On the efficiency of the hybrid and the exact second-order sampling formulations of the EnKF: A reality-inspired 3D test case for estimating biodegradation rates of chlorinated hydrocarbons at the port of Rotterdam

M.E. Gharamti[1,3], J. Valstar[2], G. Janssen[2], A. Marsman[2], and I. Hoteit[1]

[1]King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia
[2]Deltares, Princetonlaan 6, 3584 CB, Utrecht, The Netherlands
[3]Nansen Environmental and Remote Sensing Center (NERSC), Bergen 5006, Norway

Correspondence to: M.E. Gharamti (mohamad.gharamti@nersc.no)

Abstract. This study considers the assimilation problem of subsurface contaminants at the port of Rotterdam in the Netherlands. It involves the estimation of solute concentrations and biodegradation rates of four different chlorinated solvents. We focus on assessing the efficiency of an adaptive hybrid ensemble Kalman filter (EnKF-OI) and the exact second-order sampling formulation ($EnKF_{ESOS}$) for mitigating the undersampling of the estimation and observation errors covariances, respectively. A multi-dimensional and multi-species reactive transport model is coupled to simulate the migration of contaminants within a Pleistocene aquifer layer located around 25 m below mean sea level. The biodegradation chain of chlorinated hydrocarbons starting from Tetrachloroethene and ending with Vinylchloride is modelled under anaerobic environmental conditions for five decades. Yearly pseudo-concentration data is used to condition the forecast concentration and degradation rates in presence of model and observational errors. Assimilation results demonstrate the robustness of the hybrid EnKF-OI, for accurately calibrating the uncertain biodegradation rates. When implemented serially, the adaptive hybrid EnKF-OI scheme efficiently adjusts the weights of the involved covariances for each individual measurement. The $EnKF_{ESOS}$ is shown to maintain much better the parameters ensemble spread leading to more robust estimates of the state and parameters. On average, a well tuned hybrid EnKF-OI and the $EnKF_{ESOS}$ respectively suggest around 48% and 21% improved concentration estimates and around 70% and 23% improved anaerobic degradation rates, over the standard EnKF. Incorporating large uncertainties in the flow model degrades the accuracy of the estimates of all schemes. Given that the performance of the hybrid EnKF-OI depends on the quality of the background statistics, satisfactory results were obtained only when the uncertainty imposed on the background information is relatively moderate.

1 Introduction

Subsurface contamination has received significant attention in the last few decades. Consequent cleanup costs have increased the awareness of environmental issues related to contaminated fields (Appelo and Postma, 1994; Drécourt et al., 2006). Historically, it was believed that subsurface contamination could be remediated to natural background contamination levels by

digging in the soil and pumping out the contaminated groundwater. However, it was not too long before it was discovered that there were simply too many contaminated areas to completely remediate. In addition, all available cleaning technologies, including source removal, are economically not viable to fully resolve the problem (Cunningham and Berti, 1993; Starr and Cherry, 1994; Todd and Mays, 2005).

Governmental authorities are now considering another approach to remediation based on management of industrial groundwater contamination at regional scales. The idea is simply to prevent groundwater contamination from causing negative effects on humans or ecology, and to control any undesired spreading beyond the boundary of the contaminated site. In the European Water Framework Directive, an option was offered allowing groundwater aquifers to remain contaminated when remediation is too costly and when an adequate monitoring system of the contaminated area is set up (Chave, 2001; Mostert, 2003; Hering et al., 2010). This procedure relies mostly on natural attenuation of contaminant plumes without the need for a direct human intervention. This is often possible given that the size and concentration of dissolved contaminants are frequently subjected to considerable decline due to natural, and eventually human induced, biodegradation processes. The challenge is then to predict in a cost effective way what type and when contaminants may cause a risk, so monitoring and, if needed, remediation may be undertaken to prevent any unacceptable spreading beyond specific planes of compliance. One efficient way to implement such monitoring system at a regional scale is to use prediction models with monitoring data and combine them using advanced data assimilation techniques (McLaughlin, 2002; Reichle et al., 2002).

Various numerical groundwater contaminant models have been developed in the literature (e.g., Freeze and Cherry, 1979; Pollock, 1994; Dawson et al., 2004; Sun and Wheeler, 2006; Bear and Cheng, 2010). The idea behind forming such models is to simulate and predict the dynamic fluxes and energies, defined as state variables (e.g. groundwater pressure, contaminant concentration), as accurately as possible based on some selected parameters (e.g., porosity, permeability, sorption) that describe the subsurface geometry, fluid and rock properties, and surface-subsurface interactions (Moradkhani et al., 2005).

Groundwater contaminant models can be subject to several sources of uncertainties due to poorly known parameters, inputs, and boundary conditions. For instance, we often know very little about the time at which contamination started, the amount of contaminant mass present in a pure phase source zone, the location of the pure phase, and the rate at which biodegradation is taking place (Franssen and Kinzelbach, 2009; Gharamti et al., 2013). Other uncertain aspects are the heterogeneity of the parameters, such as the hydraulic conductivity, groundwater recharge and the redox state of the groundwater. Therefore, model predictions of where and when a contaminant crosses a plane of compliance, with what concentration and how long it takes before a pure phase source zone dries up, can be quite uncertain.

One way to reduce uncertainty in model predictions and parameters is to assimilate data into the model. Data assimilation (DA) methods follow a Bayesian formulation by combining prior information of a dynamical system with available measurements to obtain an analysis of the system state and parameters (Hoteit et al., 2012; Gharamti et al., 2014a). Sequential DA techniques, such as the ensemble Kalman filter (EnKF), assimilate the data as it becomes available. The EnKF (Burgers et al., 1998; Evensen, 2003) is a popular DA method in hydrology, operating in consequent forecast and an analysis steps. During the forecast, an ensemble of state realisations is run forward in time using the dynamical model. At the time of the update, a linear Kalman filter (KF) type analysis (Kalman, 1960; Gharamti et al., 2011) is applied to the ensemble members. The EnKF

is relatively simple to implement, requiring only forward integrations of the dynamical and observational models. The EnKF has been proven useful in various subsurface hydrology applications (e.g., Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008; Zhou et al., 2011; Li et al., 2012; Crestani et al., 2013; Panzeri et al., 2013; Gharamti and Hoteit, 2014). The parameters most often calibrated are those characterising the flow and the general transport of the contaminants, such as permeability and porosity. Very few applications have tackled the estimation problem of reactive modelling parameters using sequential DA techniques. Bailey et al. (2013) used the EnKF to estimate spatially variable selenium and nitrate reaction rates in near-surface agricultural soil profiles. In another study, Bailey et al. (2012) used the ensemble smoother to infer the denitrification rate constants from synthetic observations of nitrate concentrations.

It is now widely recognised that the performance of the EnKF strongly depends on the ensemble size; a large enough ensemble is required to obtain good performances. Gharamti et al. (2014b) proposed an efficient hybrid EnKF assimilation scheme for state and parameters estimation, in which the predicted EnKF statistics are complemented with predefined static background covariance in order to mitigate for filter inbreeding and undersampling (Hamill and Snyder, 2000). The hybrid filter was applied to a small-scale synthetic reactive transport model and was found computationally efficient, providing reliable estimates using fairly small ensembles (50 members). In this study, we test the hybrid EnKF with a realistic large-scale contaminant model and further extend its formulation to allow for serial processing of the observations during the analysis step. For this, the objective function involved in the adaptive scheme is designed in such a way that the weighting between the background and the filter flow-dependent statistics is adjusted for each assimilated observation. Such updating strategy could be more convenient given that observations from different sources and locations carry varying degrees of information to the system. This generalises the adaptive scheme of Gharamti et al. (2014b), allowing the weighting between the ensemble and the background covariances to change not only between different assimilation times but also for different observations at any update step.

The stochastic EnKF assimilates perturbed observations during the analysis step in order to (asymptotically) match the second moment of the KF (Burgers et al., 1998). This often introduces noise, which may become significant when the rank of the observational error covariance is larger than the ensemble size (Nerger et al., 2005). Ensemble square root filters, such as the Ensemble Transform Kalman Filter (ETKF, Bishop et al., 2001), the Singular Evolutive Interpolated Kalman filter (SEIK, Pham, 2001; Hoteit et al., 2002), and the Deterministic Ensemble Kalman Filter (DEnKF, Sakov and Oke, 2008) do not require observations perturbations. Yet, the stochastic EnKF tends to "re-Gaussianize" the ensemble distribution, which improves the stability of the filter, unlike other deterministic schemes that follow the shape of the background distribution (Lawson and Hansen, 2004). In a recent study, Hoteit et al. (2015) proposed a serial EnKF algorithm to mitigate the observation sampling errors in the EnKF. The algorithm, referred to as EnKF with exact second-order observation perturbation sampling ($\text{EnKF}_{\text{ESOS}}$), is straightforward to implement in any existing serial EnKF code, requiring only removing a single rank from the sample forecast covariance matrix to exactly match the first two moments of the KF. Compared to the EnKF and the determinstic filters, the $\text{EnKF}_{\text{ESOS}}$ was shown to provide more accurate estimates of the state of a 40-variable Lorenz-96 model. Here, we consider the $\text{EnKF}_{\text{ESOS}}$ in a realistic large-scale system and further study the impact of mitigating the observation undersampling errors in the EnKF's analysis not only on the state but also on the parameters estimates. This is the first study

addressing the application of the $\text{EnKF}_{\text{ESOS}}$ scheme for parameters estimation. The idea is to investigate whether accounting for observation sampling errors can lead to a better tuning of the unknown parameters or not.

We focus on two aspects that are known to limit the efficiency of the EnKF; namely the undersampling of the forecast errors in the forecast step and the observation errors in the analysis step. We consider an industrial groundwater contamination problem at the port of Rotterdam in the Netherlands. Many areas at the port site are contaminated due to various industrial activities. Contamination with chlorinated hydrocarbons (CH) have been detected at the port area. Reductive dechlorination process of four hazardous CH components; namely Tetrachloroethene (PCE), Trichloroethene (TCE), 1,2-Dichloroethene (DCE) and Vinyl Chloride (VC), is believed to be one of the main reactive processes taking place at the port site. We simulate this process using a coupled three-dimensional Flow-Transport-Reaction (3D-FTR) model for a single plume. The contaminant data collected in 2012 by the municipality of Rotterdam, is used for initialising the contaminant migration, which propagates to surface and deep aquifer layers (≈ 50 m below sea level). We use "synthetic" CH concentration data on a yearly basis, for a total of 50 years, to calibrate four biodegradation rates of the reaction chain. To the best of our knowledge, this is the first study in which biodegradation parameters of a reductive dechlorination process are estimated in a real-world system using a sequential DA procedure. Hydraulic parameters of the groundwater flow model are not incorporated in the assimilation system as part of the unknown parameters. They are estimated in an offline procedure using borehole data and a probabilistic approach given the associated geologic lithofacies. On top of the biodegradation, the concentrations of the components are also constrained using the EnKF, the hybrid EnKF and the $\text{EnKF}_{\text{ESOS}}$ schemes. Sensitivity analyses are performed to study the efficiency and the accuracy of the assimilation schemes under different experimental settings. The filtering schemes are evaluated based on the accuracy of the estimated solute concentrations, the handling of the posterior distributions of the biodegradation rates, and computational complexity.

The rest of this paper is organised as follows. Section 2 presents the ensemble filtering schemes. Section 3 describes the large-scale subsurface reactive transport model and its numerical implementation. Section 4 presents the assimilation setup and the experimental scenarios. Results of the assimilation experiments are presented and analysed in Section 5. Conclusions and further discussion are given in Section 6.

2 The Data Assimilation Framework

The aim of DA is to combine measured observations and a dynamical model in order to compute the best possible estimates of the past, current and future states of the system, together with estimates of the associated uncertainties (Nichols, 2010). We follow the standard discrete nonlinear dynamical system:

$$\mathbf{x}_{k+1} = \mathcal{M}_k\left(\mathbf{x}_k, \boldsymbol{\Theta}_k\right) + \boldsymbol{\eta}_{k+1}, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^{N_\mathbf{x}}$ denotes a state vector of $N_\mathbf{x}$ variables at time t_k, $\boldsymbol{\Theta}_k \in \mathbb{R}^{N_\boldsymbol{\Theta}}$ is the vector of model parameters, $\mathcal{M}_k$: $\mathbb{R}^{N_\mathbf{x}} \rightarrow \mathbb{R}^{N_\mathbf{x}}$ is the nonlinear operator that propagates the model state from t_k to t_{k+1}. $\boldsymbol{\eta}_{k+1} \in \mathbb{R}^{N_\mathbf{x}}$ is a model error accounting for model uncertainties, commonly assumed to follow a Gaussian distribution $\mathcal{N}\left(\mathbf{0}, \mathbf{Q}_{k+1}\right)$. The measurements obey the following

observational system:

$$\mathbf{y}_{k+1} = \mathcal{H}_{k+1}(\mathbf{x}_{k+1}) + \boldsymbol{\varepsilon}_{k+1}, \tag{2}$$

where $\mathbf{y}_{k+1} \in \mathbb{R}^{N_\mathbf{y}}$ is a vector of $N_\mathbf{y}$ observations at time t_{k+1}, $\mathcal{H}_{k+1}$: $\mathbb{R}^{N_\mathbf{y}} \rightarrow \mathbb{R}^{N_\mathbf{x}}$ is an observational map including grid interpolations, and could be nonlinear. The observation errors $\boldsymbol{\varepsilon}_{k+1} \in \mathbb{R}^{N_\mathbf{y}}$ are assumed Gaussian with zero mean and covariance $\mathbf{R}_{k+1}$. We also assume independent model and observation errors.

Following the Bayesian filtering problem, the objective is to evaluate the joint probability density function (pdf), i.e., $p(\mathbf{x}_k, \boldsymbol{\Theta}_k | \mathbf{y}_{0:k})$, of the system state $\mathbf{x}_k$ and the parameters $\boldsymbol{\Theta}_k$ given all available observations $\mathbf{y}_{0:k}$. The observations, $\mathbf{y}_{0:k}$, are used to update the model forecast. The updated estimate is then used to compute a future prediction. Likewise, the estimation problem can be also tackled using variational approaches that involve minimisation of a cost function (Dimet and Talagrand, 1986; Courtier et al., 1994; Hoteit et al., 2005; Altaf et al., 2013). Variational DA techniques, such as 3DVar and 4DVar, are widely used in geoscience applications. These methods look for an optimal state trajectory that best fits observational data over a time window, but do not offer an efficient framework for quantifying uncertainties in the solution. In this study, we will only consider the sequential Bayesian filtering problem.

2.1 The Ensemble Kalman Filter for State-Parameters Estimation

The computation of $p(\mathbf{x}_k, \boldsymbol{\Theta}_k | \mathbf{y}_{0:k})$ is not feasible in real applications owing to the nonlinear character of the model and observation operators in addition to the very large dimension of the subsurface flow and transport system. The ensemble Kalman filter (EnKF) is an efficient Monte Carlo method that computes an approximation of the joint pdf, using the first two moments, at reasonable computational requirements. The EnKF represents the distribution of the system using a collection of state vectors, called ensemble. Generally, the true pdf of the system might not be accessible through this Monte Carlo approximation given the finite ensemble size. We follow the state-parameters augmentation procedure (Annan et al., 2005) and denote by ψ the jointly concatenated state and parameters vector. The parameters are time-invariant so that their time-propagation function is simply the identity operator.

To illustrate, starting at time t_{k-1} from an analysis ensemble, $\left\{\boldsymbol{\psi}_{k-1}^{a,i}: \mathbf{x}_{k-1}^{a,i}, \boldsymbol{\Theta}_{k-1}^{a,i}\right\}_{i=1}^{N_e}$, which represents $p(\boldsymbol{\psi}_{k-1} | \mathbf{y}_{0:k-1})$, the EnKF propagates the dynamical model (1) to compute the forecast ensemble at the time of the next available observation, t_k. Incoming measurements are then used to update the joint ensemble. The EnKF algorithm is summarised below.

– Forecast Step: The analysis members are integrated forward in time to obtain the forecast ensemble from which we estimate the first two moments as follows:

$$\widehat{\boldsymbol{\psi}}_k^f = \begin{bmatrix} \widehat{\mathbf{x}}_k^{f,i} \\ \widehat{\boldsymbol{\Theta}}_k^{f,i} \end{bmatrix} = \frac{1}{N_e} \begin{bmatrix} \sum_{i=1}^{N_e} \mathcal{M}\left(\mathbf{x}_{k-1}^{a,i}, \boldsymbol{\Theta}_{k-1}^{a,i}\right) \\ \sum_{i=1}^{N_e} \boldsymbol{\Theta}_{k-1}^{a,i} \end{bmatrix} \equiv \frac{1}{N_e} \sum_{i=1}^{N_e} \boldsymbol{\psi}_k^{f,i}, \tag{3}$$

$$\widehat{\mathbf{P}}_k^f = \begin{bmatrix} \widehat{\mathbf{P}}_{xx}^f & \widehat{\mathbf{P}}_{x\theta}^f \\ \widehat{\mathbf{P}}_{\theta x}^f & \widehat{\mathbf{P}}_{\theta\theta}^f \end{bmatrix} \equiv \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(\boldsymbol{\psi}_k^{f,i} - \widehat{\boldsymbol{\psi}}_k^f\right) \left(\boldsymbol{\psi}_k^{f,i} - \widehat{\boldsymbol{\psi}}_k^f\right)^T. \tag{4}$$

The joint sample covariance $\widehat{\mathbf{P}}_k^f$ consists, as shown in (4), of the sample state covariance $\widehat{\mathbf{P}}_{xx}^f$, the state-parameters cross-correlation $\widehat{\mathbf{P}}_{\theta x}^f$ and the sample parameters covariance $\widehat{\mathbf{P}}_{\theta\theta}^f$ matrices. The joint state-parameters forecast estimate (mean) is denoted by $\widehat{\boldsymbol{\psi}}_k^f$. The complexity of the forecast step grows with the ensemble size. If one supposes that $\mathcal{C}_M$ is the cost for integrating the model to the next observation time, the computational requirement of the forecast step is $NN_e\mathcal{C}_M$, where N is the final simulation time (Gharamti et al., 2014a).

– Analysis Step: When the observation $\mathbf{y}_k$ becomes available, the joint forecast members $\boldsymbol{\psi}_k^{f,i}$ are updated using the Kalman-update step; i.e.

$$\boldsymbol{\psi}_k^{a,i} = \boldsymbol{\psi}_k^{f,i} + \mathbf{K}\left(\mathbf{y}_k + \boldsymbol{\epsilon}_k^i - \widetilde{\mathbf{H}}_k\boldsymbol{\psi}_k^{f,i}\right), \tag{5}$$

where $\mathbf{K} = \widehat{\mathbf{P}}_k^f\widetilde{\mathbf{H}}_k^T\left(\widetilde{\mathbf{H}}_k\widehat{\mathbf{P}}_k^f\widetilde{\mathbf{H}}_k^T + \mathbf{R}_k\right)^{-1}$ is the Kalman gain and the analysis state is:

$$\widehat{\boldsymbol{\psi}}_k^a = \frac{1}{N_e}\sum_{i=1}^{N_e}\boldsymbol{\psi}_k^{a,i} \equiv \widehat{\boldsymbol{\psi}}_k^f + \mathbf{K}\left(\mathbf{y}_k + \widehat{\boldsymbol{\epsilon}}_k - \widetilde{\mathbf{H}}_k\widehat{\boldsymbol{\psi}}_k^f\right), \qquad \text{with} \qquad \widehat{\boldsymbol{\epsilon}}_k = \frac{1}{N_e}\sum_{i=1}^{N_e}\boldsymbol{\epsilon}_k^i. \tag{6}$$

The observation perturbations, denoted by $\boldsymbol{\epsilon}_k^i$, are sampled from a Gaussian distribution of zero mean and covariance $\mathbf{R}_k$. The observational operator $\widetilde{\mathbf{H}}_k = \left[\mathbf{H}_k, \mathbf{O}\right]$, acting on the augmented state-parameter vector, is assumed linear for simplicity, and the matrix $\mathbf{O}$ is a zeros matrix. Computationally, the update step in hydrological applications is usually less demanding than the forecast step, with a complexity of $NN_eN_{\mathbf{y}}N_{\mathbf{x}} + NN_e^2\left(N_{\mathbf{x}} + N_{\boldsymbol{\Theta}}\right)$. The observations used in the update equation of (5) are processed in one single batch. In our implementation, we will consider the serial EnKF update formulation in which the observations are assimilated one at a time. The reason for this will become clear in section 2.3.

2.1.1 EnKF Limitations

The performance of the EnKF strongly depends on the accuracy of the forecast error covariance matrix $\widehat{\mathbf{P}}^f$. The errors in $\widehat{\mathbf{P}}^f$ are essentially due to: (1) model errors and the use of small ensemble sizes, and (2) propagation of errors in the sample covariance matrix $\widehat{\mathbf{P}}^a$ at the previous step. The Gaussian assumption of the system's distribution is also a limiting factor but this was proven to be less problematic (Hoteit et al., 2008). The Gaussianity of the estimates often breaks when the parameters are also included as part of the state vector during assimilation (e.g., Liu et al., 2016).

The main advantage of the ensemble approximation (Eqs. 3 and 4) is that it does not involve any linearisation and allows to represent the first two moments of the state and parameters by an ensemble of vectors (Evensen, 2003). The use of large ensembles is practically not possible and thus the sample covariance $\widehat{\mathbf{P}}_k^f$ may not well approximate the KF forecast covariance, $\mathbf{P}_k^f$. As such, the joint forecast pdf of the system's state and parameters at any time t_k is only partially sampled, which means that there exists a null subspace in the error space that is not covered by the ensemble (Song et al., 2010; Mandel et al., 2011). To mitigate this, we will a use hybrid formulation of the forecast state and parameters statistics before performing the EnKF update (e.g., Wang et al., 2007). Further details are given in section 2.2.

The limited ensemble size may also introduce noise in the update step of the EnKF when the rank of the observation error covariance is large (Hoteit et al., 2015). This is because the number of observation perturbations may not be enough to sample the observation error covariance matrix, $\mathbf{R}_k$. In addition, spurious correlations between the observation and the forecast perturbations may also introduce noise in the EnKF update (e.g., Bowler et al., 2013; Hoteit et al., 2015). To illustrate, the EnKF analysis assumes zero cross-correlations between the observation perturbations and the forecast ensemble; i.e.:

$$\sum_{i=1}^{N_e} \boldsymbol{\epsilon}_k^i \left(\boldsymbol{\psi}_k^{f,i} - \widehat{\boldsymbol{\psi}}_k^f \right)^T = \mathbf{0}. \tag{7}$$

This can be easily seen by subtracting eq. (5) from eq. (6). After arranging the terms and using eq. (4), one obtains:

$$\boldsymbol{\Delta} = \left(\mathbf{I} - \mathbf{K}\widetilde{\mathbf{H}}_k \right) \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(\boldsymbol{\psi}_k^{f,i} - \widehat{\boldsymbol{\psi}}_k^f \right) {\boldsymbol{\epsilon}_k^i}^T \mathbf{K}^T + \mathbf{K} \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \boldsymbol{\epsilon}_k^i \left(\boldsymbol{\psi}_k^{f,i} - \widehat{\boldsymbol{\psi}}_k^f \right)^T \left(\mathbf{I} - \mathbf{K}\widetilde{\mathbf{H}}_k \right)^T, \tag{8}$$

$$\widehat{\mathbf{P}}_k^a = \left(\mathbf{I} - \mathbf{K}\widetilde{\mathbf{H}}_k \right) \widehat{\mathbf{P}}_k^f \left(\mathbf{I} - \mathbf{K}\widetilde{\mathbf{H}}_k \right)^T + \mathbf{K}\mathbf{R}_k\mathbf{K}^T + \boldsymbol{\Delta}, \tag{9}$$

where $\boldsymbol{\Delta}$ is the sampling error term; not accounted for in the EnKF. Consequently, the ensemble analysis covariance matches the optimal KF covariance, $\mathbf{P}_k^a$, only when the observational sampling errors and the cross-correlation terms in $\boldsymbol{\Delta}$ are indeed zero. This can be numerically achieved by assimilating the observations serially using the so-called EnKF with exact second-order perturbations sampling, $\text{EnKF}_{\text{ESOS}}$, as will be discussed in more details in section 2.3.

2.2 The Hybrid EnKF

The hybrid EnKF and optimal interpolation (EnKF-OI) scheme was introduced as a way to mitigate for small ensemble sizes and model deficiencies in the EnKF (Hamill and Snyder, 2000). Using small ensembles results in rank deficient forecast covariance matrices, which strongly limit the fit to the observations (Song et al., 2010). Neglecting model errors might further produce small ensemble spread, and consequently unrealistic confidence in the forecast (Song et al., 2013). The standard solution for rank deficiency or covariance underestimation is to apply inflation and localization. Inflation artificially inflates the spread of the ensemble around the mean state (Hamill et al., 2001; Hoteit et al., 2002). It is also a simple way to account for neglected model errors (Pham et al., 1998). Covariance localization eliminates spurious correlations by a Schur product multiplication of the under-sampled covariance matrix with a function of local support (Houtekamer and Mitchell, 2001; Sakov and Bertino, 2011). Inflation and localization, although efficient and widely used (especially in atmosphere and ocean application), are generally model dependent and require important tuning efforts. They further do not introduce any new directions to diversify the ensemble, limiting the filter update to a small-dimensional ensemble subspace (Song et al., 2010, 2013). Moreover, global model parameters are not local quantities and therefore localization techniques might not be as straightforward (Devegowda et al., 2007). In addition, the parameters are dynamically constant quantities (static in time), and thus large ensembles are usually needed to well approximate the parameters distributions (Hendricks Franssen and Kinzelbach, 2008; Zhou et al., 2012).

The hybrid approach estimates the EnKF's forecast error covariance by a weighted sum of the ensemble covariance and a stationary covariance matrix, typically used in a variational or an optimal interpolation (OI) assimilation system. More

specifically, the background state-state and state-parameters covariances are estimated as:

$$\widetilde{\mathbf{P}}_{xx}^{\text{Hybrid}} = \alpha\widehat{\mathbf{P}}_{xx}^{\text{EnKF}} + (1-\alpha)\mathbf{P}_{xx}^{b}, \tag{10a}$$

$$\widetilde{\mathbf{P}}_{\theta x}^{\text{Hybrid}} = \beta\widehat{\mathbf{P}}_{\theta x}^{\text{EnKF}} + (1-\beta)\mathbf{P}_{\theta x}^{b}, \tag{10b}$$

where $\widehat{\mathbf{P}}_{xx}^{\text{EnKF}}$ and $\widehat{\mathbf{P}}_{\theta x}^{\text{EnKF}}$ are the sample covariance and cross-correlation matrices of the EnKF ensemble, respectively. The background covariances are denoted by $\mathbf{P}_{xx}^{b}$ and $\mathbf{P}_{\theta x}^{b}$, respectively. It was indeed shown by Hamill and Snyder (2000) that this additional stationary background covariance may help representing part of the ensemble's null space that is not described by the limited ensemble. This procedure is based on physically reliable statistics, although flow-independent, unlike inflation and localization (Wang et al., 2009). The scalar quantities α and β are weighting factors, taking values between 0 and 1.

2.2.1 Practical Implementation

The static background covariance, $\mathbf{P}_{xx}^{b}$, is often built on the basis of a long inventory of forecast errors (Wang et al., 2009). It is usually assumed to be of low-rank, r_x, and can be factorised into spectral modes using Proper Orthogonal Decomposition (POD) as follows;

$$\mathbf{P}_{xx}^{b} = \mathbf{S}\boldsymbol{\Omega}\mathbf{S}^{T} = \mathbf{S}\boldsymbol{\Omega}^{\frac{1}{2}}\left(\mathbf{S}\boldsymbol{\Omega}^{\frac{1}{2}}\right)^{T} = \widehat{\mathbf{S}}\widehat{\mathbf{S}}^{T}, \tag{11}$$

where $\mathbf{S}$ is a matrix of spectral coefficients, $\boldsymbol{\Omega}$ carries information about the associated spectral variances and $\boldsymbol{\Omega}^{\frac{1}{2}}$ is its Cholesky decomposition of $\boldsymbol{\Omega}$. The background perturbation matrix, $\widehat{\mathbf{S}}$, has r_x columns, with r_x much smaller than the number of state variables. The background state and parameters cross-covariance, $\mathbf{P}_{\theta x}^{b}$, can be also approximated by a low-rank, r_θ, matrix using singular value decomposition (SVD) if the number of parameters is not equal to the number of state variables (and thus the matrix $\mathbf{P}_{\theta x}^{b}$ is not square). This decomposition is useful in practice in order to reduce computational burden and memory storage. Accordingly, the complexity of the analysis step (referred to as $\mathcal{O}^{a}$) in the hybrid EnKF-OI scheme becomes:

$$\begin{aligned}\mathcal{O}_{\text{EnKF-OI}}^{a} &= NN_eN_{\mathbf{y}}N_{\mathbf{x}} + NN_e^2\left(N_{\mathbf{x}} + N_{\boldsymbol{\Theta}}\right) + NN_e\left(N_{\mathbf{x}}r_x + N_{\boldsymbol{\Theta}}r_\theta\right), \\ &= \mathcal{O}_{\text{EnKF}}^{a} + NN_e\left(N_{\mathbf{x}}r_x + N_{\boldsymbol{\Theta}}r_\theta\right).\end{aligned} \tag{12}$$

Given that r_x and r_θ are usually small in subsurface flow and transport problems (Gharamti et al., 2014b), the complexity of the analysis step of the hybrid EnKF-OI is only marginally larger than that of the EnKF. The complexity of the forecast step of the EnKF and the hybrid EnKF-OI is the same when both are implemented with the same ensemble size.

The weighting factors α and β need to be specified in eqs. (10a) and (10b). Careful tuning of α and β is very important (Hamill and Snyder, 2000). The simplest way is to select them based on trial and error but this can be computationally very intensive. A more efficient approach was introduced by Gharamti et al. (2014b) and consists of optimising a one-dimensional (1D) objective function at every update step of the state and the parameters. Based on Kalman's update formulation, assimilating observations causes the uncertainties in the prior estimates to shrink. Thus, using the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951), one can choose α and β that maximise the information gains at the analysis time t_k. In this study,

we opt to assimilate the observations serially and thus one can adaptively compute optimal weighting factors as follows:

$$\begin{aligned}
\underset{\alpha}{\arg\max}\ \ \mathcal{F}(\alpha) \quad &= \quad \underset{\alpha}{\arg\max}\ \ tr\left[\widetilde{\mathbf{P}}^{f}_{xx} - \widetilde{\mathbf{P}}^{a}_{xx}\right], \\
&= \quad \underset{\alpha}{\arg\max}\ \ tr\left[\widetilde{\mathbf{P}}^{f}_{xx}\mathbf{H}^{T}\left(\mathbf{H}_k\widetilde{\mathbf{P}}^{f}_{xx}\mathbf{H}_k^{T} + \mathbf{R}_k\right)^{-1}\mathbf{H}_k\widetilde{\mathbf{P}}^{f}_{xx}\right], \\
&\overset{\text{single observation}}{\equiv} \quad \underset{\alpha}{\arg\max}\ \ \frac{1}{d}\sum_{m=1}^{N_{\mathbf{x}}}\left(\mathbf{c}^{[m]}_{xx}\right)^{2},
\end{aligned} \tag{13}$$

where $tr\left[\cdot\right]$ denotes the trace of a matrix and d is a scalar quantity equivalent to observation variance $\left(\mathbf{H}_k\widetilde{\mathbf{P}}^{f}_{xx}\mathbf{H}_k^{T} + \mathbf{R}_k\right)$ when assimilating one observation. $\mathbf{c}^{[m]}_{xx}$ is the m^{th} forecast variance-component corresponding to one observed variable. Similarly, one can define the objective function for the parameters' weighting factor as follows:

$$\begin{aligned}
\underset{\beta}{\arg\max}\ \ \mathcal{G}(\beta) \quad &= \quad \underset{\beta}{\arg\max}\ \ tr\left[\widetilde{\mathbf{P}}^{f}_{\theta\theta} - \widetilde{\mathbf{P}}^{a}_{\theta\theta}\right], \\
&\overset{\text{single observation}}{\equiv} \quad \underset{\beta}{\arg\max}\ \ \frac{1}{d}\sum_{m=1}^{N_{\boldsymbol{\Theta}}}\left(\mathbf{c}^{[m]}_{\theta x}\right)^{2},
\end{aligned} \tag{14}$$

where $\mathbf{c}^{[m]}_{\theta x}$ is the forecast cross-correlation component between the m^{th} parameter and one observed variable. Such KL criterion describes the information gain from each individual observation as it reflects the difference between the prior and the posterior distributions. The interesting point here is that for each observation, different weights would be assigned to the background and the ensemble statistics. The maximisation problems in (13) and (14) are 1D and bounded, yielding minimal forecast variance after the update. In terms of implementation, we perform the optimisation, over the interval $[0, 1]$, using a computationally efficient scheme combining both golden-section search and repeated parabolic interpolation (Forsythe et al., 1977).

2.3 Exact Second-Order Observation Perturbations Sampling

The sampling error from neglecting the cross-correlation terms in eq. (9) in the EnKF analysis is generally not globally small. It is often composed of a large number of elements that can add up after successive assimilation steps (Hoteit et al., 2015). This may degrade the filter's accuracy and increases the underestimation of the analysis error covariance (Whitaker and Hamill, 2002). Furthermore, such sampling errors can propagate to subsequent steps, eventually deteriorating the performance of the filter.

In a mathematical sense, for the condition in eq. (7) to hold, the rank of the forecast perturbation matrix

$$\Psi^{f}_{k} = \left[\boldsymbol{\psi}^{f,1}_{k} - \widehat{\boldsymbol{\psi}}^{f}_{k}, \boldsymbol{\psi}^{f,2}_{k} - \widehat{\boldsymbol{\psi}}^{f}_{k}, \ldots, \boldsymbol{\psi}^{f,N_e}_{k} - \widehat{\boldsymbol{\psi}}^{f}_{k}\right] \tag{15}$$

plus the rank of $\mathbf{R}_k$ must not exceed $N_e - 1$, which is essentially the rank of Ψ^{f}_{k} (Pham, 2001). Obviously, this is not possible given that $N_{\mathbf{y}} + N_e - 1$ is always greater than $N_e - 1$. Yet, if we suppose that Ψ^{f}_{k} has a rank $N_e - 2$, then when $\mathbf{R}_k$ is scalar, it is possible to draw the observation perturbations $\boldsymbol{\epsilon}^{i}_{k}$ such that the EnKF's analysis first and second moments are exactly the same as those computed using the KF. Accordingly, Hoteit et al. (2015) proposed to remove one rank from Ψ^{f}_{k} using an SVD decomposition:

$$\boldsymbol{\psi}^{f,i}_{k} \leftarrow \boldsymbol{\psi}^{f,i}_{k} - \left(\Psi^{f}_{k}\mathbf{w}_k\right)w^{i}_{k}, \tag{16}$$

where $\mathbf{w}_k$ is the normalised right singular vector of Ψ_k^f associated with the smallest nonzero singular value. The i^{th} component of $\mathbf{w}_k$ is denoted by w_k^i and the symbol $\leftarrow$ means "replaced by." Then, assimilating the observations serially and simply choosing $\boldsymbol{\epsilon}_k^i = \sqrt{(N_e-1)\,\mathbf{R}_k} w_k^i$ would guarantee zero cross-correlations between the modified forecast perturbations and the observation perturbations. The algorithm, referred to as EnKF$_{\text{ESOS}}$, involves a recursive update for $\mathbf{w}_k$ after each update (Hoteit et al., 2015). The serial analysis procedure of the EnKF$_{\text{ESOS}}$ is summarised in the algorithm below:

While $j \leq N_{\mathbf{y}}$ **do**

1: $\widehat{z}_k = \mathbf{H}_k^{[j]} \widehat{\boldsymbol{\psi}}_k$

2: $z_k^{[j],i} = \mathbf{H}_k^{[j]} \widehat{\boldsymbol{\psi}}_k^{a,i}$

3: $\mathbf{K}^{[j]} = \sum_{i=1}^{N_e} \left(\boldsymbol{\psi}_k^{f,i} - \widehat{\boldsymbol{\psi}}_k^f\right) \left(z_k^{[j],i} - \widehat{z}_k\right) \left[\sum_{i=1}^{N_e} \left(z_k^{[j],i} - \widehat{z}_k\right) + (N_e-1)\mathbf{R}_k^{[j,j]}\right]^{-1}$

4: **For** i in $[1, 2, \ldots, N_e]$ **do**

$$\psi_k^{a,i} \leftarrow \psi_k^{a,i} + \mathbf{K}^{[j]} \left(\mathbf{y}_k^{[j]} + s\sqrt{(N_e-1)\,\mathbf{R}_k^{[j,j]}} w_k^i - z_k^{[j],i}\right)$$

5: **EndFor**

6: $\widehat{\psi}_k^a \leftarrow \dfrac{1}{N_e} \sum_{i=1}^{N_e} \psi_k^{a,i}$

7: **For** i in $[1, 2, \ldots, N_e]$ **do**

$$w_k^i \leftarrow s\sqrt{(N_e-1)\,\mathbf{R}_k^{[j,j]}} w_k^i - \left(z_k^{[j],i} - \widehat{z}_k\right) \left[\sum_{i=1}^{N_e} \left(z_k^{[j],i} - \widehat{z}_k\right)^2 + (N_e-1)\,\mathbf{R}_k^{[j,j]}\right]^{-\frac{1}{2}}$$

8: **EndFor**

EndWhile

where s is an independent plus or minus sign. The superscript $[j]$ denotes the j^{th} element and row of the given vector and matrix, respectively. The superscript $[j,j]$ denotes the element in row and column j of the associated matrix. Note that unlike the EnKF, the observation perturbations cannot be Gaussian because of the constraint they satisfy in eq. (7). In the experiments of Hoteit et al. (2015), these were shown to be almost Gaussian. In term of complexity, the EnKF$_{\text{ESOS}}$ algorithm has almost the same computational cost as that of the serial EnKF. Additional cost is required for iteratively updating the vector $\mathbf{w}_k$ and performing an SVD on $\boldsymbol{\Psi}_k$ to reduce its rank by one. Both operations are computationally almost negligible compared to the cost of integrating the subsurface model.

3 The Subsurface Model and Assimilation Experiments

3.1 The Rotterdam Port and Geology of the Area

The Port of Rotterdam is located in the Netherlands between the city of Rotterdam and the North Sea. It is the largest port of Europe covering an area of 105 km^2 and stretching over a distance of 40 km. The original geology of the area consists

of a top Holocene layer of approximately 20 m thick (Figure 1). It is composed of clay and peat with local sandy channel deposits, but in the most western part, it becomes sandier. Under the Holocene layer, there is a Pleistocene aquifer of coarse sand of approximately 10 m thick. Below lays a Pleistocene clay layer of approximately 30 m thick and a second aquifer of approximately 140 m thick. The second aquifer is saline for most of the port areas whereas the first aquifer is partly saline in the western part only. On top of the Holocene sediments, an anthropogenic layer of fine sand was added up to a level of 4 m (eastern part) to 6 m (western part) above the mean sea level. Moreover, locally a dense network of sand filled vertical drains was used in the upper part of the Holocene clay in order to speed up the settling of the clay. A large part of the industrial port area is surrounded by surface water, some of which continue to the bottom of the Holocene layer.

At the port site, more than 600 companies perform various activities such as trans-shipment of containers (coal, oil, gas, etc), storage of oils and chemicals, building/repairing ships and oil/gas rigs, distribution and transport inland, and disposal/treatment of chemical wastes. As a result of the long-term presence of these industrial activities, the soil and groundwater have become contaminated. This contamination is substantial, complex, and not limited to one particular site but affects the groundwater systems at a regional scale (Marsman et al., 2006; Ter Meer et al., 2007). Part of the contaminants are non-mobile such as heavy metals including arsene, cadmium, copper, mercury, lead and zinc. Other mobile contaminants are mineral oils, volatile aromatics, chlorinated solvents and pesticides.

3.2 Coupled 3D Subsurface Model

3.2.1 Organic Contaminants

Wells monitoring and lab analysis have concluded that groundwater at the port area is contaminated, at different depth, with varying levels of pollutants (Marsman et al., 2006). One of the major contaminants are chlorinated hydrocarbons that had entered the subsurface as Dense Non-aqueous Phase Liquids (DNAPL) and often have source zones of stagnant pure phases at considerable depth. Numerous industrial companies at the port manufacture or work with these organic molecules. Here, we simulate the degradation chain of four CH components; namely Tetrachloroethene (PCE, a.k.a perchloroethene), Trichloroethene (TCE), 1,2-Dichloroethene (DCE) and Vinyl Chloride (VC). We use plume data from a real site, but for confidentiality reasons we do not show the exact location of the site. The horizontal area of the domain is equal to 1.5 km^2, extending 1 km in the transverse direction and 1.5 km in the longitudinal direction (Figure 2). Degradation of the dissolved components takes place as chlorine atoms are subsequently replaced by hydrogen atoms under anaerobic environmental conditions (Vogel and McCARTY, 1985; Clement et al., 2000; Tobiszewski and Namieśnik, 2012). Chlorinated hydrocarbons can pose serious threat to human and environmental health (Ojajärvi et al., 2001; Lee et al., 2002, 2003).

3.2.2 Flow-Transport-Reaction Model (FTR-Model)

The subsurface model consists of three major components; namely flow, transport and reactions. First, the groundwater flow (assumed steady) is solved on a rectangular domain using MODFLOW (Harbaugh, 2005). The steady groundwater flow assumption is valid at the current port location. Temporal variations, such as tidal influences and yearly fluctuations of precipi-

tation and evapotranspiration are expected to happen, but on a small scale. Essentially, tidal influences and yearly fluctuations of precipitation and evapotranspiration are expected to be minor as the near surface groundwater levels are controlled by the drainage levels of the drainage systems in the port area (3-4 m above sea level). The deeper groundwater levels are predominantly influenced by surface water levels in the polders area (managed levels around or below sea level) and the large surface waters (approximately sea level). Temporal variations due to density driven flow are also neglected as we would expect only minor changes in the most lower part of the model domain on the time scale of 50 years. MT3DMS is used to solve the advection-dispersion based transport of the components (Zheng and Wang, 1999), in which the degradation process of the components is added based on the module within the 3D-multispecies reactive package; RT3D (Clement, 1997). The softwares are integrated in a sophisticated fortran-based tool (with graphical interface) called iMOD (Vermeulen et al., 2013). In differential form, the fate and transport of the components is modelled following:

$$\left(\phi + \rho_b k^\ell\right) \frac{\partial \mathbf{C}^\ell}{\partial t} + \lambda \phi \mathbf{C}^\ell = \nabla \cdot \left(\phi \mathcal{D} \nabla \mathbf{C}^\ell\right) - \nabla \cdot \left(\nu \mathbf{C}^\ell\right) + q_s \mathbf{C}_s^\ell + r_{\mathbf{C}}, \tag{17}$$

where ϕ is porosity, ρ_b is the bulk density of the soil, k is the distribution (sorption) coefficient, $\mathbf{C}$ is the solute concentration, λ is first-order reaction rate, $\mathcal{D}$ consists of hydrodynamic dispersion and molecular diffusion, ν denotes the Darcy velocity, q_s is the volumetric source/sink flow rate, $\mathbf{C}_s$ is the source/sink flux concentration and $r_{\mathbf{C}}$ refers to the rate of reactions. The superscript ℓ corresponds to the component number taking values between 1 and 4 in this study. Along with the basic groundwater flow and transport equations, and using the reaction operator-split strategy (Clement et al., 1998), the biological reaction kinetics are assembled as a set of ordinary differential equations as follows:

$$\frac{\partial \mathbf{C}_{\text{PCE}}}{\partial t} = -\frac{K_P \mathbf{C}_{\text{PCE}}}{R_P}, \tag{18a}$$

$$\frac{\partial \mathbf{C}_{\text{TCE}}}{\partial t} = -\frac{1}{R_T}\left(K_T \cdot \mathbf{C}_{\text{TCE}} - S_{T/P} \cdot K_P \cdot \mathbf{C}_{\text{PCE}}\right), \tag{18b}$$

$$\frac{\partial \mathbf{C}_{\text{DCE}}}{\partial t} = -\frac{1}{R_D}\left(K_D \cdot \mathbf{C}_{\text{DCE}} - S_{D/T} \cdot K_T \cdot \mathbf{C}_{\text{TCE}}\right), \tag{18c}$$

$$\frac{\partial \mathbf{C}_{\text{VC}}}{\partial t} = -\frac{1}{R_V}\left(K_V \cdot \mathbf{C}_{\text{VC}} - S_{V/D} \cdot K_D \cdot \mathbf{C}_{\text{DCE}}\right), \tag{18d}$$

where $\mathbf{C}_{\text{PCE}}, \mathbf{C}_{\text{TCE}}, \mathbf{C}_{\text{DCE}}$, and $\mathbf{C}_{\text{VC}}$ are the concentrations of the components, K_P, K_T, K_D, and K_V are first-order anaerobic degradation rate constants, $S_{T/P}, S_{D/T}$, and $S_{V/D}$ are stoichiometric yield values, and R_P, R_T, R_D, and R_V are retardation factors. Linear sorption conditions are assumed for all components.

The model domain as indicated by the blue region of Figure 2 is discretised horizontally into 20×30 grid cells of 50×50 m. In the vertical direction, we consider 120 layers each of 0.5 m thickness. The discretisation is based on the geological voxel model GeoTOP (Stafleu et al., 2011a). The top layer starts at 7.5 m above sea level, whereas the lowest layer is located at around 52.5 m below sea level. Based on different simulations conducted as part of this study, the migration of the contaminants was found to be limited to a certain depth. We thus assume that only layers $21-100$ are active. Figure 2 also shows the contaminant source (in yellow) consisting of four CH components with uniform concentration values. The plume data was obtained in January 2012 from a depth of 22.5 m below mean sea level (model layer 60), in which $C_{\text{PCE}} = 1083.0, C_{\text{TCE}} = 238.0, C_{\text{DCE}} = 633.0$, and $C_{\text{VC}} = 833.0$ μg/l. This contaminant plume is considered as the initial condition of the transport simulations in this study.

Furthermore, the PCE plume is used as a continuous contamination source and was included in the Source/Sink Mixing [SSM] package of the MT3DMS simulator. Up to this date, other time-series and well contaminant data are not accessible due to confidentiality imposed by local companies. Modelling parameters required for running the coupled FTR-Model, such as porosity and hydraulic conductivity are estimated in an offline procedure. To illustrate, the hydraulic conductivity is provided as a 3D field in the database GeoTOP. The GeoTOP for the province of South-Holland is constructed using 46000 borehole data (Busschers et al., 2010). Using the borehole data, the most probable lithostratigraphy and lithofacies have been estimated in each voxel of $100 \times 100 \times 0.5$ m. The GeoTOP further uses relations between the lithostratigraphical units and the lithofacies with parameters such as hydraulic conductivity, porosity and organic carbon content in order to provide these parameters on the voxel scale. Further details about the GeoTOP methodology in addition to application to another provinces can be found in Stafleu et al. (2011b). Table 1 outlines the mean value (averaged over all layers) for some of these parameters. We further show in Figure 3 the spatial map of the distribution coefficient of TCE averaged over the top 10 layers. The map shows larger sorption degrees in the northeast part of the domain. This gradually decreases towards the southern region.

3.3 Assimilation Experiments

3.3.1 Reference Run and Pseudo-Observations

In the scope of twin-experiments, we first conduct a reference model run using some "true" (reference) parameters and initial condition. Next, we impose different uncertainties on the model and the initial conditions, and we assimilate pseudo-observations extracted from the reference run to recover the "true" trajectory of the model. The goal is to estimate the concentration of chlorinated hydrocarbons (i.e., state variables) and their associated degradation rates (i.e., parameters):

$$\mathbf{x} = \left[\mathbf{C}^T_{\text{PCE}}, \quad \mathbf{C}^T_{\text{TCE}}, \quad \mathbf{C}^T_{\text{DCE}}, \quad \mathbf{C}^T_{\text{VC}} \right]^T, \tag{19a}$$

$$\boldsymbol{\Theta} = \left[K_P, \quad K_T, \quad K_D, \quad K_V \right]^T. \tag{19b}$$

Based on the model's configuration, the dimension of the state is $N_{\mathbf{x}} = 288{,}000$ and the parameters $N_{\boldsymbol{\Theta}} = 4$. The reference values of the anaerobic degradation rates are obtained through field and laboratory testing (Suarez and Rifai, 1999) and are given as $K_P = 0.068$, $K_T = 0.086$, $K_D = 0.004$, and $K_V = 0.153$ per day. We perform the reference run for a 50-years period using these rates along with the parameters of Table 1 and the initial conditions, $\mathbf{x}_0$, as defined above. We plot the resulting hydraulic head field at four different layers (i.e., 30, 50, 70, and 90) in Figure 4. The maps clearly show the southward and downward flow direction of the groundwater. The hydraulic head varies between 1.5 m in the center of the domain and drops to around -1 m in the southern part. The top layers, on average, have larger hydraulic heads than the deeper ones. Overall, the flow configuration indicates that the contaminant plume would follow the behaviour of the groundwater and predominantly moves vertically downwards and laterally in the southwards direction.

Based on this steady flow, we then simulated the reference transport of PCE, TCE, DCE, and VC. The time step of the transport-reaction simulator was about 11 days. To visualise the migration process, we plot in Figure 5 the contaminant evolution of PCE, TCE, DCE and VC in layers 40, 60, 80, and 100 after 50 years, respectively. As shown, the contaminant plume,

which is originally present in layer 60, has moved into deeper Pleistocene layers. After 50 years, the maximum concentration of DCE in layer 80 reaches 650 μg/l. Careful assessment of the transport process shows that the four plumes have reached the last active layer in the second aquifer; i.e., layer 100. This is mostly due to the continuous PCE contamination source located in layer 60. Contaminant concentrations in the top Holocene layers are much smaller. Laterally, the contaminant plume is seen to expand from its initial location to a distance of 1.3 km southwards.

From the reference run, we collect pseudo-observations of the concentration to use them later for assimilation. Observations are assumed available for all components from layers 30, 50, 70, and 90. From each of these four layers, 10 data points are collected and thus a total of $N_{\mathbf{y}} = 160$. Note that $N_{\mathbf{y}}$ is much smaller than the number of state variables, $N_{\mathbf{x}}$. This is usually the case in subsurface hydrology applications, given the significant and expensive cost incurred for preparing, drilling, and completing wells. The observation points are uniformly distributed throughout the domain as denoted by green triangles in Figure 2. We assume that these 160 measurements are available for assimilation on a yearly basis. We also place a control well in layer 70 around the center of the domain, particularly at the local coordinates $x = 450$ and $y = 600$ m, to monitor the concentration evolution in time. We further assume that these observations are noisy, in order to mimic realistic settings. We thus perturb them with a Gaussian noise of mean 0 and standard deviation equal to 15% of the total observation mean (averaged over the entire 50 years). During assimilation, the updated concentration values are monitored to make sure that they are non-negative. Cell values that fall out of this constraint are set to zero to obtain a physically meaningful solution (Li et al., 2012; Gharamti et al., 2013).

3.4 Initial Ensemble and Background Statistics

To initialise the filters, we perform an unconditional 50-years simulation run (referred to as free run) starting from the mean concentration of the reference model run. Thus, at time $t = 0$ the concentration of the four CH components is not only present in layer 60, but rather spread-out in all layers. In this free model run, we use around 30% larger degradation rates than the reference values. The concentration of the components was saved each 6 months. Next, we randomly select a set of N_e concentration snapshots from the free run outputs to form the state ensemble. This prior (initial) ensemble is quite far from the truth and further exhibits a relatively small spread. This is chosen in purpose to test the robustness of the assimilation schemes to challenging initial uncertainties. The initial parameters ensemble is sampled assuming a Gaussian distribution with mean equal to the reference rate values and variance 40%.

The background error statistics required for the hybrid EnKF-OI scheme are parameterised as follows. We form a set of 200 degradation realisations, as described above, and use these to perform 3-months forecasts starting from a series of initial concentrations distributed at 3-months intervals over a 50-years period, as outlined in Figure 6. To illustrate, starting from the mean concentration of the reference run, one realisation of the degradation rates $\mathbf{\Theta}_0$ is used to obtain a 3-months forecast of the concentration $\mathbf{x}_1$. Then, using $\mathbf{x}_1$ and $\mathbf{\Theta}_1$, the 3D-FTR model is integrated forward to obtain $\mathbf{x}_2$ concentration after 6 months. We continue this process until the end of the 50 years period. Then, we collect the predicted contaminant states for all components and augment them with the corresponding degradation rates in a joint matrix form. POD and SVD are then performed on the augmented concentration-degradation forecast perturbations to summarise the correlations by a small

number of orthogonal patterns (Hoteit et al., 2002; Skachko et al., 2009; Altaf et al., 2013). Consequently, the parameterisation of the background covariance matrix, $\mathbf{P}^b_{xx}$, is achieved using the leading 10 POD modes (i.e., $r_x = 10$) of this ensemble, which capture more than 98% of the total variance. Concerning the background cross-correlations, we use the first 10 singular vectors (thus, $r_\theta = 10$) to parametrise $\mathbf{P}^b_{\theta x}$ matrix. To visualise these correlations, we plot in Figure 7, as an illustration, the spatial correlation between the rate at which PCE is degrading and the concentration of the ending product of the chain, VC. Clearly, VC concentration in layer 60 exhibits the largest correlation values because of the continuous source term. Furthermore, since the groundwater flow is stronger in the downwards direction, and so is the contaminant migration, the cross-correlations in deeper layers are larger than those of the shallow layers. The background $\mathbf{P}^b_{\theta x}$ seem to vanish in the upper parts of the Holocene clay and peat layer. A consistent behaviour is observed for the remaining three degradation rates, in which the largest correlations are those associated with C_{PCE} and smallest with C_{TCE}. PCE has the highest correlation because of the continuous source zone of PCE. Any removal of PCE due to biodegradation in the source zone is directly replenished the next time step, and therefore K_P determines the total load of chlorinated hydrocarbons in the system. On the other hand, TCE has the lowest correlation because its value is high and that makes the parameter relatively insensitive to the amount of biodegradation taking place as compared to the other degradation rate constants.

State and parameters estimates of the EnKF, hybrid EnKF-OI and EnKF$_{\text{ESOS}}$ schemes are evaluated using two metrics; namely mean-squared-error (MSE) and average-ensemble-spread (AES):

$$\text{MSE} = N_{\mathbf{x}}^{-1} N_e^{-1} \sum_{j=1}^{N_e} \sum_{i=1}^{N_{\mathbf{x}}} \left(\mathbf{x}^e_{j,i} - \mathbf{x}^t_i\right)^2, \tag{20a}$$

$$\text{AES} = N_{\mathbf{x}}^{-1} N_e^{-1} \sum_{j=1}^{N_e} \sum_{i=1}^{N_{\mathbf{x}}} \left|\mathbf{x}^e_{j,i} - \hat{\mathbf{x}}^e_i\right|, \tag{20b}$$

where $\mathbf{x}^t_i$ is the true value of the variable at location i, $\mathbf{x}^e_{j,i}$ is the forecast ensemble value and $\hat{\mathbf{x}}^e_i$ is the corresponding ensemble mean. MSE measures the distance from the estimate to the truth and AES measures the spread or the uncertainty of the estimates (Hendricks Franssen and Kinzelbach, 2008).

4 Results and Discussion

In this section, we present and compare assimilation results with the Rotterdam port's 3D-FTR model using the standard EnKF, the hybrid EnKF-OI and the EnKF$_{\text{ESOS}}$ schemes. The observations are assimilated serially in all three schemes. Concentration and degradation rate estimates of the filters are compared in terms of accuracy and spread. Concentration data are assimilated every year for a total of 50 assimilation cycles. The ensemble size, N_e, is set to 48 in which batches of four members are run in parallel using Fortran's OpenMP library.

4.1 The Hybrid EnKF-OI vs the EnKF

4.1.1 Adjusting Concentration Statistics

To initiate the assimilation experiments, we first run the EnKF and the hybrid EnKF-OI, implemented using only state background statistics, i.e., using eq. (10a) and $\beta = 1$. We carry out 10 different experiments by changing the weighting factor α, for each individual run, between 0 to 1 with a step of 0.1. To visualise the resulting estimates, we plot the average MSE of the contaminant concentrations, averaged over the 4 components and in time, in Figure 8. As shown in the left panel of Figure 8, the most accurate concentration estimates are obtained using $\alpha = 0.7$. This indicates that out of the total forecast error variance, the best reconstruction of the reference contaminant solution is obtained when 30% of this variance is traced from the background statistics. Increasing or decreasing this background contribution (i.e., 30%) results in less accurate contaminant estimates. We also note that the least accurate estimates are those obtained when 90% of the ensemble statistics are built based on the static background error covariance. On average, we found that when α takes values between 0.4 and 0.9, the hybrid EnKF-OI is 16% more accurate than the EnKF.

We also study the effect of changing α on the resulting parameters' estimates. In the right panel of Figure 8, we plot the average MSE of each individual degradation rate obtained using the EnKF and 9 different hybrid EnKF-OIs using $\alpha = 0.1, 0.2, \ldots, 0.9$. We notice that the most influenced biodegradation rates are those associated with TCE and VC. In fact, K_T and K_V are the least identifiable parameters and therefore a small difference in the estimation algorithm (i.e., hybrid EnKF-OI and the EnKF) may lead to different estimates. In contrast, K_P and K_D are less sensitive to the weighting factor α. In accordance with the estimates of the contaminant concentrations, the best match for K_T is obtained using $\alpha = 0.7$. This is not the case for the other parameters; in which $\alpha = 0.3, 1$ and 0.8 resulted in the best fit to the reference degradation rates of VC, PCE and DCE, respectively. On average, K_T and K_V estimates are 13% and 35% more accurate than those of the EnKF, respectively. The key point is that complementing the state statistics, using a weighted error covariance as in (10a), does not only contribute to a better retrieval of the concentration but also helps adjusting the cross-correlations with the uncertain parameters. This is essentially the case when biodegradation is taking place at a higher rate, as in K_T and K_V, and thus the more information fed through observations, the better the state-parameters cross-correlations would be.

To better understand the performance of the hybrid EnKF-OI scheme, we further study how the uncertainties of these degradation rates are maintained in time. For this, we plot in Figure 9 the overall ensemble spread of the four degradation rates obtained using the EnKF and the hybrid EnKF-OI (all tested α's). As shown, the EnKF's spread around these 4 parameters is quickly reduced after the first 2 or 3 years. This rapid reduction of the ensemble spread, which is due to the relatively small ensemble size and large initial uncertainties, limits the ability of the filter to impose larger corrections in the future, eventually degrading the accuracy of the estimated parameters. In contrast, the hybrid EnKF-OI maintains larger uncertainties in time for different weighting factor values. As such, as α decreases from 1 to 0.1 the performance approaches that of the EnKF. For instance, after the second year, the ensemble spread of the EnKF reaches 0.03 while it remains larger in the hybrid EnKF-OI and equal to 0.09, 0.07 and 0.045, respectively. This allows the hybrid scheme to better exploit the concentration information from observations. This can be confirmed by noticing that the spread of the hybrid filter continues to decrease after 25 years

of assimilation, unlike the EnKF that does not show signs of more corrections. All in all, selecting $\alpha = 0.7$ seems to maintain enough spread for both components' concentration and degradation rates and leads to the most accurate estimates, on average. For the sake of comparison, we refer hereafter to this scenario as EnKF-OI$_{\alpha=0.7}$.

4.1.2 Adjusting Concentration and Parameters Statistics

In the following set of experiments, we fix α to 0.7 and focus on changing the weighting factor between the EnKF and background state and parameters cross-correlations; i.e., β. As in the previous section, we conduct 9 experiments in which we change β between 0.1 and 0.9. Note that the larger β is, the closer the performance is to EnKF-OI$_{\alpha=0.7}$. To analyse the results, we plot in Figure 10 the average MSE and AES of the chlorinated hydrocarbon concentrations. Compared to the previous runs that hybridise the state only, including background cross-correlations information slightly increases the spread around the ensemble mean of concentration, as observed for $0.1 < \beta < 0.3$. In terms of accuracy, varying β between 0.1 and 0.6 yields more accurate concentration estimates for all components. To illustrate, when $\beta = 0.1$ the average improvements over the EnKF and the EnKF-OI$_{\alpha=0.7}$ are around 50% and 32%, respectively. This vigorous performance suggests that using only 10% of the "flow-dependent" parameters' ensemble to characterise the pdf of the system is enough to outperform the EnKF. In essence, the background state and parameters cross-correlations seem to carry sufficient description of how the degradation rates and the concentration of each of the components are related. Consequently, only a small portion (i.e., 10%) of the online parameters' ensemble is required to obtain an accurate biodegradation picture, while the rest of the information could come from the prescribed static background statistics. This could be due to the time-independent nature of the propagation step describing the evolution of the degradation rates, thereby manifesting a minimal dependence on the flow-dependent ensemble. This observation comes in accordance with the steady-state Kalman filter (El Serafy and Mynett, 2008) that assumes time-invariant error covariance matrix as long as accurate spatial correlations are used within the so-called Kalman gain. Our experimental results suggest that the best parameter's hybrid covariance matrix is very close to a steady-state one. However, this is only for the parameters and this was not the case for the state as discussed in section (4.1.1). Following the notation introduced earlier, we refer to this scenario, hereafter, as EnKF-OI$^{\beta=0.1}_{\alpha=0.7}$

To have a better insight at the suggested performance, we plot in Figure 11 the evolution of the concentration ensemble members for all components in time. For a fair comparison, we also plot the associated reference solution, the EnKF's and the EnKF-OI$_{\alpha=0.7}$'s ensemble members. As explained in section (3.4), the initial ensemble spread is clearly far from the truth. When data is assimilated into the system, all schemes tend to move closer to the truth. By the end of the 50-years period, both EnKF and EnKF-OI$_{\alpha=0.7}$ underestimate the concentration of DCE and VC and end up with quite small ensemble spread. The EnKF-OI$^{\beta=0.1}_{\alpha=0.7}$, on the other hand, leads to the best performance, well matching the reference solution for all components. Moreover, this hybrid scheme is shown to better preserve the ensemble spread around the true final concentrations. In terms of the estimated degradation rates, we plot in Figure 12 the temporal change of MSE for each individual degradation rate as they result from the EnKF-OI$^{\beta=1.0}_{\alpha=0.1\ldots0.9}$ (top panels) and EnKF-OI$^{\beta=0.1\ldots0.9}_{\alpha=0.7}$ (bottom panels). For all rates, the EnKF-OI$^{\beta=0.1\ldots0.9}_{\alpha=0.7}$ performs much better during the first10 years, especially for K_P and K_T. Averaging in time and over all cases, EnKF-OI$^{\beta=0.1\ldots0.9}_{\alpha=0.7}$ is 33%, 17%, 33% and 15% more accurate for retrieving K_P, K_T, K_D and K_V, respectively. From these

results, one can see that the accuracy of the degradation rates tends to improve in time, except for K_D that is shown to degrade for small α and large β values. To interpret this behaviour, one should recall that K_D (and also K_V) can only be estimated correctly as long as the concentration of the source component (C_{DCE} in this case) is accurately recovered. Before reaching this, the estimates of K_D are compensated for errors in K_P and K_T.

Next, and instead of manually changing the weighting factors α and β, we follow eqs. (13) and (14) and conduct a 1D optimisation problem prior to assimilating the observations serially. The idea is to get the maximum reduction in the prior uncertainties for both the concentration and the degradation rates as a way to "optimally" exploit the information in the assimilated observations. As such, different weights can be assigned to the background and the ensemble statistics. Based on this, we plot in Figure 13 the resulting optimal α values at every assimilation cycle and for each observation. Recall that there is a total of 160 observations, such that each contaminant component is observed at 40 different locations. To better interpret the plot, we arrange these observation indices as follows: from top to bottom of the left y-axis; PCE: $1 \rightarrow 40$, TCE: $41 \rightarrow 80$, DCE: $81 \rightarrow 120$ and VC: $121 \rightarrow 160$. As can be seen from the plot, the adaptive hybrid EnKF-OI algorithm selects either 0, and thus eq. (10a) is purely based on $\mathbf{P}^b_{xx}$, or 1 so that only the ensemble covariance, $\widehat{\mathbf{P}}^{\text{EnKF}}_{xx}$, is included. When assimilating PCE, TCE and VC concentrations, the adaptive scheme tends to use the background covariance (i.e., $\alpha = 0$) for almost the first 25 years. Beyond this, the filter statistics are only based on the ensemble flow-dependent information (i.e., $\alpha = 1$). This is not surprising given the large initial uncertainties imposed on the contaminant concentrations. Once the statistics are adjusted towards the truth, the filter relies more on the correlations of the flow-dependent ensemble statistics. This performance changes when assimilating DCE observations in the sense that the filter builds its forecast error covariance mostly using static background statistics and less using the flow-dependent ensemble. This is in agreement with the results and conclusion drawn from Figure 12 in which the background information of DCE are more useful than the ensemble statistics. Averaging over the entire optimal values of α, we obtain a global $\alpha^* = 0.64$, which is quite close to the 0.7 value that resulted in the best performance in section (4.1.1). In terms of the adaptive β values, we found that maximising the difference between the prior and the posterior parameters' covariance, $\widetilde{\mathbf{P}}_{\theta\theta}$, may not always be helpful. This is because doing such maximisation can quickly diminish the ensemble spread, eventually paralysing the filter's analysis. In fact, minimising the difference[1] yielded more accurate degradation rates. To analyse this, we plot on the same figure the time-evolution of MSE of concentration when (1) maximising the information gain for both state and parameters, (2) minimising the information gain for both and (3) maximising the state's and minimising the parameters' information gain, respectively. As seen from the three curves, the best performance is obtained when the information gain for concentration is maximised and the associated parameter's one is minimised. Compared to maximising the information gain of both state and parameters, this mixed scheme now yields 37% more accurate contaminant concentrations.

[1]Minimising the difference between the prior and the posterior covariances does not mean that the filter does not apply any correction. Since the Kalman analysis equation always minimises the variance, the adaptive algorithm now acts in a way such that only the lowest minimisation possible is retrieved. Unlike standard Kalman filtering, this procedure moves at a slower pace towards the truth.

4.2 The EnKF$_{\text{ESOS}}$ vs the EnKF

In the previous section, all approaches and experimental results were intended to mitigate for the rank deficiency and the under-sampling of the ensemble's sample forecast error covariance. In the following experiment, we attempt to deal with the undersampling of the observation errors by implementing the EnKF$_{\text{ESOS}}$ algorithm presented in section (2.3). We first note that the distribution of the new observation perturbations show reasonable deviations from the prescribed Gaussian errors in the original EnKF algorithm, as has been noticed by Hoteit et al. (2015). To assess the performance of the EnKF$_{\text{ESOS}}$ against the EnKF, we study at a closer glance the contaminant maps after 50 years as estimated by the ensemble means from both schemes. Thus, we plot in Figure 14 the normalised errors for the components TCE and DCE at layers 70 and 80, respectively. These error maps are obtained by subtracting the ensemble mean concentration from the reference and then normalising the result by the average of the reference solution. One common feature in these maps is the clear underestimation of TCE and DCE in the north part of the domain. This is because the initial reference concentration is quite different from the one assigned to the initial ensemble using the free run setup as outlined in section (3.4). In time, both filtering schemes try to push the contaminant plume, which has already moved towards the southern region, upwards to match the truth. Moreover, as demonstrated in layer 70 and unlike the EnKF$_{\text{ESOS}}$, the EnKF overestimates the TCE concentration in the center of the domain, which further continues to move southwards. In layer 80 (i.e., 5 m deeper), the EnKF tends to underestimate the concentration of DCE especially in the southern part of the domain. On the other hand, a slight overestimation of this DCE concentration towards the center is suggested by the EnKF$_{\text{ESOS}}$. In general, and assessing similar patterns at other layers, the EnKF$_{\text{ESOS}}$ exhibits higher accuracy in retrieving the contaminant concentration than the EnKF. This provides further evidence that ignoring the observation sampling errors within the EnKF can indeed deteriorate the quality of the state estimates.

To study the impact of the EnKF$_{\text{ESOS}}$ on the estimates of the parameters, we examine the evolution of the approximate distribution of TCE degradation rate in time. We compare the resulting pdfs with those obtained using the EnKF after 5, 15, 30 and 45 years. On top of the pdfs, we also monitor the temporal evolution of K_T AES in Figure 15. Starting from rather flat and uncertain pdfs of K_T, both EnKF and EnKF$_{\text{ESOS}}$ correct the members of TCE degradation rate towards the truth, which is 0.086 per day. Notice that within the first 15 years, the pdfs seem to move in the wrong direction, away from the truth. This is due to the large concentrations at time 0, and thus the filter increases the degradation rates to fit the reference contaminant concentration. Beyond that and once the concentration is adjusted, the parameters from both filtering schemes begin moving closer toward the true degradation rate. However, the EnKF is seen to move faster towards the truth and further diminishes the uncertainty around K_T quite rapidly. Consequently, the resulting pdf of K_T after 45 years looks like a Kronecker delta function. This is, roughly speaking, not a very healthy assimilation system as the parameter updates become insignificant over the rest of the assimilation window. In contrast, the degradation rate obtained using EnKF$_{\text{ESOS}}$ moves at a slower pace towards the true rate maintaining large enough spread to fit the incoming observations. Compared to the EnKF, the AES suggested by the EnKF$_{\text{ESOS}}$, as shown on the left y-axis, is almost 40% to 50% higher. As a matter of fact, this performance is more trustworthy than that of the EnKF, indicating the essential need to account for observation sampling errors at the time of the analysis. Hoteit

et al. (2015) found that the ensemble spread of the EnKF$_{\text{ESOS}}$ is larger than that of the EnKF for state estimation. In here, we experienced a similar, yet more pronounced behaviour for the estimates of the parameters.

As a way to provide an overall assessment, we compare the best estimates obtained using all schemes considered earlier; i.e., the EnKF, the EnKF-OI$^{\beta=0.1}_{\alpha=0.7}$ and the EnKF$_{\text{ESOS}}$. We plot the time series of MSE for contaminant concentration and degradation rates, summed over all components, in Figure 16. Clearly the EnKF is the least accurate. Accounting for observation sampling errors yield around 21% and 23% more accurate state and parameters estimates, respectively. Tackling the rank deficiency of the EnKF results in 48% and 70% more accurate state and parameters estimates, respectively. Accordingly, addressing the issues of observation sampling errors and rank deficient forecast ensemble matrices seem to be crucial and can highly improve the accuracy of the estimates. From our experimental results and for this particular setting, resolving the rank deficiency issue appear to have the largest impact on the final estimates of the filter.

4.3 Incorporating Uncertainties in the Hydraulic Parameters

In sections 4.1 and 4.2, the static background covariance matrices were derived on the basis of perfectly known hydraulic parameters. In this section, we test the impact of incorporating uncertainties in the groundwater flow model on the performance of both the hybrid EnKF-OI and the EnKF$_{\text{ESOS}}$ schemes. Generally, such procedure is expected to alter the precise description that the hybrid scheme utilizes to relate the biodegradation rates and the components' concentrations. To this end, the GeoTOP software package, described in section 3.2.2, is used to obtain 48 different realizations for hydraulic conductivity and porosity. The realizations are built assuming a Gaussian distributed hydraulic parameters with mean equal to the 3D fields used in the reference model run (section 3.2.2) and standard deviation paramertrized in two different scenarios. We use $\sigma_m = 10\%$ of the mean in the first scenario (moderate uncertainty) and $\sigma_h = 30\%$ of the mean value in the second scenario (high uncertainty). The reason for this choice is to provide a realistic assessment of the filters under varying modelling uncertainty.

Before testing the performance of the hybrid EnKF-OI and the EnKF$_{\text{ESOS}}$, we first construct the background covariances; i.e.,$\mathbf{P}^b_{xx}$ and $\mathbf{P}^b_{\theta x}$, using a similar procedure to the one presented in section 3.4 but based on perturbed conductivity and porosity realizations (here, σ_m is used). To interpret the influence of this modelling uncertainty, we plot in Figure 17 the averaged cross-correlations of K_P rate with all four components. We observe that the dominant correlation patterns are similar to those obtained earlier using perfect flow conditions, especially in the shallow aquifer layers. In deeper layers, there are noticeable differences in the correlations of K_P and C_{TCE}. In addition, the magnitude of the new background correlations is considerably smaller. For instance, the estimated correlation value between K_P and C_{DCE} after imposing uncertainty on the hydraulic parameters has shrunk by 60%. The fact that the major spatial patterns of the background correlations were preserved and their magnitude was influenced the most is related to the nature of the perturbed hydraulic parameters. Generally, porosity and conductivity affect the speed and the movement of groundwater in the aquifer and thus the degradation process would be expected to either slow down or accelerate. Compared to the previous parameterization, in this scenario 15 POD modes were required to capture around 97% of the total variance.

With moderate uncertainty, both the hybrid EnKF-OI and the EnKF$_{\text{ESOS}}$ outperform the EnKF and recover decent estimates of the degradation parameters. However, compared to the ones obtained under perfect flow conditions, these were slightly less

accurate. Switching to the high uncertainty scenario (i.e., using σ_h), all three assimilation schemes failed to retrieve reliable estimates of the concentration and biodegradation. To illustrate, we show in Figure 18 (top panel) the temporal evolution of MSE obtained using EnKF-OI$^{\beta=1.0}_{\alpha=0.3,\ldots,0.8}$ with perfect and perturbed hydraulic parameters. The performance as shown starts degrading after 15 years, eventually leading to inaccurate concentration estimates and unreliable parameters. This suggests that quantifying the large uncertainty of the flow model alongside the reaction parameters, using any of the three schemes, might be necessary in such a challenging setup. On the other hand, under moderate uncertainty we compare the performance of the assimilation schemes to the ones obtained assuming perfect flow. For this, we plot the average gain (MSE difference), suggested by the adaptive hybrid scheme over EnKF$_{\text{ESOS}}$ estimates of the degradation rates. The adaptive hybrid EnKF-OI remains more accurate even after perturbing the hydraulic parameters. However, the proposed accuracy is roughly halved for all parameters. This is indeed related to the less reliable static background covariances, which are now subjected to uncertainties in the groundwater flow dynamics.

5 Conclusions

In this study, we examined and investigated the hybrid ensemble Kalman filter (EnKF-OI) and the second-order observation perturbations sampling (EnKF$_{\text{ESOS}}$) schemes to estimate contaminant concentration and biodegradation rates of chlorinated hydrocarbons at the port of Rotterdam. We simulated the migration problem of a single plume consisting of Tetrachloroethene (PCE), Trichloroethene (TCE), cis-1,2-Dichloroethene (DCE) and Vinyl Chloride (VC). Concentration data was used for yearly assimilation over a period of 50 years. The hybrid scheme complements the flow-dependent sample ensemble covariance of the EnKF with a prescribed static background covariance from an OI system to mitigate the undersampling of the ensembles and neglected model errors. The exact second-order sampling of the observation perturbations modifies the observation perturbations and assimilates the data one after the other, thus resolving the undersampling of the observation noise in the EnKF analysis. Challenging assimilation scenarios using a relatively small ensemble ($N_e = 48$) were presented, in which observations were processed serially. The key findings of this study and future research directions are summarised below:

1. Both the hybrid EnKF-OI and the EnKF$_{\text{ESOS}}$ successfully provide accurate concentration and degradation rate estimates. On average, a tuned hybrid EnKF-OI (using $\alpha = 0.7$ and $\beta = 0.1$) suggests 48% and 70% more accurate state and parameters estimates than those obtained using the EnKF. On the other hand, the EnKF$_{\text{ESOS}}$'s state and parameters estimates are 21% and 23% more accurate, respectively. In addition, the two schemes are easy to implement and computationally efficient requiring only a minimal change to an existing EnKF code.

2. Both filtering schemes demonstrated a better handling of the ensemble spread, for both state and parameters, avoiding collapse or false (unrealistic) confidence in the estimates, which enables better fit to the observations.

3. The hybrid scheme requires some effort to tune two weighting factors that adjust the background statistics for both state and parameters. The serial adaptive version of this scheme, which relies on maximising the information gain between the forecast and analysis for each individual observation point, seems promising. From the experiments, we found that

maximising the information gain could however possibly deplete the uncertainty within the ensemble quite rapidly. Yet, this observation may vary between systems depending on the degree and the rate of uncertainty growth. One possible solution that we tested is to minimise the information gain, and thus decrease the update impact when fitting the observations. Further, one could also build the objective function in such a way that only a portion of the information gain is maximised. For instance, an example would be to enforce the ratio between the trace of the analysis and the forecast covariance matrices to be greater than 30% meaning that at least 60% of the ensemble uncertainty is preserved.

4. Failing to account for observation undersampling errors in the standard EnKF can impact not only the quality of the state but more importantly the estimates of the parameters. In our experiments, the degradation rates obtained after assimilating the observations using the $\text{EnKF}_{\text{ESOS}}$ scheme were more accurate, more reliable and more realistic.

5. Imposing large uncertainties on the hydraulic parameters of the flow model degrades the performance of all filtering schemes. Given that the performance of the hybrid EnKF-OI depends on the quality of the static background statistics, satisfactory results were obtained only when the uncertainty imposed on the background information is relatively moderate. The $\text{EnKF}_{\text{ESOS}}$ further outperformed the standard EnKF with moderate flow uncertainty conditions.

6. Careful tuning of the hybrid EnKF-OI yields the best estimates of the concentration and the degradation rates as compared to the EnKF and the $\text{EnKF}_{\text{ESOS}}$. This manifests the importance of complementing the EnKF parameters cross-correlations with static ones.

Building a unified EnKF scheme, which tackles both the undersampling of the forecast covariance and the observation sampling errors simultaneously is an interesting line of research in the future that we will consider in our future work.

Acknowledgements. The authors would like to thank three anonymous reviewers for their constructive comments and suggestions, which improved the quality of the manuscript. The authors would also like to thank the municipality of Rotterdam for their support and cooperation throughout this project. Research reported in this publication was supported by King Abdullah University of Science and Technology (KAUST).

References

Altaf, M., Gharamti, M., Heemink, A., and Hoteit, I.: A reduced adjoint approach to variational data assimilation, Computer Methods in Applied Mechanics and Engineering, 254, 1–13, doi:10.1016/j.cma.2012.10.003, 2013.

Annan, J., Hargreaves, J., Edwards, N., and Marsh, R.: Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter, Ocean modelling, 8, 135–154, 2005.

Appelo, C. and Postma, D.: Geochemistry Groundwater and Pollution. AA Balkema, Rotterdam, Brookfield, Tech. rep., ISBN 90-5410-105-9, 536p, doi:10.1017/s0016756800011523, 1994.

Bailey, R. T., Baù, D. A., and Gates, T. K.: Estimating spatially-variable rate constants of denitrification in irrigated agricultural groundwater systems using an Ensemble Smoother, Journal of Hydrology, 468, 188–202, 2012.

Bailey, R. T., Gates, T. K., and Halvorson, A. D.: Simulating variably-saturated reactive transport of selenium and nitrogen in agricultural groundwater systems, Journal of contaminant hydrology, 149, 27–45, 2013.

Bear, J. and Cheng, A.-D.: Modeling groundwater flow and contaminant transport, vol. 23, Springer, doi:10.1007/978-1-4020-6682-5_7, 2010.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, Monthly weather review, 129, 420–436, 2001.

Bowler, N. E., Flowerdew, J., and Pring, S. R.: Tests of different flavours of EnKF on a simple model, Quarterly Journal of the Royal Meteorological Society, 139, 1505–1519, 2013.

Burgers, G., Van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, Monthly weather review, 126, 1998.

Busschers, F. S., Dubelaar, C. W., Stafleu, J., and Maljers, D.: Lithological and sand grain-size variability in the three-dimensional GeoTOP model of Zuid-Holland, Tech. Rep. uGE 100 ENG, TNO Built Environment and Geosciences, Princetonlaan 6, 3508 TA Utrecht, The Netherlands, http://www2.dinoloket.nl/nl/about/modellen/uGE%20100%20A3%20ENG%20Freek%20Busschers%20Print.pdf, 2010.

Chave, P.: The EU Water Framework Directive, 2001.

Chen, Y. and Zhang, D.: Data assimilation for transient flow in geologic formations via ensemble Kalman filter, Advances in Water Resources, 29, 1107–1122, doi:10.1016/j.advwatres.2005.09.007, 2006.

Clement, T.: RT3D, A modular computer code for simulating Reactive multispecies Transport in 3-Dimensional groundwater aquifers, Pacific Northwest National Laboratory, Richland, Washington, 1997.

Clement, T., Sun, Y., Hooker, B., and Petersen, J.: Modeling multispecies reactive transport in ground water, Groundwater Monitoring & Remediation, 18, 79–92, doi:10.1111/j.1745-6592.1998.tb00618.x, 1998.

Clement, T. P., Johnson, C. D., Sun, Y., Klecka, G. M., and Bartlett, C.: Natural attenuation of chlorinated ethene compounds: model development and field-scale application at the Dover site, Journal of Contaminant Hydrology, 42, 113–140, doi:10.1016/s0169-7722(99)00098-4, 2000.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Quarterly Journal of the Royal Meteorological Society, 120, 1367–1387, doi:10.1002/qj.49712051912, 1994.

Crestani, E., Camporese, M., Baú, D., and Salandin, P.: Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation, Hydrol. Earth Syst. Sci, 17, 1517–1531, 2013.

Cunningham, S. D. and Berti, W. R.: Remediation of contaminated soils with green plants: an overview, In Vitro Cellular & Developmental Biology-Plant, 29, 207–212, doi:10.1007/bf02632036, 1993.

Dawson, C., Sun, S., and Wheeler, M. F.: Compatible algorithms for coupled flow and transport, Computer Methods in Applied Mechanics and Engineering, 193, 2565–2580, doi:10.1016/j.cma.2003.12.059, 2004.

Devegowda, D., Arroyo, E., Datta-Gupta, A., Douma, S. G., et al.: Efficient and robust reservoir model updating using ensemble Kalman filter with sensitivity-based covariance localization, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers, doi:10.2523/106144-ms, 2007.

Dimet, F.-X. L. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus A, 38, 97–110, doi:10.3402/tellusa.v38i2.11706, 1986.

Drécourt, J.-P., Madsen, H., and Rosbjerg, D.: Calibration framework for a Kalman filter applied to a groundwater model, Advances in water resources, 29, 719–734, doi:10.1016/j.advwatres.2005.07.007, 2006.

El Serafy, G. Y. and Mynett, A. E.: Improving the operational forecasting system of the stratified flow in Osaka Bay using an ensemble Kalman filter–based steady state Kalman filter, Water resources research, 44, 2008.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean dynamics, 53, 343–367, doi:10.1007/s10236-003-0036-9, 2003.

Forsythe, G. E., Malcolm, M. A., and Moler, C. B.: Computer methods for mathematical computations, 1977.

Franssen, H. and Kinzelbach, W.: Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems, Journal of hydrology, 365, 261–274, doi:10.1016/j.jhydrol.2008.11.033, 2009.

Freeze, R. A. and Cherry, J.: Groundwater, 604 pp, 1979.

Gharamti, M. and Hoteit, I.: Complex step-based low-rank extended Kalman filtering for state-parameter estimation in subsurface transport models, Journal of Hydrology, 509, 588–600, doi:10.1016/j.jhydrol.2013.12.004, 2014.

Gharamti, M., Hoteit, I., and Sun, S.: Low-rank Kalman filtering for efficient state estimation of subsurface advective contaminant transport models, Journal of Environmental Engineering, 138, 446–457, doi:10.1061/(asce)ee.1943-7870.0000484, 2011.

Gharamti, M., Kadoura, A., Valstar, J., Sun, S., and Hoteit, I.: Constraining a compositional flow model with flow-chemical data using an ensemble-based Kalman filter, Water Resources Research, 50, 2444–2467, doi:10.1002/2013wr014830, 2014a.

Gharamti, M., Valstar, J., and Hoteit, I.: An adaptive hybrid EnKF-OI scheme for efficient state-parameter estimation of reactive contaminant transport models, Advances in Water Resources, 71, 1–15, doi:10.1016/j.advwatres.2014.05.001, 2014b.

Gharamti, M. E., Hoteit, I., and Valstar, J.: Dual states estimation of a subsurface flow-transport coupled model using ensemble Kalman filtering, Advances in Water Resources, 60, 75–88, doi:10.1016/j.advwatres.2013.07.011, 2013.

Hamill, T. M. and Snyder, C.: A hybrid ensemble Kalman filter-3D variational analysis scheme, Monthly Weather Review, 128, 2905–2919, 2000.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, Monthly Weather Review, 129, 2776–2790, 2001.

Harbaugh, A. W.: 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, USA, 2005.

Hendricks Franssen, H. and Kinzelbach, W.: Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, Water Resources Research, 44, doi:10.1029/2007wr006505, 2008.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C. K., Heiskanen, A.-S., Johnson, R. K., Moe, J., Pont, D., et al.: The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future, Science of the Total Environment, 408, 4007–4019, doi:10.1016/j.scitotenv.2010.05.031, 2010.

Hoteit, I., Pham, D.-T., and Blum, J.: A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific, Journal of Marine systems, 36, 101–127, doi:10.1016/s0924-7963(02)00129-x, 2002.

Hoteit, I., Cornuelle, B., Köhl, A., and Stammer, D.: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation, Quarterly Journal of the Royal Meteorological Society, 131, 3659–3682, doi:10.1256/qj.05.97, 2005.

Hoteit, I., Pham, D.-T., Triantafyllou, G., and Korres, G.: A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography, Monthly Weather Review, 136, 317–334, 2008.

Hoteit, I., Xiaodong, L., and Pham, D.-T.: Particle Kalman Filtering: A Nonlinear Bayesian Framework for Ensemble Kalman Filters, Monthly weather review, 140, 528–542, doi:10.1175/2011mwr3640.1, 2012.

Hoteit, I., Pham, D.-T., Gharamti, M., and Luo, X.: Mitigating observation perturbation sampling errors in the stochastic EnKF, Monthly Weather Review, 143, 2918–2936, 2015.

Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, Monthly Weather Review, 129, 123–137, 2001.

Kalman, R. E.: A new approach to linear filtering and prediction problems, Journal of basic Engineering, 82, 35–45, doi:10.1115/1.3662552, 1960.

Kullback, S. and Leibler, R. A.: On information and sufficiency, The Annals of Mathematical Statistics, pp. 79–86, doi:10.1214/aoms/1177729694, 1951.

Lawson, W. G. and Hansen, J. A.: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth, Monthly weather review, 132, 1966–1981, 2004.

Lee, L. J., Chung, C.-W., Ma, Y.-C., Wang, G.-S., Chen, P.-C., Hwang, Y.-H., and Wang, J.: Increased mortality odds ratio of male liver cancer in a community contaminated by chlorinated hydrocarbons in groundwater, Occupational and environmental medicine, 60, 364–369, doi:10.1136/oem.60.5.364, 2003.

Lee, L. J.-H., Chan, C.-C., Chung, C.-W., Ma, Y.-C., Wang, G.-S., and Wang, J.-D.: Health risk assessment on residents exposed to chlorinated hydrocarbons contaminated in groundwater of a hazardous waste site, Journal of Toxicology and Environmental Health Part A, 65, 219–235, doi:10.1080/15287390252800828, 2002.

Li, L., Zhou, H., Gómez-Hernández, J. J., and Hendricks Franssen, H.-J.: Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble Kalman filter, Journal of Hydrology, 428, 152–169, doi:10.1016/j.jhydrol.2012.01.037, 2012.

Liu, B., Gharamti, M., and Hoteit, I.: Assessing clustering strategies for Gaussian mixture filtering a subsurface contaminant model, Journal of Hydrology, 535, 1–21, 2016.

Mandel, J., Cobb, L., and Beezley, J. D.: On the convergence of the ensemble Kalman filter, Applications of Mathematics, 56, 533–541, 2011.

Marsman, A., Valstar, J., and TER MEER, J.: Risk analysis on groundwater contamination at the megasite Port of Rotterdam, IAHS-AISH publication, pp. 309–314, 2006.

McLaughlin, D.: An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering, Advances in Water Resources, 25, 1275–1286, 2002.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state–parameter estimation of hydrological models using ensemble Kalman filter, Advances in Water Resources, 28, 135–147, doi:10.1016/j.advwatres.2004.09.002, 2005.

Mostert, E.: The European water framework directive and water management research, Physics and Chemistry of the Earth, Parts A/B/C, 28, 523–527, doi:10.1016/s1474-7065(03)00089-5, 2003.

Nerger, L., Hiller, W., and Schröter, J.: A comparison of error subspace Kalman filters, Tellus A, 57, 715–735, 2005.

Nichols, N.: Mathematical concepts of data assimilation, in: Data Assimilation, pp. 13–39, Springer, doi:10.1007/978-3-540-74703-1_2, 2010.

Ojajärvi, A., Partanen, T., Ahlbom, A., Boffetta, P., Hakulinen, T., Jourenkova, N., Kauppinen, T., Kogevinas, M., Vainio, H., Weiderpass, E., et al.: Risk of pancreatic cancer in workers exposed to chlorinated hydrocarbon solvents and related compounds: a meta-analysis, American journal of epidemiology, 153, 841–850, doi:10.1093/aje/153.9.841, 2001.

Panzeri, M., Riva, M., Guadagnini, A., and Neuman, S.: Data assimilation and parameter estimation via ensemble Kalman filter coupled with stochastic moment equations of transient groundwater flow, Water Resources Research, 49, 1334–1344, 2013.

Pham, D. T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems, Monthly weather review, 129, 1194–1207, 2001.

Pham, D. T., Verron, J., and Roubaud, M. C.: A singular evolutive extended Kalman filter for data assimilation in oceanography, Journal of Marine systems, 16, 323–340, 1998.

Pollock, D. W.: User's Guide for MODPATH/MODPATH-PLOT, Version 3: A Particle Tracking Post-processing Package for MODFLOW, the US: Geological Survey Finite-difference Ground-water Flow Model, US Department of Interior, 1994.

Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, Monthly Weather Review, 130, 103–114, 2002.

Sakov, P. and Bertino, L.: Relation between two common localisation methods for the EnKF, Computational Geosciences, 15, 225–237, doi:10.1007/s10596-010-9202-6, 2011.

Sakov, P. and Oke, P. R.: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus A, 60, 361–371, 2008.

Skachko, S., Brankart, J.-M., Castruccio, F., Brasseur, P., and Verron, J.: Improved turbulent air-sea flux bulk parameters for controlling the response of the ocean mixed layer: a sequential data assimilation approach, Journal of Atmospheric and Oceanic Technology, 26, 538–555, doi:10.1175/2008jtecho603.1, 2009.

Song, H., Hoteit, I., Cornuelle, B. D., and Subramanian, A. C.: An adaptive approach to mitigate background covariance limitations in the ensemble Kalman filter, Monthly Weather Review, 138, 2825–2845, doi:10.1175/2010mwr2871.1, 2010.

Song, H., Hoteit, I., Cornuelle, B. D., Luo, X., and Subramanian, A. C.: An Adjoint-Based Adaptive Ensemble Kalman Filter, Monthly Weather Review, 141, 3343–3359, doi:10.1175/mwr-d-12-00244.1, 2013.

Stafleu, J., Maljers, D., Gunnink, J., Menkovic, A., and Busschers, F.: 3D modelling of the shallow subsurface of Zeeland, the Netherlands, Netherlands Journal of Geosciences, 90, 293–310, 2011a.

Stafleu, J., Maljers, D., Gunnink, J., Menkovic, A., and Busschers, F.: 3D modelling of the shallow subsurface of Zeeland, the Netherlands, Netherlands Journal of Geosciences, 90, 293–310, 2011b.

Starr, R. C. and Cherry, J. A.: In Situ Remediation of Contaminated Ground Water: The Funnel-and-Gate System, Groundwater, 32, 465–476, doi:10.1111/j.1745-6584.1994.tb00664.x, 1994.

Suarez, M. P. and Rifai, H. S.: Biodegradation rates for fuel hydrocarbons and chlorinated solvents in groundwater, Bioremediation Journal, 3, 337–362, doi:10.1080/10889869991219433, 1999.

Sun, S. and Wheeler, M. F.: A posteriori error estimation and dynamic adaptivity for symmetric discontinuous Galerkin approximations of reactive transport problems, Computer methods in applied mechanics and engineering, 195, 632–652, doi:10.1016/j.cma.2005.02.021, 2006.

Ter Meer, J., Hartog, N., Valstar, J., Marsman, A., and Rijnaarts, H.: Contaminant sources, pathways and receptors at the Rotterdam Port Area–a non-site-specific approach, Grundwasser und Boden, p. 71, 2007.

Tobiszewski, M. and Namieśnik, J.: Abiotic degradation of chlorinated ethanes and ethenes in water, Environmental Science and Pollution Research, 19, 1994–2006, doi:10.1007/s11356-012-0764-9, 2012.

Todd, D. K. and Mays, L. W.: Groundwater hydrology edition, Wiley, New Jersey, 2005.

Vermeulen, P., Quan, N. H., Nam, N. D. G., Van Hungd, P., Tungd, N. T., Thanhe, T. V., and Dam, R.: Groundwater modeling for the Mekong Delta using iMOD, 20th International Congress on Modelling and Simulation, Adelaide, Australia, 2013.

Vogel, T. M. and McCARTY, P. L.: Biotransformation of tetrachloroethylene to trichloroethylene, dichloroethylene, vinyl chloride, and carbon dioxide under methanogenic conditions, Applied and Environmental Microbiology, 49, 1080–1083, 1985.

Wang, X., Hamill, T. M., Whitaker, J. S., and Bishop, C. H.: A comparison of hybrid ensemble transform Kalman filter-optimum interpolation and ensemble square root filter analysis schemes, Monthly weather review, 135, 1055–1076, doi:10.1175/mwr3307.1, 2007.

Wang, X., Hamill, T. M., Whitaker, J. S., and Bishop, C. H.: A comparison of the hybrid and EnSRF analysis schemes in the presence of model errors due to unresolved scales, Monthly Weather Review, 137, 3219–3232, doi:10.1175/2009mwr2923.1, 2009.

Whitaker, J. S. and Hamill, T. M.: Ensemble data assimilation without perturbed observations, Monthly Weather Review, 130, 1913–1924, 2002.

Zheng, C. and Wang, P. P.: MT3DMS: a modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide, Tech. rep., DTIC Document, 1999.

Zhou, H., Gómez-Hernández, J. J., Hendricks Franssen, H.-J., and Li, L.: An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering, Advances in Water Resources, 34, 844–864, doi:10.1016/j.advwatres.2011.04.014, 2011.

Zhou, H., Li, L., Franssen, H.-J. H., and Gómez-Hernández, J. J.: Pattern recognition in a bimodal aquifer using the normal-score ensemble Kalman filter, Mathematical Geosciences, 44, 169–185, doi:10.1007/s11004-011-9372-3, 2012.

Table/Figure Captions:

Table 1: Different modelling parameters for the coupled flow-transport-reaction model. The values given for 3D parametric fields, such as bulk density and distribution coefficients, are the mean values from the entire 121 layers.

Figure 1: Schematic representation of the port of Rotterdam area with three main geologic layers: (i) Holocene clay and peat layer with sandy deposits (≈ 20 m thick), (ii) Pleistocene layer with coarse sand (≈ 10 m thick), and (iii) Pleistocene clay layer (≈ 30 m thick). POC1, POC2 and POC3 refer to different planes of compliance at the port site.

Figure 2: Initial configuration and geometry of the study area, located at the port of Rotterdam. The blue part is the domain area (1.5 km^2) of each layer and the yellow region is the plume of chlorinated hydrocarbon contaminants located in layer 60 at a depth of 22.5 m below the mean sea level. The green triangles indicate the measurement locations collected from layers 30, 50, 70 and 90.

Figure 3: 2D spatial configuration of sorption (distribution coefficient) for trichloroethene (TCE) averaged over the first 10 layers of the domain.

Figure 4: Groundwater (GW) hydraulic head configuration from four different active layers in the domain. The largest water head is located in the center of the domain and is equal to 1.5 m. The water head deceases in the southern part of the domain. The flow is computed using MODFLOW and plotted using iMOD's graphical interface utility.

Figure 5: Contaminant plume after 50 years for PCE, TCE, DCE and VC in layers 40, 60, 80 and 100, respectively. Vertically, the contaminant plume tend to move downwards towards the Pleistocene clay layers and the second aquifer. In the lateral direction, displacement of the plume happens southwards.

Figure 6: A sketch illustrating the procedure followed to construct the background statistics, $\mathbf{P}^b_{xx}$ and $\mathbf{P}^b_{\theta x}$. 3-months forecasts are performed starting from different initial conditions, $\mathbf{x}_{0,1,\ldots,N}$, and different degradations rate parameters, $\mathbf{\Theta}_{0,1,2\ldots,N}$, where $N = 200$ steps summing up to 50 years. The background state covariance, $\mathbf{P}^b_{xx}$, and state-parameters cross-correlations, $\mathbf{P}^b_{\theta x}$, are then constructed using the first leading modes only.

Figure 7: Individual cross-correlation terms of the background matrix $\mathbf{P}^b_{\theta x}$ associated with PCE biodegradation rate and VC concentration. The correlations are shown for all layers, assuming that the cells from each layer have been stretched in one vertical line (y-coordinate). Largest correlation is present in layer 60 where the contaminant source is located. Biodegradation in shallow layers is not as strong as in deep layers because of the downwards groundwater flow direction.

Figure 8: Left panel: Average mean square errors (MSE) of contaminant concentrations obtained using the EnKF and the hybrid EnKF-OI using $\alpha = 0.1, 0.2, \ldots, 0.9$. Right panel: Average MSE for PCE, TCE, DCE and VC biodegradation rates obtained using the EnKF and the hybrid EnKF-OI for different weight factor (α) values.

Figure 9: Time series change of average ensemble spread (AES) resulting from the EnKF and hybrid EnKF-OI using 48 members in which $\alpha = 0.1, 0.2, \ldots, 0.9$.

Figure 10: Bar-Plot (left y-axis): The AES of concentration ensemble obtained using the hybrid EnKF-OI by changing the individual weighting factors (α and β) between 0.1 and 0.9. Shown according to the right y-axis is the MSE obtained using the hybrid EnKF-OI$_{\alpha=0.7}^{\beta=0.1,\ldots,0.9}$ (triangles), the hybrid EnKF-OI$_{\alpha=0.7}$ (cross) and the EnKF (plus).

Figure 11: Forecast ensemble members of PCE, TCE, DCE, and VC concentration. The evolution of these members ($N_e = 48$) is shown for the entire 50 years. Results are obtained using the standard EnKF, the EnKF-OI$_{\alpha=0.7}$ and the EnKF-OI$_{\alpha=0.7}^{\beta=0.1}$ schemes. Solid dashed lines correspond to the reference concentration of each component.

Figure 12: Images showing the MSE of PCE, TCE, DCE and VC degradation rates in time. These are obtained using the hybrid EnKF-OI scheme for (1) different α values (top panel) and (2) different β values keeping α fixed and equal to 0.7 (bottom panel).

Figure 13: The coloured image shows, according to the left y-axis, the adaptive change in α values for each individual observation. The observation index $(1, \ldots, 160)$ is sorted such that the first 40 indices correspond to PCE measurements, the second 40 correspond to TCE, third 40 correspond to DCE and finally VC takes the last 40 indices. The yellow color indicates that no background covariance matrices have been used and the blue color suggests that only ensemble "flow-dependent" statistics are involved. The curves demonstrate the change in MSE, according to the right y-axis, in time when maximising the information gain (cyan), minimising the information gain (red) and maximising the formation for concentration and minimising it for degradation rates (green).

Figure 14: Top panel: TCE concentration and error maps in layer 70 obtained using the reference run (1$^{\text{st}}$ column), the EnKF (2$^{\text{nd}}$ column) and the EnKF$_{\text{ESOS}}$ (3$^{\text{rd}}$). Bottom panel: Same as top panel but for the concentration of DCE.

Figure 15: Left y-axis: The time evolution of the prior probability density functions corresponding to TCE degradation rate obtained using the EnKF (solid lines) and the EnKF$_{\text{ESOS}}$ (dashed lines). The reference "true" rate is given at 0.086 /day in brown color. Right y-axis: The AES of K_T suggested using the EnKF and the EnKF$_{\text{ESOS}}$.

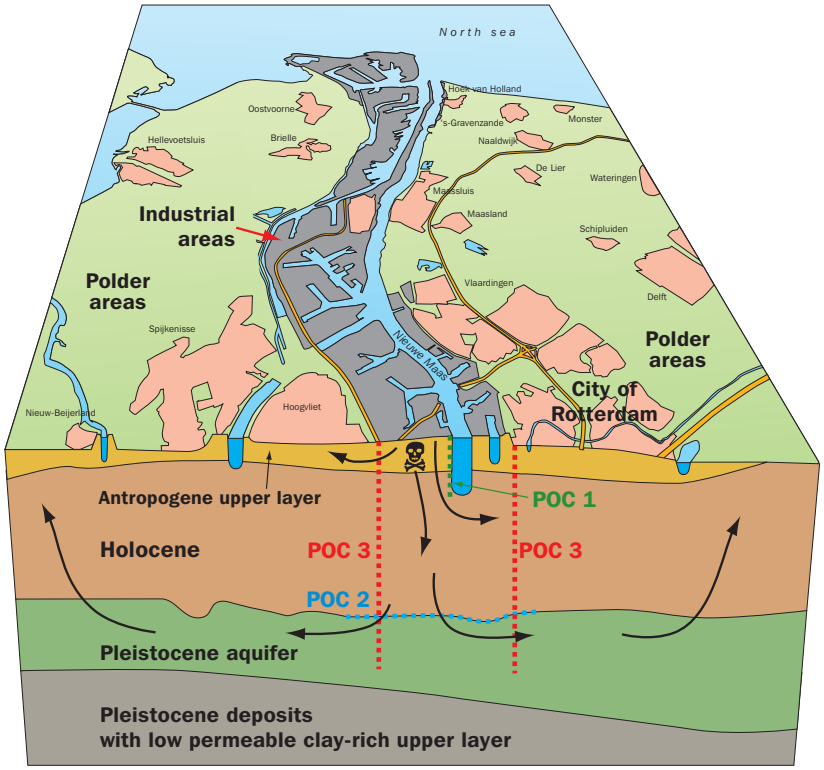
Figure 16: Left panel: Time-series of MSE for concentrations obtained using the EnKF, the EnKF$_{\text{ESOS}}$ and the hybrid EnKF-OI$_{\alpha=0.7}^{\beta=0.1}$. Right panel: Same as the left panel but for all degradation rates (y-axis is in log scale).

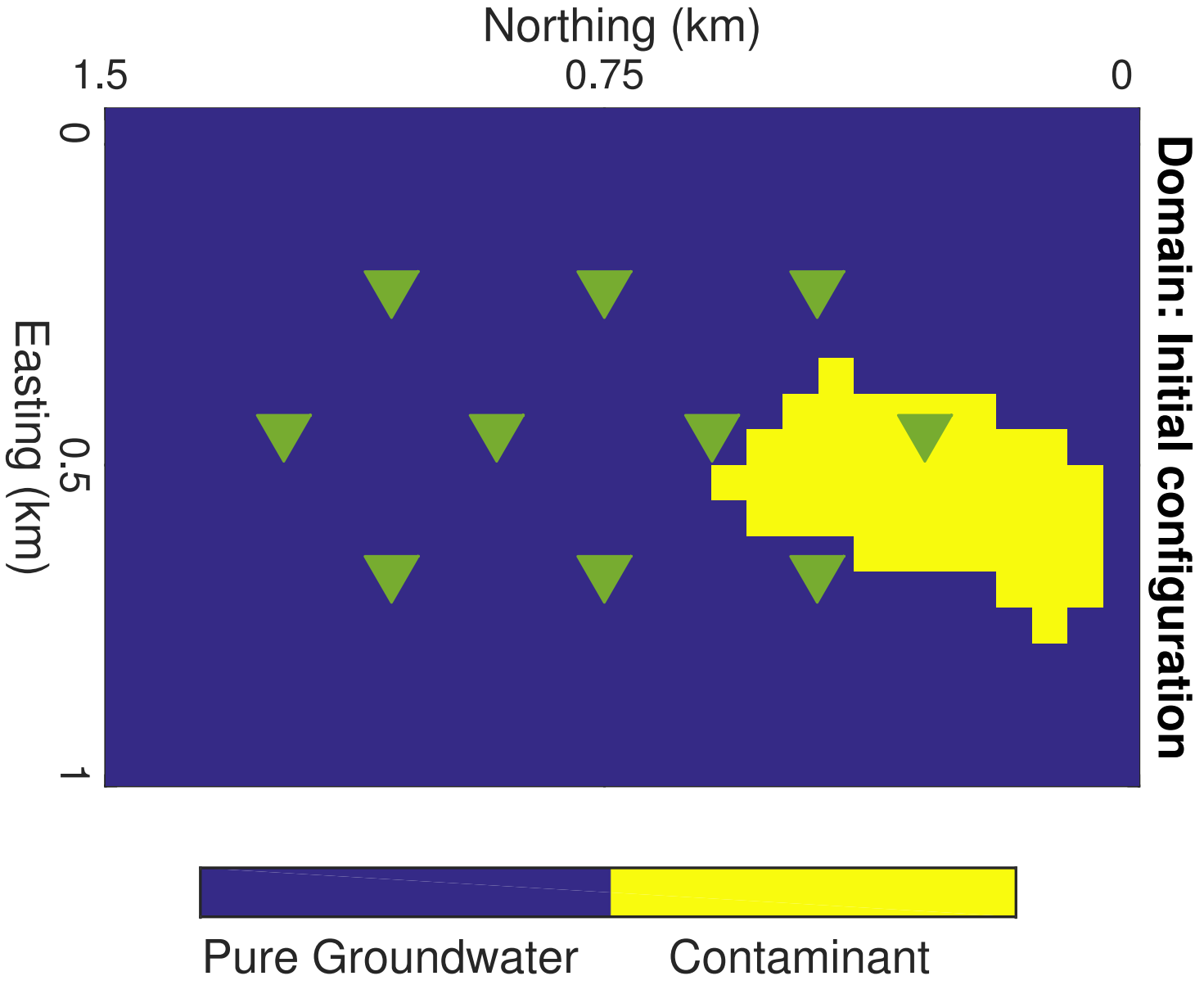
Figure 17: Cross-correlation terms of the background matrix $\mathbf{P}_{\theta x}^{b}$ associated with PCE biodegradation rate and all components' concentartions. The spatial correlations are averaged for all layers. The correlations in black are based on perfect groundwater flow model. The correlations in blue are constructed after perturbing the hydraulic conductivity and porosity of the forecast model. Largest correlation is present in layer 60 where the contaminant source is located.

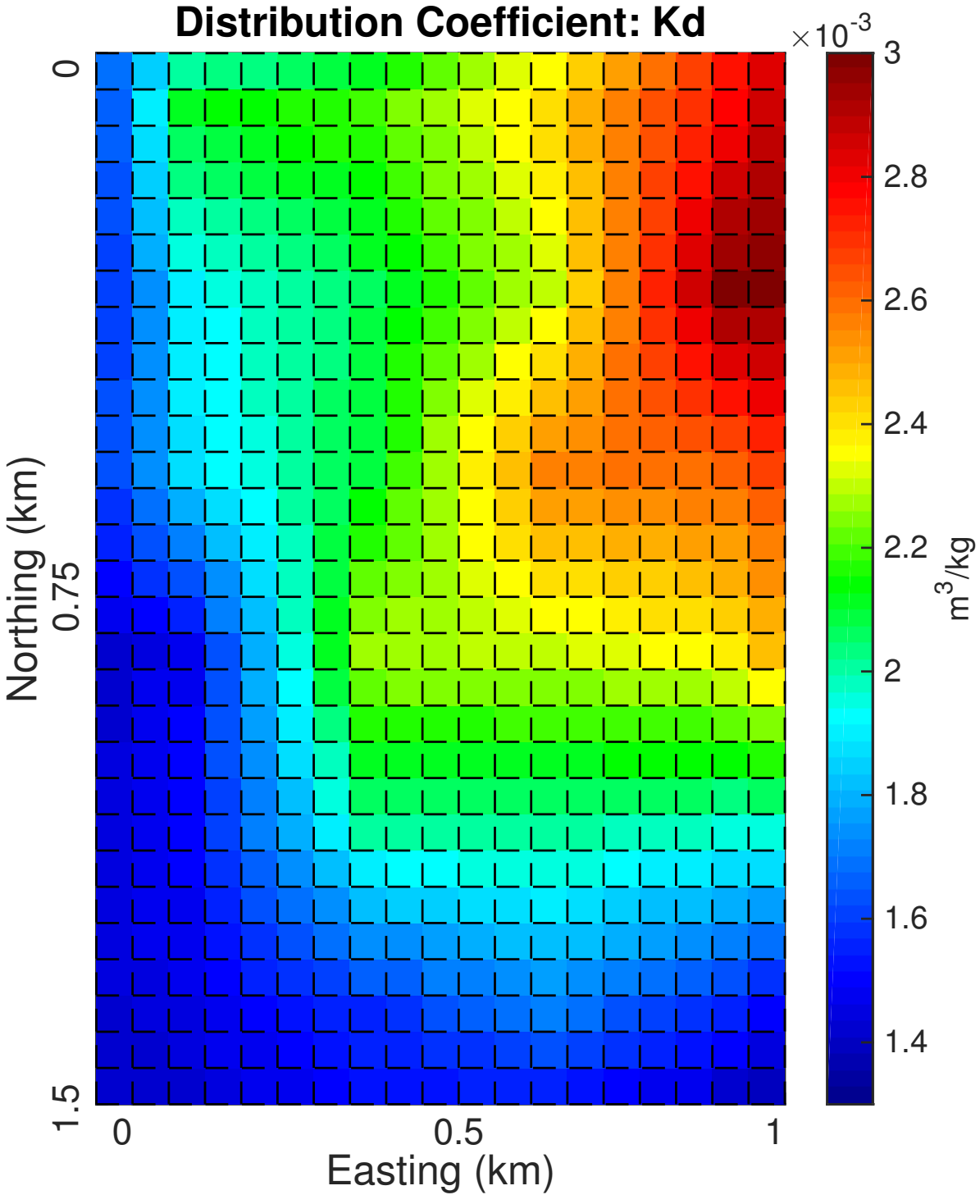
Figure 18: Top panel: Performance of the hybrid EnKF-OI$_{\alpha=0.3,0.4,\ldots,0.8}^{\beta=1.0}$ before and after perturbing the hydraulic conductivity and porosity. The performance is assessed based on the temporal evolution of the MSE (prediction) errors.

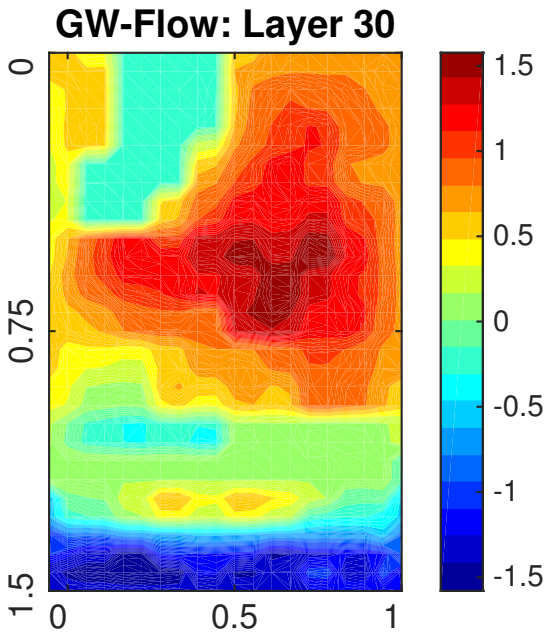
The imposed Gaussian uncertainty is high and obtained using standard deviation σ_h. Bottom panel: Average gain $\frac{1}{50}\sum_{i=1}^{50}\left(\text{MSE}_{\text{EnKF-OI}} - \text{MSE}_{\text{EnKF-ESOS}}\right)$ suggested for all degradation parameters. The improvements suggested by the adaptive EnKF-OI over $\text{EnKF}_{\text{ESOS}}$ are shown for both perfect and perturbed (using σ_m) flow scenarios.

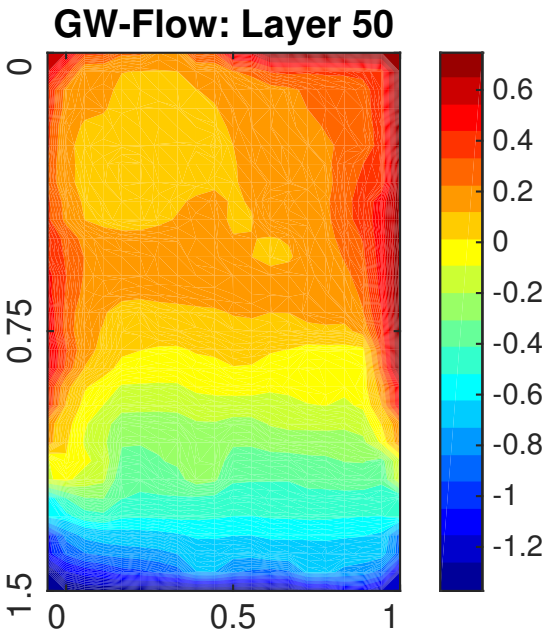
Symbol	Parameter Description	Value (unit)
ϕ	Porosity	0.30
ρ_b	Bulk density	1167 (kg/m^3)
k_{PCE}	Distribution coefficient of PCE	0.0012 (m^3/kg)
k_{TCE}	Distribution coefficient of TCE	0.0015 (m^3/kg)
k_{DCE}	Distribution coefficient of DCE	0.0014 (m^3/kg)
k_{VC}	Distribution coefficient of VC	0.0010 (m^3/kg)
κ_L	Longitudinal dispersivity	0.5 (m)
$\dot{\kappa}_T/\kappa_L$	Ratio of horizontal transverse dispersivity to longitudinal dispersivity	0.1
$\ddot{\kappa}_T/\kappa_L$	Ratio of vertical transverse dispersivity to longitudinal dispersivity	0.1
δ_m	Molecular diffusion	10^{-10} (m^2/s)

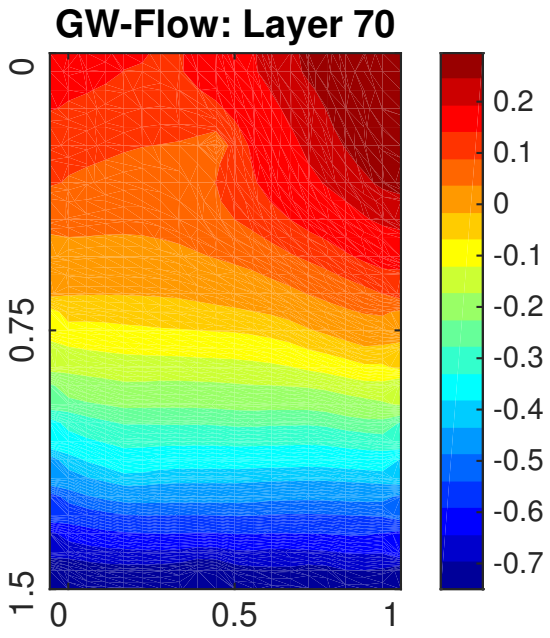

North sea
Hoek van Holland
Oostvoorne
Hellevoetsluis
Brielle
s-Gravenzande
Monster
Naaldwijk
De Lier
Wateringen
Maassluis
Maasland
Schipluiden
Industrial areas
Polder areas
Vlaardingen
Delft
Spijkenisse
Nieuwe Maas
Polder areas
City of Rotterdam
Nieuw-Beijerland
Hoogvliet
Antropogene upper layer
POC 1
Holocene
POC 3
POC 3
POC 2
Pleistocene aquifer
Pleistocene deposits
with low permeable clay-rich upper layer

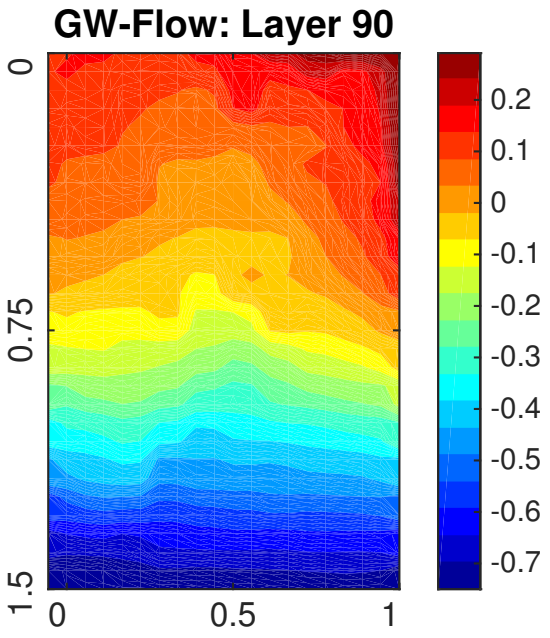

Northing (km)
1.5
0.75
0
0
0.5
1
Easting (km)
Domain: Initial configuration
Pure Groundwater
Contaminant

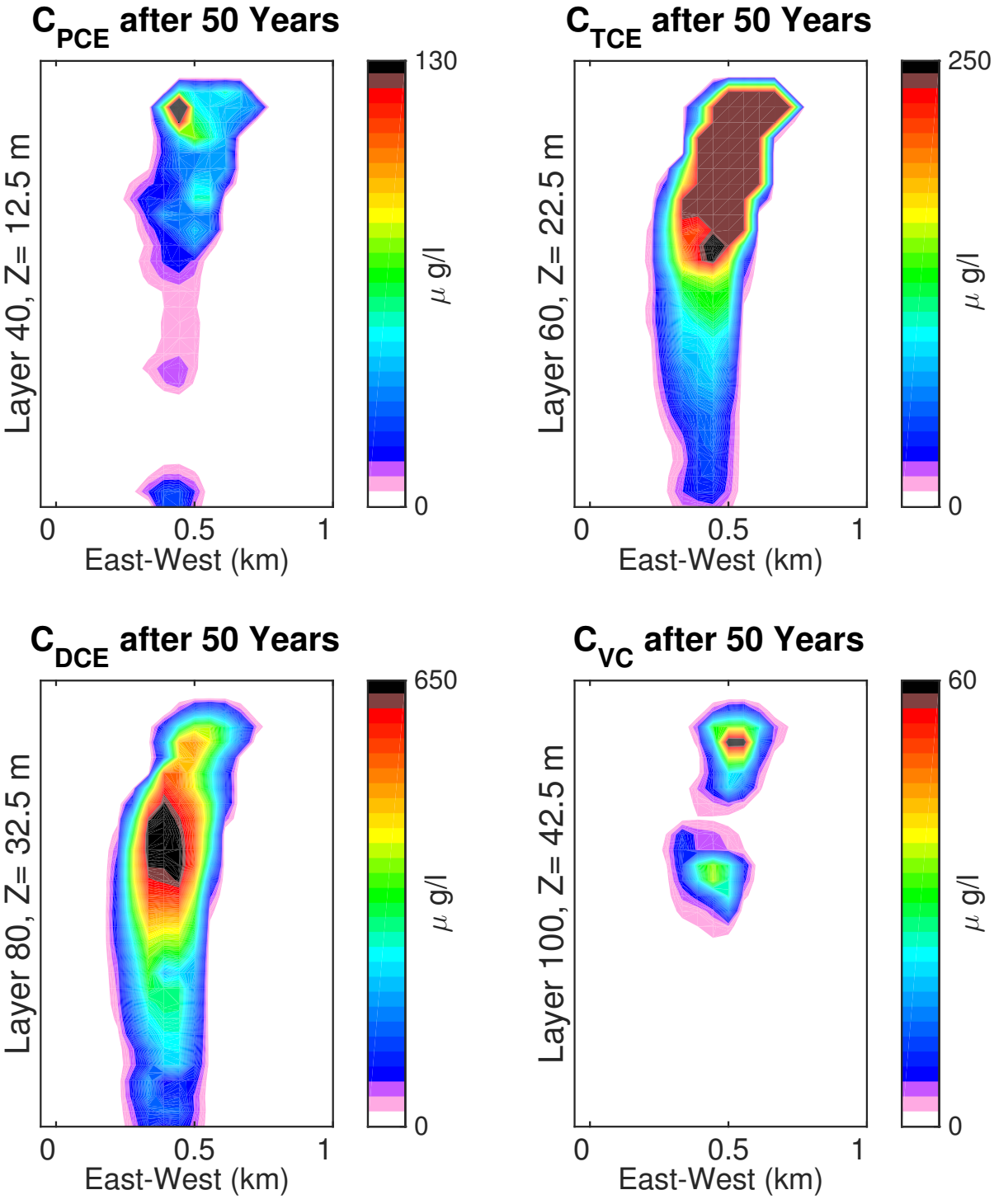

Distribution Coefficient: Kd
×10-3
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
m3/kg
Northing (km)
0
0.75
1.5
0
0.5
1
Easting (km)

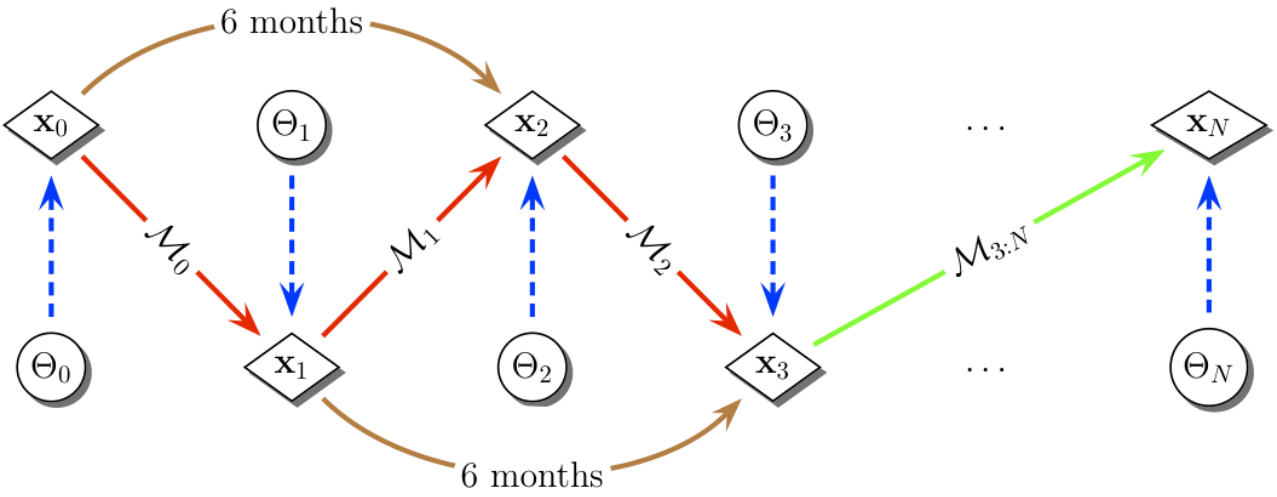

GW-Flow: Layer 30
0
0.75
1.5
0
0.5
1
1.5
1
0.5
0
-0.5
-1
-1.5

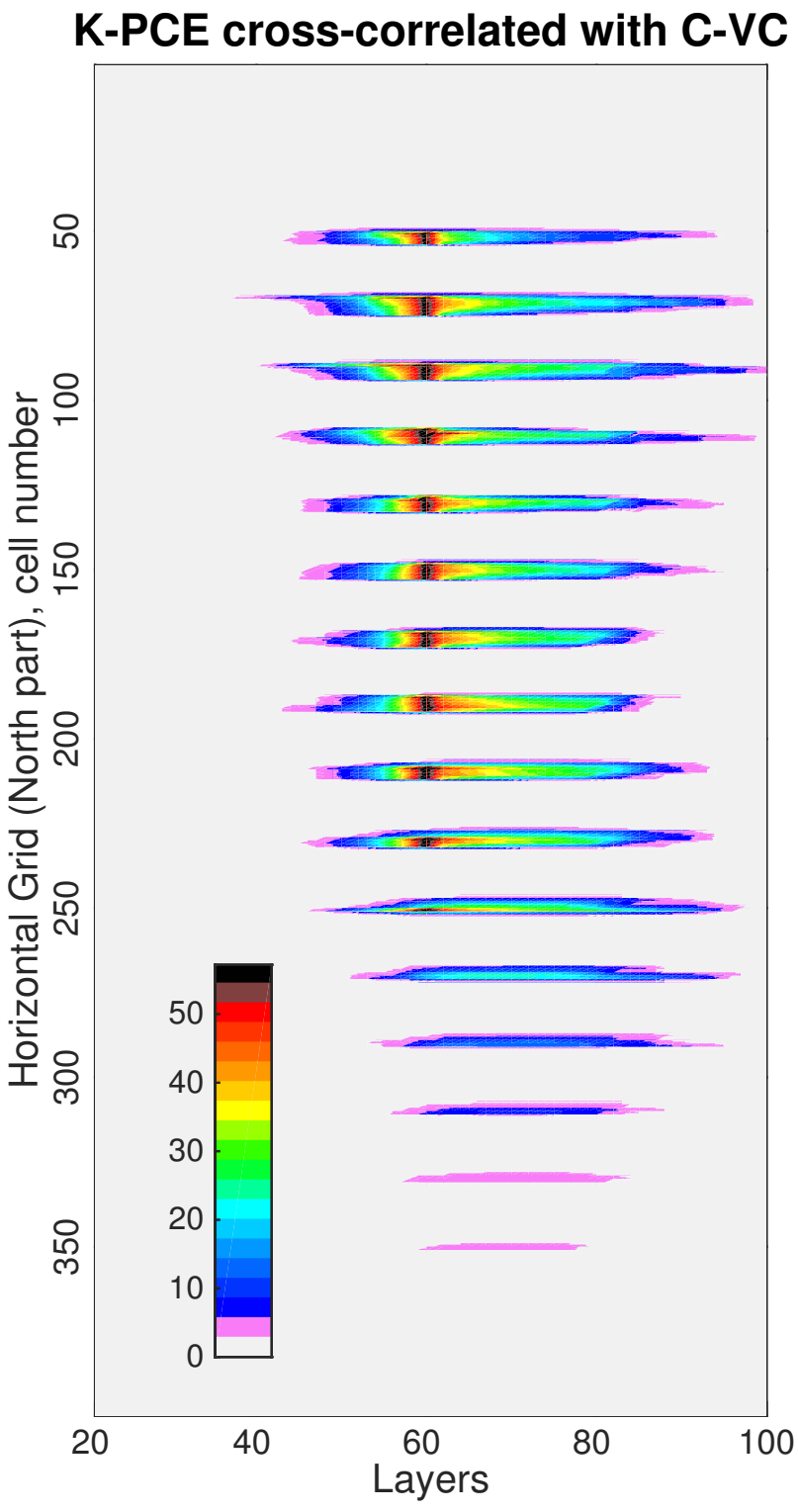

GW-Flow: Layer 50
0
0.75
1.5
0
0.5
1
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2

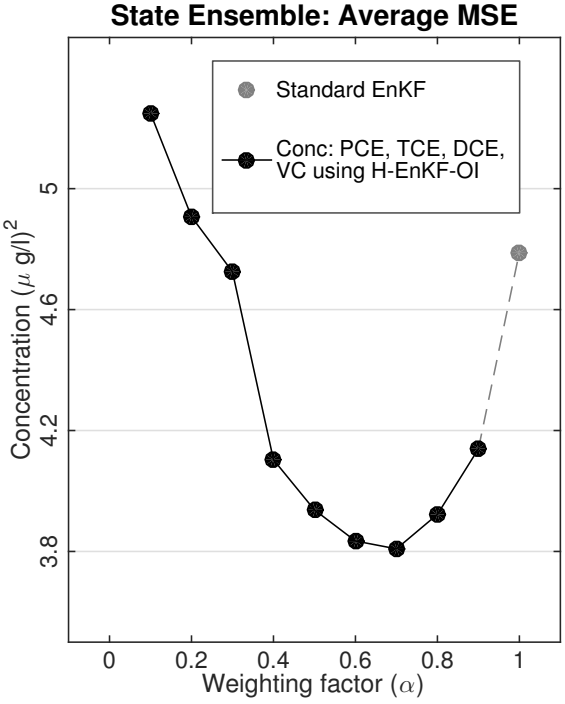

GW-Flow: Layer 70
0
0.75
1.5
0
0.5
1
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7

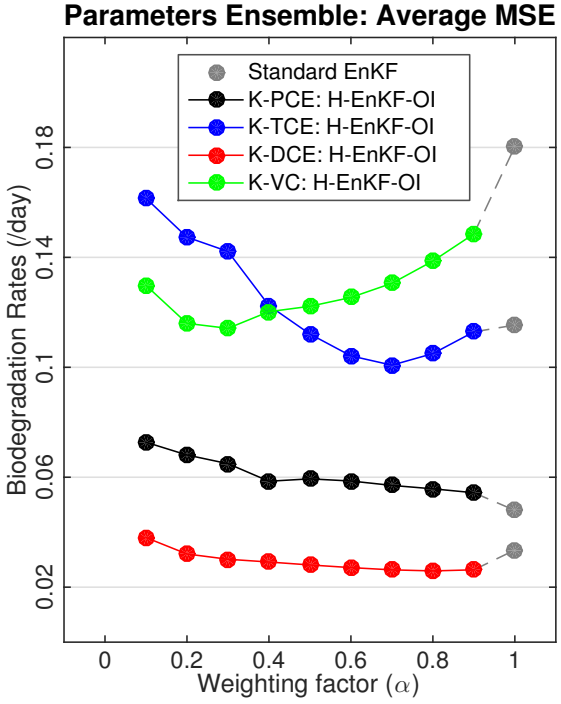

GW-Flow: Layer 90
0
0.75
1.5
0
0.5
1
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7

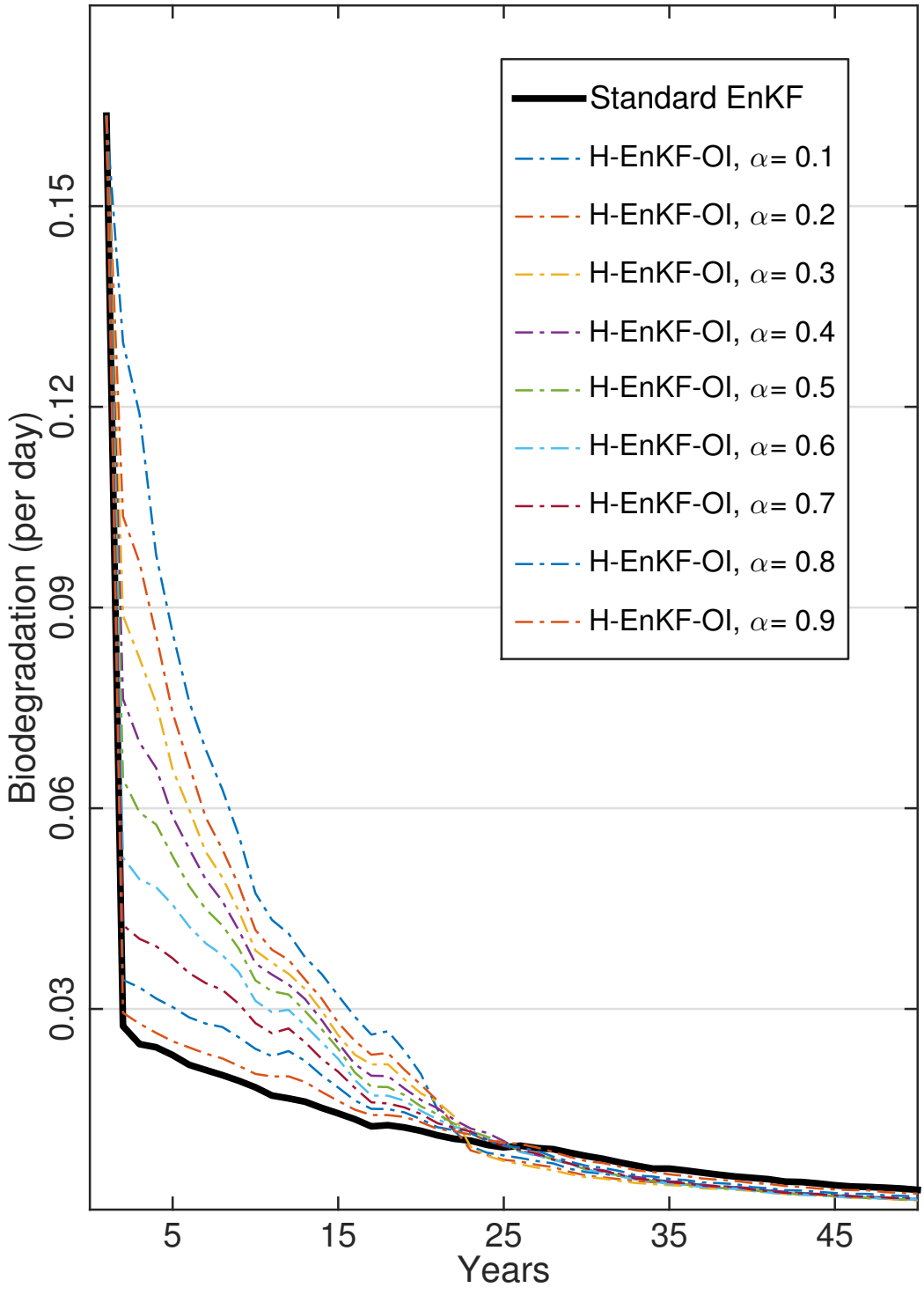

C PCE after 50 Years
Layer 40, Z= 12.5 m
East-West (km)
0
0.5
1
130
0
μ g/l
C TCE after 50 Years
Layer 60, Z= 22.5 m
East-West (km)
0
0.5
1
250
0
μ g/l
C DCE after 50 Years
Layer 80, Z= 32.5 m
East-West (km)
0
0.5
1
650
0
μ g/l
C VC after 50 Years
Layer 100, Z= 42.5 m
East-West (km)
0
0.5
1
60
0
μ g/l

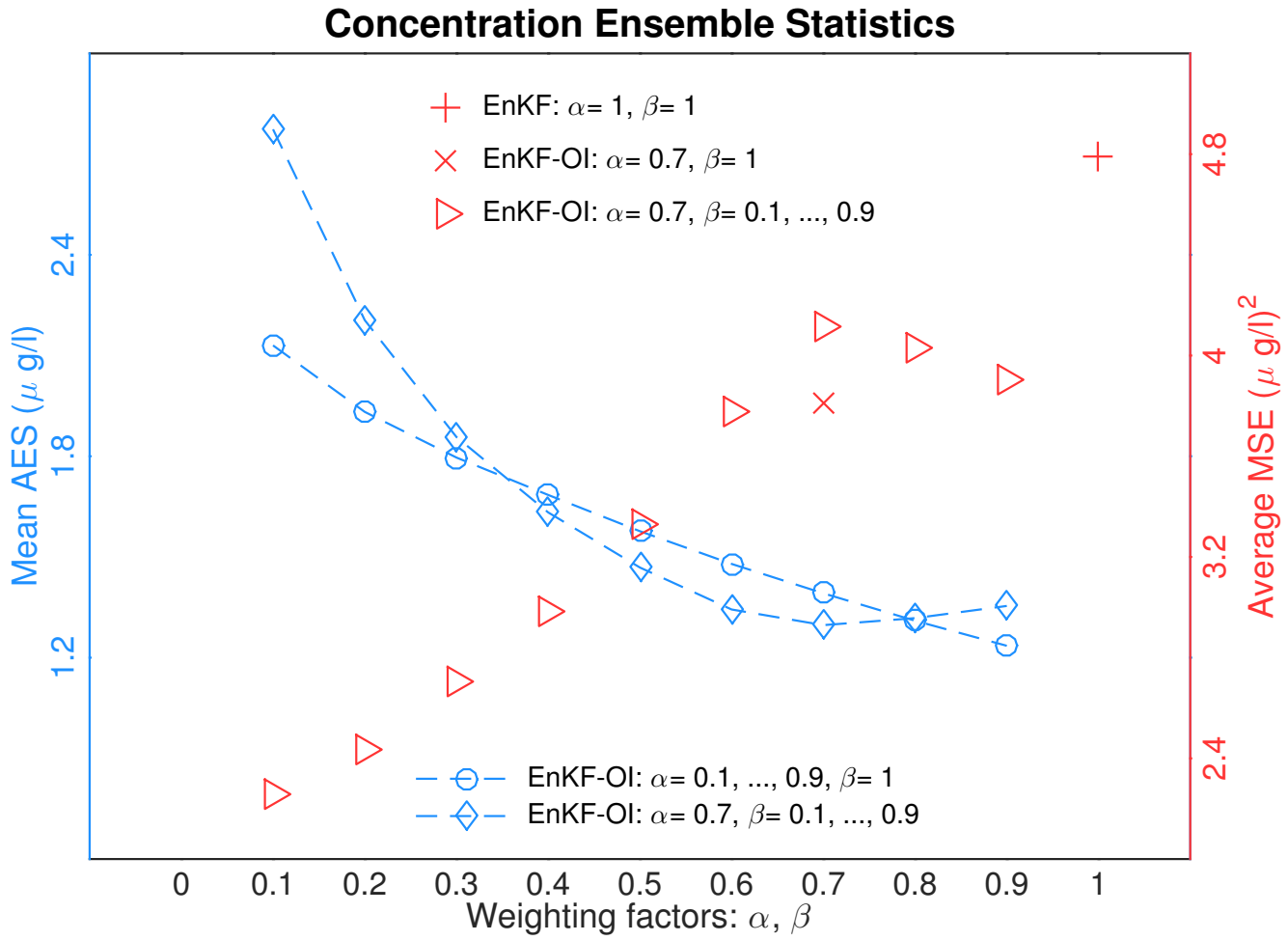

6 months
$\mathbf{x}_0$
Θ_1
$\mathbf{x}_2$
Θ_3
$\cdots$
$\mathbf{x}_N$
$\mathcal{M}_0$
$\mathcal{M}_1$
$\mathcal{M}_2$
$\mathcal{M}_{3:N}$
Θ_0
$\mathbf{x}_1$
Θ_2
$\mathbf{x}_3$
$\cdots$
Θ_N
6 months

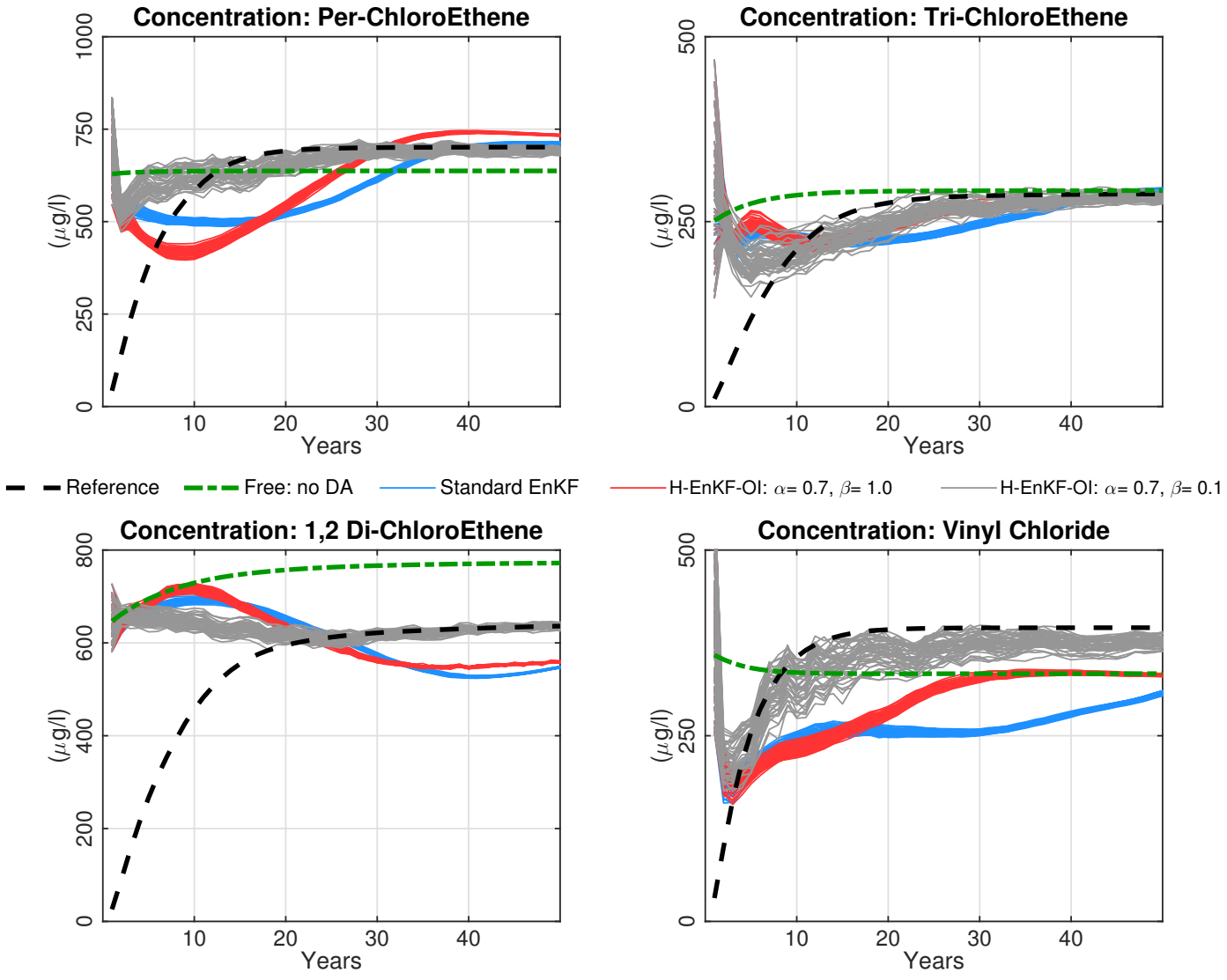

K-PCE cross-correlated with C-VC
Horizontal Grid (North part), cell number
50
100
150
200
250
300
350
50
40
30
20
10
0
20
40
60
80
100
Layers

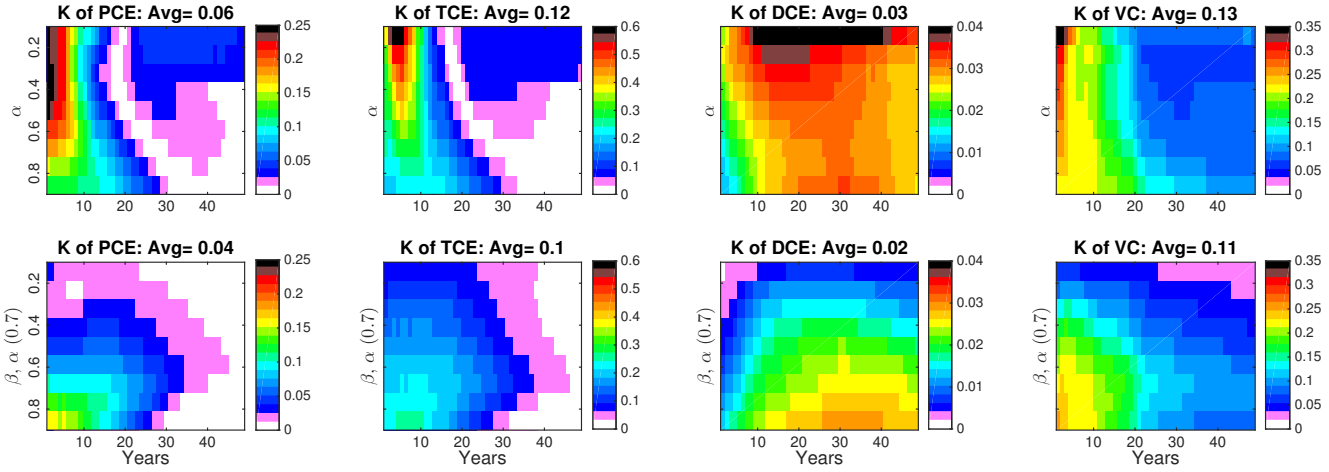

State Ensemble: Average MSE
Standard EnKF
Conc: PCE, TCE, DCE, VC using H-EnKF-OI
Concentration (μ g/l)^2
3.8
4.2
4.6
5
0
0.2
0.4
0.6
0.8
1
Weighting factor (α)

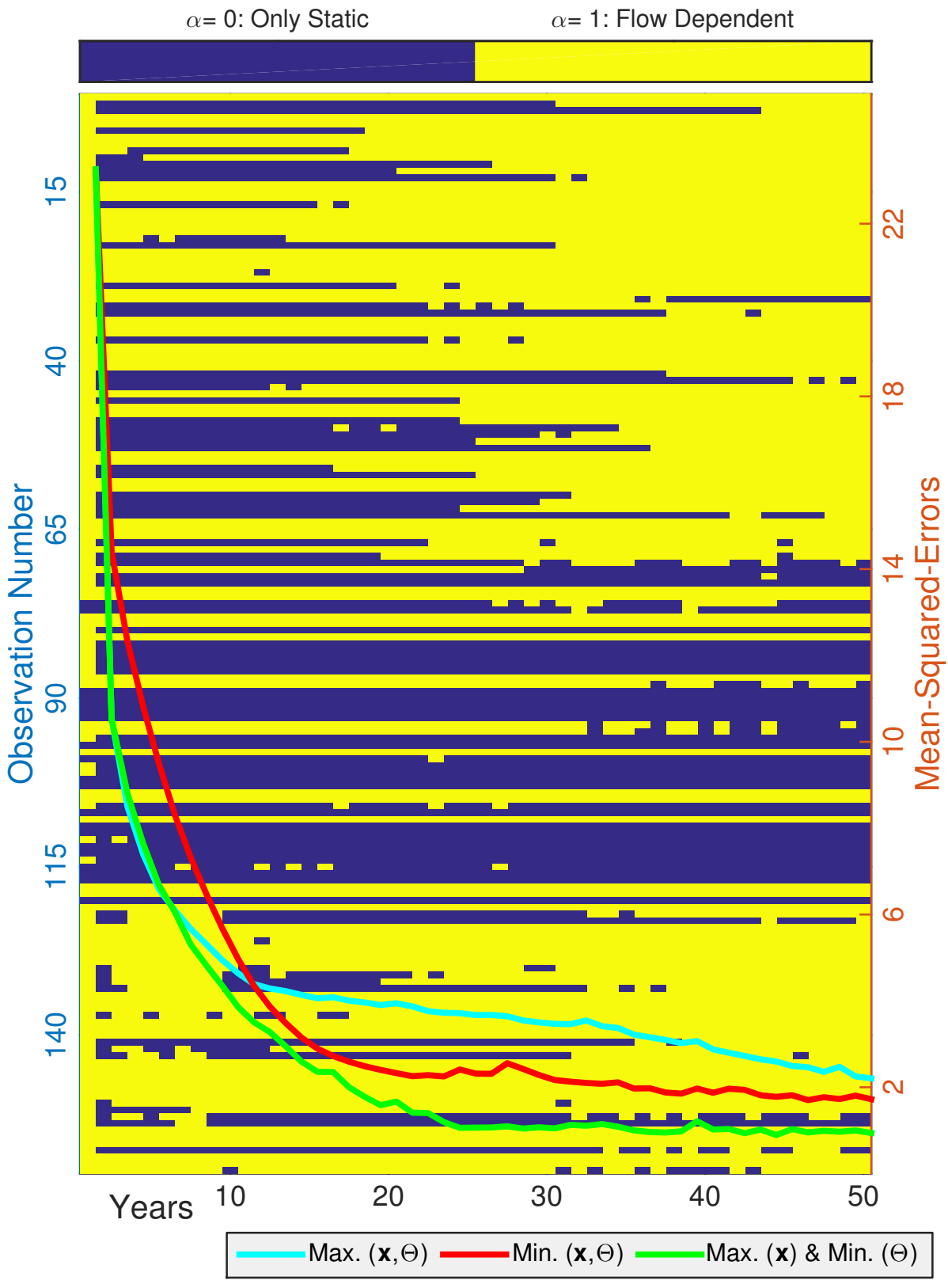

Parameters Ensemble: Average MSE
Standard EnKF
K-PCE: H-EnKF-OI
K-TCE: H-EnKF-OI
K-DCE: H-EnKF-OI
K-VC: H-EnKF-OI
Biodegradation Rates (/day)
0.02
0.06
0.1
0.14
0.18
0
0.2
0.4
0.6
0.8
1
Weighting factor (α)

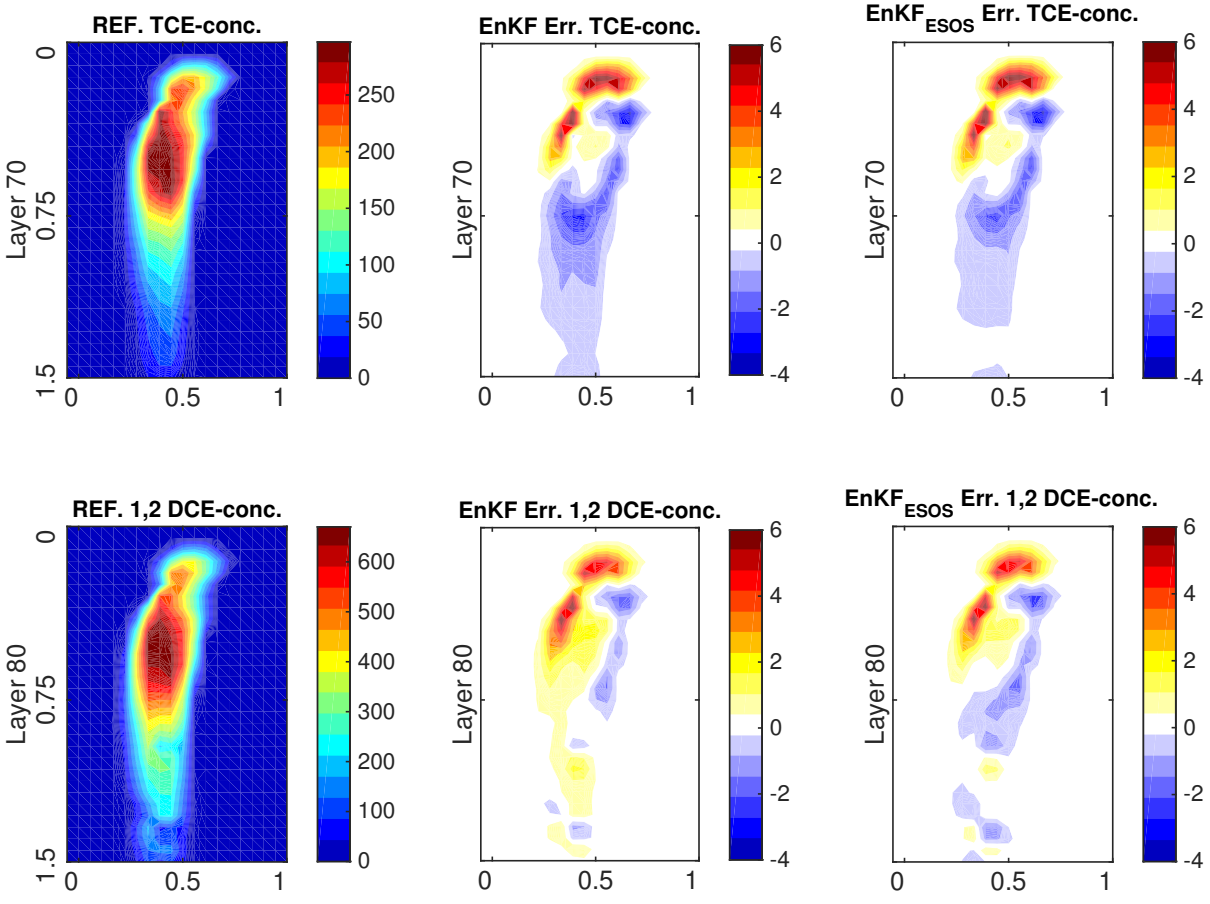

Parameter Statistics: Average-Ensemble-Spread
Standard EnKF
H-EnKF-OI, α= 0.1
H-EnKF-OI, α= 0.2
H-EnKF-OI, α= 0.3
H-EnKF-OI, α= 0.4
H-EnKF-OI, α= 0.5
H-EnKF-OI, α= 0.6
H-EnKF-OI, α= 0.7
H-EnKF-OI, α= 0.8
H-EnKF-OI, α= 0.9
Biodegradation (per day)
0.03
0.06
0.09
0.12
0.15
5
15
25
35
45
Years

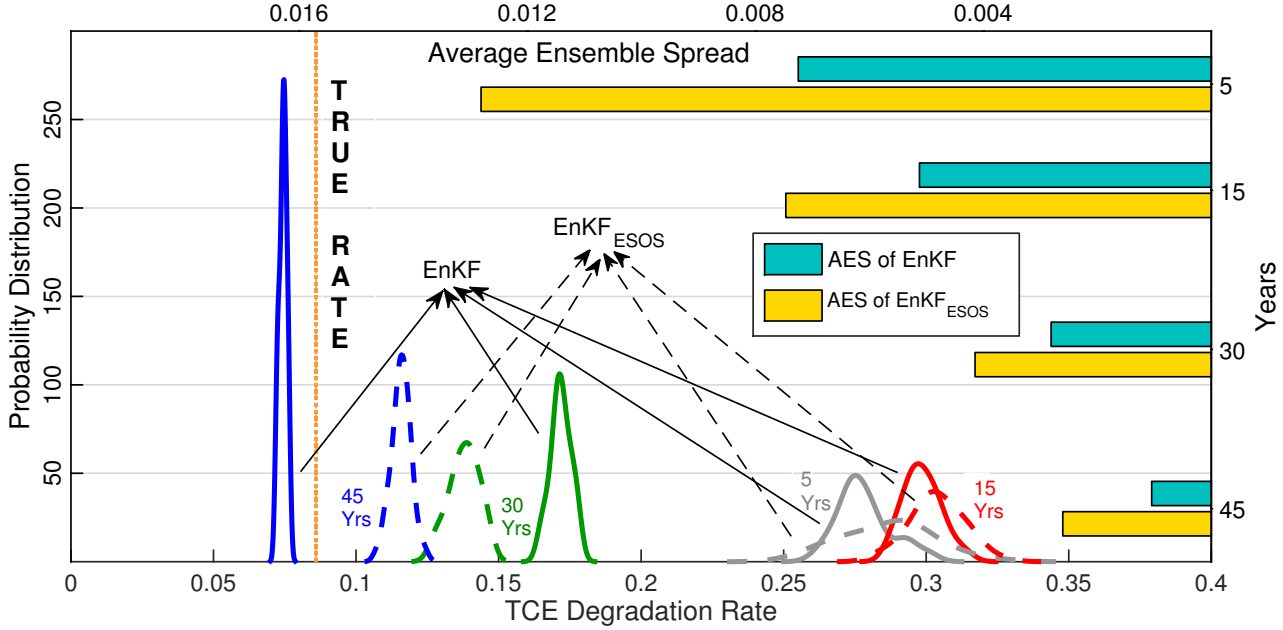

Concentration Ensemble Statistics
EnKF: α= 1, β= 1
EnKF-OI: α= 0.7, β= 1
EnKF-OI: α= 0.7, β= 0.1, ..., 0.9
EnKF-OI: α= 0.1, ..., 0.9, β= 1
EnKF-OI: α= 0.7, β= 0.1, ..., 0.9
Mean AES (μ g/l)
Average MSE (μ g/l)^2
1.2
1.8
2.4
2.4
3.2
4
4.8
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Weighting factors: α, β

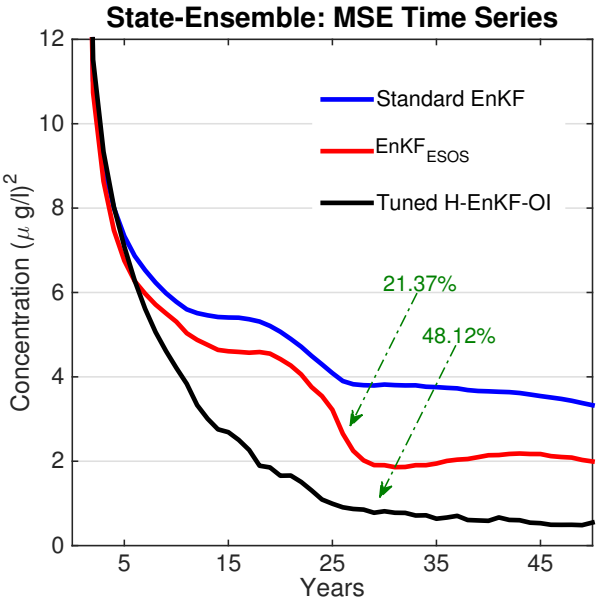

Concentration: Per-ChloroEthene
Concentration: Tri-ChloroEthene
Concentration: 1,2 Di-ChloroEthene
Concentration: Vinyl Chloride
(μg/l)
Years
Reference
Free: no DA
Standard EnKF
H-EnKF-OI: α= 0.7, β= 1.0
H-EnKF-OI: α= 0.7, β= 0.1

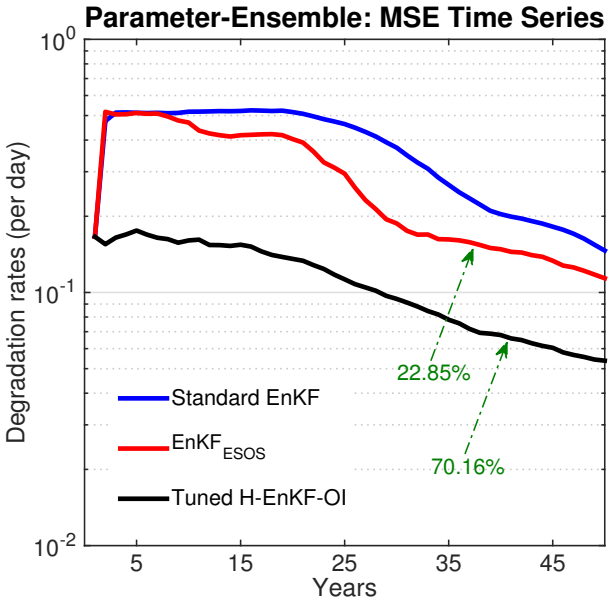

K of PCE: Avg= 0.06
K of TCE: Avg= 0.12
K of DCE: Avg= 0.03
K of VC: Avg= 0.13
K of PCE: Avg= 0.04
K of TCE: Avg= 0.1
K of DCE: Avg= 0.02
K of VC: Avg= 0.11
α
β, α (0.7)
Years
10 20 30 40
0.2 0.4 0.6 0.8
0 0.05 0.1 0.15 0.2 0.25
0 0.1 0.2 0.3 0.4 0.5 0.6
0 0.01 0.02 0.03 0.04
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35

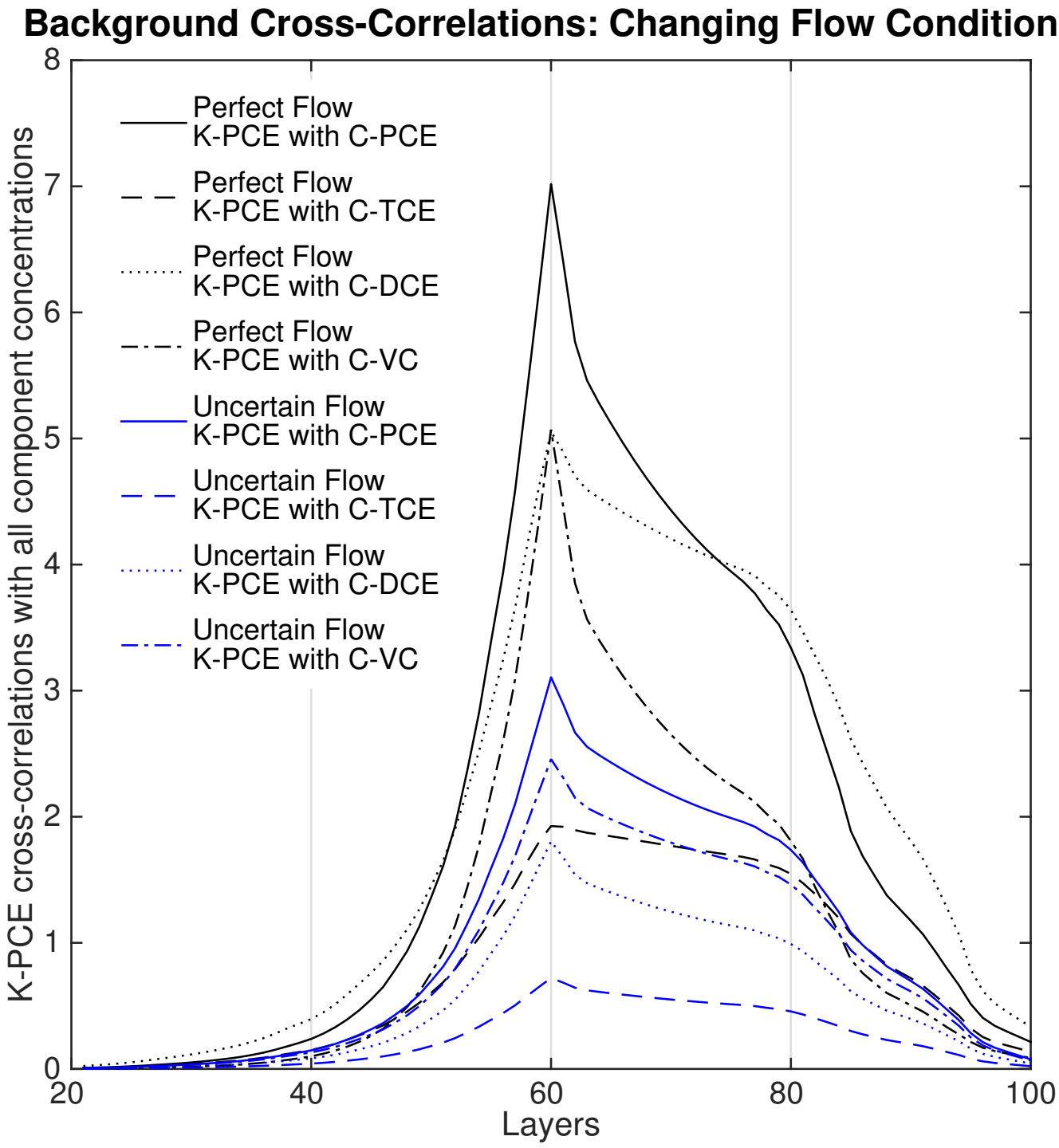

α= 0: Only Static
α= 1: Flow Dependent
Observation Number
15
40
65
90
115
140
Mean-Squared-Errors
22
18
14
10
6
2
Years
10
20
30
40
50
Max. (**x**,Θ)
Min. (**x**,Θ)
Max. (**x**) & Min. (Θ)

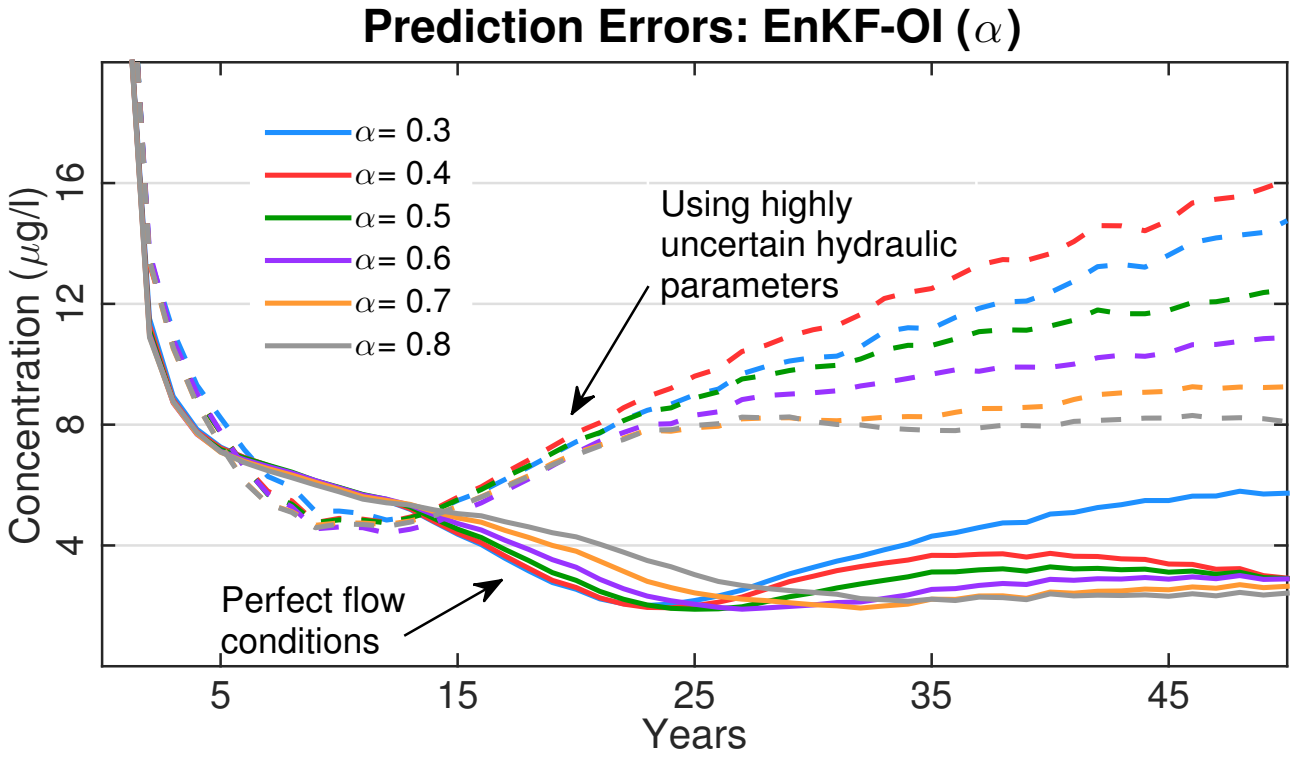

REF. TCE-conc.
EnKF Err. TCE-conc.
EnKF$_{ESOS}$ Err. TCE-conc.
REF. 1,2 DCE-conc.
EnKF Err. 1,2 DCE-conc.
EnKF$_{ESOS}$ Err. 1,2 DCE-conc.
Layer 70
Layer 80
0
0.75
1.5
0.5
1
250
200
150
100
50
600
500
400
300
6
4
2
-2
-4

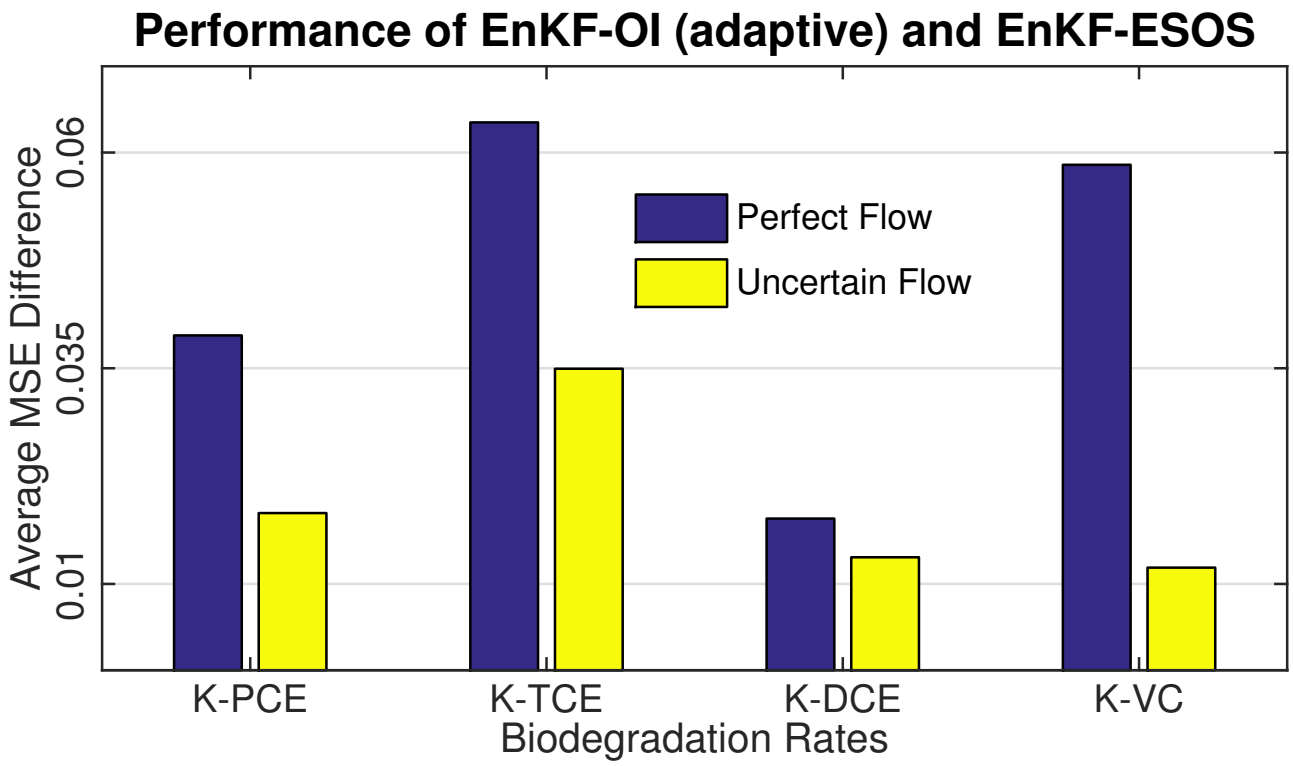

0.016
0.012
0.008
0.004
Average Ensemble Spread
TRUE RATE
EnKF
EnKF ESOS
AES of EnKF
AES of EnKF ESOS
45 Yrs
30 Yrs
5 Yrs
15 Yrs
Probability Distribution
50
100
150
200
250
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
TCE Degradation Rate
5
15
30
45
Years


State-Ensemble: MSE Time Series
Concentration (μ g/l)^2
Years
Standard EnKF
EnKF$_{ESOS}$
Tuned H-EnKF-OI
21.37%
48.12%


Parameter-Ensemble: MSE Time Series
Degradation rates (per day)
Years
Standard EnKF
EnKF$_{ESOS}$
Tuned H-EnKF-OI
22.85%
70.16%


Background Cross-Correlations: Changing Flow Conditions
K-PCE cross-correlations with all component concentrations
Layers
Perfect Flow
K-PCE with C-PCE
Perfect Flow
K-PCE with C-TCE
Perfect Flow
K-PCE with C-DCE
Perfect Flow
K-PCE with C-VC
Uncertain Flow
K-PCE with C-PCE
Uncertain Flow
K-PCE with C-TCE
Uncertain Flow
K-PCE with C-DCE
Uncertain Flow
K-PCE with C-VC
0
1
2
3
4
5
6
7
8
20
40
60
80
100


Prediction Errors: EnKF-OI (α)
α= 0.3
α= 0.4
α= 0.5
α= 0.6
α= 0.7
α= 0.8
Using highly uncertain hydraulic parameters
Perfect flow conditions
Concentration (μg/l)
4
8
12
16
5
15
25
35
45
Years


Performance of EnKF-OI (adaptive) and EnKF-ESOS
Perfect Flow
Uncertain Flow
Average MSE Difference
0.01
0.035
0.06
K-PCE
K-TCE
K-DCE
K-VC
Biodegradation Rates