Accepted Manuscript

A comparison of deterministic and stochastic approaches for regional scale inverse modelling on the Mar del Plata aquifer

M. Pool, J. Carrera, A. Alcolea, E. Bocanegra
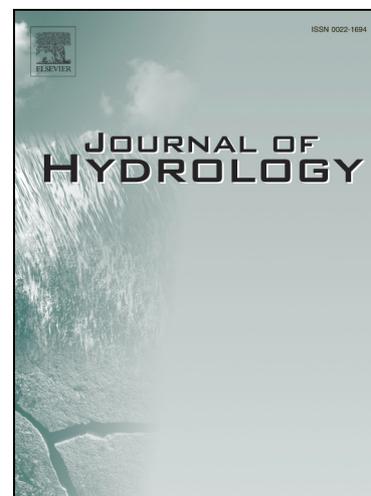
Please cite this article as: Pool, M., Carrera, J., Alcolea, A., Bocanegra, E., A comparison of deterministic and stochastic approaches for regional scale inverse modelling on the Mar del Plata aquifer, *Journal of Hydrology* (2015), doi: http://dx.doi.org/10.1016/j.jhydrol.2015.09.064

A comparison of deterministic and stochastic approaches for regional scale inverse modelling on the Mar del Plata aquifer.

M. Pool[a], J. Carrera[a], A. Alcolea[b], E. Bocanegra[c]

[a]*GHS, Institute of Environmental Assessment and Water Research (IDAEA), CSIC, Spain*
[b]*TK Consult AG, Zürich, Switzerland*
[c]*Instituto de Geología de Costas y del Cuaternario. UNMDP. Argentina*

Abstract

Inversion of the spatial variability of transmissivity (T) in groundwater models can be handled using either stochastic or deterministic (i.e., geology-based zonation) approaches. While stochastic methods predominate in scientific literature, they have never been formally compared to deterministic approaches, preferred by practitioners, for regional aquifer models. We use both approaches to model groundwater flow and solute transport in the Mar del Plata aquifer, where seawater intrusion is a major threat to freshwater resources. The relative performance of the two approaches is evaluated in terms of (i) model fits to head and concentration data (available for nearly a century), (ii) geological plausibility of the estimated T fields, and (iii) their ability to predict transport. We also address the impact of conditioning the estimated fields on T data coming from either pumping tests interpreted with the Theis method or specific capacity values from step-drawdown tests. We find that stochastic models, based upon conditional estimation and simulation techniques, identify some of the geological features (river deposit channels and low transmissivity regions associated to quartzite outcrops) and yield better fits to calibration data than the much simpler geology-based deterministic model, which cannot properly address model structure uncertainty. However, the latter demonstrates much greater robustness for predicting sea water intrusion and for incorporating concentrations as calibration data. We attribute the poor performance, and underestimated uncertainty, of the stochastic simulations to estimation bias introduced by model errors. Qualitative geological information is extremely rich in identifying large-scale variability patterns, which are identified by stochastic models only in data rich areas, and should be explicitly included in the calibration process.

Keywords: Real regional aquifer, deterministic-stochastic approaches, pilot points, plausible calibration.

1. Introduction

Groundwater model parameters are usually heterogeneous and uncertain. Fortunately, a large number of measurements of model outputs (most often heads and concentrations) are usually available, which favors formulating models in an inverse problem framework. The availability of model independent tools for inversion (Doherty et al., 1994;

Poeter and Hill, 1998) has fostered its common use in hydrological practice. Still, groundwater inversion suffers from a number of problems (non-uniqueness, instability, computational cost, etc.), as illustrated by the periodic reviews of the topic (Yeh, 1986; Carrera, 1987; McLaughlin and Townley, 1996; Zimmerman et al., 1998; de Marsily et al., 1999; Carrera et al., 2005; Franssen et al., 2009; Zhou et al., 2014). Amongst them, spatial variability is likely the most difficult problem to address. This work focusses on the characterization of spatial variability of transmissivity (T).

Two sets of approaches can be identified in dealing with spatial variability and the conceptual prior information: deterministic and stochastic. Stochastic approaches rely on treating T as a random field and estimating its properties from prior estimates of T (and/or other model parameters) and available state variables (i.e., heads (h) and concentrations (ω) measurements). Initially, the problem was formulated as finding the 'best field', in the sense of minimum variance or the expected field conditioned to measurements (Clifton and Neuman, 1982; Kitanidis and Vomvoris, 1983; Rubin and Dagan, 1987b). Estimation uncertainties derived from these solutions were too optimistic (Carrera and Glorioso, 1991). This prompted the development of conditional simulation methods (Gómez-Hernández et al., 1997), and more recently moment equations approaches (Hernández et al., 2003, 2006). All these methods assume stationarity of the T field. That is, they all assume that, prior to measurements, nothing is known about variability patterns, so that they effectively disregard geological knowledge. Although much work has been done on geologically based geostatistics (e.g., Winter and Tartakovsky, 2002; Winter et al., 2003; Riva et al., 2006), most applications of geostatistical inversion are based on this assumption of stationarity.

The deterministic approach relies on the assumption that the patterns of spatial variability are known from geological or geophysical information, so that the whole unknown field can be expressed in terms of a limited number of uncertain parameters. The process of expressing the field as a function of parameters is called parametrization. The most widely used parametrization method is zonation, which consists of subdividing the model domain into a number of regions (zones). Zones are typically (though not necessarily) homogeneous with a single effective parameter value (e.g., Carrera and Neuman, 1986; Barlebo et al., 2004). Therefore, in contrast to stochastic methods, the spatial zonation pattern, normally inferred from the available geological information, is prescribed explicitly in the deterministic approach. The main advantage of zonation is its flexibility to incorporate geological or geophysical data available in the form of maps (sedimentary deposits, paleochannels, water conducting features, etc). However, the identification of zones is a subjective and hard-to-systematize task. In fact, an inappropriate definition of zones is transmitted to errors in the model structure and is often a main cause of failure in actual applications (Sun et al., 1998). Some efforts have been devoted to alleviate the effect of errors in the geometry of zones (e.g., Gaganis and Smith, 2006; Roggero and Hu, 1998). However, no well-defined approach has emerged as generally accepted. Another disadvantage of the de-

2

terministic approach results from implicitly neglecting the effect of small scale heterogeneity on flow and, especially, on solute transport. Thus, the calibration process using the geostatistical approach captures the structural fine detail better than that which can be captured using a limited number of zones. However, despite of these problems, zonation remains the method of choice in hydrological practice, especially in regional or basin-scale groundwater models (e.g., Senger and Fogg, 1987; Guymon and Yen, 1990; Castro et al., 1998; Walvoord et al., 1999; Shavit and Furman, 2001; Best and Lowry, 2014; Ala-aho et al., 2015; Nocchi and Salleolini, 2013, etc).

The geostatistical approach has been applied successfully to cases where geological information is not strong enough to allow predefining patterns of spatial variability. This is certainly the case in synthetic aquifers (e.g., Yeh and Liu, 2000; Kowalsky et al., 2004; Zhu and Yeh, 2005, 2006; Alcolea et al., 2006b; Hao et al., 2008; Franssen et al., 2009; Riva et al., 2010, 2011) and laboratory sandboxes (e.g., Liu et al., 2002, 2007; Illman et al., 2007, 2009). Field applications are restricted to relatively small scale problems, where geology is not very binding, with well defined stresses and responses. These include hydraulic test interpretation (e.g., Meier et al., 2001; Vesselinov and Neuman, 2001; Li et al., 2005; Hernández et al., 2006; Rubin et al., 2010; Murakami et al., 2010; Janetti et al., 2010; Bianchi Janetti et al., 2010; Berg and Illman, 2011, 2013, 2015), well capture zone delineation (e.g., Vassolo et al., 1998; Kunstmann et al., 2002; Harrar et al., 2003; Riva et al., 2006), river-aquifer interaction (e.g., Rötting et al., 2006), or coastal aquifers (Alcolea et al., 2007, 2009) and others (Barlebo et al., 2004; Franssen and Kinzelbach, 2008; Vesselinov et al., 2001a,b; Chen et al., 2012).

Few stochastic studies have been carried out for large-scale problems. Clifton and Neuman (1982) and Rubin and Dagan (1987a,b) used a stochastic inversion approach to model the Avra Valley aquifers under steady-state conditions. Rubin and Dagan (1988) and Rubin et al. (1990) demonstrated the applicability of the geostatistical approach in the Israeli Coastal Aquifer and the Rio Maior aquifer, respectively. These examples proved useful in advancing the method during its early stages. However, they did not really demonstrate its validity, which is usually the case in real-word large scale applications because little data are available to test independently the model. In fact, to date, there is a very important lack of real-world applications of stochastic theories and approaches at large scale (Dagan, 2002; Neuman, 2004; Renard, 2007). Recent exceptions include the works of Jardani et al. (2012) and Dausman et al. (2015). This lack is much more marked when it comes to stochastic inversion of flow and transport data. We argue that the problem lies in ignoring geological information, which is much richer at large than at small scale. That is, patterns of geological variability (and, therefore, hydraulic variability) are usually well known at regional scale. Ignoring them would be poor practice. Thus, when facing a regional-scale model, professionals find that the easily accessible stochastic approaches fail to incorporate the geological information and prefer deterministic approaches. Instead, geological information at the local scale (<1km) is usually much less specific, and patterns of variability

3

cannot be stated a priori. In such cases, parsimony justifies stationarity as the prior model, which explains the broad use and success of small scale stochastic models.

Much effort has been dedicated to overcome this limitation of stochastic approaches. Perhaps, the most successful attempts are based on categorizing heterogeneity in terms of hydrogeological facies. Models of transition probabilities based on Markov chains (TPMC) analyze spatial variability and generate equally-likely realizations of geological units or facies. TPMC methods are a powerful geostatistical approach to estimate the spatial distribution of geological units using categorical indicator variables (e.g., Carle and Fogg, 1996, 1997; Fogg and Carle, 1998; Ritzi, 2000; Elfeki and Dekking, 2001; Park et al., 2004; Ritzi et al., 2004; Zhang et al., 2006; Li, 2007a,b; Dai et al., 2007; Zhang and Li, 2008; Ye and Khaleel, 2008; Khaninezhad et al., 2012a,b). These approaches have been generalized using multiple-point geostatistics and connectivity concepts (see Renard and Allard (2013) for a recent review). In practice, these approaches involve generating a large number of lithofacies distributions and/or hydraulic conductivity fields and rejecting those that fail to honor observed heads (e.g., Sakaki et al., 2009; Zhou et al., 2012; Alcolea and Renard, 2010; Berg and Illman, 2011; Khodabakhshi and Jafarpour, 2013). Other approaches aimed at reproducing actual variability patterns are based on geophysical data to compensate for the scarcity of in situ hydrological measurements and to improve the accuracy of spatial heterogeneity characterization (e.g., Rubin et al., 1992; Copty et al., 1993; Hyndman et al., 1994; Hubbard et al., 1997; Ezzedine et al., 1999; Hubbard and Rubin, 2000; Chen et al., 2001; Doro et al., 2013). Although all these techniques look promising, their application to real regional systems is still in developmental stages.

In light of the above considerations, the first relevant question is whether the geostatistical approach is suitable for modeling real regional aquifers or whether in these cases the inverse problem is best handled in a deterministic framework.

A complementary key question when using geostatistical inversion approaches is the source of T data. These are generally obtained from long-term pumping tests, which are expensive (and thus scarce). Alternatively, additional T data can be derived from specific capacity (pumping rate, Q, divided by drawdown, s). Specific capacity (SC) is the parameter most often provided by drillers from step-drawdown tests to characterize the performance of a well. Hence, SC data are often much more abundant than T data. Clifton and Neuman (1982) and Ahmed and Marsily (1987) argued that the estimation of a T field is improved by using both pumping test and specific capacity data. On the other hand, Meier et al. (1999) demonstrated that these two types of measurements usually represent different scales. Specific capacity measurements strongly depend on the transmissivity at the closest vicinity of the well whereas T values inferred from pumping test are sensitive to the test duration and tend towards the effective transmissivity with time. Therefore, a major challenge faced by hydrogeologists is the upscaling of T data from well-scale to aquifer-

scale, while taking into account the geological context. This prompts us to question the suitability of these two types of T measurements to model regional aquifers.

This study is aimed at addressing the two above questions. First, what are the relative merits of deterministic and geostatistical approaches for the modeling of flow and transport in real regional aquifers?, and second, are pumping and specific capacity measurements appropriate for the modeling of regional groundwater models?. In an attempt to answer these questions, we apply the two approaches and use both types of T data for the modeling of the Mar del Plata aquifer where data are available for nearly a century.

2. Study area

Mar del Plata is located in the northeastern coast of Argentina. The city grew at a rate of 10000 new inhabitants per year during the second half of the 20th century. This population growth caused an increase in groundwater pumping, the main source of freshwater. Intensive groundwater exploitation led to seawater intrusion. As a result, pumping was extended to the north. The original goal of the model was to assess the consequences of further increasing in pumping.

The aquifer extends over some 1000 km^2 and is traversed by several rivers; the Vivorat, the Cueros, the Seco, the Santa Elena and the Tapera from north to south, respectively (Figure 1). The main aquifer comprises Quaternary loess-like sediments that conformably overlie the quartzitic Palaeozoic bedrock at the south of the domain, and the Miocene clayey-sandy sediments at the north. Aquifer deposits were formed in response to the combined forcing of wind-deposited accumulation sediments, basement subsidence and fluvial processes.

Loess deposits are usually homogeneous, which is indeed the case at the north of the system, where the aquifer is unconfined. However, at the south, where the city is located, the local quartzitic hills acted as a barrier to wind transport and as a source for fluvial sediments. Therefore, the urban zone presents a number of discontinuous levels of silt and sand which are interbedded with thin layers of clayey silt. Thus, the main aquifer in Mar del Plata city is semi-confined and consists largely of sandy or silt-sandy sediments. It is assumed that the natural drainage network has not undergone dramatic changes during the past. Relict fluvial channels were infilled with sediments, as sea level rose, producing paleochannels. To the north, the Cueros and the Santa Elena paleochannels were infilled with high permeability sediments because these channels were sourced from eolian sands. These paleochannels act as preferential pathways for groundwater flow and solute transport, and are separated by the interfluves which contain sequences of silt clayey sediments. To the south, the Tapera and Chacra paleochannels were infilled with lower permeability sediments because they could have been derived from the Los Padres hills. The main aquifer in the southern portion is a semi-confined unit with a maximum thickness of 40 meters, separated from a shallow aquifer by an aquitard consisting of silt clayey sediments. The thickness of this aquitard decreases towards north, where the

5

aquifer becomes unconfined with a thickness of about 80 meters.

Infiltration from rain and irrigation return flow are the main inputs to the aquifer. The average annual recharge due to local rainfall has increased from 55.6 hm^3 per year in 1913 to 90.6 hm^3 per year in 2005. The increase concentrates in recent years, which may reflect climate change. A significant proportion of recharge in the urban area comes from losses of the water supply and sewer networks. Some lateral recharge enters the aquifer from the quartzitic hills. Pumping is the main discharge.

Groundwater flows mostly south-eastwards from the north-western hills toward the coast. Part of the recharge is pumped, leading to depressed heads in the central part of the aquifer. The remainder flows to the sea, with a small proportion discharging to rivers. A summary of the main inflows/outflows in the Mar del Plata aquifer from 1913 to 2005 is displayed in Table 1.

Transmissivity measurements (m^2/d) from long-term pumping tests are available at 109 locations. Furthermore, we estimated the transmissivity at the same locations from step-drawdown tests using a linear relation from the specific capacity values (SC). This type of measurement has been used for conditioning in the past (Ahmed and Marsily, 1987) and yields better information about the local value of T than pumping tests (Meier et al., 1999). Available data also include storage coefficient values estimated from long-term pumping tests at 21 locations, pumping records at 315 wells, monthly head records at 315 wells and 148 piezometers from 1913 to 2005, 7293 hydrochemical data, 3190 annual rainfall data, drillers' textural descriptions and geophysical data at 100 wells.

3. Methodology

We compare the performance of deterministic, geology based, and stochastic, head data based, models on the Mar del Plata aquifer. In the stochastic approach, we carry out both conditional estimations (CE) and conditional simulations (CS), using the regularized pilot points approach of Alcolea et al. (2006b). For each approach, we carry out two exercises. First, we use head and T data (from either pumping tests or specific capacity) for the calibration. The calibrated models are then used to simulate seawater intrusion to test the prediction capabilities of each approach. Second, we calibrate against both head and concentration data to test the robustness of each approach to the inclusion of new data.

3.1. Model set-up

The model domain is bounded by (1) a no-flow water divide to the north, (2) a prescribed flow boundary at the southern and western boundaries representing the contact with the quartzitic hills, where flow is assumed to be controlled by recharge, and (3) prescribed head at the coast.

A two-dimensional horizontal model is used because regional flow is essentially horizontal as the aquifer thickness (40-80 m) is small compared with the horizontal extent (some 1000 km^2) and because all available data come from wells screened throughout the aquifer thickness (vertically integrated). The aquifer is divided in two layers at the urban area because (1) the deep portions of the aquifers are separated from shallow levels by silty deposits, and (2) wells pump only at the deepest portions to prevent pollution from shallow groundwater. As a result of deep pumping, head differences along the vertical are sizeable in the urban area, and modeling salinization of pumping wells at the city requires extending the aquifer below the sea. The finite element grid consists of 7814 nodes, arranged in 17363 triangular elements of variable size, and it is divided into two horizontal layers. The upper layer depicts the shallower aquifer (only in the urban area) and the lower layer represents the main aquifer. These two layers are connected by 1D elements representing the aquitard. Note that a full three dimensional grid could be used to model the aquitard, but it would have been more CPU consuming (see, e.g., Bredehoeft and Pinder, 1970; Neuman et al., 1982). The submerged portion of the main (deep) aquifer is modeled by extending the lower layer 5 km seawards in the urban area. The mesh is refined in the urban area and near the main pumping areas to ensure accuracy (average size 60m x 60m). The element size increases to the north where few data are available (average size 1000m x 1000m). Transmissivity was treated as spatially variable as discussed in the following sections. Storage coefficient was assumed spatially constant because significant spatial variability of this aquifer property is not expected in this kind of sediments. Two different values for the storage coefficient (S) have been adopted: $S=0.1$ for the shallow aquifer and for the main aquifer at the north where the aquifer becomes unconfined, and $S=7\cdot10^3$ for the main aquifer in the urban area.

Recharge from precipitation was calculated from a root zone mass balance, with potential evapotranspiration evaluated according to the Penman-Monteith equation, using meteorological data from the stations Mar del Plata and La Serrana, situated in the urban area and Los Padres hills, respectively. Five zones of recharge are defined: the plain area (recharge due to rainfall), the irrigated zone (recharge due to rainfall and irrigation return flow), and three urban zones that account for losses of the water supply and sewer networks. While these zones can be defined deterministically, variability within each zone should be restricted to scales much smaller than the model (e.g., roads) and was neglected here. Although the aquifer receives lateral inflow from the quartzitic hills Los Padres, the actual amount is small, see Table 1.

Groundwater pumping from 1913 to 2005 is simulated as a prescribed time dependent sink term at each well (total of 107 nodes representing pumping wells) or incorporated into the model over sets of nodes (a total of 27 zones were considered), where pumping wells are closely spaced. All wells are screened in the main aquifer. Surface - groundwater interactions at rivers and Los Padres lake are simulated as drains using Cauchy boundary conditions, with mean river elevation as external heads. Preliminary runs showed that models results were not very sensitive to

7

leakage factors, which were fixed to 0.2 m^2/d per unit length of river. In practice, this implies that rivers acted close to prescribed head. It is assumed that river bed shapes did not change over time.

Density variations have been neglected because the aquifer bottom is essentially flat and the thickness small. Under these conditions, buoyancy forces are small compared to viscous forces which are large for regional aquifers. Therefore, we have adopted an approach similar to that of Iribar et al. (1997) and Vázquez-Suñé et al. (2006). That is, the shore is represented by imposing a freshwater head equivalent to a zero saltwater head for flow and an advective flux condition for transport. This implies that the salt mass fraction equals that of seawater for inflowing portions of the boundary or the resident mass fraction for outflowing portions (see, e.g., Voss and Souza, 1987; Frind, 1982).

The transmissivity field was calibrated for the period from 1913 to 2005, with a monthly temporal discretization.

3.2. Calibration Procedure

Transmissivity was calibrated by transient inverse modeling. Following Medina and Carrera (2003), the optimum set of model parameters minimizes an objective function that penalizes (1) the differences between measured and calculated heads and/or concentrations and (2) the offsets of calibrated model parameters from their corresponding prior estimates. In this study the objective function was defined as

$$F = F_h + \beta_\omega F_\omega + \mu_j F_j \tag{1}$$

with

$$F_h = (\mathbf{h} - \mathbf{h}^*)^t \mathbf{V}_h^{-1} (\mathbf{h} - \mathbf{h}^*) \quad \text{(OBJHEAD)} \tag{1.1}$$

$$F_\omega = (\omega - \omega^*)^t \mathbf{V}_\omega^{-1} (\omega - \omega^*) \quad \text{(OBJCON)} \tag{1.2}$$

$$F_j = (\mathbf{p}_j - \mathbf{p}_j^*)^t \mathbf{V}_{\mathbf{p}_j}^{-1} (\mathbf{p}_j - \mathbf{p}_j^*) \tag{1.3}$$

In these equations $\mathbf{x}$ ($\mathbf{h}, \omega$ or $\mathbf{p}_j$) represents the vector of computed values, with measurements (or prior estimates) $\mathbf{x}^*$ and covariance matrix $\mathbf{V_x}$ with $\mathbf{x} = h$ for heads, $\mathbf{x} = \omega$ for concentrations and $\mathbf{x} = j$ for estimated model parameters, where the subscript $j = 1$ denotes pilot points linked to transmissivities, $j = 2$ to storage coefficients, etc. In the present study, only transmissivity values were calibrated ($j = 1$). Thus this subscript is omitted hereinafter. β_ω and μ are weighting scalars correcting errors in the specification of $\mathbf{V_x}$. The objective function is minimized with respect to

8

model parameters for given values of β_ω and μ. We assumed that the standard deviation of head residuals is 1 m (3.5% maximum drawdown). Therefore, $\mathbf{V_h} = \mathbf{I}$. Concentration residuals were assumed heteroscedastic, with standard deviation equal to max (10% ω, ω_{min} mg/L) where $\omega_{min} = 100$ mg/L (except for the observation wells TP8 and TP32 where $\omega_{min} = 500$ mg/L). The covariance matrix of transmissivities depends on the heterogeneity assumptions and is described in the next section. Maximum likelihood estimation, MLE, is used to find the optimum values of β_ω and μ, as described by Alcolea et al. (2006b). Eight values of μ were tested, ranging from $5 \cdot 10^3$ to 10^{-2} (lower values were discarded due to convergence and stability problems). High values of μ result in a poor identification of heterogeneity because the objective function (1) is controlled by the by the plausibility term (1.3) and the calibrated parameters tend towards the prior information which is usually smooth. On the contrary, small values of μ tend to disregard the plausibility term, thus risking instability. The optimum value of μ was chosen as the one yielding the maximum of the expected likelihood of the parameters given the data, equivalent to the minimum of the support function S_2 defined by Medina and Carrera (2003), which reads as

$$S_2 = N + \ln|\mathrm{H}| + N \ln \frac{F}{N} - n_\omega \ln \beta_\omega - \sum_{j=1}^{\mathrm{ntypar}} k_j \ln \mu_j \tag{2}$$

where, N is the total number of data, n_ω and k_j are the number of concentration measurements and parameter type j data (up to the number of types of model parameters being optimized, 'ntypar', equal to 1 in this case, as only T was estimated), respectively, and H is the first order approximation of the Hessian matrix of the objective function at the end of the optimization process. Using this criterion, we chose $\mu=1$ and $\beta_\omega=1$ as optimal, which are the values to be expected if the covariance matrices are appropriate. Figure 3 displays the support function S_2 obtained by conditional estimation with Theis and SC data versus the weighting factors μ and β_ω. When only head data are used, S_2 is minimum for $\mu=1$. Using this plausibility weight, the optimal β_ω is 0.1 for SC data and 1 for Theis data. We chose $\beta_\omega=1$ for consistency and because the difference between 0.1 and 1 in the SC case is small. Actually, even larger weights for concentration data could be obtained when adopting larger values of μ (not shown here).

Two different cases are considered regarding conditioning data. First, the transmissivity fields are conditioned to transmissivity and heads. We evaluate these fields by comparing actual concentration measurements to blind predictions of solute transport from 1913 to 2005. Second, we added concentration data to the calibration. That is, the transmissivity fields were conditioned to transmissivity, head and concentration data. Although we could not assess independently the validity of this calibration because we were using all available data, we evaluated the quality of the fits to available concentration variation measurements and the conceptual validity of the resulting T fields.

9

3.3. Handling of spatial variability

3.3.1. Deterministic approach

The transmissivity zonation chosen in the deterministic model is based on data integration and geological review. That is, we base our zonation upon (1) drillers' textural descriptions, (2) geophysical logs, and (3) a sedimentation concept that suggests EW lineations. Zone boundaries are chosen to separate areas of significantly different transmissivity as observed in the true field. The actual delineation of the paleochannels was based on sedimentological columns of the wells coming from drillers descriptions and geophysical logs.

Eight transmissivity zones are defined: four paleochannels representing the relict fluvial channels; the interfluves zone; the main aquifer in the urban area, which is confined; the shallow aquifer, upper layer in the urban area and the piedmont zone, which corresponds to coarse piedmont deposits. Prior estimates of these eight parameters were obtained from pumping tests and geological information. Additionally, the vertical conductivity of the aquitard in the urban zone is also estimated.

3.3.2. Geostatistical approach

The regularized pilot points method (PPM) (Alcolea et al., 2006b) was used for solving inverse problems with stochastically variable T fields. The PPM has been the subject of some criticism (Huang et al., 2011) on the basis that it does not provide sufficient flexibility to accommodate all the fluctuations required to fit the data. This argument might be valid in the original PPM, where only a few pilot points were used in order to ensure stability (de Marsily et al., 1984; Ramarao et al., 1995). To avoid this limitation, Alcolea et al. (2006a,b) suggest using as large a number of pilot points as computationally feasible. A rule of thumb is to use two or three pilot points per correlation range (Gómez-Hernández et al., 1997), which can be relaxed in areas where few data are available. The fact is that the numerous methods available for inversion yield very similar results in comparison tests (Zimmerman et al., 1998; Franssen et al., 2009).

Actual inversion can be performed to obtain conditional estimation (CE) or conditional simulation (CS) of T fields. The chosen method depends on whether one is interested in seeking a single 'best' model (i.e., CE, or expected value of logT conditioned to measurements), which may be the most appropriate for comparison with the deterministic approach, or in obtaining an ensemble of T fields that reproduce small scale variability and honor measurements, which may be most appropriate for performing uncertainty analyses based on Monte Carlo methods. T at every element is obtained through kriging (CE), or through conditional simulation. These fields are then perturbed by estimating T at pilot points so as to minimize (1) (see Alcolea et al., 2006a,b, for details). The CE is the best estimate in the sense of minimum variance of T, whereas each CS yields an equally likely random T field. As such,

10

each simulation reproduces small scale variability and the ensemble of them quantifies uncertainty. We obtained 15 simulations conditioned to all T and head measurements and another 15 conditioned additionally to concentration measurements.

Three non-regular networks of pilot points were tested, containing 133, 304 and 403 pilot points. Calibration errors decrease substantially when using more than 133 pilot points. However, little improvement was obtained with 403 points. Therefore, we used 304 pilot points, which leads to enough accurate characterizations of the T field while keeping the CPU effort tolerable. This choice was also supported by the expected likelihood (Medina and Carrera, 2003). Pilot point locations are depicted by dots in Figure 5. Note that their locations do not coincide with T measurement points. This choice will cause all T fields to honor T measurements, which are implicitly assumed to be error free, but which is the best option given that the experimental variograms did not display a nugget effect (see Figure 2).

3.3.3. Theis vs specific capacity data

T measurements arising from the Theis interpretation of pumping tests (Theis) and from specific capacity (SC) using a linear relation (T= 1.4 Q/s, Razack and Huntley (1991)), are available at 109 locations. The probability density function of both sets of measurements is close to lognormal. This is in agreement with results from other aquifers (see, e.g., Delhomme, 1976; Ahmed and Marsily, 1987). However, contrary to these, our T and SC values were not significantly correlated (Figure 2). Also, contrary to most applications, our Theis and SC measurements are co-located. While conventional wisdom would call for ignoring SC data, we opted for testing the use of SC data instead of Theis data on the assumption that they may represent better the local variability (Meier et al., 1999). In short, we carry out our calculations twice, first, with Theis data and, second, with SC data.

Note that Theis T values are higher than SC T values (log T means of 2.98 and 2.53, respectively). We attribute this to the scale dependent growth of T (see, e.g., Sánchez-Vila et al., 1996; Martinez-Landa and Carrera, 2005) and to the fact that long-term pumping test T values are representative of a region much larger than SC T values. Both sets of measurements show a very low variance, which facilitates parameter estimation (Yeh, 1986), and reflects the geological nature of the aquifer. This little variability can be observed in the estimated variograms of both data sets (actually, of their $\log_{10}$ transformations), see Figure 2. Both of them are exponential with low sill (0.034 and 0.036 for T from Theis and and SC, respectively). The ranges are 4770 and 3510 m for Theis and SC, respectively.

4. Results

The water balance obtained from the deterministic model is displayed in Table 1. Similar outputs were obtained from all the stochastic simulations. Note that Santa Elena and La Tapera rivers changed from losing to gaining

11

conditions at the end of the 60s. This change reflects the decrease in pumping at the city area due to the salinization of pumping wells and the increase of rainfall. Note also the increase in recharge in the urban area from 1950 to 2005, which is caused by the increase in leakage of the water supply and sewerage systems, and occurs in many cities (Vázquez-Suñé et al., 2010)

Figure 4 summarizes the results obtained in terms of F_h (OBJHEAD, Eq. 1.1) and F_ω (OBJCON, Eq. 1.2) to assess the model fits for all runs with the two approaches. Only 6 out of 15 conditional simulations are shown for each set of runs, as they are quite similar (Figure 4a,c). Results obtained when calibrating only against heads are presented for three different values of μ (100,10 and 1). For those runs, prediction capability is assessed by means of F_ω. Results obtained when calibration to head, concentration and transmissivity data is carried out are presented for $\mu=1$ and $\beta_\omega=1$. The columns with red triangles depict the scores of the deterministic approach.

The first observation that becomes apparent from Figure 4 is that a satisfactory fit between model outputs and field measurements is obtained with all the models, even for the simulation yielding the worst fit to measured head and concentration data.

When calibrating with heads (Figure 4a), very similar fits are obtained by the deterministic and all stochastic models with $\mu=100$. As expected, lower values of the weighting factor μ in the stochastic models yield better fits, and the best fit is attained when $\mu=1$. Note that CE and CS approaches yield very similar results regardless of the conditioning T data. Note also that an improved fit does not mean a better model. As mentioned above, we use the expected likelihood for selecting the optimal value of μ.

Transport predictions (Figure 4c) are generally quite accurate. However, contrary to expectations, the models that were optimal for flow ($\mu=1$) perform worst for transport. The best transport predictions are obtained by the models calibrated with $\mu=10$ when using Theis data, and comparable results are obtained for the $\mu=100$ and $\mu=10$ models when using SC data.

When calibrating against h and ω data (Figure 4b,d), we would expect a slight worsening of F_h and a significant improvement of F_ω because head fits need to be slightly sacrificed to achieve accurate fits to concentration data. As it turns out, this happens only for the stochastic models calibrated with Theis data. However, F_h for the deterministic model is unaffected when including concentration data in the calibration, while F_ω decreases considerably.

Simulation results show that the calibrated T distribution obtained from the deterministic model is geologically realistic and consistent with the prior information embedded in the conceptual model described in section 2. The T fields generated by all calibration runs with $\mu=1$ are displayed in Figures 5 and 6. Figure 5 displays the deterministic field obtained by the zonation approach and the 'single best' one by conditional estimation. It is worth noting that calibrated transmissivity values in the deterministic model are not significantly different when calibrating against

12

heads and when calibrating against both heads and concentrations. Figure 6 illustrate three simulations of T fields conditioned to T and h and to T, h and ω. Overall, the stochastic fields reproduce the main patterns of the geological model. Specifically, most models reproduce the hight T channels corresponding to Los Cueros, Seco and Santa Elena river, and the low T channel of La Tapera river. They also tend to reproduce the high T zone of piedmont deposits. In fact, all models but those conditioned to Theis, head and concentration data depict several features that were not included in the deterministic model: two low T zones at the southern and north-east of the city aquifer (red squares in Figure 5). These may be caused by outcropping quartzite deposits that were not included in the deterministic model. The definition of these variability patterns is somewhat sharper for the models conditioned to SC data than for the models conditioned to Theis data.

Figure 7 illustrates the effect of increasing μ on the transmissivity fields obtained by conditional estimation and simulation. As expected, large values of the weighting factor cause the T fields to become too smooth. The general shape of the paleochannels is reproduced by almost all stochastic simulations, and transmissivity values are adequate. Note that the fields with μ=10, optimum value for predicting transport, vaguely capture the main features of the geology-based model.

Model fits to measured heads and concentrations are shown in Figures 8, 9 and 10. The fits to measured heads from the deterministic and stochastic methods using Theis T data and a plausibility weight of μ=10 are displayed in Figure 8. Satisfactory fits are obtained by both stochastic (CE and CS) and deterministic (zonation) approaches. The small misfit (average residual is 1.3m, a 4.4% of the maximum drawdown) can be attributed to the inherent uncertainty of pumping rate records. The corresponding transport predictions are shown in Figure 9. It can be observed that all models tend to overestimate concentrations. The goodness of the blind predictions is most prominent at low chloride concentrations. At high concentrations, only the shape of the measured breakthrough curves but not the actual concentration values are reproduced. We attribute this to the fact that transport predictions depend on features of the system beyond the level of detail that can be attained by calibrating against head data only (Moore and Doherty, 2006). Figure 10 displays the fit to concentration data when both head and concentration data (besides Theis T data) are used for the calibration with the optimum plausibility weights (μ=1 and β_ω=1). Note that adding concentrations as calibration date sets helps to improve concentration fits.

5. Discussion

We have compared deterministic (zonation) and stochastic (regularized pilot points) approaches for transmissivity (T) inversion at the Mar del Plata regional aquifer. The T zonation patterns defined in the deterministic model were based on data integration and geological review. Spatial variability of T was characterized by the regularized pilot

13

points method using two available sets of transmissivity data, i.e., from pumping test (Theis) and from specific capacity (SC). Two different inverse problems have been considered: (1) T fields conditioned to T (either Theis or SC) and head data, which were then used to predict transport, and (2) T fields conditioned to T, head and concentration data to assess whether the addition of concentrations data renders a more accurate spatial characterization of the T field.

The two approaches were evaluated in terms of (1) the fit to head (F_h) and concentration (F_ω) measurements, (2) the plausibility of the estimated T fields, and (3) the prediction capabilities when calibrating to T and head data only.

The following observations can be made after analyzing the results:

1) Head fits were quite good and slightly better for stochastic than for deterministic runs. However, the T fields obtained with each method are quite different. This single observation points to the high degree of non-uniqueness of the solution to the inverse problem in that different T fields yielded similar head fits. Note that differences lie in the modeling approach (SC versus Theis data, calibrating against heads or concentrations, value of μ), but within each approach, differences are minimal. However, within each modeling approach, all conditional simulations look very similar. This is somewhat disheartening in view of (1) the large number of head data, (2) the length of the calibration period, (3) the high and well characterized forcing terms imposed on the system, and (4) the relative homogeneity of this aquifer. In most situations, few of these conditions are met and, therefore, an even higher degree of non-uniqueness should be expected. Still, if we had been interested in using the model to assess future groundwater management strategies or to evaluate the uncertainties on transport calculations (Renard and Jeannee, 2008), none of the T fields should have been discarded.

2) The estimation of the weighting factor μ turned out to be critical. The optimum value of μ ($\mu = 1$) yielded good head fits during calibration but led to poor predictions of concentration. This suggests a certain degree of over-parametrization. Indeed, stochastic T fields obtained with $\mu=1$ (Figures 5 and 6) fluctuate much in space. These fluctuations are dramatically smoothed out when increasing μ (Figure 7). This is not surprising because large plausibility weights tend to bias the solution towards kriging, effectively reducing degrees of freedom (Alcolea et al., 2006b). In contrast, the deterministic model is not very sensitive to the plausibility weight. In this case, the optimum value for the deterministic model is attained at $\mu=1$, as it was for all stochastic models. A close analysis of all stochastically calibrated models for a given μ reveals three observations. First, all conditional simulations (CS) yield very similar values of F_h and the results are slightly lower than those obtained by conditional estimation (CE). Second, the fits obtained using Theis T data are slightly but consistently better than those achieved using SC data. Third, all T fields are quite similar. This result was also obtained by Alcolea et al. (2006a), but it is disturbing because it leads to a false sense of certainty. Uncertainty is much larger than suggested by the set of conditional simulations.

3) Blind predictions of transport are generally quite accurate for all flow calibrated models. However, contrary

14

to expectations, the best matches to concentration data were obtained with T fields obtained with high plausibility weights. This may reflect a certain degree of instability (i.e., overfitting when $\mu=1$) or model errors, which favors stable calibration procedures (methods to identify μ have been tested on synthetic examples, which are free of model errors). Instability implies large fluctuations of estimated parameters, which lead to (a) anomalously large (or low) parameter estimates and (b) large 'jumps' of parameter values over small distances. Instabilities may result in very small calibrations errors, but may seriously degrade the model ability for accurate and realistic transport predictions.

4) In stark contrast to head calibrations, slightly better concentration predictions were obtained with conditional estimations than with conditional simulations (Figure 4c). This suggests that the type of small scale variability simulated during flow calibration may not be the most appropriate for transport. Although similar transmissivity fields are obtained using Theis and SC T data during head calibration, Theis models led to slightly but consistently better transport predictions than SC models.

5) Theoretically, calibration against head and concentration data, should cause a slight sacrifice in head fit to achieve a good fit to concentration data. This effect is only evident in the stochastic models using transmissivity values from pumping tests. However, in these cases there are several noticeable differences between the transmissivity fields obtained by calibration of heads and those by calibration of heads and concentrations. The latter loose large scale heterogeneity patterns but lead to better match to measured concentration data. Stochastic models with transmissivity data from specific capacity yield values for both F_h and F_ω similar to those calculated when calibrating heads only. Several reasons may explain this anomalous behavior. First, transmissivity from specific capacity represents the local transmissivity at the closest vicinity of the well, and seawater intrusion can be more sensitive to large scale heterogeneity patterns, which are better captured by Theis T data. Moreover, transmissivity values from specific capacity data are smaller than those from pumping test data (Figure 2), which may be relevant when calibrating head and concentration data. Another likely reason is that the resulting inverse problems with specific capacity data do not have unique solutions. Thus, many different transmissivity fields may yield equally good fits to the available measurements. By contrast, calibration of the deterministic model improves the concentration fit while maintaining an adequate match to head data, without large changes in transmissivity values. These results clearly support that the deterministic model is the most robust, in the sense that it was the least sensitive to calibration data, although the head and concentration fits are not the best.

6) All calibrated transmissivity fields display some consistency with the conceptual geological model. Specifically, the overall shape of the paleochannels is reproduced to some extent by all stochastic models. At the north, Los Cueros and Santa Elena paleochannels were infilled with high permeability sediments, causing high transmissivity values in these zones. On the contrary, the sediment deposition infilling La Tapera and Chacra paleochannels could have been

15

derived from the Los Padres hills. Thus, these zones are attributed with low transmissivity values by all stochastic and deterministic models. The geometry of the paleochannels produced by the stochastic models is slightly different to the one we proposed for the deterministic model. The discussion of this last observation is subtle. The existence and roles of paleochannels result from a geological reasoning. However, defining their shape is a very subjective task (specially near the shore, where T and head data are scarce). On the one hand, the fact that stochastic models capture the overall geometry of the paleochannels supports the applicability of these methods where data are abundant. Identifying the presence of these paleochannels from noisy and highly fluctuating data is non-trivial and demonstrates that stochastic inversion can capture features that are not evident from data. On the other hand, the fact that these paleochannels are identified with a different connection to the coast where data are scarce points to the limitations of our simplistic geostatistical model (stationarity of T).

The stochastic models also identify consistently two low T zones located at the NE and SE of the city (identified by red squares in Figure 5) that were not included in the deterministic model. These coincide with underground quartzite (it outcrops in several locations), which is known by local geologists, so that few wells are available in the area. We could have modified the deterministic model to account for them, but would have made the comparison unfair. As it is, it illustrates a problem with deterministic approaches. Errors in the proposed zonation will go unnoticed, underscoring the inability of deterministic models for properly accounting for errors in model structure.

7) Most stochastic transmissivity fields obtained by conditional simulation and conditional estimation approaches yield similar fits to h and ω measurements. But stochastic simulations look alike in all cases, which implies a clear underestimation of uncertainty. We attribute the similarity of all simulations to biases introduced by model errors. Model errors are unavoidable in real situations and represent simplifications of reality that may not be appropriate. Examples include boundary conditions (e.g., the sea side boundary condition may be more complex than assumed), the relatively simple assumption of homogeneous recharge zones might not be accurate, etc. Deterministic models, having much less degrees of freedom, appear to be less sensitive to these simplifications. In fact, the robustness of the deterministic model results suggests that such errors were not large.

6. Conclusions

The conclusions from this work are subtle. Stochastic models led to quite good head fits and reproduced some of the connectivity patterns implicit in the geological model. However, the fitted models were frail, in the sense that they changed dramatically in response to changes in model assumptions (value of the plausibility weight, use of head or concentration data, use of T or SC data). The changes were generally restricted to areas where data are scarce. This is true both at the northern portion of the model and near the coast. The latter was especially relevant for transport

16

models because connectivity to the sea is what controls sea water intrusion. The deterministic models yielded almost as good fits to heads and concentrations as the stochastic models, and estimated T fields were not very sensitive to the incorporation of new calibration data sets. Therefore, we must conclude that the deterministic models of the Mar del Plata aquifers are more robust and reliable than stochastic models.

It has been observed that transmissivity data inferred from pumping tests (Theis) yields better fits than those obtained using specific capacity data (SC). However, models conditioned to SC data yielded sharper definitions of expected variability patterns than those conditioned to Theis T data, which is consistent with the assumption that SC data provide good estimates of local T. This observation, together with the fact that SC data will be generally much more abundant than Theis data, definitely supports the use of SC data in regional aquifer modeling.

The most disturbing result of the stochastic calibration exercise was their misleading certainty. Stochastic simulations aim at assessing uncertainty and the range of possible model predictions under future scenarios. Yet, all stochastic T fields simulated under a given assumption (i.e., μ value or transmissivity data type) were very similar. This would have led modelers to a false feeling of confidence and to assume that data quality and quantity are good enough to dramatically reduce uncertainty. As it turned out, estimated fields changed dramatically when modifying data (SC versus Theis, h versus h and ω) or plausibility weight. This behavior must be qualified as stubborn and jumpy. We attribute the stubbornness to model errors inherent in any real application (e.g., wrong plausibility set, errors in data, in types of boundary conditions, etc.). We assume that stochastic methods correct the effect of model errors by modifying the T field, but the correction will be always the same (hence the similarity of the CS T fields). However, the effect of model errors (and the required 'correction' in the T field) will be extremely sensitive to data and to the simulated phenomenon, which explains the jumpy behaviour (abrupt changes in estimation in response to small changes in assumptions).

This exercise was singular in the sense that a relatively large head data set was available over a long period of time. Point values of T were also unusually numerous. These features should have favored stochastic calibration. The fact that stochastic models did not perform way better than the deterministic method of zonation in this case, points to the importance of geological information and suggests that deterministic methods will produce better overall regional models than stationary stochastic methods whenever geological information is relevant.

This finding should not be taken as an unqualified support of deterministic models. Clearly, the main disadvantage of deterministic models lies in their inability to produce reliable estimates of uncertainty (parameter covariances are heavily constrained by geometry and result far too optimistic). That is, errors in the proposed zonation are not acknowledged in the deterministic approach, and can only be inferred through careful residual analysis (large errors in a portion of the model may suggest the need to modify zonation). Therefore, deterministic models are not the solution

when reliable estimates of uncertainty are required.

We would like to conclude by insisting on the stochastic models ability to capture patterns of variability consistent with geology (paleochannels, possible reduction in T at the south and northeast of the city caused by quartzitic deposits) in areas where data are abundant. We presume that the poor performance of stochastic models was caused by the fact that no information was provided on how to extend these features. This suggests that the future lies in coercing stochastic models to geologically based information, so that hard data will constrain the model where abundant and 'soft' geological data will constrain the model elsewhere.

Acknowledgments

References

Ahmed, S., Marsily, G.D., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resour. Res. 23(9), 1717–37. doi:10.1029/WR023i009p01717.

Ala-aho, P., Rossi, P.M., Isokangas, E., Klove, B., 2015. Fully integrated surfacesubsurface flow modelling of groundwaterlake interaction in an esker aquifer: Model verification with stable isotopes and airborne thermal imaging. Journal of Hydrology 522, 391–406. doi:10.1016/j.jhydrol.2014.12.054.

Alcolea, A., Carrera, J., Medina, A., 2006a. Inversion of heterogeneous parabolic-type equations using the pilot points method. International Journal for Numerical Methods in Fluids 51(10), 963–80.

Alcolea, A., Carrera, J., Medina, A., 2006b. Pilot points method incorporating prior information for solving the groundwater flow inverse problem. Adv. Water Res. 29, 1678–1689. doi:10.1016/j.advwatres.2005.12.009.

Alcolea, A., Castro, E., Barbieri, M., Carrera, J., Bea, S., 2007. Inverse modeling of coastal aquifers using tidal response and hydraulic tests. Ground Water 45, 711–22. doi:10.1111/j.1745-6584.2007.00356.x.

Alcolea, A., Renard, P., 2010. Blocking moving window algorithm: Conditioning multiple-point simulations to hydrogeological data. Water Resources Research 46(8). doi:10.1029/2009WR007943.

Alcolea, A., Renard, P., Mariethoz, G., Bertone, F., 2009. Reducing the impact of a desalination plant using stochastic modeling and optimization techniques. Journal of Hydrology 365, 275 –88. doi:10.1016/j.jhydrol.2008.11.034.

Barlebo, H.C., Hill, M.C., Rosbjerg, D., 2004. Investigating the macrodispersion experiment (made) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model. Water Resour. Res. 40. doi:10.1029/2002WR001935.

Berg, S.J., Illman, W., 2011. Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system. Water Resources Research 47. doi:10.1029/2011WR010616.

Berg, S.J., Illman, W.A., 2013. Field study of subsurface heterogeneity with steady-state hydraulic tomography. Ground Water 51, 29–40. doi:10.1111/j.1745-6584.2012.00914.x.

Berg, S.J., Illman, W.A., 2015. Comparison of hydraulic tomography with traditional methods at a highly heterogeneous site. Ground Water 53, 71–89. doi:10.1111/gwat.12159.

Best, L.C., Lowry, C.S., 2014. Quantifying the potential effects of high-volume water extractions on water resources during natural gas development: Marcellus shale, {NY}. Journal of Hydrology: Regional Studies 1, 1 – 16. doi:10.1016/j.ejrh.2014.05.001.

Bianchi Janetti, E., Riva, M., Straface, S., Guadagnini, A., 2010. Stochastic characterization of the montalto uffugo research site (Italy) by geostatistical inversion of moment equations of groundwater flow. Journal of hydrology 381, 42–51. doi:10.1016/j.jhydrol.2009.11.023.

Bredehoeft, J.D., Pinder, G.F., 1970. Digital analysis of areal flow in multiaquifer groundwater systems: A quasi three-dimensional model. Water Resources Research 6, 883–8. doi:10.1029/WR006i003p00883.

Carle, S.F., Fogg, G.E., 1996. Transition probability-based indicator geostatistics. Math. Geol. .28(4), 453–76. doi:10.1007/BF02083656.

Carle, S.F., Fogg, G.E., 1997. Modeling spatial variability with one and multidimensional continuous-lag markov chains,. Math. Geol., 29(7),, 891–918. doi:10.1023/A:1022303706942.

Carrera, J., 1987. State of the art of the inverse problem applied to the flow and solute transport equations. Groundwater Flow and Quality Modelling 224, 549–83. doi:10.1007/978-94-009-2889-3-31.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., Slooten, L., 2005. Inverse problem in hydrogeology. Hydrogeology Journal 13, 206–22. doi:10.1007/s10040-004-0404-7.

Carrera, J., Glorioso, L., 1991. On geostatistical formulations of the groundwater flow inverse problem. Adv. Water Res. 14 (5), 273–83. doi:10.1016/0309-1708(91)90039-Q.

Carrera, J., Neuman, S., 1986. Estimation of aquifer parameters under transient and steady state conditions: 2. uniqueness, stability, and solution algorithms. Water Resour Res 22, 211–27. doi:10.1029/WR022i002p00211.

Castro, M.C., Jambon, A., de Marsily, G., Schlosser, P., 1998. Noble gases as natural tracers of water circulation in the paris basin 1. measurements and discussion of their origin and mechanisms of vertical transport in the basin. Water Resour. Res. 34, 2443–66. doi:10.1029/98WR01956.

Chen, J.S., Hubbard, S., Rubin, Y., 2001. Estimating the hydraulic conductivity at the south oyster site from geophysical tomographic data using bayesian techniques based on the normal linear regression model. Water Resour. Res. 37, 1603–13. doi:10.1029/2000WR900392.

Chen, X., Murakami, H., Hahn, M., Zachara, J., Rubin, Y., 2012. Three-dimensional bayesian geostatistical aquifer characterization at the hanford 300 area using tracer test data. Water Resources Re. 48. doi:10.1029/2011wr010675.

Clifton, P., Neuman, S., 1982. Effects of kriging and inverse modeling on conditional simulation of the Avra valley Aquifer in southern Arizona. Water Resour. Res. 18, 1215–34. doi:10.1029/WR018i004p01215.

Copty, N., Rubin, Y., Mavko, G., 1993. Geophysical-hydrological identification of field permeabilities through bayesian updating. Water Resour. Res., 29(8), 2813–25. doi:10.1029/93wr00745.

Dagan, G., 2002. An overview of stochastic modeling of groundwater flow and transport: From theory to applications. EOS Transactions AGU 83, 621–5. doi:10.1029/2002EO000421.

Dai, Z., Wolfsberg, A., Lu, Z., R. Ritzi, J., 2007. Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix. Geophys. Res. Lett. 34. doi:10.1029/2007gl031608.

Dausman, A.M., Doherty, J., Langevin, C.D., Dixon, J., 2015. Hypothesis testing of buoyant plume migration using a highly parameterized variable-density groundwater model at a site in florida, usa. Hydrogeology Journal 18, 147–60. doi:10.1007/s10040-009-0511-6.

Delhomme, J.P., 1976. Application de ia theorie des variables regionalisees dans les sciences de l'eau. These de docteur-ingenieur, Universite Pierre et Marie Curie.

Doherty, J., Brebber, L., Whyte, P., 1994. Pest: Model-independent parameter estimation. Watermark Computing, Corinda, Australia. .

Doro, K., Leven, C., Cirpka, O., 2013. Delineating subsurface heterogeneity at a loop of river steinlach using geophysical and hydrogeological

methods. Enviromental Earth Sciences 69 (2), 335–48. doi:10.1007/s12665-013-2316-0.

Elfeki, A., Dekking, M., 2001. A markov chain model for subsurface characterization: Theory and applications. Math. Geol., 33(5), 569–89. doi:10.1023/A:1011044812133.

Ezzedine, S., Rubin, Y., Chen, J., 1999. Bayesian method for hydrogeological site characterization using borehole and geophysical survey data: Theory and application to the lawrence livermore national laboratory superfund site. Water Resour. Res. 35(9), 2671–83. doi:10.1029/1999wr900131.

Fogg, G. E., C.D.N., Carle, S.F., 1998. Geologically based model of heterogeneous hydraulic conductivity in an alluvial setting. Hydrogeol. J., 6(1), 131–43. doi:10.1007/s100400050139.

Franssen, H.J.H., Alcolea, A., Riva, M., Bakr, M., van der Wiel, N., Stauffer, F., Guadagnini, A., 2009. A comparison of seven methods for the inverse modelling of groundwater flow. application to the characterisation of well catchments. Adv. Water Res. 32, 851–72. doi:10.1016/j.advwatres.2009.02.011.

Franssen, H.J.H., Kinzelbach, W., 2008. Real-time groundwater flow modeling with the ensemble kalman filter: Joint estimation of states and parameters and the filter inbreeding problem. Water Resour. Res. 44(9). doi:10.1029/2007wr006505.

Frind, E., 1982. Simulation of long-term transient density-dependent transport in groundwater. Adv. Water Resour. 5, 73–88. doi:10.1016/0309-1708(82)90049-5.

Gaganis, P., Smith, L., 2006. Evaluation of the uncertainty of groundwater model predictions associated with conceptual errors: A per-datum approach to model calibration. Advances In Water Resources 29, 503–14. doi:10.1016/j.advwatres.2005.06.006.

Gómez-Hernández, J., Sahuquillo, A., Capilla, J., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data - i. theory. Journal of Hydro 203(1-4), 162–74. doi:10.1016/s0022-1694(97)00098-x.

Guymon, G.L., Yen, C.C., 1990. An efficient deterministic-probabilistic approach to modeling regional groundwater-flow .2. application to owens-valley, california. Water Resources Research 26, 1569–81. doi:10.1029/89WR03182.

Hao, Y., Yeh, T.C.J., Xian, J., Illman, W.A., Ando, K., Hsu, K.C., Lee, C.H., 2008. Hydraulic tomography for detecting fracture zone connectivity. Ground Water 46, 183–92. doi:10.1111/j.1745-6584.2007.00388.x.

Harrar, W.G., Sonnenborg, T.O., Henriksen, H.J., 2003. Capture zone, travel time, and solute-transport predictions using inverse modeling and different geological models. Hydrogeology Journal 11, 536–48. doi:10.1007/s10040-003-0276-2.

Hernández, A.F., Neuman, S.P., Guadagnini, A., Carrera, J., 2003. Conditioning mean steady state flow on hydraulic head and conductivity through geostatistical inversion. Stochastic Environmental Research and Risk Assessment 17, 329–38. doi:10.1007/s00477-003-0154-4.

Hernández, A.F., Neuman, S.P., Guadagnini, A., Carrera, J., 2006. Inverse stochastic moment analysis of steady state flow in randomly heterogeneous media. Water Resour. Res. 42. doi:10.1029/2005wr004449.

Huang, S., Wen, J., Yeh, T., Lu, W., Juan, H., Tseng, C., Lee, J., Chang, K., 2011. Robustness of joint interpretation of sequential pumping tests: Numerical and field experiments. Water Resources Research 47. doi:10.1029/2011WR010698. w10530.

Hubbard, S., Rubin, Y., 2000. Hydrogeological parameter estimation using geological data: A review of selected techniques,. J. Contam. Hydro!., 45, 3–34. doi:10.1016/s0169-7722(00)00117-0.

Hubbard, S., Rubin, Y., Majer, E., 1997. Ground penetrating radar assisted saturation and permeability estimation in bimodal systems,. Water Resour. Res. 33(5), 971–90. doi:10.1029/96wr03979.

Hyndman, D.W., Harris, J.M., Gorelick, S.M., 1994. Coupled seismic and tracer test inversion for aquifer property characterization. Water Resour. Res., 30(7), 1965–77. doi:10.1029/94wr00950.

Illman, W., Liu, X., Craig, A., 2007. Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms. Journal of Hydrology 341, 222–34. doi:10.1016/j.jhydrol.2007.05.011.

Illman, W.A., Liu, X., Takeuchi, S., Yeh, T.C.J., Ando, K., Saegusa, H., 2009. Hydraulic tomography in fractured granite: Mizunami underground research site, japan. Water Resources Research 45. doi:10.1029/2007WR006715. w01406.

Iribar, V., Carrera, J., Custodio, E., Medina, A., 1997. Inverse modelling of sea water intrusion in the Llobregat delta deep aquifer. J Hydrol 198, 226–44. doi:10.1016/s0022-1694(96)03290-8.

Janetti, E.B., Riva, M., Straface, S., Guadagnini, A., 2010. Stochastic characterization of the montalto uffugo research site (italy) by geostatistical inversion of moment equations of groundwater flow. Journal of Hydrology 381, 42–51. doi:10.1016/j.jhydrol.2009.11.023.

Jardani, A., Dupont, J.P., Revil, A., Massei, N., Fournier, M., Laignel, B., 2012. Geostatistical inverse modeling of the transmissivity field of a heterogeneous alluvial aquifer under tidal influence. Journal of Hydrology 472, 287–300. doi:10.1016/j.jhydrol.2012.09.031.

Khaninezhad, M.M., Jafarpour, B., Li, L., 2012a. Sparse geologic dictionaries for subsurface flow model calibration: Part i. inversion formulation. Advances in Water Resources 39, 106 –21. doi:http://dx.doi.org/10.1016/j.advwatres.2011.09.002.

Khaninezhad, M.M., Jafarpour, B., Li, L., 2012b. Sparse geologic dictionaries for subsurface flow model calibration: Part ii. robustness to uncertainty. Advances in Water Resources 39, 122 –36. doi:http://dx.doi.org/10.1016/j.advwatres.2011.10.005.

Khodabakhshi, M., Jafarpour, B., 2013. A bayesian mixture-modeling approach for flow-conditioned multiple-point statistical facies simulation from uncertain training images. Water Resources Research 49, 328–42. doi:10.1029/2011WR010787.

Kitanidis, P., Vomvoris, E., 1983. A geostatistical approach to the inverse problem in groundwater modelling (steady state) and one-dimensional simulations. Water Resour. Res. 19(3), 677–90. doi:10.1029/wr019i003p00677.

Kowalsky, M.B., Finsterle, S., Rubin, Y., 2004. Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. Advances in Water Resources 27(6), 583 –99. doi:10.1016/s0309-1708(04)00045-4.

Kunstmann, H., Kinzelbach, W., Siegfried, T., 2002. Conditional first-order second-moment method and its application to the quantification of uncertainty in groundwater modeling. Water Resour. Res. 38. doi:10.1029/2000wr000022.

Li, W., 2007a. A fixed-path markov chain algorithm for conditional simulation of discrete spatial variables. Math. Geol. 39(2), 159–76. doi:10.1007/s11004-006-9071-7.

Li, W., 2007b. Markov chain random fields for estimation of categorical variables. Math. Geol. 39(3), 321–35. doi:10.1007/s11004-007-9081-0.

Li, W., Nowak, W., Cirpka, O.A., 2005. Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown. Water Resour. Res. 41. doi:10.1029/2004WR003874.

Liu, S.Y., Yeh, T.C.J., Gardiner, R., 2002. Effectiveness of hydraulic tomography: Sandbox experiments. Water Resour. Res. 38. doi:10.1002/wrcr.20519.

Liu, X., Illman, W.A., Craig, A.J., Zhu, J., Yeh, T..C.J., 2007. Laboratory sandbox validation of transient hydraulic tomography. Water Resour. Res. 43. doi:10.1029/2006WR005144.

de Marsily, G., Delhomme, J.P., Delay, F., Buoro, A., 1999. 40 years of inverse problems in hydrogeology. Comptes Rendus De L Academie Des Sciences Serie Ii Fascicule A-sciences De La Terre Et Des Planetes 329, 73–87. doi:10.1016/S1251-8050(99)80208-0.

de Marsily, G., Lavedan, C., Bouchere, M., Fasanino, G., 1984. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. Geost. for Natural Resour. Charact. 182, 831–49. doi:10.1007/978-94-009-3701-7-16.

Martinez-Landa, L., Carrera, J., 2005. An analysis of hydraulic conductivity scale effects in granite (full-scale engineered barrier experiment (febex), grimsel, switzerland). Water Resour. Res. 41. doi:10.1029/2004WR003458.

McLaughlin, D., Townley, L.R., 1996. A reassessment of the groundwater inverse problem. Water Resour. Res. 32, 1131–61. doi:10.1029/96WR00160.

Medina, A., Carrera, J., 2003. Geostatistical inversion of coupled problems: dealing with computational burden and different types of data. J Hydrol 281, 251–64. doi:10.1016/s0022-1694(03)00190-2.

21

Meier, P., Medina, A., Carrera, J., 2001. Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. Ground Water 39, 10–7. doi:10.1111/j.1745-6584.2001.tb00346.x.

Meier, P.M., Carrera, J., Sanchez-Vila, X., 1999. A numerical study on the relationship between transmissivity and specific capacity in heterogeneous aquifers. Ground Water, 37, 611617. doi: 10.1111/j.1745–6584.1999.tb01149.x.

Moore, C., Doherty, J., 2006. The cost of uniqueness in groundwater model calibration. Adv. Water Res. 29(4), 605–23. doi:10.1016/j.advwatres.2005.07.003.

Murakami, H., Chen, X., Hahn, M.S., Liu, Y., Rockhold, M.L., Vermeul, V.R., Zachara, J.M., Rubin, Y., 2010. Bayesian approach for three-dimensional aquifer characterization at the hanford 300 area. Hydrology and Earth System Sciences 14, 1989–2001. doi:10.5194/hess-14-1989-2010.

Neuman, S.P., 2004. Stochastic groundwater models in practice. Stochastic Environmental Research and Risk Assessment 18, 268–70. doi:10.1007/s00477-004-0192-6.

Neuman, S.P., Preller, C., Narasimhan, T.N., 1982. Adaptive explicit-implicit quasi three-dimensional finite element model of flow and subsidence in multiaquifer systems. Water Resources Research 18, 1551–61. doi:10.1029/WR018i005p01551.

Nocchi, M., Salleolini, M., 2013. A 3d density-dependent model for assessment and optimization of water management policy in a coastal carbonate aquifer exploited for water supply and fish farming. Journal of Hydrology 492, 200–18. doi:http://dx.doi.org/10.1016/j.jhydrol.2013.03.048.

Park, Y.J., Sudicky, E.A., McLaren, R.G., Sykes, J.F., 2004. Analysis of hydraulic and tracer response tests within moderately fractured rock based on a transition probability geostatistical approach,. Water Resour. Res. 40. doi:10.1029/2004wr003188.

Poeter, E.P., Hill, M.C., 1998. Ucode, a computer code for universal inverse modeling. U.S. Geological Survey Water-Resources Investigations Report 98-4080, Denver, Colorado, USA. .

Ramarao, B.S., Lavenue, A.M., de Marsily, G., Marietta, M.G., 1995. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields .1. theory and computational experiments. Water Resources Research 31, 475–93. doi:10.1029/94WR02258.

Razack, M., Huntley, D., 1991. Assessing transmissivity from specific capacity in a large and heterogeneous alluvial aquifer. Ground 29(6), 856–61. doi:10.1111/j.1745-6584.1991.tb00572.x.

Renard, F., Jeannee, N., 2008. Estimating transmissivity fields and their influence on flow and transport: The case of champagne mounts. Water Resour. Res. 44. doi:10.1029/2008WR007033.

Renard, P., 2007. Stochastic hydrogeology: What professionals really need? Ground Water 45(5), 531–41. doi:10.1111/j.1745-6584.2007.00340.x.

Renard, P., Allard, D., 2013. Connectivity metrics for subsurface flow and transport. Adv. Water Res. 51, 168–96. doi:10.1016/j.advwatres.2011.12.001.

Ritzi, R., 2000. Behavior of indicator variograms and transition probabilities in relation to the variance in lengths of hydrofacies. Water Resour. Res. 36(11), 3375–81. doi:10.1029/2000wr900139.

Ritzi, R., Dai, Z., Dominic, D., Rubin, Y.N., 2004. Spatial correlation of permeability in cross-stratified sediment with hierarchical architecture. Water Resour. Res. 40. doi:10.1029/2003wr002420.

Riva, M., Guadagnini, A., De Gaspari, F., Alcolea, A., 2010. Exact sensitivity matrix and influence of the number of pilot points in the geostatistical inversion of moment equations of groundwater flow. Water Resour Res 46(11). doi:10.1029/2009wr008476.

Riva, M., Guadagnini, L., Guadagnini, A., Ptak, T., Martac, E., 2006. Probabilistic study of well capture zones distribution at the lauswiesen field site. J. Contam. Hydrol. 88, 92–118. doi:10.1016/j.jconhyd.2006.06.005.

Riva, M., Panzeri, M., Guadagnini, A., Neuman, S.P., 2011. Role of model selection criteria in geostatistical inverse estimation of statistical data- and model-parameters. Water Resour Res 47. doi:10.1029/2011WR010480.

22

Roggero, F., Hu, L., 1998. Gradual deformation of continuous geostatistical models for history matching. In annual technical conference SPE 49004.

Rötting, T.S., Carrera, J., Bolzicco, J., Salvany, J.M., 2006. Stream-stage response tests and their joint interpretation with pumping tests. Ground Water 44, 371385. doi: 10.1111/j.1745–6584.2005.00138.x.

Rubin, Y., Chen, X., Murakami, H., Hahn, M., 2010. A bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields. Water Resour. Res. 46. doi:10.1029/2009WR008799.

Rubin, Y., Dagan, G., 1987a. Stochastic identification of transmissivity and effective recharge in steady groundwater-flow .1. theory. Water Resour. Res. 23, 1185–92. doi:10.1029/WR023i007p01185.

Rubin, Y., Dagan, G., 1987b. Stochastic identification of transmissivity and effective recharge in steady groundwater-flow .2. case-study. Water Resour. Res. 23, 1193–200. doi:10.1029/WR023i007p01193.

Rubin, Y., Dagan, G., 1988. Stochastic-analysis of boundaries effects on head spatial variability in heterogeneous aquifers .1. constant head boundary. Water Resour. Res. 24, 1689–97. doi:10.1029/WR024i010p01689.

Rubin, Y., Ferreira, J.P.L., Rodrigues, J.D., Dagan, G., 1990. Estimation of the hydraulic parameters of the Rio-maior Aquifer in Portugal by using stochastic inverse modeling. Journal of Hydrology 118, 257–79. doi:10.1016/0022-1694(90)90262-V.

Rubin, Y., Mavko, G., Harris, J., 1992. Mapping permeability in heterogeneous aquifers using hydrologic and seismic data. Water Resour. Res. 28(7), 1809–16. doi:10.1029/92wr00154.

Sakaki, T., Frippiat, C.C., Komatsu, M., Illangasekare, T.H., 2009. On the value of lithofacies data for improving groundwater flow model accuracy in a three-dimensional laboratory-scale synthetic aquifer. Water Resour. Res. 45. doi:10.1029/2008wr007229.

Sánchez-Vila, X., Carrera, J., Girardi, J., 1996. Scale effects in transmissivity. Journal of Hydrology 183, 1 – 22. doi:10.1016/S0022-1694(96)80031-X.

Senger, R.K., Fogg, G.E., 1987. Regional underpressuring in deep brine aquifers, palo-duro basin, texas .1. effects of hydrostratigraphy and topography. Water Resources Research 23, 1481–93. doi:10.1029/WR023i008p01481.

Shavit, U., Furman, A., 2001. The location of deep salinity sources in the israeli coastal aquifer. Journal of Hydrology 250, 63–77. doi:10.1016/S0022-1694(01)00406-1.

Sun, N.Z., Yang, S.L., Yeh, W.W.G., 1998. A proposed stepwise regression method for model structure identification. Water Resources Research 34, 2561–72. doi:10.1029/98WR01860.

Vassolo, S., Kinzelbach, W., Schafer, W., 1998. Determination of a well head protection zone by stochastic inverse modelling. Journal of Hydrology 206, 268–80. doi:10.1016/S0022-1694(98)00102-4.

Vázquez-Suñé, E., Abarca, E., Carrera, J., Capino, B., Gamez, D., Pool, M., Simo, T., Badle, F., Ninerola, J.M., Ibanez, X., 2006. Groundwater modelling as a tool for the european water framework directive (wfd) application: The Llobregat case. Phys. Chem. Earth 31, 1015–29.

Vázquez-Suñé, E., Carrera, J., Tubau, I., Sánchez-Vila, X., Soler, A., 2010. An approach to identify urban groundwater recharge. Hydrology and Earth System Sciences 14, 2085–97. doi:10.5194/hess-14-2085-2010.

Vesselinov, V.V., Neuman, S.P., 2001. Numerical inverse interpretation of single-hole pneumatic tests in unsaturated fractured tuff. Ground Water 39, 685–95. doi:10.1111/j.1745-6584.2001.tb02358.x.

Vesselinov, V.V., Neuman, S.P., Illman, W.A., 2001a. Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 1. methodology and borehole effects. Water Resources Research 37, 3001–17. URL: http://dx.doi.org/10.1029/2000WR000133, doi:10.1029/2000WR000133.

Vesselinov, V.V., Neuman, S.P., Illman, W.A., 2001b. Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 2. equivalent parameters, high-resolution stochastic imaging and scale effects. Water Resources Research 37, 3019–41.

doi:10.1029/2000WR000135.

Voss, C.I., Souza, W.R., 1987. Variable density flow and transport simulation of regional aquifers containing a narrow freshwater-saltwater transition zone. Water Resour. Res. 26, 2097–106. doi:10.1029/wr023i010p01851.

Walvoord, M.A., Pegram, P., Phillips, F.M., Person, M., Kieft, T.L., Frederickson, J.K., McKinley, J.P., Swenson, J.B., 1999. Groundwater flow and geochemistry in the southeastern san juan basin: Implications for microbial transport and activity. Water Resources Research 35, 1409–24. doi:10.1029/1999WR900017.

Winter, C., Tartakovsky, D., Guadagnini, A., 2003. Moment differential equations for flow in highly heterogeneous porous media. Surveys in Geophysics 24, 81–106. doi:10.1023/A:1022277418570.

Winter, C.L., Tartakovsky, D.M., 2002. Groundwater flow in heterogeneous composite aquifers. Water Resources Research 38, 1–11. doi:10.1029/2001WR000450.

Ye, M., Khaleel, R., 2008. A markov chain model for characterizing medium heterogeneity and sediment layering structure. Water Resour. Res., 44. doi:10.1029/2008wr006924.

Yeh, T.C.J., Liu, S.Y., 2000. Hydraulic tomography: Development of a new aquifer test method. Water Resour. Res., 36, 2095–105. doi:10.1029/2000WR900114.

Yeh, W.W.G., 1986. Review of parameter-identification procedures in groundwater hydrology - the inverse problem. Water Resour. Res., 22, 95–108. doi:10.1029/WR022i002p00095.

Zhang, C., Li, W., 2008. A comparative study of nonlinear markov chain models for conditional simulation of multinomial classes from regular samples,. Stochastic Environ. Res. Risk Assess., 22(2), 217 230, doi:10.1007/s00477–007–0109–2.

Zhang, H., Harter, T., Sivakumar, B., 2006. Nonpoint source solute transport normal to aquifer bedding in heterogeneous, markov chain random fields,. Water Resour. Res. 42. doi:10.1029/2004wr003808.

Zhou, H., Gmez-Hernndez, J.J., Li, L., 2014. Inverse methods in hydrogeology: Evolution and recent trends. Advances in Water Resources 63, 22 – 37. doi:http://dx.doi.org/10.1016/j.advwatres.2013.10.014.

Zhou, H., Gómez-Hernández, J., Li, L., 2012. A pattern-search-based inverse method. Water Resour. Res. 48. doi:10.1029/2011wr011195.

Zhu, J.F., Yeh, T.C.J., 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. Water Resour. Res. 41. doi:10.1029/2004WR003790.

Zhu, J.F., Yeh, T.C.J., 2006. Analysis of hydraulic tomography using temporal moments of drawdown recovery data. Water Resour. Res. 42. doi:10.1029/2005WR004309.

Zimmerman, D.A., de Marsily, G., Gotway, C.A., Marietta, M.G., Axness, C.L., Beauheim, R.L., Bras, R.L., Carrera, J., Dagan, G., Davies, P.B., Gallegos, D.P., Galli, A., Gomez-Hernandez, J., Grindrod, P., Gutjahr, A.L., Kitanidis, P.K., Lavenue, A.M., McLaughlin, D., Neuman, S.P., RamaRao, B.S., Ravenne, C., Rubin, Y., 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. Water Resour. Res., 34, 1373–413. doi:10.1029/98WR00003.

Table 1: Averaged water mass balance in Mar del Plata aquifer from 1913 to 2005, hm^3 per year

	Main inflow/outflow	1913-1950	1950-1970	1970-2005
Input	Recharge from rainfall	55.6	74.9	90.6
	Recharge in urban area	5.2	8.7	29.8
	Recharge from irrigation	6.86	10.85	17.68
	Lateral inflow	3.7	5.8	7.17
	Pumping	-8.2	-50.6	-100.4
Output	Los Cueros river	-5.51	-11.2	-10.78
	Seco river	-9.29	-20.37	-18.95
	Santa Elena river	-2.27	-5.46	1.19
	La Tapera river	-2.29	-3.63	2.46
	Los Padres lake	3.79	-1.61	-0.26
	Seaside shallow aquifer urban area	1.62	1.9	1.71
	Seaside deep aquifer urban area	-2.51	1.67	-0.57
	Seaside deep aquifer	-22.97	-25.78	-16.37
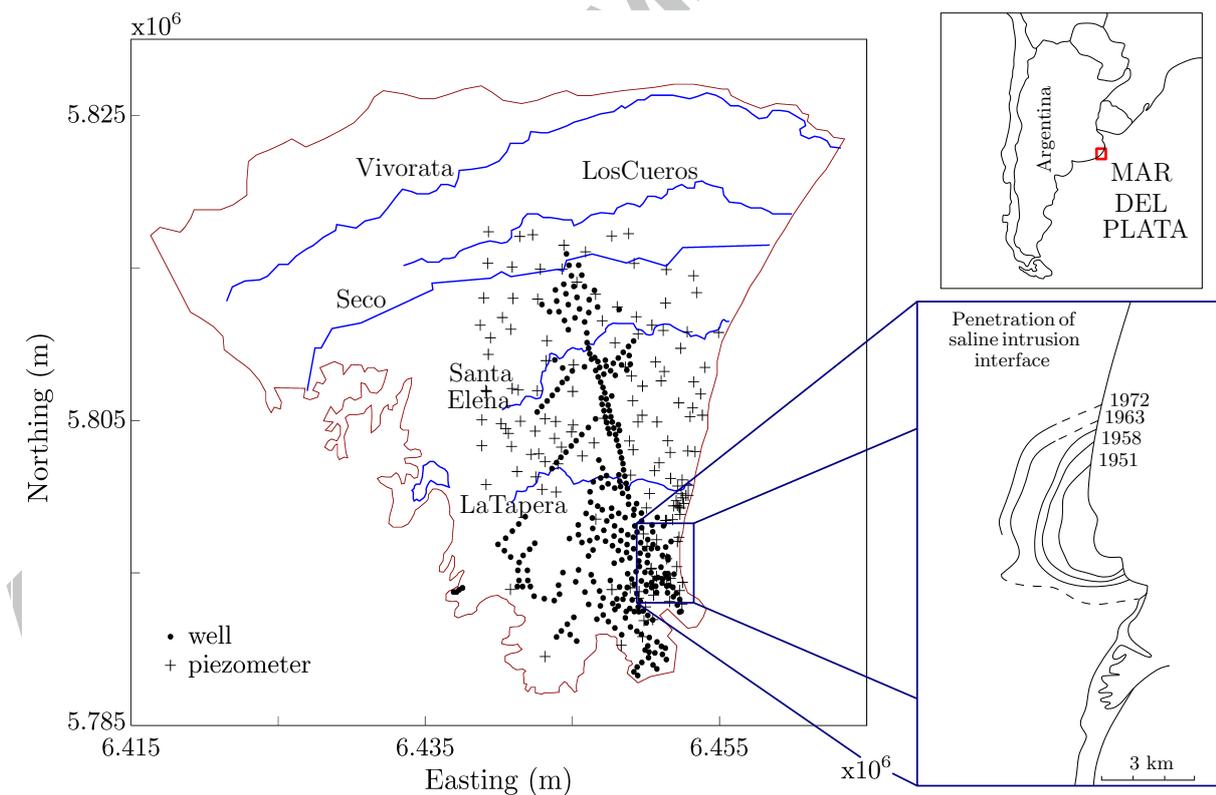
Figure 1: Upper right: Location of the study area. Left: model domain, main rivers (blue lines), and observation points. Dots represent pumping wells and crosses depict piezometers. Lower right: Penetration of the seawater intrusion front between 1951 and 1972. In all maps, the origin is set to (6415 km, 5785 km) in Lambert II extended coordinates, which corresponds to 3805'S latitude and 5800'W longitude.

25
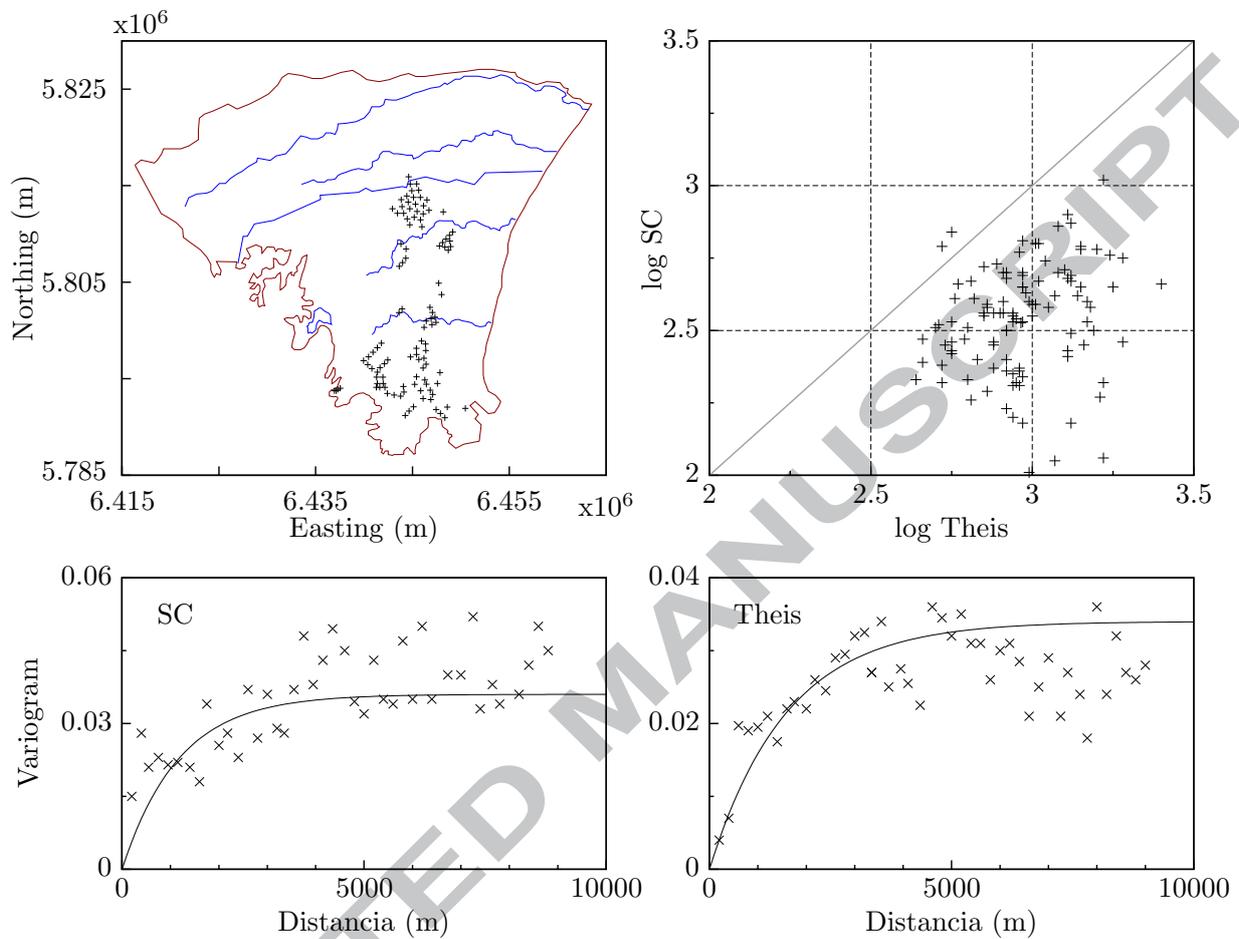
Figure 2: Location of wells with transmissivity data arising from the interpretation of pumping test (Theis) and specific capacity (SC), Log-log plot of Theis versus SC data and calculated variograms. Note that transmissivities from Theis interpretation are consistently larger than those derived from specific capacity.
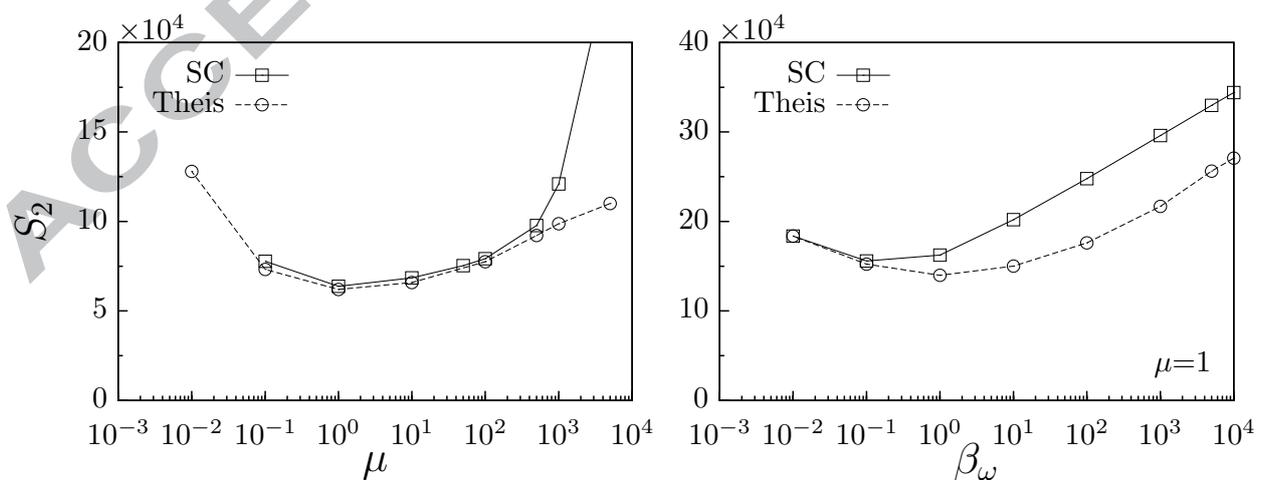
Figure 3: Discrimination criteria S_2 (eq.2), with respect to (left) the plausibility weight (μ) when using only head data and (right) the concentration data weight (β_ω) when using concentration and head data for $\mu=1$.

26
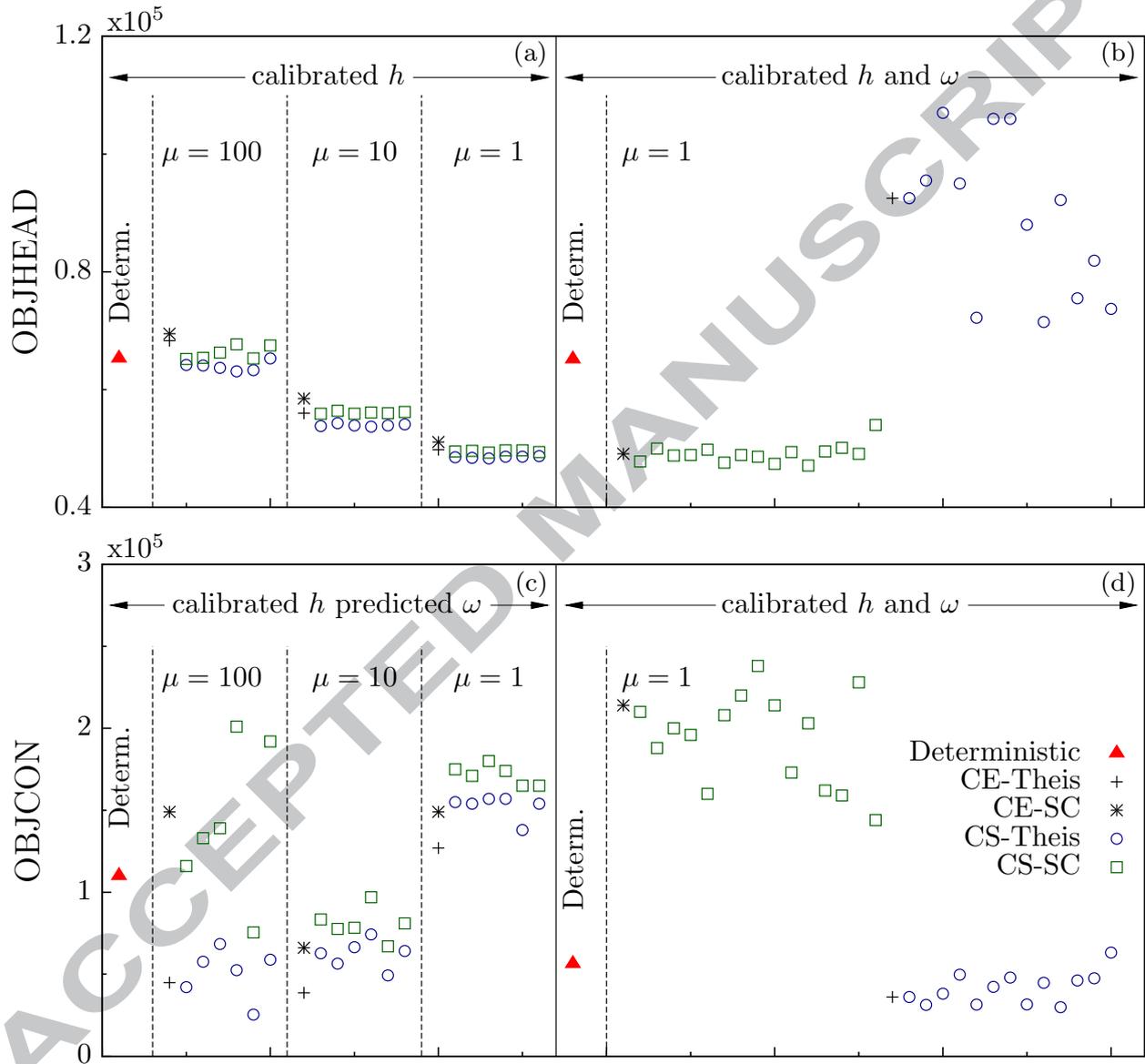
Figure 4: Objective function measuring the fit between calculated and measured heads (OBJHEAD, above) and concentrations (OBJCON, below) for the calibration and prediction runs discussed in this paper. Note that models performing best during calibration (μ=1 was the optimum) did not lead to best predictions of concentrations. The deterministic model (i.e., zonation) predictions ranked around the average of the stochastic models. Among these, conditional estimation (CE) predictions were always slightly better than those of conditional simulations (CS), even though their calibration errors are slightly worse.

deterministic   CE − SC   CE − Theis
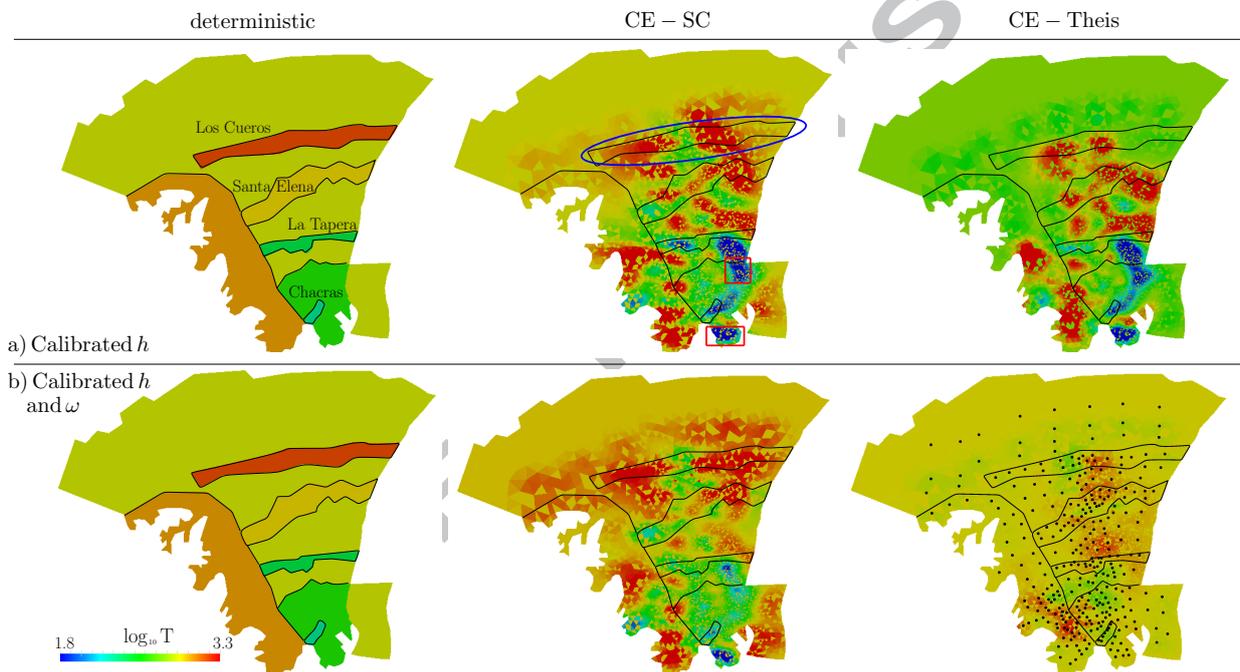
Figure 5: Optimum transmissivity fields (m²/d) obtained, from left to right, from the deterministic (geologically-based) model and conditional estimation of transmissivity fields using SC T data and Theis T data with the optimum plausibility weights ($\mu=\beta_\omega=1$), when (a, upper row) the T fields are conditioned to T and heads, or (b, lower row) the T fields are conditioned to T, head and concentration data (lower row). In the lower right panel, dots represent the pilot points distribution. Note that the conditional estimations honor the paleochannels (with a slightly modified geometry and continuity, see blue ellipse for the Los cueros stream), and quatzitic outcrops (red squares).
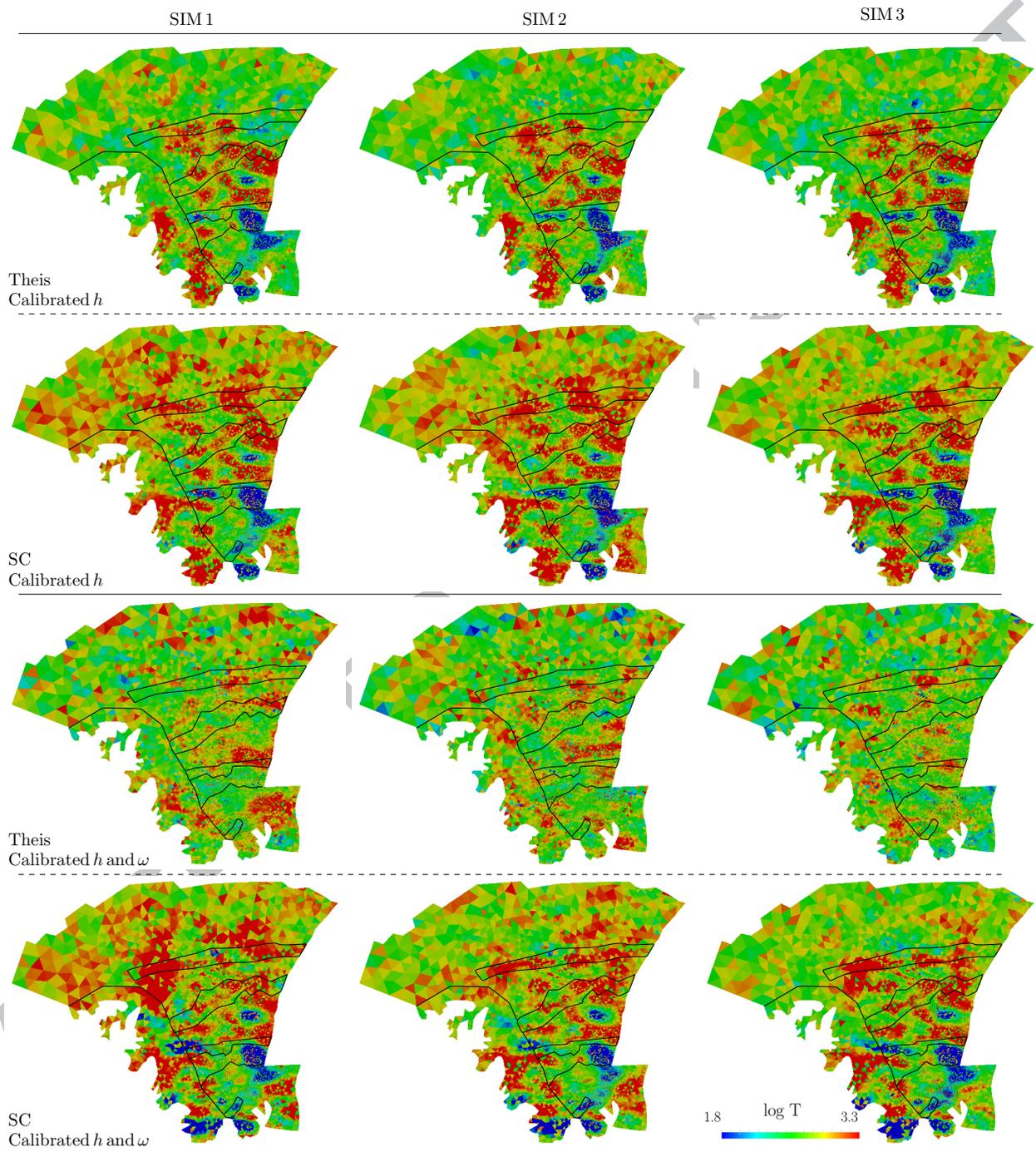
Figure 6: Transmissivity fields (m^2/d) conditioned to either h above or h and ω (below) data and point values of T derived from Theis test interpretation or from specific capacity data (SC) with the optimum plausibility weights ($\mu=\beta_\omega=1$). Note that, except for the simulations considering Theis T data, the stochastic T fields reproduced paleochannel structures and quartzitic outcrops, as in the case of conditional estimations.
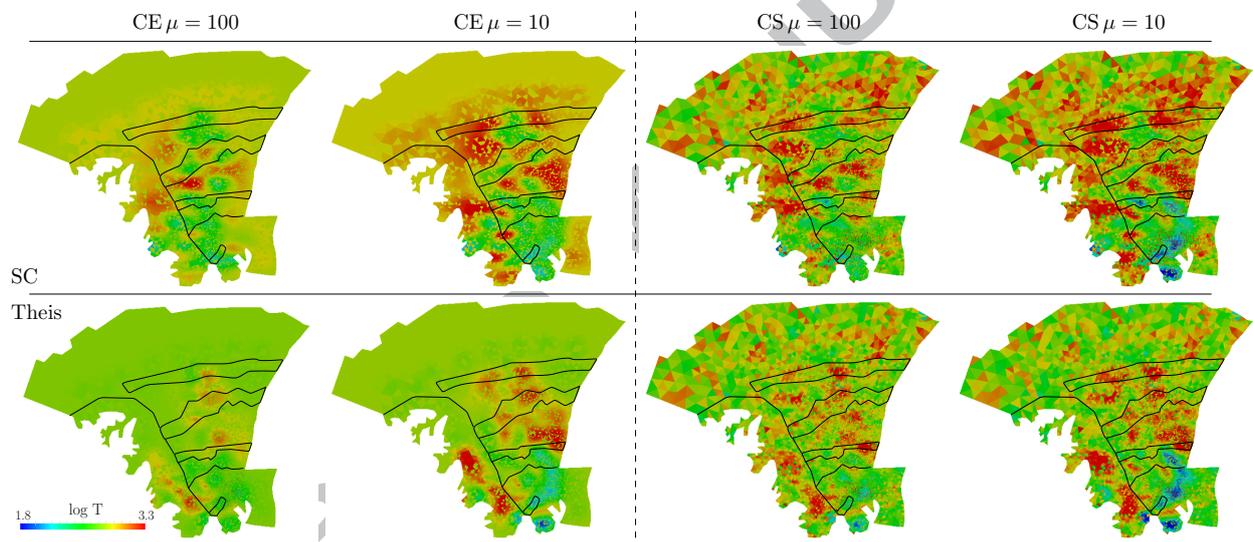
Figure 7: Conditional estimation (CE, right) and one conditional simulation (CS, left) of transmissivity fields (m^2/d) conditioned to heads and SC (above) or Theis (below) data for increased plausibility weights μ. Note that the paleochannels are still detected, albeit less sharply than in Figures 5 6, especially in the conditional simulation runs.

30
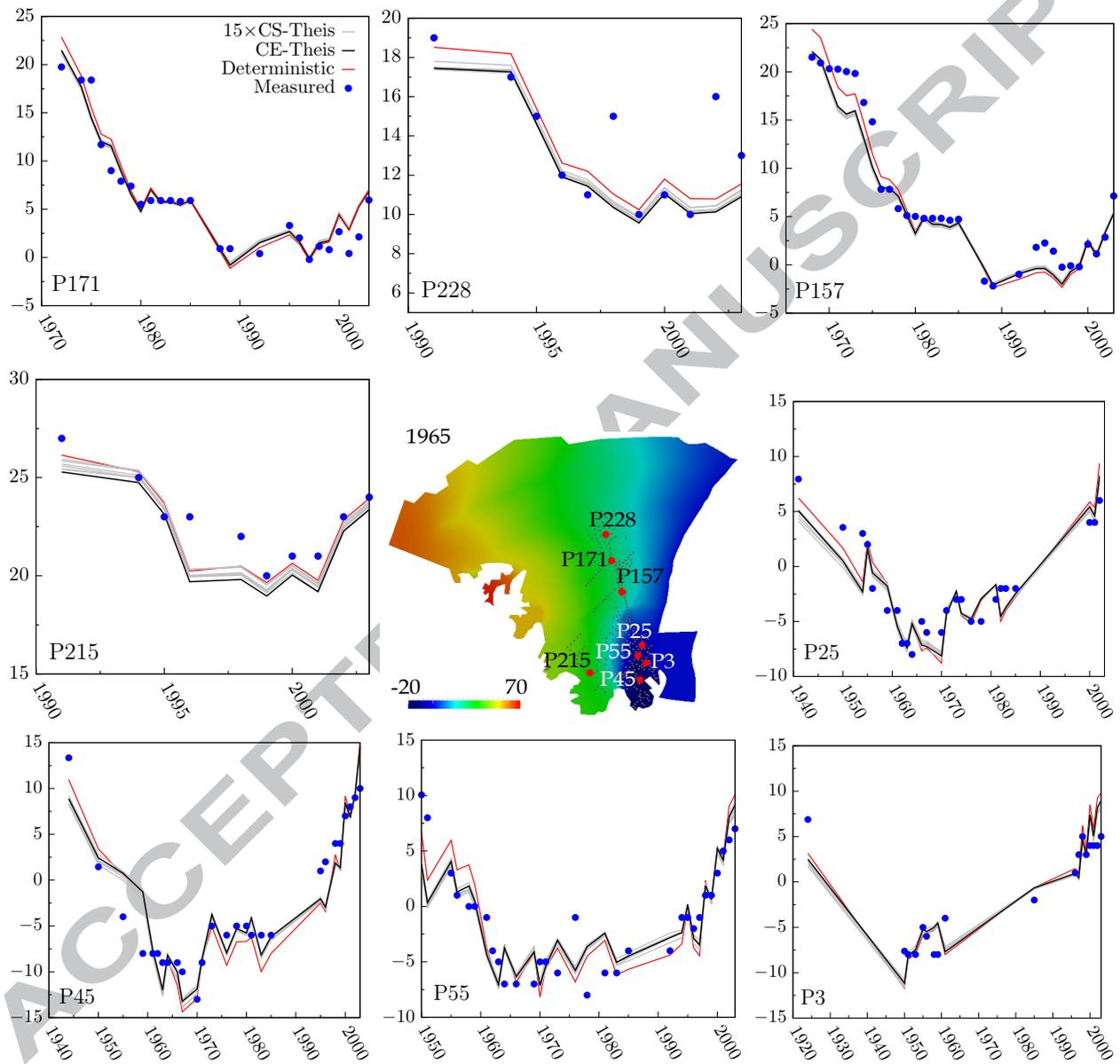
Figure 8: Piezometric map (m) calculated for December 1965 together with the calculated and measured heads at eight selected observation points when calibrating to T and head data only (μ=10 and β_ω=1). Red lines represent calculated heads from the deterministic model, black lines the single best solution by conditional estimation (EC), grey lines the conditional simulations and blue dots depict the measured head data used for calibration.

31
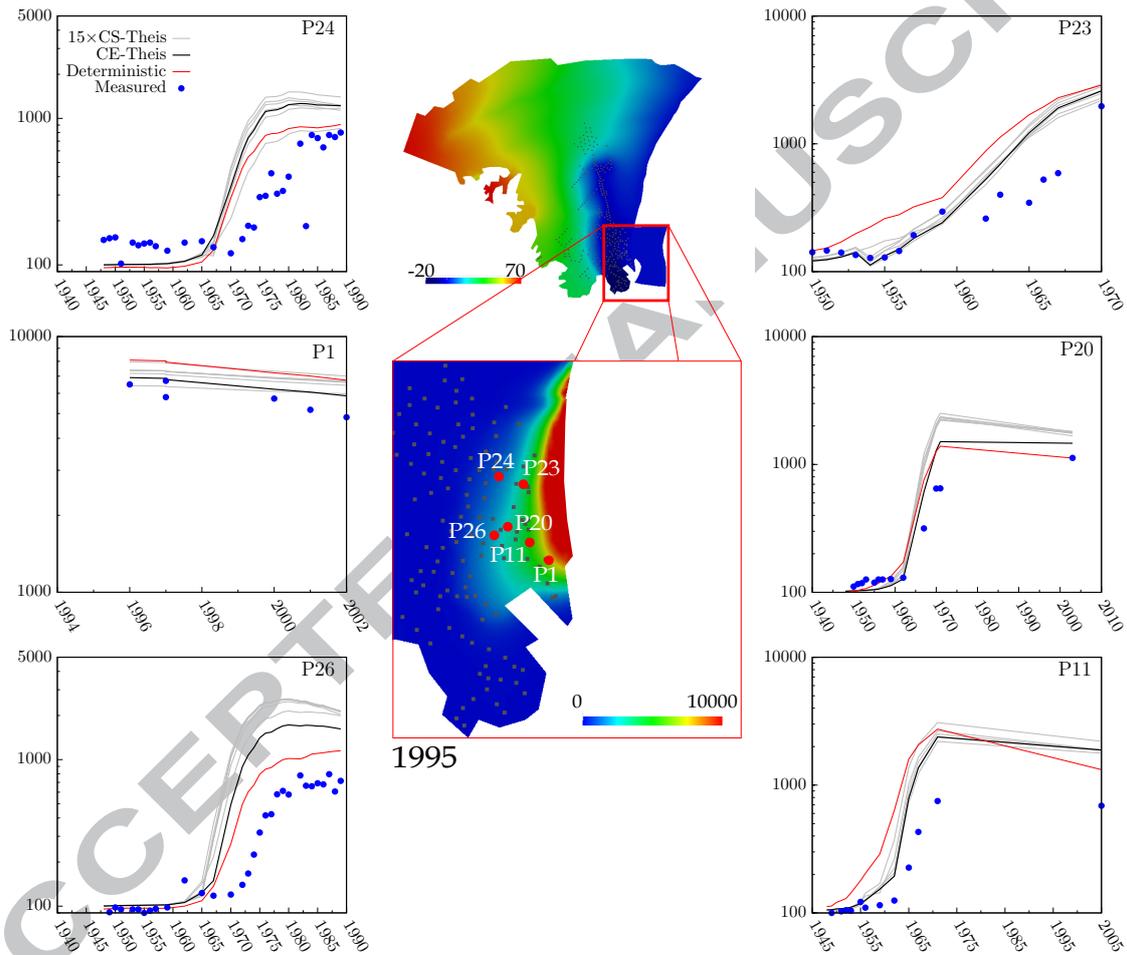
Figure 9: Blind predictions of concentrations. Chloride concentration (mg/l) and piezometric maps (m) for December 2001 obtained from the deterministic model. Calculated and measured concentrations at eight selected observation points (μ=10 and β_ω=1).

32
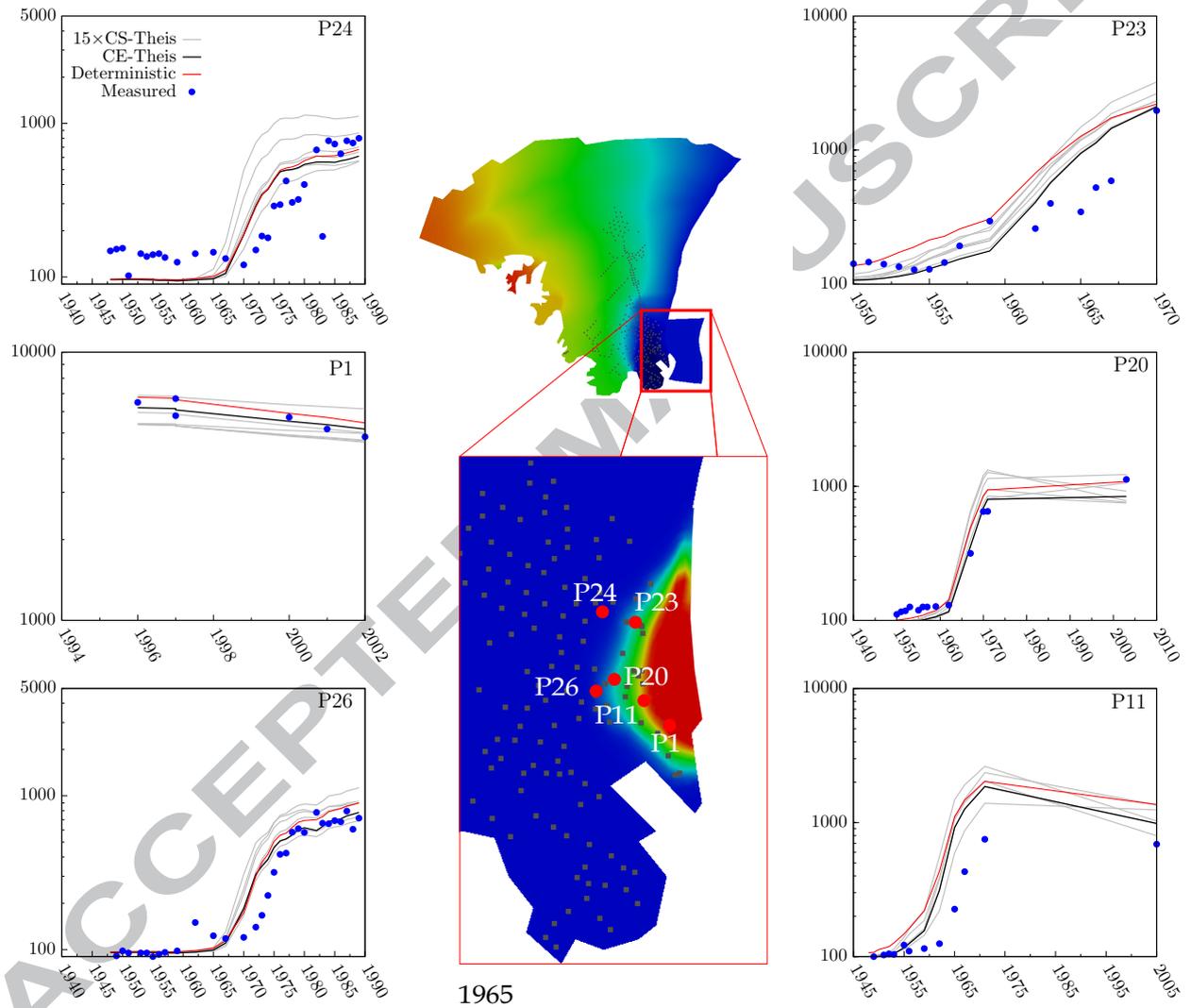
Figure 10: calibration against both h and ω data. Chloride concentration (mg/l) and piezometric maps (m) for December 2001 obtained from the deterministic model. Calculated and measured concentrations at eight selected observation points ($\mu=\beta_\omega=1$).

33

- Modeling of flow and transport in real regional aquifers.

- Relative merits of deterministic and stochastic approaches: regional-scale calibration.

- Impact of using T data from large and small-scale tests on the calibration process.