Direct prediction of spatially and temporally varying physical properties

from time-lapse electrical resistance data

Hermans Thomas, Stanford University, Department of Geological Sciences.

Oware Erasmus, The State University of New-York at Buffalo, Department of Geology

Caers Jef, Stanford University, Department of Geological Sciences.

Corresponding author

Hermans Thomas, Stanford University, Department of Geological Sciences, 450 Serra Mall,

Building 320, Room 118.

Phone : 1-650-7232223

E-mail : thermans@stanford.edu

Running head:

Direct forecast from TL resistance data

Keypoints:

1. We use time-lapse geophysical data to directly forecast physical property in the subsurface
2. We use a prior set of models to establish a linear relationship between data and forecast in a reduced dimension space

20   3. We estimate the full posterior distribution of the physical property with the established

21   linear relationship for observed data

Abstract

Time-lapse applications of electrical methods have grown significantly over the last decade. However, the quantitative interpretation of tomograms in terms of physical properties, such as salinity, temperature or saturation, remains difficult. In many applications, geophysical models are transformed into hydrological models, but this transformation suffers from spatially and temporally varying resolution resulting from the regularization used by the deterministic inversion. In this study, we investigate a prediction-focused approach (PFA) to directly estimate subsurface physical properties with electrical resistance data, circumventing the need for classic tomographic inversions. First, we generate a prior set of resistance data and physical property forecast through hydrogeological and geophysical simulations mimicking the field experiment. We reduce the dimension of both the data and the forecast through principal component analysis in order to keep the most informative part of both sets in a reduced dimension space. Then, we apply canonical correlation analysis to explore the relationship between the data and the forecast in their reduced dimension space. If a linear relationship can be established, the posterior distribution of the forecast can be directly sampled using a Gaussian process regression where the field data scores are the conditioning data. In this paper, we demonstrate PFA for various physical property distributions. We also develop a framework to propagate the estimated noise level in the reduced dimension space. We validate the results by a Monte Carlo study on the posterior distribution and demonstrate that PFA yields accurate uncertainty for the cases studied.

Keywords: electrical resistance tomography, time-lapse, prediction-focused approach, direct forecast, inversion

1. Introduction

Over the last few decades, hydrogeophysics has emerged as a new, interdisciplinary, field of sciences that aims at better understanding of hydrological processes through geophysical methods. Geophysical surveys are designed to acquire data related to the subsurface such as subsurface structures (e.g., conceptualization of the aquifer), hydraulic property spatial distributions (such as porosity and hydraulic conductivity) or the monitoring of physical properties of an aquifer (e.g., saturation, salinity or concentration, temperature) [e.g., *Rubin and Hubbard*, 2005; *Vereecken et al.*, 2007; *Binley et al.*, 2015; *Linde et al.*, 2015] .

Among the large number of hydrogeophysical methods (see *Binley et al.* [2015] for a review), the number of time-lapse applications of electrical methods has grown significantly (see *Singha et al.* [2015] for a review). The electrical resistivity of many fluid bearing sediments and rocks is largely dominated by the volumetric fluid content and the fluid electrical conductivity. Time-lapse electrical methods are therefore particularly well-suited for the monitoring of dynamic processes such as the evolution of saturation in the vadose zone [e.g., *Binley et al.*, 2002; *Koestel et al.*, 2008], solute transport [e.g., *Doetsch et al.,* 2012; *Robert et al.*, 2012], heat transport [e.g., *Hermans et al.,* 2014; *Arato et al.*, 2015], contaminant remediation operations [e.g., *Truex et al.*, 2013; *Johnson et al.*, 2015; *Masy et al.* 2015] or CO_2 sequestration [e.g., *Carrigan et al.*, 2013; *Auken et al.*, 2014]. The goal of hydrogeophysics is not just to quantify the change in resistivity caused by those processes, but also to infer the related change in physical properties such as water content and saturation, salinity or temperature. Knowing the spatio-temporal distribution of such physical properties is important for forecasting/managing purposes in water resources, agriculture or geothermal applications. Examples include the need to irrigate when the moisture content is too small, pumping control when salt water intrusion reaches the capture zone of a well, the adaptation of heat injection or extraction if the temperature reaches a given threshold in a geothermal system. In addition, regardless of any application, the quantification

3

of the physical properties constitutes a major and fundamental area of research within hydrogeophysics to study subsurface dynamics [*Singha et al.*, 2015].

The most common strategy is to transform geophysical parameters into physical or hydrologic properties through a petrophysical relationship from field or laboratory measurements [*Singha et al.*, 2015]. The geophysical parameters are most often obtained through a deterministic regularized inversion. However, the deterministic regularized approach used to invert geophysical data often yields a smeared distribution of the hydrologic or physical parameters. Indeed, the solution of the geophysical inverse problem is non-unique and results obtained from deterministic inversion sometimes lack geological realism [*Linde et al.*, 2015]. Many authors have observed poor recovery of mass-balance in tracing experiments [e.g., *Singha and Gorelick*, 2005; *Müller et al.*, 2010] and over- or underestimation of the physical properties due to over-smoothing of the geophysical image [e.g., *Vandenborght et al.*, 2005; *Hermans et al.*, 2015]. Some authors have proposed alternative constraints to the standard smoothness constraint in order to reduce smoothing and better estimate physical properties [*Doetsch et al.*, 2010; *Fiandaca et al.*, 2015; *Caterina et al.*, 2014; *Nguyen et al.*, 2016; *Hermans et al.*, 2016], but the inherent limitations of deterministic inversion remain.

One major issue with the classic deterministic inversion is the spatially varying resolution of the inversion method that prevents the direct use of petrophysical relationships at the field scale [*Day-Lewis et al.*, 2005] and strongly depends on the regularization operator. To overcome this limitation, *Moysey et al.* [2005] and *Singha and Moysey* [2006] proposed Monte Carlo simulations of numerical analogs of the monitoring experiment to derive a field-scale cell-by-cell petrophysical relationship, within the modeling grid. Similarly, *Oware et al.* [2013] used hydrological analog models to derive, through proper orthogonal decomposition (POD), a physically-based regularization operator able to reproduce physically plausible property distributions. An underlying limitation of deterministic approaches is their lack of proper

94   uncertainty quantification in the distribution of geophysical parameters, despite that the solution

95   is known to be non-unique. *Audebert et al.* [2014] developed a method based on multiple

96   deterministic inversions with various regularization parameters and clustering to better

97   delineate zone(s) affected by resistivity changes; but their uncertainty concerns only the zone

98   affected by resistivity changes and not their amplitude.

99   The issue of variable resolution also extends to the time dimension. Time-lapse inversion

100   schemes with temporal regularization constraints (4D deterministic inversion) have been

101   developed by *Kim et al.* [2009] and *Karaoulis et al.* [2011]. Even if they generally ensure

102   coherent variations of geophysical parameters with time, they still suffer from the limitation

103   arising from regularization.

104   Another approach is the use of summary statistics of the physical property (temporal and spatial

105   moments) which may be sufficient to identify advection and dispersion effects [e.g., *Kemna et*

106   *al.*, 2002; *Vandenborght et al.*, 2005; *Koestel et al.*, 2008]. However, the regularization in the

107   inversion process still influences the estimation of spatial moments [*Day-Lewis et al.*, 2007].

108   The inverse problem may be formulated in terms of plume geometry [*Ben Hadj Miled and*

109   *Miller*, 2009; *Pidlisecky et al.*, 2011; *Laloy et al.*, 2012]. In this case, the number of model

110   parameters is reduced by using the first spatial moments of the geophysical parameter

111   distribution, making the inverse problem overdetermined and allowing the efficient use of

112   Markov chain Monte Carlo (McMC) methods [*Laloy et al.*, 2012]. However, the limited number

113   of spatial moments used for inversion limits the accuracy of the inverted model in case of

114   complex geological distributions of the physical properties.

115   The estimated geophysical images or spatial moments, with their inherent limitations, may

116   subsequently be used, possibly with a petrophysical transformation, to calibrate a

117   hydrogeological model [e.g., *Binley et al.*, 2002; *Briggs et al.*, 2013]. The modeler may account

for the resolution limitations of tomograms by adequately weighting the geophysics-derived information [*Hermans et al.*, 2012] in order to limit the influence of imperfect estimates of physical quantities in the calibration process. For example, filtering schemes have been proposed to account for the varying resolution of tomograms [*Caterina et al.*, 2013] and applied in hydrogeophysical inversion [*Beaujean et al.*, 2014].

An alternative approach is to seek correlation in time between geophysical and hydrological data. Time series and spectral analysis can provide quantitative information on dynamic processes occurring in aquifers [*Johnson et al.*, 2012; *Wallin et al.*, 2013]. However, the method still suffers from the same limitations regarding the inversion itself, such as variable resolution and lack of (hydro)geological realism.

Finally, coupled inversion approaches have gained popularity in the last few years [e.g., *Hinnell et al.*, 2010; *Irving and Singha*, 2010]. In this procedure, geophysical data are directly included in the hydrologic inversion process, suppressing the need for a geophysical inversion and post-inversion petrophysical transform. The hydrologic model output is used as input for geophysical forward modeling through a petrophysical relationship and the corresponding geophysical response is compared to the observed data. Coupled inversion approaches are efficient when only a few hydrologic parameters are calibrated but might be difficult to apply in complex and heterogeneous cases [*Christensen et al., 2016*], because it strongly depends on the hydrological conceptual model itself and its inherent assumptions (boundary conditions, prior). The method also becomes computationally demanding when used within a stochastic framework to quantify uncertainty [*Irving and Singha*, 2010].

In this paper, we propose a prediction-focused approach (PFA) [e.g., *Satija and Caers*, 2015; *Scheidt et al., 2015*] to forecast the spatio-temporal change in physical properties using the changes in resistance collected during time-lapse electrical resistivity tomography (ERT),

142   without explicitly inverting for the resistivity distribution. Note that in this paper we keep the

143   terminology "forecast" used in previous studies even if the physical property is not estimated

144   in the future but simultaneously to the data. Similarly to *Moysey et al.* [2005] and *Oware et al.*

145   [2013], we use analog or prior models to simulate the hydrogeophysical experiment. The use

146   of aquifer analogs ensures geological realism in the approach, which is essential to produce

147   physically-plausible results [*Linde et al.*, 2015]. However, instead of using a regularization-

148   based framework, we aim to directly derive the full posterior distribution of physical properties

149   from time-lapse ERT, without any iterative inversion. To achieve this, we establish a direct

150   multivariate statistical model between the data variables (changes in resistance) and the forecast

151   variables (variations in physical property) in the section where data were collected. We then

152   use the statistical model to establish a full posterior uncertainty quantification of the forecast

153   given the actual observed data. Although our method is Bayesian, it differs from previous

154   Bayesian methods developed in hydrogeophysics and hydrogeology [e.g., *Rubin et al.*, 2010;

155   *Murakami et al.*, 2010, *Chen et al.*, 2012; *Linde et al.*, 2015]. In PFA, the likelihood is computed

156   in a reduced dimension statistical framework. Moreover, the likelihood is not formulated based

157   on the model parameters but based on the target (physical) variables. PFA directly builds a joint

158   bivariate distribution between data and forecast (target variables) in reduced dimensional space.

159   From the joint distribution, posterior distribution of forecast given data can then be readily

160   sampled, without additional forward model evaluations or inversion.

161   *Satija and Caers* [2015] applied the PFA to forecast the contamination in a pumping well from

162   concentration measurements upstream from this well, i.e. data and forecast were punctual time

163   series of the same property. They showed that PFA yields results similar to the rejection

164   sampler. In this work, we extend the PFA to the use of geophysical data to forecast subsurface

165   physical properties. We also demonstrate that the method is successful to forecast complex

166   spatio-temporal series. A topic not covered by *Satija and Caers* [2015] is the propagation of

167 noise through PFA. This issue is an important concern for geophysical data sets. We propose a

168 specific framework based on Monte Carlo simulations to prevent erroneous estimates due to

169 not accounting for the proper noise propagation. With several case examples, we demonstrate

170 the robustness, methodological and computational simplicity of the approach.

171 **2. Methods**

172 **2.1. Overview of the methodology**

173 The principle of the PFA is to directly estimate the posterior distribution of the physical

174 property $\boldsymbol{h}$ of interest $f_H(\boldsymbol{h}|\boldsymbol{d_{obs}})$ without explicitly transforming the observed geophysical

175 data $\boldsymbol{d_{obs}}$ into a distribution of geophysical parameters. We therefore avoid inverting the time-

176 lapse geophysical data sets and using a petrophysical transform *after* inversion that can be

177 space-dependent and uncertain [*Singha et al.*, 2015]. The key of the PFA is to work in a reduced

178 dimension space to establish a statistical model between geophysical data and physical

179 properties to be able to estimate the full posterior distribution of the forecast for some observed

180 data. The PFA developed for this specific case involves six steps:

181 1) generation of prior sets of data and forecast, sampled from the prior distribution,

182 2) dimension reduction of the data and forecast variables using principal component analysis,

183 3) canonical correlation analysis to establish a multivariate correlation between data and

184 forecast variables,

185 4) Gaussian process regression (kriging) with the observed data to derive the conditional mean

186 and covariance matrix of the forecast variables (in the canonical component space),

187 5) sampling the Gaussian model,

188 6) back-transform of the samples into the original high-dimension space.

189 Steps 1 to 3 correspond to the training stage, where the relationship between data and forecast

190 in the reduced dimension space is learned from forward simulations on the prior set of models.

191 Steps 4 to 6 correspond to the forecast stage, where the full posterior distribution of the

192 predictions is deduced from the observed data.

193 **2.2. Generating a prior set of models**

194 Establishing a statistical model between data and forecast requires knowing both for some

195 known subsurface realizations. A realization is the distribution of subsurface parameters

196 required to simulate data and forecast. We generate a prior set of n realizations of the

197 subsurface $\boldsymbol{m}$, here through geostatistical modeling, and simulate (equation 1) the model

198 response $\boldsymbol{h}$ corresponding to the monitored process, here subsurface flow and transport

199
$$\boldsymbol{h} = g_H(\boldsymbol{m}) \qquad (1)$$

200 where g_H is the physics-based forward model operator. The model response $\boldsymbol{h}$ is generally the

201 evolution of a given physical property at l different locations at t different time-steps in the

202 subsurface. In the prior set, $\boldsymbol{h}$ is therefore a high dimension array of n realizations of $l \times t$

203 variables and constitutes the prior distribution of the forecast.

204 Then, for each realization of this prior, the corresponding geophysical data $\boldsymbol{d}$ are computed

205 using a geophysical forward model g_D

206
$$\boldsymbol{d} = g_D(\boldsymbol{h}, \boldsymbol{m}) \qquad (2)$$

207 For ERT, the data are changes in resistance at t different time steps (subtraction with the

208 background) for u different electrode configurations. In the prior set, $\boldsymbol{d}$ is therefore an array of

209 n realizations of $u \times t$ variables, also high dimensional.

210 Since $\boldsymbol{h}$ is a physical property and not a geophysical parameter, a petrophysical relationship

211 linking both is used in g_D. In time-lapse studies, the relations between the change of physical

212 properties and geophysical parameters are relatively well-known and more straightforward

213 [*Linde et al.*, 2015]. The petrophysical relationship can vary depending on the values of $\boldsymbol{m}$,

214 which is therefore a parameter of equation 2. Note that the PFA in general remains valid if data

215 and forecast are not co-located and not simultaneous (the dimension t can be different for data

216 and forecast), i.e. the methodology can be applied to a broader range of problems than time-

217 lapse geophysical inversion [*Satija and Caers*, 2015].

218 **2.3. Dimension reduction of the data and forecast using principal component analysis**

219 Principal component analysis (PCA) is a multivariate analysis technique that aims at

220 transforming a set of observed variables into a set of linearly uncorrelated variables, called

221 principal components [*Krzanowski*, 2000]. Typically, the first k components of the PCA

222 decomposition, with $k \ll n$, explain almost all the variance of the data set. By keeping only those

223 k first dimensions, one thus achieves an efficient dimension reduction. PCA is based on the

224 eigen-decomposition of a matrix and has been applied on training images by *Oware et al.* [2013]

225 to derive a physically-based operator for the inversion of ERT data. The method is also at the

226 basis of truncated singular value decomposition which is widely applied to regularize under-

227 determined inverse problems as those encountered in hydrogeology [e.g. *Tonkin and Doherty*,

228 2005].

229 We first apply independently PCA on both the data and forecast variables to reduce their

230 dimension to p and q, respectively. The reduced data and forecast variables after PCA are

231 denoted by $\boldsymbol{d^f}$ (p x 1) and $\boldsymbol{h^f}$ (q x 1).

232 **2.4. Linearization using canonical correlation analysis between data and forecast**

233   The relationship between d^f and h^f is often non-linear, due to the petrophysical relationships

234   and underlying partial differential equations. Canonical correlation analysis (CCA) is a

235   multivariate analysis method that can be used to transform the relationships between pairs of

236   vector variables with n replicates into a set of independent linearized relationships between

237   pairs of scalar [*Krzanowski*, 2000]. Mathematically, CCA finds the coefficients in a_i (p x 1)

238   and b_i (q x 1) with $i = 1 \dots m$ that maximize the correlation between the linear combinations

239    $a_i^T d^f$ and $b_i^T h^f$. The resulting linear combinations are denoted by d_i^c and h_i^c, and called the

240   canonical variates of d^f and h^f. The coefficient of correlation between their n scalar values is

241   called the canonical correlation. The vectors a_i (p x 1) and b_i (q x 1) are orthogonal to each

242   other. The maximum number of canonical variate pairs m that can be found is the minimum of

243   the ranks of d^f and h^f. Since p and q are generally much smaller than n, m corresponds to the

244   minimum of p and q. The canonical transformation yields two reduced data variables d^c and

245   h^c, such that

$$d^c = d^f A^T \qquad (3)$$

246

$$h^c = h^f B^T \qquad (4)$$

247

248   with $A = [a_1 \ \dots \ a_m]$ and $B = [b_1 \ \dots \ b_m]$.

249   The canonical transformation can be found through the eigen-decomposition of the sample

250   covariance matrix [*Krzanowski*, 2000]. If $p > q$, i.e. more dimensions are kept in the data, this

251   transformation is reversible

$$h^f = B^{-1} h^{cT}, \qquad (5)$$

252

253   This back transformation to the original space is essential for the last step of the process.

254   **2.5. Gaussian process regression**

255 We use Bayes' rule in formulating the problem of estimating the posterior distribution of $\boldsymbol{h}$ for

256 some observed data $\boldsymbol{d_{obs}}$

$$f_H(\boldsymbol{h}|\boldsymbol{d_{obs}}) = kf_H(\boldsymbol{h})L(\boldsymbol{h}|\boldsymbol{d_{obs}}) \qquad (6)$$

258 where $f_H(\boldsymbol{h})$ is the prior distribution, $L(\boldsymbol{h}|\boldsymbol{d_{obs}})$ is a likelihood function and k is a

259 proportionality constant. PCA and CCA enable reducing a set of high-dimensional variables

260 $(\boldsymbol{d},\boldsymbol{h})$ to another set of much reduced dimension variables$(\boldsymbol{d^c}, \boldsymbol{h^c})$, linearly related and whose

261 axes are orthogonal. By back transformation of $\boldsymbol{h^c}$ in the original space (equation 5), the

262 forecasting problem (equation 6) is tantamount to forecasting $\boldsymbol{h^c}$ from $\boldsymbol{d^c_{obs}}$. Our goal is thus to

263 obtain the posterior distribution

$$f_H(\boldsymbol{h^c}|\boldsymbol{d^c_{obs}}) = kf_H(\boldsymbol{h^c})L(\boldsymbol{h^c}|\boldsymbol{d^c_{obs}}) \qquad (7)$$

265 If the canonical correlation of the first components of those reduced dimension variables is

266 large enough, i.e., that a linear relationship between the respective components of $\boldsymbol{d^c}$ and $\boldsymbol{h^c}$

267 exists, then a Gaussian process regression can be used to solve the problem of estimating $\boldsymbol{h^c}$

268 from $\boldsymbol{d^c_{obs}}$ [*Tarantola*, 2005].

269 The framework described by *Tarantola* [2005] assumes a multivariate Gaussian prior

270 distribution of $\boldsymbol{h^c}$, with the mean $\overline{\boldsymbol{h^c}}$ and covariance $\boldsymbol{C_{Hc}}$ estimated from the sample

271 distribution. If the marginal distributions of $\boldsymbol{h^c}$ are not Gaussian, then a histogram transform

272 needs to be performed. We now model the relationship between $\boldsymbol{d^c}$ and $\boldsymbol{h^c}$ through the linear

273 forward model operator $\boldsymbol{G}$

$$\boldsymbol{d^c} = \boldsymbol{G}\boldsymbol{h^c} \qquad (8)$$

275   Because the data $\boldsymbol{d_{obs}}$ may be contaminated with noise, its reduced dimension transform $\boldsymbol{d_{obs}^c}$

276   has observational error $\boldsymbol{C_c^d}$. We will see in the next section how we can model $\boldsymbol{C_c^d}$. The

277   likelihood (in equation 7) is assumed to be multivariate Gaussian

$$L(\boldsymbol{h^c}|\boldsymbol{d_{obs}}) = \exp\left(-\tfrac{1}{2}(\boldsymbol{Gh^c} - \boldsymbol{d_{obs}^c})^T \boldsymbol{C_c^{d}}^{-1}(\boldsymbol{Gh^c} - \boldsymbol{d_{obs}^c})\right) \qquad (9)$$

279   hence, the posterior distribution is also multivariate Gaussian. The conditional mean and

280   covariance have the following analytical expressions:

$$\widetilde{\boldsymbol{h_c}} = \left(\boldsymbol{G^T} \boldsymbol{C_c^{d}}^{-1}\boldsymbol{G} + \boldsymbol{C_{Hc}^{-1}}\right)^{-1}\left(\boldsymbol{G^T} \boldsymbol{C_c^{d}}^{-1}\boldsymbol{d_{obs}^c} + \boldsymbol{C_{Hc}^{-1}}\overline{\boldsymbol{h^c}}\right) \qquad (10)$$

$$\widetilde{\boldsymbol{C_H^c}} = \boldsymbol{G^T} \boldsymbol{C_c^{d}}^{-1}\boldsymbol{A} + \boldsymbol{C_{Hc}^{-1}} \qquad (11)$$

283   Similarly to *Hansen et al.* [2014] and *Satija and Caers* [2015], we use a multivariate Gaussian

284   error model obtained from the residual error in the linear regression (equation 8) for the prior

285   set to account for the non-perfectly linear relationship between $\boldsymbol{d^c}$ and $\boldsymbol{h^c}$ generally observed

286   for complex systems. The solution in (10) and (11) is modified by adding the covariance term

287   of the error model $\boldsymbol{C_T}$ to the data observational error $\boldsymbol{C_c^d}$, and the mean error $\overline{\boldsymbol{\epsilon}}$ to $\boldsymbol{d_{obs}^c}$. We refer

288   to *Satija and Caers* [2015] for a detailed description of the Gaussian error model.

289   An obvious advantage of the methodology is that a Gaussian distribution is uniquely defined

290   by its mean and covariance (equations 10 and 11), and that sampling a Gaussian distribution is

291   straightforward. It is therefore simple to generate *s* samples $\{\boldsymbol{h_1^c}, \boldsymbol{h_2^c}, \dots, \boldsymbol{h_s^c}\}$ of the posterior

292   distribution $f_H(\boldsymbol{h^c}|\boldsymbol{d_{obs}})$, each sample being composed of *q* components. By back

293   transformation of CCA (equation 5), the *q* components of $\boldsymbol{h^f}$ are retrieved for the *s* samples.

294   Then, those *q* components (first *q* principal components) are used to estimate *s* samples of the

295   original high-dimension physical property $\boldsymbol{h}$, by multiplying the reduced components by the

296   inverse matrix of the eigenvectors of the PCA decomposition. Since the vectors are

297 orthonormal, the inverse matrix is equal to its transpose. Those s samples are then used to

298 estimate the statistical properties of the posterior distribution $f_H(\boldsymbol{h}|\boldsymbol{d}_{obs})$.

299 **2.6. Noise correlation in PCA space**

300 Solving equations 10 and 11 requires the definition of the covariance matrix of observational

301 error in the canonical space $\boldsymbol{C}_c^d$. The canonical space is related to the principal component space

302 through linear relationships (equation 3), and therefore $\boldsymbol{C}_c^d$ is linearly related to the error

303 covariance in principal component space $\boldsymbol{C}_f^d$ through

$$304 \qquad \boldsymbol{C}_c^d = \boldsymbol{A}\boldsymbol{C}_f^d\boldsymbol{A}^T \qquad (12)$$

305 The error covariance in the principal component space expresses the error that is made on the

306 estimation of the principal component scores due to the noise on the data. If we express the

307 observed data $\boldsymbol{d}_{obs}$ as a function of the true data $\boldsymbol{d}_{true}$ corrupted by some noise perturbation $\boldsymbol{\epsilon}$

$$308 \qquad \boldsymbol{d}_{obs} = \boldsymbol{d}_{true} + \boldsymbol{\epsilon} \qquad (13)$$

309 the error on the estimation of principal components can be similarly expressed as

$$310 \qquad \boldsymbol{\epsilon}^f = \boldsymbol{d}^f_{obs} - \boldsymbol{d}^f_{true} \qquad (14)$$

311 However, the noise observed in the data cannot be easily transformed in the principal

312 component space. Indeed, the different electrode configurations are affected to various degrees

313 by the noise, depending on the strength of the measured signal. The variable transformation in

314 PCA will also affect differently the principal components which have decreasing variance

315 contributions and are affected differently by noise. Hence obtaining some analytical expression

316 for $\boldsymbol{C}_f^d$ is unfeasible. We here propose to use Monte Carlo experiments to estimate the effect of

317 noise in the principal and canonical component space:

1) we use established methods to estimate an error model ϵ (equation 13), such as reciprocal measurements in electrical methods [*LaBrecque et al.*, 1996; *Flores Orozco et al.*, 2012];

2) for the *first* of the *n* samples of the prior set, we replace the noise free data set by noisy data, with a level of noise ϵ (equation 13);

3) we perform PCA on the prior set with the *first* sample contaminated by noise;

4) we compare the principal components of the *first* sample for noisy and noise free data set to deduce ϵ_1^f (equation 14);

5) we repeat step 2 to 4 for all other *n-1* samples of the prior set;

6) we use the obtained set of errors in principal component space $\{\epsilon_1^f \ \epsilon_2^f \ ... \ \epsilon_n^f\}$ to estimate C_f^d;

7) we estimate C_c^d (equation 12).

The estimated covariance matrix in the canonical space can now be used in equations 10 and 11 to solve the Gaussian process regression.

3. Applications

In this section, we demonstrate the ability of PFA in a multi-time step example simulating a heat tracer experiment. We will test the robustness of the method for noisy data and various distributions of the reference forecast. For all examples, from *n* realizations (see equations (1) and (2)), we choose one as the reference solution and the corresponding data as the actual observed data d_{obs}. We forecast the posterior distribution after which we can compare the results with the reference solution.

A comparison between PFA and POD [*Oware et al.*, 2013] is presented in the supporting information files (Text S1 and Figure S1).

3.1. Multi time-steps thermal tracer experiment

The methodology is applied here for forecasting a heat tracer experiment. The heat tracer experiment mimics the injection of heated water ($\Delta T = 40°C$) into a confined alluvial aquifer for 24 hours at a rate of 3 m^3/h. The tracer flows in the aquifer under forced gradient conditions, induced by pumping at 30 m^3/h in a well located at 20 m from the injection. At 8 m from the injection well, centered on the direction linking injection and pumping well, a cross-borehole ERT panel is considered. The distance between the two boreholes is 4.5 m (Figure 1A). The boreholes are equipped with 13 electrodes spaced every 50 cm from 0.5 to 6.5 m deep, covering the alluvial deposits (Figure 1B).

3.2. Prior set

The alluvial deposits are represented by two facies representing sand lenses (30%) in a gravel (70 %) background. The sand lenses are represented by ellipsoids elongated in the direction of flow with a maximum length of 5 m, a maximum width of 3 m, and a thickness of 2 m. The SNESIM algorithm [*Strebelle*, 2002] is used to simulate 500 unconditional realizations of the aquifer deposits on a 40 x 60 x 14 grid with dimensions 1 m x 1 m x 0.5 m.

For each realization, the heat injection and transport are simulated with HydroGeoSphere [*Therrien et al.*, 2010] using transient, fully-saturated flow. The model was extended 20 m in each direction to reduce the effect of boundary conditions on the flow. The boundary conditions are imposed head mimicking a small natural gradient from the injection to the pumping well and constant temperatures. Flow and transport parameters are considered constant, i.e. the variability in the prior is only related to spatial variability in the distributions of the facies. Only the hydraulic conductivity depends on the facies, it is constant and equal to 10^{-4} and 5×10^{-2} m/s for sand and gravel, respectively.

364 The forecasted property h is composed of the variations of temperature in the ERT panel at 20

365 time-steps (every 6 hours during 5 days), simulated by the flow and transport model. We use

366 the standard relationship between temperature and water electrical conductivity [*Campbell et*

367 *al.*, 1948] to derive the corresponding change in electrical conductivity.

368
$$\frac{\sigma_T}{\sigma_{25}} = m_f(T - 25) + 1 \quad (15)$$

369 where σ_T is the water electrical conductivity at temperature T (in °C) and m_f is the fractional

370 change of conductivity by °C. We use $m_f = 0.02$ °C^{-1} in accordance with values found in the

371 literature [e.g., *Hermans et al.*, 2014] and $\sigma_{25} = 0.04$ S/m. Both can be easily measured on a

372 water sample. The background resistivity of sand and gravel are respectively 160 and 120

373 Ohm.m and we used the ratio of Archie's law between two time-steps to link the change in

374 water and bulk electrical conductivity. This allows expressing the temperature-dependence of

375 electrical conductivity without defining the formation factor [*Hermans et al., 2014*].

376 Based on these resistivity distributions, an ERT resistance data set is simulated for each time

377 step with a combination of AB-MN and AM-BN bipole-bipole configurations [*Zhou and*

378 *Greenhalgh*, 2000]. The total number of resistance data is 1436. Finally, we compute the change

379 in resistance by subtracting at each time step the resistance corresponding to the background

380 resistivity. This constitutes our data set d.

3.3. Posterior distribution of the temperature change from noise-free data

382 The choice of the dimensions q and p to retain in PCA is driven by the represented variance

383 [e.g., *Oware and Moysey*, 2014]. The dimension q kept for the forecast must therefore represent

384 enough variance to propose realistic samples of the physical property. For low dimension

385 forecasts such as a time series at a given location, explaining more than 99% of the variance is

386 feasible [*Satija and Caers*, 2015]. However, for high dimension forecasts such as those

387 considered here with two spatial dimensions and one temporal dimension, a lower cut-off is

388 expected.

389 Part of the data may be uninformative about the forecast, hence it is not always necessary to

390 retain 99% or more of the variance of the data to choose the dimension p. However, we

391 generally observe an increase in correlation coefficients in CCA when more dimensions are

392 kept in the data. To ensure the back transform of CCA, p has also to be higher than q. As will

393 be shown in the next section, the choice of the dimension should rather be dictated by

394 considerations on the noise level and kept to a level sufficiently low compared to the number

395 of models in the prior set.

396 The noise free data is composed of 20 series of 1436 resistance changes, i.e., a total of 28720

397 dimensions for each of the 500 realizations. However, resistance data exhibit a very high level

398 of redundancy and 90% of the spatiotemporal variance is represented by the first 3 principal

399 components, 99% by the first 12 components and 99.95% by the first 50 dimensions. In this

400 example, we keep the first 50 dimensions in the data in a reduced data set $\boldsymbol{d^f}$.

401 The forecast is composed of the value of temperature changes at the 28 x 26 cells (25 cm x 25

402 cm) of the grid used for forward ERT modeling, giving a total of 14560 dimensions for the 499

403 different realizations of the prior set (the reference forecast is not considered). 90% of the

404 variance is explained by the first 8 principal components, 95% by the first 14 and 99% by the

405 first 35 dimensions. We keep the first 14 dimensions in the reduced forecast $\boldsymbol{h^f}$. This ensures a

406 reliable spatio-temporal reconstruction of the forecast with a relatively low number of

407 dimensions.

408 CCA is applied on the reduced data sets in order to maximize the correlation between the

409 components of the data and forecast variables (Figure 2). In this case, the proposed dimension

410 reduction is efficient: the correlation coefficient for the four first dimensions, carrying most of

411   the variance, is above 0.99. For higher dimensions, a linear relationship is still observed until

412   dimension 14, supporting the use of the Gaussian regression.

413   We generate 300 samples of the posterior distribution $f_H(\boldsymbol{h^c}|\boldsymbol{d^c_{obs}})$ which were subsequently

414   back transformed to the physical space. Four samples of the distribution are shown with the

415   reference solution at 4 selected time steps in Figure 3. The samples successfully identify the

416   temporal behavior of the temperature distribution as well as the location and relative spread of

417   temperature changes. Most differences between samples are related to the value of the

418   maximum change in temperature. Almost all samples show an abrupt decrease in temperature

419   at 6 m depth, as observed in the reference solution.

420   The spatial and temporal characteristics of the posterior distribution are summarized in Figure

421   4. The spatiotemporal mean of the samples slightly overestimates the maximum change in

422   temperature but clearly identifies the zone affected by a change in temperature. The standard

423   deviation shows that most variability is related to the time-step exhibiting maximum change

424   (i.e., t = 36 h). The maximum value of the standard deviation is observed at the location of the

425   abrupt temperature decrease, which shows that it is a region difficult to resolve. Figure 4B

426   shows the mean temperature in the panel at the different time steps for the prior, the posterior

427   and the reference distributions. This illustrates that the ERT data is very efficient in reducing

428   uncertainty on the mean temperature in the panel compared to the prior distribution. The PFA

429   also clearly identifies the temporal behavior of the tracer, because the maximum change in

430   temperature is not observed at the same time step for the reference as for most of the models in

431   the prior set (36 h instead of 30 h).

432   **3.4. Posterior distribution of the temperature change from noisy data**

433   We perform Monte Carlo simulations with noisy data to evaluate how the noise is propagated

434   into the PCA component scores. Figure 5 shows the estimated covariance matrix for 2 different

435   cases. The first case corresponds to a 1% random Gaussian noise added to the changes in

436   resistance. The error thus increases with the change in resistance

437   If the error on the measured resistance is composed of systematic and random components, the

438   first component is cancelled out when using data difference, but the random component can be

439   additive. If the actual resistance between two-time steps does not change, but that the random

440   component does, the resulting error for a zero change in resistance can be as high as for a larger

441   change in resistance.

442   The second case therefore considers a level of noise independent of the change in resistance.

443   The error distribution is symmetric around a zero mean with a 0.034 Ohm standard deviation,

444   the proportion of large errors decreases exponentially: 50% of the absolute errors are below

445   0.004 Ohm and 90% below 0.0483 Ohm. The histogram of the error distribution is presented

446   in the supplementary file (Figure S2). The median absolute change in resistance for the whole

447   data set is 0.018 Ohm; therefore, the ratio of the median absolute error on the median absolute

448   resistance change is 22 %. The median of the relative error distribution is 18%. The overall

449   error level is thus much larger than in the first case

450   An illustration of the corresponding noisy data is given in the supporting information figure S2.

451   The latter case will be referenced as "Independent noise" in the next sections. It is important to

452   note that the proposed methodology can accommodate any noise description or any noise model

453   as it does not rely on any specific assumption concerning the noise distribution. An additional

454   case, with a 10% random Gaussian noise added to the changes in resistance is presented in the

455   supplementary file (Figures S3 and S4). The interpretation of this case is relatively similar to

456   the "Independent noise" below.

457   For the 1% Gaussian noise case (first column of Figure 5), the PCA error covariance matrix C_f^d

458   is sparse, with only a few non-zero entries along the diagonal. The low level of noise has little

459   effect on the estimation of PCA scores in the first 50 dimensions. The CCA error covariance

460   matrix C_c^d is computed through equation 12 and shows values close to 0, smaller than the values

461   observed in the error model covariance matrix C_T whose entries increase with the dimensions,

462   as the correlation decreases. This means that the noise has a negligible effect compared to the

463   error made in the linear regression. The 1% Gaussian noise does not influence the posterior

464   distribution in this case. For the "Independent noise" (second column of Figure 5), the entries

465   of C_f^d are significantly larger (note the change in the color scale), meaning that the error made

466   on the estimated PCA scores is no longer negligible. C_c^d entries are one order of magnitude

467   above C_T. Note that C_T is the same as in the first case (only the color scale is different), as it

468   only depends on the error in the regression model and thus on the number of dimensions used

469   in CCA. An inappropriate estimation of C_c^d may lead to a poor estimation of the posterior. Note

470   that the estimated C_f^d is almost a diagonal matrix, which is expected since the principal

471   components are orthogonal and thus uncorrelated.

472   The analysis of C_f^d clearly shows that all components are not affected in the same fashion by

473   the noise. In the "Independent noise" case, except for the 4th dimension, the ten first components

474   are well estimated. The error on the PCA scores is particularly high between dimensions 25 to

475   40. One possibility is therefore to reduce the number of dimensions in the data set. For example,

476   keeping the 25 first dimensions in the data (3rd column in Figure 5) still explains 99.7 % of the

477   variance and drastically reduces the values of C_c^d. Since less dimensions are used, the linear

478   regression is weaker and the components of C_T are slightly higher than when 50 dimensions are

479   kept, such that C_c^d is of the same order of magnitude as C_T, but only higher dimensions are

480   affected.

481   The effects of the choice of the data dimensions and the estimation of the error covariance

482   matrix are illustrated in Figures 6 and 7 with the mean and standard deviation of the spatial

483   temperature change for noisy data sets ("Independent noise"). It appears that ignoring the

484   covariance matrix for noisy data when 50 dimensions are used (2nd column in Figures 6 and 7)

485   yields results contaminated by strong artifacts. The mean of the posterior distribution misses

486   the spatial patterns and shows no temporal coherence. Due to noise, the poor estimation of the

487   CCA scores leads to an unacceptable estimation of the forecast. This stresses the importance of

488   estimating the noise level, as it is the case for standard tomographic inversion. Using the

489   appropriate estimation of $\boldsymbol{C_c^d}$ (3rd column) stabilizes the PFA with a mean solution correctly

490   estimating the spatial and temporal trends but with a higher standard deviation compared to the

491   noise free case (1st column), as expected.

492   Using only 25 dimensions in the data enables reducing the impact of noise (4th and 5th columns).

493   Both solutions with and without estimation of $\boldsymbol{C_c^d}$ renders results similar to the noise free case.

494   Given that less dimensions are used, $\boldsymbol{C_T}$ has larger values which tends to increase the standard

495   deviation of the posterior. Taking into account the error covariance matrix has little effect on

496   both mean and standard deviations, as the poorly estimated dimensions are not considered in

497   CCA.

498   The proposed noise propagation procedure identifies clearly the risk for the prediction to be

499   affected by artifacts related to the noise on the data. By analyzing the error covariance matrix

500   in PCA scores, the components affected by the noise are easily identified. They can be

501   disregarded or properly processed by using the covariance matrix for prediction. In this specific

502   case, the conservative choice is to reduce the dimensions of data variables. Indeed, the noise

503   introduced when a high number of dimensions is used is counterbalanced by an estimated

504   covariance matrix with larger components, due to the higher error made on the estimation of

505   PCA scores. This leads to a higher uncertainty on the temperature estimate.

506   **3.5. Applications to high and low temperature references**

22

507   As shown in Figure 4B, the prior distribution spans a broad range of maximum mean change in

508   temperature, from 0.5 up to 3.5°C. In the previous application, the reference lies in the middle

509   of the prior distribution of the forecast, and therefore many realizations in the prior set are

510   relatively similar to the reference. In this section, we apply the PFA to forecast the temperature

511   for reference realizations at the extreme ends of the prior distribution of the forecast with

512   relatively high and low temperature changes respectively. The data correspond to the

513   "Independent noise" case (see section 3.4) and we use the 25 first PCA components of the data

514   variables.

515   For a high maximum temperature change (Figures 8 and 10), the zone affected by temperature

516   changes, its spreading and its maximum amplitude are correctly identified by the samples of

517   the posterior distributions and the mean. The temporal behavior is also coherent with the

518   reference solution. We observe somewhat more variations in the posterior distribution (higher

519   standard deviation) which are related to the higher amplitude of the anomaly.

520   The low temperature change is more difficult to image (Figures 9 and 10). Indeed, low

521   variations in temperature yield low change in resistance so that the measured signal is low. The

522   case displayed in Figure 9 has a maximum change in temperature of 1°C. In addition, it shows

523   a relatively early arrival of the tracer. A change of 1°C corresponds to a change in resistivity of

524   only about 2%, which is often considered as not interpretable in field-scale time-lapse ERT

525   [e.g., *Robert et al.*, 2012; *Doetsch et al.*, 2012; *Chretien et al.*, 2014; *Hermans et al.*, 2012,

526   2015].

527   The two displayed samples identify the early arrival, with a maximum temperature slightly

528   higher than the reference. Since some samples of the posterior also display lower maximum

529   temperature changes, the mean solution is relatively close to the reference. The location of the

530   anomaly is correctly imaged. Given the low temperature change, the uncertainty in the posterior

23

531   is higher than for the two first examples, showing that many temperature distributions can be

532   coherent with the data.

533   Figure 10 shows the mean temperature in the panel at the different time steps for the prior, the

534   posterior and the reference distributions for the cases illustrated in Figures 8 and 9. In both

535   cases, the reference lies within the posterior distribution whose distributions are much narrower

536   than the prior. This illustrates that the use of ERT data is very efficient in reducing uncertainty

537   related to the mean temperature in the panel, even though the "hot" and "cold" references are

538   above the 90% quantile and below the 10% quantile of the prior distribution, respectively. ERT

539   data also allow the temporal behavior of the tracer to be clearly identified.

540   For the high temperature case, the reference response is between quantiles 10% and 50% of the

541   posterior. Because many samples are generated between the 10% and 50% quantiles, these two

542   quantiles are very close. This can be related to the small number of realizations sampled in the

543   high temperature region of the prior distribution of the forecast by the prior set (the reference

544   is actually above the 98% quantile of the prior distribution).

545   For the low temperature case, noisy data may produce small decrease in temperature ($\Delta T < 0$),

546   limited to a few tenth of degrees for some of the samples. This is an indirect effect of the

547   Gaussian regression where values of $\boldsymbol{h}_c^d$ are sampled that are not observed in the data. The PFA

548   is not worse than classical inversions, since similar behaviors are observed in standard

549   inversions showing positive resistivity changes even if only negative changes are expected. It

550   highlights the fact that sampling the forecast in the tail of a distribution is challenging and might

551   require additional models in the prior set.

552   **4. Validation**

553   In this section, we validate the posterior distribution of the forecasted physical properties. *Satija*

554   *and Caers* [2015] and *Scheidt et al.* [2014] validated their direct forecasting methodology in a

555   hydrogeological context with rejection sampling. Here, we propose, as an alternative, to

556   validate the PFA through a Monte Carlo experiment. In a second step, we verify that the

557   sampled posterior distributions are consistent with the geophysical data.

558   **4.1. Monte Carlo experiments**

559   The objective of the Monte Carlo experiments is to show that the posterior distributions

560   obtained for many references integrate to the prior distribution. This will show that PFA is an

561   appropriate method to sample the posterior distribution.

562   Equation 6 expresses the general form of the posterior distribution given the prior distribution

563   and a likelihood function. If we generate Monte Carlo samples to explore the entire data space

564   by subsequently using every sample of the prior set as observed data $\boldsymbol{d_{obs}}$, and integrate the

565   posterior distributions of the forecast we have:

$$\int f_H(\boldsymbol{h}|\boldsymbol{d})f_D(\boldsymbol{d})d\boldsymbol{d} = \int \frac{L(\boldsymbol{h}|\boldsymbol{d})}{f_D(\boldsymbol{d})}f_H(\boldsymbol{h})f_D(\boldsymbol{d})d\boldsymbol{d} \qquad (16)$$

567   where $f_D(\boldsymbol{d})$ is the prior distribution of the data. The prior distribution of the forecast is

568   independent of $\boldsymbol{d}$

$$\int f_H(\boldsymbol{h}|\boldsymbol{d})f_D(\boldsymbol{d})d\boldsymbol{d} = f_H(\boldsymbol{h}) \int L(\boldsymbol{h}|\boldsymbol{d})d\boldsymbol{d} \qquad (17)$$

570   and the integral of the likelihood over the data space is equal to 1 so that we obtain

$$\int f_H(\boldsymbol{h}|\boldsymbol{d})f_D(\boldsymbol{d})d\boldsymbol{d} = f_H(\boldsymbol{h}) \qquad (18)$$

572   Equation 18 states that integrating the *s* samples of the posterior for the *n* realizations of the

573   prior set should give a distribution equal to the prior distribution of the forecast. This offers a

574   way to validate that the PFA yields physically plausible samples and to study its validity

575 regarding the number of dimensions kept in the CCA. If equation 18 is not verified, one should

576 question the number of principal components in the data and forecast used in the CCA.

577 This Monte Carlo validation is illustrated in Figure 11 for 25 and 50 dimensions kept in the

578 data and 14 in the forecast. The curves correspond to some of the quantiles of the prior and

579 integrated posterior distribution. In both cases, the integrated posterior distributions are almost

580 exactly similar to the prior distribution.

581 **4.2. Consistency of the sampled posterior of forecast with data**

582 Our methodology is defined statistically in a reduced dimension space, such that the Gaussian

583 regression process (or linear inversion) implies that the sampled h^c are coherent with the

584 observed reduced data d_{obs}^c. Therefore an objective function (deterministic inversion) or a

585 likelihood based on the actual observed geophysical data, is not required. There is also no need

586 for any regularization or to run a forward simulation at any step of the PFA, except to build the

587 prior set. In fact, the entire procedure can be perfectly parallelized in terms of the forward

588 modeling.

589 Nevertheless, we verify the consistency of the sampled posterior with the actual geophysical

590 data [e.g., *Lochbühler et al.*, 2014]. However, since the solution is expressed in terms of

591 physical property (temperature) and not resistivity and that the data are expressed in terms of

592 change in resistance, there is no direct way to perform this validation. To overcome this issue,

593 we use the following workflow:

594 1) the resistances of the background data set are deterministically inverted to obtain a

595   distribution of resistivity ρ_{BG}, we do not use the true resistivity because they would be

596   unknown in practice;

2) through the petrophysical relationship used for the creation of the prior set (equation

   15), we transform, for each time step, the change in temperature ΔT into a change in

   resistivity $\Delta \rho$ that we add to ρ_{BG} to get ρ_{TL}; note that there is no issue of resolution-

   dependent petrophysical relationship, since no constraint on the model was imposed to

   derive ΔT;

3) we compute the resistances for each time step and deduce the data, i.e. the changes in

   resistance $\Delta R_{calc} = d_{calc}$;

4) we calculate the data misfit between d_{true}, the noise free reference data set, and d_{calc}

   through the root-mean-square error (RMS) normalized by the value of RMS between

   noise free and noisy data

$$\eta_{RMS} = \frac{\sqrt{\sum_i^{ut}(d_{calc,i}-d_{true,i})^2}}{\sqrt{\sum_i^{ut}(d_{obs,i}-d_{true,i})^2}} \tag{19}$$

   where u is the number of resistance data and t the number of time steps in the monitoring

   process;

5) we repeat steps 2 to 4 for the s samples of the posterior.

The above workflow was applied to the samples of the posterior distribution corresponding to

the 3 reference realizations, i.e., intermediate, high and low mean temperature changes, shown

respectively in Figures 6, 8 and 9 with 25 dimensions kept in the data and the "Independent

noise" level. The posterior consists of 300 samples and the histogram of their normalized error

η_{RMS} is shown in Figure 12. For the three cases, the mode of the histogram is around 1, i.e., an

error corresponding to the noise level in the data. This shows that the proposed methodology is

able to generate realistic distribution of the physical properties of interest in accordance with

geophysical data, without explicitly running the geophysical forward operator during the

inversion process.

5. Conclusions

We propose to solve the estimation of physical properties with geophysical data with a direct forecast approach, the PFA, as an alternative to existing deterministic inversions broadly used in geophysics. The PFA enables directly estimating the spatio-temporal posterior distribution of physical properties, such as salinity or temperature, to be directly sampled from geophysical time-series. The method is based on generating a prior set of models mimicking the experiment of interest. After dimension reduction of the data and forecast with principal component analysis and canonical correlation analysis, we use a Gaussian regression in the reduced dimension space to sample the posterior distribution. Then, the samples in the reduced space are back transformed to the original space.

We applied the PFA in a heat tracer experiment for a broad range of amplitude variations. The methodology is able to correctly identify the spatio-temporal behavior of the tracer, as well as the amplitude of the maximum change. The spatio-temporal distribution of temperature is complex. We retained 14 dimensions explaining 95% of the variance. It signifies that small-scale features of the forecast corresponding to higher dimensions cannot be reproduced by the method. The corresponding loss of resolution can be assessed by applying this dimension reduction on models from the prior set.

We proposed to analyze the propagation of noise in the reduced component space through Monte Carlo simulations to estimate the error covariance matrix in the reduced dimension space. This step is necessary to stabilize the inversion and shows that reducing the number of dimension in the data can drastically increase the robustness of the method regarding noise.

We validated the approach by verifying the consistency of the samples of the posterior distribution with the geophysical data and the prior distribution. However, the geophysical parameters are never explicitly inverted in the PFA.

644 In contrast to hydrogeophysical modeling using standard regularized geophysical inversion
645 techniques, the method is (i) stochastic, i.e., the solution is composed of samples of the full
646 posterior distribution and not a single deterministic solution; (ii) it solves the spatio-temporal
647 physical property distribution and not the geophysical parameter; (iii) it does not require the
648 definition of a regularization term which is recognized to be the major element preventing
649 quantitative integration of geophysical data in hydrogeological models. Using physically-based
650 simulations to build the prior set of model, PFA yields physically plausible forecasts. The key
651 point is to build a prior set in accordance with the expected response.

652 The major advantage of the methodology is to avoid any regularization procedure required in
653 deterministic inversions. The solutions therefore do not suffer from any spatial bias, spatially
654 and temporally varying resolution or uncertainty in the post-inversion petrophysical
655 transformation. Petrophysics is only required to build the prior set. After this step, there is no
656 need to run any forward hydrogeological or geophysical simulation. Therefore, the PFA enables
657 to quickly generate the full posterior distribution of the problem compared to classic stochastic
658 inversion methods. This remains true for problems with a large number of parameters, as long
659 as it is possible to efficiently reduce their dimensions.

660 Except for building the prior set, the PFA does not require a flow and transport model to
661 generate the posterior distribution. It is thus more flexible than coupled inversion schemes and
662 much faster than stochastic McMC methods. It is therefore a good alternative when only
663 monitoring results are needed or to improve sequential inversion scheme with a physically
664 plausible distribution of the physical property, integrating uncertainty of the estimate.

665 Nevertheless, the method is limited by the choice of prior parameters involved in the
666 construction of the prior set itself, i.e., on the underlying hydrogeological model and the
667 petrophysical relationship. Indeed, a satisfactory solution will be found only if the actual data

668   lies in the empirical prior distribution of the prior set. Therefore, the method should not be used

669   for extrapolation. If the observed data lies outside of the prior distribution, it signifies that the

670   data is not consistent with the prior. However, this is a limitation for all Bayesian methods,

671   since the Bayesian model requires for the prior to capture data to avoid prior-data inconsistency.

672   PFA has the advantage over other techniques to clearly identify such inconsistency using the

673   joint distribution in the reduced dimension space. Observed data corresponding to extremes of

674   the data prior distribution might be difficult to forecast. In such a case, it might be necessary to

675   generate more realizations in the vicinity of the observed data and therefore increase the number

676   of models in the prior set. Similarly, more models might be needed for prior considering more

677   uncertain parameters such as geological scenarios, boundary conditions or continuous hydraulic

678   conductivity distributions. However, the computational cost of the method would remain lower

679   than stochastic McMC methods.

680   Extending the distribution of prior parameters in the prior set to allow uncertainty on sensitive

681   parameters such as the petrophysical relationships, the hydraulic conductivity of the deposits

682   and the heterogeneity would ensure a larger data prior distribution and therefore allow an

683   application of the prior set to a broader range of observed data. Although physically-based flow

684   and transport simulations were used to generate the prior set in this study, the PFA applied to

685   geophysical data could also work with a prior set generated differently (e.g., with a

686   geostatistical tool). Creating the prior set would be faster but at the expense of geological

687   realism. The proposed developments will be investigated in future work and tested on a real-

688   world case study.

689   **Acknowledgement**

References

699　Arato, A., J. Boaga, C. Comina, M. De Seta, E. Di Sipio, A. Galgaro, N. Giordano, and G.

700　　　　Mandrone (2015), Geophysical monitoring for shallow geothermal applications - Two

701　　　　Italian case histories, *First Break*, *33*(8), 75–79.

702　Audebert, M., R. Clément, N. Touze-Foltz, T. Günther, S. Moreau, and C. Duquennoi (2014),

703　　　　Time-lapse ERT interpretation methodology for leachate injection monitoring based

704　　　　on multiple inversions and a clustering strategy (MICS), *Journal of Applied*

705　　　　*Geophysics*, *111*, 320–333, doi:10.1016/j.jappgeo.2014.09.024.

706　Auken, E., J. Doetsch, G. Fiandaca, A. V. Christiansen, A. Gazoty, A. G. Cahill, and R.

707　　　　Jakobsen (2014), Imaging subsurface migration of dissolved CO2 in a shallow aquifer

708　　　　using 3-D time-lapse electrical resistivity tomography, *Journal of Applied Geophysics*,

709　　　　*101*, 31–41, doi:10.1016/j.jappgeo.2013.11.011.

710　Beaujean, J., F. Nguyen, A. Kemna, A. Antonsson, and P. Engesgaard (2014), Calibration of

711　　　　seawater intrusion models: Inverse parameter estimation using surface electrical

712　　　　resistivity tomography and borehole data, *Water Resources Research*, *50*(8), 6828–

713　　　　6849, doi:10.1002/2013WR014020.

714　Ben Hadj Miled, M. K., and E. L. Miller (2007), A projection-based level-set approach to

715   enhance conductivity anomaly reconstruction in electrical resistance tomography,

716   *Inverse Problems*, *23*(6), 2375–2400, doi:10.1088/0266-5611/23/6/007.

717   Binley, A., P. Winship, L. J. West, M. Pokar, and R. Middleton (2002), Seasonal variation of

718   moisture content in unsaturated sandstone inferred from borehole radar and resistivity

719   profiles, *Journal of Hydrology*, *267*(3), 160–172.

720   Binley, A., S. S. Hubbard, J. A. Huisman, A. Revil, D. A. Robinson, K. Singha, and L. D.

721   Slater (2015), The emergence of hydrogeophysics for improved understanding of

722   subsurface processes over multiple scales, *Water Resources Research*, *51*, 3837–3866,

723   doi:10.1002/2015WR017016.

724   Briggs, M. A., F. D. Day-Lewis, J. B. T. Ong, G. P. Curtis, and J. W. Lane (2013),

725   Simultaneous estimation of local-scale and flow path-scale dual-domain mass transfer

726   parameters using geoelectrical monitoring, *Water Resources Research*, *49*(9), 5615–

727   5630, doi:10.1002/wrcr.20397.

728   Campbell, R., C. Bower, and L. Richards (1948), Change of electrical conductivity with

729   temperature and the relation of osmotic pressure to electrical conductivity and ion

730   concentration in soil extracts, *Soil Sci. Soc. Am. J.*, *13*, 66–69.

731   Carrigan, C. R. et al. (2013), Electrical resistance tomographic monitoring of CO2 movement

732   in deep geologic reservoirs, *International Journal of Greenhouse Gas Control*, *18*,

733   401–408, doi:10.1016/j.ijggc.2013.04.016.

734   Caterina, D., J. Beaujean, T. Robert, and F. Nguyen (2013), A comparison study of different

735   image appraisal tools for electrical resistivity tomography, *Near Surface Geophysics*,

736   *11*, 639–657, doi:10.3997/1873-0604.2013022.

737   Caterina, D., T. Hermans, and F. Nguyen (2014), Case studies of incorporation of prior

738   information in electrical resistivity tomography: comparison of different approaches,

739   *Near Surface Geophysics*, *12*, 451–465, doi:10.3997/1873-0604.2013070.

740   Chen, X., H. Murakami, M. S. Hahn, G. E. Hammond, M. L. Rockhold, J. M. Zachara, and Y.

741   Rubin (2012), Three-dimensional Bayesian geostatistical aquifer characterization at

742   the Hanford 300 Area using tracer test data, *Water Resources Research*, *48*(6),

743   W06501, doi:10.1029/2011WR010675.

744   Chrétien, M., J. F. Lataste, R. Fabre, and A. Denis (2014), Electrical resistivity tomography to

745   understand clay behavior during seasonal water content variations, *Engineering*

746   *Geology*, *169*, 112–123, doi:10.1016/j.enggeo.2013.11.019.

747   Christensen, N.K., S. Christensen, and T.P.A. Ferre (2016), Testing alternative uses of

748   electromagnetic data to reduce the prediction error of groundwater models, *Hydrology*

749   *and Earth System Sciences*, *20*(5), 1925-1946, doi:10.5194/hess-20-1925-2016.

750   Day-Lewis, F. D., K. Singha, and A. Binley (2005), Applying petrophysical models to radar

751   travel time and electrical resistivity tomograms: Resolution-dependent limitations,

752   *Journal of Geophysical Research*, *110*, B08206, doi:10.1029/2004JB003569.

753   Day-Lewis, F. D., Y. Chen, and K. Singha (2007), Moment inference from tomograms,

754   *Geophysical Research Letters*, *34*(22), L22404, doi:10.1029/2007GL031621.

755   Doetsch, J., N. Linde, and A. Binley (2010), Structural joint inversion of time-lapse crosshole

756   ERT and GPR traveltime data, *Geophysical Research Letters*, *37*(24), L24404,

757   doi:10.1029/2010GL045482.

758   Doetsch, J., N. Linde, T. Vogt, A. Binley, and A. G. Green (2012), Imaging and quantifying

759   salt-tracer transport in a riparian groundwater system by means of 3D ERT

760   monitoring, *Geophysics*, *77*(5), B207–B218, doi:10.1190/geo2012-0046.1.

761   Fiandaca, G., J. Doetsch, G. Vignoli, and E. Auken (2015), Generalized focusing of time-

762   lapse changes with applications to direct current and time-domain induced polarization

763   inversions, *Geophysical Journal International*, *203*(2), 1101–1112.

764   Flores Orozco, A., A. Kemna, and E. Zimmermann (2012), Data error quantification in

765   spectral induced polarization imaging, *Geophysics*, *77*(3), E227–E237,

766   doi:10.1190/GEO2010-0194.1.

767   Hansen, T. M., K. S. Cordua, B. H. Jacobsen, and K. Mosegaard (2014), Accounting for

768   imperfect forward modeling in geophysical inverse problems — Exemplified for

769   crosshole tomography, *Geophysics*, *79*(3), H1–H21, doi:10.1190/geo2013-0215.1.

770   Hermans, T., A. Vandenbohede, L. Lebbe, and F. Nguyen (2012), A shallow geothermal

771   experiment in a sandy aquifer monitored using electric resistivity tomography,

772   *Geophysics*, *77*(1), B11–B21, doi:10.1190/geo2011-0199.1.

773   Hermans, T., F. Nguyen, T. Robert, and A. Revil (2014), Geophysical Methods for

774   Monitoring Temperature Changes in Shallow Low Enthalpy Geothermal Systems,

775   *Energies*, *7*(8), 5083–5118, doi:10.3390/en7085083.

776   Hermans, T., S. Wildemeersch, P. Jamin, P. Orban, S. Brouyère, A. Dassargues, and F.

777   Nguyen (2015), Quantitative temperature monitoring of a heat tracing experiment

778   using cross-borehole ERT, *Geothermics*, *53*, 14–26,

779   doi:10.1016/j.geothermics.2014.03.013.

780   Hermans, T., A. Kemna, and F. Nguyen (2016), Covariance-constrained difference inversion

781   of time-lapse electrical resistivity tomography data, *Geophysics*, *81*(5).

782 Hinnell, A. C., T. P. . Ferré, J. A. Vrugt, J. A. Huisman, S. Moysey, J. Rings, and M. B.

783   Kowalsky (2010), Improved extraction of hydrologic information from geophysical

784   data through coupled hydrogeophysical inversion, *Water Resources Research*, *46*,

785   W00D40, doi:10.1029/2008WR007060.

786 Irving, J., and K. Singha (2010), Stochastic inversion of tracer test and electrical geophysical

787   data to estimate hydraulic conductivities, *Water Resources Research*, *46*, W11514,

788   doi:10.1029/2009WR008340.

789 Johnson, T. C., L. D. Slater, D. Ntarlagiannis, F. D. Day-Lewis, and M. Elwaseif (2012),

790   Monitoring groundwater-surface water interaction using time-series and time-

791   frequency analysis of transient three-dimensional electrical resistivity changes, *Water

792   Resources Research*, *48*(7), W07506, doi:10.1029/2012WR011893.

793 Johnson, T. C., R. J. Versteeg, F. D. Day-Lewis, W. Major, and J. W. Lane (2015), Time-

794   Lapse Electrical Geophysical Monitoring of Amendment-Based Biostimulation,

795   *Groundwater*, *53*(6), 920–932, doi:10.1111/gwat.12291.

796 Karaoulis, M., J. H. Kim, and P. Tsourlos (2011), 4D active time constrained resistivity

797   inversion, *Journal of Applied Geophysics*, *73*, 25–34.

798 Kemna, A., B. Kulessa, and H. Vereecken (2002), Imaging and characterisation of subsurface

799   solute transport using electrical resistivity tomography (ERT) and equivalent transport

800   models, *Journal of Hydrology*, *267*(3-4), 125–146.

801 Kim, J. H., M. J. Yi, S. G. Park, and J. G. Kim (2009), 4-D inversion of DC resistivity

802   monitoring data acquired over a dynamically changing earth model, *Journal of

803   *Applied Geophysics*, *68*(4), 522–532.

804   Koestel, J., A. Kemna, M. Javaux, A. Binley, and H. Vereecken (2008), Quantitative imaging

805   of solute transport in an unsaturated and undisturbed soil monolith with 3-D ERT and

806   TDR, *Water Resources Research*, *44*(12), W12411, doi:10.1029/2007WR006755.

807   Krzanowski, W. J. (2000), *Principles of multivariate analysis: A user's perspective*, Oxford

808   Statistical Series 22, Revised Edition., Oxford University Press, New York.

809   LaBrecque, D. J., M. Miletto, W. Daily, A. Ramirez, and E. Owen (1996), The effects of

810   noise on Occam's inversion of resistivity tomography data, *Geophysics*, *61*(2), 538–

811   548.

812   Laloy, E., N. Linde, and J. A. Vrugt (2012), Mass conservative three-dimensional water tracer

813   distribution from Markov chain Monte Carlo inversion of time-lapse ground-

814   penetrating radar data, *Water Resources Research*, *48*(7), W07510,

815   doi:10.1029/2011WR011238.

816   Linde, N., P. Renard, T. Mukerji, and J. Caers (2015), Geological realism in hydrogeological

817   and geophysical inverse modeling: A review, *Advances in Water Resources*, *86*, 86–

818   101, doi:10.1016/j.advwatres.2015.09.019.

819   Lochbühler, T., G. Pirot, J. Straubhaar, and N. Linde (2014), Conditioning of Multiple-Point

820   Statistics Facies Simulations to Tomographic Images, *Mathematical Geosciences*,

821   *46*(5), 625–645, doi:10.1007/s11004-013-9484-z.

822   Masy, T., D. Caterina, O. Tromme, B. Lavigne, P. Thonart, S. Hiligsmann, and F. Nguyen

823   (2016), Electrical resistivity tomography to monitor enhanced biodegradation of

824   hydrocarbons with Rhodococcus erythropolis T902.1 at a pilot scale, *Journal of*

825   *Contaminant Hydrology*, *184*, 1–13, doi:10.1016/j.jconhyd.2015.11.001.

826   Moysey, S., K. Singha, and R. Knight (2005), A framework for inferring field-scale rock

827   physics relationships through numerical simulation, *Geophysical Research Letters*,

828   *32*(8), L08304, doi:10.1029/2004GL022152.

829   Müller, K., J. Vanderborght, A. Englert, A. Kemna, J. A. Huisman, J. Rings, and H.

830   Vereecken (2010), Imaging and characterization of solute transport during two tracer

831   tests in a shallow aquifer using electrical resistivity tomography and multilevel

832   groundwater samplers, *Water Resources Research*, *46*(3), W03502,

833   doi:10.1029/2008WR007595.

834   Murakami, H., X. Chen, M. S. Hahn, Y. Liu, M. L. Rockhold, V. R. Vermeul, J. M. Zachara,

835   and Y. Rubin (2010), Bayesian approach for three-dimensional aquifer

836   characterization at the Hanford 300 Area, *Hydrology and Earth System Sciences*,

837   *14*(10), 1989–2001, doi:10.5194/hess-14-1989-2010.

838   Nguyen, F., A. Kemna, T. Robert, and T. Hermans (2016), Data-driven selection of the

839   minimum-gradient support parameter in time-lapse focused electric imaging,

840   *Geophysics*, *81*(1), A1–A5, doi:10.1190/GEO2015-0226.1.

841   Oware, E. K., and S. M. J. Moysey (2014), Geophysical evaluation of solute plume spatial

842   moments using an adaptive POD algorithm for electrical resistivity imaging, *Journal

843   of Hydrology*, *517*, 471–480, doi:10.1016/j.jhydrol.2014.05.054.

844   Oware, E. K., S. M. J. Moysey, and T. Khan (2013), Physically based regularization of

845   hydrogeophysical inverse problems for improved imaging of process-driven systems,

846   *Water Resources Research*, *49*(10), 6238–6247, doi:10.1002/wrcr.20462.

847 Pidlisecky, A., K. Singha, and F. D. Day-Lewis (2011), A distribution-based parametrization

848   for improved tomographic imaging of solute plumes, *Geophysical Journal*

849   *International*, *187*(1), 214–224, doi:10.1111/j.1365-246X.2011.05131.x.

850 Robert, T., D. Caterina, J. Deceuster, O. Kaufmann, and F. Nguyen (2012), A salt tracer test

851   monitored with surface ERT to detect preferential flow and transport paths in

852   fractured/karstified limestones, *Geophysics*, *77*, B55–B67.

853 Rubin, Y., and S. Hubbard (2005), *Hydrogeophysics*, Water Science and Technology Library,

854   Springer, Doordrecht, the Netherlands.

855 Rubin, Y., X. Chen, H. Murakami, and M. Hahn (2010), A Bayesian approach for inverse

856   modeling, data assimilation, and conditional simulation of spatial random fields,

857   *Water Resources Research*, *46*(10), W10523, doi:10.1029/2009WR008799.

858 Satija, A., and J. Caers (2015), Direct forecasting of subsurface flow response from non-linear

859   dynamic data by linear least-squares in canonical functional principal component

860   space, *Advances in Water Resources*, *77*, 69–81, doi:10.1016/j.advwatres.2015.01.002.

861 Scheidt, C., P. Renard, and J. Caers (2015), Prediction-Focused Subsurface Modeling:

862   Investigating the Need for Accuracy in Flow-Based Inverse Modeling, *Mathematical*

863   *Geosciences*, *47*(2), 173–191, doi:10.1007/s11004-014-9521-6.

864 Singha, K., and S. M. Gorelick (2005), Saline tracer visualized with three-dimensional

865   electrical resistivity tomography: Field-scale spatial moment analysis, *Water*

866   *Resources Research*, *41*(5), W05023, doi:10.1029/2004WR003460.

867 Singha, K., and S. Moysey (2006), Accounting for spatially variable resolution in electrical

868   resistivity tomography through field-scale rock-physics relations, *Geophysics*, *71*(4),

A25–A28, doi:10.1190/1.2209753.

Singha, K., F. D. Day-Lewis, T. Johnson, and L. Slater (2015), Advances in interpretation of subsurface processes with time-lapse electrical imagin, *Hydrological Processes*, *29*(6), 1549–1576.

Strebelle, S. (2002), Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics, *Mathematical Geology*, *34*(1), 1–21.

Tarantola, A. (2005), *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Therrien, R., R. McLaren, E. Sudicky, and S. Panday (2010), *HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport*, Groundwater Simulation Group, Waterloo, ON, Canada.

Tonkin, M. J., and J. Doherty (2005), A hybrid regularized inversion methodology for highly parameterized environmental models, *Water Resources Research*, *41*(10), n/a-n/a, doi:10.1029/2005WR003995.

Truex, M. J., T. C. Johnson, C. E. Strickland, J. E. Peterson, and S. S. Hubbard (2013), Monitoring Vadose Zone Desiccation with Geophysical Methods, *Vadose Zone Journal*, *12*(2), 14, doi:10.2136/vzj2012.0147.

Vanderborght, J., A. Kemna, H. Hardelauf, and H. Vereecken (2005), Potential of electrical resistivity tomography to infer aquifer transport characteristics from tracer studies: A synthetic case study, *Water Resources Research*, *41*, W06013, doi:10.1029/2004WR003774.

Vereecken, H., A. Binley, G. Cassiani, A. Revil, and K. Titov (2007), *Applied*

891   *Hydrogeophysics*, NATO Science Series IV. Earth and Environmental Sciences,

892   Springer, Doordrecht, the Netherlands.

893 Wallin, E. L., T. C. Johnson, W. J. Greenwood, and J. M. Zachara (2013), Imaging high stage

894   river-water intrusion into a contaminated aquifer along a major river corridor using 2-

895   D time-lapse surface electrical resistivity tomography, *Water Resources Research*,

896   *49*(3), 1693–1708, doi:10.1002/wrcr.20119.

897 Zhou, B., and S. A. Greenhalgh (2000), Cross-hole resistivity tomography using different

898   electrode configurations, *Geophysical prospecting*, *48*(5), 887–912.

899

FIGURE CAPTIONS

901 Figure 1. A) Experimental set-up of the heat tracing test. B) Electrode location in the ERT

902 panel.

903 Figure 2. Canonical correlation analysis of the principal components of the change in resistance

904 data (50 dimensions) and change in temperature (14 dimensions). A linear relation is clearly

905 visible with correlation coefficient above 0.9 for the first eleven components. The black lines

906 show the position of the observed data.

907 Figure 3. Refererence model and 4 samples of the posterior at 4 different time steps. The

908 samples correctly catch the spatial and temporal behavior of the temperature distribution. Most

909 variations between samples are related to the amplitude of the maximum temperature.

910 Figure 4. A) Reference model, mean and standard deviation of the posterior distribution at

911 different time steps. The mean solution slightly overestimates the maximum temperature but

912 the reference solution lies within the posterior. The standard deviation amplitude is higher at

913 the location and time of the maximum temperature change. B) The posterior distribution of the

914 mean temperature in the panel is displayed with the prior and reference distribution, showing

915 that the proposed solution is a very good estimator of the mean temperature in the panel (top).

916 10, 50 and 90% quantiles of the distribution (bottom).

917 Figure 5. Error covariance matrix in PCA space (top), in CCA space (middle row) and Gaussian

918 regression error model covariance matrix (bottom) for Monte Carlo simulations with change in

919 resistance data contaminated by 1% random Gaussian noise (left), "Independent noise" (middle

920 column) and "Independent noise" with only the first 25 dimensions considered.

921 Figure 6. Mean of the posterior distribution at four different time steps for various cases: 50

922 dimensions in the data and noise free data (col. 1), 50 dimensions in the data, noisy data but

923 $C_c^d = 0$ (col. 2), 50 dimensions in the data, noisy data and C_c^d estimated with Monte Carlo

924 simulations (col. 3), 25 dimensions in the data, noisy data but $C_c^d = 0$ (col. 4), 25 dimensions

925 in the data, noisy data and C_c^d estimated with Monte Carlo simulations (col. 5).

926 Figure 7. Standard deviation of the posterior distribution at four different time steps for various

927 cases: 50 dimensions in the data and noise free data (col. 1), 50 dimensions in the data, noisy

928 data but $C_c^d = 0$ (col. 2), 50 dimensions in the data, noisy data and C_c^d estimated with Monte

929 Carlo simulations (col 3.), 25 dimensions in the data, noisy data but $C_c^d = 0$ (col 4.), 25

930 dimensions in the data, noisy data and C_c^d estimated with Monte Carlo simulations (col. 5).

931 Figure 8. Reference model for a high maximum temperature change (col. 1), 2 samples of the

932 posterior (col. 2 and 3), mean and standard deviation of the posterior distribution (col. 4 and 5).

933 Figure 9. Reference model for a low maximum temperature change (col. 1), 2 samples of the

934 posterior (col. 2 and 3), mean and standard deviation of the posterior distribution (col. 4 and 5).

935 Figure 10. Posterior distribution (10, 50 and 90% quantiles) of the mean temperature in the

936 panel with the prior and reference distribution for (A) the high and (B) low temperature models.

937 The bottom row sums up the distributions with the 10, 50 and 90% quantiles for the two cases.

938 Figure 11. The integral of posterior distribution of the temperature in the ERT panel is compared

939 to the prior distribution using the 10, 25, 50, 75 and 90% quantiles, for (A) 25 dimensions and

940 (B) 50 dimensions in the data. The distributions are almost exactly similar.
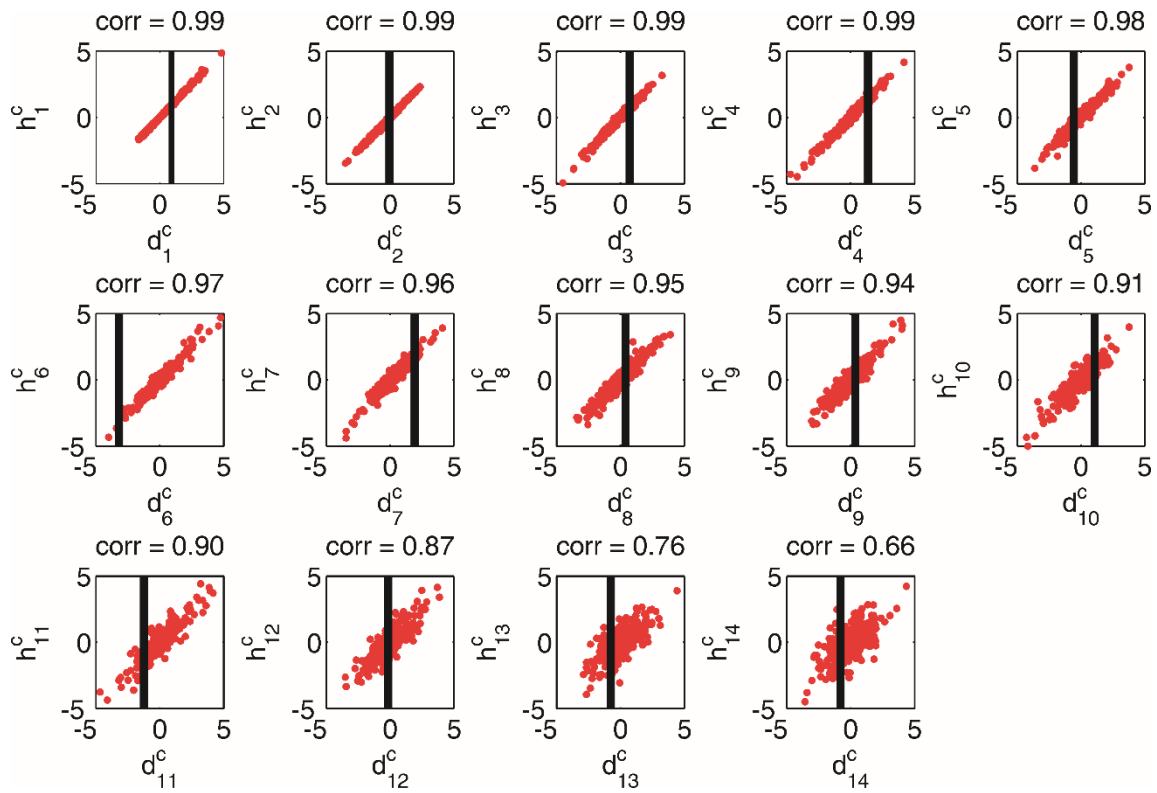
941 Figure 12. Normalized RMS value of the samples of the posterior distribution for (A) the

942 reference model with intermediate temperature of Figures 6, (B) high temperature of Figure 8

943 and (C) low temperature of Figure 9. The observed values are within the range of samples fitted

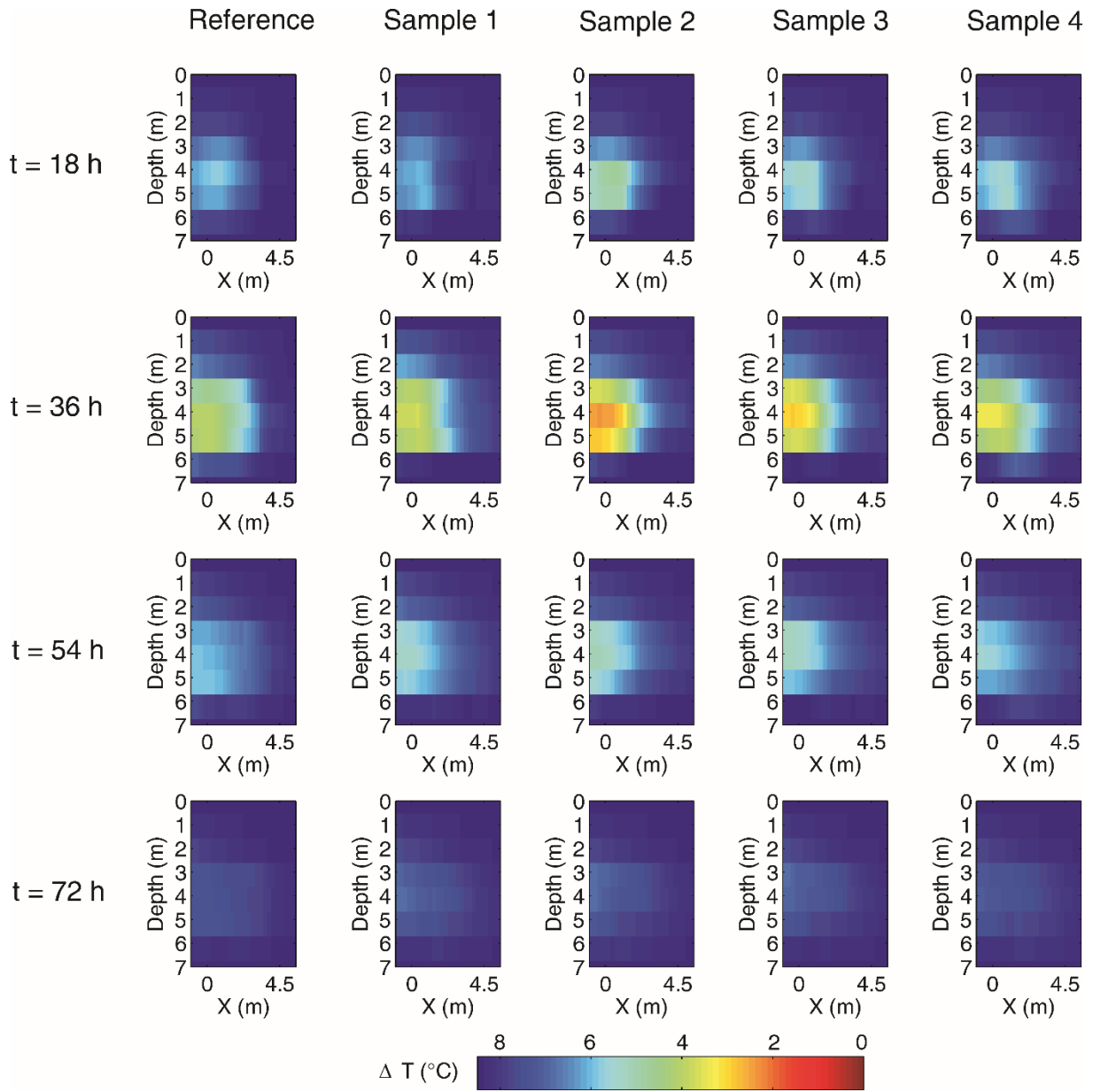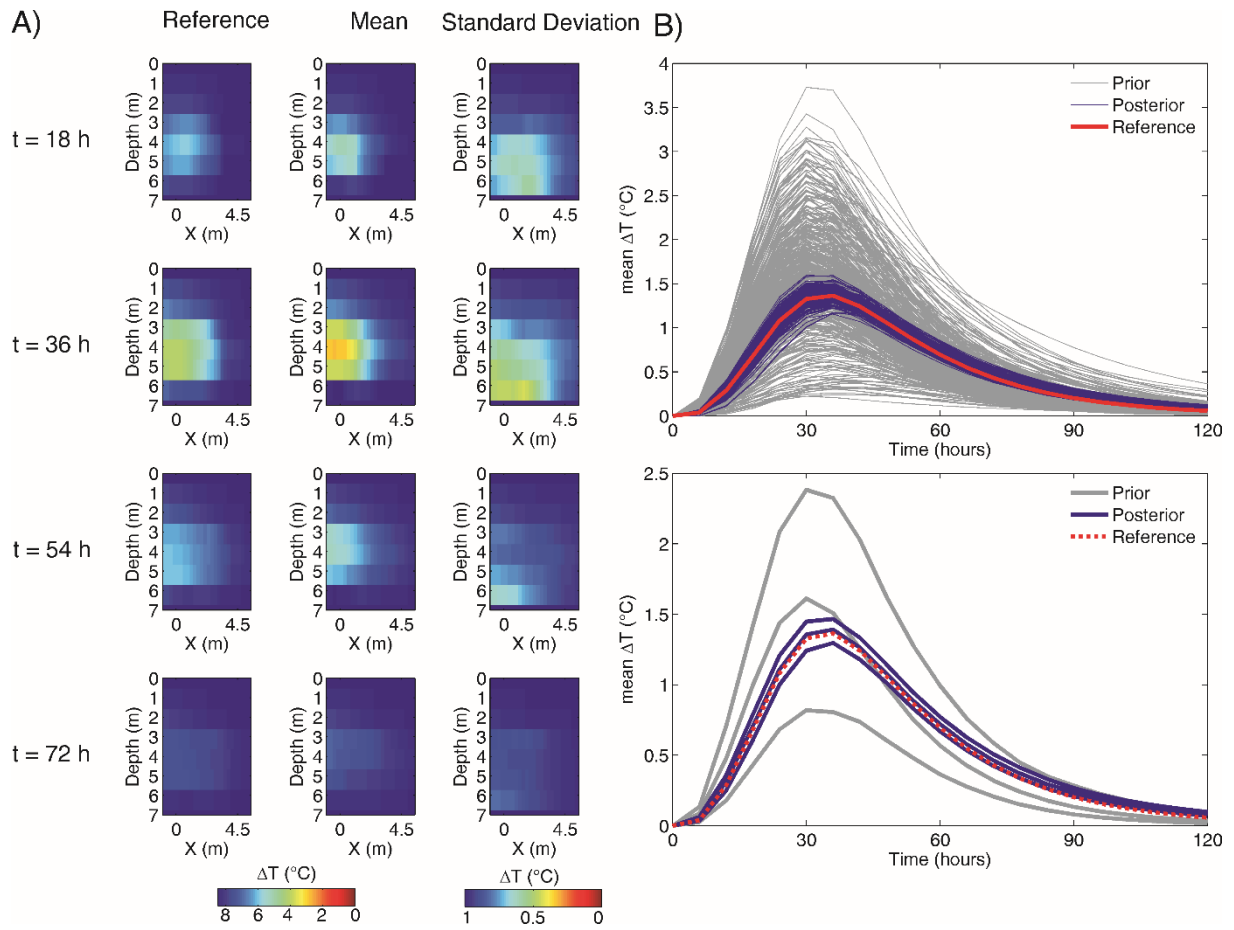944 to their noise level in standard inversion.

945

946

Figure 1

948
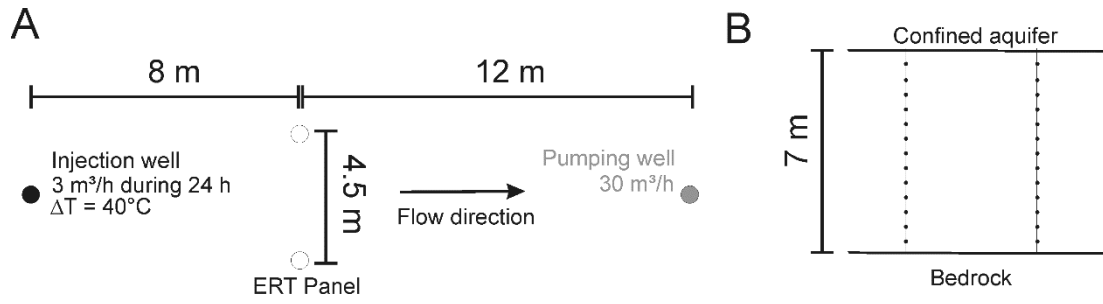
Figure 2
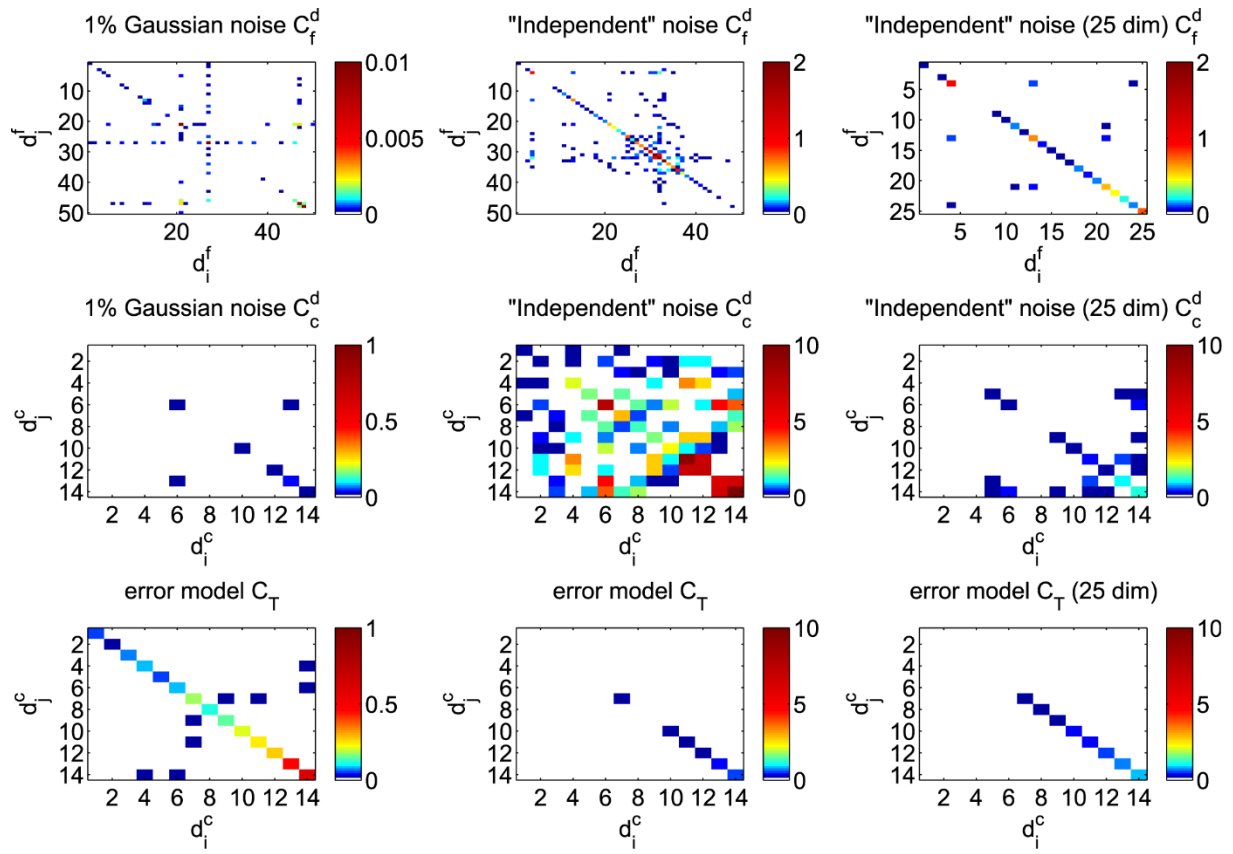
950

951   **Figure 3**

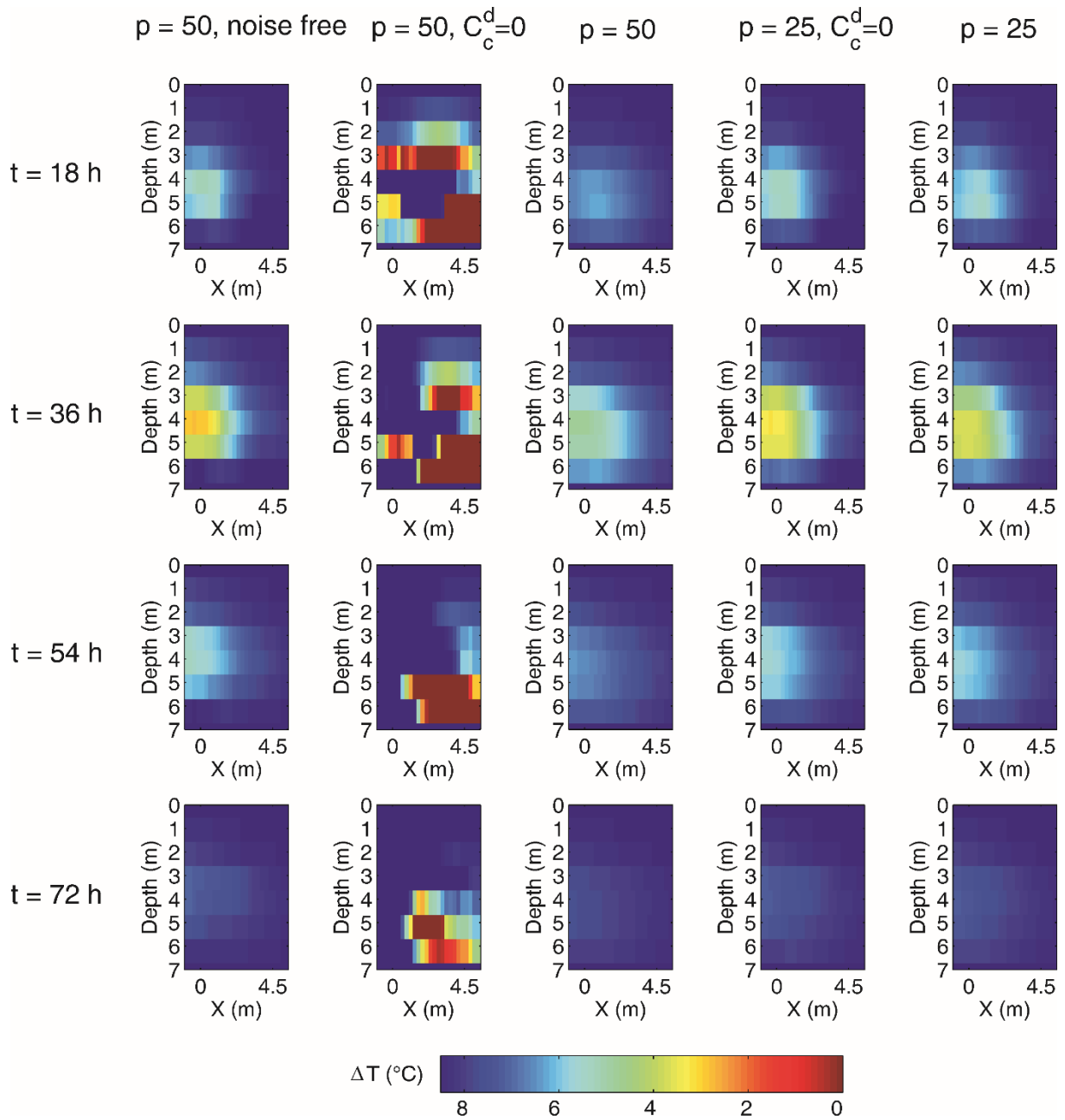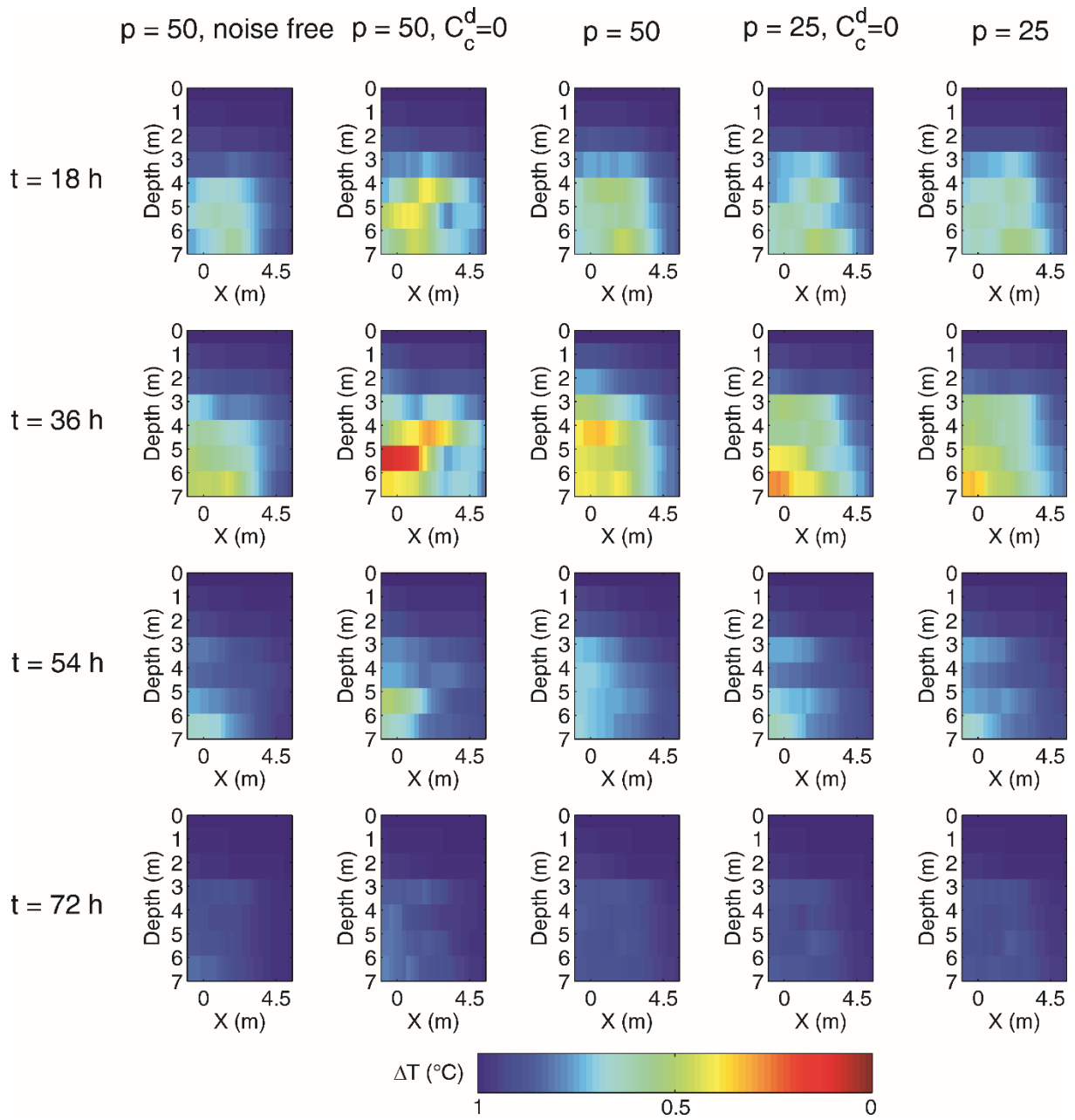952

Figure 4

954

Figure 5

46

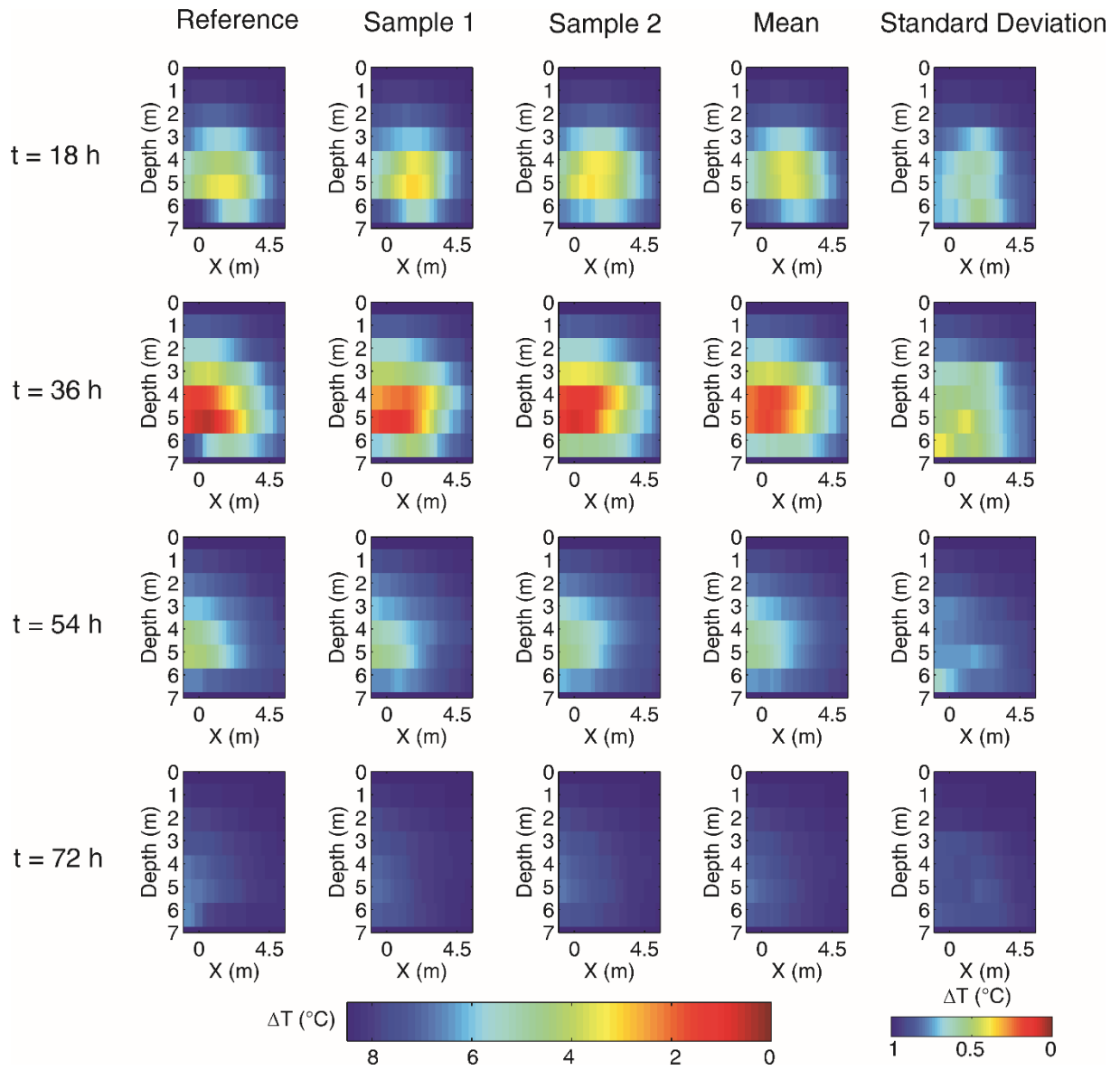p = 50, noise free p = 50, $C_c^d=0$ p = 50 p = 25, $C_c^d=0$ p = 25

t = 18 h

t = 36 h

t = 54 h

t = 72 h

ΔT (°C)

8   6   4   2   0

956

957   **Figure 6**

47

958

959 **Figure 7**

Columns: Reference, Sample 1, Sample 2, Mean, Standard Deviation

Rows: t = 18 h, t = 36 h, t = 54 h, t = 72 h

Axes: Depth (m) 0–7, X (m) 0–4.5

ΔT (°C) colorbar: 8, 6, 4, 2, 0
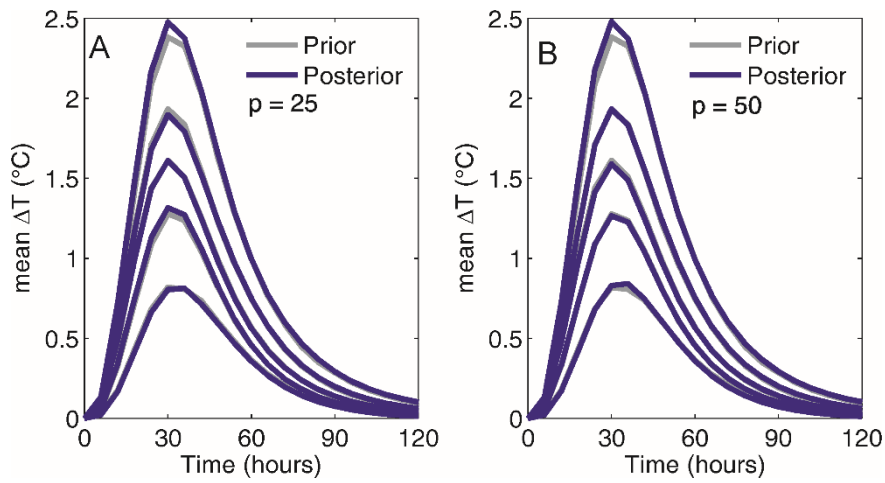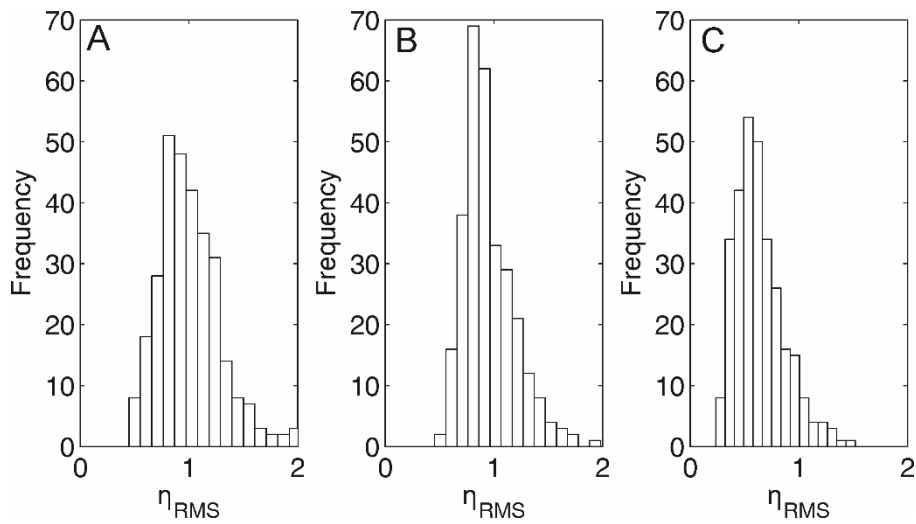
ΔT (°C) colorbar: 1, 0.5, 0

960

Figure 8

49

Figure 9

965

Figure 10

966

967

Figure 11

968

969

Figure 12

970

51