.60	36.67	30.57	34.18	23.54	32.31
$RMSE_{3<wt<7}$	2.00	1.63	2.52	2.22	1.83	2.05	1.55	1.81
$BIAS_{7<wt<16}$	33.95	30.49	59.29	14.63	28.26	17.73	30.99	30.92
$RMSE_{7<wt<16}$	4.33	3.73	5.93	4.25	4.05	4.13	3.76	4.01
$BIAS_{16<wt<30}$	30.43	22.25	73.62	11.82	20.17	12.60	22.79	24.10
$RMSE_{16<wt<30}$	7.80	5.82	17.15	4.31	5.59	4.29	5.89	6.21
$BIAS_{wt<30}$	5.19	−0.15	39.72	−5.43	−2.53	−5.55	−0.97	1.25
$RMSE_{wt<30}$	11.56	10.12	38.64	12.87	10.92	12.71	10.42	10.19
$BIAS_{wt}$	15.04	5.98	39.66	11.83	9.11	9.72	5.09	11.63
$RMSE_{wt}$	2.76	2.38	7.21	2.91	2.54	2.81	2.40	2.45
$BIAS_{pump}$	−7.81	−130.27	1.88	4.25	2.68	4.07	−57.40	−53.28
$RMSE_{pump}$	0.05	0.18	0.04	0.06	0.05	0.05	0.09	0.08
$BIAS_{drain}$	1.51	48.74	−11.89	6.30	3.83	6.44	24.74	43.21
$RMSE_{drain}$	0.03	0.06	0.05	0.03	0.03	0.03	0.04	0.06

RMSE is root mean square error [m], Eq. (18). BIAS is expressed as a percentage relative to the observations, Eq. (19). Subscripts denote the range of observations used to calculate each criterion. For example, $BIAS_{wt<2}$ gives the bias for predicting observed water table depths less than 2 m below the land surface. MO_1–MO_4 denote optimal solutions obtained by the multi-objective optimization, as discussed in Section 3.4. MO_1, minimum Euclidean distance in the normalized space of the three main objectives; MO_2, minimum $RMSE_{wt<2}$ and absolute $BIAS_{wt<2}$ less than 5%; MO_3, minimal $RMSE_{2<wt<3}$ and absolute $BIAS_{2<wt<3}$ less than 5%; MO_4, minimum Euclidean distance in the normalized two-objective space of $RMSE_{wt<3}$ and $RMSE_{2<wt<3}$.

approach of case (1), or by identifying the single parameter set from the Pareto set solutions that is nearest to the origin within the Euclidean normalized three-objective space (case 5). Cases (2)–(4) consider a single main objective, ignoring the other two objectives. When selecting case (6), the user is focusing on the best possible prediction of shallow water tables only (<2 m), whereas the user that selects case (7) is interested in the intermediate water table depths (2–3 m depth). Case (8) presents a compromise solution between (6) and (7).

First, we conclude that both pumping and drainage objectives are generally well predicted by most parameter sets, except when selecting case (2) that only optimizes the groundwater table data. The most accurate drainage predictions were obtained for case (1), minimizing the single-objective function of Eq. (17). The parameter set associated with case (1) was already discussed in Section 3.1. This parameter set gives good results for groundwater pumping and drainage, but an under-prediction of water tables deeper than 2 m and an over-prediction of water tables shallower than 2 m, also as discussed earlier. The parameter set that minimizes the $RMSE_{wt}$ is different from all the others in Table 3, in that it has the largest values for f_{ET} and K_{Corc}. The result is deeper simulated water tables than in case (1), leading to smaller under-predictions for deeper water tables, but worse over-predictions for the shallow water tables (Table 7). This indicates again the significant trade-offs between water table groups. Optimizing the pumping RMSE in case (3) leads to very small values for specific yield (Table 3), and hence simulated water tables that are too shallow, as evidenced by the large positive BIAS for the water table predictions and a negative BIAS for drainage in Table 7. This also eliminates the over-prediction of the shallow water tables (<2 m). The parameter set for case (4), which minimizes $RMSE_{drain}$, is characterized by a small value for irrigation efficiency (0.74) in the shallow water table area, $IE_{shallow}$. In other words,

irrigation water is shifted from the deep (>2 m) to the shallow water table areas (<2 m), resulting in a smaller trade-off in simulating these two water table groups. Case (5) represents a trade-off between the three main objectives, which results in a parameter set and model performance intermediate to cases (2)–(4). In case (6) the focus is specifically on the shallow water tables (<2 m). This results in a similarly small value for $IE_{shallow}$, as in case (4), where the drainage RMSE was minimized. Again, minimizing the shallow water tables results in bad results for the other water table groups, and vice versa. When attention shifts to the water tables between 2 and 3 m deep, case (7), the optimum parameter values change such that water tables decline: this is accomplished by increasing $IE_{shallow}$, K_{Corc} and f_{ET} (Table 3). Given the trade-off between the objectives in case (6) and (7), it is worthwhile to consider a compromise solution between the two, as in case (8). It results in parameter values ($IE_{shallow}$, K_{Corc}, f_{ET}) and performance statistics ($RMSE_{wt<2}$ and $RMSE_{2<wt<3}$) that lie in between cases (6) and (7).

Finally, in Fig. 13 we show a comparison between predicted groundwater table maps for October 1988 using the parameter sets for cases (1), (2), and (8), and a map of interpolated water table depths measured at individual wells in October 1988. The observed map (Fig. 13A) shows that water tables are shallow in the north-eastern part of the study area, dropping in the south-western direction. It is clear that the parameter set for case (1) significantly under-predicts water table depths. As discussed, this is caused by the emphasis that the single-objective function, Eq. (17), puts on groundwater pumping, which results in parameter values that cause water tables to rise: small values for f_{ET}, K_{Corc}, and $S_{y,C}$ (Table 3). On the other hand, the parameter set which minimizes $RMSE_{wt}$, case (2), results in a smaller than observed shallow water table extent. Finally, the results for case (8) provide an intermediate prediction between cases (1) and (2), with generally a good resemblance to the observed pattern, especially in the area where observations are abundant (see Fig. 1A). Not surprisingly, none of the parameter sets performs well in the northern area of the model, since observations in that region are scarce.

4. Summary and conclusions

We presented the calibration of a regional distributed subsurface water flow model for a 1400 km^2 agricultural area in the western San Joaquin Valley of California, using two global optimization algorithms. Ten model parameters were selected that characterized the regional hydrology, comparing measured with predicted spatially distributed groundwater table depth measurements, district-average groundwater pumping estimates, and subsurface drainage data. For the single-objective optimization approach, all three measurement types were combined in a single-objective function. In the second approach, the independence of the three measurement types was maintained by formulating a three-objective optimization problem. The single-objective optimization approach resulted in well-identified calibration parameters with small parameter uncertainty bounds. However, at least three parameter values approached their physically-realistic bounds during the optimization, thereby questioning their validity as effective parameters. Also, the normalization method that was used weighted each data type with respect to the number of available data points, thereby emphazing the pumping and drainage data. In the multi-objective approach, each of the three objectives was treated independently, so that no preference was given to any of the three data types. Within a single optimization run, a Pareto set of solutions was generated that included the end-members, representing the optimal results for each objective separately. We demonstrated that there was considerable trade-off between pumping and water table predictions, but that the drainage data optimizations were slightly independent of the other two objectives. Results showed the large parameter uncertainty of the Pareto set, indicating the presence of significant errors related to model structure, measurements, and model input. We showed that the magnitude of prediction uncertainties associated with the Pareto parameter uncertainty is large for making water table predictions, but much smaller for drainage and pumping predictions. Trade-offs between fitting shallow and deep water tables were revealed by considering additional performance criteria for model evaluation, namely BIAS and RMSE values for six water table depth groups. These results point to two possible strategies for

improving model predictions of spatially distributed water tables: (1) remove some of the trade-offs by spatially distributing the model parameters, especially f_{ET} and K_{Corc}, and (2) perform a multi-objective optimization with additional criteria for the water table groups. Further work will focus on the evaluation of the calibrated model using independent data not used for the calibration here.

Acknowledgements

The authors acknowledge financial support by the USDA Funds for Rural America, project no. 97-362000-5263. We acknowledge the support of HydroGeologic, Inc. by providing us with a beta version of the MODHMS code. The Earth Life Sciences and Research Council (ALW) supported the investigations of the University of Amsterdam co-author with financial aid from the Netherlands Organization for Scientific Research (NWO).

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration, Guidelines for Computing Crop Water Requirements FAO Irrigation and Drainage Paper 56. FAO, Rome, Italy pp. 135–158.

Ayars, J.E., Schrale, G., 1990. Irrigation Efficiency and Regional Subsurface Drain Flow on the West Side of the San Joaquin Valley. USDA/ARS, Fresno, CA. 120 pp.

Belitz, K., Heimes, F.J., 1990. Character and Evolution of the Groundwater Flow System in the Central Part of the Western San Joaquin Valley, California USGS Water Supply Paper 2348.

Belitz, K., Phillips, S.P., 1995. Alternative to agricultural drains in California's San Joaquin Valley: results of a regional-scale hydrogeologic approach. Water Resour. Res. 31 (8), 1845–1862.

Belitz, K., Phillips, S.P., Gronberg, J.M., 1993. Numerical Simulation of Groundwater Flow in the Central Part of the Western San Joaquin Valley, California USGS Water Supply Paper 2396, Sacramento, CA. 69 pp.

Beven, K., 1993. Prophecy, reality and uncertainty in distributed hydrological modeling. Adv. Water Resour. 16, 41–51.

Beven, K., 2001. How far can we go in distributed hydrological modeling. Hydrol. Earth Syst. Sci. 5, 1–12.

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

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modeling: a review. Hydrol. Proc. 9, 251–290.

Boyle, D.P., Gupta, H.V., Sorooshian, S., 2000. Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods. Water Resour. Res. 36 (12), 3663–3674.

Burt, C.M., Styles, S., 1994. Grassland Basin Irrigation and Drainage Study. Irrigation Training and Research Center, San Luis Obispo, CA.

Davis, G.H., Coplen, T.B., 1989. Late Cenozoic Paleohydrogeology of the Western San Joaquin Valley, California, as Related to Structural Movements in the Central Coast Ranges GSA Special Paper 234, Boulder, CO. 40 pp.

Desbarats, A.J., 1998. Scaling of constitutive relationships in unsaturated heterogeneous media, a numerical investigation. Water Resour. Res. 34 (6), 1427–1435.

Doherty, J., 2001. Improved calculations for dewatered cells in MODFLOW. Ground Water 39 (6), 863–869.

Duan, Q., Gupta, V.K., Sorooshian, S., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.

Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R., 2003. Calibration of Watershed Models. American Geophyiscal Union, Washington, DC.

Fio, J.L., 1994. Calculation of a Water Budget and Delineation of Contributing Sources to Drainflows in the Western San Joaquin Valley, California, USGS Open File Report 94-45, Sacramento, CA, 39 pp.

Grismer, M.E., Gates, T.K., 1988. Estimating saline water table contributions to crop water use for irrigation-drainage system design. California Agric. 42 (2), 23–24.

Gronberg, J.M., Belitz, K., 1992. Estimation of a Water Budget for the Central Part of the Western San Joaquin Valley, California, USGS Water Resources Investigations Report 91–4195, 22 pp.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Towards improved calibration of hydrologic models: multiple noncommensurable measures of information. Water Resour. Res. 34, 751–763.

Gupta, H.V., Bastidas, L., Vrugt, J.A., Sorooshian, S., 2003. Multiple criteria global optimization for watershed model calibration, in calibration of watershed models, in: Duan, Q. et al. (Ed.), Water Sci. Appl. Ser, vol. 6. AGU, Washington, DC, pp. 125–132.

Hopmans, J.W., Šimůnek, J., Romano, N., Durner, W., 2002. Simultaneous determination of water transmission and retention properties—inverse methods, in: Dane, J.H., Topp, G.C. (Eds.), Methods of Soil Analysis, Part 4, Physical Methods SSSA Book Series No. 5. SSSA, Madison, WI, pp. 963–1008.

Huyakorn, P.S., Wu, Y.-S., Park, N.S., 1994. An improved sharp-interface model for assessing NAPL contamination and remediation of groundwater systems. J. Contam. Hydrol. 16, 203–234.

HydroGeoLogic Inc., 2001. MOD-HMS, A Comprehensive MOD-FLOW-Based Hydrologic Modeling System, Herndon, VA.

Johnson, A.I., Moston, R.P., Morris, D.A., 1968. Physical and Hydrologic Properties of Water-Bearing Materials in Subsiding Areas in Central California USGS Professional Paper 497-A.

Kahrl, W.L. (Ed.), 1979. California Water Atlas. Governor's Office of Planning and Research, Sacramento, CA. 113 pp.

Karpouzos, D.K., Delay, F., Katsifarakis, K.L., de Marsily, G., 2001. A multi-population genetic algorithm to solve the inverse problem in hydrogeology. Water Resour. Res. 37, 2291–2302.

Kool, J.B., Parker, J.C., 1988. Analysis of the inverse problem for transient unsaturated flow. Water Resour. Res. 24, 817–830.

Laudon, J., Belitz, K., 1991. Texture and depositional history of late Pleistocene-Holocene alluvium in the central part of the western San Joaquin Valley, California. Bull. Assoc. Eng. Geol. 28 (1), 73–88.

Madsen, H., 2003. Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives. Adv. Water Resour. 26, 205–216.

McDonald, M.G., Harbaugh, A.H., 1988. A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model Techniques of Water Resources Investigations of the United States Geological Survey, Book 6. US Geological Survey, Washington, DC (Chapter A1).

McLaughlin, D., Townley, L.R., 1996. A reassessment of the groundwater inverse problem. Water Resour. Res. 32, 1131–1161.

Medina, A., Carrera, J., 1996. Coupled estimation of flow and solute transport parameters. Water Resour. Res. 32, 3063–3076.

Panday, S., Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. Adv. Water Resour. 27, 361–382.

Refsgaard, J.C., 1997. Parameterization, calibration and validation of distributed hydrological models. J. Hydrol. 198, 69–97.

Refsgaard, J.C., Henriksen, H.J., 2004. Modeling guidelines—terminology and guiding principles. Adv. Water Resour. 27, 71–82.

Refsgaard, J.C., Thorsen, M., Jensen, J.B., Kleeschulte, S., Hansen, S., 1999. Large-scale modeling of groundwater contamination from nitrate leaching. J. Hydrol. 221, 117–140.

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 1998. Neural network analysis for hierarchical prediction of soil water retention and saturated hydraulic conductivity. Soil Sci. Soc. Am. J. 62, 847–855.

Šimůnek, J., Šejna, M., van Genuchten, M.Th., 1998. The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media, Version 2.0, IGWMC—TPS—70, International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 202 pp.

Šimůnek, J., Jarvis, N.J., van Genuchten, M.Th., Gärdenäs, A., 2003. Review and comparison -of models for describing non-equilibrium and preferential flow and transport in the vadose zone. J. Hydrol. 272, 14–35.

Snyder, R.L., Lanini, B.J., Shaw, D.A., Pruitt, W.O., 1989. Using Reference Evapotranspiration (ET_0) and Crop Coefficients to Estimate Crop Evapotranspiration (ET_c) for Agronomic Crops, Grasses, and Vegetable Crops. California Department of Water Resources, Leaflet 21427, Sacramento, CA pp. 1–12.

van Dam, J.C., Stricker, J.N.M., Droogers, P., 1994. Inverse method for determining soil hydraulic functions from multi-step outflow experiments. Soil Sci. Soc. Am. J. 58, 647–652.

van Genuchten, M.Th., 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44, 892–898.

Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S., 2003a. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters. Water. Resour. Res. 39 (8), 1201.

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

Vrugt, J.A., Schoups, G., Hopmans, J.W., Young, C., Wallender, W.W., Harter, T.H., Bouten, W., 2004. Inverse modeling of large-scale spatially-distributed vadose zone properties using global optimization, Water Resour. Res. 40, W06503, doi: 10.1029/2003WR002706.

Weiss, R., Smith, L., 1998. Parameter space methods in joint parameter estimation for groundwater flow models. Water Resour. Res. 34, 647–661.

Wildenschild, D., Jensen, K.H., 1999. Numerical modeling of observed effective flow behavior in unsaturated heterogeneous sands. Water Resour. Res. 35, 29–42.

Wood, B.D., Kavvas, M.L., 1999. Ensemble-averaged equations for reactive transport in porous media under unsteady flow conditions. Water Resour. Res. 35, 2053–2068.

Yapo, P.O., Gupta, H.V., Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. J. Hydrol. 204, 83–97.

Zhang, D., 2002. Stochastic Methods for Flow in Porous Media: Coping with Uncertainties. Academic Press, San Diego, CA.