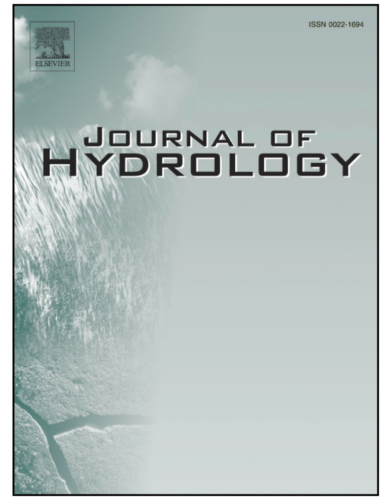
Journal Pre-proofs

Research papers



A framework for parameter estimation using sharp-interface seawater intrusion models

Cécile Coulon, Alexandre Pryet, Jean-Michel Lemieux, Ble Jean Fidele Yrro, Abderrezak Bouchedda, Erwan Gloaguen, Jean-Christophe Comte, J. Christian Dupuis, Olivier Banton

PII: S0022-1694(21)00556-4
DOI: https://doi.org/10.1016/j.jhydrol.2021.126509
Reference: HYDROL 126509

To appear in: *Journal of Hydrology*

Received Date: 17 February 2021
Revised Date: 20 April 2021
Accepted Date: 27 May 2021

A framework for parameter estimation using sharp-interface seawater intrusion models

Cécile Coulon[a,b,c], Alexandre Pryet[d], Jean-Michel Lemieux[a,b,c], Ble Jean Fidele Yrro[e], Abderrezak Bouchedda[e], Erwan Gloaguen[e], Jean-Christophe Comte[f], J. Christian Dupuis[a], Olivier Banton[g]

[a] Département de géologie et de génie géologique, Université Laval, 1065 avenue de la Médecine, Québec (Québec) G1V 0A6, Canada

[b] Centre québécois de recherche sur l'eau, 1065 avenue de la Médecine, Québec (Québec) G1V 0A6, Canada

[c] Centre d'études nordiques, 2405 rue de la Terrasse, Université Laval, Québec (Québec) G1V 0A6, Canada

[d] EA 4592 Géoressources et Environnement, Bordeaux INP and Univ. Bordeaux Montaigne, ENSEGID, 1 allée F. Daguin, 33607 Pessac, France

[e] Institut national de la recherche scientifique, Centre Eau Terre Environnement (INRS-ETE), Québec, Canada

[f] School of Geosciences, University of Aberdeen, Aberdeen, United Kingdom

[g] UMR 1114 EMMAH Avignon Université – INRAE, Avignon, France

Corresponding author: Cécile Coulon, cecile.coulon.1@ulaval.ca

Email addresses:

- alexandre.pryet@ensegid.fr

- jean-michel.lemieux@ggl.ulaval.ca
- ble_jean_fidele.yrro@ete.inrs.ca
- abderrezak.bouchedda@ete.inrs.ca
- erwan.gloaguen@ete.inrs.ca
- jc.comte@abdn.ac.uk
- christian.dupuis@ggl.ulaval.ca

Keywords

Seawater Intrusion, Sharp Interface, Parameter Estimation, Uncertainty Analysis, Data Worth, Geophysics

1 Introduction

In coastal areas and islands, seawater intrusion is a major challenge for groundwater management. Numerical models of seawater intrusion are heavily relied on for groundwater management decision-support (Werner et al., 2013), as they can help quantify current and future freshwater resources and support the design of optimal pumping scenarios.

Seawater intrusion models are generally based on variable density codes, which simulate mixing between fresh and saline groundwater. As these codes solve the coupled, non-linear groundwater flow and advective-dispersive solute transport equations, they require a fine vertical discretization of the simulated domain and the resulting models are computationally expensive. Excessive model run times have severely limited the possibility of parameter estimation, i.e. "*automated trial-and-error calibration"* (Anderson et al., 2015), especially for large-scale regional models. While parameter estimation is routinely performed for hydrogeological inversions, it has remained scarce for seawater intrusion models (Carrera et al., 2010; Werner et al., 2013). Although several studies have recently carried out parameter estimation with variable density models, with methods to reduce model run times (e.g. Ataie-Ashtiani et al., 2013; Dentoni et al., 2014) or using mixed manual-automated calibration strategies (e.g. Meyer et al., 2019; Siarkos and Latinopoulos, 2016), manual trial-and-error calibration is still often applied (e.g. Holding and Allen, 2015; Post et al., 2018c). However, manual calibration has shortcomings which can be crucial for decision-support models. While manual calibration can be subjective and does not necessarily lead to the optimum parameter set, regularized parameter estimation can lead to the minimum error variance parameter set, which allows for predictions of minimum error variance (Anderson et al., 2015). In addition, parameter estimation allows for quantitative uncertainty analysis, which is critical for model-based decision-making (Delottier et al., 2016; Hunt et al., 2020).

Sharp-interface codes, like SHARP (Essaid, 1990) or the SWI2 package for MODFLOW-2005 (Bakker et al., 2013), neglect mixing processes and simulate a sharp boundary (or interface) between freshwater and saltwater. Sharp-interface codes are well adapted to regional seawater intrusion modeling (Reilly and Goodman, 1985) and their range of applicability has been explored by Llopis-Albert and Pulido-Velazquez (2014). As they do not solve the solute transport equation, run times are significantly shorter. In a synthetic case, using the sharp-interface code SWI2 instead of the variable density code SEAWAT (Langevin et al., 2008) reduced run times from three hours to a few seconds (Dausman et al., 2010b). The fast run times afforded by sharp-interface codes have made these practical for coastal pumping optimization (Dhar and Datta, 2009; Kopsiaftis et al., 2019) and pave the way for parameter estimation. However, even with sharp-interface models, parameter estimation remains far from common practice. Manual calibration of such models is still widely used (e.g. Babu et al., 2018; Dokou and Karatzas, 2012; Gingerich, 2002; Pappa et al., 2017) and on occasion, the SWI2 package has been implemented in a previously calibrated MODFLOW groundwater model without further calibration (e.g. Baalousha, 2016; Walter et al., 2016). Only very rarely has parameter estimation been carried out (Hughes and White, 2014; Rotzoll et al., 2016), such that guidelines and case studies are lacking which could otherwise be used to help perform this task. Quantification of the predictive uncertainty of real-world seawater intrusion models also remains scarce (Werner et al., 2013).

One of the major knowledge gaps for the parameter estimation of sharp-interface models concerns the type of observations to be included. Currently, apart from the works of Hughes and White (2014), who used both head and flow rate observations, or Babu et al. (2018), who derived the thickness of the freshwater lens from nested monitoring wells, most sharp-interface model calibrations have been constrained by groundwater levels alone. However, it is known that using head observations alone is insufficient to constrain the inversion problem uniquely

(Anderson et al., 2015), and salinity observations are commonly used in variable density models (Shoemaker, 2004). Literature on data assimilation for sharp-interface models is therefore limited, and it is unclear which observations should be used, what processing is required and what weighting strategy should be used to account for contrasting measurement errors. Few resources are available to guide both data assimilation and parameter estimation in sharp-interface models and considering that these are crucial for decision-support modeling (Doherty and Moore, 2020), more investigations are warranted in this area. Exploring the links between data and models was also identified by Werner et al. (2013) as a key area of research for seawater intrusion modeling.

The objective of this paper is to present a framework for parameter estimation using a regional sharp-interface seawater intrusion model. A model was developed using the SWI2 package for MODFLOW-2005, which has shown efficient run times (Dausman et al., 2010b). A diverse dataset was assembled using typical coastal aquifer observations. Groundwater head observations were extracted from shallow wells, deep open wells and pumping wells. Observations of the freshwater-seawater interface, further referred to as interface observations, were extracted from deep open wells, and from time-domain electromagnetic (TDEM) and electrical resistivity tomography (ERT) surveys. The uncertainty of these 6 observation groups was quantified and accounted for in parameter estimation through the weighting strategy, as recommended for PEST (Doherty, 2004). Linear predictive uncertainty analysis was conducted for 2 types of forecasts: the total freshwater volume and interface elevations at pumping wells. An examination of residuals (i.e. the difference between simulated and observed values) then provided insight on the capacity of all 6 observation types to constrain calibration, while a data worth analysis explored their value for reducing predictive uncertainties. This modeling and inversion framework was developed for a real-world case in the Magdalen Islands (Quebec, Canada), and the findings of the study can help guide data collection efforts in other coastal aquifers where decision-support models are needed.

2 Study area

The Magdalen Islands (Quebec, Canada) form an archipelago located in the middle of the Gulf of Saint-Lawrence (Fig. 1), with an area of 200 km^2 and approximately 12,000 permanent inhabitants (Statistics Canada, 2017). Due to the lack of surface water resources and the prohibitive cost of seawater desalinization (Chaillou et al., 2012), groundwater is the only drinking water source. In addition, a high water demand during the summer enhances the risk of saltwater upconing under pumping wells. Local decision-makers are therefore strongly concerned by the capacity and management of their groundwater resources (BAPE, 2013). Grande Entrée Island, lying within the archipelago, was considered for this study as it is one of the most vulnerable islands due its shallow freshwater-seawater interface.

2.1 Conceptual model

Grande Entrée Island has an area of 8.5 km^2 and is surrounded by water from both the Gulf of Saint-Lawrence and lagoons. It has a relatively flat topography, with land elevations reaching at most 43 m above local mean sea level and a gently sloping seafloor (between 5 and 15‰ up to 1 km from the island). The nearby weather station indicated an average precipitation rate of 1040 mm/yr between 1981 and 2010, with relatively uniform rates throughout the year. Past studies have estimated the potential evapotranspiration as approximately 500 mm/yr (Dessureault and Simard, 1970) and recharge as 25% to 40% of total precipitation, i.e. 230 to 380 mm/yr (Leblanc, 1994, Poulin, 1977, Sylvestre, 1979b).

The main geological unit, both onshore and offshore (Fig. 1), is a red Permian eolian sandstone with large cross-bedding features (Brisebois, 1981; Rabeau and Thériault, 2013) belonging to the

Étang-des-Caps Member (Cap-aux-Meules formation) and estimated to be about 300 meters thick (Brisebois, 1981). Onshore, Quaternary unconsolidated sediments overlie the Permian sandstone: glacial sediments fill a paleovalley in the middle of the island (fine sand with traces of silt and gravel) and sand dunes lie on the outskirts. While the paleovalley reaches a thickness of approximately 110 meters at the center of the island, the thickness of the sand dunes is not well known (a thickness of 10 to 15 m is observed to the west).

The red Permian sandstone is the main aquifer formation, intercepted by all nine municipal wells and by most industrial and domestic wells. A number of aquifer tests in the archipelago have shown this formation to be heterogeneous, with a high hydraulic conductivity ($4 \cdot 10^{-5}$ m/s on average). The hydraulic conductivity of the other geological formations is less well known. The sand dunes are considered highly permeable (Sylvestre, 1979a) whereas the sparse aquifer tests in the glacial sediments yield a lower average hydraulic conductivity of $1 \cdot 10^{-5}$ m/s. Municipal pumping started in 2013 and freshwater is distributed to households, institutions and industries. Using data on municipal water use and individual consumption estimates, it was determined that an additional 80 m^3/day is being pumped by domestic wells. Industrial groundwater pumping was neglected in the study. Before installation of the municipal wells, groundwater pumping was mostly uncontrolled, leading to several episodes of saltwater well contamination which are not well documented.

2.2 Monitoring

Few historic observations are available on the Magdalen Islands, whether for head, salinity or pumping rates. Among the available data, automated meters have recorded pumping rates, water levels and electrical conductivity at all municipal wells, every minute, since mid-2014. Data gaps in the records are frequent because of technical issues. At five deep, open monitoring wells (Fig. 1), loggers have recorded pressure on an hourly basis since mid-2016 and downhole electrical conductivity and temperature profiles are carried out once to twice a year using a graduated water conductivity meter. Few head and salinity observations are available outside this monitoring network. Additional head observations were collected as part of the study (from manual measurements and pressure transducers), and electrical conductivity profiles were obtained from two other deep, open wells.

The results of two geophysical surveys were also used: an ERT campaign from 2004, which delineated the glacial palaeovalley and mapped the shape of the interface along nine transects perpendicular to the coast (Madelin'Eau, 2004), and a TDEM campaign carried out in 2019. The location of all observations used for the study is shown Fig. 1. The electrical conductivity measurements from the deep wells show that the transition zone from fresh to saline groundwater is relatively narrow, between 5 and 15 m wide. It is also shallow (on average 45 m below local mean sea level), suggesting a freshwater lens which does not intersect the bottom of the aquifer formation (Fig. 2). Inter- and intra-annual fluctuations of the transition zone are limited.

Fig. 1 Simplified geological map of the Grande Entrée Island and locations of head and freshwater-seawater interface observations. All pumping wells are drilled into the red Permian sandstone. Multiple observations are available, including interface observations derived from TDEM (time-domain electromagnetics) and ERT (electrical resistivity tomography) surveys.

Fig. 2 Conceptual model: schematic cross-section perpendicular to the island and example of a downhole electrical conductivity (EC) profile in one of the island's deep, open wells. In the freshwater lens, the relatively narrow transition zone from freshwater (FW) to saltwater (SW) is approximated as a sharp freshwater-seawater interface. Elevation is expressed in meters above local sea level (masl).

3 Methods

3.1 Seawater intrusion numerical model

This study used the SWI2 seawater intrusion package (Bakker et al., 2013) of the finite-difference MODFLOW-2005 groundwater model (Harbaugh, 2005). SWI2 was developed specifically to simulate regional seawater intrusion. Besides neglecting diffusion and dispersion effects, SWI2 does not require vertical discretization as an aquifer can be represented by a single layer containing several zones of constant (or linearly-varying) density. This sharp-interface code was chosen because the narrow transition zone observed in deep open wells (Fig. 2) suggests that diffusion and dispersion are less important than advection. Also, its short simulation times allow to efficiently run multiple simulations in the context of parameter estimation. Finally, model development and execution can be scripted in Python using the FloPy package (Bakker et al., 2016; 2020). This was advantageous because the whole framework, from data preprocessing to parameter estimation and uncertainty analysis, was developed in Python. This workflow was proven efficient for collaborative modeling (Shuler and Mariner, 2020) and to improve the transparency and reproducibility of decision-support models (White et al., 2020).

SWI2 successively solves two modified versions of MODFLOW-2005's groundwater flow continuity equation, which are each adjusted with pseudo-source terms representing density variations. These equations, detailed by Bakker et al. (2013), are rewritten here for a single-layer model with two constant-density zones (freshwater and seawater), separated by a unique interface (Fig. 3). Eq. 1 is solved for the whole saturated model domain:

$$\nabla(T\nabla h_f) = S\frac{\partial h_f}{\partial t} - \gamma + R \qquad Eq.1$$

where *T* is the transmissivity of the aquifer (m^2/s), h_f is the freshwater head at the water surface (m), *S* is the storage coefficient (dimensionless), *γ* is a source term (m/s) and *R* is a pseudo-source term representing the flux caused by density variations below the water table (m/s). Eq. 2 is then solved for the portion of the model domain below the interface:

$$(v_2 - v_1)\,\nabla(T\nabla\zeta) = n_e\frac{\partial \zeta}{\partial t} - \gamma_2 + R_2 \qquad Eq.2$$

where v_1 and v_2 are respectively the dimensionless densities of freshwater and seawater, *ζ* is the interface elevation approximating the 50-percent seawater salinity contour (m), n_e is the effective porosity (dimensionless), γ_2 represents all source terms beneath the interface (m/s) and R_2 is a pseudo-source term representing the flux caused by density variations below the interface (m/s). At the end of each MODFLOW timestep, after freshwater heads and interface elevations are updated, the horizontal movement of the interface is computed using a tip-and-toe tracking algorithm (Bakker et al., 2013).

Fig. 3 Implementation of the SWI2 sharp-interface code in the study site: schematic cross-section perpendicular to the island. The aquifer is represented by a single layer in which model cells contain constant-density freshwater and seawater zones separated by a sharp interface. The freshwater lens is delimited by the freshwater head at the top of the aquifer (h_f) and the interface elevation (*ζ*). A general head boundary (GHB) condition is used to simulate exchanges between the aquifer and the sea.

3.2 Model implementation and parameterization

A regular, 20 m x 20 m model grid was built for the island using the Python QGridder package. The model's active cells extended seaward up to 1 km from the coast. The aquifer formation was represented as a single layer containing two constant-density zones (freshwater and seawater), separated by an interface representing the 50-percent seawater salinity contour (Fig. 3). The bottom elevation of the aquifer was set to -300 m relative to local mean sea level.

Since insufficient pumping and observation timeseries were available, the model was calibrated assuming steady-state conditions with a mean pumping rate. This choice was in line with the objective of the study to simulate long term trends in seawater intrusion, rather than reproducing seasonal variability. A 5.5-year reference period was selected (mid-2014 to 2019), constrained by the availability of municipal pumping data at all wells, and during which pumping conditions were approximately the same. This choice affected the parameterization strategy. As they have no influence in steady-state conditions, specific yield and effective porosity values were considered to be homogeneous over the whole model domain. A mixed parameterization scheme was then used for the hydraulic conductivity field. The sand dunes and glacial sediments were assigned homogeneous hydraulic conductivities (zones of piecewise constancy), while the onshore Permian sandstone was parameterized using 52 pilot points distributed along a regular grid with a 500 m spacing. Hydraulic conductivities at model cells were determined by kriging of pilot point values based on an exponential variogram, with a range equivalent to 3 times the pilot point spacing. Since the aquifer was simulated as a single model layer, when several geological units overlapped (Fig. 1) an equivalent horizontal transmissivity was inferred from the arithmetic mean of hydraulic conductivities weighted by unit thicknesses. At the exception of a buffer around the coast, offshore pilot points were tied, effectively implying a homogeneous hydraulic conductivity for the offshore Permian sandstone. Prior information on hydraulic conductivity was based on a compilation of aquifer tests and existing literature (Freeze and Cherry, 1979), and prior values and ranges are shown in Table 1.

All boundary conditions were averaged over the 5.5-year reference period. A homogeneous recharge representing approximately 40% of total precipitation (900 mm/yr) was implemented for onshore cells (Table 1). This was supported by the small seasonal fluctuations observed in groundwater levels. A general head boundary (GHB) condition was implemented for offshore cells to represent freshwater head at the ocean bottom (Fig. 3). With GHB boundaries, flows between the aquifer and the sea are controlled by the seafloor elevation, sea level, the ratio between freshwater and seawater densities (respectively 1000 and 1025 kg.m^{-3}) and the hydraulic conductivity of the seabed (Hughes and White, 2014; Eq. 25). The seabed was assigned prior information close to that of the Permian sandstone (Table 1). Municipal pumping was implemented using the MNW2 package (Revised Multi-Node Well – Konikow et al., 2009), in order to assimilate water level observations (Section 3.3) and domestic pumping was implemented using the WEL package.

It has been shown that sharp-interface models (including SWI2), which assume saltwater to be static, tend to overestimate seawater intrusion (Dausman et al., 2010b). An empirical correction factor was developed by Pool and Carrera (2011) to correct this effect and the Lu and Werner (2013) modified version of this correction factor was implemented in the model (Eq. 3):

$$\varepsilon^* = \varepsilon\left[1 - \left(\frac{\alpha_T}{b}\right)^{1/4}\right] \qquad Eq. 3$$

where ε* is the corrected density ratio (dimensionless), α_T is the transverse (vertical) dispersivity (m), b is the aquifer thickness (m), ρ_f and ρ_s are respectively freshwater and seawater densities (kg/m^3) and ε is the density ratio (dimensionless) given by:

$$\varepsilon = \frac{\rho_s - \rho_f}{\rho_f} \qquad Eq. 4$$

For a transverse dispersivity of zero, the original and corrected density ratios are identical and the correction factor has no effect. As transverse dispersivity is difficult to characterize, it was considered as an adjustable parameter (Table 1), with a prior information based on a previous model of the island (Lemieux et al., 2015) and existing literature (Gelhar et al., 1992). All parameter distributions were assumed to be Gaussian and upper and lower bounds represented the 95% confidence interval (i.e. the mean ± 2 times the standard deviation).

Table 1 Prior and posterior parameter distributions, described by the mean and the 95% confidence interval (C.I.). Distributions are assumed normal for recharge and log-normal for all other parameters. The posterior hydraulic conductivity of all pilot points is specified as the average of all pilot point values, however the posterior 95% C.I. varies with each pilot point.

	Prior distribution		**Posterior distribution**	
Parameters	Mean	95% C.I.	Mean	95% C.I.
$K_{\text{sand dunes}}$ (m/s)	5×10^{-3}	5×10^{-5} - 5×10^{-1}	5×10^{-4}	5×10^{-6} - 5×10^{-2}
$K_{\text{sandstones}}$ (m/s)	4×10^{-5}	3×10^{-6} - 6×10^{-4}	3×10^{-4}	4×10^{-5} - 2×10^{-3}
$K_{\text{sandstones}}$ (m/s) (pilot points)	4×10^{-5}	3×10^{-6} - 6×10^{-4}	2×10^{-4}	-
$K_{\text{glacial sediments}}$ (m/s)	1×10^{-5}	1×10^{-7} - 1×10^{-3}	1×10^{-4}	3×10^{-6} - 3×10^{-3}
K_{seabed} (m/s)	2×10^{-5}	2×10^{-7} - 2×10^{-3}	9×10^{-6}	5×10^{-6} - 2×10^{-5}
Recharge (mm/yr)	380	180 - 580	547	401-693
Transverse dispersivity α_T (m)	1×10^{-1}	1×10^{-3} - 10	6×10^{-3}	3×10^{-4} - 1×10^{-1}

3.3 Observations

All observations were averaged over the reference period. Freshwater head observations were extracted from 3 types of locations: shallow wells, deep open wells and municipal pumping wells. Interface observations were derived from 3 sources of information: deep open wells, TDEM and ERT surveys. All processing steps are summarized in Figure 4. The uncertainty associated with all 6 observation types was then estimated.

Freshwater head observations

Water levels and pressures recorded at wells were converted to heads. Comparing simulated freshwater heads to observed heads requires an additional preprocessing step: the conversion of all measured heads to freshwaters heads (Bakker et al., 2013). This procedure, detailed by Post et al. (2018b), requires knowledge on the average water density in the well (Eq. 5):

$$h_{\mathrm{f}} = \frac{\rho_{\mathrm{a}}}{\rho_{\mathrm{f}}} h - \frac{\rho_{\mathrm{a}} - \rho_{\mathrm{f}}}{\rho_{\mathrm{f}}} z_{\mathrm{b}} \qquad Eq. 5$$

where h_{f} is the freshwater head (masl), h is the measured head (masl), ρ_{a} is the average density in the water column between the first and last density measurements (kg/m^3) and z_{b} is the elevation of the bottom of the well screen or open interval (masl). Heads measured in freshwater wells (i.e. shallow wells and pumping wells) were directly equal to freshwater heads, as the average density was equal to freshwater density. However, point water heads measured at deep open wells needed to be converted to freshwater heads (Fig. 4). Features of this calculation are presented in Appendix A and Table A.1.

At pumping wells, the comparison of observed heads with the values obtained with the relatively coarse model grid required an extra postprocessing step with the MNW2 package. Simulated heads were corrected for the difference between cell size and well radius, based on the Thiem (1906) steady-state flow equation (Konikow et al., 2009). At steady state and using the prior parameter set (Table 1), heads at pumping wells were on average 0.1 m lower than those simulated at the cell (with differences ranging from 0.001 m to 0.4 m). This average value dropped to 0.05 m when using the posterior parameter set, because of higher hydraulic conductivity values. Although in this study, the correction was relatively small, magnitudes will increase with increasing pumping rate and cell size, and with decreasing hydraulic conductivity and well radius (Eq. 4, Konikow et al., 2009).

Direct interface observations at deep wells

Interface elevations were extracted from downhole electrical conductivity profiles acquired in deep wells with large open or screened intervals. As the transition zone between freshwater and seawater spans a dozen meters, an objective method was required to extract interface elevations from all profiles. In their sharp-interface manual calibration, Babu et al. (2018) extracted an elevation close to the top of the transition zone, from the specific conductance of 2,500 µS/cm. However it was decided to extract an interface elevation close to the midpoint of the transition zone, as SWI2 simulates the 50% seawater salinity (Bakker et al., 2013). For each electrical conductivity profile, an error function (*erf*) was adjusted to the data points near the transition zone and the inflection point of this function was defined as the interface elevation.

Geophysical interface observations

Interface elevations were extracted from inverted TDEM and ERT geophysical data and were used as indirect interface observations for the hydrogeological inversion. This approach was preferred to a coupled hydrogeophysical inversion, in which hydrogeological and geophysical models are linked and inverted sequentially or simultaneously (Comte and Banton, 2007; Herckenrath et al., 2013; Steklova and Haber, 2016). These allow to use directly the geophysical observations instead of inverted geophysical model results, but are computationally demanding and therefore have mostly been applied on synthetic, small-scale or structurally simple regional models rather than complex, large-scale models.

Inversion of the one-dimensional TDEM measurements was conducted using the CSIRO Airbeo codes (Chen and Raiche, 1998; Raiche et al., 1985) and formation resistivity and thickness were adjusted for a three-layer subsurface model. From top to bottom, these layers represented the unsaturated zone, the freshwater-saturated zone and the seawater-saturated zone. The top of the seawater-saturated layer was defined as the freshwater-seawater interface elevation. Inversion of the two-dimensional ERT data was conducted using the RES2DINV software (Loke and Dahlin, 2002), which adjusted and smoothed the formation resistivity of subsurface model blocks of 2.5 m to 5 m user-defined thicknesses. The elevation of different threshold resistivities (from 2 to 15 Ω.m) was extracted at each model block, and a visual inspection showed that the threshold resistivity of 15 Ω.m yielded ERT interface elevations most consistent with the other interface observations. This value is close to the 14 Ω.m threshold which was chosen by Meyer et al. (2019) to extract interface elevations from time-domain airborne electromagnetic data, and which was based on EU drinking water guidelines (Jørgensen et al., 2012). ERT-derived interface observations were resampled to 80 m, to increase the statistical independence of the values obtained at model blocks while maintaining a good description of the interface's spatial variability. As the ERT survey was conducted outside of the reference period and under different pumping conditions, all data points within 100 m of current or past pumping wells were removed. The remaining points were preserved, since the interface showed minor temporal variability and no long-term trend (Section 2). A final visual analysis confirmed that all interface observation types were consistent and allowed the identification and removal of several outliers from the TDEM dataset. Geophysical surveys provided more interface observations than wells.

Fig. 4 Summary of the processing steps necessary for the assimilation of freshwater heads and interface elevations. All initial data are associated with uncertainties and each processing step adds supplementary uncertainty.

3.4 Observation uncertainties

For each data type, the total uncertainty was derived from the sum of variances associated with independent sources of uncertainty (assumed to be Gaussian). For these independent sources of uncertainty, the 95% confidence interval (C.I.) of measured values was assessed, and corresponding standard deviation values were then inferred by dividing the 95% C.I. by 4 (Table 2). Total uncertainties reflected the "*level of uncertainty in reproduction of observations*" (Fienen et al., 2010), including measurement and structural error. Sources of uncertainty and total uncertainties are summarized in Tables 2 and 3, respectively. Methods are described for freshwater head observations, direct interface observations at deep wells and finally for geophysical interface observations.

Freshwater head observations

For freshwater head observations, the total uncertainty σ_{hf} (m) was calculated as follows:

$$\sigma_{\mathrm{hf}} = \sqrt{(\sigma_{\mathrm{hfm}}{}^2 + \sigma_{\mathrm{temp}}{}^2 + \sigma_{\mathrm{pump}}{}^2)} \qquad Eq.6$$

where σ_{hfm} is the measurement uncertainty associated with the freshwater head (m), σ_{temp} is the uncertainty due to temporal aggregation over the reference period (m) and σ_{pump} (m) is the uncertainty associated with the reproduction of heads at pumping wells. The calculation of σ_{hfm} depended on the type of well. Where measured heads and freshwater heads were identical (i.e. shallow and municipal wells), σ_{hfm} was equal to the uncertainty of the measured head σ_{hm}, which encompassed operator error, inaccuracies of the measurement devices and of the elevation survey, and errors resulting from the conversion of water levels (or pressures) to heads. The 95% C.I. of head measurements was estimated at 0.15 m (Table 2). In deep open wells, the conversion of measured heads to freshwater heads propagated additional uncertainties to σ_{hfm}, stemming from uncertainties on the average water density ($\sigma_{\rho a}$) and on the bottom elevation of the open interval (σ_{zb}). The 95% confidence intervals were inferred from fluctuations of ρ_a and from field knowledge, respectively, and calculations of σ_{hfm} were performed following the method described by Post et al. (2018a). σ_{temp} was estimated by calculating the standard deviation of the mean (Appendix B). σ_{pump} was only implemented for head observations at municipal wells. A 95% C.I. of 0.5 m was chosen to account for modeled-to-measured misfit at pumping wells, which resulted in a similar uncertainty between deep wells and pumping wells, considering a null temporal aggregation (Table 3).

Direct interface observations at deep wells

Similarly, the total uncertainty of direct interface observations, σ_ζ (m), was calculated as follows:

$$\sigma_\zeta = \sqrt{(\sigma_{\mathrm{ECm}}{}^2 + \sigma_{\mathrm{temp}}{}^2 + \sigma_{\mathrm{s}}{}^2)} \qquad Eq. 7$$

where σ_{ECm} is the measurement uncertainty associated with the electrical conductivity profile (m) and σ_s is the uncertainty related to the definition of the interface elevation ('spatial' uncertainty – m). σ_{ECm} reflected operator error, inaccuracies of the conductivity measurement devices (resulting from imperfect calibration, instrument drift, varying accuracy) and of the elevation survey. A 95% C.I. of 0.2 m was assumed for electrical conductivity elevations (Table 2). σ_s was evaluated as one-sixth of the transition zone width (Table 3). While deep open wells are influenced by vertical borehole flows, which can lead to artificial electrical conductivity profiles (Rushton, 1980; Shalev et al., 2009), it was assumed that the relatively large σ_s values accounted for these flows. Total standard deviation values associated with direct interface observations averaged 2.11 m (Table 3) and therefore 95% confidence intervals nearing 10 m.

Geophysical interface observations

Uncoupled hydrogeophysical inversions propagate errors into the hydrogeological models (Hinnell et al., 2010). Uncertainties associated with the inverted geophysical data result from measurement and elevation errors, from parameters of the geophysical inversion (e.g. smoothness constraints), from non-unique hydrogeological interpretations (dependence of resistivity on lithology, saturation, solute concentration) and from electric or electromagnetic noise. A global, heuristic uncertainty was attributed to the geophysical interface elevations, which were considered more uncertain than direct interface observations at wells. ERT interface observations were considered more uncertain than TDEM interface observations because TDEM has better depth resolution than ERT for mapping conductive layers such as seawater layers (Christiansen et al., 2006). In addition, the smoothness constraint of the ERT inversion and the lack of resolution with depth of ERT images could result in missing the interface by a few meters. Contrary to TDEM data, the use of ERT interface observations also required the definition of a threshold resistivity. The 95% confidence intervals of the TDEM and ERT interface observations were set to 15 m and 20 m, respectively, to reflect the relative confidence in all three interface observations types.

Total uncertainty reflected the level of confidence in different observations groups (Table 3). On average, the uncertainty of freshwater heads at shallow wells (low σ_{hfm}, high σ_{temp}) was close to that of heads at pumping wells (high σ_{pump}, low σ_{temp}). Freshwater heads at deep open wells were more uncertain, as conversion of point water head to freshwater head resulted in a high σ_{hfm}. Interface observations derived from deep wells were more uncertain than head observations (high σ_s). These direct interface observations were less uncertain than TDEM observations, and ERT observations were the most uncertain dataset. With these uncertainties, all interface observations were consistent across the island. Quantification of measurement uncertainties was based on existing methods when available. However, this process required making a certain amount of choices, based on in-depth knowledge of the study site and fieldwork methods and on expert judgment. This is further discussed in Section 5.1.

Table 2 Individual sources of uncertainty in the observation dataset. The standard deviation values (designated by σ_i notations) were obtained by dividing the 95% confidence interval (C.I.) by 4. Errors are assumed to follow independent Gaussian distributions with a mean of zero.

Sources of uncertainty	**95% C.I.**	**Standard deviation**	
Measured head (m)	0.15	σ_{hm}	0.0375
Modeled-to-measured misfit at pumping well (m)	0.5	σ_{pump}	0.125
Average water density (kg.m^{-3})	8	$\sigma_{\rho a}$	2
Bottom elevation of the open interval (m)	0.15 – 4	σ_{zb}	0.0375 – 1
Elevation of an electrical conductivity measurement (m)	0.2	σ_{ECm}	0.05
TDEM-derived interface elevation (m)	15	$\sigma_{\zeta\ TDEM}$	3.75
ERT-derived interface elevation (m)	20	$\sigma_{\zeta\ ERT}$	5

Table 3 Uncertainties associated with freshwater head (h_f) and interface elevation (ζ) observations, in increasing order. The total uncertainty (σ) is a function of independent sources of uncertainty such as measurement uncertainties (σ_m), pumping in a model cell (σ_{pump}**),**

temporal aggregation (σ_{temp}) or spatial definition of the saltwater interface (σ_s). Given that settings vary slightly from well to well, average values are provided. σ is calculated using Eq. 6 and 7. The signal-to-noise ratio is equal to the mean absolute observation value (1.5 masl for freshwater heads and -44 masl for interface elevations) divided by σ.

Observation group	**Number of observations**	**σ_m (m)**	**σ_{pump} (m)**	**σ_{temp} (m)**	**σ_s (m)**	**σ (m)**	**Signal-to-noise ratio**
h_f shallow wells	4	0.0375	-	0.1	-	0.11	13
h_f pumping wells	9	0.0375	0.125	0.0002	-	0.13	11
h_f deep wells	7	0.15	-	0.06	-	0.17	9
ζ deep wells	7	0.05	-	0.33	2.08	2.11	21
ζ TDEM	48	-	-	-	-	3.75	12
ζ ERT	87	-	-	-	-	5	9

3.5 Parameter estimation

The model was calibrated under steady-state conditions representative of the reference period. For numerical reasons, the solution was obtained after stabilization of a long transient simulation with constant boundary conditions. Each model run started with an initial run without the SWI2 package, to compute a steady-state distribution of heads. These were used as the initial head distribution for a model run with the SWI2 package, using the Ghyben-Herzberg equation (Post et al., 2018a) to compute the initial interface elevation. This second run stretched out for 1000 years, to allow the freshwater lens to reach a steady state under the average stresses prescribed and for the parameter set tested. The total run time was short, around 7 min on a desktop computer (1.9 GHz Intel Core i7®) including model run, pre- and post-processing operations.

Parameter estimation was conducted with PEST (Doherty, 2015), which uses the Gauss-Levenberg-Marquardt algorithm to minimize the square-weighted differences between simulated and measured values. PEST was selected because this algorithm is particularly adapted to highly parameterized models, such as the model developed in the study; to computationally expensive models and to regularized inversion problems (Doherty, 2004). In addition, pre- and post-processing of PEST files can be readily implemented in Python using the PyEMU library (Python framework for Environmental Modeling Uncertainty analyses), which also offers multiple tools for model-independent uncertainty analysis (White et al., 2016). This was consistent with the overall strategy of developing a complete framework in Python (Section 3.1). The PEST-HP code was selected from the PEST suite as it is designed specifically to improve inversion performance when model runs are parallelized (Doherty, 2020).

The model's 56 hydraulic conductivities, recharge and transverse dispersivity were adjusted during parameter estimation. Singular value decomposition was used to regularize the inversion. Prior information on the 58 parameters (Table 1) was incorporated using first-order Tikhonov regularization (preferred value) for all parameters and second-order Tikhonov regularization (preferred homogeneity) for the pilot points (Doherty et al., 2010). A total of 162 freshwater head and interface observations was used to constrain parameter estimation. To avoid overfitting, the weighting of the regularization objective function was conducted as detailed by Doherty (2015), with weights defined as the inverse of the standard deviation.

3.6 Linear-based uncertainty analysis and data worth

A first-order, second-moment (FOSM) uncertainty analysis was conducted using PyEMU. In a linearized model, prior parameter uncertainty and epistemic uncertainty (related to measurement errors) are propagated to the posterior parameter set and then to model predictions (Fienen et al., 2010; White et al., 2016). All parameter distributions and measurement errors are assumed to be Gaussian, implying that forecast distributions are also Gaussian. While the linear-based analysis is approximative, it is less computationally expensive than nonlinear methods and still provides insight into forecast uncertainty and data worth (Brunner et al., 2012; Hill et al., 2016; Nolan et al., 2015). Even with the relatively short model run times afforded by SWI2, nonlinear methods based on random sampling such as Monte Carlo simulations would be unfeasible. The high dimensionality of the model would require a very high number of model runs. Furthermore, the linear assumption was proven to be reasonable, as the integrity of model sensitivities used for the linearization of the model was verified beforehand using the JACTEST utility of PEST (Doherty, 2004).

Two types of forecasts were considered for the analysis: the volume of freshwater (a global forecast) and the interface elevation at pumping wells (local forecasts). Both types of forecasts were of interest for groundwater management, under current and future pumping and climate conditions. The importance of model parameters in forecast uncertainty was quantified by examining the decrease in forecast uncertainty as a result of parameters being considered as perfectly known (Fienen et al., 2010). The worth of different observation groups was evaluated by examining the decrease in prior forecast uncertainty as a result of progressively adding these observation groups to an initially empty calibration dataset. The worth of each observation group was therefore considered independently from the others.

4 Results

4.1 Parameter estimation

The parameter estimation procedure ended after 8 calibration iterations, necessitating 1847 model runs. The initial objective function of 4709 was decreased to a final value of 1175. Using a total of 70 cores (at 2.1GHz), the procedure ended after 10.5 hours. Summary statistics for each observation group are provided in Table 4, allowing to assess the fit to available observations: the root-mean square error (RMSE) is the average of the squared residuals and the mean error (ME) is the mean difference of the residual errors, with residuals defined as the difference between simulated and observed values. Freshwater head residuals at shallow and municipal wells had small RMSE values compared to the average observed value (1.5 masl), although the presence of an outlier (identified Fig. 5a) increased the average RMSE and ME values for the municipal wells. This outlier was most probably linked to a technical issue with the automated meter, but was kept for transparency. For both groups, no bias was identified in simulated heads, as indicated by an equal distribution of values around the 1:1 diagonal line in Figure 5a and small ME values (Table 4). The RMSE values for interface observations were small to intermediate (6 to 11 m), compared to the average observed interface elevation (-44 masl), and Figure 5b shows TDEM and ERT interface residuals scattered around the 1:1 diagonal line. Small ME values indicated little bias in simulated interface elevations (Table 4). However, it can be noted that the highest ME values, whether for freshwater head or interface observations, were for observations made at deep open wells (Table 4). Almost systematically, simulated heads and interface elevations were respectively lower and deeper than the observed values (Fig. 5). For freshwater head observations at these deep wells, the RMSE value was high and bias was clearly identifiable.

Table 4 Summary statistics of the calibration: root-mean-square error (RMSE) and mean error (ME) of the residuals for each freshwater head (h_f) and interface (ζ) observation group. RMSE values are small to intermediate compared to the order of magnitude of observed values. The statistics for heads at municipal pumping wells are high because of an outlier.

	$h_{\text{f shallow wells}}$	$h_{\text{f pumping wells}}$	$h_{\text{f deep wells}}$	$\zeta_{\text{deep wells}}$	ζ_{TDEM}	ζ_{ERT}	Total
RMSE (m)	0.1	0.6	0.5	6.3	10.8	7.0	7.4
ME (m)	0.05	-0.3	-0.4	-3.7	-2.3	1.0	0.4

Fig. 5 Scatter plots of simulated to observed data: (A) freshwater heads and (B) interface elevations. The 1:1 diagonal line represents equal simulated and observed values. Bias is noticeable for freshwater heads and interface elevations at deep open wells. An outlier is clearly identifiable within the freshwater head observations at pumping wells (panel A).

The final parameters were consistent with prior information, as is shown by the posterior parameter values being included in the prior 95% confidence intervals (Table 1). Figure 6a shows the final hydraulic conductivity field post-calibration. A lower hydraulic conductivity zone, which was not predicted by the geological map (Fig. 1), arose in the south of the island in an area where all interface observations were deeper. The final transverse dispersivity α_T had a low but hydrogeologically reasonable value, resulting in a corrected density ratio of 0.022 instead of 0.025 (substituting parameters from Table 1 into Eq. 3). The transverse dispersivity parameter was found to be uncorrelated to other model parameters (by analysis of the correlation coefficient matrix – White et al., 2016). New maps of the interface elevation (Fig. 6b) and of freshwater lens thickness were generated by these optimum parameters, and can be used by groundwater managers to support decision-making.

Fig. 6 Post-calibration maps: (A) hydraulic conductivity field and (B) freshwater-seawater interface elevations. The interface is relatively shallow on the island. The general head boundary

(GHB) condition is implemented on all cells between the coastline (full line, panel B) and model boundaries (dotted line, panel B).

4.2 Uncertainty analysis and data worth

The uncertainty in the posterior parameter set was reduced through parameter estimation, as is shown Table 1 by the decrease in the 95% confidence intervals. In areas parameterized using pilot points, the uncertainty of the hydraulic conductivity field was reduced near observations but remained close to the prior uncertainty far from observations. The uncertainty of model forecasts was noticeably reduced through parameter estimation, as shown by large reductions in predictive uncertainties (Fig. 7).

Fig. 7 Prior and posterior probability distributions of the model forecasts: (A) total freshwater volume and (B) interface elevation in the cell containing pumping well no. 1 (ζ_{muni} 1). Distributions are represented by 95% confidence intervals. The trend in panel B is representative of the other pumping wells.

The first part of the data worth analysis considered the importance of parameters in forecast uncertainty (Section 3.5). The analysis showed that hydraulic conductivities were the dominant source of forecast uncertainty for all forecasts (Fig. 8). For predictions of the freshwater volume, recharge was a small but non-negligible source of uncertainty while transverse dispersivity had a minimal contribution (Fig. 8a). For interface elevations at municipal wells, both recharge and transverse dispersivity had minimal contributions to total forecast uncertainty (Fig.8b).

Fig. 8 Percent decrease in posterior forecast uncertainty (standard deviation σ_{post}) when one parameter group is considered fully known: (A) total freshwater volume and (B) interface elevation in the model cell containing pumping well no. 1 (ζ_{muni} 1). The hydraulic conductivity (K)

field accounts for the majority of uncertainty reduction, while recharge (*R*) and transverse dispersivity (α_T) play a smaller role. Panel B is representative of the other municipal wells.

The second part of the data worth analysis considered the importance of observation groups in reducing prior forecast uncertainty (Section 3.5). For all the forecasts evaluated, the analysis revealed that interface observations, and particularly geophysical observations, were most effective to reduce predictive uncertainty (Fig. 9a, b). For the freshwater volume (Fig. 9a), using only one group of geophysical interface observations, whether ERT or TDEM, resulted in a larger predictive uncertainty reduction (around 85%) than using all freshwater head observations combined (70% reduction). This also indicates that the observation dataset was redundant: using less data, small predictive uncertainties could also have been obtained. For interface elevations at pumping wells, while data worth varied slightly depending on the well, interface observations were systemically responsible for the top two uncertainty reductions (Fig. 9b) and for 7 municipal wells out of 9, the geophysical interface observations occupied this rank.

For all local forecasts (at pumping wells), freshwater heads from deep wells were systematically the least effective observations to reduce predictive uncertainties (Fig. 9a). For the freshwater volume, they were equally informative as all the other freshwater head observations (Fig. 9a). When looking at individual observation worth, the observations closest to the wells were more informative of interface elevations at pumping wells (Fig. A.1).

Fig. 9 Percent decrease in prior forecast uncertainty (standard deviation σ_{prior}) when one or several observation groups is added to the initially empty calibration dataset: (A) total freshwater volume and (B) interface elevation in the model cell containing pumping well no. 1 (ζ_{muni} 1). Interface observations, particularly geophysical observations, lead to a considerable decrease in prior forecast uncertainties. The order in panel B varies depending on the well.

5 Discussion

The discussion reviews the procedure used for measurement uncertainty quantification, the results of parameter estimation and the findings of the linear-based uncertainty analysis. Additional points are ultimately discussed, regarding the use of different interface observations and limits to the study. Findings and recommendations are summarized in Table 5.

5.1 Observation uncertainties

The quantification of measurement uncertainties was a challenging process, because many sources of uncertainty were not truly known and a certain amount of subjective choices had to be made. The final uncertainty values that were used for the study reflected site-specific considerations. For instance, using $\sigma_{\rho a}$ = 2 kg/m^3, σ_{zb} = 0.0375 – 1 m and σ_{hm} = 0.0375 m (Table 2) resulted in deep well freshwater head uncertainties around 0.15 m (σ_{hfm}, Table 3). This is higher than the values obtained by Post et al. (2018b) at their study site (σ_{hfm} = 0.02 – 0.08 m), resulting from the choices $\sigma_{\rho a}$ = 1 kg/m^3, σ_{zb} = 0.01 m and σ_{hm} = 0.02 m. For interface observations, the uncertainty on the location of the 50% seawater salinity contour will increase with the width of the transition zone. At this study site the transition zone was narrow, so the estimated uncertainties might be in the lower range compared to other coastal areas with larger transition zones (for example due to more heterogenous and/or lower hydraulic conductivity geological formations). For direct interface observations from deep open wells, the uncertainty depends on the manner in which the width of the transition zone is defined. For ERT-derived interface observations, a more heterogeneous system might also make the extraction of a threshold resistivity more challenging, resulting in more uncertain observations (Section 5.4). Total uncertainty values were also affected by temporal aggregation uncertainty, which was specific to each well. The uncertainty values reflected model-specific considerations. An uncertainty for modeled-to-measured misfit at pumping wells was indeed defined to account for the structural error introduced by the finite-difference model and coarse grid (20 m x 20 m). This might not have been necessary with a refined finite-difference grid or a finite-element model mesh.

While absolute values of uncertainties are site- and model-specific, what is most important is that total uncertainties reflect the relative level of confidence in each observation type. Heads measured directly at shallow wells were less uncertain than freshwater heads derived from deep open wells. Heads at shallow wells were also less uncertain than heads simulated at pumping wells by the finite-difference model, although uncertainty linked to modeled-to-measured misfit at pumping wells (σ_{pump}, Table 2) could be explored in more depth. Head observations were less uncertain than saltwater interface observations. Direct interface observations from deep, open wells were less uncertain than geophysical interface observations, although this assumption could be challenged for wide transition zones and downhole profiles heavily affected by borehole flows. TDEM-derived interface observations were less uncertain than ERT-derived interface observations. These general trends, which are linked to the nature of the measurements (or model) and to the amount of pre- and post-processing associated with the measurements, will likely be the same in other studies. Therefore, the relative weighing scheme will likely be similar, which is determining for the results of the inversion and of the data worth analysis (Section 5.3).

5.2 Parameter estimation

The final parameter set was hydrogeologically reasonable and conformed to the conceptual model. Recharge was worth approximately 60% of total precipitation, which is higher than past estimates of 25% to 40% of total precipitation (Section 2.1) but seems to be more consistent with the negligible runoff observed on the whole island (no streams or surface water). The hydraulic conductivity of the sand dunes was greater than values for the Permian sandstone, which were in turn greater than the hydraulic conductivity of the seabed and of the glacial sediments (Table 1). In the final hydraulic conductivity map, the glacial paleovalley delineated in the geological map (Fig. 1) was overshadowed by a lower conductivity region that arose from parameter estimation (Fig. 6A). More generally, the thickness-averaged hydraulic conductivity in model cells containing glacial sediments or sand dunes remained close to the hydraulic conductivity of the sandstone, as the sandstone had a predominant thickness compared to the overlying formations.

The signal-to-noise ratio, i.e. the ratio of a signal to the level of background noise, was defined as the average measured head (or interface) value to the total standard deviation. The signal-to-noise ratio of the observations was low (Table 3) because uncertainties were high compared to the magnitude of the observations, which made parameter estimation challenging. As the uncertainties of coastal aquifer observations are high, due to many factors highlighted in Section 3.3, they should not be underestimated to avoid overfitting parameters to measurement error (which would reduce the predictive capacity of the model). However, conservative uncertainty estimates resulted in uncertainties so high that the observations could not be reliably differentiated from the noise and were unable to constrain model parameters. Thus, the uncertainties defined in our framework aimed to balance these effects. Since calibrated parameters were consistent with prior information, a suitable model-to-observation fit was obtained for the majority of observation groups and no overall model bias was noted for both freshwater head and interface observations, parameter estimation was considered successful. The implementation of weighted Tikhonov regularization prevented PEST from excessive reduction of model-to-measurement misfit.

Important observations were made when independently examining the residuals of observation groups. The assimilation of heads at pumping wells in the regional model was successful through the use of the MNW2 package with the Thiem (1906) correction (Section 3.3). Head and interface observations from deep open wells were biased (Fig. 5), possibly because of vertical flows (Shalev et al., 2009). The displacement of the salinity profile in a well due to borehole flows could indeed affect both the measurement of the interface elevation and the calculation of the freshwater head (through a modification of the average water density in the well, Eq. 5). For instance, an upward flow could result in a shallower observed interface elevation and a higher observed freshwater head (because of an artificially higher average water density) than the ones simulated for the aquifer, as shown respectively in Figures 5B and 5A. Therefore, parameter estimation should not be conducted against data from deep open wells alone, as this could bias model calibration. Characterizing the vertical flows through temperature or flow profiles could help evaluate the magnitude of the bias. Furthermore, acquiring and preprocessing heads at deep open wells was costly and time-consuming, but these were unable to constrain the calibration because of their low signal-to-noise ratios (Table 3). In contrast, not only were head observations from shallow wells easier to process, but they were more beneficial for calibration because of their higher signal-to-noise ratios. Finally, the dispersion of TDEM and ERT interface residuals (Fig. 5b) showed this data to be noisy and it was considered that fitting parameters to the mean of the geophysical observations (i.e. targeting little model bias) was an acceptable target rather than trying to fit each individual observation.

Parameter estimation resulted in a non-null correction factor for the density ratio. However, it should be noted that the applied correction factor was developed for lateral seawater intrusion (without upconing effects) and generally has not been used for freshwater lenses (Werner et al., 2017), in which both longitudinal and transverse dispersivity affect seawater intrusion.

5.3 Uncertainty analysis and data worth

This framework allowed to quantify the uncertainty of model forecasts of interest to water managers, and the large uncertainty reduction of forecasts during parameter estimation (Fig. 7) then demonstrated that model forecasts were constrained by the calibration process. This shows the importance of the parameter estimation framework in a decision-support context. It should be noted that the prior uncertainty of the forecasts was already informed by data, before the parameter estimation process was undertaken. Before assimilating the information contained in the calibration dataset, multiple site characterization data and expert knowledge were assimilated to develop the conceptual model and the parameterization scheme, and to inform the prior parameter values. This may explain why the observation dataset appeared to be redundant (Fig. 9), as the forecasts were already informed by the work preceding calibration.

The results of the data worth analysis are somewhat specific to the context, as data worth is dependent on observation uncertainty, which can be site-specific and model-specific (Section 5.1), but also on the number and location of observations relative to the forecasts and to aquifer configuration. However, this investigation is valuable because data worth analyses of seawater intrusion models have traditionally focused on variable density and mostly synthetic models (Baker, 2010; Dausman et al., 2010a; Sanz and Voss, 2006; Shoemaker, 2004) and some conclusions can be generalized.

It was found that further characterization of the hydraulic conductivity field would most reduce forecast uncertainties, while further characterization of recharge and of transverse dispersivity (as a correction factor) would be less beneficial (Fig. 8). However, because of scaling effects and parameterization assumptions, it is difficult to quantify how field measurements can reduce prior parameter uncertainty (White et al., 2016) so conclusions are more easily drawn regarding the worth of observations. Interface observations were essential to reduce predictive uncertainties (Fig. 9), even though they are much more uncertain than head observations (Table 3). This was expected for interface elevations at pumping wells, as the observations and predictions are of the same nature (interface elevations). Because pilot point parameterization was implemented onshore, the worth of direct *vs* TDEM *vs* ERT observations then depended on which observations were closest to the wells (Fig. A.1). Freshwater head observations from deep wells were the least effective observations for reducing predictive uncertainty, because of the high uncertainty resulting from conversion to freshwater head. For the total freshwater volume, interface observations were also crucial to reduce predictive uncertainties, with geophysical surveys being most informative (Fig. 9a). This is because the geophysical surveys provided a much greater number of interface observations compared to the total number of wells, and they provided observations for areas on the island otherwise uncharacterized by the wells (Fig. 1), giving an extensive view of the shape of the freshwater-seawater interface on the island.

5.4 Additional considerations on coastal aquifer observations

The data worth analysis (Section 5.2) showed that interface observations closest to the pumping wells were most informative of predictions of the interface at these wells. However, other aspects need to be considered for the design of a data collection strategy. For instance, it has been shown that deep open wells drilled near pumping wells present a risk of saltwater contamination for the pumping wells (Rotzoll, 2010). Also, TDEM data points can generally not be acquired too close to pumping wells, as they are affected by electromagnetic noise due to pumping and fencing installations. In our case, acquiring additional ERT transects close to the pumping wells might be useful to obtain additional interface observations.

The analysis of model residuals showed that freshwater head and interface observations from deep open wells were biased. However, having at least one deep open well on the study site was essential to estimate the approximate width of the transition zone: in the present study, a narrow transition zone oriented the choice of a sharp-interface model. It was also essential to observe the temporal variability of the transition zone, and in the study small variability led to the assimilation of ERT interface data outside of the reference period. Finally, for the assimilation of ERT interface observations, having at least one deep open well was important to choose a threshold resistivity defining ERT-derived interface elevations.

In order to assimilate ERT interface observations, it was critical to have at least one other type of interface observation (e.g. from deep open wells or from a TDEM survey) to define the threshold resistivity (Section 3.3). This threshold could depend on the lithology and choosing an arbitrary threshold with no means of verification could have biased the ERT interface observations. For example, using a threshold resistivity of 5 Ω.m instead of 15 Ω.m yielded mean interface elevations of -60 masl rather than -43 masl. Additionally, the choice of a fixed threshold relied on the reasonable assumption that the sandstone aquifer was relatively homogeneous and that resistivity spatial variations were due to salinity variations only, however reliable identification of such a threshold could be challenging in more heterogeneous aquifer systems (González-Quirós and Comte, 2020). It appears that the only interface observations that could have been used alone were TDEM data inverted with a limited-layer model. This might be the best alternative to constrain the calibration of sharp-interface seawater intrusion models, in cases where the interface depth is within the range of the depth of investigation and where the land cover is not too urbanized. Including interface data from several sources made the identification of TDEM outliers easier, and the uncertainty of the TDEM observations was defined based on the uncertainty of the other interface observations. More generally, assimilating multiple interface observation types (at least two) seemed essential, due to the numerous uncertainties and possible biases associated with each of them. Having an area where all interface observations coexisted (e.g. having geophysical data points near a deep open well) was valuable to check for consistency, uncertainty and biases.

Finally, for future water level collection efforts, it was found that installing loggers in shallow freshwater wells was more beneficial than installing loggers in deep open wells, where the total uncertainty would remain high due to the uncertainty on average water density (Table 3). The assimilation of flow observations was not considered for this study, as none were available (no streams, no tracer tests). Hughes and White (2014), through the calculation of composite parameter sensitivities, inferred that their model parameters were informed by the head and flow observations in their dataset. In future research, it would be interesting to quantify the worth of flow observations, including observations of submarine groundwater discharge, for model calibration and for reducing predictive uncertainties.

A limit of this framework is that the interface elevation forecasts at cells containing pumping wells are not directly representative of the true interface elevation below wells. Just as the drawdown at pumping wells is averaged over the cell area, the upconing of the interface under the well is also averaged over the cell area. This effect was considered for simulated heads (with the MNW2 package), but the simulated interface should be corrected for this as well. Local hydraulic conductivities near pumping wells may also not be represented accurately by the regional model. Finally, the modeled interface should be corrected from neglected dispersion and diffusion effects, which are no longer negligible under pumping wells. These interface values should therefore be interpreted as indicative values. However, we still believe this regional model can prove a useful and informative tool for groundwater management decision-support. The impact of uncertain storage parameters on forecast uncertainty was not accounted for, therefore this will need to be considered for transient simulations.

Table 5 Main conclusions on coastal aquifer observations, for data collection, parameter estimation and data worth.

Observations	Main conclusions
Freshwater heads (*h_f*)	1. Acquire and assimilate h_f observations from shallow wells in priority (high signal-to-noise ratios), compared to h_f observations from deep open wells (time-consuming preprocessing, bias, low signal-to-noise ratios) 2. If their number is limited, placing pressure loggers in shallow wells reduces total uncertainty σ_{hm} more than for deep open wells 3. If available, assimilate h_f observations from pumping wells using the MNW2 package and the Thiem (1906) correction
Interface elevations (*ζ*)	1. Implement a correction factor (e.g. Lu and Werner, 2013) to correct for the overestimation of seawater intrusion by the sharp-interface model 2. Acquire ζ observations, as they are valuable to reduce model predictive uncertainty. TDEM and ERT surveys are especially valuable 3. Acquire ζ observations as close as possible to pumping wells to lower the uncertainty on pumping well interface predictions (considering the risk of saltwater contamination posed by deep open wells and electromagnetic noise near pumping installations) 4. Acquire ζ observations over different portions of the study area to lower the uncertainty on the freshwater volume prediction

	5. Assimilate at least two ζ observation types, as ζ observations at deep open wells are biased (vertical flows), ζ observations from ERT can be biased (if the threshold resistivity is incorrectly defined) and all ζ observations have a low signal-to-noise ratio
	6. Have an area where all ζ observations coexist to check for consistency, uncertainty and bias
	7. Use ζ observations at deep open wells and/or ζ observations from TDEM to define a threshold resistivity for ζ observations from ERT
	8. Geophysical data is noisy: during parameter estimation, aim for no model bias rather than fitting each observation individually
	9. Have at least one deep monitoring well on the study site, to guide the choice of the model and data assimilation
All	1. Coastal aquifer observations have a low signal-to-noise ratio: evaluate uncertainties of observation groups adequately and implement weighted Tikhonov regularization, to avoid overfitting to measurement errors (if uncertainties are too low) while allowing flexibility for parameter estimation (if uncertainties are too high)
	2. Parameter estimation should not be conducted against data from deep open wells alone, as this data is biased by vertical flows which could bias model calibration

6 Conclusions

Using multiple head and interface observations from various well types and geophysical surveys, parameter estimation of the sharp-interface model was carried out successfully, and provided the basis for a linear-based uncertainty analysis. It was demonstrated that parameter estimation led to an important decrease in predictive uncertainty for two important decision-support model forecasts: the volume of freshwater and interface elevations near municipal pumping wells. The methodology that was developed in the study is relatively straightforward, showing that parameter estimation and linear uncertainty analysis could be carried out more systematically for regional sharp-interface models developed for decision-support. The complete framework is highly reproducible as it was scripted using Python (open-source and documented packages) and it is shared in the Supplementary Material. It could be implemented in multiple other coastal areas, as it was developed for a common hydrogeological setting (an unconfined aquifer), it used typical coastal aquifer observations from wells and geophysical surveys and it examined typical seawater intrusion model forecasts.

The analysis of residual errors and a data worth analysis provided further insight on data assimilation for sharp-interface models. Interface observations were critical to reduce predictive uncertainties, especially geophysical observations as they provided a large number of data points and a wide spatial coverage. While deep open wells were essential to select a sharp-interface approach (through the identification of a narrow transition zone), preprocessed heads and interface observations from these wells were biased, which deterred their reproduction by the model. All coastal aquifer observations had a low signal-to-noise ratio, requiring a careful evaluation of measurement uncertainties. These findings can help guide future data assimilation and data collection efforts in similar contexts.

To conclude, this study highlighted several advantages of the sharp-interface approach for modeling regional seawater intrusion, compared to the variable density approach. Fast model run times allowed to conduct parameter estimation (yielding minimum error variance predictions) and uncertainty analysis (quantifying predictive uncertainties and their sources). Also, in relatively homogeneous aquifers with a narrow transition zone, extracting interface observations from geophysical data is more straightforward, and likely as reliable, than extracting salinity observations from geophysical data, as is usually done for the calibration of variable density models. Further applications of this sharp-interface approach are being explored, for example its use for municipal pumping optimization and to explore climate projections. In future research, predictive uncertainties could be evaluated using non-linear uncertainty analysis methods. The uncertainty of the interface elevation modeled at municipal wells is being explored in more detail. Although this methodology was developed for a freshwater lens, the findings are transferable to continental settings (with lateral seawater intrusion only) and the location of the toe of the saltwater wedge could be explored as an additional model forecast.

7 Appendices

Fig. A.1 Percent decrease in prior forecast uncertainty (standard deviation σ_{prior}) when an individual observation is added to the initially empty calibration dataset**,** for the interface elevation in the model cell containing municipal well no. 1 (ζ_{muni} 1).

Appendix A Conversion of point water heads to freshwater heads in deep open wells

Downhole electrical conductivity and temperature profiles are used to estimate water density profiles, using the UNESCO 1980 equation of state (Post, 2012). The average density of the water column is then estimated using Eq. (A.1) (Post et al., 2018b):

$$\rho_{\mathrm{a}} = \frac{\int_{z_1}^{z_{\mathrm{n}}} \rho(z_{\mathrm{d}})\, dz_{\mathrm{d}}}{D} \qquad Eq.\ (A.1)$$

where ρ_{a} is the average density in the water column between the first and last density measurements (kg/m^3), the numerator represents the integration of density measurements ρ (kg/m^3) at elevations z_{d} (masl), between the first and last density measurements (at elevations z_1 and z_{n}), and D is the distance between the first and last density measurements (m). The average density ρ_{a} is then used in Eq. 5. Table A.1 summarizes the principal parameters intervening in Eq. 5 for the study site's 7 deep open wells.

Table A.1 Conversion of measured heads to freshwater heads in the island's deep open wells and associated uncertainties. Freshwater heads (h_{f}) are calculated from measured heads (h), average water density (ρ_{a}) and the bottom elevation of the open or screened interval (z_{b}), using Eq. 5. The uncertainties $\sigma_{\rho a}$, σ_{zb}, and σ_{hm} are defined in Table 2 and σ_{hfm} is calculated following the method in Post et al. (2018b). Freshwater heads are systematically higher than point water heads and the highest freshwater heads are obtained at wells intersecting larger portions of saline groundwater. Wells pz01, pz02, pz03 and pz04 are located in a transect perpendicular to the coast and a seaward horizontal gradient can be observed after conversion to freshwater heads.

Well name	**$\rho_a \pm \sigma_{\rho a}$ (kg/m^3)**	**$z_b \pm \sigma_{zb}$ (masl)**	**$h \pm \sigma_{hm}$ (masl)**	**h_f (masl)**	**σ_{hfm} (m)**
pz01	1009 ± 2	-96.05 ± 0.0375	1.24 ± 0.0375	2.14	0.20
pz02	1009 ± 2	-76.78 ± 0.0375	1.12 ± 0.0375	1.97	0.16
pz03	1005 ± 2	-59.11 ± 0.0375	1.07 ± 0.0375	1.59	0.12
pz04	1008± 2	-56.21 ± 0.0375	0.81 ± 0.0375	1.58	0.12
pz05	1011 ± 2	-76.28 ± 0.0375	0.87 ± 0.0375	1.73	0.16
pz07	1003 ± 2	-84.71 ± 1	1.15 ± 0.0375	1.44	0.17
pz08	1010 ± 2	-73.84 ± 1	0.7 ± 0.0375	1.43	0.15

Appendix B Calculation of temporal aggregation uncertainty

The temporal aggregation uncertainty σ_{temp} represents the uncertainty in a mean value, resulting from averaging observations over a given time period. It is estimated by calculating the standard deviation of the mean. This method was described by Hughes and Hase (2010) and is rewritten here for head observations. It is the same method for interface observations. For a given well, a mean head value $\bar{h}$ (m) can be calculated from the arithmetic mean of individual head observations acquired at different times (Eq. (A.2)):

$$\bar{h} = \frac{1}{N}\sum_{i=1}^{N} h_i \qquad Eq.\ (A.2)$$

where N is the total number of head observations h_i (m) made at the well during the time period. The standard deviation σ_{N-1} (m) of the head observations at the well can be calculated using Eq. (A.3):

$$\sigma_{N-1} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(h_{\mathrm{i}} - \bar{h})^2} \qquad Eq.\ (A.3)$$

The less the number of head observations available at the well, the greater the uncertainty in the calculated mean. A standard deviation of the mean $\sigma_{\bar{h}}$ (m), also named standard error, can be calculated to evaluate this uncertainty, using Eq. (A.4):

$$\sigma_{\bar{h}} = \frac{\sigma_{N-1}}{\sqrt{N}} \qquad Eq.\ (A.4)$$

As the number of head observations in the well increases, the uncertainty in the mean $\sigma_{\bar{h}}$ decreases. However, for small sample sizes, Eq. (A.4) cannot be used, as this would result in a standard deviation of the mean equal to the standard deviation of the measurements. A threshold of six observations was chosen, over which the error on $\sigma_{\bar{h}}$ is smaller than 32% (Hughes and Hase, 2010, Eq. 2.8) i.e. $\sigma_{\bar{h}}$ continues to reflect a 68% confidence interval. Under this threshold, the uncertainty in the mean was defined as an average uncertainty σ_{a} (m). σ_{a} represents the global variability of head observations in all wells and was calculated as the square-root of the mean of all head variances in the model (Hughes and Hase, 2010). Therefore, the uncertainty due to temporal aggregation σ_{temp} (m) was defined using Eq. (A.5):

$$\begin{cases} \sigma_{temp} = \sigma_{\bar{h}} & if\ N \geq 6 \\ \sigma_{temp} = \sigma_{a} & if\ N < 6 \end{cases} \qquad Eq.\ (A.5)$$

Acknowledgments

Funding : This work was supported by Quebec’s Ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCC) [project « Acquisition de connaissances sur les eaux souterraines dans la région des Îles-de-la-Madeleine » (Groundwater characterization project in the Magdalen Islands region)]; and the Fonds québécois de la recherche sur la nature et les technologies (FRQNT) [International internship program accessed through CentrEau, the Quebec Water Research Center]. The authors would like to thank the Municipality of Les Îles-de-la-Madeleine for providing pumping datasets and information on current and historical groundwater management. They would also like to thank the team at Université Laval working on the Magdalen Islands project, for their help acquiring datasets and for field logistics, John Molson, for proofreading, and finally the two anonymous reviewers for their valuable comments. The authors would also like to thank Vincent Post for discussions on deep open boreholes, and Francesca Lotti and John Doherty for discussions on seawater intrusion modeling and data assimilation. J-C Comte and O Banton acknowledge the financial support from the Fonds d'Action Québécois pour le Développement Durable for the ERT data collection, undertaken as part of the Madelin'Eau consortium (Ageos-Enviro'Puits-Hydriad), and further thank the Municipality of Les Îles-de-la-Madeleine for fieldwork logistical and technical support.

References

Anderson, M.P., Woessner, W.W., Hunt, R.J., 2015. Applied groundwater modeling: simulation of flow and advective transport. Academic press.

Ataie-Ashtiani, B., Rajabi, M.M., Ketabchi, H., 2013. Inverse modelling for freshwater lens in small islands: Kish Island, Persian Gulf. Hydrological Processes, 27(19): 2759-2773. DOI:10.1002/hyp.9411

Baalousha, H.M., 2016. The potential of using beach wells for reverse osmosis desalination in Qatar. Modeling Earth Systems and Environment, 2(2). DOI:10.1007/s40808-016-0151-5

Babu, R., Park, N., Yoon, S., Kula, T., 2018. Sharp Interface Approach for Regional and Well Scale Modeling of Small Island Freshwater Lens: Tongatapu Island. Water, 10(11). DOI:10.3390/w10111636

Baker, R.A., 2010. Locating nested monitoring wells to reduce model uncertainty for management of a multilayer coastal aquifer. Journal of Hydrologic Engineering, 15(10): 763-771.

Bakker, M. et al., 2020. FloPy v3.3.2. U.S. Geological Survey Software Release, 26 June 2020. DOI:http://dx.doi.org/10.5066/F7BK19FH

Bakker, M. et al., 2016. Scripting MODFLOW Model Development Using Python and FloPy. Ground Water, 54(5): 733-739. DOI:10.1111/gwat.12413

Bakker, M., Schaars, F., Hughes, J.D., Langevin, C.D., Dausman, A.M., 2013. Documentation of the seawater intrusion (SWI2) package for MODFLOW. US Geological Survey Techniques and Methods, Book, 6.

Brisebois, D., 1981. Lithostratigraphie des strates permo-carboniferes, de l'archipel des Iles de la Madeleine. Direction générale des énergies conventionnelles.

Brunner, P., Doherty, J., Simmons, C.T., 2012. Uncertainty assessment and implications for data acquisition in support of integrated hydrologic models. Water Resources Research, 48(7). DOI:10.1029/2011wr011342

Bureau d’audience publique sur l’environnement (BAPE), 2013. Les effets liés à l’exploration et l’exploitation des ressources naturelles sur les nappes phréatiques aux Îles-de-la-Madeleine, notamment ceux liés à l’exploration et l’exploitation gazière (Effects of natural resources exploration and exploitation on groundwater in the Magdalen Islands, including effects related to gas exploration and exploitation), Rapport d’enquête et d’audiences publiques.

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. DOI:10.1007/s10040-009-0524-1

Chaillou, G. et al., 2012. Synthèse de l’état des connaissances sur les eaux souterraines aux Îles-de-la-Madeleine - Impacts de l’exploration et de l’exploitation des ressources naturelles sur celles-ci (Summary of existing knowlegde on groundwater resources in the Magdalen Islands - Impacts of natural resource exploration and exploitation on these), Université du Québec à Rimouski, Département de biologie, chimie et géographie.

Chen, J., Raiche, A., 1998. Inverting AEM data using a damped eigenparameter method. Exploration Geophysics, 29(1): 128-132.

Christiansen, A.V., Auken, E., Sørensen, K., 2006. The transient electromagnetic method, Groundwater geophysics. Springer, pp. 179-225.

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). DOI:10.1029/2007gl029981

Dausman, A.M., Doherty, J., Langevin, C.D., Sukop, M.C., 2010a. Quantifying data worth toward reducing predictive uncertainty. Ground Water, 48(5): 729-40. DOI:10.1111/j.1745-6584.2010.00679.x

Dausman, A.M., Langevin, C., Bakker, M., Schaars, F., 2010b. A comparison between SWI and SEAWAT–the importance of dispersion, inversion and vertical anisotropy. Proceedings of SWIM, 21: 2010.

Delottier, H., Pryet, A., Dupuy, A., 2016. Why Should Practitioners be Concerned about Predictive Uncertainty of Groundwater Management Models? Water Resources Management, 31(1): 61-73. DOI:10.1007/s11269-016-1508-2

Dentoni, M., Deidda, R., Paniconi, C., Qahman, K., Lecca, G., 2014. A simulation/optimization study to assess seawater intrusion management strategies for the Gaza Strip coastal aquifer (Palestine). Hydrogeology Journal, 23(2): 249-264. DOI:10.1007/s10040-014-1214-1

Dessureault, R., Simard, G., 1970. Hydrogéologie des Îles de la Madeleine (Hydrogeology of the Magdalen Islands), Gouvernement du Québec, Ministère des Richesses naturelles, Service de l'hydrogéologie.

Dhar, A., Datta, B., 2009. Saltwater intrusion management of coastal aquifers. I: linked simulation-optimization. Journal of Hydrologic Engineering, 14(12): 1263-1272.

Doherty, J., 2004. PEST: Model-Independent Parameter Estimation User Manual, 3338. Watermark Numerical Computing, Brisbane, Australia, 393 pp.

Doherty, J., 2015. Calibration and uncertainty analysis for complex environmental models. Watermark Numerical Computing Brisbane, Australia.

Doherty, J., 2020. PEST_HP, PEST for highly parallelized computing environments. Watermark Numerical Computing.

Doherty, J., Moore, C., 2020. Decision Support Modeling: Data Assimilation, Uncertainty Quantification, and Strategic Abstraction. Ground Water, 58(3): 327-337. DOI:10.1111/gwat.12969

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

Dokou, Z., Karatzas, G.P., 2012. Saltwater intrusion estimation in a karstified coastal system using density-dependent modelling and comparison with the sharp-interface approach. Hydrological Sciences Journal, 57(5): 985-999. DOI:10.1080/02626667.2012.690070

Essaid, H.I., 1990. The computer model SHARP, a quasi-three-dimensional finite-difference model to simulate freshwater and saltwater flow in layered coastal aquifer systems, 90. Department of the Interior, US Geological Survey.

Fienen, M.N., Doherty, J.E., Hunt, R.J., Reeves, H.W., 2010. Using prediction uncertainty analysis to design hydrologic monitoring networks: example applications from the Great Lakes water availability pilot project, U. S. Geological Survey.

Freeze, R.A., Cherry, J.A., 1979. Groundwater.

Gelhar, L.W., Welty, C., Rehfeldt, K.R., 1992. A critical review of data on field-scale dispersion in aquifers. Water resources research, 28(7): 1955-1974.

Gingerich, S.B., 2002. Geohydrology and Numerical Simulation of Alternative Pumping Distributions and the Effects of Drought on the Ground-Water Flow System of Tinian, Commonwealth of the Northern Mariana Islands, US Department of the Interior, US Geological Survey.

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. DOI:10.1016/j.advwatres.2020.103695

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

Herckenrath, D. et al., 2013. Calibrating a salt water intrusion model with time-domain electromagnetic data. Ground Water, 51(3): 385-97. DOI:10.1111/j.1745-6584.2012.00974.x

Hill, M.C. et al., 2016. Practical Use of Computationally Frugal Model Analysis Methods. Ground Water, 54(2): 159-70. DOI:10.1111/gwat.12330

Hinnell, A.C. et al., 2010. Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion. Water Resources Research, 46(4). DOI:10.1029/2008wr007060

Holding, S., Allen, D.M., 2015. From days to decades: numerical modelling of freshwater lens response to climate change stressors on small low-lying islands. Hydrology and Earth System Sciences, 19(2): 933-949. DOI:10.5194/hess-19-933-2015

Hughes, I.G., Hase, T.P.A., 2010. Measurements and their uncertainties: a practical guide to modern error analysis. Oxford University Press.

Hughes, J.D., White, J.T., 2014. Hydrologic conditions in urban Miami-Dade County, Florida, and the effect of groundwater pumpage and increased sea level on canal leakage and regional groundwater flow.

Hunt, R.J., Fienen, M.N., White, J.T., 2020. Revisiting "An Exercise in Groundwater Model Calibration and Prediction" After 30 Years: Insights and New Directions. Ground Water, 58(2): 168-182. DOI:10.1111/gwat.12907

Jørgensen, F. et al., 2012. Transboundary geophysical mapping of geological elements and salinity distribution critical for the assessment of future sea water intrusion in response to sea level rise. Hydrology and Earth System Sciences, 16(7): 1845-1862. DOI:10.5194/hess-16-1845-2012

Konikow, L.F., Hornberger, G.Z., Halford, K.J., Hanson, R.T., 2009. Revised multi-node well (MNW2) package for MODFLOW ground-water flow model. US Geological Survey Techniques and Methods, 67.

Kopsiaftis, G., Christelis, V., Mantoglou, A., 2019. Comparison of Sharp Interface to Variable Density Models in Pumping Optimisation of Coastal Aquifers. Water Resources Management, 33(4): 1397-1409. DOI:10.1007/s11269-019-2194-7

Langevin, C.D., Thorne Jr, D.T., Dausman, A.M., Sukop, M.C., Guo, W., 2008. SEAWAT version 4: a computer program for simulation of multi-species solute and heat transport. 2328-7055, Geological Survey (US).

Leblanc, Y., 1994. Analyse et modélisation numérique de huit puits de production sur l'Île du Cap-aux-Meules, Îles-de-la-Madeleine (Analysis and Numerical Modeling of Eight

Production Wells on the Island of Cap-aux-Meules, Magdalen Islands), Université Laval, Quebec, Canada.

Lemieux, J.-M. et al., 2015. Simulating the impact of climate change on the groundwater resources of the Magdalen Islands, Québec, Canada. Journal of Hydrology: Regional Studies, 3: 400-423. DOI:10.1016/j.ejrh.2015.02.011

Llopis-Albert, C., Pulido-Velazquez, D., 2014. Discussion about the validity of sharp-interface models to deal with seawater intrusion in coastal aquifers. Hydrological Processes, 28(10): 3642-3654. DOI:10.1002/hyp.9908

Loke, M.H., Dahlin, T., 2002. A comparison of the Gauss–Newton and quasi-Newton methods in resistivity imaging inversion. Journal of Applied Geophysics, 49(3): 149-162. DOI:10.1016/s0926-9851(01)00106-9

Lu, C., Werner, A.D., 2013. Timescales of seawater intrusion and retreat. Advances in Water Resources, 59: 39-51. DOI:10.1016/j.advwatres.2013.05.005

Madelin'Eau, 2004. Gestion des eaux souterraines aux Îles-de-la-Madeleine, un défi de développement durable – Rapport final; délivré à la Municipalité des Iles-de-la-Madeleine (Groundwater management in the Magdalen Islands, a challenge for sustainable developement - Final report; delivered to the Municipality of the Magdalen Islands).

Meyer, R., Engesgaard, P., Sonnenborg, T.O., 2019. Origin and Dynamics of Saltwater Intrusion in a Regional Aquifer: Combining 3-D Saltwater Modeling With Geophysical and Geochemical Data. Water Resources Research, 55(3): 1792-1813. DOI:10.1029/2018wr023624

Nolan, B.T. et al., 2015. Data worth and prediction uncertainty for pesticide transport and fate models in Nebraska and Maryland, United States. Pest Manag Sci, 71(7): 972-85. DOI:10.1002/ps.3875

Pappa, A., Dokou, Z., Karatzas, G.P., 2017. Saltwater intrusion management using the SWI2 model: application in the coastal aquifer of Hersonissos, Crete, Greece. Desalination and Water Treatment, 99: 49-58. DOI:10.5004/dwt.2017.21550

Pool, M., Carrera, J., 2011. A correction factor to account for mixing in Ghyben-Herzberg and critical pumping rate approximations of seawater intrusion in coastal aquifers. Water Resources Research, 47(5). DOI:10.1029/2010wr010256

Post, V.E., 2012. Electrical conductivity as a proxy for groundwater density in coastal aquifers. Ground Water, 50(5): 785-92. DOI:10.1111/j.1745-6584.2011.00903.x

Post, V.E., Houben, G.J., van Engelen, J., 2018a. What is the Ghijben-Herzberg principle and who formulated it? Hydrogeology Journal, 26(6): 1801-1807.

Post, V.E.A., Banks, E., Brunke, M., 2018b. Groundwater flow in the transition zone between freshwater and saltwater: a field-based study and analysis of measurement errors. Hydrogeology Journal, 26(6): 1821-1838. DOI:10.1007/s10040-018-1725-2

Post, V.E.A., Bosserelle, A.L., Galvis, S.C., Sinclair, P.J., Werner, A.D., 2018c. On the resilience of small-island freshwater lenses: Evidence of the long-term impacts of groundwater abstraction on Bonriki Island, Kiribati. Journal of Hydrology, 564: 133-148. DOI:10.1016/j.jhydrol.2018.06.015

Poulin, M., 1977. Étude hydrogéologique des Îles de Grosse-Île et de Grande-Entrée Îles-la-Madeleine (Hydrogeological study of the Grosse-Île and Grande Entrée Islands, Magdalen islands), Service des eaux souterraines, Ministère des richesses naturelles, Gouvernement du Québec.

Rabeau, O., Thériault, R., 2013. Modélisation géologique 3D des Îles-de-la-Madeleine.

Raiche, A.P., Jupp, D.L.B., Rutter, H., Vozoff, K., 1985. The joint use of coincident loop transient electromagnetic and Schlumberger sounding to resolve layered structures. Geophysics, 50(10): 1618-1627. DOI:10.1190/1.1441851

Reilly, T.E., Goodman, A.S., 1985. Quantitative analysis of saltwater-freshwater relationships in groundwater systems—a historical perspective. Journal of Hydrology, 80(1-2): 125-160.

Rotzoll, K., 2010. Effects of groundwater withdrawal on borehole flow and salinity measured in deep monitor wells in Hawai'i-implications for groundwater management.

Rotzoll, K., Izuka, S.K., Nishikawa, T., Fienen, M.N., El-Kadi, A.I., 2016. Quantifying effects of humans and climate on groundwater resources of Hawaii through sharp-interface modeling. AGUFM, 2016: H23E-1594.

Rushton, K., 1980. Differing positions of saline interfaces in aquifers and observation boreholes. Journal of Hydrology, 48(1-2): 185-189.

Sanz, E., Voss, C.I., 2006. Inverse modeling for seawater intrusion in coastal aquifers: Insights about parameter sensitivities, variances, correlations and estimation procedures derived from the Henry problem. Advances in Water Resources, 29(3): 439-457. DOI:10.1016/j.advwatres.2005.05.014

Shalev, E. et al., 2009. Biased monitoring of fresh water-salt water mixing zone in coastal aquifers. Ground Water, 47(1): 49-56. DOI:10.1111/j.1745-6584.2008.00502.x

Shoemaker, W.B., 2004. Important observations and parameters for a salt water intrusion model. Ground Water, 42(6): 829-840.

Shuler, C.K., Mariner, K.E., 2020. Collaborative groundwater modeling: Open-source, cloud-based, applied science at a small-island water utility scale. Environmental Modelling & Software, 127. DOI:10.1016/j.envsoft.2020.104693

Siarkos, I., Latinopoulos, P., 2016. Modeling seawater intrusion in overexploited aquifers in the absence of sufficient data: application to the aquifer of Nea Moudania, northern Greece. Hydrogeology Journal, 24(8): 2123-2141. DOI:10.1007/s10040-016-1455-2

Statistics Canada, 2017. Les Îles-de-la-Madeleine, MÉ [Census subdivision], Quebec and Les Îles-de-la-Madeleine, TÉ [Census division], Quebec (table). Census Profile. 2016 Census, Statistics Canada Catalogue no. 98-316-X2016001, Ottawa.

Steklova, K., Haber, E., 2016. Joint hydrogeophysical inversion: state estimation for seawater intrusion models in 3D. Computational Geosciences, 21(1): 75-94. DOI:10.1007/s10596-016-9595-y

Sylvestre, M., 1979a. Carte hydrogéologique des Îles de la Madeleine (Hydrogeological map of the Magdalen Islands), Gouvernement du Québec, Ministère des Richesses naturelles, Direction générale des eaux. Service des eaux souterraines., Quebec.

Sylvestre, M., 1979b. Étude par modèle mathématique des nappes souterraines de la Grosse-Île et de l'île de Grande-Entrée, Îles-la-Madeleine (Study of the groundwater resources of the Grosse-Ile and Grande Entrée Islands using a mathematical model, Magdalen islands), Gouvernement du Québec, Ministère des Richesses naturelles, Direction générale des eaux. Service des eaux souterraines.

Thiem, G., 1906. Hydrologische Methoden: Dissertation zur Erlangung der Würde eines Doktor-Ingenieurs durch die Königliche Technische Hochschule zu Stuttgart, JM Gebhardt's verlag.

Walter, D.A., McCobb, T.D., Masterson, J.P., Fienen, M.N., 2016. Potential effects of sea-level rise on the depth to saturated sediments of the Sagamore and Monomoy flow lenses on Cape Cod, Massachusetts. 2016-5058, Reston, VA. DOI:10.3133/sir20165058

Werner, A.D. et al., 2013. Seawater intrusion processes, investigation and management: Recent advances and future challenges. Advances in Water Resources, 51: 3-26. DOI:10.1016/j.advwatres.2012.03.004

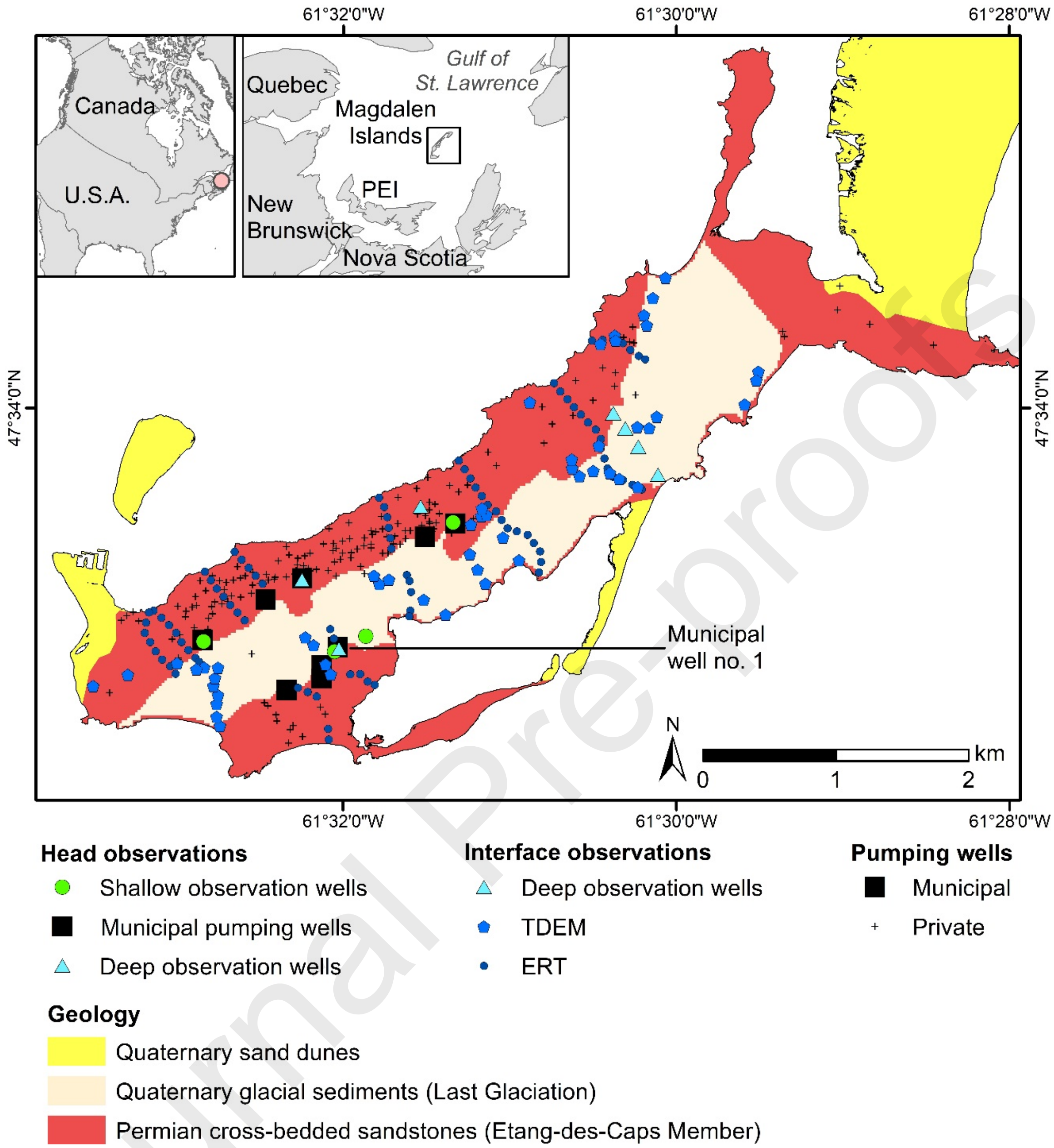
Werner, A.D., Sharp, H.K., Galvis, S.C., Post, V.E.A., Sinclair, P., 2017. Hydrogeology and management of freshwater lenses on atoll islands: Review of current knowledge and research needs. Journal of Hydrology, 551: 819-844. DOI:10.1016/j.jhydrol.2017.02.047

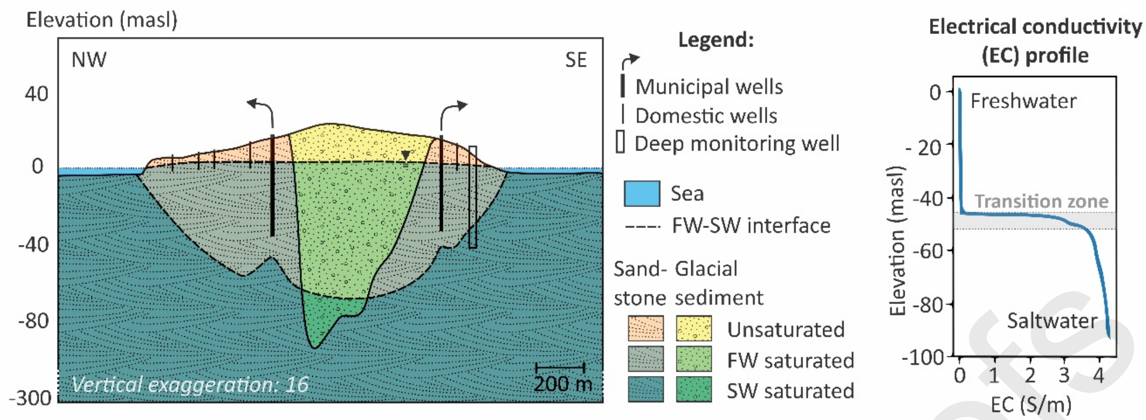
White, J.T., Fienen, M.N., Doherty, J.E., 2016. A python framework for environmental model uncertainty analysis. Environmental Modelling & Software, 85: 217-228. DOI:10.1016/j.envsoft.2016.08.017

White, J.T. et al., 2020. Toward Reproducible Environmental Modeling for Decision Support: A Worked Example. Frontiers in Earth Science, 8. DOI:10.3389/feart.2020.00050

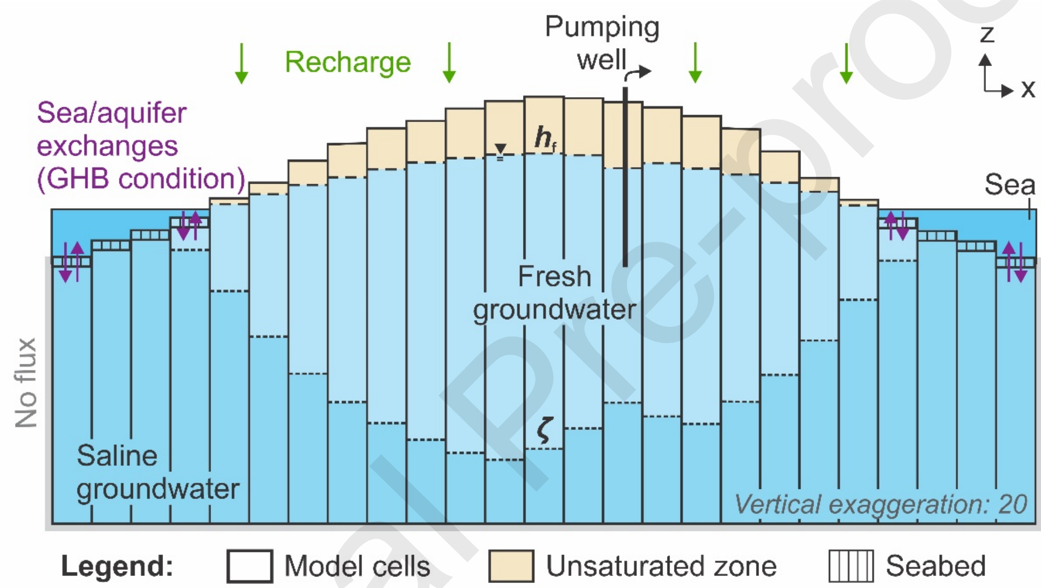
Authorship statement

Cécile Coulon: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, review & editing, Visualization. **Alexandre Pryet:** Conceptualization, Methodology, Software, Investigation, Resources, Writing – original draft, review & editing. **Jean-Michel Lemieux:** Conceptualization, Methodology, Investigation, Resources, Writing – original draft, review & editing, Supervision, Project administration, Funding acquisition. **Ble Jean Fidele Yrro:** Software, Formal analysis, Investigation. **Abderrezak Bouchedda:** Software, Formal analysis, Investigation, Writing – review & editing. **Erwan Gloaguen:** Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Jean-Christophe Comte:** Resources, Writing – review & editing. **J. Christian Dupuis:** Resources, Writing – review & editing, Supervision. **Olivier Banton:** Resources.

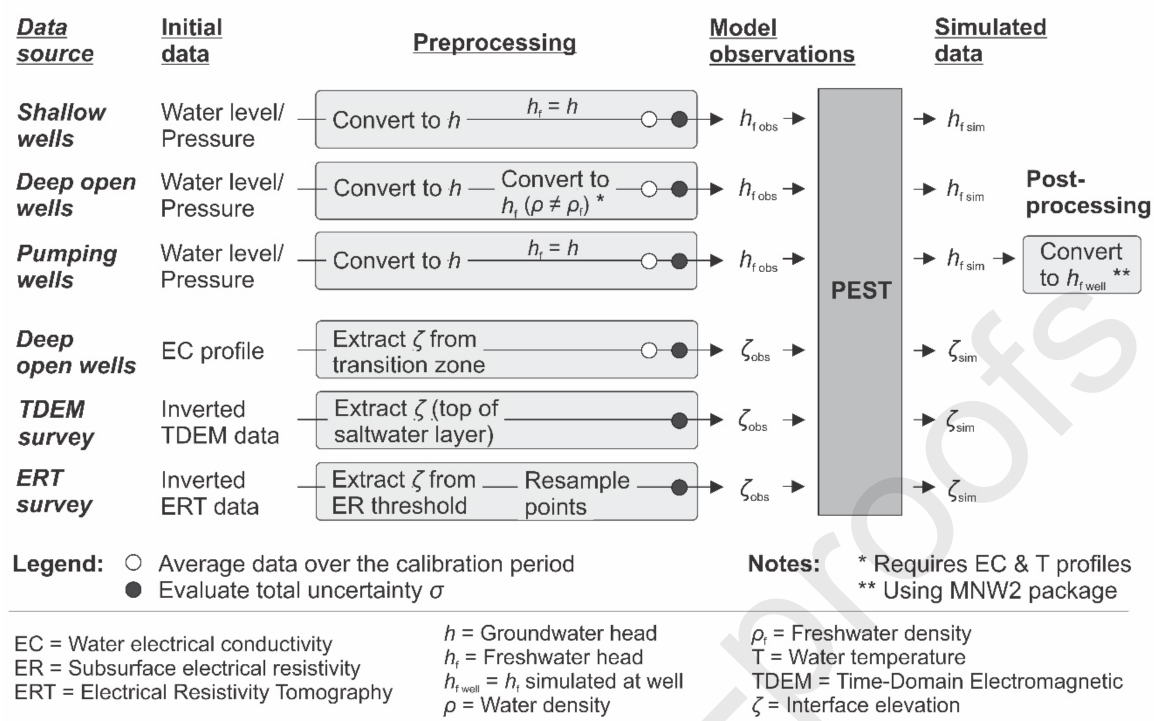

61°32'0"W
61°30'0"W
61°28'0"W
47°34'0"N
Canada
U.S.A.
Quebec
Gulf of St. Lawrence
Magdalen Islands
New Brunswick
PEI
Nova Scotia
Municipal well no. 1
N
km
0
1
2
Head observations
Shallow observation wells
Municipal pumping wells
Deep observation wells
Interface observations
Deep observation wells
TDEM
ERT
Pumping wells
Municipal
Private
Geology
Quaternary sand dunes
Quaternary glacial sediments (Last Glaciation)
Permian cross-bedded sandstones (Etang-des-Caps Member)



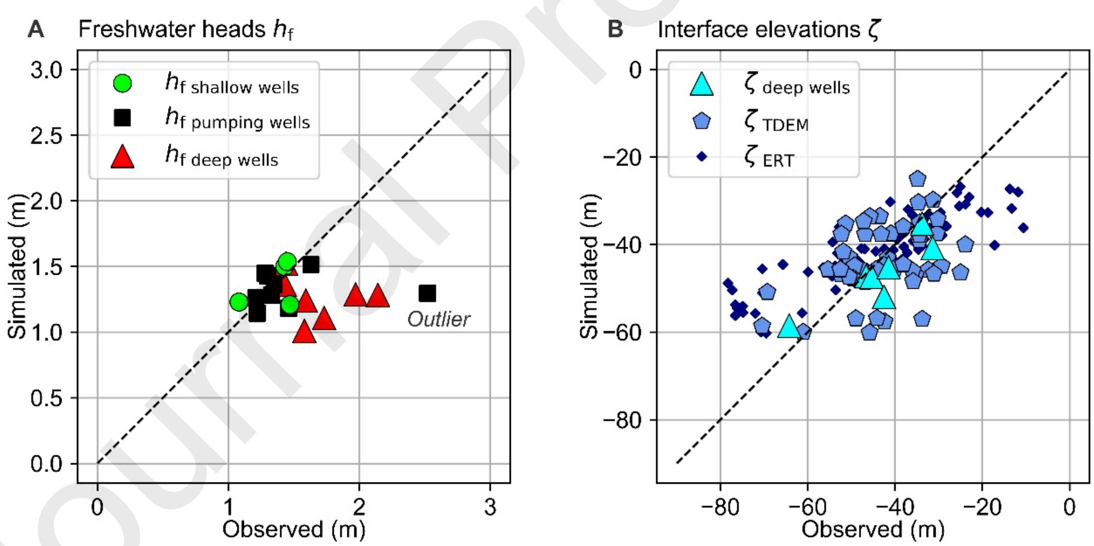
Elevation (masl)
NW
SE
40
0
-40
-80
-300
Vertical exaggeration: 16
200 m
Legend:
Municipal wells
Domestic wells
Deep monitoring well
Sea
FW-SW interface
Sand-stone
Glacial sediment
Unsaturated
FW saturated
SW saturated
Electrical conductivity (EC) profile
Freshwater
Transition zone
Saltwater
Elevation (masl)
0
- 20
-40
-60
-80
-100
0 1 2 3 4
EC (S/m)



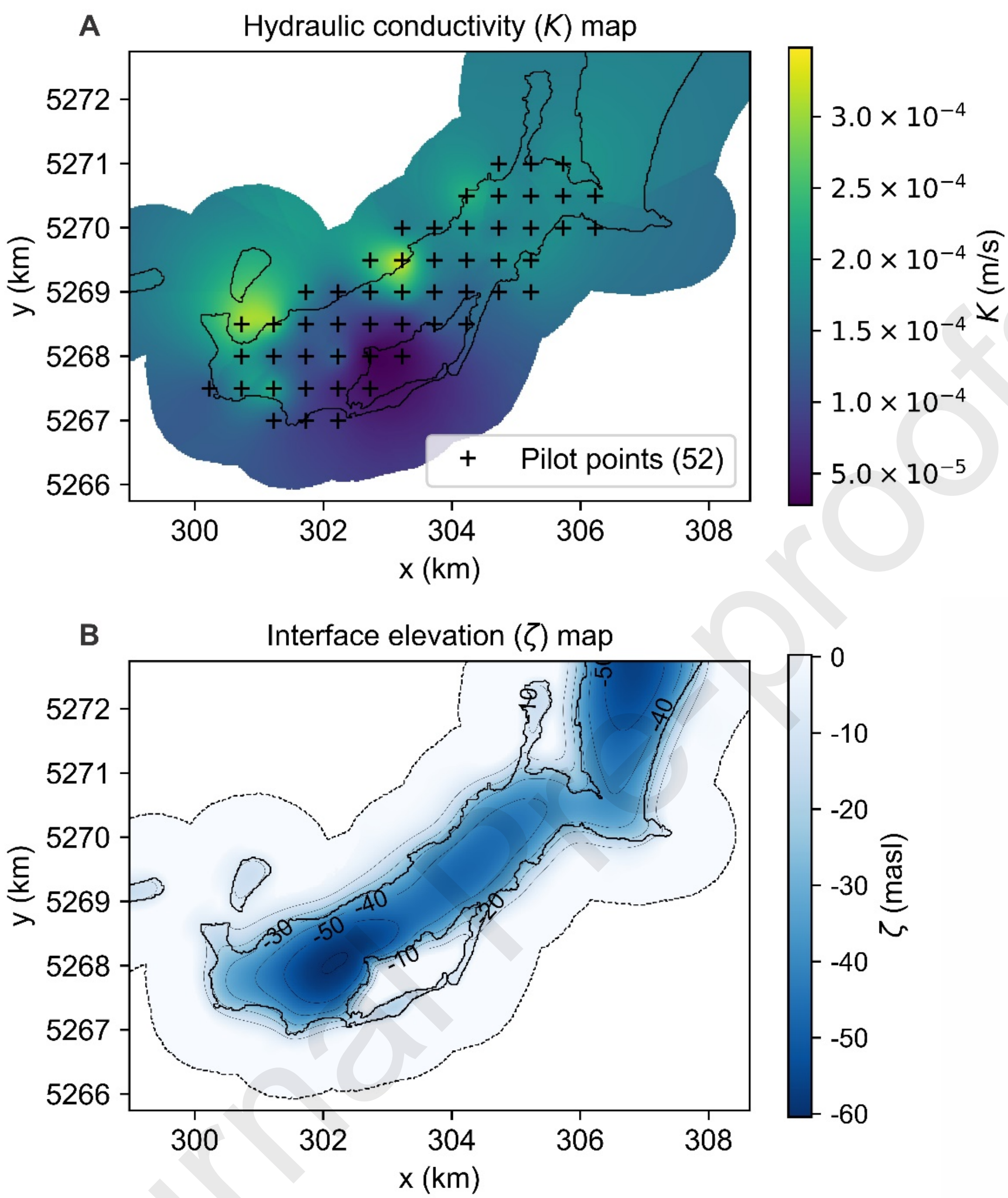
Recharge
Pumping well
z
x
Sea/aquifer exchanges (GHB condition)
h_f
Sea
Fresh groundwater
No flux
Saline groundwater
ζ
Vertical exaggeration: 20
Legend:
Model cells
Unsaturated zone
Seabed

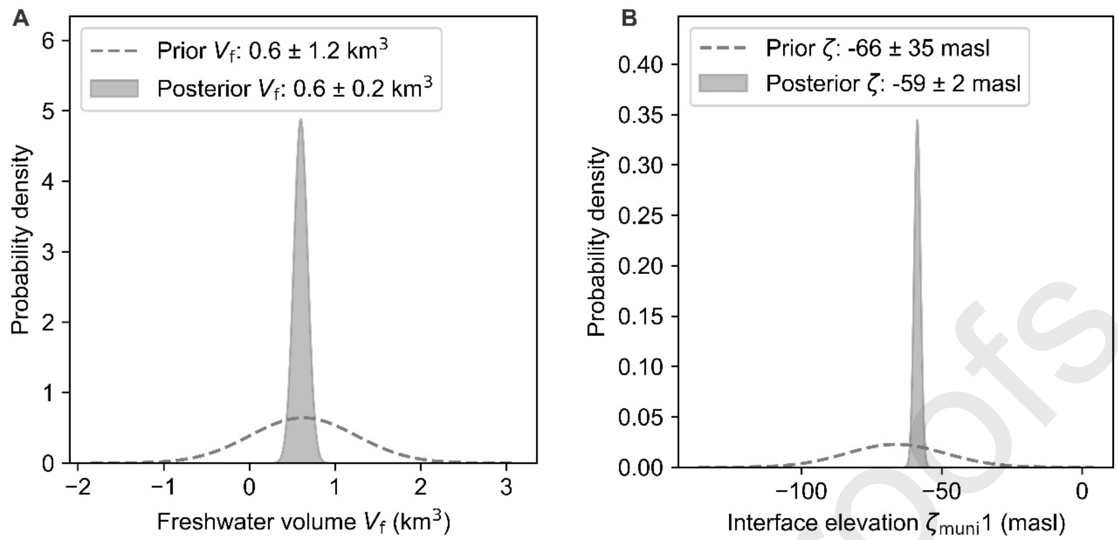
1119


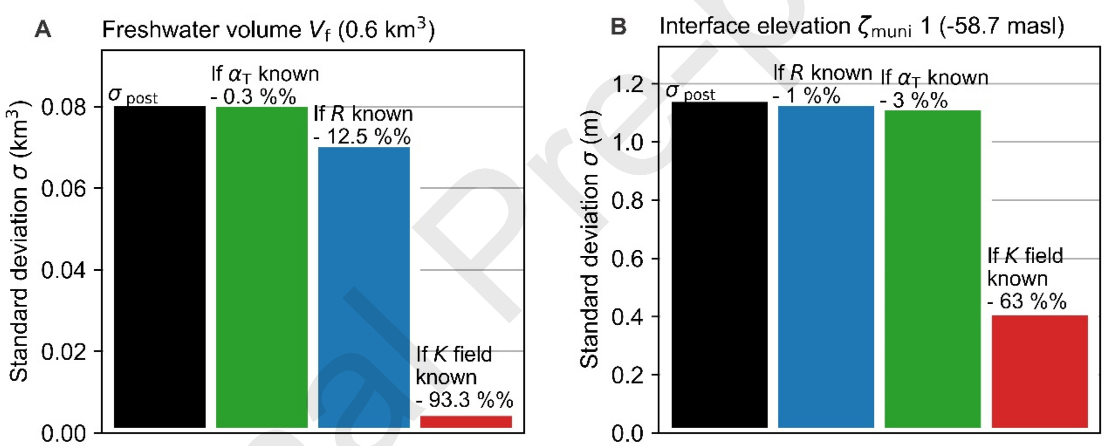
Data source
Initial data
Preprocessing
Model observations
Simulated data
Shallow wells
Water level/ Pressure
Convert to h
h_f = h
Deep open wells
Water level/ Pressure
Convert to h
Convert to h_f (ρ ≠ ρ_f) *
Pumping wells
Water level/ Pressure
Convert to h
h_f = h
Deep open wells
EC profile
Extract ζ from transition zone
TDEM survey
Inverted TDEM data
Extract ζ (top of saltwater layer)
ERT survey
Inverted ERT data
Extract ζ from ER threshold
Resample points
PEST
Post-processing
Convert to h_f well **
Legend:
Average data over the calibration period
Evaluate total uncertainty σ
Notes:
* Requires EC & T profiles
** Using MNW2 package
EC = Water electrical conductivity
ER = Subsurface electrical resistivity
ERT = Electrical Resistivity Tomography
h = Groundwater head
h_f = Freshwater head
h_f well = h_f simulated at well
ρ = Water density
ρ_f = Freshwater density
T = Water temperature
TDEM = Time-Domain Electromagnetic
ζ = Interface elevation

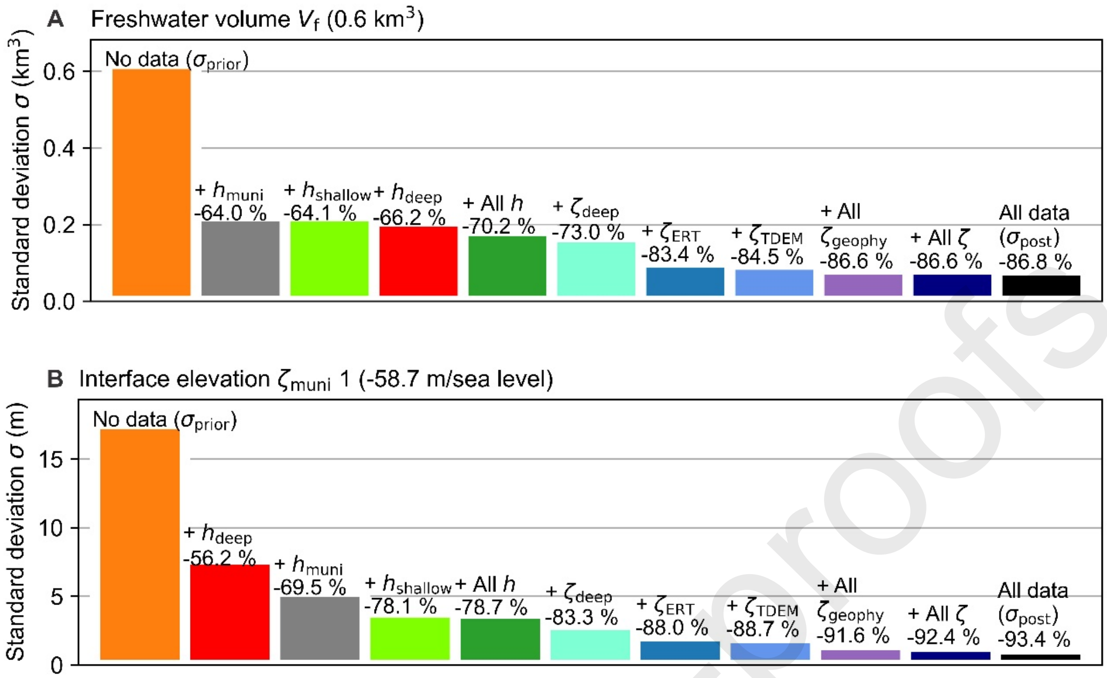
1120


A
Freshwater heads h_f
h_f shallow wells
h_f pumping wells
h_f deep wells
Outlier
Simulated (m)
Observed (m)
B
Interface elevations ζ
ζ deep wells
ζ TDEM
ζ ERT
Simulated (m)
Observed (m)

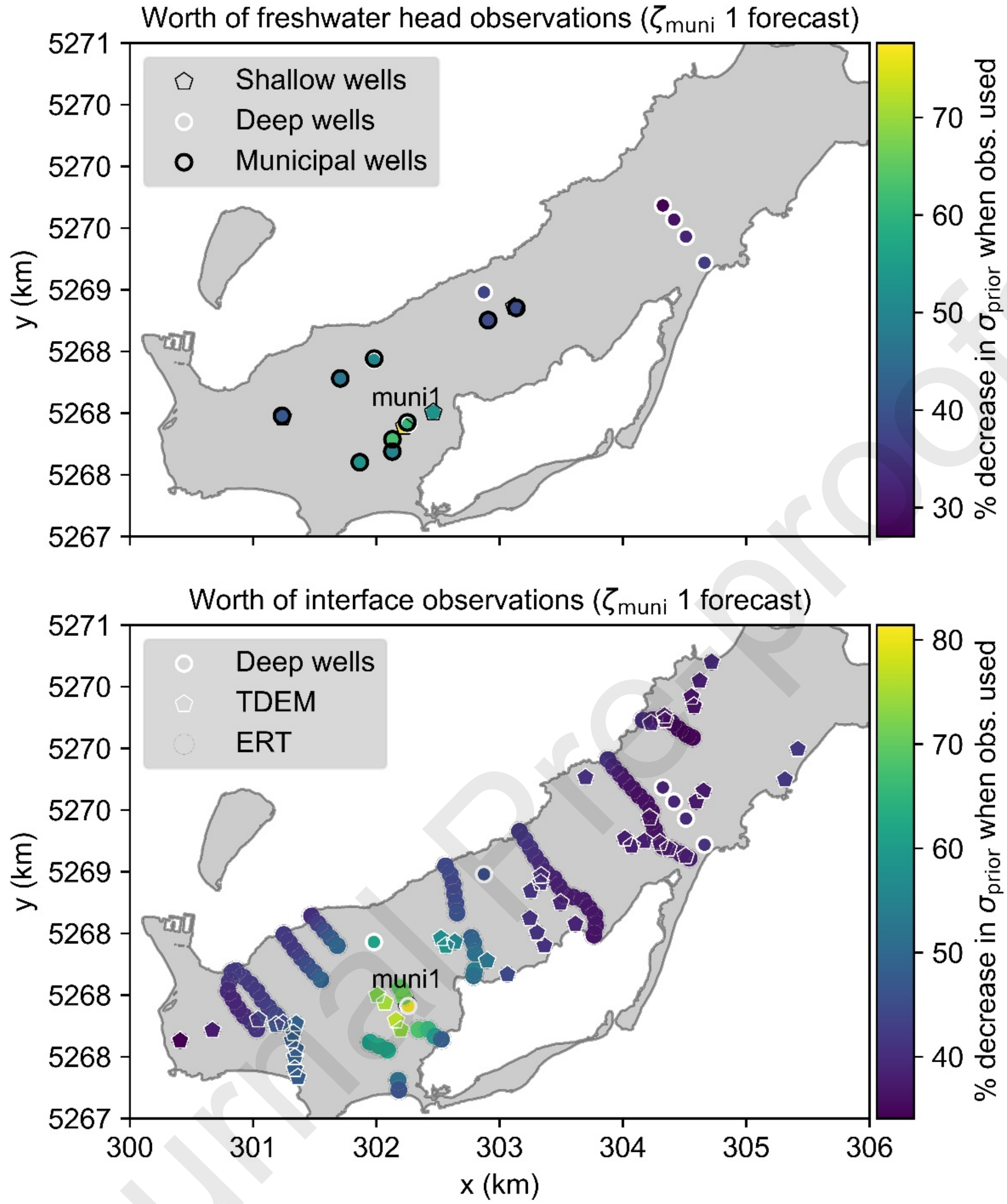

A
Hydraulic conductivity (K) map
y (km)
5272
5271
5270
5269
5268
5267
5266
300
302
304
306
308
x (km)
Pilot points (52)
3.0 × 10⁻⁴
2.5 × 10⁻⁴
2.0 × 10⁻⁴
1.5 × 10⁻⁴
1.0 × 10⁻⁴
5.0 × 10⁻⁵
K (m/s)
B
Interface elevation (ζ) map
-10
-50
-40
-30
-50
-40
-20
-10
0
-10
-20
-30
-40
-50
-60
ζ (masl)


A
Prior V_f: 0.6 ± 1.2 km^3
Posterior V_f: 0.6 ± 0.2 km^3
Probability density
Freshwater volume V_f (km^3)
B
Prior ζ: -66 ± 35 masl
Posterior ζ: -59 ± 2 masl
Probability density
Interface elevation ζ_{muni}1 (masl)


A
Freshwater volume V_f (0.6 km^3)
Standard deviation σ (km^3)
σ_{post}
If α_T known - 0.3 %%
If R known - 12.5 %%
If K field known - 93.3 %%
B
Interface elevation ζ_{muni} 1 (-58.7 masl)
Standard deviation σ (m)
σ_{post}
If R known - 1 %%
If α_T known - 3 %%
If K field known - 63 %%

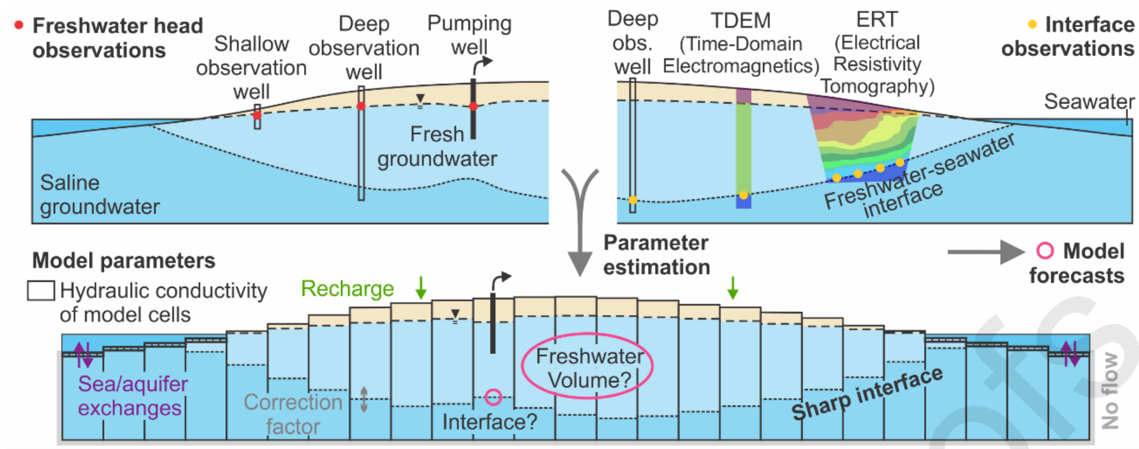

A Freshwater volume V_f (0.6 km^3)
Standard deviation σ (km^3)
No data (σ_{prior})
+ h_{muni} -64.0 %
+ $h_{shallow}$ -64.1 %
+ h_{deep} -66.2 %
+ All h -70.2 %
+ ζ_{deep} -73.0 %
+ ζ_{ERT} -83.4 %
+ ζ_{TDEM} -84.5 %
+ All ζ_{geophy} -86.6 %
+ All ζ -86.6 %
All data (σ_{post}) -86.8 %
B Interface elevation ζ_{muni} 1 (-58.7 m/sea level)
Standard deviation σ (m)
No data (σ_{prior})
+ h_{deep} -56.2 %
+ h_{muni} -69.5 %
+ $h_{shallow}$ -78.1 %
+ All h -78.7 %
+ ζ_{deep} -83.3 %
+ ζ_{ERT} -88.0 %
+ ζ_{TDEM} -88.7 %
+ All ζ_{geophy} -91.6 %
+ All ζ -92.4 %
All data (σ_{post}) -93.4 %

1124



Worth of freshwater head observations (ζ_{muni} 1 forecast)
Shallow wells
Deep wells
Municipal wells
muni1
y (km)
% decrease in σ_{prior} when obs. used
Worth of interface observations (ζ_{muni} 1 forecast)
Deep wells
TDEM
ERT
x (km)

1125



Highlights

- A regional seawater intrusion model was built using the SWI2 sharp-interface code
- Head and interface observations from wells and geophysical surveys were assimilated
- Fast run times enabled parameter estimation and linear-based uncertainty analysis
- Parameter estimation reduced the uncertainty of decision-support model forecasts
- Geophysical interface observations were essential to reduce predictive uncertainty