Hydrogeophysical model calibration and uncertainty analysis via full integration of PEST/PEST++ and COMSOL

Andrés González-Quirós[1*], Jean-Christophe Comte[1]

[1] School of Geosciences, University of Aberdeen, Aberdeen, UK

Highlights

- *We present a strategy for hydrogeophysical inversion and uncertainty analysis.*
- *PEST/PEST++ and COMSOL Multiphysics are fully integrated.*
- *The approach is applied to electrical resistivity monitoring in a coastal aquifer.*
- *The coupled inversion achieves better delineation of seawater intrusion.*
- *The iterative Ensemble Smoother is used for uncertainty analysis.*

Abstract

Calibration of groundwater models is frequently limited by a lack of direct hydrogeological data. Non-intrusive geophysical methods are increasingly used to provide higher spatio-temporal resolution datasets for identification of hydrological processes and estimation of hydraulic properties. In groundwater model calibration performed through joint or coupled hydrogeophysical inversion, the hydrogeological datasets are supplemented with auxiliary geophysical data. In this work we propose a methodological approach to perform coupled inversion by integrating the calibration software PEST/PEST++ with COMSOL Multiphysics using MATLAB. The strategy provides multiple options for calibration and uncertainty analysis relevant for a broad range of environmental models. To illustrate the approach, we show a hydrogeophysical application in which electrical resistivity is used jointly with borehole data for the identification of seawater intrusion in a coastal aquifer.

Keywords: *Hydrogeophysics, Multiphysics modelling, Coupled inversion, Uncertainty analysis.*

1 Introduction

Groundwater management and decision making is frequently supported by outcomes from numerical models (Doherty and Simmons, 2013; Ferré, 2017). To have predictive capabilities numerical

groundwater models are traditionally calibrated (e.g., Anderson et al., 2015) and further evaluated to quantify prediction uncertainty (Linde et al., 2017). The uncertainty analysis can be conducted either using a deterministic (error propagation analysis) or stochastic approach, even avoiding the necessity of a pre-calibrated model (e.g., Scheidt et al., 2018; Hermans et al., 2018, 2019). Model calibration and uncertainty analysis in groundwater modelling, as for other subsurface systems, is frequently hampered by the scarcity of direct hydrogeological data, such as head or concentration measurements from wells or boreholes. As drilling is expensive, intrusive and, generally, provides spatially scattered point information, key groundwater processes and spatial distribution of hydrogeological properties are not fully captured, more so in heterogeneous aquifer systems (e.g., Zheng and Gorelick, 2003).

Geophysical methods have been, and are, routinely used in hydrogeological studies to provide indirect information of the subsurface (Kirsch, 2006). Traditionally and most often, geophysical methods are processed and interpreted independently and used qualitatively in the initial stages of the groundwater modelling workflow, typically for geometrical delineation of geological bodies and structures or the mapping of hydrological features (e.g. Mastrocicco et al., 2010) or processes of interest, such as the location of the water table (Buchanan, S., & Triantafilis, J., 2009), the identification of saltwater intrusion (De Franco et al., 2009), soil moisture variations (Chambers et al., 2014) or the tracking of contaminant plumes (Gasperikova et al., 2012) .

In the last two decades, there has been a growing interest in a more quantitative use of geophysical information in hydrogeology including their integration in groundwater models, eventually leading to the emergence of a new scientific sub-discipline known as hydrogeophysics (Binley et al., 2015). At the core of the hydrogeophysical approach lies the use or definition of petrophysical relationships (e.g., Archie, 1942 and its variants; Slater, 2007), commonly derived at the laboratory. These are required to link geophysical properties (e.g. electrical resistivities) to hydrogeological parameters (e.g. storage properties; Mezquita-Gonzalez et al., 2021) or state variables (e.g. water salinity content; Klotzsche et al., 2018).

Previous authors have shown limitations of the capability of geophysical inversion to accurately resolve subsurface features (Day-Lewis et al., 2005; Singha and Moysey, 2006), typically providing a

53 blurry, smoother, and sometimes distorted representation of reality, as well as significant uncertainties
54 in the values of the parameter measured. Although alternative regularization approaches may improve
55 the inverted model (Hermans et al., 2012; Bouchedda et al., 2017; Thibaut et al., 2021), recent studies
56 (e.g. Revil et al., 2017; Brunetti and Linde, 2018; González-Quirós and Comte, 2020) have also shown
57 the importance of conceptual and structural errors in the hydrogeophysical workflow, such as the
58 assumption of homogeneity in heterogeneous systems when parameterizing the petrophysical model or
59 the incorrect selection of the petrophysical model itself. As a result, the quantitative use of information
60 derived from geophysical inversion in groundwater model parametrization (i.e. direct mapping) or
61 calibration could lead, if incorrectly applied, to unrealistic or erroneous property distributions or
62 estimations, and consequently, inaccurate hydrological interpretations. Awareness of these limitations
63 and incorporation of their associated uncertainty in the workflow (e.g. Beaujean et al., 2014; Hermans
64 and Irving, 2017) have been shown to reduce some of these errors associated with the use quantitative
65 geophysical information in hydrogeophysical applications.

66 To overcome these limitations, it was proposed (e.g. Lebbe, 1999) that a more appropriate strategy
67 for an efficient integration of geophysics in groundwater modelling calibration is through a coupled
68 hydrogeophysical modelling. In it, both the geophysical and hydrogeological observations, rather than
69 inversion results, are simultaneously used as model verification datasets (Comte and Banton, 2007), or
70 in an automated way through coupled hydrogeophysical inversion in which geophysical and
71 hydrogeological measurements are used as input datasets for automatic groundwater model calibration
72 (Hinnell et al., 2010; Linde and Doestch, 2016). While the approach is still subject to practical and
73 structural errors, their main advantage is that the significantly extended multiphysical observation
74 datasets provide additional constraints in the estimation of the ranges and distributions of hydrological
75 parameters and variables (or quantities), and in the resolution of groundwater processes, which are the
76 main objectives in hydrogeological studies (Linde et al., 2015). Coupled modelling approaches start
77 with the simulation of the hydrological model for some defined distribution of hydrological properties
78 and boundary conditions. A petrophysical relationship, which can be uncertain (e.g. Irving and Singha,
79 2010), is used to obtain a spatial distribution of geophysical properties from which the forward

geophysical model is simulated to obtain the geophysical response at locations of interest. Model predicted geophysical and hydrologic observations are compared with field measurements through an iterative procedure where the hydrological-geophysical model properties are sequentially adjusted until an acceptable fit is obtained. In the coupled inversion approaches the iterative comparison and minimisation of differences between modelled and observed values is automated.

A key difficulty for development and application of coupled hydrogeophysical inversion for real world applications is the requirement to code or program the necessary governing equations of both, the groundwater and -at least one- geophysical problems linked with the appropriate petrophysical relationship. Previous authors have used specific software –e.g. MODFLOW (Harbaugh et al., 2005), SEAWAT (Langevin et al., 2008), SUTRA (Voss and Provost, 2002) or FEFLOW (Diersch, 2013) for the groundwater problem, and RES2DMOD/RES2DINV (Loke, 2018), BERT (Rücker et al., 2006) or R2 (Binley and Kemna, 2005) for the geophysical problem– to solve the hydrogeophysical model sequentially (e.g. Herckenrath et al., 2013; Kang et al., 2019) or have developed solutions for hydrogeophysical inversion that usually require the development or programming of in-house codes for specific applications of (Pollock and Cirpka, 2012; Steklova and Haber, 2017). On this sense, open-source libraries under continuous development, such as SimPEG (Cockett et al., 2015) or PyGIMLI (Rücker et al., 2017), provide a wide range of solutions to perform coupled hydrogeophysical inversion or be integrated with a groundwater model.

However, although coupled modelling is usually achievable for models with simple geometries, representing scenarios that can be straightforwardly conceptualized, and for small multiphysical calibration datasets, implementation in real-world aquifers targeting complex hydrological processes and/or large datasets (e.g. Comte et al., 2017) can be time consuming and would benefit from use of automatic calibration and a much higher degree of flexibility for geometry, meshing, parametrization and implementation of the governing equations and boundary conditions. Additionally, multiphysical model calibration and uncertainty analysis demands a robust framework —aside from the mathematical engine itself— to handle the multiple types of observation datasets and program the specific

106 relationships between parameter types with geologically realistic distributions of properties (Linde et
107 al., 2015).

108 In this work, we propose a methodology to implement a fully coupled multiphysical inversion for
109 the calibration and uncertainty analysis in hydrogeophysical environmental models. The strategy is
110 achieved by integrating a fully coupled hydrogeophysical forward model developed with the
111 commercial finite-element software COMSOL Multiphysics® and the model-independent calibration
112 software PEST (Doherty, 2020) and PEST++ (White et al., 2020), widely used by the groundwater
113 community. The coupling of PEST and COMSOL was performed previously by Halloran et al. (2019),
114 who developed a Java interface for communication. Here we follow a different strategy for
115 communication using MATLAB instead of Java, and we include some novelties that were not
116 implemented in COMPEST, such as the use of pilot points for spatial parameterization, Tikhonov
117 regularization, application of the novel iterative Ensemble Smoother from PEST++, parallelization and
118 a first application of the strategy to a hydrogeophysical problem.

119 The paper is organised as follows. In section 2.1 we present some basic concepts and theory behind
120 PEST and PEST++. In section 2.2 we explain the methodological workflow for integration of
121 PEST/PEST++ and COMSOL with model parallelization. In section 2.3 we describe the
122 hydrogeophysical modelling and inversion and present the example model. In section 3 we show the
123 hydrogeophysical example application results: in section 3.1 the results of a Monte Carlo forward
124 coupled hydrogeophysical model, in section 3.2 the results of model calibration and in section 3.3 the
125 results of the uncertainty analysis obtained with application of the iterative ensemble smoother. Finally,
126 in section 4, we discuss the capabilities of the multiphysical inversion for calibration of environmental
127 models.

2 Theory, Software and Numerical Methods

2.1 Model Calibration with PEST/PEST++

2.1.1 PEST and PEST++

The model-independent calibration software PEST (Doherty, 2020) has been widely used since the 90s by the groundwater community for model calibration. PEST relies on the use of input and output files to interact with any numerical model. It includes a suite of functionalities for model parameterization, data conversion or setup of the calibration workflow, including parallelization. PEST++ (White et al., 2020) development began in 2009 within the USGS and contained much (but not all) of the capabilities of PEST. During the last decade more functionalities such as global sensitivity or uncertainty analysis, have been incorporated. Both codes and accompanying utilities can be freely download in their dedicated webpages.

Below we explain some basic principles of the parameter estimation in environmental sciences for which PEST and PEST++ provide a wide range of solutions. For more details we refer to their respective and extensive support documentation for details (Doherty, 2020; White et al., 2020).

2.1.2 Highly Parameterized Parameter Estimation

The parameter estimation or inverse problem in environmental numerical modelling aims to determine a reduced and finite set of parameters, **p** (for example permeabilities defined at the discretized model grid), that agree within a mathematical norm with the set of historical observations, **h** (for example hydraulic heads collected in the field)

$$\mathbf{h} = \mathbf{X}\mathbf{p} + \boldsymbol{\varepsilon}, \tag{1}$$

X is a matrix that represents the action of the model in the parameters and **ε** is the noise associated with the measurements. For a non-linear problem, and omitting the noise term, the previous expression is often found as

$$\mathbf{h} = \mathbf{X}(\mathbf{p}). \quad (2)$$

2.1.3 GLM inversion

From the equation above we can define a vector of residuals, **r**, as the differences between the field measurements and the outputs of the model

$$\mathbf{r} = \mathbf{h} - \mathbf{Xp}, \quad (3)$$

and an objective function as the sum of weighted squared residuals such as

$$\Phi_d = (\mathbf{h} - \mathbf{Xp})^{\boldsymbol{t}}\mathbf{Q}(\mathbf{h} - \mathbf{Xp}), \quad (4)$$

where **Q** is the weight matrix

$$\mathbf{Q} = \sigma_r^2 C^{-1}(\varepsilon), \quad (5)$$

σ_r^2 is a proportionality constant known as reference variance or variance of unit weight (Doherty, 2015) and $\boldsymbol{C}(\boldsymbol{\varepsilon})$ the covariance matrix that characterises the measurement noise.

Our objective is to find values for **p** improving the fit achieved between model outputs and observations, that is, to iteratively lower the objective function, Φ_d. PEST minimizes the objective function by using the Gauss-Levenberg-Marquardt (GLM) algorithm (Doherty, 2015). For each iteration PEST searches an improvement in fit by modifying the model parameters using the formula

$$\mathbf{p} - \mathbf{p_0} = (\mathbf{J^tQJ} + \boldsymbol{\lambda}\mathbf{I})^{-1} + \mathbf{J^tQr}, \quad (6)$$

where $\mathbf{p_0}$ are the parameter values at the start of the iteration, $\boldsymbol{\lambda}$ is the Marquardt lambda (Levenberg, 1944; Marquardt, 1963) and **J** the Jacobian matrix**,** which is filled using finite-difference approximation of the partial first derivative of the simulated observations, **s** (model simulated equivalent to the field observations, **h**), with respect to the parameters, **p**,

$$\mathbf{J}[s_i, p_j] = \frac{\partial s_i}{\partial p_j} \approx \frac{\Delta s_i}{\Delta p_j}. \quad (7)$$

The equation above means that to fill the Jacobian matrix it is necessary to run the model at least as many times as model estimable parameters, which is an important computational limitation for highly parameterized models and longer computational runs.

2.1.4 Pilot points

For models with thousands, or even millions of elements, the estimation of properties for every cell of the grid using equation (2) is computationally unachievable. We require a reduced parameter set for which we can establish a set of spatial relationships. The parametrization of groundwater models using pilot points has been widely used since the first works of De Marsily et al. (1984). Instead of trying to estimate the parameters at every cell of the discretized model domain, parameters are estimated in a smaller set of discrete locations, the pilot points, which values are interpolated to the model cells in which the domain has been discretized (Doherty et al., 2010).

2.1.5 Tikhonov regularization

Inverse problems in environmental sciences are usually non-unique (e.g. Zhou et al., 2014). Tikhonov regularization (Tikhonov and Arsenin, 1977) is implemented in PEST to achieve a unique solution (for which the functional attains its minimal) by introducing either preferred values or relationships between the estimable parameters expressed as a series of expert knowledge observations, $\mathbf{h_r}$. A regularization objective function can be defined such as

$$\Phi_r = (\mathbf{h_r} - \boldsymbol{Z}\mathbf{p})^t \mathbf{Q_r} (\mathbf{h_r} - \boldsymbol{Z}\mathbf{p}), \quad (8)$$

Here $\boldsymbol{Z}$ represents the effect of the "regularization model" in the parameters. $\mathbf{Q_r}$ is a user-provided weight matrix, different from $\mathbf{Q}$, which represents the strength of expert belief in the regularization observations, $\mathbf{h_r}$ (Doherty, 2015).

When adding the regularization term, the final objective function to minimise is the sum of the measurement or data objective function, Φ_d, and the regularization objective function, Φ_r, with μ^2 a regularization weight factor.

$$\Phi = (\mathbf{h} - \mathbf{Xp})^t\mathbf{Q}(\mathbf{h} - \mathbf{Xp}) + \mu^2[(\mathbf{h_r} - \boldsymbol{Z}\mathbf{p})^t\mathbf{Q_r}(\mathbf{h_r} - \boldsymbol{Z}\mathbf{p})], \quad (9)$$

Or expressed in compacted form

$$\Phi = \Phi_d + \mu^2\Phi_r\,, \quad (10)$$

which is the total objective function that PEST aims to minimize in regularization mode.

2.1.6 Iterative Ensemble Smoother

When using PEST as described above, the result of the calibration procedure is a unique property field with minimum error variance that fits the observation datasets. This strategy has some limitations when using a model for predictive purposes (Doherty, 2015) because (1) it usually provides a smoother property field in which fine scale heterogeneities -which might have an important impact in the predictions- are unnoticed, and (2) the unique estimated property field cannot be used alone to perform an uncertainty analysis. An available technique for uncertainty analysis using the PEST suite is to follow the known as Null-Space Monte Carlo (NSMC) strategy (Tonkin and Doherty, 2009; Herckenrath et al, 2011). NSMC allows for estimation of calibrated-constrained property fields that can be used with predictive purposes but requires to perform the calibration to obtain a model that is next disturbed to obtain an ensemble of calibrated-constrained models used with predictive purposes. The calibration step can be an important limitation for highly parameterized models and long forward running times.

PEST++ (White et al., 2020) includes the iterative Ensemble Smoother (iES) (White, 2018) with which it is possible to obtain a number of calibrated-constrained parameter fields that can be used for predictive uncertainty analysis. An important advantage of the iES is that alleviates the computational cost for highly parameterized models as the Jacobian matrix does not have to be filled with finite difference approximation of the partial first derivative at every iteration. At the contrary, the iES only requires as many models runs per iteration as members of realizations of the ensemble. Instead of using

finite difference approximation, Chen and Oliver (2013) proposed to use a Jacobian matrix obtained empirically using the following equation

$$\mathbf{J} \approx \left(C^{-1}(\varepsilon)\right)^{1/2} \Delta \mathrm{s}\, \Delta \mathrm{p}^{-1} \left(C^{-1}(p)\right)^{-1/2}, \tag{11}$$

where

$$\Delta \mathrm{s} = \frac{\left(C^{-1}(\varepsilon)\right)^{-1/2}(s-\bar{s})}{\sqrt{N_E-1}}, \tag{12}$$

$$\Delta \mathrm{p} = \frac{\left(C^{-1}(p)\right)^{-1/2}(p-\bar{p})}{\sqrt{N_E-1}}. \tag{13}$$

In the equations above, $C^{-1}(\varepsilon)$ is the covariance matrix of measurement noise and $C^{-1}(p)$is the prior parameter covariance matrix, $\bar{s}$ and $\bar{p}$ denote mean simulated and parameter values of the ensemble and N_E is the number of realizations in the ensemble.

As the number of parameters has no effect on the computational cost in the update of the model, a much finer heterogeneity detail with higher density of pilot points (i.e. as many as cells or element in the mesh of the model) can be introduced to account for smaller scale variability. As a result, it is possible to work with models with thousands of parameters requiring just some hundreds of runs for a model until it is calibrated (White, 2018). An advantage of using the iES is that the uncertainty analysis comes at no extra cost of the calibration. Because the results of application of the iES are not a unique minimum variance property field, but an ensemble of acceptably calibrated parameter fields that can be evaluated for uncertainty analysis purposes. On the contrary, and in general, calibrated models using iES do not reach as best fitting as those using the GLM algorithm (White, 2018).

For further theoretical details of the iES we refer to Chen and Oliver (2013) and to White (2018) and White et al. (2020) for use of the PEST++ suite and instructions for the additional inputs required in the PEST control file to define the control variables.

2.2 COMSOL - PEST/PEST++ Integration

2.2.1 General Workflow

The fully coupled inversion (FCI) procedure was implemented through integrating the model-independent calibration software PEST (Doherty, 2020) and PEST++ (White et al., 2020) with MATLAB and COMSOL (**Fig. 1**). Previously, Halloran et al. (2019) presented COMPEST, an interface built in Java to link PEST and COMSOL, that they satisfactorily applied to a case of isotopic fractionation of groundwater contaminants (Halloran et al., 2021). In this work we follow a different strategy, using MATLAB for connection instead of Java, and with some important additions, including the use of pilot points for spatial parametrization, Tikhonov regularization, singular value decomposition, parallelization to reduce computational burden and the integration with PEST++ that provides global sensitivity and uncertainty analysis.

In COMSOL, we built the forward coupled hydrogeophysical model with additional pre- and post-processing computations performed using MATLAB (González-Quirós et al., 2019). We used MATLAB and the LiveLink for MATLAB as a linking platform between COMSOL and PEST/PEST++. Appropriate COMSOL license is necessary to implement the workflow.

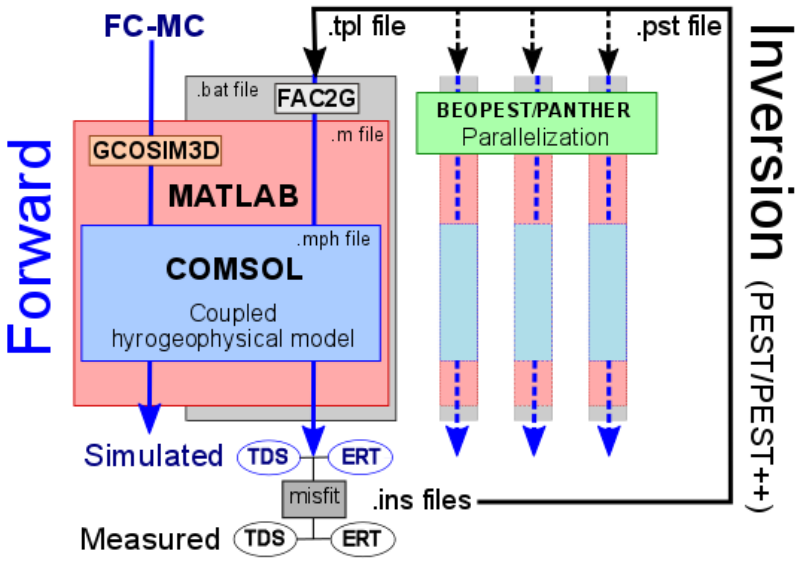


***Fig. 1.**- Flowchart of integration. Communication between MATLAB and COMSOL is achieved using the ComsolServer and Livelink for MATLAB. Template (.tpl) and instruction (.ins) files are required to*

communicate PEST with the relevant model inputs and output files. TDS (total dissolved solids) and ERT (electrical resistivity tomography) represent the observation datasets. FC-MC indicates the forward fully coupled Monte Carlo workflow explained in section 2.3.2.

2.2.2 PEST/PEST++ input files

Both PEST and PEST++ require, at least, three types of files (**Fig. 1**): instructions (.ins), template (.tpl) and control (.pst). The instructions file(s) contain information that point to model output for comparison between simulated values and the observed datasets. The template file includes the information of the parameters to be calibrated that are used as input files in the model run. The control file is the core of the PEST/PEST++ workflow as it contains the options of the model calibration process, the observation data, groups and weights, parameter ranges and groups, the name of instructions and template files and the file to run the forward model (wrapped in a .bat file). Specific details to fill the relevant PEST files can be found in its extensive and complete documentation (Doherty, 2020) For the additional options necessary in the PEST++ control file to use the iterative Ensemble Smoother we refer to White et al. (2020).

For implementation of pilot points and regularization, additional applications of the PEST suite (Doherty, 2020) are necessary. After defining the desired pilot point locations in the model domain, we used the functionalities PPK2FAC and FAC2G to generate a set of kriging factors from the desired geostatistical model and for interpolation from the pilot point locations to a defined finer grid. The generation of kriging factors using PPK2FAC is performed once before the start of the calibration, while FAC2G functionality is required before every model run and included as part of the workflow in the batch file to run the model (COMLINE in the PEST control file).

2.2.3 Parallelization

To speed up the procedure we used parallelization through the BEOPEST (Schreuder 2009), for the model calibration using PEST, and PANTHER parallel run manager (Welter et al., 2019), for use with PEST++. For each thread, a ComsolServer™ was launched with a different port assigned. Then, each worker was connected to the designated port using the LiveLink for MATLAB® application. The

COMSOL-MATLAB forward model was wrapped in an .m file and the order to run it was included in the COMLINE batch file.

2.3 Application to Hydrogeophysical Inversion in Saltwater Intrusion Modelling

2.3.1 Geophysical Monitoring of Saltwater Intrusion

As an example application we show how electrical resistivity and borehole hydrogeological observations can be used with the proposed methodology to calibrate and perform a sensitivity and uncertainty analyses of saltwater intrusion in a multiphysical model that simulates a synthetic heterogeneous coastal aquifer (**Fig. 2**). Electromagnetic methods (in the broad sense and including electromagnetic and electrical resistivity techniques among others) are well-established geophysical investigation techniques in coastal studies because of their sensitivity to saltwater content (Jiao and Post, 2019). Electrical resistivity imaging (ERI), in particular, with the popularisation of multi-channel instrumentation enabling rapid high-resolution data acquisition allows hydrogeologists to image saltwater intrusion patterns over distances of meters to thousands of meters (e.g. Goebel et al., 2017; Comte et al., 2017; Costall et al., 2018, 2020).

2.3.2 Forward Coupled Modelling

The reference groundwater model for forward and inverse hydrogeophysical modelling is a model of variable density flow and salt transport in a synthetic heterogeneous coastal aquifer based on a modification of the well-known Henry's problem (Henry, 1964) with the following boundary conditions: towards the coast, a constant hydrostatic pressure condition (no tides or waves) is imposed in the steeped boundary of the aquifer with a constant saltwater concentration; inland is imposed a constant freshwater inflow of $1.75 \cdot 10^{-3}$ $m \cdot d^{-1}$ and at the top and bottom a zero flux condition is imposed. **Table 1** compiles the characteristics of the model.

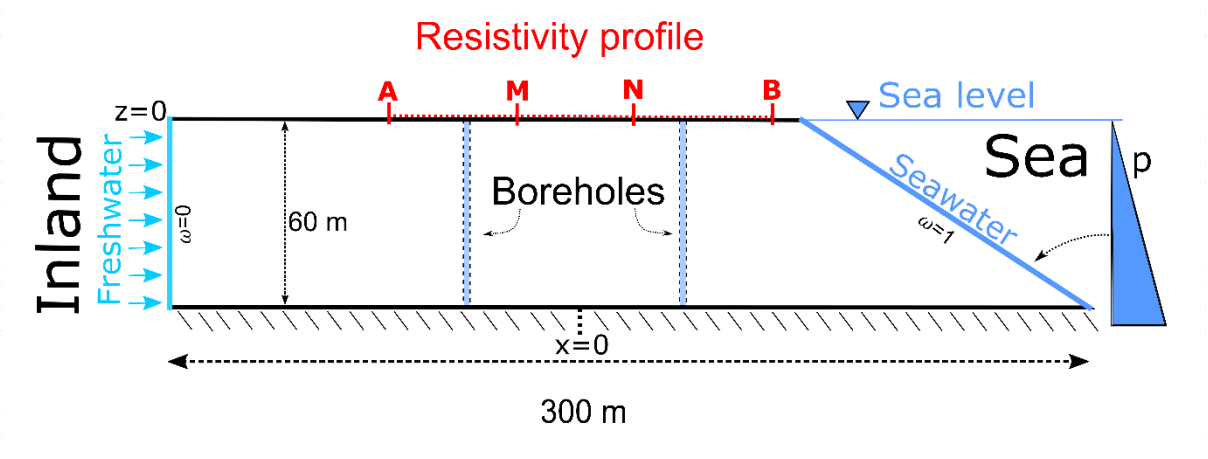


Fig. 2 *Characteristics of the modelled scenario. A, B and M, N represent, respectively, the current and potential electrodes of a Wenner-α quadripole. ω is a relative salt mass fraction, 1 for seawater.*

Table 1. *Groundwater model parameters.*

Parameter	Symbol	Value	Unit
Saltwater density	ρ_s	1025	kg m^{-3}
Freshwater density	ρ_0	1000	kg m^{-3}
Saltwater salinity	TDS	35	g l^{-1}
Freshwater electrical conductivity	σ_{Fw}	500	µS cm^{-1}
Saltwater electrical conductivity	σ_{Sw}	50000	µS cm^{-1}
Average Mean Hydraulic conductivity	K_h	1×10^{-5}	m s^{-1}
Effective porosity	ϕ	0.33	-
Molecular diffusion	D_m	1×10^{-9}	m^2 s^{-1}
Longitudinal dispersivity	α_L	5	m
Transversal dispersivity	α_T	0.5	m

A petrophysical model —which we assumed perfectly describes the relationship between the groundwater and electrical models— was used to transfer the groundwater variables and parameters after solution of the groundwater model into electrical resistivities. In this work we used the Waxman and Smits (1968) petrophysical relationship and associated petrophysical equations (Appendix A) to link the groundwater and electrical resistivity models. Finally, a synthetic surface electrical resistivity acquisition was performed on the surface with a Wenner-alpha array of 72 electrodes with 2 m spacing centred at coordinate x=0 (**Fig. 2**). The fully coupled model was implemented in COMSOL and solved sequentially in a single run following the strategy explained in González-Quirós et al. (2019) and González-Quirós and Comte. (2020). Two different meshes were used to solve each problem, a fine

mesh for the groundwater problem (19726 elements) and a coarser mesh but refined around the electrode locations for the ERT problem (7376 elements).

First, we performed a Monte Carlo (FC-MC) analysis to obtain the forward hydrogeophysical response of 500 random realizations of heterogenous log-permeability multiGaussian fields generated using the software GCOSIM3D (Gómez-Hernández and Journel, 1993). We used a spherical variogram with a range of 50 m and a sill of 1. For each permeability scenario we simulated the groundwater (flow and salinity concentration) and electrical responses (apparent resistivities). The fully coupled Monte Carlo analysis workflow is illustrated as FC-MC in **Fig. 1**. Within MATLAB, GCOSIM3D was run to generate the random permeability fields that were used by COMSOL for simulation of the forward response applying the petrophysical relationships explained in the Appendix A.

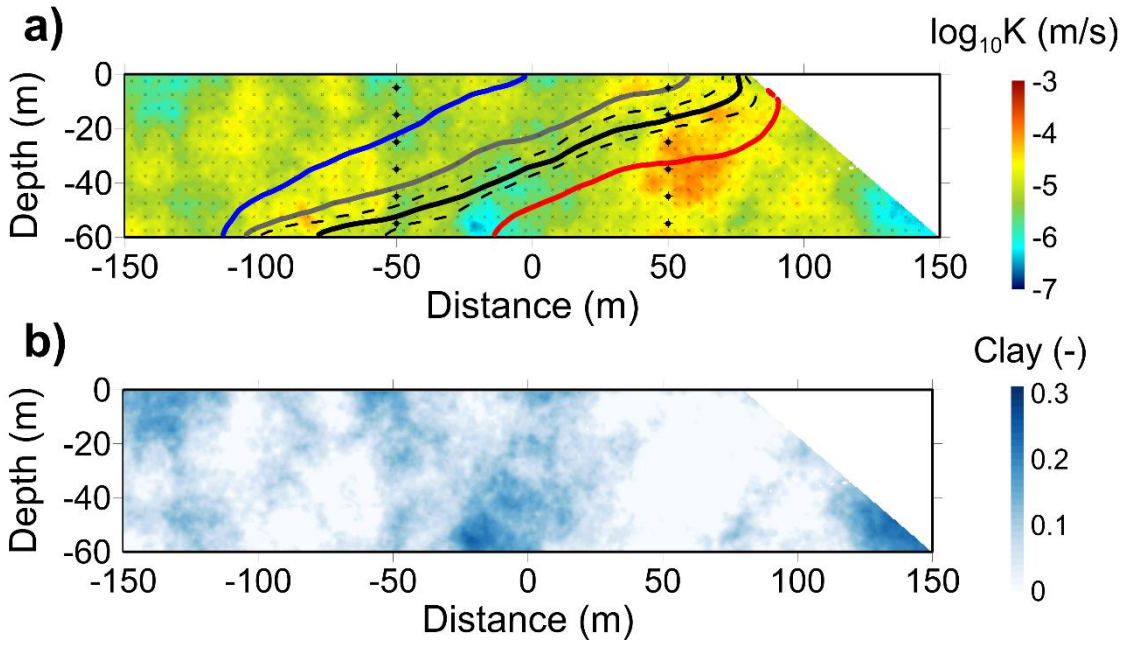


Fig. 3. *Computed hydraulic conductivity* ***(a)*** *and clay fraction* ***(b)*** *for one selected reference model (one of the 500 permeability stochastic realizations). In* ***(a)****: Blue contour is the DWL and continuous iso-contours are for relative concentrations of 0.1 (grey), 0.5 (black) and 0.99 (red). The width of the mixing zone is computed between the contours 0.25 and 0.75 (dashed black contours). Black diamonds are observation points in the two boreholes and crosses are pilot point locations.*

Calculated salinities from the 500 models were analysed using saltwater intrusion indicators. For the analysis of the models we were mainly interested in the position of the contour associated with the DWL (for Drinking Water Limit) taken at 0.5 g l^{-1} (USEPA, 2009). Values with salinity concentrations above

this limit are considered as poor quality or undrinkable water. We also considered the width, or spread, of the mixing zone between the iso-concentrations 0.25 and 0.75 (**Fig. 3**).

2.3.3 Observation datasets

From the set of solved models explained in section 2.3.2, we selected a scenario with large saltwater intrusion as the "true" model (**Fig. 3**). For the calibration data set we used 12 observation points for salt concentration (**Fig. 3**), expressed as TDS in kg·m^{-3}, from different depths of 2 boreholes at coordinates $x = -50$ m and $x = 50$ m, and the apparent resistivities (in Ohm·m, or Ω·m) from 828 ERT quadripoles of the 72-electrode Wenner-alpha array. For the concentration points we assumed that the measured data at the two boreholes are those of the aquifer at the same locations.

Theoretical simulated responses, salinity concentrations and logarithm of apparent resistivities, from the considered true model at the observation locations, were disturbed with uncorrelated Gaussian noise with standard deviation of 5% to define the observation dataset. These two sets of data were used to form the measurement objective function, which is calculated as the weighted squared residuals, r_i, between the observations, here the noise contaminated data. The observation weights, w_i, were assigned as the inverse of the variance of the observation error (Hill and Tiedeman, 2007). TDS and ERT (apparent resistivity) data were assigned to two different observation groups. To assign the same contribution from each of the observation groups we used the PWTADJ1 application from the PEST suite (Doherty, 2020).

2.3.4 Calibration with PEST-GLM

To perform the calibration of the coupled hydrogeophysical inversion model, we first used PEST (Doherty, 2020) for the estimation of permeability in 636 pilot points uniformly distributed in the simulation domain (**Fig. 3**). Although the GLM algorithm was used in COMPEST (Halloran et al., 2019), the use of pilot points for spatial parametrization in COMSOL is a novelty of this work. A computational limitation, especially for highly parameterized models and long forward running times, is that the GLM algorithm requires to fill the Jacobian matrix at every iteration by computing a finite-difference approximation of the partial first derivative, which is directly related with the number of

parameters, in our case the same number of pilot point locations, 636. To reduce the computational burden, we used Singular Value Decomposition (SVD) with a reduced set of super-parameters. To proceed, we computed once the full Jacobian, which required a total of 636 model runs, as many as the number of pilot points. After running the SUPCALC application (Doherty, 2020) we chose the suggested maximum number of 100 super-parameters, which was an important reduction from the initial 636 pilot-points, and therefore, a decrease of the computational cost in the inversion. Finally, we ran the SVD-Assist (Doherty, 2020) to generate new control and input files that were used for the inversion following the same approach explained above and shown in **Fig. 1**.

2.3.5 Iterative Ensemble Smoother (PEST++ iES) and Uncertainty Analysis

For application of the PEST++ iterative ensemble smoother (iES) to the same hydrogeophysical saltwater intrusion scenario described in section 2.1 we used the same observation datasets described in section 2.3.3. To take advantage of the capabilities of parameter space evaluation of the iES, and aiming to recover finer scale features, we increased the number of pilot points to 4650, distributed randomly, but using the same geostatistical model explained in the previous sections.

With PEST++-iES the uncertainty analysis comes at no extra cost to the model calibration. Aiming to perform a post-calibration uncertainty analysis to a quantity of models comparable with the one used in the stochastic example described in section 2.3.2, we applied the iES to an initial ensemble of 500 random parameter fields. However, the initial fields of the ensemble were generated automatically by PEST++ and are different than the 500 models used in the Monte Carlo analysis that were generated using GCOSIM3D. At every iteration, the ensemble was updated as explained in section 2.1.6 and references therein. We used the last ensemble of models for the uncertainty analysis using as metrics the same saltwater intrusion indicators (DWL position and penetration and mixing zone width) explained in section 2.3.2. Additionally, we considered as the best model of the ensemble the one with the best fit (lowest objective function) and computed the ensemble average and variance for the estimated hydraulic properties.

3 Results

3.1 Stochastic Groundwater Modelling Results

Each single forward coupled model was computed in less than 80 seconds in an Intel® Core™ i5-8500 with 40 GB RAM. For the forward stochastic modelling exercise, from the total of 500 model realizations of the random field, 8 failed to converge (i.e., less than 2%), mainly because of a particular arrange of low hydraulic conductivities in the density-dependent flow model.

Representative computed results are compiled in **Fig. 4**. **Fig. 4a** shows the positions of the drinking water limit contour (TDS=0.5 $kg \cdot m^{-3}$) for the resolved 492 models. The penetration of the toe position of the DWL spans for almost 200 m. A more detailed statistical representation of the toe penetration is shown in the histogram of **Fig. 4b**. In this figure we present the distribution of computed distances from the coordinate x = - 150 m (defined at the bottom left corner of the model) to the toe position of the DWL at 5 m intervals. To represent statistics of the ensemble of models we computed a mean distance of 117.1 m, median of 119.1 m and a standard deviation of 40.9 m. In the histogram of **Fig. 4c** we compile the computed values of width (or spread) of the mixing zone –distance between the iso-contours of concentrations 0.25 and 0.75— at the bottom of the aquifer. Representative statistics of the ensemble of models were mean mixing zone width 19.3 m, median 15.7 m, and standard deviation 12.5 m. The computed spread of the mixing zone for the chosen reference model was 45.7 m. The results of the reference model are represented in all figures with a red line. The reference model was chosen for being more challenging for identification during the inversion procedure because representing a statistically less frequent scenario. This model was characterised by a DWL toe position of 37.5 m from x = -150 and a width of the mixing zone of 45.6 m.

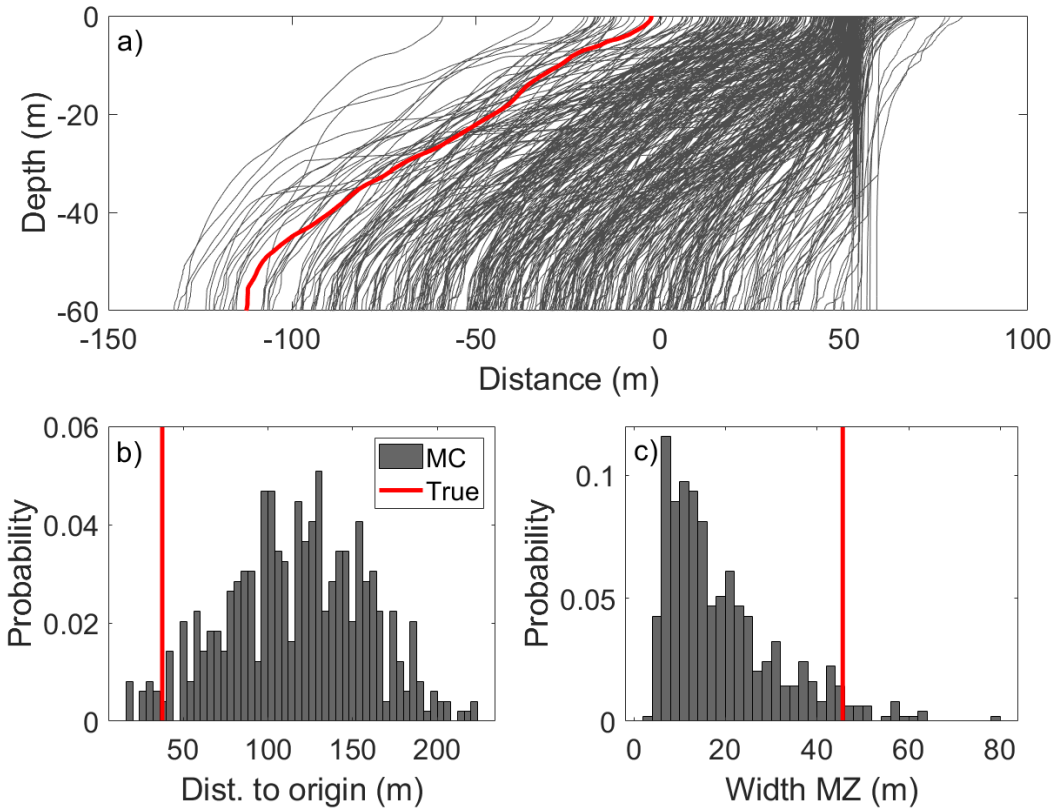


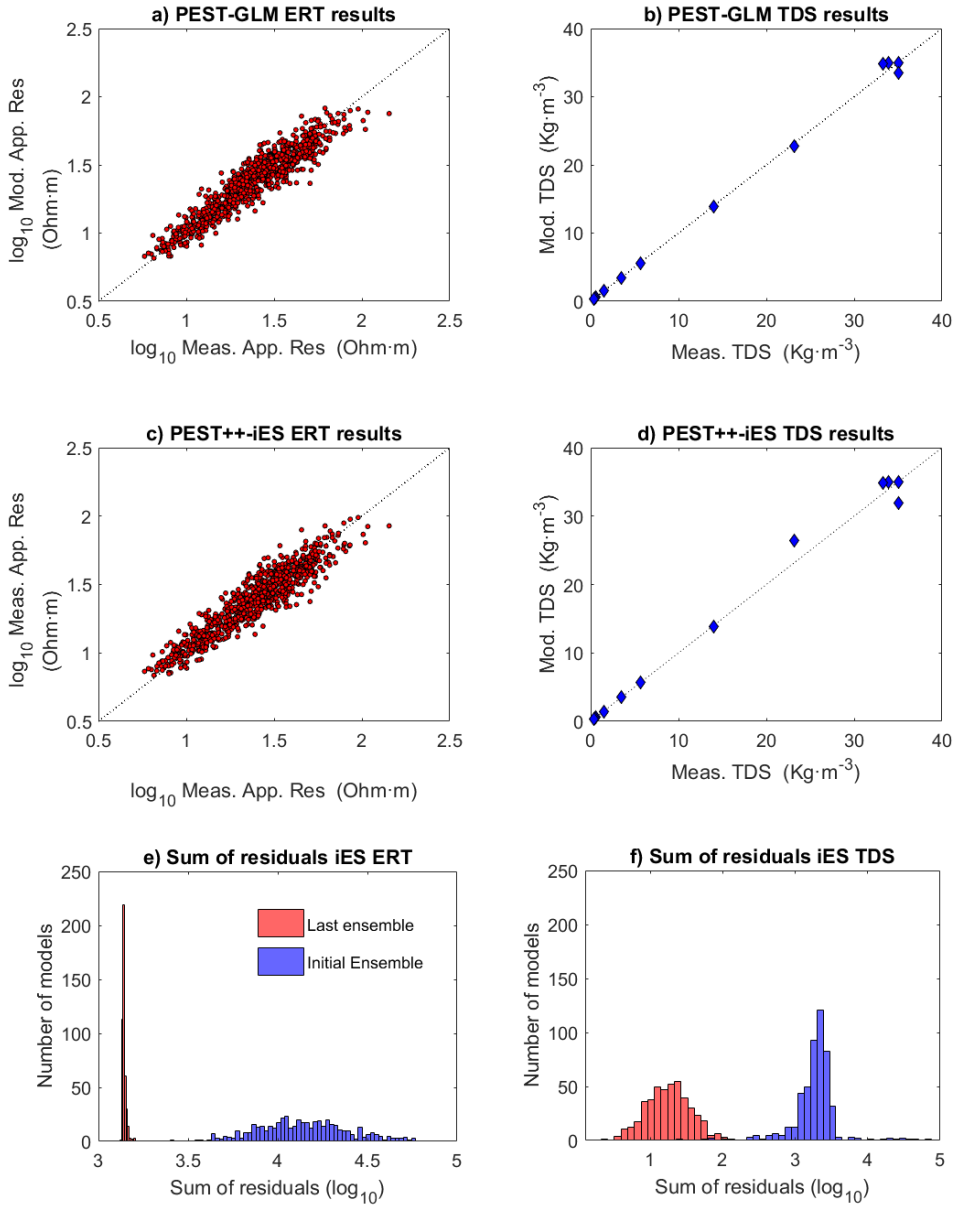
Fig. 4. *Results from the 492 solved models.* ***(a)*** *DWL contour position.* ***(b)*** *Histogram of distances from coordinate x=-150 m to the DWL at the toe (refer to Figure 2)* ***(c)*** *Width of the mixing zone at the toe. Red lines indicate the reference model in all graphs (a-c).*

3.2 Fully Coupled Hydrogeophysical Inversion Results

Fig. 5 shows the calibration results after application of PEST-GLM. We show data fitting between measured and modelled (calibrated) apparent resistivities (**Fig. 5a**) and total dissolved solids (TDS) (**Fig. 5b**). To reach these results PEST was run with Tikhonov regularization and SVD (100 superparameters) and required 20 iterations and 3293 model runs (plus the 636 model runs to compute the initial Jacobian matrix before the SVD implementation). After evaluation of the evolution of the objective function we observed an acceptable fit after 8 iterations and 841 model runs (**Fig. 6**).

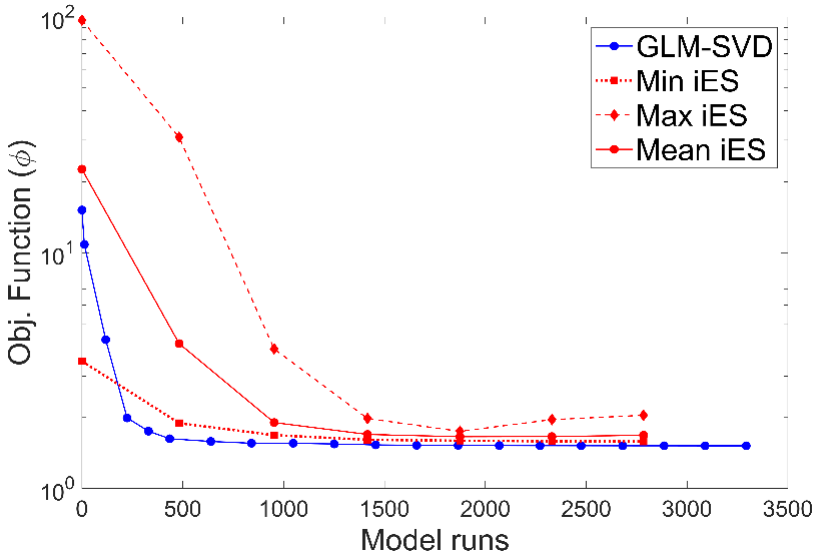
The iES approach evaluated over an initial ensemble of 500 permeability fields was run for 6 iterations but reached a good fit after 4 iterations and 1874 model runs (**Fig. 6**). From the initial ensemble, 53 models were discarded by PEST++ along the procedure. That is, when model run failure is encountered (e.g., because of convergence problems) PEST++-iES drops one parameter set from the

ensemble before continuing to the next iteration (White, 2020). The remaining 447 models of the final ensemble were used to perform an uncertainty analysis. In **Fig. 5** we show the results of the model with best data fit (min Φ) for measured and modelled (calibrated) apparent resistivities (**Fig. 5c**) and total dissolved solids (TDS). The latest (**Fig. 5d**) is slightly worse than the results applying the PEST-GLM, especially for some high values of TDS. Distribution of sum of squared residuals between the initial and last ensemble is shown or each of the observation groups, ERT and TDS in **Fig. 5e** and **Fig. 5f** respectively.



418

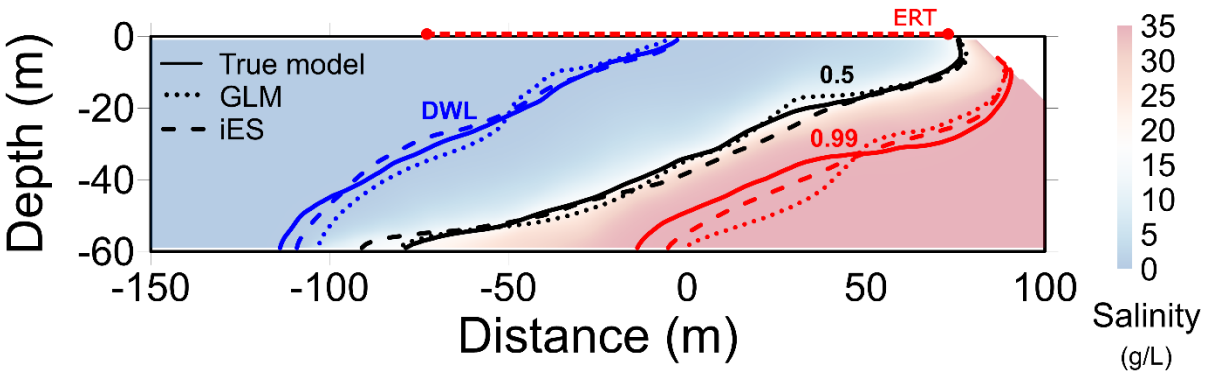
419 ***Fig. 5.*** *Modelled (calibrated) vs. measured apparent resistivities (in Ohm·m)* ***(a, c)*** *and TDS (in Kg·m^{-3})*
420 ***(b, d)*** *for calibration using PEST-GLM* ***(a, b)****, and best model fit using the iES* ***(c, d).*** *Distribution of the*
421 *sum of squared residuals for the two observation groups for the initial and last ensemble of models* ***(e,***
422 ***f)****.*



423

424 ***Fig. 6.*** *Evolution of the objective function (normalized with number of observations) with model runs (each*
425 *marker represents an iteration) for GLM mode and ensemble minimum, maximum and mean in iES mode. For*
426 *reference, one model run with 6 models in parallel and no GUI is computed in ~80 s.*

427 In **Fig. 7** we compare the recovered position of the saltwater-freshwater mixing zone between the
428 calibrated models and the true position. Representative saltwater concentration iso-contours show that
429 the mixing zone was well resolved, especially around the centre of the domain (coordinate $x = 0$ m)
430 and near the surface, which are the most sensitive regions to the observation datasets. Some
431 discrepancies were observed in the toe for the penetration of the DWL and the 0.1 iso-salinity contour
432 around coordinate $x = -100$ m and in the delineation of the 0.99 salinity concentration. We obtained
433 a better correspondence in the position of the representative iso-contours using the iES method. This
434 might be derived from the introduction of heterogeneities out of the regions of data sensitivity, which
435 is an important difference between the iES and the GLM approaches. It must also be noted that, rather
436 than a single model, the iES resulted in an ensemble of models that provide a range of uncertainty in
437 the position of the different salinity contours (this is illustrated in **Fig. 9** for the DWL).



438

439 ***Fig. 7.*** *True (continuous) and recovered (GLM-dotted and iES-dashed) representative position of the*
440 *DWL, 0.5 and 0.99 salinity iso-contours. ERT profile and borehole locations are shown for reference.*

441 **Fig. 8** shows a comparison between true and best estimate distribution of permeabilities using both
442 the GLM (**Fig. 8b**) and iES (**Fig. 8c**). The use of a much smaller number of pilot points and the minimum
443 variance strategy of the GLM workflow is well noted in the resolution of the recovered sections shown
444 compared with the iES. The GLM strategy is prone to create artifacts in the sensitive regions to fit
445 within the noise level. Some regions were acceptably well recovered, especially in the shallower 20 m
446 along the domain of investigation of the ERT data in the centre of the domain. Examples of well
447 recovered are the low permeability shallow regions near the surface at coordinates $x = 0$ m and $x =$
448 -50 m or the higher permeability regions at $x = -25$ m and $x = 50$ m.

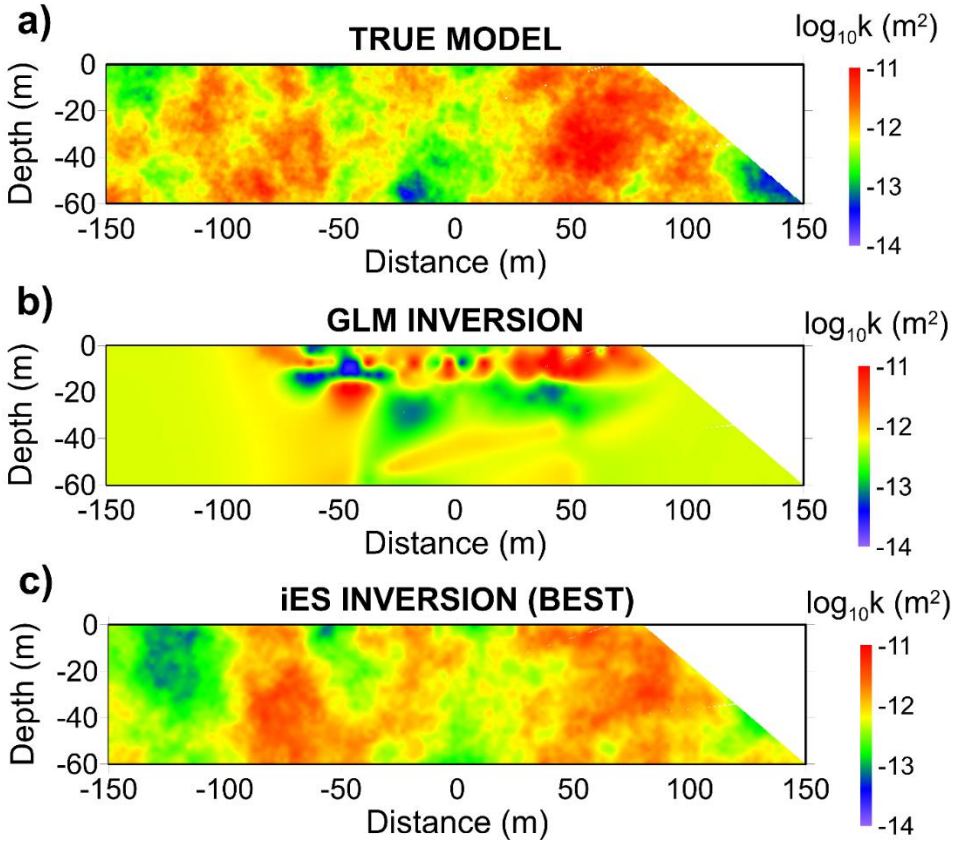


Fig. 8. *True* ***(a)*** *and recovered permeability distribution after inversion with* ***(b)*** *GLM and* ***(c)*** *the iterative ensemble smoother.*

The inversion found some problems with the extreme values of highest and lowest permeabilities at deeper locations, for example around coordinates $[50, -40]$ and $[-20, -55]$ respectively, which are in the limits of investigation of the electrical resistivity profile (**Fig. 10**). This was especially problematic for the GLM inversion that resulted in inaccurate results below 20 m depth and introduced two shifted higher and lower permeability patches at $[-50, -20]$ and $[-25, -25]$ and "bulls eye" type inversion artifacts. Also, because the GLM inversion started from a homogeneous distribution of permeabilities, it only introduced heterogeneity progressively in the central regions of the domain, related with data coverage, especially ERT. Regions towards inland from $x = -75$ m and towards the sea from $x = 75$ m remained homogeneous after the inversion.

3.3 Uncertainty Analysis

Fig. 9 shows representative results (blue lines) for the ensemble of 447 models (of the initial set of 500) obtained after 4 iterations. For comparison purposes we show (in grey) the results of the unconstrained Monte Carlo analysis presented in in section 3.1. With a red line we show the reference model and with a green line we indicate the best model estimate (the one presented in section 3.2).

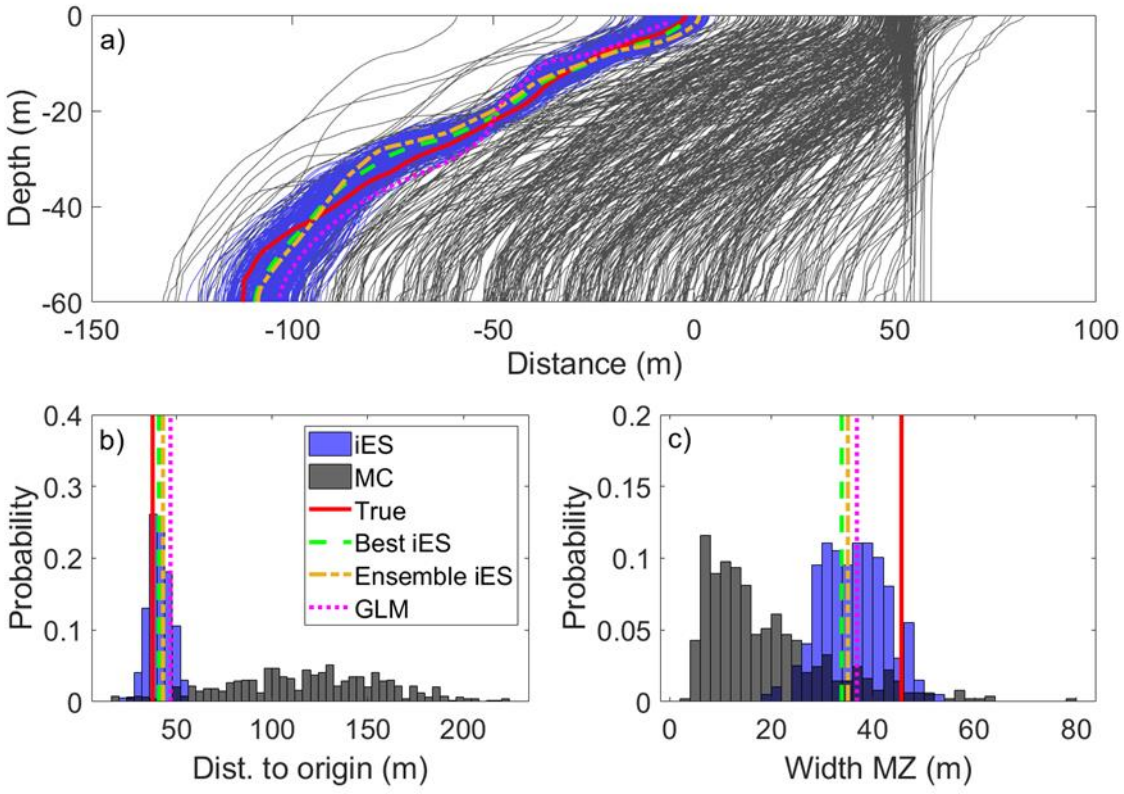


Fig. 9. *Results of the uncertainty analysis, in blue, from the last iteration of the ensemble of realizations obtained after application of the ES.* ***(a)*** *DWL contour position.* ***(b)*** *Histogram of distances from coordinate x=-150 to the DWL toe.* ***(c)*** *Toe width of the mixing zone. Red line is true reference model, orange is the best model of the ensemble, green ensemble average and magenta results of PEST-GLM. We show, for comparison, the results of the Monte Carlo analysis (MC) from fig* ***Fig. 4****.*

Fig. 9a shows the position of the DWL contours. Best model estimate is 40.8 m from the origin, while the computed ensemble mean is 42.9 m, with a standard deviation of 5.2 m (**Fig. 9b**). The difference with the true toe location (37.5 m) is 3.3 and 5.4 m, for the true and ensemble mean, respectively. The width of the mixing zone for the best estimate was 33.8 m, while the computed

ensemble mean was 35 m, with a standard deviation of 5.9 m. Differences with the true width of the mixing zone at the toe (45.6 m) were 14.1 and 10.6 m, for the true and ensemble mean, respectively.

Fig. 10 illustrates the differences in the distribution of permeability between the true model (**Fig. 10a**) and the ensemble mean (**Fig. 10b**) computed from all the models of the final iES ensemble. Some regions were acceptably well recovered, especially in the central domain of investigation of the ERT data. Examples of well recovered regions are the low permeability region near the surface at coordinates $x = 0$ and $x = -50$. In **Fig. 10c** it is shown the variance for the ensemble of realizations. The region with lower values is located near the surface at the centre of the domain, where the data coverage, especially ERT, is higher.

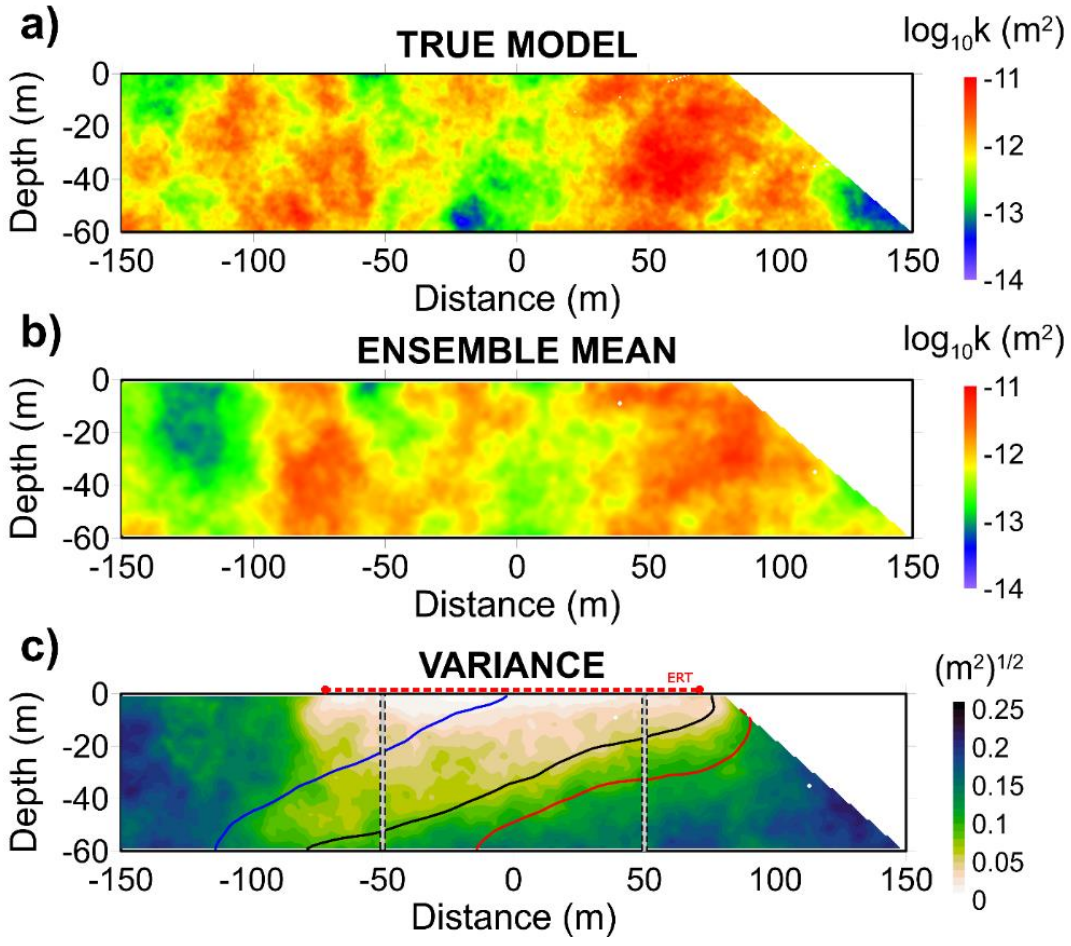


Fig. 10 (a) *True (reference) model,* ***(b)*** *ensemble mean and* ***(c)*** *variance of permeability (log_{10}). ERT profile, boreholes location and contour lines representing true DWL, 0.5 and 0.99 salinities are shown for reference.*

4 Discussion

4.1 Comparison with standalone ERT inversion

It is well known that stand-alone inversion of geophysical data is a blurry, usually a smoother, representation of reality in which small scale features of the subsurface are not well recovered because of resolution limitations (Day-Lewis et al., 2005; Singha and Moysey, 2006). Additionally, there is a loss of resolution with depth that, when studying saltwater intrusion in coastal aquifers, has resulted in discrepancies when comparing tomograms obtained from surface ERI with borehole salinity data (e.g. Palacios et al., 2020). As a result, when interpreting ER tomograms in terms of salinities there is an effect of overdispersion (González-Quirós and Comte, 2020) that can be an important source of bias in uncoupled hydrogeophysical inversion.

Fig. 11 shows the comparison between the true resistivity distribution (**Fig. 11a**), the results of the coupled inversion with both the PEST-GLM (**Fig. 11b**) and the PEST++-iES (**Fig. 11c**) strategies, and a stand-alone ERT inversion (performed here with the widely tested software BERT; Günther and Rücker, 2015) using the same geophysical dataset and measurement error.

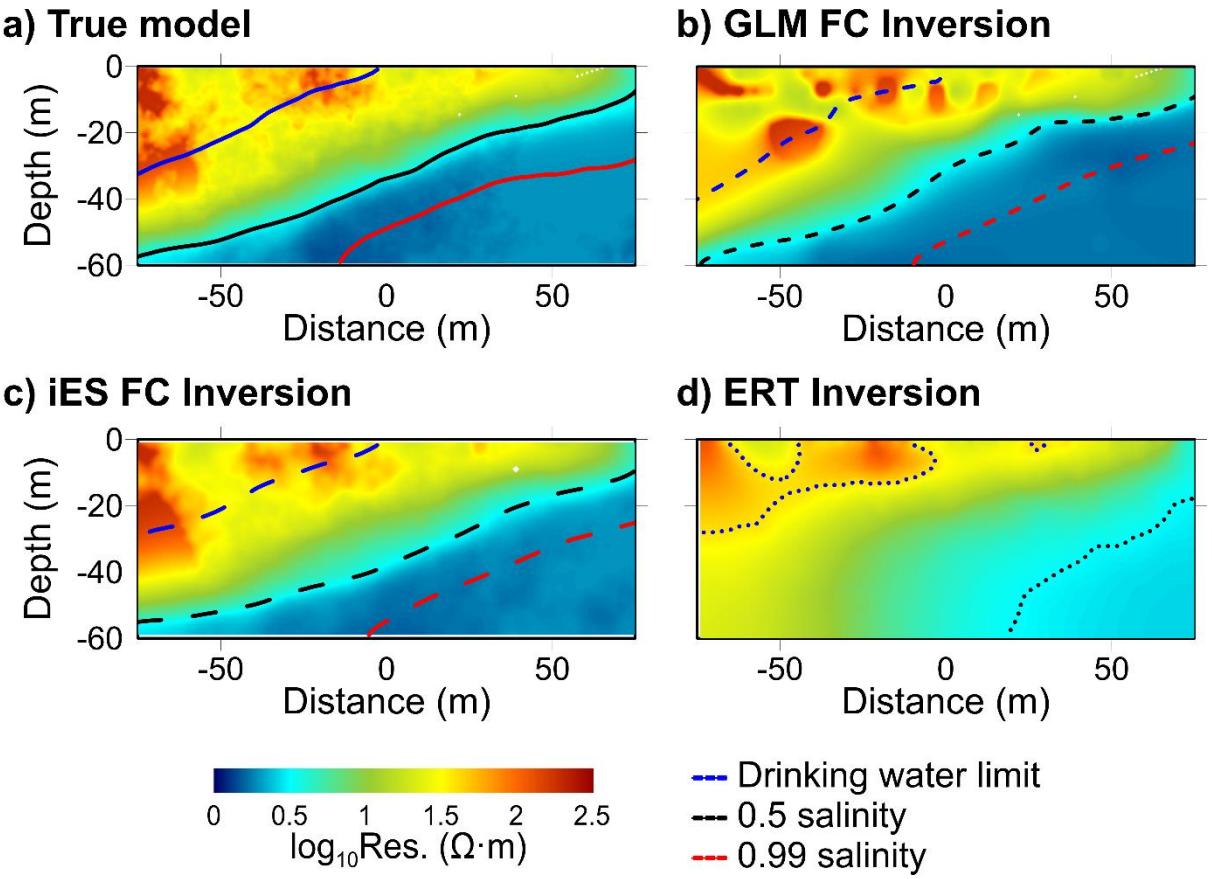
With PEST-GLM the inversion recovers some of the heterogeneities in the near surface, mainly in the higher resistivity (low salinity) regions, and the two patches of higher resistivity located around coordinate $x = -70$ m. It also introduces other small heterogeneities near the surface that are not observed in the true model, and which allow PEST to improve the model fitting. The observation datasets are not sensitive to changes in some zones of the aquifer and, as a result, the inversion cannot recover some of the higher resistivity zones located towards inland or in the saline water regions. These zones are recovered with better resolution by PEST++-iES (**Fig. 10c**) showing the importance of using appropriate prior information to constraint the inversion. Even though, both fully coupled inversion routines provide reliable results regarding salinity content, especially in the most conductive regions.

The stand-alone ERT inversion (**Fig. 11d**) shows a smoother distribution of resistivities, especially in the mixing zone where the transition between fresh and saltwater aims to be identified. This complicates the delineation of the mixing zone and the identification of the position of the saltwater

intrusion. When compared with a stand-alone ERT inversion, the fully coupled hydrogeophysical inversion shows a much better correspondence and definition of the saltwater-freshwater mixing zone, even when slightly underestimates the toe location. Additionally, it overrides the use of constant resistivity value to delineate the saltwater intrusion, as the coupled model directly provides a solution of the groundwater modelling scenario. This is important as it has been shown that using a constant resistivity threshold for delineation of the mixing zone may introduce large biases in heterogeneous aquifers (González-Quirós and Comte, 2020), especially when assuming a homogeneous distribution of petrophysical properties.

By application of a coupled hydrogeophysical inversion, the groundwater model acts as a physically-based constraint for the geophysical model, and therefore more realistic distributions of state variables can be obtained. This however requires an adequate, and sometimes more complex, forward model that considers other elements (e.g., the effect in the geophysical measurements of variable saturation near the surface, temperature variations or the boundary conditions of the groundwater model) not required in stand-alone ERI inversion. These characteristics that are difficult to quantify in real conditions might have a strong influence in the results.

It must also be noted that in this work we used a Wenner-alpha array for having low noise in coastal settings. Other arrays, or a combination of different arrays, with higher sensitivity to the target structures and higher sampling density would further improve inversion results. In addition, the choice of the petrophysical model, here assumed to be known and certain, is another important source of error and uncertainty in hydrogeophysical applications (Irving and Singha, 2010; Brunetti and Linde, 2018; Tso et al., 2019; González-Quirós and Comte, 2020).



535

536 ***Fig. 11.*** *Log10 electrical resistivity (Ohm.m):* ***(a)*** *True model,* ***(b)*** *fully coupled inversion using PEST-GLM,*
537 ***(c)*** *fully coupled inversion using the iterative ensemble smoother and* ***(d)*** *stand-alone ERT inversion.*
538 *Continuous lines in (a) show the true representative salinity iso-values. Recovered salinity iso-contours from*
539 *fully coupled inversion are shown with short and long dash lines for GLM (b) and iES (c) respectively. In* ***(d),***
540 *salinity contours, shown with dotted lines, are obtained from petrophysical transformation assuming a*
541 *homogeneous model (as explained in González-Quirós and Comte, 2020).*

542 ## 4.2 Comparison of the Coupled Inversion Strategies

543 In this work we used to calibration strategies, GLM and iES, available from the suites of PEST and
544 PEST++ respectively. Both offer distinct strategies and are different in their capabilities and limitations.

545 As local optimization method, the GLM inversion results in a single minimum variance property
546 field, with a smoother distribution of properties, conditioned by the number and location of the pilot
547 points, which results resemble those of the smooth (or Occam's) type inversion (Constable et al., 1987)
548 widely used in geophysical applications (**Fig. 11**). In addition, we observed some artifacts that are well
549 known in the literature such as "bulls eye" and overfitting (**Fig. 8b**), associated to the amount and

distribution of pilot points (Doherty et al., 2010). In this work we used a homogeneous distribution of pilot points aiming for the inversion to be computationally treatable. The use of a very large number of pilot points is an important limitation of the GLM workflow due to the computational effort required to calculate the Jacobian. This could be an important drawback for calibration of multiphysical models that require long computational times. Further analysis towards an optimization number and location of pilot points when using surface ERT data could improve the results of the coupled inversion using PEST-GLM. An additional limitation is that the result provided by the GLM is a unique smooth property field that should not be used on its own to perform an uncertainty analysis, requiring additional efforts and strategies such as the Null Space Monte Carlo method. On the other hand, the GLM produced a good estimation of the DWL and the mixing zone (**Fig. 7**) with lower number of model runs necessary to achieve and acceptable model fit than the iES (**Fig. 6**).

The iES strategy, on the contrary, allows for the evaluation of highly parametrized models solving the restriction in the parameter space, and therefore in computational cost, for the calculation of the Jacobian. With this advantage it is possible to introduce a much larger number of pilot points resulting in the recovering of finer scale changes in the distribution of properties. Even more, the resulting ensemble of acceptably calibrated models can be used for a meaningful multi-model uncertainty analysis accounting for non-uniqueness at no extra computational cost and without the requirement of performing or implement additional strategies. On the other hand, the iES approach is based on some assumptions (linearity, Gaussian distributions of the prior) that might not be valid n some scenarios and may require different approaches (e.g., Irving and Singha, 2010) to obtain an accurate uncertainty quantification. Model fit achieves is also not as good as the one obtained with the GLM (**Fig. 6**). However, the better fit in the GLM strategy is achieved by creating numerical artifacts (**Fig. 8b**), which might be not desirable either. Further analysis is necessary to evaluate the impacts of, among others, changing the number of members in the iES ensemble, using alternative amounts and locations of pilot points, or assigning different weights to the multiphysical observation groups.

Finally, as both have advantages and limitation, the coupling workflow presented in this paper proves that they can be used complementary and applied in the same study with no further difficulties

577 due to the correspondence between file formats and implementation steps. This provides a framework
578 for application of a wide range of solutions of the PEST/PEST++ suite, including global optimization
579 and global sensitivity analysis (Morris and Sobol) methods, not discussed here, but that can be
580 implemented following the workflow presented in this work with multiphysical models developed in
581 COMSOL.

582 ## 4.3 Advantages of the Multiphysical Approach

583 The methodology presented in this work aims to provide a methodological framework for full
584 integration of geophysical information for calibration and uncertainty analysis in hydrogeological
585 studies for decision making support (Ferré, 2017; Doherty and Moore, 2020). In this regard, geophysics
586 could be used either for conceptual model testing (Linde, 2014; Brunetti et al., 2017; Lopez-Alvis et
587 al., 2019; Enemark et al., 2020) as an affordable and nimble monitoring alternative for evaluation of
588 hydrological properties (e.g., storage), aquifer boundary conditions (e.g., recharge, pumping, interaction
589 with surface water bodies) or alternative forecasts (e.g. to track the evolution of a contaminant plume
590 and compare with different model predictions).

591 We have identified two main advantages of fully integrating geophysics in the calibration. First, and
592 more importantly, our primary objective of characterization of the freshwater-saltwater mixing zone
593 has been achieved very satisfactorily. The position of the DWL is very well recovered throughout the
594 model domain both using GLM and iES strategies (**Fig. 7**). Secondly, the geophysical data is able to
595 constraint and identify better the spatial distribution of hydraulic properties near the surface, especially
596 in the freshwater domains (**Fig. 10c**). Further evaluations on the possibilities of using this information
597 to identify and constrain structural characteristics of the aquifer, complemented with multiphysical
598 uncertainty analysis, offers a promising framework for predictive evaluation and risk assessment in
599 coastal aquifers and other hydrogeological settings.

600 Even more, the strategy allows to follow a multi-scale, multi-dimension strategy; that is, the
601 multiphysical forward model can be solved in COMSOL in different domains that are coupled using ad
602 hoc strategies and operators. An example was shown for an hydrogravity model by González-Quirós

603 and Fernández-Álvarez (2014), who used a 2D domain to solve the groundwater problem and an
604 extruded 3D domain to solve, coupled, the gravity problem. Alternative strategies that will require
605 imaginative solutions, for example to solve the ERT problem in 2.5D in a reduced scale as a subproblem
606 of a larger regional 3D groundwater model, are possible and computationally more affordable.

607 Finally, and importantly, the framework enables to assimilate additional hydrological and
608 geophysical datasets in order to provide additional constraints, such as groundwater heads (which were
609 not used in this work but are commonly available data in most real field studies) and complementary
610 geophysical techniques (electromagnetics, potential field, etc).

611 4.4 Limitations and Future Research

612 Previous authors (Hinnell et al., 2010; Camporese et al., 2015) have identified some key limitations
613 for the application of coupled hydrogeophysical inversion. Among them, the definition of the
614 conceptual model, the use of the appropriate petrophysical relationship and the computational effort
615 traditionally required for coupled hydrogeophysical simulations. In this work, for simplicity, we have
616 assumed that the dimensions of the domain, the boundary conditions, the structural model and the
617 petrophysical relationship were perfectly known. Future research is needed to address these limitations
618 or hurdles in the application of the FCI that we discuss briefly below. In addition, the impact of variably
619 saturated media near the surface in the geophysical measurements can have an important impact in the
620 results if disregarded (González-Quirós and Fernández-Álvarez, 2021). Variable saturation can be
621 included in the proposed methodology with some additional modelling and computational effort.

622 The conceptualization problem has gained considerable attention during the last decade (e.g.,
623 Enemark et al., 2019). Previous authors (Carrera et al., 2010) have shown that errors in the
624 conceptualization can render an inverse routine inadequate, even when using the most sophisticated
625 methodology available. Additionally, the multiGaussian model used in this work, even when easier to
626 implement has limitations to identify some type of structures (e.g., high permeability channels) that can
627 be key flow and contaminant transport paths (Gómez-Hernández and Wen, 1998) in many
628 environmental applications. Recent applications (e.g., Kang et al., 2019) have shown that coupled

hydrogeophysical inversion can be performed satisfactorily to characterize solute transport in non-Gaussian fields. Implementation in the routine presented in this work is under the scope of our research.

The groundwater community has been more prone for many years to evaluate parameter uncertainty, disregarding the structural uncertainty of the model (Refsgaard et al., 2012). In the hydrogeophysics modelling and inversion workflow, the use and application of the petrophysical model is a key element. However, petrophysical uncertainty has been widely ignored (Linde et al., 2017) even when it can lead to overconfident favourable impressions of the capability of geophysics for parameter estimation and for interpretation of results of the hydrogeological modelling (Brunetti et al., 2017). The evaluation and quantification of petrophysical uncertainty are increasingly receiving more attention in the literature (Brunetti and Linde, 2018; Tso et al., 2019; Mezquita-González et al., 2020). In many hydrogeophysical applications, computational burden can prevent a full Bayesian approach for uncertainty quantification, but recent strategies (e.g., Hermans et al., 2018, 2019; Kang et al., 2019; Tso et al., 2020 or this work) have shown that is possible to perform uncertainty analysis using an appropriate ensemble of hydrogeophysical models with an acceptable computational effort. Furthermore, the use of parallelization shown in this work reduces computational cost and, although not done in this example application, the workflow could be implemented in a computer cluster with the adequate COMSOL licensing.

Finally, it must be noted that the coupled multiphysical inversion with integration of COMSOL with PEST/PEST++ is not limited nor restricted to hydrogeophysical applications like the presented as an example in this study. The modelling flexibility of COMSOL together with the open-source model-independent capabilities of PEST/PEST++ allows researchers for imaginative approaches to solve a broad range of environmental problems at multiple scales (from lab-scale to regional scales) integrating not only geophysical methods, but also other types of observations, such as remote sensing, that are not traditionally fully incorporated quantitatively in the modelling workflow. The methodology is also not restricted to stationary models as the example shown here; transient modelling and time-lapse data can be incorporated into the workflow.

5 Conclusions

The full integration of geophysical data in the groundwater modelling workflow has been identified as a step forward to provide more reliable models used for decision support. Fully coupled hydrogeophysical inversion has been proposed as one of the solutions to improve the integration of geophysics in hydrogeological studies. However, its use in real-world settings has been limited because of the difficulty of simulating complex field conditions, the unavailability of flexible strategies and the computational limitations, which have prevented the widespread use by the hydrogeological community. In this work we have shown an efficient solution to perform fully coupled hydrogeophysical inversion with integration of PEST/PEST++ and COMSOL.

The integration of COMSOL and PEST/PEST++ using MATLAB provides a solution for model calibration and uncertainty analysis of multiphysical models. COMSOL is a well-known and powerful finite element software that provides a high degree of flexibility for the implementation of complex geometries and the coupled simulation of multiple physics. PEST has been the standard calibration software in the groundwater community for decades, is open-source and freely available, and in conjunction with PEST++ provide a powerful toolbox for calibration, sensitivity, and uncertainty analysis in environmental modelling. Finally, the methodological approach is not restricted to hydrogeophysical models and can be extended to multiphysical modelling and a broad number of environmental applications using the flexibility of COMSOL.

Software Availability

PEST (Doherty, 2020) is freely available in its dedicated webpage https://pesthomepage.org/.

PEST++ (White et al., 2020) is accessible and available in the US Geological Survey webpage https://www.usgs.gov/software/pest-parameter-estimation-code-optimized-large-environmental-models .

MATLAB is a commercial numerical computing platform (https://uk.mathworks.com/, or the respective webpage for each county). Version 2018a was used in this work.

COMSOL Multiphysics is a commercial finite element software (https://www.comsol.com/) COMSOL version 5.4b was used. *Livelink for Matlab* license is necessary to establish the connection between PEST/PEST++ and COMSOL. The *ACDC* and *Subsurface Flow* modules are required for the hydrogeophysical simulation presented in the example. A *Floating Network Licence* is necessary for parallelization.

Codes and examples used in this work are available in the A.G. Quiros' GitHub page https://github.com/AndresGQuiros.

Acknowledgments

Andrés González Quirós is the recipient of a Royal Society – Newton International Fellowship (NIF\R1\182210), hosted at the University of Aberdeen. We thank the Scottish Funding Council/Scottish Alliance for Geoscience, Environment and Society for seed funding the development of the project. We would also like to thank the Associate Editor Tim Green, Thomas Hermans, Landon Halloran and one anonymous reviewers for their comments and suggestions in the revision process.

Appendix A

In this work we used the Waxman and Smits (1968) petrophysical relationship, defined as

$$\sigma_{bulk} = \frac{1}{F}(\sigma_w + BQ_v) \quad \text{(A1)}$$

In this equation σ_w is the fluid conductivity [S m^{-1}] and F is the formation factor, ($F = 1/\phi^m$) with ϕ the porosity and m [-] the cementation exponent. B is the equivalent counterion mobility $B_0\left[1 - 0.6e^{(-\sigma_w/0.013)}\right]$, $B_0 = 4.78 \times 10^{-8}$the maximum counterion mobility. Finally, the excess of surface charge per unit pore volume, Q_v [meq·ml^{-1}], is calculated as (Revil et al., 1998)

$$Q_v = \rho_g\left(\frac{1-\phi}{\phi}\right)CEC \quad \text{(A2)}$$

699 Here, *CEC* [meq·g^{-1}] is the cation exchange capacity which value is proportional to clay content and
700 depends on the nature of the clay minerals (Revil et al., 1998). To compute the CEC we used the
701 expression

$$CEC = \varphi_w^{cl} \times CECi \tag{A3}$$

702 In this model we used a threshold to define a lower permeability value of clay-free sand using the
703 expression (Revil and Cathles, 1999):

$$k_{sd} = \frac{d_{sd}^2 (\phi_{sd})^{3m_{sd}}}{24} \tag{A4}$$

704 Beyond this value we considered that clay was present. In the equation above d_{sd}^2 is the grain
705 diameter of sand; ϕ_{sd} the porosity of sand and m_{sd} is the sand cementation exponent. We assigned
706 values of $d_{sd}^2 = 2 \times 10^{-4}$ m, $\phi_{sd} = 0.32$, $m_{sd} = 2$ and $m_{sd} = m_{cs} = 2$ (Power et al. 2013; Kang et
707 al., 2019). We applied the following equations to compute the clay content (Revil and Cathles, 1999)

$$\begin{cases} k < k_{sd} & Cl = \dfrac{k_{sd}^{\frac{1}{3m_{cs}}} - k^{\frac{1}{3m_{cs}}}}{k_{sd}^{\frac{1}{3m_{cs}}} \left(\frac{1-\phi_{cl}}{\phi_{sd}}\right)} \\ k > k_{sd} & Cl = 0 \end{cases} \tag{A5}$$

708 Total porosity was computed applying this expression (Berg, 1995)

$$\phi = \phi_{sd}(1 - Cl) + \phi_{cl}\, Cl \tag{A6}$$

709 Mass fractions of clay (used in equation A3 to compute CEC) and sand grains, φ_w^{cl} and φ_w^{sd} were
710 calculated using the equations (Power et al., 2013; Kang et al., 2019)

$$\varphi_w^{cl} = \frac{\mathrm{Cl}(1-\phi_{cl})\rho_{cl}}{\mathrm{Cl}(1-\phi_{cl})\rho_{cl} + (1-Cl)(1-\phi_{sd})\rho_{sd}} \tag{A7}$$

$$\varphi_w^{sd} = 1 - \varphi_w^{cl} \quad \text{(A8)}$$

In the study we assumed a constant density of mineral grains for sand and clay particles, $\rho_{cl} = \rho_{sd} = \rho_g = 2650$ kg·m^{-3} and cconstant temperature of 25 ºC.

The petrophysical relationship is fully integrated in COMSOL. The models are solved in every run with a direct spatial correspondence established between the hydraulic parameters (permeability and porosity), the hydrogeological variables (salinity), the petrophyical relationships (clay fraction) and the geophysical parameter (bulk resistivity).

Additionally, the following equation was used for conversion of fluid electrical conductivity into total dissolved solids (TDS) (Jiao and Post, 2019)

$$TDS = \mathrm{k}_c\, \sigma_w \quad (14)$$

We used a value of k_c=0.7 obtained by using the relationship $\mathrm{k}_c = TDS_{SW}/\sigma_{Sw}$ (Jiao and Post, 2019) where $TDS_{SW} = 35000$ mg l^{-1} (35 kg·m^{-3}) and $\sigma_{Sw} = 50000$ µS·cm^{-1} (5 S·m^{-1}) are the total dissolved solids and electric conductivity of saltwater.

References

Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2015). Applied groundwater modeling: simulation of flow and advective transport. Academic press, 630 p. https://doi.org/10.1016/C2009-0-21563-7

Archie, GE (1942), The electrical resistivity log as an aid in determining some reservoir characteristics, Trans. Am. Inst. Min. Metall. Pet. Eng., **146**(1), 54–62, https://doi.org/10.2118/942054-G.

Beaujean, J., Nguyen, F., Kemna, A., Antonsson, A., & Engesgaard, P. (2014). Calibration of seawater intrusion models: Inverse parameter estimation using surface electrical resistivity tomography and borehole data. Water Resources Research, 50(8), 6828-6849. https://doi.org/10.1002/2013WR014020

Berg, C. R., (1995). A simple effective-medium model for water saturation in porous rocks: Geophysics, **60**, 1070–1080, https://doi.org/10.1190/1.1443835

Binley, A., & Kemna, A. (2005). DC resistivity and induced polarization methods. In Hydrogeophysics (pp. 129-156). Springer, Dordrecht.

Binley, A., Hubbard, S. S., Huisman, J. A., Revil, A., Robinson, D. A., Singha, K., & Slater, L. D. (2015). The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. Water Resources Research, **51**(6), 3837-3866. https://doi.org/10.1002/2015WR017016

Bouchedda, A., Giroux, B., & Gloaguen, E. (2017). Constrained ERT Bayesian inversion using inverse Matérn covariance matrix, 52. https://doi.org/10.1190/geo2015-0673.1

Brunetti, C., Linde, N., & Vrugt, J. A. (2017). Bayesian model selection in hydrogeophysics: Application to conceptual subsurface models of the South Oyster Bacterial Transport Site, Virginia, USA. Advances in Water Resources, **102**, 127-141. https://doi.org/10.1016/j.advwatres.2017.02.006

Brunetti, C., & Linde, N. (2018). Impact of petrophysical uncertainty on Bayesian hydrogeophysical inversion and model selection. Advances in Water Resources, **111**, 346-359. https://doi.org/10.1016/j.advwatres.2017.11.028

Buchanan, S., & Triantafilis, J. (2009). Mapping water table depth using geophysical and environmental variables. Groundwater, 47(1), 80-96. https://doi.org/10.1111/j.1745-6584.2008.00490.x

Camporese, M., Cassiani, G., Deiana, R., Salandin, P., & Binley, A. (2015). Coupled and uncoupled hydrogeophysical inversions using ensemble Kalman filter assimilation of ERT-monitored tracer test data. Water Resources Research, **51**(5), 3277-3291. https://doi.org/10.1002/2014WR016017

Carrera, J., Hidalgo, J. J., Slooten, L. J., & Vázquez-Suñé, E. (2010). Computational and conceptual issues in the calibration of seawater intrusion models. Hydrogeology Journal, **18**(1), 131-145. https://doi.org/10.1007/s10040-009-0524-1

Chambers, J. E., Gunn, D. A., Wilkinson, P. B., Meldrum, P. I., Haslam, E., Holyoake, S., Kirkham, M.,Kuras, O., Merritt, A. & Wragg, J. (2014). 4D electrical resistivity tomography monitoring of soil moisture dynamics in an operational railway embankment. Near Surface Geophysics, **12**(1), 61-72. https://doi.org/10.3997/1873-0604.2013002

Chen, Y. & Oliver, D. S. (2013). Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history-matching and uncertainty quantification. Computational Geosciences. **17**(4), 689–703. https://doi.org/10.1007/s10596-013-9351-5

Cockett, R., Kang, S., Heagy, L. J., Pidlisecky, A., & Oldenburg, D. W. (2015). SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications. Computers & Geosciences, **85**, 142-154. https://doi.org/10.1016/j.cageo.2015.09.015

Comte, J. C., & Banton, O. (2007). Cross-validation of geo-electrical and hydrogeological models to evaluate seawater intrusion in coastal aquifers. Geophysical Research Letters, **34**(10). https://doi.org/10.1029/2007GL029981

Comte, J. C., Wilson, C., Ofterdinger, U. and González-Quirós, A. (2017). Effect of volcanic dykes on coastal groundwater flow and saltwater intrusion: A field-scale multiphysics approach and parameter evaluation. *Water Resources Research*, **53**(3), 2171-2198. https://doi.org/10.1002/2016WR019480

Constable, S. C., Parker, R. L., & Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. Geophysics, **52**(3), 289-300. https://doi.org/10.1190/1.1442303

Costall, A., Harris, B., & Pigois, J. P. (2018). Electrical resistivity imaging and the saline water interface in high-quality coastal aquifers. Surveys in Geophysics, **39(**4), 753-816. https://doi.org/10.1007/s10712-018-9468-0

781 Costall, A. R., Harris, B. D., Teo, B., Schaa, R., Wagner, F. M., & Pigois, J. P. (2020). Groundwater
782 throughflow and seawater intrusion in high quality coastal aquifers. Scientific reports, 10(1), 1-
783 33. https://doi.org/10.1038/s41598-020-66516-6

784 Day-Lewis, F. D., K. Singha, and A. M. Binley (2005), The application of petrophysical models to radar
785 and electrical resistivity tomograms: Resolution dependent limitations, J. Geophys. Res., **110**,
786 B08206, doi:10.1029/2004JB003569

787 de Franco, R., Biella, G., Tosi, L., Teatini, P., Lozej, A., Chiozzotto, B, Giada, M. , Rizzetto, F. , Claude,
788 C. , Mayer, A. , Bassan, V. & Gasparetto-Stori , G. (2009). Monitoring the saltwater intrusion by
789 time lapse electrical resistivity tomography: The Chioggia test site (Venice Lagoon, Italy).
790 Journal of Applied Geophysics, **69**(3-4), 117-130. https://doi.org/10.1016/j.jappgeo.2009.08.004

791 de Marsily, G., Lavedan, C., Boucher, M., and Fasanino, G. (1984). Interpretation of interference tests
792 in a well field using geostatistical techniques to fit the permeability distribution in a reservoir
793 model, in Verly, G., David, M., Journel, A.G., and Marechal, A., Geostatistics for natural
794 resources characterization: NATO Advanced Study Institute, ser. C 182, p. 831–849.

795 Diersch, H. J. G. (2013). FEFLOW: finite element modeling of flow, mass and heat transport in porous
796 and fractured media. Springer Science & Business Media.

797 Doherty, J. (2015). Calibration and uncertainty analysis for complex environmental models. Watermark
798 Numerical Comp

799 Doherty, J. (2020). PEST, Model-independent Parameter Estimation, 7th edition. User Manual,
800 Watermark Numerical Computing.

801 Doherty, J. E., Fienen, M. N., & Hunt, R. J. (2010). Approaches to highly parameterized inversion:
802 Pilot-point theory, guidelines, and research directions. US Geological Survey scientific
803 investigations report, 5168, 36.

804 Doherty, J., & Moore, C. (2020). Decision Support Modeling: Data Assimilation, Uncertainty
805 Quantification, and Strategic Abstraction. Groundwater, **58**(3), 327-337.
806 https://doi.org/10.1111/gwat.12969

807 Doherty, J., & Simmons, C. T. (2013). Groundwater modelling in decision support: reflections on a
808 unified conceptual framework. Hydrogeology Journal, **21**(7), 1531-1537.
809 https://doi.org/10.1007/s10040-013-1027-7

810 Enemark, T., Peeters, L. J., Mallants, D., & Batelaan, O. (2019). Hydrogeological conceptual model
811 building and testing: A review. Journal of Hydrology, **569**, 310-329.
812 https://doi.org/10.1016/j.jhydrol.2018.12.007

813 Enemark, T., Peeters, L., Mallants, D., Flinchum, B., & Batelaan, O. (2020). A systematic approach to
814 hydrogeological conceptual model testing, combining remote sensing and geophysical data.
815 Water Resources Research, **56**(8), e2020WR027578. https://doi.org/10.1029/2020WR027578

816 Ferré, T. P. (2017). Revisiting the relationship between data, models, and decision-making.
817 Groundwater, **55**(5), 604-614. https://doi.org/10.1111/gwat.12574

818 Gasperikova, E., Hubbard, S. S., Watson, D. B., Baker, G. S., Peterson, J. E., Kowalsky, M. B., Smith,
819 M. & Brooks, S. (2012). Long-term electrical resistivity monitoring of recharge-induced
820 contaminant plume behavior. Journal of Contaminant Hydrology, **142-143**, 33-49.
821 https://doi.org/10.1016/j.jconhyd.2012.09.007

822 Goebel, M., Pidlisecky, A., & Knight, R. (2017). Resistivity imaging reveals complex pattern of
823 saltwater intrusion along Monterey coast. Journal of Hydrology, **551**, 746-755.
824 https://doi.org/10.1016/j.jhydrol.2017.02.037

825 Gómez-Hernández, J. J., & Journel, A. G. (1993). Joint Sequential Simulation of MultiGaussian Fields.
826 In: Soares A. (eds) Geostatistics Tróia '92. Quantitative Geology and Geostatistics, 5. Springer,
827 Dordrecht.

828 Gómez-Hernández, J. J., & Wen, X. H. (1998). To be or not to be multi-Gaussian? A reflection on
829 stochastic hydrogeology. Advances in Water Resources, **21**(1), 47-61.
830 https://doi.org/10.1016/S0309-1708(96)00031-0

González-Quirós, A., & Fernández-Álvarez, J.P. (2014) Simultaneous Solving of Three-Dimensional Gravity Anomalies Caused by Pumping Tests in Unconfined Aquifers. Mathematical Geosciences, **46**, 649–664. https://doi.org/10.1007/s11004-014-9539-9

González-Quiros, A., Comte, J., Rubio-Melendi, D., & Fernandez-Alvarez, J. (2019). Towards fully coupled finite element modelling of DC resistivity in complex seawater intrusion scenarios. In 81st EAGE Conference and Exhibition 2019, London, UK. https://doi.org/10.3997/2214-4609.201900972

González-Quirós, A., & Comte, J. C. (2020). Relative importance of conceptual and computational errors when delineating saltwater intrusion from resistivity inverse models in heterogeneous coastal aquifers. Advances in Water Resources, **144**, 103695. https://doi.org/10.1016/j.advwatres.2020.103695

González-Quirós, A., & Fernández-Álvarez, J. P. (2021). Use of microgravity for identification of delayed gravity drainage and conceptual model selection in unconfined aquifers. Journal of Hydrology, 597, 126285. https://doi.org/10.1016/j.jhydrol.2021.126285

Günther, T. and Rücker, C. (2015). Boundless Electrical Resistivity Tomography BERT 2–user tutorial.

Halloran, L. J., Brunner, P., & Hunkeler, D. (2019). COMPEST, a PEST-COMSOL interface for inverse multiphysics modelling: Development and application to isotopic fractionation of groundwater contaminants. Computers & Geosciences, **126**, 107-119. https://doi.org/10.1016/j.cageo.2019.02.001

Halloran, L. J., Vakili, F., Wanner, P., Shouakar-Stash, O., & Hunkeler, D. (2021). Sorption-and diffusion-induced isotopic fractionation in chloroethenes. Science of The Total Environment, 788, 147826. https://doi.org/10.1016/j.scitotenv.2021.147826

Harbaugh, A. W. (2005). MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process (pp. 6-A16). Reston, VA: US Department of the Interior, US Geological Survey.

856 Henry, H. R. (1964). Effects of dispersion on salt encroachment in coastal aquifers, in" Seawater in
857 Coastal Aquifers". US Geological Survey, Water Supply Paper, **1613**, C70-C80.

858 Herckenrath, D., Langevin, C. D., & Doherty, J. (2011). Predictive uncertainty analysis of a saltwater
859 intrusion model using null-space Monte Carlo. Water Resources Research, **47**(5), W05504.
860 https://doi.org/10.1029/2010WR009342

861 Herckenrath, D., Fiandaca, G., Auken, E., & Bauer-Gottwein, P. (2013). Sequential and joint
862 hydrogeophysical inversion using a field-scale groundwater model with ERT and TDEM data.
863 Hydrology and Earth System Sciences, **17**(10), 4043-4060. https://doi.org/10.5194/hess-17-
864 4043-2013

865 Hermans, T., Vandenbohede, A., Lebbe, L., Martin, R., Kemna, A., Beaujean, J., & Nguyen, F. (2012).
866 Imaging artificial salt water infiltration using electrical resistivity tomography constrained by
867 geostatistical data. Journal of Hydrology, 438-439, 168-180.
868 https://doi.org/10.1016/j.jhydrol.2012.03.021

869 Hermans, T., & Irving, J. (2017). Facies discrimination with ERT using a probabilistic methodology:
870 effect of sensitivity and regularization. Near Surface Geophysics, 15, 13-25.
871 https://doi.org/10.3997/1873-0604.2016047

872 Hermans, T., Nguyen, F., Klepikova, M., Dassargues, A., & Caers, J. (2018). Uncertainty quantification
873 of medium-term heat storage from short-term geophysical experiments using Bayesian evidential
874 learning. Water Resources Research, 54(4), 2931-2948. https://doi.org/10.1002/2017WR022135

875 Hermans, T., Lesparre, N., De Schepper, G., & Robert, T. (2019). Bayesian evidential learning: a field
876 validation using push-pull tests. Hydrogeology Journal, 27(5), 1661-1672.
877 https://doi.org/10.1007/s10040-019-01962-9

878 Hill, M.C. and Tiedeman, C.R. (2007). Effective groundwater model calibration: With analysis of data,
879 sensitivities, predictions, and uncertainty. Wiley and Sons, New York, New York, 455 p.

880 Hinnell, A. C., Ferré, T. P. A., Vrugt, J. A., Huisman, J. A., Moysey, S., Rings, J. and Kowalsky, M. B.
881 (2010). Improved extraction of hydrologic information from geophysical data through coupled

882 hydrogeophysical inversion. Water Resources Research, **46**(4), W00D40.
883 https://doi.org/10.1029/2008WR007060

884 Irving, J., & Singha, K. (2010). Stochastic inversion of tracer test and electrical geophysical data to
885 estimate hydraulic conductivities. Water Resources Research, 46(11).
886 https://doi.org/10.1029/2009WR008340

887 Jiao, J., & Post, V. (2019). Coastal Hydrogeology. Cambridge University Press.
888 https://doi.org/10.1017/9781139344142

889 Kang, X., Shi, X., Revil, A., Cao, Z., Li, L., Lan, T., & Wu, J. (2019). Coupled hydrogeophysical
890 inversion to identify non-Gaussian hydraulic conductivity field by jointly assimilating
891 geochemical and time-lapse geophysical data. Journal of Hydrology, **578**, 124092.
892 https://doi.org/10.1016/j.jhydrol.2019.124092

893 Kirsch, R. (2006). Groundwater geophysics: A Tool for Hydrogeology. Berlin: Springer.

894 Klotzsche, A., Jonard, F., Looms, M. C., van der Kruk, J., & Huisman, J. A. (2018). Measuring soil
895 water content with ground penetrating radar: A decade of progress. Vadose Zone Journal, **17**(1),
896 1-9. https://doi.org/10.2136/vzj2018.03.0052

897 Langevin, C. D., Thorne Jr, D. T., Dausman, A. M., Sukop, M. C., & Guo, W. (2008). SEAWAT version
898 4: a computer program for simulation of multi-species solute and heat transport (No. 6-A22).
899 Geological Survey (US).

900 Lebbe, L. (1999). Parameter identification in fresh-saltwater flow based on borehole resistivities and
901 freshwater head data. Advances in Water Resources, 22(8), 791-806.
902 https://doi.org/10.1016/S0309-1708(98)00054-2

903 Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares.
904 Quarterly of Applied Mathematics, **2**(2), 164-168.

905 Linde, N. (2014). Falsification and corroboration of conceptual hydrological models using geophysical
906 data. Wiley Interdisciplinary Reviews: Water, **1**(2), 151-171. https://doi.org/10.1002/wat2.1011

907 Linde, N., Renard, P., Mukerji, T. and Caers, J. (2015). Geological realism in hydrogeological and
908 geophysical inverse modeling: A review. Advances in Water Resources, **86**, 86-101.
909 https://doi.org/10.1016/j.advwatres.2015.09.019

910 Linde, N., & Doetsch, J. (2016). Joint inversion in hydrogeophysics and near-surface geophysics. In:
911 Integrated imaging of the Earth: Theory and applications, 218, 119.
912 https://doi.org/10.1002/9781118929063.ch7

913 Linde, N., Ginsbourger, D., Irving, J., Nobile, F., & Doucet, A. (2017). On uncertainty quantification
914 in hydrogeology and hydrogeophysics. Advances in Water Resources, **110**, 166-181.
915 https://doi.org/10.1016/j.advwatres.2017.10.014

916 Loke, M.H. (2018) Rapid 2-D resistivity and IP inversion using the least-squares method.
917 www.geotomosoft.com.

918 Lopez-Alvis, J., Hermans, T., & Nguyen, F. (2019). A cross-validation framework to extract data
919 features for reducing structural uncertainty in subsurface heterogeneity. Advances in Water
920 Resources, 133, 103427. https://doi.org/10.1016/j.advwatres.2019.103427

921 Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. Journal
922 of the Society for Industrial and Applied Mathematics, **11**(2), 431-441.

923 Mastrocicco, M., Vignoli, G., Colombani, N., & Zeid, N. A. (2010). Surface electrical resistivity
924 tomography and hydrogeological characterization to constrain groundwater flow modeling in an
925 agricultural field site near Ferrara (Italy). Environmental Earth Sciences, **61**(2), 311-322.
926 https://doi.org/10.1007/s12665-009-0344-6

927 Mezquita-González, J. A., Comte, J. C., Legchenko, A., Ofterdinger, U., & Healy, D. (2021).
928 Quantification of groundwater storage heterogeneity in weathered/fractured basement rock
929 aquifers using electrical resistivity tomography: sensitivity and uncertainty associated with
930 petrophysical modelling. Journal of Hydrology, **593**, 125637.
931 https://doi.org/10.1016/j.jhydrol.2020.125637

Palacios, A., Ledo J.J., Linde N., Luquot, L., Bellmunt, F., Folch, A., Marcuello,A., Pilar Queralt, P., Pezard, P.A., Martínez, L., del Val, L., Bosch, D. & Carrera, J (2020). Time-lapse cross-hole electrical resistivity tomography (CHERT) for monitoring seawater intrusion dynamics in a Mediterranean aquifer. Hydrology and Earth System Sciences. **24**(4), 2121-2139. https://doi.org/10.5194/hess-24-2121-2020

Pollock, D., & Cirpka, O. A. (2012). Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by electrical resistivity tomography. Water Resources Research, **48**(1), W01505. https://doi.org/10.1029/2011WR010779

Power, C., Gerhard, J.I., Tsourlos, P., Giannopoulos, A., (2013). A new coupled model for simulating the mapping of dense nonaqueous phase liquids using electrical resistivity tomography. Geophysics, **78**, 1–15. https://doi.org/10.1190/geo2012-0395.1.

Refsgaard, J. C., Christensen, S., Sonnenborg, T. O., Seifert, D., Højberg, A. L., & Troldborg, L. (2012). Review of strategies for handling geological uncertainty in groundwater flow and transport modeling. Advances in Water Resources, **36**, 36-50. https://doi.org/10.1016/j.advwatres.2011.04.006

Revil, A., L. M. Cathles III, S. Losh, and J. A. Nunn (1998), Electrical conductivity in shaly sands with geophysical applications, J. Geophys. Res., **103**(B10), 23,925–23,936, doi:10.1029/98JB02125.

Revil, A., Cathles III, L.M., 1999. Permeability of shaly sands. Water Resour. Res. 35 (3), 651–662. https://doi.org/10.1029/98WR02700

Revil, A., Coperey, A., Shao, Z., Florsch, N., Fabricius, I. L., Deng, Y., Delsman, J.R., Pauw, P.S., Karaoulis, M., de Louw, P.G.B., van Baaren, E.S., Dabekaussen, W., Menkovic, A. & Gunnink, J.L. (2017). Complex conductivity of soils. Water Resources Research, **53**(8), 7121-7147. https://doi.org/10.1002/2017WR020655

Rücker, C., Günther, T., & Spitzer, K. (2006). Three-dimensional modelling and inversion of dc resistivity data incorporating topography—I. Modelling. Geophysical Journal International, 166(2), 495-505. https://doi.org/10.1111/j.1365-246X.2006.03010.x

Rücker, C., Günther, T., & Wagner, F. M. (2017). pyGIMLi: An open-source library for modelling and inversion in geophysics. Computers & Geosciences, 109, 106-123. https://doi.org/10.1016/j.cageo.2017.07.011

Scheidt, C., Li, L., & Caers, J. (Eds.). (2018). Quantifying uncertainty in subsurface systems (Vol. 236). John Wiley & Sons. https://doi.org/10.1002/9781119325888

Schreuder, W. A. (2009). Running BeoPEST. In Proceedings of the 1st PEST Conference.

Singha, K., & Moysey, S. (2006). Accounting for spatially variable resolution in electrical resistivity tomography through field-scale rock-physics relations. Geophysics, 71(4), A25-A28. https://doi.org/10.1190/1.2209753

Slater, L. (2007). Near Surface Electrical Characterization of Hydraulic Conductivity: From Petrophysical Properties to Aquifer Geometries—A Review. Surveys in Geophysics, 28(2-3), 169-197. https://doi.org/10.1007/s10712-007-9022-y

Steklova, K., & Haber, E. (2017). Joint hydrogeophysical inversion: state estimation for seawater intrusion models in 3D. Computational Geosciences, **21**(1), 75-94. https://doi.org/10.1007/s10596-016-9595-y

Thibaut, R., Kremer, T., Royen, A., Kim Ngun, B., Nguyen, F., & Hermans, T. (2021). A new workflow to incorporate prior information in minimum gradient support (MGS) inversion of electrical resistivity and induced polarization data. Journal of Applied Geophysics, 187, 104286. https://doi.org/10.1016/j.jappgeo.2021.104286

Tikhonov, A. N., & Arsenin, V. Y. (1977). Solutions of ill-posed problems. New York, Halsted, 258 p.

Tonkin, M., & Doherty, J. (2009). Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. Water Resources Research, **45**(12). https://doi.org/10.1029/2007WR006678

Tso, C. H. M., Kuras, O., & Binley, A. (2019). On the field estimation of moisture content using electrical geophysics: the impact of petrophysical model uncertainty. Water Resources Research, **55**(8), 7196-7211. https://doi.org/10.1029/2019WR024964

Tso, C. H. M., Johnson, T. C., Song, X., Chen, X., Kuras, O., Wilkinson, P., Uhleman, S., Chambers, J. & Binley, A. (2020). Integrated hydrogeophysical modelling and data assimilation for geoelectrical leak detection. Journal of Contaminant Hydrology, **234**, 103679. https://doi.org/10.1016/j.jconhyd.2020.103679

United States Environmental Protection Agency (USEPA). (2009). National recommended water quality criteria. Office of Water, Office of Science and Technology (4304T), 22.

Voss, C. I., & Provost, A.M., (2002), SUTRA, A model for saturated-unsaturated variable-density ground-water flow with solute or energy transport, U.S. Geological Survey Water-Resources Investigations Report 02-4231, 291 p.

Waxman, M. H., & Smits, L. J. M. (1968). Electrical conductivities in oil-bearing shaly sands. Society of Petroleum Engineers Journal, **8**(2), 107-122. https://doi.org/10.2118/1863-A

Welter, D., Doherty, J. and Egan, C., (2019). Manual for the PANTHER Parallel Run Manager. Watermark Numerical Computing,

White, J. T. (2018). A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. Environmental Modelling & Software, **109**, 191-201. https://doi.org/10.1016/j.envsoft.2018.06.009

White, J. T., Hunt, R. J., Doherty, J. E., and Fienen, M. N. (2020). PEST++ version 5, a parameter estimation and uncertainty analysis software suite optimized for large environmental models. U.S. Geological Survey Techniques and Methods Report.

Zheng, C., & Gorelick, S. M. (2003). Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale. Ground Water, **41**(2), 142-155. https://doi.org/10.1111/j.1745-6584.2003.tb02578.x

Zhou, H., Gómez-Hernández, J. J., & Li, L. (2014). Inverse methods in hydrogeology: Evolution and recent trends. Advances in Water Resources, 63, 22-37. https://doi.org/10.1016/j.advwatres.2013.10.014