Advances in Water Resources 109 (2017) 170–180

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

A probabilistic collocation based iterative Kalman filter for landfill data assimilation

Zheng Qiang[a], Xu Wenjie[b], Man Jun[a], Zeng Lingzao[a,*], Wu Laosheng[c]

[a] Institute of Soil and Water Resource and Environmental Science, Zhejiang Provincial Key Laboratory of Agricultural Resources and Environment, Zhejiang University, Hangzhou, 310058, China
[b] MOE Key Laboratory of Soft Soils and Geoenvironmental Engineering, Institute of Civil Engineering, Zhejiang University, Hangzhou, 310058, China
[c] Department of Environmental Sciences, University of California, Riverside, CA 92521, USA

ARTICLE INFO

Keywords:
Data assimilation
Landfill
Iterative Kalman filter
Polynomial chaos

ABSTRACT

Accurate forecast of landfill gas (LFG) transport has remained as an active research area, due to the safety and environmental concerns, as well as the green energy potential. The iterative ensemble Kalman filter (IEnKF) has been used to characterize the heterogeneous permeability field of landfills. As a Monte Carlo-based method, IEnKF requires a sufficiently large ensemble size to guarantee its accuracy, which may result in a huge computational cost, especially for large-scale problems. In this study, an efficient probabilistic collocation based iterative Kalman filter (PCIKF) is developed. The polynomial chaos expansion (PCE) is employed to represent and propagate the uncertainties, and an iterative form of Kalman filter is used to assimilate the measurements. To further reduce the computational cost, only the zeroth and first-order ANOVA (analysis of variance) components are kept in the PCE approximation. As demonstrated by two numerical case studies, PCIKF shows significant superiority over IEnKF in terms of accuracy and efficiency. The developed method has the potential to reliably predict and develop best management practices for landfill gas production.

1. Introduction

Sanitary landfill is a main way to dispose municipal solid waste around the world. Strict environmental rules and urgent demands for best management practices have contributed to considerable research in modeling the dynamics of landfill gas (LFG), with the purpose of minimizing potential hazards that are associated with its generation and emission (El-Fadel et al., 1997). Meanwhile, LFG, which is mainly composed of CH_4 and CO_2, is also a promising source of renewable green energy (Nyns and Gendebien, 1993), providing an additional incentive for the study of LFG dynamics.

Numerical models have been widely used to characterize the LFG transport (EI-Fadel et al., 1996; Xi and Xiong, 2013). For accurate prediction, it is crucial to identify the model parameters in the governing equations. However, the landfill properties are strongly heterogeneous (Zacharof and Butler, 2004). As these parameters are difficult to determine, they are often indirectly inferred from monitoring data using data assimilation methods.

Among the many data assimilation methods, the ensemble Kalman filter (EnKF) has become the most popular one (Houtekamer et al., 2005; Chen and Zhang, 2006; Evensen et al., 2007; Gharamti et al., 2013; Xue and Zhang, 2014). Generally, EnKF is based on the Gaussian assumption, which makes it inherently suitable for tackling Gaussian linear problems (Evensen, 2007). To improve the performance of EnKF in estimating non-Gaussian parameters, different methods have been proposed. For example, EnKF has been combined with the level set technique for history matching of facies distribution (Chang et al., 2010). The normal-score transformation-based EnKF has been proposed to estimate the logarithmic conductivity of bimodal aquifers (Zhou et al., 2011).

In subsurface flow problems, the logarithmic permeability field is commonly modeled as Gaussian random field. However, the governing equations for multiphase flow in porous media are nonlinear. Furthermore, the observation operator may be nonlinear. Thus, the states and observations may be non-Gaussian distributed even if the input parameters are Gaussian. To improve the performance in strongly nonlinear problems, various iterative forms of EnKF (IEnKF) have been proposed (Aanonsen et al., 2009; Li and Reynolds, 2007; Wen and Chen, 2006). The key in IEnKF is to use ensemble members to calculate the sensitivities, which are obtained by solving adjoint equations in traditional approaches (Wu et al., 1999). Furthermore, iterations from the initial time (Zafari et al., 2006) or the last assimilation step

* Corresponding author.
E-mail address: lingzao@zju.edu.cn (L. Zeng).

http://dx.doi.org/10.1016/j.advwatres.2017.09.015
Received 12 January 2017; Received in revised form 12 September 2017; Accepted 14 September 2017
Available online 19 September 2017

(Wen and Chen, 2006) are employed to avoid the inconsistence between the updated parameters and states in EnKF. For strongly nonlinear problems, an adaptively adjusted iteration step length is usually required (Gu and Oliver, 2007).

No matter which form of EnKF is chosen, the sampling errors in generating the initial ensemble are inevitable. Fishman (1996) reported that the error convergence rate is $1/\sqrt{N_e}$, where N_e is the ensemble size. Thus the accuracy of EnKF largely depends on the ensemble size. On the other hand, the requirement for a large N_e will result in a huge computational cost, especially for large-scale problems. Therefore, a compromise has to be made between the accuracy and affordable computational cost in practical applications.

To alleviate the computational burden, one promising approach is to employ reduced-order modeling techniques (Asher et al., 2015). He et al. (2013) combined EnKF with trajectory piecewise linearization (TPWL). The Karhunen–Loeve (KL) expansion was employed as a dimension reduction tool to parameterize the random parameter field with a relatively small number of Gaussian variables. Then the TPWL was used to enrich the ensemble at a low cost. In this work, we employ the polynomial chaos expansion (PCE), which has gained popularity in uncertainty quantification (Saad and Ghanem, 2009). With the PCE, stochastic information is expressed by the orthogonal polynomials of random variables. Then different methods, e.g., stochastic Galerkin projection (Ghanem and Spanos, 2003), regression method (Isukapalli et al., 1998), and probabilistic collocation method (PCM) (Tatang et al., 1997; Sun et al., 2013) can be employed to calculate the PCE coefficients. It has been shown that, PCE-based methods exhibit faster convergence compared with Monte Carlo-based methods (Xiu, 2010).

For efficient data assimilation, a probabilistic collocation based Kalman filter (PCKF), which combines the Kalman filter with PCM, has been proposed by Zeng and Zhang (2010). Since PCM is a non-intrusive method, any existing simulator can be directly used to solve the governing equations with collocational parameter realizations. It should be pointed out that the computational cost of this method drastically increases with the number of PCE terms. Therefore, the superiority of PCKF over EnKF is only valid under certain conditions (Zeng et al., 2011). To further improve the performance, the adaptive probabilistic collocation based Kalman filter (APCKF) has been developed based on the combination of PCKF and functional analysis of variance (ANOVA) (Li et al., 2014). In APCKF, a few important low-dimensional orthogonal polynomial functions are adaptively identified. It has been shown that, based on a restart scheme, APCKF outperforms traditional EnKF in data assimilation for unsaturated flow (Man et al., 2016). To better handle the nonlinearity and further improve the accuracy, a more sophisticated scheme, i.e., the iterative form of EnKF with an adaptively adjusted analysis step length (Gu and Oliver, 2007), may be needed.

A landfill is a large-scale porous medium, and the gas generation and transport dynamics therein are nonlinear. To better predict the LFG transport, Li et al. (2012) used EnKF to estimate the heterogeneous permeability field of a single-phase gas transport model. However, in addition to the gas phase, landfills usually contain leachate from different sources, such as the precipitation, intra-particle water released from wastes when undergoing biochemical degradation and compression by self-gravity, and recirculation of leachate. When the leachate level gets higher, it becomes more difficult to collect gas from the landfill, because the relative gas permeability reduces drastically as the water content increases in a multiphase flow system. Therefore, it is important to consider the liquid-gas coupling, especially in China where municipal solid wastes are usually with high water content (Chen et al., 2010; Yao et al., 2015).

Motivated by recent progresses in EnKF and multiphase flow modeling for landfill, we develop a probabilistic collocation based iterative Kalman filter (PCIKF) for data assimilation of liquid-gas flow in landfills. The performance of this method is tested with numerical cases and compared with that of IEnKF. The remainder of this paper is organized as follows: Section 2 presents the mathematical models for liquid and gas flow in a sanitary landfill. Section 3 provides algorithm details of PCIKF, and illustrative examples are presented in Section 4. Finally, some conclusions are given in Section 5.

2. System model

In this study, we focus on landfill gas transport with the existence of leachate. The governing equations are shown below (Pinder and Gray, 2008)

$$\phi\rho_w \frac{\partial S_w}{\partial p_c}\frac{\partial p_c}{\partial t} + \nabla q_w = 0 \tag{1}$$

$$-\phi\rho_g \frac{\partial S_w}{\partial p_c}\frac{\partial p_c}{\partial t} + \phi(1-S_w)\rho_g\frac{\partial \rho_g}{\partial p_g}\frac{\partial p_g}{\partial t} + \nabla q_g = 0 \tag{2}$$

where ϕ is the porosity [-]; p_c [ML^{-1}T^{-2}] is the capillary pressure, defined as the difference between the gas pressure p_g [ML^{-1}T^{-2}] and liquid pressure p_w [ML^{-1}T^{-2}], i.e.,

$$p_c = p_g - p_w \tag{3}$$

The phase fluxes q_w and q_g [ML^{-2}T^{-1}] can be calculated through the modified Darcy's law

$$q_w = -\rho_w\frac{k_{ini}k_{r,w}}{\mu_w}(\nabla p_g - \nabla p_c + \rho_w g) \tag{4}$$

$$q_g = -\rho_g\frac{k_{ini}k_{r,g}}{\mu_g}(\nabla p_g + \rho_g g) \tag{5}$$

where k_{ini} is the intrinsic permeability [L^2]; μ_w and μ_g are the viscosities for liquid and gas [ML^{-1}T^{-1}], respectively; the relative permeabilities of liquid phase $k_{r,w}$ and gas phase $k_{r,g}$ are described by the Mualem approach (Mualem, 1976)

$$k_{r,w} = S_e^{1/2}[1-(1-S_e^{m/(m-1)})^{(m-1)/m}]^2 \tag{6}$$

$$k_{r,g} = (1-S_e)^{1/2}(1-S_e^{m/(m-1)})^{2(1-1/m)} \tag{7}$$

where m is the shape factor; the capillary pressure p_c varies with the effective water saturation S_e [L^{3}L^{-3}] as follows (van Genuchten, 1980)

$$p_c = p_e(S_e^{m/(m-1)}-1)^{1/m} \tag{8}$$

$$S_e = \frac{S_w - S_{wr}}{1-S_{wr}-S_{gr}} \tag{9}$$

where p_e is the gas entry pressure; S_{wr} and S_{gr} are the residual saturations of the liquid and gas phases, respectively.

3. Methods

3.1. Iterative ensemble Kalman filter (IEnKF)

EnKF is essentially a Monte Carlo-based variant of the standard Kalman filter (Evensen, 2007). Assume that the response of a system can be simulated by

$$\mathbf{d} = F(\mathbf{x}) \tag{10}$$

where $F(\cdot)$ represents the system model; $\mathbf{d}$ is the model output; $\mathbf{x}$ denotes the joint state vector consisting of model parameters $\mathbf{m}$ and state variables $\mathbf{r}$

$$\mathbf{x} = [\mathbf{m}\mathbf{r}]^T \tag{11}$$

Usually, model parameters $\mathbf{m}$ and state variables $\mathbf{r}$ are updated simultaneously, which makes EnKF very computationally efficient (Anderson, 2001). However, for a strongly nonlinear problem, it may be impossible to guarantee the consistence between the updated state variables $\mathbf{r}$ and model parameters $\mathbf{m}$ without re-solving the nonlinear