Ensemble modeling of the two-dimensional stochastic confined groundwater flow through the evolution of the hydraulic head's probability density function

By

JOAQUIN MEZA ARAYA
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

M. Levent Kavvas, Chair

Ali Ercan

Boris Jeremic

Committee in Charge

2022

Abstract

Ensemble modeling of the two-dimensional stochastic confined groundwater flow through the evolution of the hydraulic head's probability density function

by

Joaquin Meza

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Davis

Levent Kavvas, Chair

Groundwater storage in aquifers has become a vital water source due to the water scarcity in the last years. However, aquifer systems are full of uncertainties, which inevitably propagate throughout the calculations, mainly reducing the reliability of the model output. This study develops a novel two-dimensional stochastic confined groundwater flow model. The proposed model is developed by linking the stochastic governing partial differential equations through their one-to-one correspondence to the nonlocal Lagrangian-Eulerian form of the Fokker-Planck equation (LEFPE). In the form of the LEFPE, the resulting deterministic governing equation describes the spatio-temporal evolution of the probability density function of the state variables in the confined groundwater flow process. This probability evolution is performed by one single numerical realization rather than requiring thousands of simulations in the Monte Carlo simulation. Consequently, the ensemble groundwater flow process's mean and standard deviation behavior can be modeled under uncertainty in the transmissivity field and recharge and/or pumping conditions. In addition, an appropriate numerical method for its solution is subsequently devised. Then, LEFPE's solution is presented, discussed, and

illustrated through numerical examples. Furthermore, they are compared against the results obtained through the Monte Carlo simulations to evaluate the performance of the LEFPE framework. Results suggest that the proposed model appropriately characterizes the ensemble behavior in confined groundwater systems under uncertainty in the transmissivity field.

Key words: Groundwater modelling; Uncertainty quantification; Stochastic partial differential equation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Groundwater stored in aquifers has become a vital water source in the face of emerging droughts in many world regions. Moreover, this water scarcity is expected to aggravate due to both global population and economic growth. Therefore, the value of groundwater will increase as water availability decreases in various regions of the world with climate change, making groundwater management a fundamental tool. Anderson and Woessner (2015) stated that Groundwater Models are a well-suited tool for predicting complex aquifer systems' behavior. Nevertheless, groundwater modeling is challenging since its predictive power predominantly depends on the input data representativity, commonly requiring large samples to estimate the hydraulic parameters.

Improving the representativity of the input data and, consequently, groundwater models' predictive power is complicated mainly for two reasons: first, there is always an incomplete understanding of the site conceptual model (Xia et al., 2019). Second, limited knowledge of the hydrogeological parameter values, typically due to the limited availability of suitable observations (Li et al., 2003; Yeh et al., 1992). Beyond, these uncertainties inevitably propagate throughout the model calculations, mainly reducing the reliability of the model output.

Groundwater systems are highly heterogeneous in space, resulting in significant hydraulic parameter value variations. In particular, hydraulic conductivity is one of the most sensitive to spatial variation (Lu and Zhang, 2005). Moreover, this parameter is the primary parameter affecting the output accuracy of groundwater flow and hydraulic transport models (Bakshevskaya and Pozdnyakov, 2013; Kitanidis, 1997; Zhu et al., 2016) since it controls both advective and dispersive transport (Neuman, 1990). Hydraulic conductivity depends on various factors such as the size of pores, geological structure, and connectivity. Furthermore, it is not easy to accurately estimate this parameter with commonly used methods (Wu and Zeng, 2013). In addition, aquifer properties are scale-dependent, and can vary over many orders of magnitude in typical aquifer systems (Dagan, 1986; Sudicky, 1986). Hence, subsurface heterogeneity's characterization needs to be incorporated into the modeling in order to be able to predict groundwater flow and contaminant transport in groundwater environments accurately.

To date, the uncertainty quantification in groundwater modeling has been carried out mainly by two different techniques, (1) Monte Carlo simulations (Freeze, 1975; Refsgaard et al., 2012; Tonkin and Doherty, 2009) and (2) perturbation methods (Connell, 1995; Li et al., 2003; Ma et al., 2009; Smith and Freeze, 1979; Xia et al., 2019). The importance of quantifying the uncertainty inherent to any groundwater system relies on evaluating the risks coming from the heterogeneity and the lack of information on design and management (Renard, 2007). Furthermore, stochastic approaches align with a typical decision-making agency's expectation that predictions should be accompanied by uncertainty measures that allow risk assessment (Rajaram, 2016).

Monte Carlo (MC) simulations repeatedly solve the deterministic governing equations

for a large number of equally likely realizations of the model parameters (e.g., hydraulic conductivity) to achieve multiple realizations. Then, the ensemble of solutions is further used to make statistical estimations. Furthermore, the MC approach is well known as the most robust approach for uncertainty evaluation, as well as the benchmark to validate other methods (Scharffenberg and Kavvas, 2011). While the MC method may offer a robust approach for estimating uncertainty in groundwater flow and transport, it is computationally demanding, and therefore, its application is restricted (Connell, 1995).

Alternatively, the regular perturbation method incorporates variability into the model using a different approach. This method decomposes the state variables into a mean plus a perturbation. By design, the regular perturbation has a zero mean and a variance equal to the original variable variance. Even though the regular perturbation method could offer significant savings in computation over MC methods (Townley and Wilson, 1985; T. C. Yeh, 1992), its accuracy is directly related to the magnitude of the process variance (Connell, 1995). Therefore, solutions by the regular perturbation approach may be poor approximations for highly heterogeneous ($\sigma^2 \gg 1$) aquifers (Gotovac et al., 2009). Moreover, using regular perturbation approaches results in a closure problem, where the equation for a specific moment requires information about higher moments' behavior (Keese, 2003). Hence, one can close the system of equations only by employing some ad-hoc assumptions.

Even though stochastic groundwater modeling has developed considerably in the last fifty years, these techniques are not standard tools in practice. Why the market has not adopted stochastic analysis is of substantial debate (Dagan, 2002; Gelhar, 1986; Renard, 2007; Rubin et al., 2018). The gap between theory and practice in stochastic modeling of groundwater is attributed to diverse factors, such as (1) economic constraints and lack of regulations (Rubin

et al., 2018), (2) the need to construct statistical models for heterogeneity fields from limited data (Dagan, 2002), (3) the assumptions and simplifications adopted in most theoretical analyses (Renard, 2007). However, all authors agree that (4) the excessive computational requirement of available stochastic numerical methods is one of the primary factors limiting these techniques' application.

Meanwhile, in order to avoid the drawbacks of Monte Carlo Simulations and regular perturbation methods, this study proposes a novel methodology to solve the expected system behavior in a single simulation. This proposed methodology upscales the governing stochastic differential equations from a point-scale (at which they are valid) to a field scale. Thus, the conservation equations describing the groundwater flow in confined aquifers are consistent with the scale of the grid areas over which they describe the hydrologic process. Furthermore, through this general framework, not only the mean and the variance of the targeted state random variable/function can be estimated, but also the probability density function (PDF) of the process, which evolves in space and time.

Ensemble averaging has been a popular approach in hydrology to upscale both linear (Gelhar and Axness, 1983; Rubin and Dagan, 1989; Wood and Kavvas, 1999) and non-linear (Dogrul et al., 1998; Mantoglou and Gelhar, 1987; Tayfur and Kavvas, 1994) hydrologic processes. To upscale the stochastic governing equations, they are averaged to become deterministic. Thus, statistical descriptions are used to represent the values of the stochastic parameters. However, most studies performing the ensemble averaging technique used the regular perturbation method, which only works for small fluctuations in the dependent variables.

A general ensemble average conservation equation was developed by Kavvas (2003) to upscale the point-scale conservation equations to the scale of computational grid areas. Thus, the hydrologic processes' probabilistic and mean behavior can be estimated. To carry out this task, the one-to-one correspondence between the governing equation under uncertainty (that represents a hydrologic process) and a mixed Eulerian-Lagrangian form of the Fokker-Planck equation (LEFPE) to the second-order was employed (Kavvas, 2003). The resulting "master key" equation allows dealing with uncertainties in the parameters and the forcing terms in their corresponding point-scale governing equations. Therefore, employing this approach, it is possible to estimate the probability density function's (PDF) time-space evolution for any nonlinear or linear hydrologic process.

Within this framework, the proposed methodology has been successfully applied to many hydrologic processes. Kim et al., (2005) used the methodology to model the one-dimensional root-water uptake under uncertainty in the saturated soil vadose zone. Cayar and Kavvas (2009) modeled the effect of uncertainty in the hydraulic conductivity for the one-dimensional horizontal confined groundwater flow. They transformed the governing PDE into an ordinary differential equation using the Boltzmann transformation and Lie Group theory. Ercan and Kavvas (2012) used this methodology to explore variability in the channel properties and lateral flow conditions to upscale the kinematic open-channel flow governing equation. Tu et al. (2019, 2020) applied the methodology for describing the ensemble behavior for the 1D and 2D solute transport in open channel flow under uncertain flow and solute loading conditions. Even though this stochastic framework was initially developed to tackle uncertainties for hydrologic processes, it has also been used in the geomechanics field. Jeremić et al. (2007) and Sett et al., (2007) used the "master key" equation for the 1-D elastic-plastic constitutive rate equations to quantify the uncertainty in material parameters. Later, Kara-

piperis et al. (2016) applied this stochastic upscaling framework together with stochastic Garlekin techniques to solve 1-D elastoplastic boundary value problems with non-Gaussian parametric uncertainty.

Even though MC and Perturbation methods have been widely used to tackle groundwater flow in confined aquifers, they present relevant drawbacks regarding their computational time and their coefficient of variation, respectively. Thus, the proposed method has the goal of overcoming these mentioned issues. To the authors' knowledge, this is the first study to attempt to estimate the ensemble behavior in a confined aquifer based on the cumulant expansion of the stochastic governing partial differential equation. This novel approach estimates the ensemble mean and variance in one shot, instead of multiple simulations, by considering the stochasticity in multiple parameters.

In this study, a stochastic model for confined aquifer flow is proposed, motivated by the probabilistic responses of aquifers to the influence of spatial uncertainties. This general framework was carried out using the above-discussed methodology proposed by Kavvas (2003). Thus, the effect of uncertainties in the transmissivity field and the source term are accounted for. Consequently, the ensemble average for this dynamical system is estimated, determining its probabilistic behavior. In the form of a deterministic partial differential equation, the resulting expression illustrates the complete description of the spatio-temporal evolution of the hydraulic head's PDF in confined aquifers in one single simulation. Furthermore, the derived LEFPE corresponding to the confined groundwater flow process is described by employing illustrative examples.

Chapter 2

Governing equations for unsteady confined groundwater flow

This chapter describes the governing equation which controls the subsurface flow in confined aquifers, including the assumptions used for its derivation. Under the specified initial and boundary conditions, this governing equation at the point scale can be solved. Later on, employing the method of characteristics, the governing PDE is recast into a system of ordinary differential equations (ODEs). This ODE system will be crucial in the next section to upscale the governing groundwater flow equation from the point scale to the field scale and then obtain its ensemble behavior.

At the end of this chapter, a numerical discretization of the resulting Groundwater flow equation is presented, followed by the details of the numerical scheme used for solving this problem.

2.1 Derivation of the point-scale confined groundwater flow equation

The groundwater flow equation in a confined aquifer is developed in this section to understand the assumptions considered when this mathematical model is used. Hydraulic groundwater flow theory in a fully saturated geologic porous media is based on the mass balance principle combined with Darcy's equation. Therefore, an elemental control volume in a porous medium is considered (Figure 2.1) to define the mass that enters or leaves this control volume. Thus, the mass balance principle can be used over this elemental control volume.

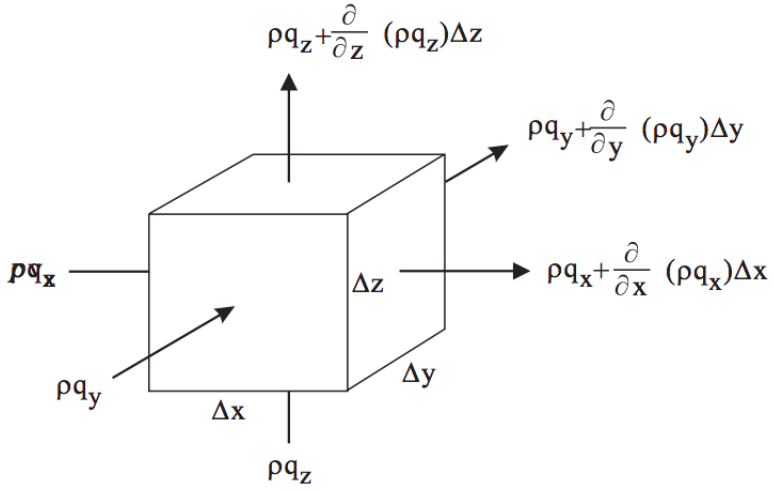


Figure 2.1: Fluid mass balance in a elementary volume with sides Δx, Δy and Δz.

Therefore, the rate of change of fluid mass inside the control volume must equal the difference in the fluid mass flux into and out of the control volume. In mathematical terms, this principle could be expressed as follows.

$$\frac{\partial(\rho q_i)}{\partial x_i} = \frac{\partial m}{\partial t} \tag{2.1}$$

Where,

q_i $[LT^{-1}]$ are fluxes in the i directions.

m=$\theta\rho\Delta x\Delta y\Delta z$ $[M]$ is the fluid mass inside the control volume at any point in time.

θ [1] represents the aquifer porosity.

ρ $[ML^{-3}]$ is the water density.

Then, by using the definition of fluid mass inside the control volume (m), the right side of Equation (2.1) can be written as

$$\frac{\partial m}{\partial t} = \frac{\partial \theta \rho}{\partial t} = \rho \frac{\partial \theta}{\partial t} + \theta \frac{\partial \rho}{\partial t} = \left(\rho \frac{\partial \theta}{\partial p} + \theta \frac{\partial \rho}{\partial p} \right) \frac{\partial p}{\partial t} \tag{2.2}$$

Where p $[ML^{-1}T^{-2}]$ is the water pressure. It is then assumed that the grains' volume does not change, which means that the porosity must change in the elastic range instead. In mathematical terms, this assumption is expressed as follows.

$$\alpha = -\frac{1}{V_T} \frac{\partial V_T}{\partial \sigma} \approx -\frac{1}{V_T} \frac{\partial V_V}{\partial \sigma} \tag{2.3}$$

Where,

σ is the static stress which in addition with the pressure must be balanced by the total stress $[ML^{-1}T^{-2}]$.

V_T represents the total volume, which is equal to $V_S + V_V$ $[L^3]$.

V_S and V_V are grain an pore volume, respectively $[L^3]$.

α is the amount of water released from the vertical compression per unit bulk volume per unit change in intergranular stress.

Then, considering the following equality $V_V = \theta \cdot V_T$, implies that

$$\frac{\partial V_S}{\partial \sigma} = 0 \qquad \Rightarrow \frac{1}{V_T}\frac{\partial V_V}{\partial \sigma} = \frac{1}{1-\theta}\frac{\partial \theta}{\partial \sigma} \qquad \Rightarrow \alpha = \frac{1}{1-\theta}\frac{\partial \theta}{\partial \sigma} = \frac{1}{1-\theta}\frac{\partial \theta}{\partial p} \tag{2.4}$$

On the other hand, it is assumed that the water in the aquifer is slightly compressible. Thus, the coefficient of compressibility of water, β (1) is defined as

$$\beta = -\frac{1}{V_V}\frac{\partial V_V}{\partial p} = -\frac{1}{m/\rho}\frac{\partial m/\rho}{\partial \rho} = -\rho\frac{-\frac{\partial \rho}{\partial p}}{\rho^2} = \frac{1}{\rho}\frac{\partial \rho}{\partial p} \tag{2.5}$$

Now, by using equation (2.4) and (2.5) into equation (2.2), we get

$$\frac{\partial(\rho\theta)}{\partial t} = \rho(\alpha(1-\theta)+\beta\theta)\frac{\partial p}{\partial t} = g\rho^2(\alpha(1-\theta)+\beta\theta)\frac{\partial h}{\partial t} \tag{2.6}$$

Where g is gravity $[L^1T^{-2}]$ and h is the hydraulic head $[L]$. In addition to the assumptions previously mentioned, the horizontal stresses are neglected. Thus, the right side of the Equation (2.1) can be rewritten as.

$$\frac{\partial m}{\partial t} = S_s\frac{\partial h}{\partial t} \tag{2.7}$$

Where $S_s = g\rho^2(\alpha(1-\theta)+\beta\theta)$ is defined as the specific storage $[L^{-1}]$. Moreover, in groundwater hydrology, Darcy's law states that the flux of water q $[LT^{-1}]$ through a unit surface $[L^2]$ is proportional to the gradient of the hydraulic head $\frac{\partial h}{\partial x}$ and a physical parameter, K $[LT^{-1}]$, termed hydraulic conductivity. The hydraulic conductivity depends on the fluid and porous medium. Mathematically, Darcy's law can be expressed as

$$q_i = -K_{ij}\frac{\partial h}{\partial x_j} \tag{2.8}$$

This expression, known as Darcy's law, can be used to manipulate the left-hand side of Equation (2.1). Assuming that spatial changes in density are negligible (i.e., terms of the form $\rho\frac{\partial q_i}{\partial x_i}$ are much more significant than terms of the form $q_i\ \frac{\partial \rho}{\partial x}$. Additionally, incorporating a source or sink term Q_s $[L^3/(TL^3)]$, which could be due to pumping/injection of a volume flux $[L^3/T]$ per volume $[L^3]$, we can obtain the following Equation.

$$\frac{\partial}{\partial x_i}\left(K_i(\overline{x})\frac{\partial h(\overline{x},t)}{\partial x_i}\right) + Q_s(\overline{x}) = S_S(\overline{x})\frac{\partial h(\overline{x},t)}{\partial t} \tag{2.9}$$

Subject to boundary conditions on $\Gamma = \Gamma_N \cup \Gamma_D$, which are defined by

$$h(\overline{x}) = h_D(\overline{x})\ on\ \Gamma_D \tag{2.10}$$

$$\left(K_i(\overline{x},t)\frac{\partial h(\overline{x},t)}{\partial x_i}\right)\cdot n = q_N(\overline{x})\ on\ \Gamma_N \tag{2.11}$$

where $\overline{x} = \{x, y, z\}$ is the position vector $[L]$, t is time [T], $h_D(\overline{x})$ is the prescribed head at boundary Γ_D, $q_N(\overline{x})$ is the flux at boundary Γ_N, and n is a unit vector normal to Γ_N. Equation (2.9) is the governing equation for transient flow through an anisotropic and heterogeneous porous groundwater system with the hydraulic conductivity tensor oriented along the principal directions.

Then, it is possible to model the three-dimensional aquifer's flow as a horizontal two-

dimensional flow by neglecting vertical fluxes and averaging properties over the aquifer thickness. These simplifications are called the Dupuit-Forchheimer approximation. Under this assumption, the equation for two-dimensional horizontal flow in a confined aquifer can be expressed as (Bear and Verruijt, 1987)

$$\frac{\partial}{\partial x_i}\left(T_i(\underline{x})\frac{\partial h(\underline{x},t)}{\partial x_i}\right) + Q_s(\underline{x}) = S(\underline{x})\frac{\partial h(\underline{x},t)}{\partial t} \tag{2.12}$$

where $\underline{x} = \{x, y\}$ is the two-dimensional position vector $[L]$, $S(\underline{x},t) = S_S(\underline{x},t) \cdot b(\underline{x})$ is the storativity [1], $b(\underline{x})$ is the confined aquifer thickness $[L]$ and $T_i(\underline{x})$ is the transmissivity $[L^2T^{-1}]$, which is defined by

$$T_i(\underline{x}) = \int_0^{b(\underline{x})} K_i(\overline{x},t)dz \tag{2.13}$$

2.2 Transforming the point-scale groundwater flow equation into the characteristic form

The previous section derived the governing equation that controls the groundwater flow (equation 2.12). However, to obtain the ensemble behavior through the methodology proposed, the governing equation needs to be recast into a system of ordinary differential equations. For this analysis, the water table height h is considered a function of space and time. Then, by expanding the spatial derivates, equation (2.12) can be expressed as

$$\underbrace{\frac{\partial T_i(\underline{x})}{\partial x_i}\frac{\partial h(\underline{x},t)}{\partial x_i}}_{Convective\ terms} + \underbrace{T_i(\underline{x})\frac{\partial^2 h(\underline{x},t)}{\partial x_i^2}}_{Diffusive\ terms,\ DT(\underline{x},t)} + Q_s(\underline{x}) = S(\underline{x})\frac{\partial h(\underline{x},t)}{\partial t} \tag{2.14}$$

It assumed that at the regional scale, the stochasticity of the random groundwater field in

equation (2.14) is mainly due to the stochasticity in the convective term, while the diffusive terms $DT(\underline{x},t)$, may contribute only to the mean behavior (Kavvas and Karakas, 1996; Tu et al., 2019, 2020a). This assumption will be assessed later by comparing the results from LEFPE against Monte Carlo solutions. Thus, equation (2.14) may be recast into the form

$$\frac{\partial h(\underline{x},t)}{\partial t} - \frac{1}{S(\underline{x})}\frac{\partial T_i(\underline{x})}{\partial x_i}\frac{\partial h(\underline{x},t)}{\partial x_i} = \frac{1}{S(\underline{x})}\left[DT(\underline{x},t) + Q_s(\underline{x})\right] \tag{2.15}$$

Through the method of characteristics, the governing PDE (2.15) may be transformed into a system of stochastic ordinary differential equations (Pletcher et al., 1997). By following this method, the characteristic equations for unsteady confined groundwater flow can be expressed as

$$\eta_1 = \frac{\partial x(t)}{\partial t} = -\frac{1}{S(\underline{x})}\frac{\partial T_x(\underline{x})}{\partial x}\frac{\partial h(\underline{x},t)}{\partial x} \tag{2.16}$$

$$\eta_2 = \frac{\partial y(t)}{\partial t} = -\frac{1}{S(\underline{x})}\frac{\partial T_y(\underline{x})}{\partial y}\frac{\partial h(\underline{x},t)}{\partial x} \tag{2.17}$$

$$\eta_3 = \frac{\partial h(\underline{x},t)}{\partial t} = \frac{1}{S(\underline{x})}\left[DT(\underline{x},t) + Q_s(\underline{x})\right] \tag{2.18}$$

Equations (2.16) and (2.17) are known as the characteristic equations, which represent two velocity expressions that can be used to determine the evolution of the stochastic characteristic path $\underline{x}(t)$. On the other hand, equation (2.18) is called the compatibility equation, which describes the process behavior of the state variable $h(\underline{x},t)$ along $\underline{x}(t)$. In the $\underline{x}$-t domain, the characteristic equations will represent curves along $\underline{x}(t)$ where the information propagates through the solution domain. For convenience, the Ordinary Differential Equation (ODE) system may be expressed.

$$\frac{\partial \boldsymbol{H}(\underline{x}, t)}{\partial t} = \boldsymbol{\eta}(H, A, f) \tag{2.19}$$

with an initial condition

$$\boldsymbol{H}(\underline{x}, t = 0) = \boldsymbol{H_0} \tag{2.20}$$

where $\boldsymbol{H} = \{x, y, h\}^T$ is the vector of all state variables of the hydrologic system of equations, A is the tensor of parameters, f is the vector of forcing functions, and $\boldsymbol{\eta} = \{\eta_1, \eta_2, \eta_3\}$ is the governing function vector.

2.3 Numerical Solution of the deterministic confined groundwater flow equation

A form of the diffusion equation describes the groundwater flow governing equation for a confined aquifer (Equation 2.9). This linear, second-order, and parabolic PDE has a limited number of analytical solutions. Moreover, analytical solutions for real groundwater systems are even more restricted. The main reason for the inadequacy of the analytical solutions is the heterogeneity that characterizes this kind of natural system (Karatzas, 2017). Therefore, numerical solutions are widely used to handle more complicated real aquifer systems (Baalousha, 2008).

Finite Differences Method (FDM) and Finite Element Method (FEM) are the most popular numerical techniques used to solve the governing equation (2.12). Even though some authors support one or the other method, the choice between FDM and FEM relies mainly on the user's preference (Karatzas, 2017; Simpson and Clement, 2003). According to Gray

(1984) , each method has its own unique features that could be desirable for a specific application. This study uses a generalized control-volume finite-difference (CVFD) method. This method was chosen since it is the same approach used by the most widely used groundwater flow model in consultancy and for research proposes, MODFLOW (Hughes et al., 2017).

2.3.1 Formulation and Solution of the Control-Volume Finite-Difference Equation

Flows Between Cells

In this numerical scheme, the flows between cells are estimated by a discrete form of Darcy's Law. Thus, the flow into cell n from cell m is given by

$$Q_{n,m} = \left(\frac{\overline{T}_{n,m} \Delta w_{n,m}}{L_{n,m} + L_{m,n}} \right) (h_m - h_n) = C_{n,m}(h_m - h_n) \tag{2.21}$$

where $Q_{n,m}$ is the flow rate into cell n from cell m (Figure 2.2.a), $\Delta w_{n,m}$ is the width of the face through which flow occurs, h_i is the head at node i, $L_{n,m}$ is the distance from the center of cell n to its shared face with cell m, $C_{n,m}$ is the conductance between cells n and m and $\overline{T}_{n,m}$ is the distance weighted harmonic mean of transmissivity of two half cells, which is given by

$$\overline{T}_{n,m} = \frac{L_{n,m} + L_{m,n}}{\frac{L_{n,m}}{T_{n,m}} + \frac{L_{m,n}}{T_{m,n}}} \qquad T_{n,m} = K_{n,m} \Delta v_{n,m} \tag{2.22}$$

In simple words, $T_{n,m}$ represents the transmisivity in cell n in the direction of cell m, $K_{n,m}$ is the hydraulic conductivity of cell n in the direction of cell m and $\Delta v_{n,m}$ is the height of the face through which flow occurs (Figure 2.2.b).

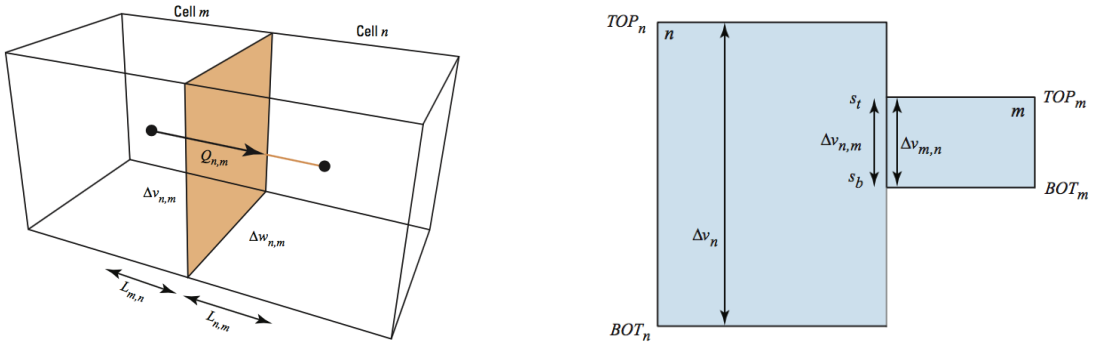


Figure 2.2: Schematic representation of fluxes between cells. (a) Diagram showing flows into cell n from cell m. (b) A cross-section diagram shows cell n connected to vertically staggered cells, including cell m. *source: MODFLOW 6 Documentation*

Water Balance on a Cell

To solve the governing groundwater flow equation (equation 2.9), the CVFD scheme is used. To perform this task, equation (2.9) is solved in a discrete way, in which differential control volumes are approximated by cells with a finite volume.

In order to derive the CVFD discrete set of equations, a water balance on each cell must be done. The balance considers groundwater flows between cells, flows from and to any sources, sinks, and storage. Then, by considering a constant density of groundwater, the continuity equation for the cell n is given by

$$\sum_{m \in \eta_n} Q_{n,m} + Q_{n,s} = Q_{storage} \tag{2.23}$$

where η_n are the number of cells connected to cell n, $Q_{n,m}$ is the flow rate into cell n from cell m, $Q_{n,s}$ is the flow rate of sources and sinks into cell n, and $Q_{storage}$ is the change

in volume of water stored in cell n, which can be discretely approximated as follows

$$Q_{storage} = S_{Sn} V_n \frac{\Delta h_n}{\Delta t} \tag{2.24}$$

where S_{Sn} is the specific storage of cell n, V_n is the volume of cell n, and Δh_n is the change in the head in cell n over a time interval Δt.

Temporal Discretization

The temporal rate of change of the hydraulic head is estimated discretely. Even though it is only a first-order approximation, a backward-difference approach was used to perform this estimation. The scheme selected was chosen since it is always numerically stable, which means that numerical errors do not grow and propagate through time (Hughes et al., 2017). Thus, equation 2.24 can be expressed as

$$Q_{storage} = S_{Sn} V_n \frac{h_n^{t+1} - h_n^t}{\Delta t} \tag{2.25}$$

Formulation of Control-Volume Finite-Difference equations for Solution

Combining equations (2.21), (2.22), (2.23) and (2.25), the generalized CVFD flow equation for cell n is obtained as follow

$$\sum_{m \in \eta_n} C_{n,m}(h_m^k - h_n^k) + Q_{n,s} = S_{Sn} V_n \frac{h_n^k - h_n^{k-1}}{\Delta t} \tag{2.26}$$

Thus, an equation of this type (equation 2.26) can be written for each cell in the grid where h_n is unknown. Therefore, a system of N equations and N unknowns is obtained. Then, the system can be solved simultaneously, subject to appropriate initial and boundary

conditions.

Rewriting the equation (2.26) by grouping on the left-hand side of the equation all terms containing unknown heads at the current time step, and on the right, all terms that are independent of the unknown heads at the current time step. The resulting equation is expressed as follows

$$\left(\sum_{m\in\eta_n} C_{n,m} + \frac{S_n V_n}{\Delta t}\right) h_n^k - \sum_{m\in\eta_n} C_{n,m} h_m^k = Q_{n,s} + \frac{S_{Sn} V_n}{\Delta t} h_n^{k-1} \tag{2.27}$$

Finally, the entire system of equations in the form of equation (2.27), which consist of one equation for the head of each cell in the grid, can be written in a matrix form (**Ah=b**). This system of equations implies that the matrix (**A**) must be inverted to obtain the hydraulic head vector for future time steps.

Chapter 3

Ensemble-averaged equations for the stochastic unsteady confined groundwater flow

In this section, we introduce a novel methodology to solve the stochastic confined groundwater flow equation and a well-suited numerical scheme to solve the resulting partial differential equation (PDE) in the form of a LEFPE. The proposed methodology will allow obtaining the statistical properties of the hydraulic head for a confined groundwater system in only one shot instead of multiple simulations through the Monte Carlo approach. Thus, the hydraulic head's time-space evolutionary probability density function (PDF) will be obtained. The following derivation assumes uncertainty in the transmissivity random field and the source term. Nevertheless, stochasticity in the storativity and the boundary conditions could be also incorporated following similar steps.

3.1 Fokker-Planck equation overview

The classical Fokker-Planck equation (FPE) is a second-order differential equation that describes the statistical behavior of the system under analysis. Meanwhile, solving stochastic differential equations gives the solution of the system's state variable; the FPE is solved for the probability density function (PDF) of the state variables. Therefore, this equation governs the time evolution of the state variables' PDF. For a multivariate process, the FPE can be approximated locally by an Itô process driven by the standard Wiener process as follows

$$d\mathbf{X}_t = \boldsymbol{\mu}(\mathbf{X}_t, t)\, dt + \boldsymbol{\sigma}(\mathbf{X}_t, t)\, d\mathbf{W}_t, \tag{3.1}$$

Where $\mathbf{X}_t$ and $\boldsymbol{\mu}(\mathbf{X}_t, t)$ are $n-dimensional$ random vectors, $\boldsymbol{\sigma}(\mathbf{X}_t, t)$ is an $n \times m$ matrix and $\mathbf{W}_t$ is an $m-dimensional$ Wiener process. By considering that $\boldsymbol{\mu}(\mathbf{X}_t, t)$ and $\boldsymbol{\sigma}(\mathbf{X}_t, t)$ satisfy a global Lipschitz and linear growth conditions, the corresponding FPE to the Itô stochastic differential equation (Equation 3.1) can be expresed as

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i} \left[\mu_i(\mathbf{x}, t) p(\mathbf{x}, t)\right] + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i\, \partial x_j} \left[D_{ij}(\mathbf{x}, t) p(\mathbf{x}, t)\right] \frac{\partial p(\mathbf{x}, t)}{\partial t} \tag{3.2}$$

with drift vector $\boldsymbol{\mu}\,(\mu_1, ..., \mu_N)$ and diffusion tensor represented by

$$D_{ij}(\mathbf{x}, t) = \frac{1}{2} \sum_{k=1}^{M} \sigma_{ik}(\mathbf{x}, t) \sigma_{jk}(\mathbf{x}, t). \tag{3.3}$$

The FPE is intimately connected to the corresponding Ito stochastic differential equation,

being this correspondence unique (Gardiner, 2009). However, solving the FPE presents advantages over their corresponding Ito SPE. The Ito SDE has stochastic parameters in the first place, while in the FPE, the drift vector and the diffusion tensor are deterministic. Therefore, the resulting FPE is deterministic too. Secondly, the FPE is a linear equation, which is more straightforward to solve than non-linear equations. Finally, the solution of the FPE results in the state variables' PDF, allowing us to estimate any statistical moment and provide the ensemble behavior in one shot.

3.2 Derivation of the Fokker-Planck equation for confined groundwater flow

Kavvas (2003) described a system of point-scale conservation equations, which can be expressed in mathematical terms for a dynamical system as

$$\frac{\partial \boldsymbol{H}(\overline{\boldsymbol{x}}, t)}{\partial t} = \boldsymbol{\eta}(\boldsymbol{H}, \boldsymbol{A}, \boldsymbol{f}; \overline{\boldsymbol{x}}, t) \tag{3.4}$$

Subject to the initial condition

$$\boldsymbol{H}(\overline{\boldsymbol{x}}, 0) = \boldsymbol{H_0} \tag{3.5}$$

where $\boldsymbol{H}(\overline{\boldsymbol{x}}, t)$ is the vector of all state variables of the hydrologic system of equations, $\boldsymbol{A}(\overline{\boldsymbol{x}}, t)$ is the tensor of parameters; $\boldsymbol{f}(\overline{\boldsymbol{x}}, t)$ is the vector of forcing functions; $\overline{\boldsymbol{x}}$ is the vector of spatial locations, and t is the time.

From a different point of view, $\boldsymbol{H}(\overline{\boldsymbol{x}}, t)$ could be interpreted as a point in the n-dimensional $\boldsymbol{H}$-space. Therefore, equation (3.4) would determine each point's velocity in the mentioned

space. Conceptually, this could be interpreted as the path describing equation (3.4)'s solution for a given deterministic initial condition $\boldsymbol{H_0}$. Nonetheless, the initial condition may be described by a density $\rho(\boldsymbol{H}, t=0)$ in the $\boldsymbol{H}-phase$ space to describe a cloud of deterministic initial conditions. Thus, the phase density $\rho(\boldsymbol{H}, t)$ evolves according to a continuity equation, representing the conservation of all these points in the $\boldsymbol{H}$-space.

Then, by applying a second-order cumulant expansion, Kavvas (2003) developed the general Lagrangian-Eulerian Fokker-Planck Equation (LEFPE) form of the dynamical system established in equation (3.4) to the exact second-order as follows

$$\frac{\partial P[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t]}{\partial t} = -\frac{\partial}{\partial H_j}\Bigg\{P[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t]\Bigg(\left\langle \eta_j\left[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t), \boldsymbol{A}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t), \boldsymbol{f}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t)\right]\right\rangle \tag{3.6}$$

$$+\int_0^t Cov_0\left[\frac{\partial \eta_j\left[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t), \boldsymbol{A}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t), \boldsymbol{f}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t)\right]}{\partial H_i}; \eta_j\left[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}, t-s), \boldsymbol{A}(\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}, t-s), \boldsymbol{f}(\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}, t-s)\right]\right] ds\Bigg)\Bigg\}$$

$$+\frac{1}{2}\frac{\partial^2}{\partial H_j \partial H_i}\Bigg\{2P[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t]\int_0^t Cov_0\left[\begin{matrix}\eta_j\left[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t), \boldsymbol{A}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t), \boldsymbol{f}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t)\right]; \\ \eta_i\left[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}, t-s), \boldsymbol{A}(\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}, t-s), \boldsymbol{f}(\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}, t-s)\right]\end{matrix}\right] ds\Bigg)\Bigg\}$$

where $P[\boldsymbol{H}(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t]$ is the probability density of the vector of state variables at the location $\overline{\boldsymbol{x}}_{\boldsymbol{t}}$ and at time t, the operator $\langle\cdot\rangle$ is the ensemble average operator, s is a time displacement and $Cov_0[\cdot]$ is the time-ordered covariance function (Van Kampen, 1974), which can be expressed as

$$Cov_0\left[\eta_i(\boldsymbol{x}, t_1); \eta_j(\boldsymbol{x}, t_2)\right] = \left\langle \eta_i(\boldsymbol{x}, t_1); \eta_j(\boldsymbol{x}, t_2)\right\rangle - \left\langle \eta_i(\boldsymbol{x}, t_1)\right\rangle; \left\langle \eta_j(\boldsymbol{x}, t_2)\right\rangle \tag{3.7}$$

Equation (3.6) is a mixed Eulerian-Lagrangian partial differential equation. Note that

while the real space location $\overline{\boldsymbol{x}}_{\boldsymbol{t}}$ is known, the lagrangian location $\overline{\boldsymbol{x}}_{\boldsymbol{t-s}}$ is unknown. However, the lagrangian location can be calculated from the known location by using the Lie operator (Kavvas and Karakas, 1996), as follows

$$\overline{\boldsymbol{x}}_{\boldsymbol{t-s}} = \overleftarrow{exp}\left[-\int_{t-s}^{t} \langle v_l(\boldsymbol{x}_{\boldsymbol{\tau}}, \tau)\rangle \, d\tau\right] \overline{\boldsymbol{x}}_{\boldsymbol{t}} \tag{3.8}$$

where $\overleftarrow{exp}$ is the time-ordered exponential, and v_l is determined from the characteristics curve equations corresponding to the system's governing equation. However, Kavvas and Karakas (1996) proposed a first-order approximation for equation (3.8), which is easier to estimate. This approximation can be expressed as

$$\overline{\boldsymbol{x}}_{\boldsymbol{t-s}} = \overline{\boldsymbol{x}}_{\boldsymbol{t}} - \int_{t-s}^{t} \langle v_l(\boldsymbol{x}_{\boldsymbol{\tau}}, \tau)\rangle \, d\tau \tag{3.9}$$

By solving the LEFPE (equation 3.6), the quantitative description of the state variables' probabilistic behavior of a system (e.g., hydraulic head for an aquifer system) can be obtained under appropriate initial and boundary conditions. In order to estimate the state variables' ensemble behavior through the LEFPE for a confined groundwater system, the general method detailed in Equation (3.6) is applied. Thus, this study will solve the probability density function of the hydraulic head $h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t)$, which can be written as shown in Equation 3.10

$$
\begin{aligned}
\frac{\partial P(x_t, y_t, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)}{\partial t} = &- \frac{\partial}{\partial x}\Bigg[P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\Bigg\{\langle \eta_{1,t}\rangle + \int_0^t Cov_0\left[\frac{\partial \eta_{1,t}}{\partial x}; \eta_{1,t-s}\right] ds \\
&+ \int_0^t Cov_0\left[\frac{\partial \eta_{1,t}}{\partial y}; \eta_{2,t-s}\right] ds + \int_0^t Cov_0\left[\frac{\partial \eta_{1,t}}{\partial h}; \eta_{3,t-s}\right] ds\Bigg\}\Bigg] \\
&- \frac{\partial}{\partial y}\Bigg[P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\Bigg\{\langle \eta_{2,t}\rangle + \int_0^t Cov_0\left[\frac{\partial \eta_{2,t}}{\partial y}; \eta_{2,t-s}\right] ds \\
&+ \int_0^t Cov_0\left[\frac{\partial \eta_{2,t}}{\partial x}; \eta_{1,t-s}\right] ds + \int_0^t Cov_0\left[\frac{\partial \eta_{2,t}}{\partial h}; \eta_{3,t-s}\right] ds\Bigg\}\Bigg] \\
&- \frac{\partial}{\partial h}\Bigg[P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\Bigg\{\langle \eta_{3,t}\rangle + \int_0^t Cov_0\left[\frac{\partial \eta_{3,t}}{\partial h}; \eta_{3,t-s}\right] ds \\
&+ \int_0^t Cov_0\left[\frac{\partial \eta_{3,t}}{\partial x}; \eta_{1,t-s}\right] ds + \int_0^t Cov_0\left[\frac{\partial \eta_{3,t}}{\partial y}; \eta_{2,t-s}\right] ds\Bigg\}\Bigg] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{1,t}; \eta_{1,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{2,t}; \eta_{2,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial h^2}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{3,t}; \eta_{3,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial x \partial y}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{1,t}; \eta_{2,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial x \partial h}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{1,t}; \eta_{3,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial y \partial x}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{2,t}; \eta_{1,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial y \partial h}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{2,t}; \eta_{3,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial h \partial x}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{3,t}; \eta_{1,t-s}\right] ds\right] \\
&+ \frac{1}{2}\frac{\partial^2}{\partial h \partial y}\left[2P(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, h(\overline{\boldsymbol{x}}_{\boldsymbol{t}}), t)\int_0^t Cov_0\left[\eta_{3,t}; \eta_{2,t-s}\right] ds\right]
\end{aligned}
\tag{3.10}
$$

Equation (3.10) has the form of an advection-diffusion equation. The first three terms are advection terms, while the nine last expressions are the diffusion terms. Liang and Kavvas (2008) define the expected values of eta's functions $\langle \eta_{i,t} \rangle$ as the mean advection coefficients and the integrals of the ordered covariance eta's functions η as the advection correction terms.

Even though equation (3.10) is a linear and deterministic PDE, it still presents substantial challenges for its solution. Therefore, some simplifications without losing the prime physical characteristics are examined. One of the complications arises in estimating the covariance terms in the advection coefficients. However, the advection correction terms can be considered negligible compared to the mean advection terms. This assumption is based on the expected η_i magnitudes, which are much larger than those of the integral terms (Kavvas and Wu, 2002; Liang and Kavvas, 2008; Tu et al., 2019). Mathematically, this can be expressed as follows

$$\langle \eta_{1,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{1,t}}{\partial x}; \eta_{1,t-s} \right] ds \qquad \langle \eta_{1,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{1,t}}{\partial y}; \eta_{2,t-s} \right] ds$$

$$\langle \eta_{1,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{1,t}}{\partial h}; \eta_{3,t-s} \right] ds \qquad \langle \eta_{2,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{2,t}}{\partial y}; \eta_{2,t-s} \right] ds$$

$$\langle \eta_{2,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{2,t}}{\partial x}; \eta_{1,t-s} \right] ds \qquad \langle \eta_{2,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{2,t}}{\partial h}; \eta_{3,t-s} \right] ds$$

$$\langle \eta_{3,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{3,t}}{\partial h}; \eta_{3,t-s} \right] ds \qquad \langle \eta_{3,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{3,t}}{\partial x}; \eta_{1,t-s} \right] ds$$

$$\langle \eta_{3,t} \rangle >> \int_0^t Cov_0 \left[\frac{\partial \eta_{3,t}}{\partial y}; \eta_{2,t-s} \right] ds$$

Regarding diffusion terms, they will be classified into two categories. The first is the terms involving the ordered covariance between the same functions. The second one is the terms that consider the ordered covariance between two different functions or also called cross-covariance dispersion terms. Liang and Kavvas (2008) showed that the autocovariance of the η function of one state variable is considerably larger in magnitude than the covariance between any two different η functions. Therefore, all cross-covariance terms are neglected from equation (3.10) to simplify calculations. However, this assumption could break down depending on the behavior of the η functions. For η functions with similar periodicity and close frequencies, the cross-covariance terms could become more comparable in magnitude to the autocovariance values, contrary to the previous simplification (Dib and Kavvas, 2018). The resulting nonlocal LEFPE after simplifications previously mentioned is shown in equation (3.12) below

$$
\begin{aligned}
\frac{\partial P(h,x,y;t)}{\partial t} =& \frac{1}{S(x,y)}\frac{\partial}{\partial x}\left[P(h,x,y;t)\cdot\left\langle\frac{\partial T_x(\overline{\boldsymbol{x}}_{\boldsymbol{t}})}{\partial x}\right\rangle\right]+ \\
& \frac{1}{S(x,y)}\frac{\partial}{\partial y}\left[P(h,x,y;t)\cdot\left\langle\frac{\partial T_y(\overline{\boldsymbol{x}}_{\boldsymbol{t}})}{\partial y(t)}\right\rangle\right] \\
& -\frac{1}{S(x,y)}\frac{\partial}{\partial h}\left[P(h,x,y;t)\left\{\left[\ DT(\langle h(x,y;t)\rangle)+\langle G(x,y;t)\rangle\right]\right\}\right] \\
& +\frac{1}{S^2(x,y)}\frac{\partial^2}{\partial x^2}\left[P(h,x,y;t)\cdot\int_0^t Cov_0\left[\frac{\partial T_x(h,x,y;t)}{\partial x};\frac{\partial T_x(h,x,y;t-s)}{\partial x}\right]ds\right] \\
& +\frac{1}{S^2(x,y)}\frac{\partial^2}{\partial y^2}\left[P(h,x,y;t)\cdot\int_0^t Cov_0\left[\frac{\partial T_y(h,x,y;t)}{\partial y};\frac{\partial T_y(h,x,y;t-s)}{\partial y}\right]ds\right] \\
& +\frac{1}{S^2(x,y)}\frac{\partial^2}{\partial h^2}\left[P(h,x,y;t)\cdot\int_0^t Cov_0\big[G(h,x,y;t);G(h,x,y;t-s)\big]ds\right]
\end{aligned}
\tag{3.11}
$$

In equation (3.11), transmissivity and the source terms are treated as random functions, while storativity is assumed deterministic. Accordingly, storativity was moved outside of the expectations and ordered covariance terms. This assumption is in agreement with several studies, which indicate that storativity has low variability (Castagna et al., 2011; Dagan and Rubin, 1988; Freeze, 1975; Meier et al., 1998). Nonetheless, the inclusion of storativity as a random function is straightforward.

Moreover, it is essential to note that the diffusive term $DT(h(x,y;t))$ is not considered in the last term, together with $G(h,x,y;t)$. The previous assumption can explain this absence that the diffusive terms $DT(h(x,y;t))$ at the point-scale fundamentally contribute only to the mean behavior. Therefore, since $DT(h(x,y;t))$ is deemed deterministic, it will imply a null correlation against any random function.

Following the notation established by Garabedian (2009), the final form of the LEFPE (equation 3.11) can be written as shown in equation (3.12). D^i denotes the drift coefficient, while F^i expresses diffusion coefficients. In addition, $P(h,x,y;t)$ and $S(x,y)$ are substituted by P and S, respectively, to increase the readability and simplicity of large equations. The final form of the LEFPE (equation 3.11) for a well-defined heterogeneous confined aquifer can be written in a more concise form as follow

$$\begin{aligned}\frac{\partial P}{\partial t} =& + \frac{1}{S_s}\frac{\partial}{\partial x}(D^x P) + \frac{1}{S_s}\frac{\partial}{\partial y}(D^y P) + \frac{1}{S_s}\frac{\partial}{\partial h}(D^h P) \\ &+ \frac{1}{S^2}\frac{\partial^2}{\partial x^2}(F^x P) + \frac{1}{S^2}\frac{\partial^2}{\partial y^2}(F^y P) + \frac{1}{S^2}\frac{\partial^2}{\partial h^2}(F^h P)\end{aligned} \tag{3.12}$$

where the coefficients D's and F's in equation (3.12) can be expressed by

$$D^x = \left\langle \frac{\partial T_x(\overline{\boldsymbol{x}}_{\boldsymbol{t}})}{\partial x(t)} \right\rangle \tag{3.13}$$

$$D^y = \left\langle \frac{\partial T_y(\overline{\boldsymbol{x}}_{\boldsymbol{t}})}{\partial y(t)} \right\rangle \tag{3.14}$$

$$D^h = -\frac{1}{S_s}\left[D(\langle h(\overline{\boldsymbol{x}}_{\boldsymbol{t}})\rangle) + \langle G(\overline{\boldsymbol{x}}_{\boldsymbol{t}}, t)\rangle\right] \tag{3.15}$$

$$F^x = \int_0^t Cov_0\left[\frac{\partial T_x(h,x,y;t)}{\partial x}; \frac{\partial T_x(h,x,y;t-s)}{\partial x}\right] ds \tag{3.16}$$

$$F^y = \int_0^t Cov_0\left[\frac{\partial T_y(h,x,y;t)}{\partial y}; \frac{\partial T_y(h,x,y;t-s)}{\partial y}\right] ds \tag{3.17}$$

$$F^h = \int_0^t Cov_0\big[G(h,x,y;t); G(h,x,y;t-s)\big] ds \tag{3.18}$$

Equation (3.12) represents the resulting nonlocal Lagrangian-Eulerian Fokker-Planck equation, which is a deterministic second-order PDE. The probability density function for the hydraulic head in the $h-x-y$ phase through time can be obtained based on the solution of this nonlocal LEFPE. Hence, the spatio-temporal evolution of the hydraulic head's mean and standard deviation behavior (in the horizontal two-dimensional x-y space) can also be easily obtained, as well as any statistical moment.

3.3 Numerical solution for the of the Lagrangian-Eulerian Extension of the Fokker-Planck equation (LEFPE)

Once the LEFPE is derived (equation 3.12), it must be solved using a well-suited numerical scheme. However, significant computational resources are required to solve this equation, especially to compute the PDF, which has significant tails (Johnson et al., 1997). To date, most of the studies have used the Finite Element Method (Galán et al., 2007; Král and Náprstek, 2017; Masud and Bergman, 2005) and the Finite Difference Method (Fok et al., 2002; Qian et al., 2019; Sepehrian and Radpoor, 2015) to solve the conventional FPE numerically.

According to Masud and Bergman (2005), solving the conventional FPE for high dimensional problems (≥ 3) requires fine resolution in the time integration scheme. This requirement is even more critical at early times in the simulation to accurately resolve features in the rapidly evolving probability flow. The difficulties at the beginning of the simulation are associated with sharp changes occurring in small spatial regions, which can strongly influence the global properties of the system (Wei, 2000). Moreover, numerical algorithms can be highly sensitive to sharp gradients and quickly lead to numerically induced spatial and/or temporal chaos (Ablowitz et al., 1996).

Pichler et al. (2013) addressed comparative studies between the Finite Element Method (FEM) and Finite Difference Method (FDM), motivated by the high dimensionality in engineering dynamical systems. In this study, four widely used Finite Element and Finite Difference schemes to solve the transient FPE were reviewed by means of various numerical examples. They stated that FEM is preferable over FDM in terms of accuracy. However,

the FEM is only suitable for dimension ≤ 3 due to the significant numerical effort required. Park and Petrosian (1996) examined six Finite Difference Methods, three fully implicit and three semi-implicit commonly used to solve the FPE. Each method was evaluated in terms of its stability, accuracy, and efficiency. They concluded that the most robust FDM was the fully implicit Chang-Cooper method (Chang and Cooper, 1970).

Nowadays, the Chang-Cooper method is still considered one of the most popular tools to tackle the FPE due to its desirable properties (Buet and Dellacherie, 2010; Buet and Thanh, 2007; Butt, 2021; Larsen et al., 1985). This scheme ensures second-order accuracy, the positiveness of the solution, the conservation of the probability mass, and the exact representation of the analytical solution upon equilibration. While having the above-mentioned good characteristics, the Chang-Cooper scheme is based on the drift coefficients' positiveness ($D^i > 0$). Therefore, its solution is not guaranteed for the LEFPE derived in the previous section. In equation (3.12), the drift coefficients could take negative values for this particular problem since their sign depends on the transmissivity field, the hydraulic head curvature, flux conditions, and the source term. Hence, different from Chang and Cooper's scheme, an appropriate numerical scheme must be selected to tackle this problem.

The classical FPE closely resembles an advection-diffusion equation (ADE), widely studied in the computational fluid dynamics (CFD) field. Hence, several researchers have used common numerical schemes used in the CFD field to solve the conventional FPE. Even though the ADE is a linear PDE, difficulties arise due to its dual nature. If the ADE is diffusion-dominated, the equation behaves as a second-order parabolic equation, while if the ADE is advective-dominated, this equation behaves as a first-order hyperbolic equation. Therefore, to accurately solve the ADE, the numerical scheme needs to handle the mixed

parabolic-hyperbolic behavior of the physical system (Yeh et al., 1992). Further details about the numerical method used to solve the advection-diffusion equation can be found in Zheng and Bennett (2002).

Motivated by the drift coefficients' negativity in their problem, Ohara et al. (2008) used an explicit Finite Volume Method (FVM) to solve the FPE for the snowmelt process. The FVM was used to handle the boundary conditions that present potential difficulty for the snowmelt process. Ohara et al. (2008) pointed out that there is no universal scheme to solve the LEFPE since the selected numerical scheme depends on each particular problem.

Ceyhan and Kavvas (2018) developed a one-dimensional local FPE for the transient confined groundwater flow to a well. They used the explicit Lax-Wendroff scheme to solve the advective term, which is a second-order FDM. Additionally, a flux delimiter was included to mitigate artificial oscillation while preserving sharp concentration fronts using the upwind method (Shu and LeVeque, 1991). In the case of the diffusive term, an implicit second-order centered difference approximation was used, which resulted in a one-dimensional implicit-explicit (IMEX) scheme.

Selecting implicit schemes for solving diffusive terms and explicit schemes for advective terms is a common practice in the CFD field (Ascher et al., 1995; Chaudhry et al., 2015). This flexibility of mixing these schemes allows specialized numerical methods for systems formed by operators with different time scales. Refer to Ascher et al. (2006) for additional details about IMEX methods to solve PDEs.

ULTIMATE QUICKEST (Leonard, 1991) is another widely used explicit numerical

scheme to tackle the advection-diffusion equation in the CFD field. Cahyono (1993) tested more than 30 schemes used for discretizing advection terms numerically. He concluded that the ULTIMATE QUICKEST (UQ) was the most attractive since it was more general than the other schemes considered. Moreover, this scheme has not just been successfully applied in open-channel systems (Kashefipour and Zahiri, 2010; Yoshimura and Fujita, 2019) and subsurface hydrology (B. Lin and Falconer, 1997; Neumann et al., 2011) but also for solving the LEFPE (Tu et al., 2019). Tu et al. (2019) developed a LEFPE for a one-dimensional solute transport model under uncertain open-channel flow conditions. The UQ and an explicit second-order centered difference approximation were used to discretize the advective and diffusive terms. Their results suggest that the UQ numerical scheme can generally handle both small and significant variability in the flow fields with satisfactory results. Therefore, this numerical scheme could be used to solve advective-dominated as well as diffusive-dominated problems. While Leonard (1991) developed and applied the UQ scheme to a one-dimensional problem, Tu et al. (2020) generalized it to solve a two-dimensional LEFPE.

The ULTIMATE QUICKEST is based on the QUICKEST (Quadratic Upwind Interpolation for Convective Kinematics with Estimated Streaming Terms) scheme and the ULTIMATE (Universal Limiter for Transient Interpolation Modeling of the Advective Transport Equation) limiter. The QUICKEST scheme (Leonard, 1979) is an explicit third-order accurate upwind explicit scheme. Leonard (1979) designed this scheme for reducing numerical oscillations and truncations problems on highly advective unsteady flows. This scheme was derived using a finite difference approach for temporal discretization and a control volume approach for spatial discretization. This derivation implies calculating the discrete fluxes entering and leaving the control volume, estimated using an average quadratic upstream interpolation over a time increment. Even though the QUICKEST method has little numerical

dispersion and a large stability region, even minimal oscillations can corrupt the solution, especially near sharp gradients (Leonard, 1991; Neumann et al., 2011).

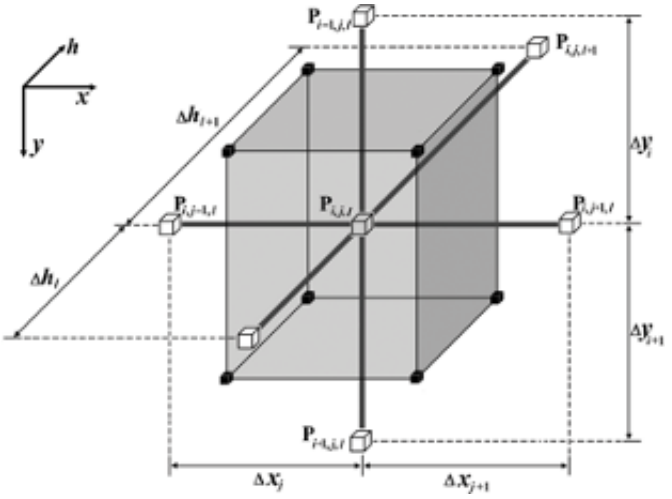


Figure 3.1: Schematic grid cell representation used to solve the LEFPE. The probability density function values P are defined at the cell center, whereas the probability fluxes are defined in the cell faces.

Leonard (1991) proposed the ULTIMATE (the Universal Limiter for Transient Interpolation Modeling of the Advective Transport Equations) to overcome the problem associated with the numerical oscillations by the QUICKEST scheme. ULTIMATE is a total variation diminution (TVD) algorithm that uses flux limiters to preserve the local monocity of the solution, even close to the sharp discontinuities and shocks. Thus, the local monotonicity is preserved, which is crucial for our problem since negative values would indicate negative probabilities, losing our problem's physical meaning. This scheme was derived using the Normalized Variation Diagram (NVD) concept, based on normalized variables to set the nodal boundaries to suppress oscillations. Several authors have indicated improved stability near steep fronts due to non-oscillatory behavior and reduced numerical dispersion by using ULTIMATE (Harris et al., 2002; B. Lin and Falconer, 1997; Neumann et al., 2011). In

addition to the ULTIMATE limiter's desirable characteristics, this algorithm involves minimal extra computational load and is particularly attractive for point sources (deterministic conditions) in the Lagrangian-Eulerian Fokker-Planck Equation.

The numerical method selected to solve the three-dimensional LEFPE shown by equation (3.12) needs to achieve two main objectives: (1) to have a potential to be parallelized, and also (2) it needs to control the numerical diffusion. The importance of parallelizing the numerical scheme relies on the high computational times required to solve multidimensional FPE (Masud and Bergman, 2005). In addition, parallelization will allow dealing with larger domains in space and probability sense. Furthermore, the main goal of this study is to estimate the ensemble mean and variance of the hydraulic head under uncertain parameters. Numerical schemes that cause numerical diffusion would overestimate the variance. Therefore, accurate numerical simulation without numerical dispersion plays a vital role in the goodness of the results.

In this study, the UQ is adopted to solve the advective terms. Like Tu et al. (2020), who generalized the UQ to solve a two-dimensional LEFPE, an attempt is made here to generalize this scheme to three-dimensional problems. An explicit scheme for advective terms was preferred over an implicit scheme for mainly two reasons: (1) implicit schemes for advection tend to be more numerically diffusive than explicit schemes (Dawson, 1995), and (2) implicit schemes for advection lead to a nonsymmetric matrix. In general terms, a nonsymmetric matrix is more challenging to invert (Ascher et al., 1995), and convergence of many iterative solvers will suffer (Wang et al., 2015). However, the stability region imposed by the Courant-Friedrich-Lewy (CFL) condition is a severe restriction of standard explicit numerical methods. This restriction implies a much smaller time step than permitted by accuracy

considerations (Casulli, 1990; Dehghan, 2004).

A typical practice to rectify the CFL condition's limitation consists of using an implicit discretization for diffusive terms in the ADE. Thus, the resulting implicit-explicit (IMEX) scheme will be less restrictive regarding the time step required for stability (Bürger et al., 2020). In addition to the desirable properties of IMEX schemes previously mentioned, it is essential to mention that the resulting linear 7-diagonal system of equations is symmetric, sparse, and strictly diagonally dominant, with positive elements on the main diagonal and negative elements elsewhere (Ascher et al., 1995). Therefore, this system is positive define, and it has a unique solution. Furthermore, a 7-diagonal sparse system of equations can be solved very efficiently by preconditioned conjugate gradient methods (Casulli, 1990).

Then, by applying forward finite difference formulas for the time derivative and the previously discussed schemes for the first and second spatial derivatives on a staggered grid, as shown in figure (3.1), the proposed LEFPE (equation 3.12) can be discretized by the following IMEX scheme:

$$\begin{aligned}\frac{P^{t+1}_{i,j,l}-P^{t}_{i,j,l}}{\Delta t} &= F^{x,t}\frac{P^{t}_{i-1/2,j,l}-P^{t}_{i+1/2,j,l}}{\Delta x_i}+\frac{(D^xP)^{t+1}_{i-1,j,k}+(D^xP)^{t+1}_{i,j,k}+(D^xP)^{t+1}_{i+1,j,k}}{\Delta x_i^2}\\ &+F^{y,t}\frac{P^{t}_{i,j-1/2,l}-P^{t}_{i,j+1/2,l}}{\Delta y_j}+\frac{(D^yP)^{t+1}_{i,j-1,k}+(D^yP)^{t+1}_{i,j,k}+(D^yP)^{t+1}_{i,j+1,k}}{\Delta y_j^2}\\ &+F^{h,t}\frac{P^{t}_{i,j,l-1/2}-P^{t}_{i,j,l+1/2}}{\Delta h_l}+\frac{\left(D^hP\right)^{t+1}_{i,j,k-1}+\left(D^hP\right)^{t+1}_{i,j,k}+\left(D^hP\right)^{t+1}_{i,j,k+1}}{\Delta h_l^2}\end{aligned} \tag{3.19}$$

where Δt, Δx_i, Δy_j, and Δh_l are discretization intervals for t, x, y, h dimensions, respec-

tively; the subscripts $\{i, j, l\}$ denote the nodes in x, y, and h domains, while the superscript t denotes the current time step of the probability density P, and $t + 1$ denotes the next time step. Note that in Equation (3.19), each subscript $\{i, j, l\}$ followed by a ± 0.5 may be interpreted as the probability flux or probability current, whereas P is considered as the state variable.

It is important to note that while the drift coefficients $D^{x,t}$, $D^{y,t}$, and diffusion coefficients $F^{x,t}$, $F^{y,t}$, and $F^{h,t}$ depend on external input such as the transmissivity random field and the source term, the drift coefficient $D^{h,t}$ depends on the curvature in the $x-y$ space at the current time t. Hence, we estimate this term in the current time 't' through the solution obtained in the last time '$t-1$' by assuming that small changes in the curvature occur during a time step Δt. Then, the diffusive term $DT(h(x, y; t))$ can be estimated as follows

$$\begin{aligned} DT(h(x,y;t)) &= T_x(\underline{x})\frac{\partial^2 h(\underline{x},t)}{\partial x^2} + T_y(\underline{x})\frac{\partial^2 h(\underline{x},t)}{\partial y^2} \\ &\approx T^x_{i,j}\frac{h^{t-1}_{i-1,j} - 2h^{t-1}_{i,j} + h^{t-1}_{i+1,j}}{\Delta x_i^2} + \approx T^y_{i,j}\frac{h^{t-1}_{i,j-1} - 2h^{t-1}_{i,j} + h^{t-1}_{i,j+1}}{\Delta y_i^2} \end{aligned} \tag{3.20}$$

Due to its control-volume formulation for spatial gradients, this scheme is sensitive to the flow direction. Hence, the probability fluxes of each face consider the flux direction to carry out the calculations. Thus, the probability fluxes in the advection terms, represented by $P_{i-1/2,j,l}$, $P_{i+1/2,j,l}$, $P_{i,j-1/2,l}$, $P_{i,j+1/2,l}$, $P_{i,j,l-1/2}$ and $P_{i,j,l+1/2}$ are estimated based on the one-dimensional ULTIMATE QUICKEST scheme. Assuming that the velocity in node $\{i, j, l\}$ is positive in the $x-direction$ ($F^{x,t} > 0$), the flux at the face $\{i = 1/2, j, l\}$ can be summarized as follows

1. Compute the upwind-biased second order difference $CURV$ and the normalization difference DEL as:

$$CURV_j = C_j = P^t_{i+1,j,l} - 2P^t_{i,j,l} + P^t_{i-1,j,l} \tag{3.21}$$

$$DEL_j = D_j = P^t_{i+1,j,l} - P^t_{i-1,j,l} \tag{3.22}$$

2. Set the probability flux as

$$P^t_{i+1/2,j,l} = \begin{cases} P^t_{i,j,l} & : \ |C_j| \geq |D_j| \\ \frac{P^t_{i+,j,l}P^t_{i,j,l}}{2} + |c_x|\frac{P^t_{i+,j,l}P^t_{i,j,l}}{2} - \frac{1-c_x^2}{6} & : \ |C_j| < |D_j| \end{cases} \tag{3.23}$$

3. Compute the reference face value:

$$P^t_{ref-i} = P^t_{i-1,j,l} + \frac{P^t_{i,j,l} - P^t_{i-1,j,l}}{c_x} \tag{3.24}$$

4. If $DEL_i > 0$, limit the fluxes according to

$$if \ P^t_{i+1/2,j,l} \leq P^t_{i,j,l} \implies P^t_{i+1/2,j,l} = P^t_{i,j,l} \tag{3.25}$$

Then

$$if \ P^t_{i+1/2,j,k} \geq min\left\{P^t_{ref-i}; P^t_{i+1,j,l}\right\} \implies P^t_{i+1/2,j,l} = min\left\{P^t_{ref-i}; P^t_{i+1,j,l}\right\} \tag{3.26}$$

5. Similar to step (4), if $DEL_i < 0$, limit the fluxes according to

$$if \ P^t_{i+1/2,j,l} > P^t_{i,j,l} \implies P^t_{i+1/2,j,l} = P^t_{i,j,l} \tag{3.27}$$

Then

$$if\ P^t_{i+1/2,j,k} < max\left\{P^t_{ref-i}; P^t_{i+1,j,l}\right\} \implies P^t_{i+1/2,j,l} = max\left\{P^t_{ref-i}; P^t_{i+1,j,l}\right\} \quad (3.28)$$

Like the flux calculation at face $P_{i+0.5,j,l}$, the fluxes at the other five faces can be estimated by following steps (1) to (5). Once all probability fluxes faces are estimated, they need to be substituted into the discretized LEFPE (equation 3.19). Thus, the probability density function at the next step can be calculated by solving the resulting system of linear equations. Moreover, the probability mass in the $x-y-h$ space must be conserved since the proposed LEFPE is a transport equation of the evolutionary probability density function. Therefore, reflecting boundary conditions are applied to the boundaries in the $x-y-h$ space. Furthermore, since an explicit numerical scheme is used to deal with the advective terms, the Courant–Friedrich–Lewy (CFL) condition is required at each computational node to achieve stable solutions. The CFL condition can be expressed in mathematical terms as follows

$$Courant\ Number = C = \frac{D^x \cdot \Delta t}{\Delta x_i} + \frac{D^y \cdot \Delta t}{\Delta y_j} + \frac{D^h \cdot \Delta t}{\Delta h_l} \leq C_{max} \leq 1 \quad (3.29)$$

The solution of this LEFPE results in the joint PDF of the state variables in the x-y-h space. Thus, it is possible to obtain a hydrologic system's ensemble behavior and variability defined by the point-scale stochastic confined groundwater flow equation under uncertainty in the transmissivity field and the source term. However, the proposed methodology may also be expanded to problems with uncertainties in storativity and boundary conditions.

Chapter 4

Application to groundwater flow problems

4.1 Statistical Characteristics of groundwater systems

Subsurface environments are complex natural systems that usually present a high degree of spatial heterogeneity in their properties. They are influenced by meteorological conditions, topography, geological structures, human activities, and vegetation, among others. Hence, these systems are full of uncertainties that need to be considered in the modeling process to obtain reliable solutions. The uncertainties in groundwater modeling are mainly attributed to factors such as (1) errors in the conceptualization of the system (Bredehoeft, 2003; Rojas et al., 2008), (2) uncertainty derived from observation error as well as the limited accuracy of instruments (Wu and Zeng, 2013) and (3)uncertainties due to the scarcity or the lack of measurements in space and time. Furthermore, an extra amount of uncertainty is inherent in any modeling process.

Soil system properties vary spatially and temporally, and direct or indirect techniques are used to measure them. Unfortunately, the estimation obtained by different approaches provides substantially different hydraulic parameter estimates (Morbidelli et al., 2017). The discrepancy in the estimations using different methods can be explained due to the flow geometry, sample size, soil conditions, and installation procedures (Lai et al., 2012). Moreover, it is well documented in the literature that groundwater parameters are scale-dependent (Mu et al., 2020; Sudicky, 1986), making the measurements even more complex.

Not only has much attention been spent on measuring hydraulic properties in the field, but also on estimating them in locations with no information due to observation technology and cost constraints. To overcome these limitations, the restricted geological data available is usually analyzed by employing geostatistical tools, such as semivariograms (Lin et al., 2017), transition probabilities (Carle and Fogg, 1996), and patterns of geological heterogeneities (Fleckenstein et al., 2006) to

4.1.1 Modeling uncertainty in the hydraulic conductivity

The geostatistical interpretation of the measured conductivity data is usually considered a log-normal probability distribution with an assumed exponential correlation model (Alecsa et al., 2020; Bohling et al., 2016; Woodbury and Sudicky, 1991). In addition, to represent the correlation structure in the conductivity data, a common practice is to model the log-conductivity transformed data with an exponential correlation structure (Baye et al., 2013). However, other authors have considered the same correlation model just for conductivity without applying the log transformation (de Brito Fontenele et al., 2014; Rahman et al., 2008; Razack and Lasm, 2006; Rotzoll and El-Kadi, 2008).

Stochastic hydrogeology's common assumption is that log-conductivity data follows an exponential correlation structure, and two factors could explain this supposition. The first one came from the stochastic flow-transport theory (Dagan, 1982; Gelhar and Axness, 1983), where the flow and transport equations are expanded by using power series, and then a first-order approximation is attempted. The solutions obtained by this method take advantage of the Gaussian properties of the $ln(K)$ data, which depend only on the mean and the autocorrelation of log-conductivity data. Furthermore, geostatistical techniques (e.g., kriging) also have focused on the spatial structure of the log-transformed conductivity since prediction performances by using kriging are generally worse when the sample distribution displays a strong skewness (Ahmed and Marsily, 1987; Patriarche et al., 2005). However, Razack and Lasm (2006) stated that variograms of hydraulic conductivity (which tend to follow a lognormal distribution) are more structured than the variograms of log-conductivity (which are normally distributed), raising the question of whether data normality is necessary or not in the geostatistical characterization.

In this study, the method proposed by Zárate-Miñano and Milano (2016) is used to carry out the transmissivity random field generation. This method was originally developed to construct wind speed models, and it involves the use of stochastic differential equations (SDEs) by using the equivalence between the Langevin and the Fokker–Planck equations and the regression theorem. This study considers a log-normal distribution for transmissivity since this assumption is the most used to model transmissivity fields in subsurface hydrology. To generate the transmissivity random field, the only information needed to build the models is the mean and variance of the transmissivity as well as its correlation length. Thus, it is possible to generate stochastic processes with a specific probability distribution and exponentially decaying autocorrelation function. This model can be expressed as follows:

$$dT(x) = a\left[T(x), x\right] \cdot dx + b\left[T(x), x\right] \cdot dW(x) \tag{4.1}$$

where the drift term $a\left[T(x), x\right]$, and the diffusion term $b\left[T(x), x\right]$ are expressed as follows

$$a\left[T(x), x\right] = -\alpha \left(T(x) - exp\left[\mu + \frac{\sigma^2}{2}\right]\right) \tag{4.2}$$

$$b\left[T(x), x\right] = \sqrt{b_1\left[T(x), x\right] \cdot b_2\left[T(x), x\right]} \tag{4.3}$$

with

$$b_1\left[T(x), x\right] = \sqrt{2\pi} \cdot \alpha \cdot \sigma \cdot T(x) \cdot exp\left[\mu + \frac{\sigma^2}{2} + \frac{\left[ln(T(x)) - \mu\right]^2}{2\sigma^2}\right] \tag{4.4}$$

$$b_2\left[T(x), x\right] = erf\left(\frac{\mu + \sigma^2 - ln(T(x))}{\sqrt{2}\sigma}\right) - erf\left(\frac{\mu - ln(T(x))}{\sqrt{2}\sigma}\right) \tag{4.5}$$

where $W(x)$ represents a standard Wiener process, $erf(\cdot)$ is the error function, T is transmissivity, α is the autocorrelation coefficient, and μ and σ are the mean and the standard deviation of the natural logarithm of variable T, respectively.

To generate the transmissivity random field implies the numerical integration of the stochastic partial differential equation (4.1). To solve the SPDE (4.1) numerically, the implicit Milstein scheme is used (Mil'shtejn, 1975). This scheme is based on the truncated Ito-Taylor expansion, and it has a convergence error of order 1 ($\mathcal{O}(h)$) in both the weak and the strong sense. This scheme can be expressed as follows:

$$T(x+\Delta h) = T(x) + \frac{\Delta h}{2} \cdot \Big[a\left[T(x+\Delta h), x+\Delta h\right] + a\left[T(x), x\right] \Big] \tag{4.6}$$

$$+ b\left[T(x), x\right] \cdot \Delta W(x) \cdot \frac{1}{2} b \cdot \left[T(x), x\right] \cdot \frac{\partial b\left[T(x), x\right]}{\partial T(x)} \cdot \Big[(\Delta W(x))^2 - \Delta h \Big]$$

where Δh is the integration spatial step and $\Delta W(x) \sim \mathcal{N}(0, \Delta h)$ are random increments of the Wiener process.

4.2 Application of the Monte Carlo Approach

Deterministic models are characterized by variables that are uniquely determined by specific parameters. However, uncertainties inherent to any groundwater model render the deterministic governing equation (equation 2.12) into a stochastic partial differential equation (SPDE), challenging to solve. Just a few of these SPDEs have an analytical solution. Therefore, numerical techniques are standard tools to tackle this issue.

The Monte Carlo simulation is the most widely used approach to consider uncertainties in the analysis. This approach consists in generating a series of samples from the statistical distribution of the input parameters. Then, deterministic equations for each set of realized parameter values need to be solved. Once a sufficient number of realizations are solved in a deterministic fashion, those results can be used to determine the statistical properties of the system (Hassan et al., 1998; Jafari et al., 2016).

The MC approach is accepted as the most robust technique used for benchmarking purposes and for evaluating the system's uncertainties (Scharffenberg and Kavvas, 2011). However, this approach has a high computational cost. This limitation is more relevant for large

and complex systems (Hayley, 2017). These complex systems are composed of an intensive number of variables, and the inter-relationships between them increase the system's mathematical complexity.

This study uses the MC approach to assess the results obtained through the LEFPE methodology. The main program, which solves the groundwater flow equation in a deterministic fashion, was coded on Matlab. The number of realizations used in the MC simulation varies for each case. A convergence criterion was adopted to establish the number of realizations needed for the MC simulation. The criterion evaluates the relative standard deviation error for N and $N+2000$ realizations. If their relative error is more significant than 1%, other 2000 realizations are incorporated until satisfying a change of less than 1% in the standard deviation. This analysis was performed over the standard deviation since the mean tends to converge quickly.

Figure (4.1) shows the relative standard deviation error for Applications I and II, which will be discussed in more detail in section 5. While Application I consider an uncorrelated transmissivity field (blue line), Application II considers a correlated transmissivity field (red line). Simulations indicate that around 90.000 and 110.000 samples are needed to satisfy the criterion defined previously for the uncorrelated and correlated fields, respectively. Furthermore, the results suggest that a higher number of samples are needed for correlated fields compared to the uncorrelated fields. This result is consistent with results in section 5 since the ensemble standard deviation is higher for the correlated field. It is important to note that this analysis was carried out on a one-dimensional problem. Therefore, this number for higher-dimensional problems would increase considerably as well as its computational times.

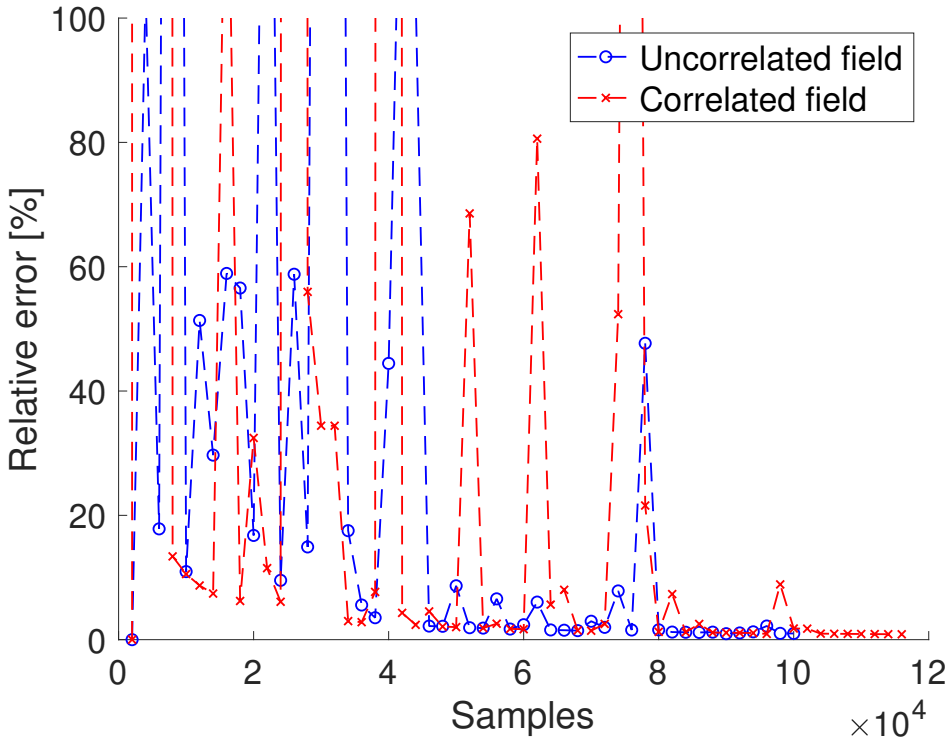


Figure 4.1: Relative standard deviation error for MC simulations by considering uncorrelated and correlated transmissivity field for Application I and II.

4.2.1 Verification of the Deterministic Numerical Solution

In order to verificate the capabilities of the numerical scheme used to solve the governing equation (2.9) in chapter 2, a set of three benchmark tests are performed. These numerical tests include analytical solutions for the drawdown of two-dimensional radial flow towards a discharge well in an infinite aquifer (Theis, 1935; Thiem, 1906) and a grid and time convergence analysis. Thus, the numerical solution is compared against the analytical solution, which is used as a reference.

Test 1

Theis (1935) presented an analytical expression for transient drawdown of two-dimensional radial flow towards a discharge well in an infinite homogeneous aquifer. This solution gives

the hydraulic head as function of the distance from the well (r) and time (t) as follows

$$h(r,t) = h_0 - \frac{Q}{4\pi T} W(u) \tag{4.7}$$

where,

$$u = \frac{r^2 S}{4Tt} \qquad W(u) = -0.577216 - ln(u) - \sum_{i=0}^{\infty} \frac{(-u)^i}{i \cdot i!} \tag{4.8}$$

where h_0 is the initial hydraulic head, Q is the discharge rate of the well, T is the aquifer transmissivity, $W(u)$ is the well function, and S is the storativity. The domain used consists of a square of 2500 [m] by 2500 [m], where each grid was 50 [m] by 50 [m].

Q	20 $[m^3/d]$
h_0	20 [m]
T	1 $[m^2/d]$
S	0.001

Table 4.1: Parameter values for the MC numerical test 1 and 2.

Figures (4.2.a), (4.2.b) and (4.2.c) show the transient flow calculations for a distance r equal to 100, 200 and 300 [m] from the well center respectively and the analytical solution proposed by Theis (1935). Results suggest a near-perfect agreement for short discharge times. As time increases, the hydraulic head computed by the numerical scheme proposed reaches a steady-state value depending on the distance from the well's center, while Theis' solution indicates an ongoing drawdown. The difference between the analytical and the numerical solution for long-time simulations can be explained by the influence of boundaries in the numerical simulation. Additionally, it is possible to note that observation points closer

to the border "feel" the boundary effect earlier than observation points farther from the boundary. This fact explains why Figure (4.2.a) has the best agreement in the water table height between the analytical solution and the numerical scheme.

Test 2

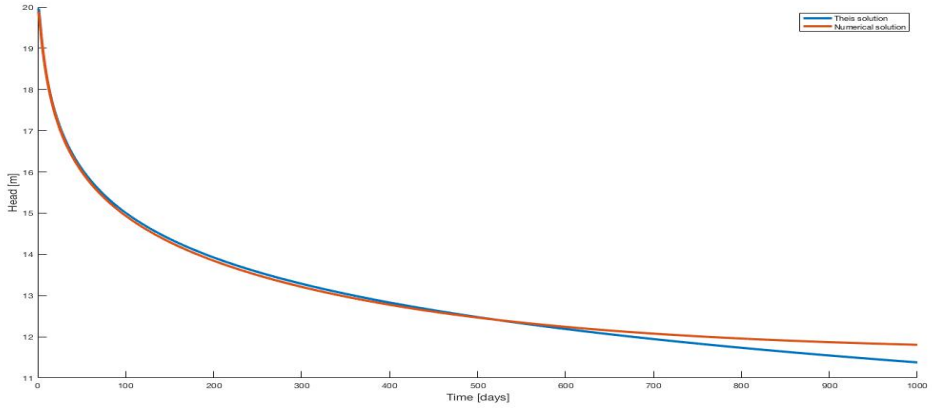
Theis' solution cannot reproduce the steady-state solution for long discharge times. However, Thiem (1906) proposed an analytical formulation for steady-state radial flow conditions, where the hydraulic head is given by

$$h(r) = h_0 - \frac{Q}{2\pi T} ln\left(\frac{R}{r}\right) \tag{4.9}$$

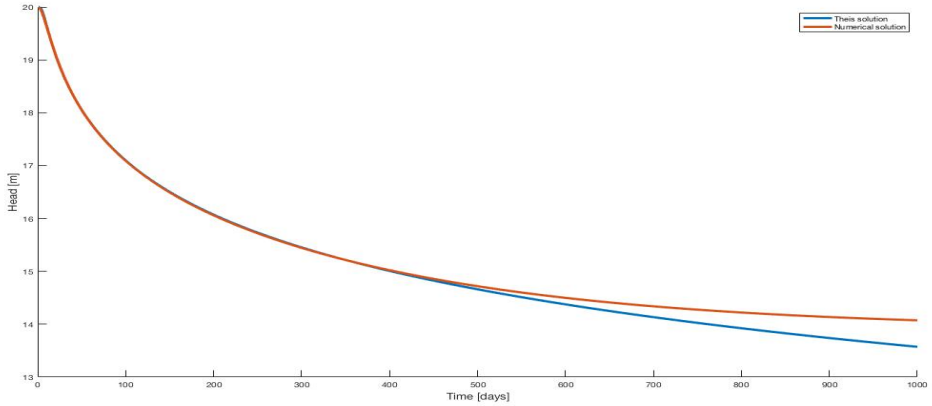
where R is the radius of influence, i.e. the distance at which the head equals h_0. To perform the numerical experiment, it was considered the same values used in test 1 (Table 4.1) and a radius of influence equal to 1300 [m]. The steady-state results obtained by the numerical scheme proposed were compared to the result of the Thiem's analytical formulation in Figure 4.3. Results show a good agreement with Thiem's solution along the whole model domain.

Test 3

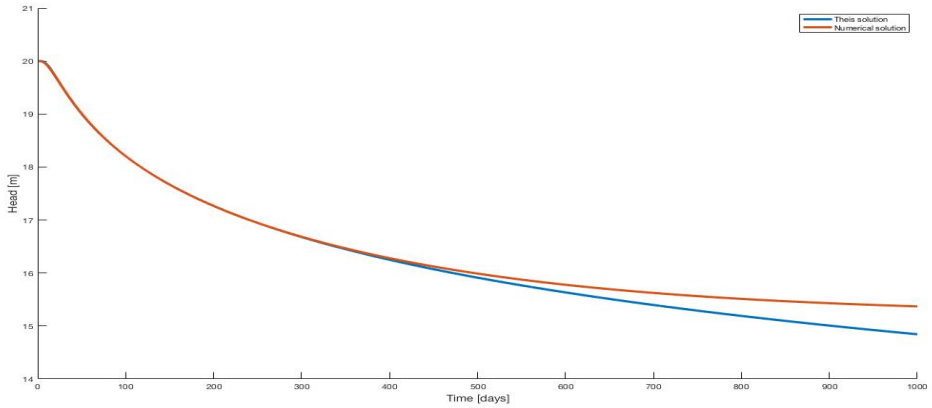
A grid and time convergence study is performed to verify the code and analyze the quantification of errors introduced during its application. This numerical test considers a rectangular domain with side lengths L and H. The mesh and the time step were successively refined, according to table (4.3), until a sufficient level of accuracy was obtained. The analytical



(a) r = 100 [m]



(b) r = 200 [m]



(c) r = 300 [m]

Figure 4.2: Test 1: Head variation over time at (a) r=100 [m], (b) r=200 [m] and (c) r=300 [m].

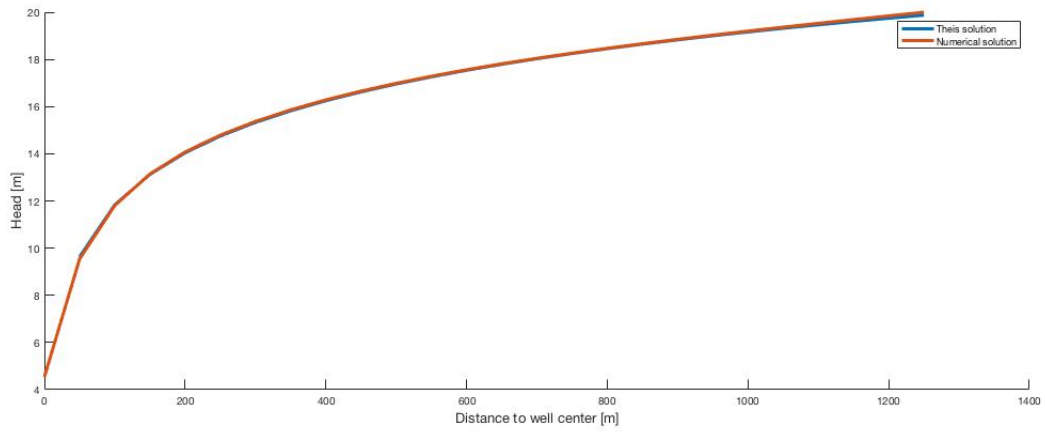


Figure 4.3: Test 2: Head variation over distance from the well center at steady-state conditions.

solution used to carry out this analysis considers a unique transmissivity and storativity for the whole domain. Thus, the governing equation (2.12) can be expressed as follows

$$\frac{\partial h}{\partial t} = \frac{T}{S}\frac{\partial^2 h}{\partial x^2} + \frac{T}{S}\frac{\partial^2 h}{\partial y^2} \qquad 0 \leq x \leq L, \quad 0 \leq y \leq H \tag{4.10}$$

with

$$h(0,y,t) = 0[m] = h(L,y,t), \qquad h(x,0,t) = 0[m], \qquad \frac{\partial h}{\partial y}(x,H,t) = 0 \qquad h(x,y,0) = 100[m]$$

which has an analytical solution given by

$$h(x,y,t) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1600}{(2n-1)(2m-1)\pi^2} \cdot sin\left(\frac{(2n-1)\pi x}{L}\right) \cdot sin\left(\frac{(2m-1)\pi y}{2H}\right) \cdot exp\left(-\left(\frac{(2m-1)^2}{4H^2} + \frac{(2n-1)^2}{L^2}\right)\frac{T}{S}\pi^2 t\right) \tag{4.11}$$

L	6 [m]
H	4 [m]
T/S	1 $[m^2/d]$

Table 4.2: Parameter values for the MC numerical test 3.

Parameters used to perform the numerical experiment are shown in Table 4.2. Numerical results were compared to the analytical solution at t=1 [d] by monitoring maximum errors versus mesh-time length scale shown in Table 4.3 and Figure 4.4.

dx = dy [m]	1	1/2	1/4	1/8	1/16	1/32	1/64	1/128	1/256
dt [d]	1	1/2	1/4	1/8	1/16	1/32	1/64	1/128	1/256
max error	21.892	13.588	7.788	4.209	2.195	1.121	0.567	0.285	0.143

Table 4.3: Discretization parameter values and maximum errors for the MC numerical test 3.

Results suggest that the discrete solution converges to the exact solution as the mesh spacing and time step are reduced in calculations. Additionally, space and time order of the scheme used in this study is two and one respectively. Therefore, an order one for this numerical scheme is expected. That hypothesis can be verified in the Figure 4.4, where the effective order is one since it is possible to note that the slope of the plot (blue line) tends to one as the grid size and time step decrease.

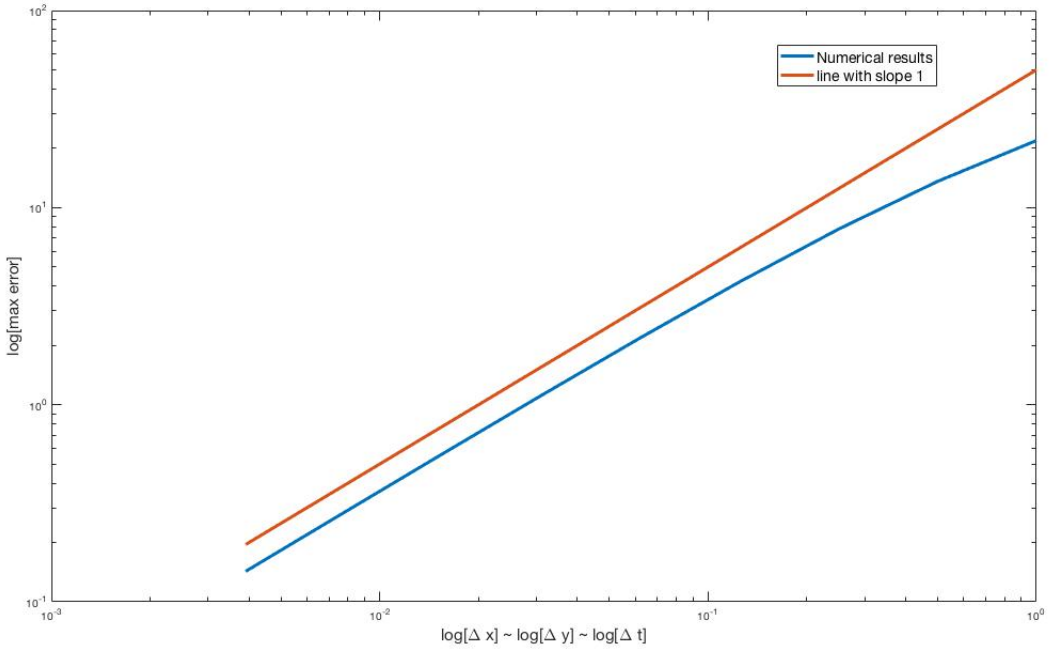


Figure 4.4: Test 3: Maximum error as a function of space and time

4.3 Application of the Proposed Fokker-Planck Equation Methodology

Unlike the MC approach, the proposed stochastic framework solves the PDF evolution of the state variables in space and time in only one simulation. However, this particular LEFPE has one extra dimension compared to the governing equation used to carry out the MC simulation. This extra dimension needs to be handled appropriately to keep reasonable computational times. For this reason, the numerical scheme presented in the previous section to solve the FPE was implemented in C++, a low-level language. Furthermore, the OpenMP library (Chandra et al., 2001) is used to parallelize part of the code, which explicitly solves the advection terms. The parallelization is just applied to the advection term solver since that part is the bottleneck in the code flow.

4.3.1 Validation of the Numerical scheme proposed to solve the LEFPE

In order to validate the scheme, an advection-dispersion problem in three dimensions was carried out to assess the performance of the numerical solution. To perform the validation, a unit Gauss impulse centered at $(x_0 = 0.3, y_0 = 0.3, z_0 = 0.3)$ was used as initial condition, represented by

$$C(x, y, z, t = 0) = exp\left(-\frac{(x - x_0)^2}{C^x} - \frac{(y - y_0)^2}{C^y} - \frac{(z - z_0)^2}{C^h}\right) \tag{4.12}$$

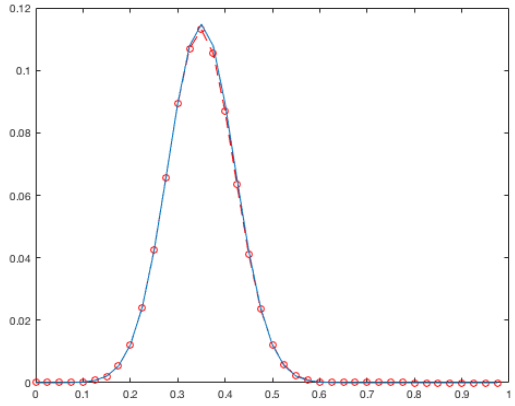
where the exact solution according to Shukla and Tamsir (2018) is given by

$$C(x, y, z, t) = \frac{1}{(1 + 4 \cdot t)^{3/2}} \cdot exp\left(-\frac{(x - B^x t - x_0)^2}{C^x(1 + 4t)} - \frac{(y - B^y t - y_0)^2}{C^y(1 + 4t)} - \frac{\left(z - B^h t - z_0\right)^2}{D^h(1 + 4t)}\right) \tag{4.13}$$

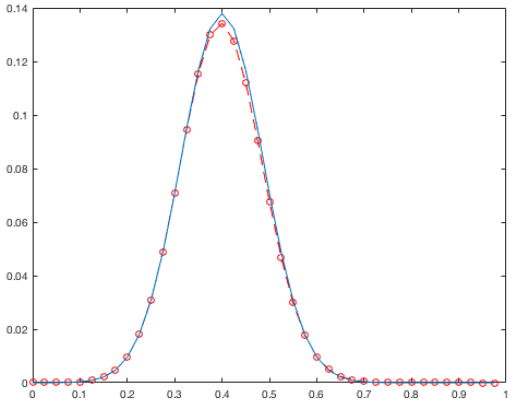
The discretization of the spatial coordinates was established as $\Delta x = \Delta y = \Delta h = 0.025[m]$ and the simulation time between 0 and 1 seconds with a time step of $\Delta t = 0.025[s]$. The numerical solution was obtained with parameters $B^x = B^y = B^h = 0.2$ and $C^x = C^y = C^h = 0.01$.

A comparison between the scheme proposed and the theoretical values are shown in Figure (4.5). The comparison shows a good agreement between the exact analytical solution and numerical solution, ensuring that the numerical scheme appropriately describes the evolution of this advective-diffusive problem.

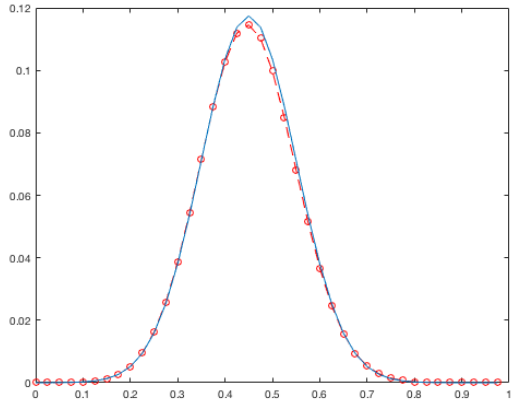
Figure (4.5) illustrates that the mean location of the plume matches the theoretical values given by the drift coefficients, although a small presence of numerical diffusion is observed for the numerical scheme. The numerical diffusion observed could be due to the mesh size



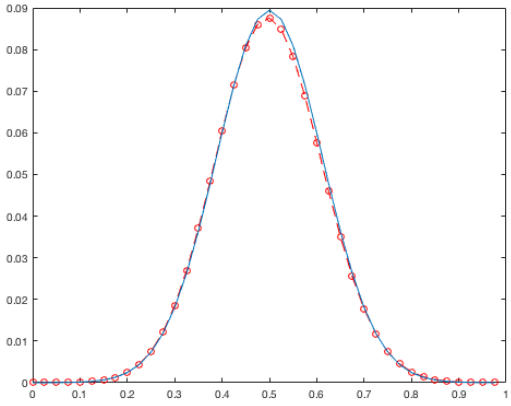
(a) t = 0.25 [s]



(b) t = 0.5 [s]



(c) t = 0.75 [s]



(d) t = 1.0 [s]

Figure 4.5: Comparisons between numerical and exact analytical solutions along face x=y, z=0.5 [m] at different times.

since the numerical scheme proposed is susceptible to this parameter. However, the results show that the numerical scheme has the same shape-evolution as the theoretical value, where the mean's magnitudes are almost identical.

4.3.2 Initial and Boundary Conditions

The LEFPEs' solution is numerically obtained by solving the partial differential equation (3.12). Therefore, the initial and boundary conditions of the problem need to be correctly established to solve this PDE. These conditions must be translated from the real space $x-h$ to the probability space $P-x-h$.

According to Kavvas (2003), deterministic boundary and initial conditions can be represented Dirac delta functions in the probabilistic space. This function is also known as the unit impulse, infinitely peaked at one location while it is zero everywhere else. Since it is impossible to numerically represent a Dirac delta pulse with infinite value at only one position, it is estimated as a block that has a very high height with a small bottom base (in a three-dimensional framework). Therefore, in the LEFPE approach, deterministic conditions can be expressed by means of

$$P(x=x^*,t=t^*)=\delta(h-h^*)=\begin{cases} 0 & : \ h\neq h^* \\ \infty & : \ h=h^* \end{cases} \tag{4.14}$$

In this study, initial conditions are considered deterministic, which means a known hydraulic head at the beginning of the simulation. For the case of the boundary conditions, they are also treated as deterministic. However, they can be represented by a known hydraulic head (Dirichlet B.C.) or flux (Neumann B.C.) at the boundary. In the case of Dirichlet B.C, the treatment is similar to the representation of the initial conditions. Nevertheless, the case for Neuman B.C. in the LEFPE needs to be carefully analyzed in agreement with its

governing PDE at that location (Diaz 2019).

4.3.3 Incorporation of the transmissivity random field statistics into the LEFPE framework.

The variability in the hydraulic head generated by the transmissivity field is one of the critical objectives of this study. Hence, the statistical characteristics of the transmissivity random field need to be incorporated into the stochastic model proposed. Hence, the advection (D^i) and diffusion (F^i) terms from equation (3.12) need to be correctly estimated.

Application I considers an uncorrelated random field. Even though this is not realistic, it will be helpful to compare it against a correlated one. Thus, exploring how the hydraulic head behavior is modified by including a spatial correlation structure in the transmissivity field will be possible. An uncorrelated random field means the process has a very short memory. In other words, the values next to the point under analysis do not depend on their neigh-borhood. Therefore, the transmissivity random field can be estimated as delta-correlated. In mathematical terms, this assumption simplifies the estimation of the diffusion coefficient (F^i) in equation (3.12), which involves estimating the ordered covariance terms. Then, the ordered-covariance term in the LEFPE is simplified as follows.

$$\begin{aligned}\int_0^t Cov_0\left[\frac{\partial T_x(h,x,y;t)}{\partial x};\frac{\partial T_x(h,x,y;t-s)}{\partial x}\right]ds &= \int_0^t Cov_0\left[\frac{\partial T_x(h,x,y;t)}{\partial x};\frac{\partial T_x(h,x,y;t-s)}{\partial x}\right]\delta(s)ds \\ &= \int_0^t Cov_0\left[\frac{\partial T_x(h,x,y;t)}{\partial x};\frac{\partial T_x(h,x,y;t)}{\partial x}\right]ds \\ &= Var\left[\frac{\partial T_x(h,x,y;t)}{\partial x}\right]\end{aligned}$$

Through this assumption, the ordered covariance term associated with the diffusion term (F^x) in equation (3.11) collapses to the variance. Therefore, the non-local LEFPE reduces to a less complex local FPE. On the other hand, a correlated transmissivity field implies calculating the covariance term in equation (). The incorporation of the correlation is made by using the exponential model denoted by equation (4.1). By combining equation (4.1) into (3.12), the diffusion term (F^i) is represented according to the following equation.

$$\begin{aligned}
&\int_0^t Cov\left[\frac{dT(x_t)}{dx_t}; \frac{dT(x_{t-s})}{dx_{t-s}}\right] ds \\
&\quad = \int_0^t Cov\Big[a\left[T(x_t), x_t\right] + b\left[T(x_t), x_t\right]\xi(x_t); a\left[T(x_{t-s}), x_{t-s}\right] + b\left[T(x_{t-s}), x_{t-s}\right]\xi(x_{t-s})\Big] ds \\
&\quad = \int_0^t Cov\Big[a\left[T(x_t), x_t\right]; a\left[T(x_{t-s}), x_{t-s}\right]\Big] ds + smaller\ order\ terms \qquad (4.15) \\
&\quad \approx \alpha^2 \cdot \sigma^2 \cdot \int_0^t e^{-\alpha|x_t - x_{t-s}|} ds
\end{aligned}$$

While the Eulerian location x_t is known, the Lagrangian location x_{t-s} is not. However, by substituting equations (2.16), (2.17), (2.18) into (3.9) is possible to estimate the Lagrangian unknown location x_{t-s} numerically (Liang and Kavvas, 2008). In addition, from equation (4.15) is possible to note that the integral converges to a certain value as the distance between x_t and x_{t-s} increase. Mathematically speaking, this can be explained by $\lim_{x\to\infty} \exp\left[-x\right] = 0$

Chapter 5

Numerical Results and Discussion

This chapter evaluates the performance of the proposed upscaling technique for the groundwater flow process in confined aquifers by comparing results from corresponding Monte Carlo simulations. To perform this task, three different numerical tests are carried out to illustrate the potential and capabilities of the novel proposed stochastic model. Application I and II are one-dimensional problems. While application I assumes an uncorrelated transmissivity random field, application II considers the correlation structure. In addition, application III corresponds to a two-dimensional extension of Application II.

5.1 Application I - One-dimensional uncorrelated transmissivity case

The first application is performed on a one-dimensional problem. This hypothetical problem involves an aquifer of 1000 [m] long and a width of 15 [m], as presented in Figure (5.1). For this case, the transmissivity is taken as the only random field in the groundwater flow equation (2.12), while the storativity is considered deterministic. On the other hand, both initial and boundary conditions are taken as deterministic by following the steps described in section 4.3.2.

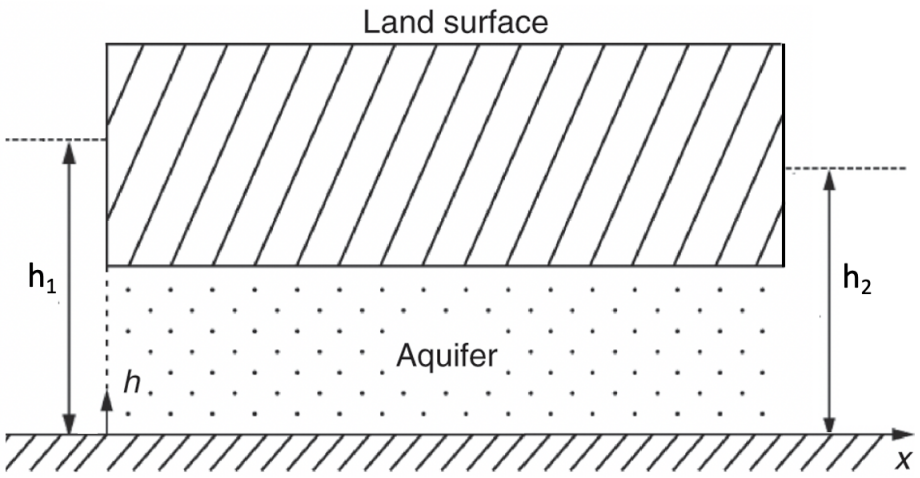


Figure 5.1: Hypothetical problem: One-dimensional confined aquifer.

This case assumes a lack of correlation structure in the transmissivity field. Even though this assumption does not accurately represent reality, it is used to explore the capabilities of the proposed methodology. In addition, application I will be used to compare against application II, which includes the correlation structure in the transmissivity field for the same problem.

In this case, the transmissivity random field was assumed as log-normal distributed. A transmissivity mean equal to $\mu_T = 3 \cdot 10^{-2}$ $[m^2/s]$ is used, which is a representative hydraulic conductivity value ($K_x = 2 \cdot 10^{-3}[m/s]$) for gravel aquifers (Domenico and Schwartz, 1998). In terms of their variability, a coefficient of variation equal to $CV = 1.5$ was considered. According to Todd (1980), typical values for specific storage S_s in confined aquifers range from $5 \cdot 10^{-5}$ $[1/m]$ to $5 \cdot 10^{-3}$ $[1/m]$. Thus, numerical tests in this study consider a specific storage S_S equal to 10^{-4} $[1/m]$, which is consistent with the range reported by Todd (1980)

.

Boundary conditions were modeled as deterministic Dirichlet boundary conditions (or first-type) along $x = 0$ and $x = 1000$ [m], equal to 22 [m] and 20 [m], respectively. The initial condition is set as deterministic too, such as $h(t = 0)$ is equal to 20 [m], which results in a flux directed along the $x-axis$. The statistics of the physical parameters and the model properties are presented in Table (5.1).

Parameter	Symbol	Value
Transmissivity mean	μ_T	$3 \cdot 10^{-2}[m^2/s]$
Coefficient of Variation	CV	1.5
Specific storage	S_s	10^{-4}
Aquifer length	L	1000 [m]
Aquifer width	B	15 [m]
I.C. at t=0	h_0	20 [m]
B.C. at x=0	h_1	22 [m]
B.C. at x=L	h_2	20 [m]
x-grid size	Δx	2 [m]
h-grid size(1)	$\Delta h^{(1)}$	1 [cm]
correlation length(2)	$L_x^{(2)}$	100 [m]
Simulation time	T	3 [hours]

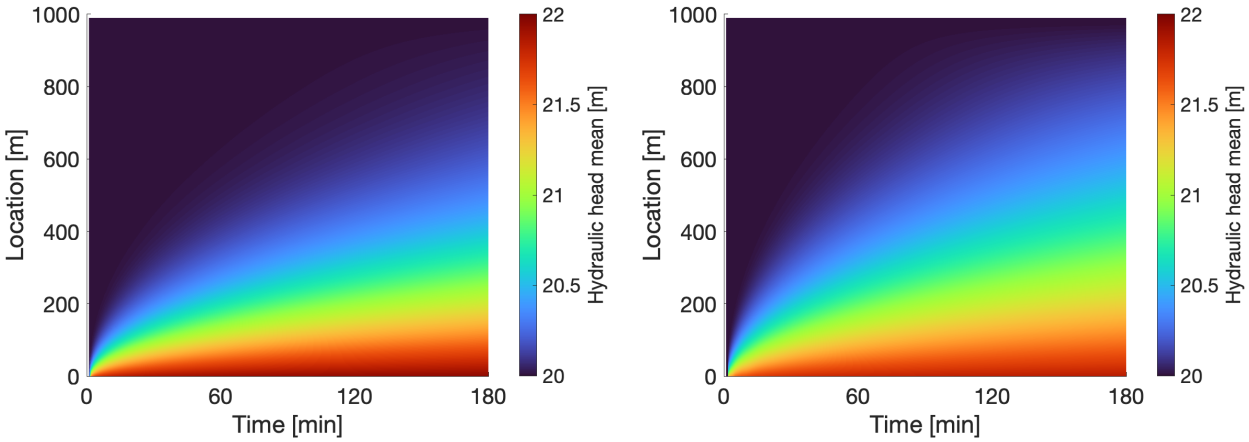
(1) It just applies to the LEFPE methodology
(2) Application II

Table 5.1: Physical parameters values for applications I and II.

Regarding the numerical discretization for the MC and LEFPE, both models were implemented using the same spatial grid $\Delta x = 2[m]$. In the case of the time discretization, while the time step in the MC simulation was fixed ($\Delta t = 1[s]$), the FPE was variable to satisfy the Courant constraint (Equation 3.29). This time step restriction is crucial at the beginning of the simulation due to the high gradients in the probability domain $P - x - h$, which add extra numerical diffusion to the system. These high gradients are due to incorpo-

rating deterministic conditions by means of Dirac delta pulses. However, as time goes, the simulation tends to smooth the probability mass in the $x-h$ plane, reducing the probability gradients.

Moreover, this variable time step helps reduce the computation time since the LEFPE has an extra dimension compared to the original groundwater flow equation, implying a higher computational cost. This extra dimension is associated with the hydraulic head in the probability space, and it shows up due to the probabilistic nature of the LEFPE. Since the state variable in this equation is the PDF (P) instead of the groundwater flow equation in which the state variable is the hydraulic head (h).



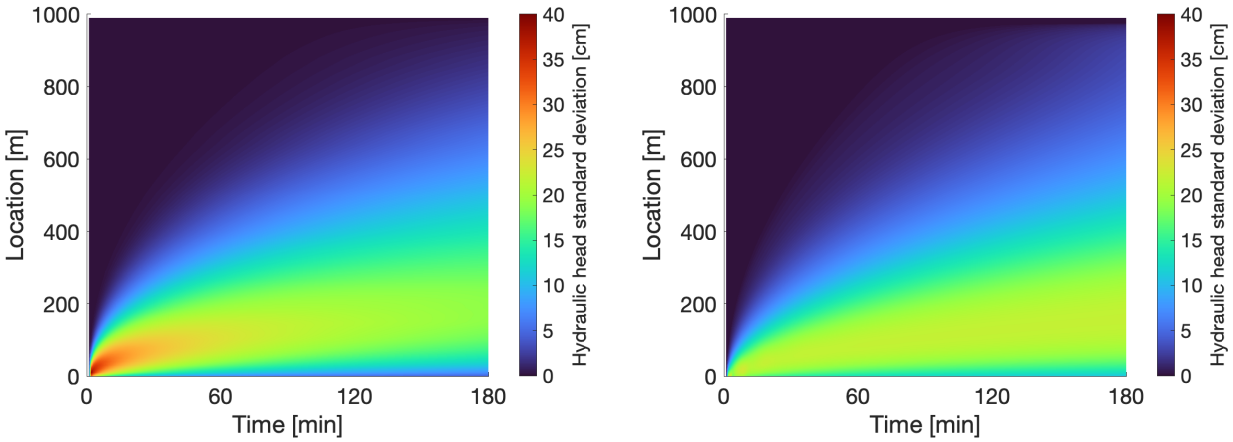
(a) MC (b) FPE

Figure 5.2: Plan view of the mean Hydraulic head over position and time for the results obtained by the LEFPE method and those obtained by the MC method for application I.

Figure (5.2) shows the ensemble average of the hydraulic head in time and space for both the LEFPE methodology and the MC approach. In the comparison between figures (5.2.a) and (5.2.b), it is clear that the ensemble average hydraulic head computed by the

LEFPE method appropriately resembles the one obtained from the MC simulations, sharing the same behavior in time and space.

In addition, to evaluate the performance of the LEFPE methodology, an assessment of the ensemble standard deviation of the hydraulic head is performed. Figure (5.3) shows the standard deviation in time and space for both approaches. Results suggest that the standard deviation has similar spatio-temporal behavior in both methodologies. However, they show some discrepancies at early times, where the MC approach results in higher standard deviation values compared to the results from the proposed stochastic framework.



(a) MC (b) FPE

Figure 5.3: Plan view of the standard deviation of the Hydraulic head over position and time for the results obtained by the LEFPE method and those obtained by the MC method for application I.

The combination of two factors could explain the high values for the standard deviation obtained through the MC approach at early times. These factors are (1) the proximity to B.C. and (2) the uncorrelated random field used in this case. Since we use an uncorrelated transmissivity random field, the cell next to the boundary condition could significantly differ

in the transmissivity value with respect to its neighborhood. Therefore, the conductance between those cells (estimated by equation 2.21) has high variability for MC simulations.

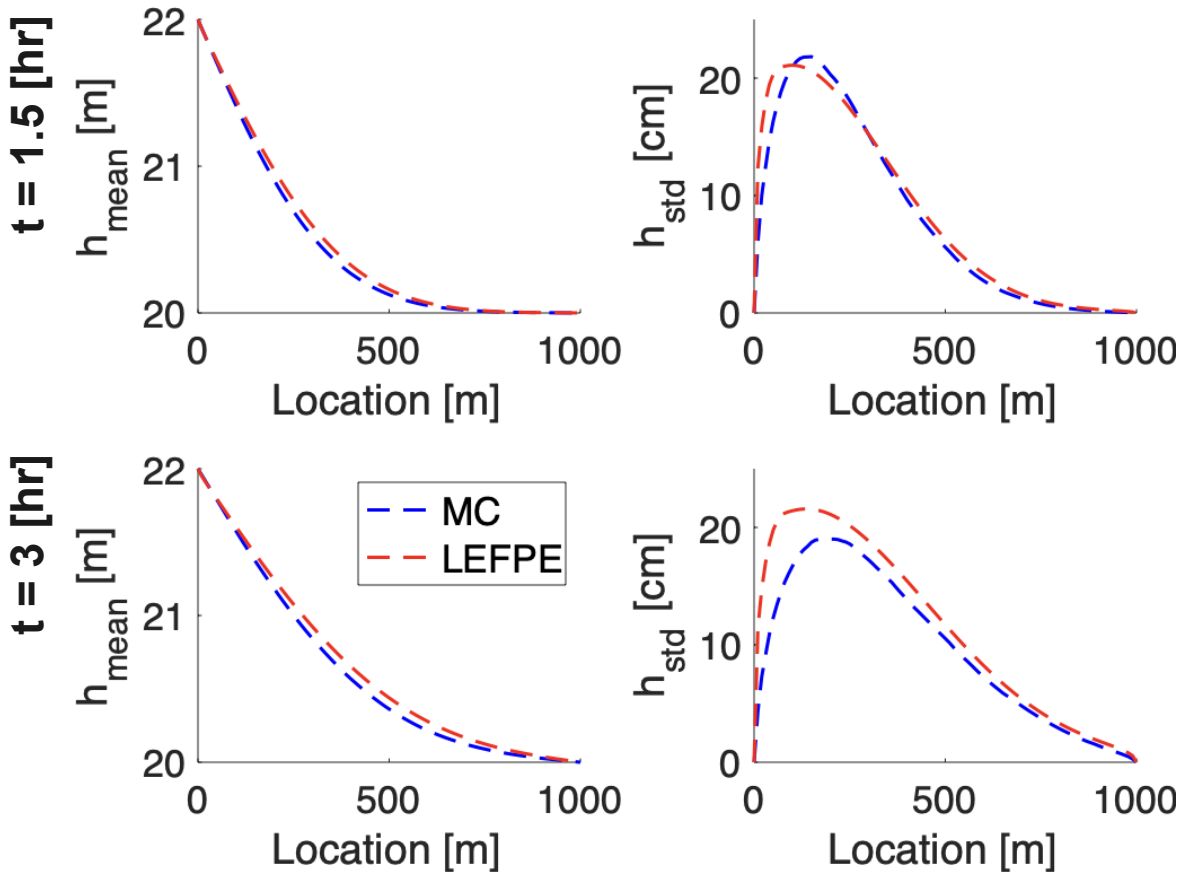


Figure 5.4: Comparison of the ensemble average and variability by LEFPE method and the MC approach as a function of location at different times for Application I.

In addition, at the beginning of the simulation ($t = 0$) the hydraulic gradient is the highest ($\Delta h = 22[m] - 20[m]$) during the simulation and gradually decreases in time as time goes on. Therefore, the flux obtained by the Darcy equation (equation 2.8) will also have higher variability, decreasing gradually as the hydraulic gradient does. Therefore, the resulting fluxes have a high variability too. This high variability in the fluxes produced by these mentioned factors generates relevant discrepancies in the hydraulic head for early times next to the B.C., resulting in a high standard deviation value. On the other hand, the

LEFPE approach cannot constrain these high standard deviation values at the beginning of the simulation. Similar issues were faced by Diaz (2019), where the standard deviation estimation suffers from a lack of accuracy next to the boundary condition.

Figure (5.4) shows the ensemble mean and standard deviation for application I at simulation time $t = \{T/2; T\}$. The left panel in this figure indicates that the mean hydraulic head computed by the LEFPE methodology matches the MC results, with a slight underestimation in the area with the higher curvature. For the standard deviation, the general behavior is well caught as well. However, the right panel of the figure (5.4), which shows the standard deviation, confirms an overestimation of this parameter. Furthermore, a higher deviation from the MC simulation occurs next to the boundary condition. The possible cause of these differences between both approaches is assessed in the following application, which includes a correlation structure in its transmissivity field.

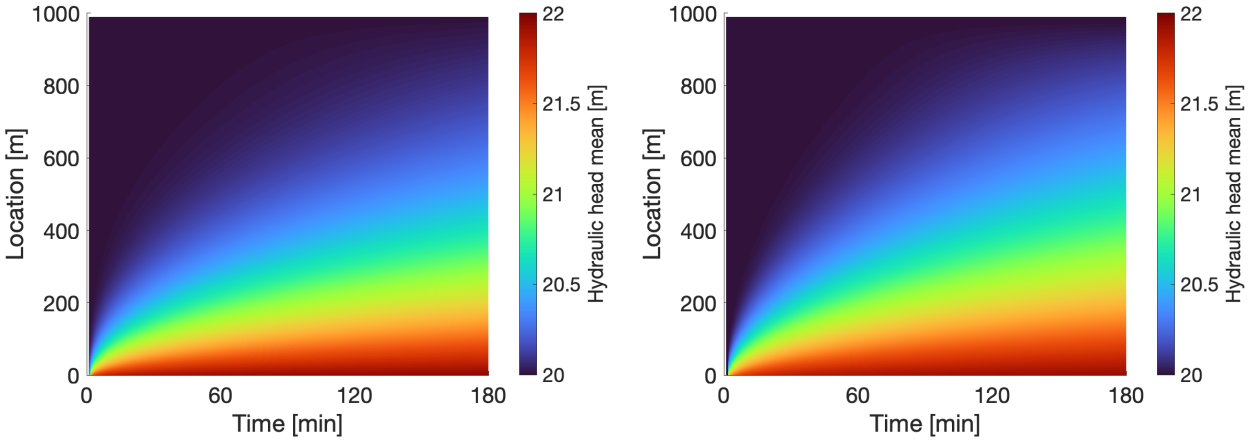
5.2 Application II - One-dimensional correlated transmissivity case

For application II, a hypothetical problem similar to application one was chosen. Nevertheless, the only difference is the correlation structure in the transmissivity random field. While application I assumes a lack of correlation, application II includes the correlation structure in the transmissivity field. This numerical experiment used an exponential model to include this structure, with a correlation length equal to $L_x = 100\ [m]$, represented by equation (4.1).

Furthermore, the coefficient of variation is the same in the experiment performed in application I ($CV = 1.5$). Thus, it is possible to explore how different the aquifer's probabilistic behavior is under both correlated and uncorrelated transmissivity random fields. The statis-

tics of the physical parameters and the model properties are presented in Table (5.1).

Figure (5.5) shows the computed hydraulic head's mean of the LEFPE and MC approaches. These figures indicate a satisfactory agreement, similar to application I. However, the estimation obtained from the LEFPE methodology tends to slightly overestimate the hydraulic head (light blues) next to the boundary at $x = 1000[m]$. This overestimation occurs due to the numerical diffusion coming from that boundary, which moves the probability mass through the positive direction of the $h - axis$, since h domain is limited by $19.9\ [m] < h < 22.1\ [m]$.

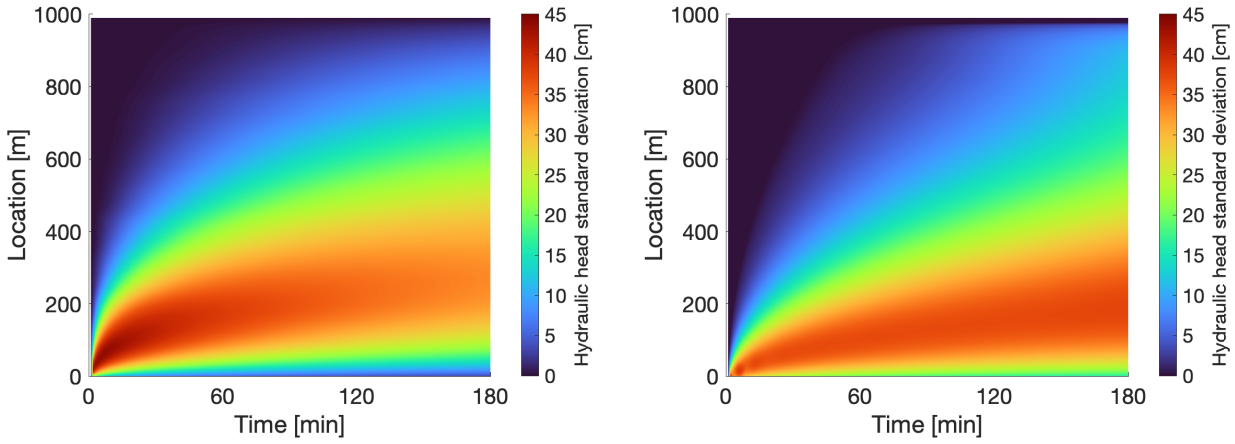


(a) MC (b) FPE

Figure 5.5: Plan view of the mean hydraulic head over position and time for the results obtained by the LEFPE method and those obtained by the MC method for application II.

The corresponding standard deviation profiles for this case are shown in figure (5.6). The shapes of the standard deviation profiles by the LEFPE are spatially and temporally consistent with those by the MC approach. However, the hydraulic head variability is overestimated by the LEFPE methodology, being this deviation more important next to the boundaries. Even though the LEFPE was derived by employing a second-order cumulant

expansion, it is essential to note that it tends to smooth its solutions (Ohara, 2003). This behavior of the LEFPE reported by Ohara (2003) makes the PDF flatter, which artificially increases the standard deviation estimation.



(a) hydraulic head' standard deviation - MC

(b) hydraulic head' standard deviation - FPE

Figure 5.6: Plan view of the standard deviation of the Hydraulic head over position and time for the results obtained by the LEFPE method and those obtained by the MC method for application II.

Figure (5.7) shows the hydraulic head's ensemble average and standard deviation. These plots show the mean and the standard deviation behavior at simulation time equal to $t = \{T/2; T\}$ across the space domain. The ensemble average behavior represented by the left panel from the figure (5.7) indicates an overestimation by the LEFPE methodology against the MC approach. This overestimation suggests higher water fluxes through the $x - direction$. In the probability space $P - x - h$, the advection coefficient B^h (Equation 3.15) in the $h - direction$ could be overestimated. Advection correction terms (showed in equation (3.10) could explain the increment of this parameter. Since they were neglected, assuming low correlations between their variables, the advection correction term could be slightly incremented. Hence, the artificial increments generated by this simplification could explain the overestimation of the hydraulic head's mean estimation.

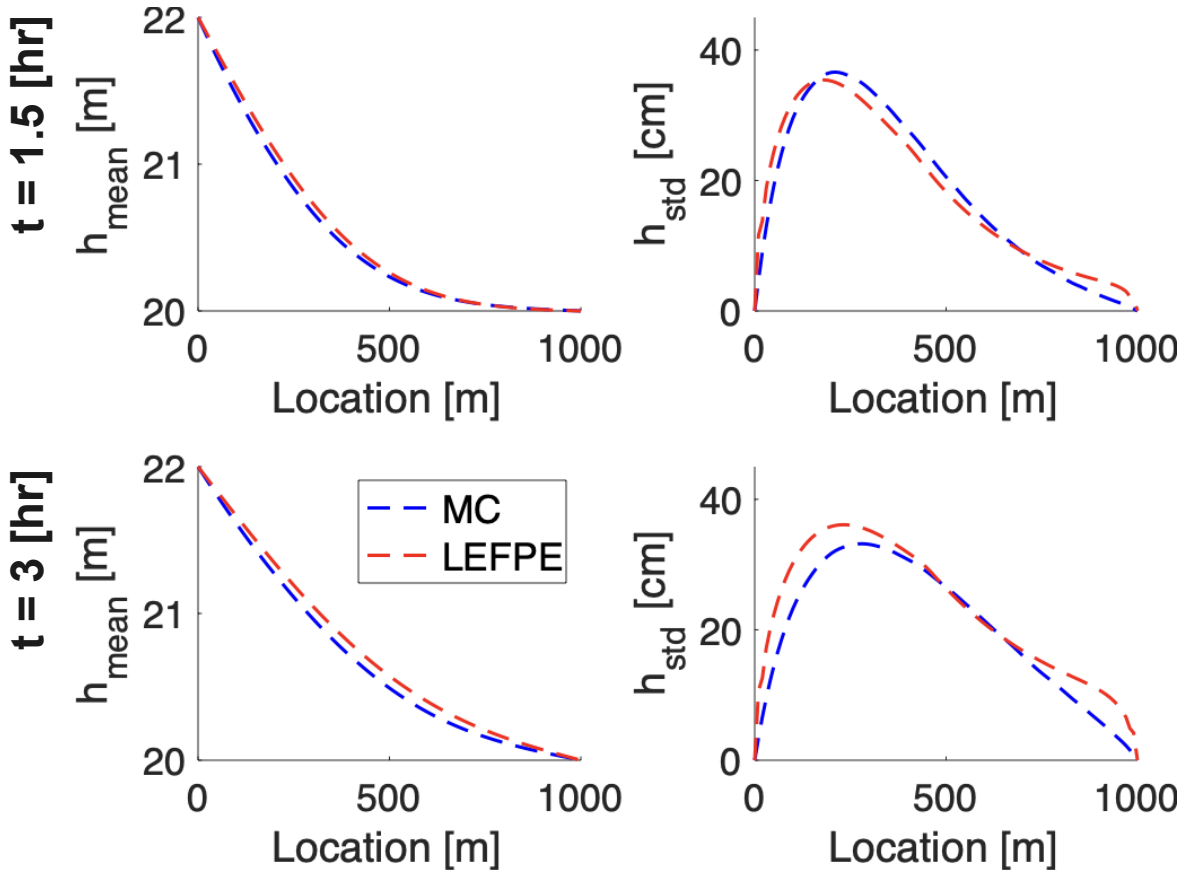


Figure 5.7: Comparison of the ensemble average and variability by LEFPE method and the MC approach as a function of location at different times for Application II.

The right panel of figures (5.7) compares the standard deviations of hydraulic heads computed by Monte Carlo and FPE approaches at two different times. These figures show that LEFPE can constrain its evolution but overestimate this statistical parameter. Moreover, the overestimation is exacerbated next to the boundaries. The increment in overestimation at the boundaries can be attributed to the numerical scheme. Even though flux limiters were used to solve the advection part of the LEFPE, there is still numerical diffusion when high gradients are faced. Since the boundary conditions in the probability domain were represented as a Dirac delta pulse, high gradients are inherent in the boundary neighborhood.

The comparison between applications 1 and 2 results in interesting findings regarding their ensemble hydraulic head. The correlated and uncorrelated random field presents a similar behavior to their time and space evolution. This similarity indicates that the correlation or lack of correlation in the transmissivity field does not considerably affect the ensemble behavior of the system. The LEFPE framework well captures this behavior. Nevertheless, there are differences in terms of their variability. Figures (5.3) and (5.6) show that the lack of correlation structure in the transmissivity random field underestimates the hydraulic head's standard deviation.

This difference can be qualitatively explained by the correlation structure in the transmissivity field. Relative points in this field should be more similar in a correlated random field, presenting more minor discrepancies between each other. If we think from a groundwater perspective, every transmissivity sample field can be interpreted as "preferential paths", generating preferential flow. On the other hand, an uncorrelated transmissivity field, in which values have no relation to their neighborhood, would smooth the conductance between cells, averaging the flux between them. Therefore, the fluxes would be more uniform, reducing the variability in the hydraulic head.

In general terms, results from figures (5.2) to (5.7) indicate that the proposed stochastic framework performs satisfactorily in predicting the ensemble hydraulic head mean and variability. Results from Applications I and II suggest that the spatio-temporal evolution of these statistical parameters is well simulated. However, the LEFPE methodology tends to overestimate the standard deviation. This overestimation is addressed in more detail in Application III.

5.3 Application III - Two-dimensional extension of application II

Application III consists of a two-dimensional extension of application II, represented by figure (5.8). The physical and statistical parameters have similar characteristics as application II but are extended to the $x-y$ domain. The boundary conditions at $x = y = 0$ are considered deterministic Dirichlet B.C, while at $x = y = 2000[m]$ are assumed as absorbing boundaries to limit the area of computation. An exponential correlation model was used to represent the transmissivity random field, shown in equation (4.1). Nevertheless, the correlation is independent along each spatial dimension. This independence in the transmissivity field occurs by the use of independent SPDE in the x and y direction, which is inconvenient for the used transmissivity generation model. The parameters used to carry out this application can be found in the table (5.2).

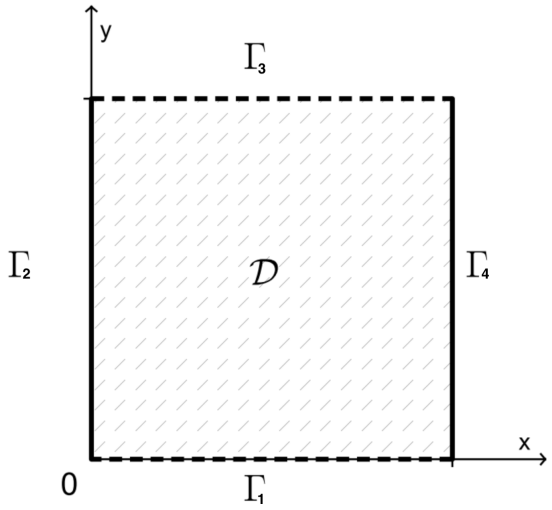


Figure 5.8: Illustration of the 2D physical domain for Application III.

Figure (5.10) show the hydraulic head's ensemble average at the simulation time $t =$

$\{T/3; 2T/3; T\}$ for the two-dimensional case. These figures indicate water fluxes through the x and y axes. Fluxes are due to the higher hydraulic gradient next to the boundaries at $x = 0$ and $y = 0$. The comparison shows a good match in terms of the ensemble mean of the hydraulic head. However, it is possible to note an overestimation next to $x = y = 2000[m]$. For $t = 1$ $[h]$, the hydraulic head follows an "L" pattern, but as time goes on, it changes to a "C" pattern. This behavior is probably due to the influence of absorbing boundaries, which tend to advect probability mass through the positive direction of the h-axis, increasing the ensemble mean at the boundaries.

Parameter	Symbol	Value
Transmissivity mean	$\mu_{T_x} = \mu_{T_y}$	$3 \cdot 10^{-2}[m^2/s]$
Coefficient of Variation	CV	1.5
correlation length$^{(2)}$	$L_x = L_y$	100 [m]
Specific storage	S_S	10^{-4}
Aquifer length	$L_x = L_y$	2000 [m]
Aquifer width	B	15 [m]
I.C. at t=0	h_0	20 [m]
B.C. at x=0	Γ_1	22 [m]
B.C. at y=0	Γ_2	22 [m]
B.C. at y=L	Γ_3	absorbing
B.C. at x=L	Γ_4	absorbing
grid size	$\Delta x = \Delta y$	5 [m]
h-grid size$^{(1)}$	$\Delta h^{(1)}$	1 [cm]
Simulation time	T	3 [hours]

$^{(1)}$ It just applies to the LEFPE methodology

Table 5.2: Physical parameters values for application III.

Regarding the hydraulic head variability, the standard deviation is shown in figure (5.10) as a function of the location at specific times. This figure indicates that the LEFPE methodology correctly portrays the ensemble variability, providing a good representation of the

patterns. Nevertheless, this statistical parameter is again overestimated next to the boundaries. Moreover, this simulation shows standard deviations close to 6 [cm]. Instead, the one-dimensional case of application II shows values close to 45 [cm]. This global standard deviation reduction is attributed to the increment of the spatial dimensions with respect to application II. However, the inclusion of extra spatial dimensions implies an extra drift term, which add extra numerical diffusion to the system.

Figures (5.11) and (5.12) show the ensemble average and variability for the hydraulic head at three different times at lines $x = y$ and $x = 2y$, respectively. From the left panel of both figures, it can be easily seen that the proposed stochastic framework's mean behavior of the hydraulic head and the Monte Carlo simulation indicate good matches, improving this estimation considerably compared to application II. However, at line $x = 2y$, this difference is more important than the mean at line $x = y$. In the variability (right panel) case, the LEFPE methodology shows excellent agreement with only light deviations concerning the MC simulations. Moreover, it is possible to note how the variability decreases as time goes on, and the LEFPE methodology predicts that behavior exceptionally. Therefore, the proposed stochastic approach can constrain the standard deviation reduction in time for the two-dimensional case. However, the excellent agreement loses quality as we walk away from $x = y$ through the boundaries. Indicating how sensitive is this method to the boundary conditions.

The overestimation of the ensemble variability of the proposed LEFPE methodology over the MC simulations can be attributed to three factors. The first one is associated with the (1) derivation of the LEFPE. Even though a second-order cumulant expansion was used for its derivation, the solution tends to be smooth (Ohara, 2003). Second, (2) the numerical

diffusion inherent to any numerical model, especially with advection schemes. The previous factors mentioned diffusing the state variable P, making it flatter and consequently increasing its variability. In addition, (3) modeling deterministic conditions as Dirac delta pulses introduce high gradients into the system. These high gradients result in extra numerical diffusion in these areas compared to zones far from the boundaries.

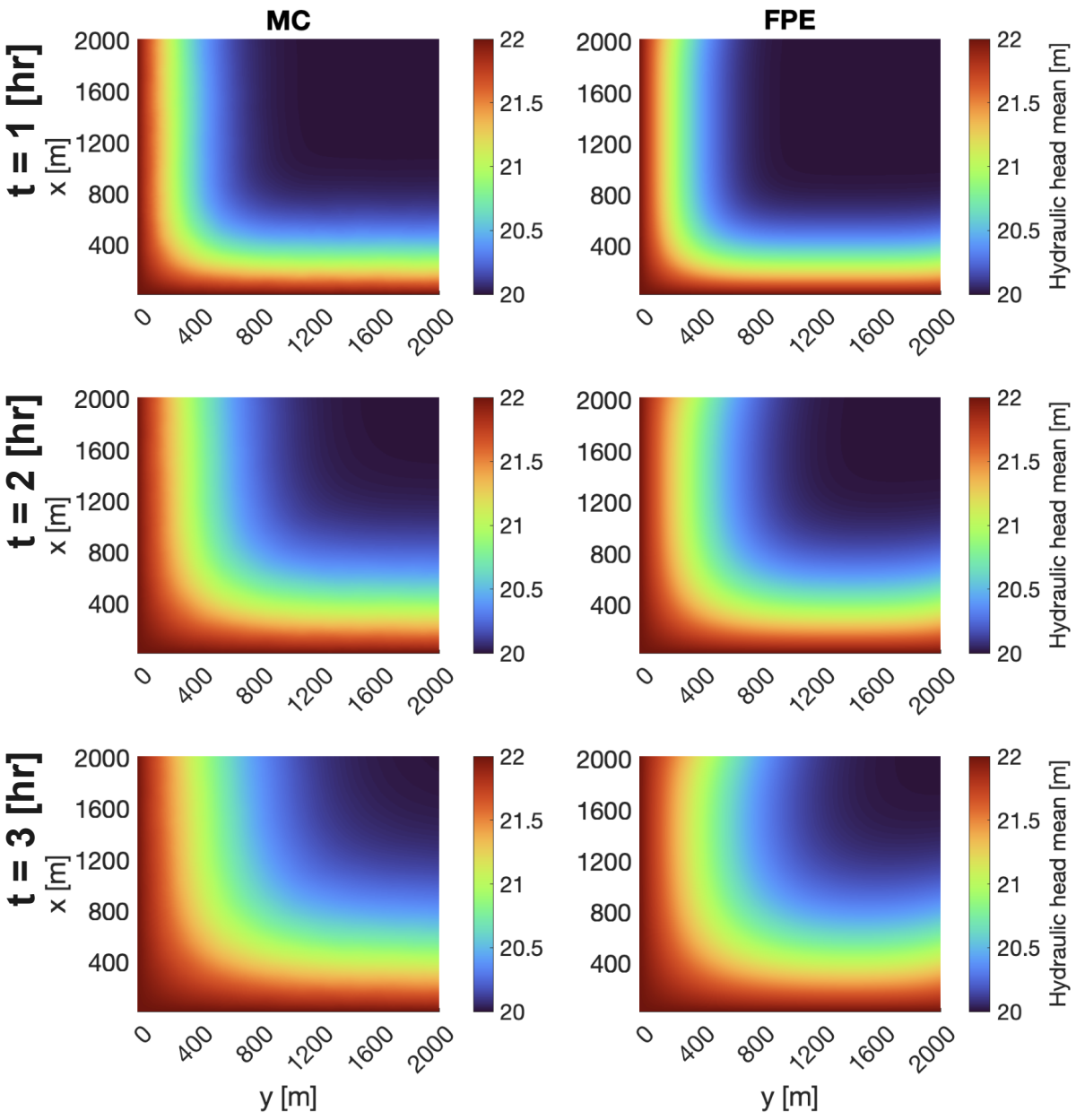


Figure 5.9: Plan view of the Hydraulic head's ensemble average over position and time for the results obtained by the LEFPE method and those obtained by the MC method for application III.

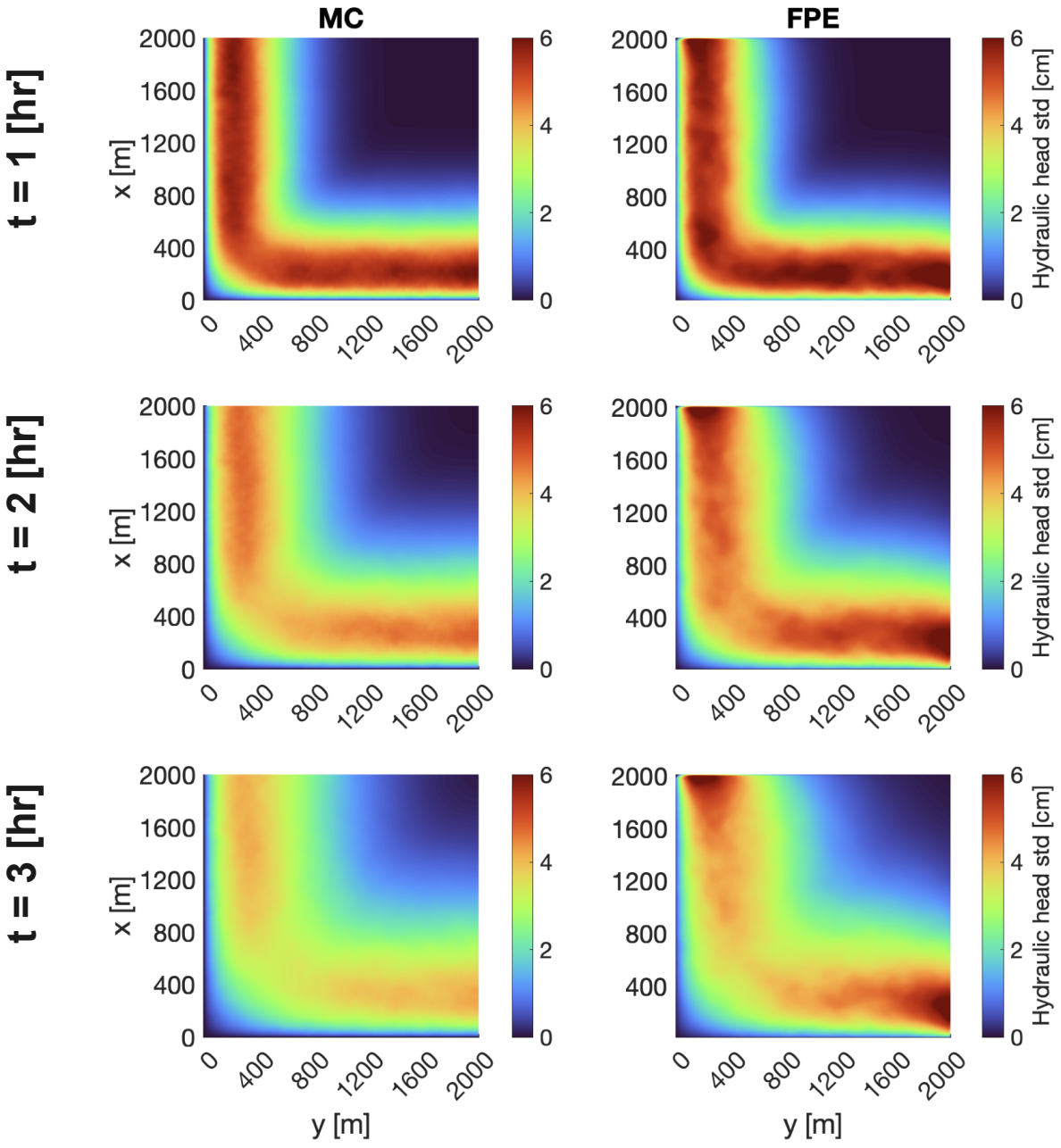


Figure 5.10: Plan view of the Hydraulic head's ensemble variability over position and time for the results obtained by the LEFPE method and those obtained by the MC method for application III.

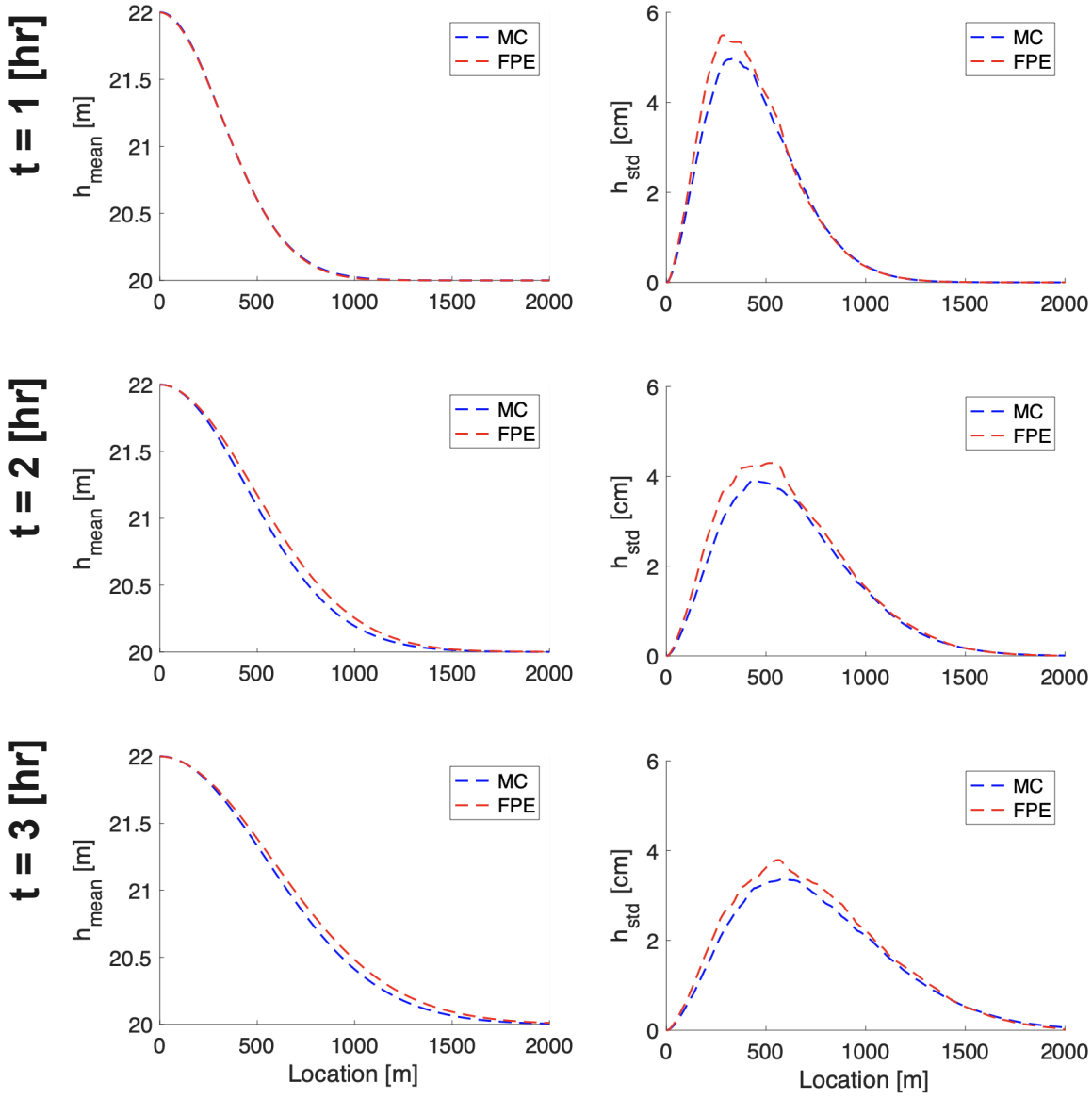


Figure 5.11: Comparison of the ensemble average and variability by LEFPE method and the MC approach at location $x = y$ at different times.

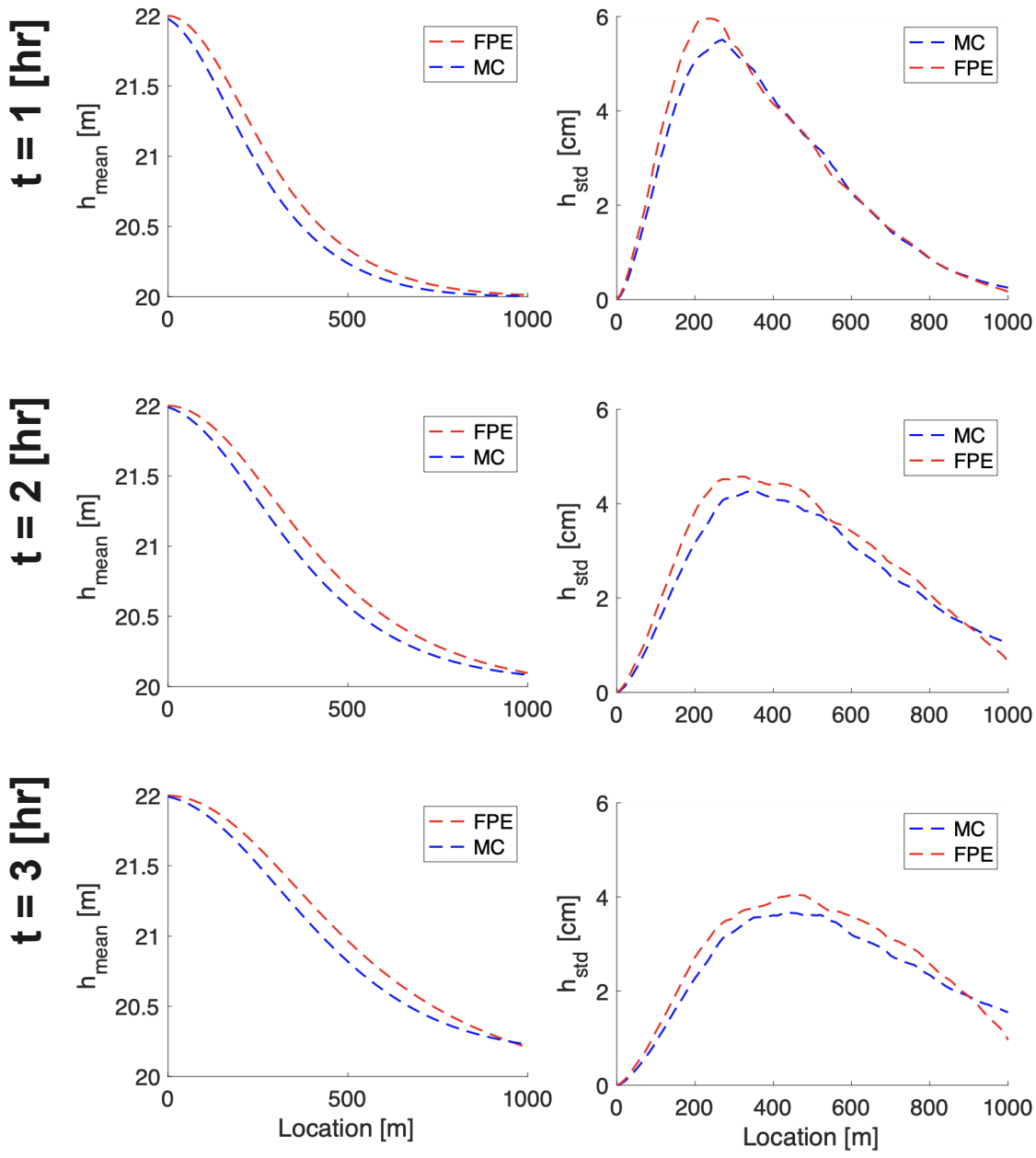


Figure 5.12: Comparison of the ensemble average and variability by LEFPE method and the MC approach at location $x = y$ at different times.

Chapter 6

Summary and Conclusions

Due to the heterogeneity in geomaterials and the stochasticity of the source process, the governing subsurface flow equation in confined aquifers at the point-scale becomes a stochastic PDE at the field scale. Therefore, this governing equation needs to be upscaled to the corresponding field scale to predict its behavior correctly. One of the most popular approaches to carry out this task has been developing their ensemble average forms. Although the MC and Perturbation methods have been widely used for upscaling the groundwater flow equation in confined aquifers, they have essential drawbacks regarding computational cost and small variance constraints.

For this purpose, this study developed a second-order expression for the mean and probabilistic behavior of the groundwater flow equation in confined aquifers. The resulting expression (equation 3.12) is a deterministic PDE in the form of a Lagrangian-Eulerian Fokker-Planck equation (LEFPE), while the original governing equation, described by equation (9), was a stochastic PDE. The solution of the LEFPE under the appropriate initial and boundary conditions describes the time-space evolution of the hydraulic head's PDF. Therefore, it is possible to determine the ensemble averages and variances of the hydraulic head in con-

fined aquifers in one shot, unlike numerical analysis based on a Monte Carlo approach. The ensemble average is then obtained from a standard expectation operation using the system's already calculated PDF.

Obtaining the probabilistic ensemble behavior by a single simulation is one of the most advantageous characteristics of the proposed methodology. This characteristic is fundamental in (1) high-dimensional problems and (2) simulations where tails' responses are essential to estimate accurately. Thus, the computational time could be reduced significantly in both cases since larger sample sizes are required for Monte Carlo simulations. Furthermore, the LEFPE approach does not suffer from the "closure problem" associated with the traditional perturbation approach or limitations associated with high random parameter variability.

The resulting linear and deterministic LEFPE was discretized by using the explicit ULTIMATE QUICKEST algorithm (Leonard, 1991) for the advective terms, while an implicit second-order centered difference approximation was used for diffusive terms. Thus, the LEFPE is solved numerically using an implicit-explicit (IMEX) scheme to determine the confined aquifer system's ensemble behavior and variability. The election of the numerical scheme was based on two points. The scheme needed the potential to be parallelized to speed up the calculations and to model large domains. In addition, the numerical scheme had to control the numerical dispersion. The latter is a crucial factor since extra diffusion in the system would artificially increase its variance estimation.

Results obtained by solving the LEFPE were compared against the Monte-Carlo solutions to evaluate the performance of the proposed methodology. Despite minor differences, the comparison showed a good agreement in terms of the ensemble mean and variance in

the solution domain, characterizing the first and second moments of the hydraulic head in confined aquifers. Thus, the LEFPE solution gives a good representation of the general patterns, the decreasing and increasing trend, and the ranges of the ensemble mean and variability.

Even though the proposed stochastic framework can effectively capture the shapes of the standard deviation behavior, the ensemble variability tends to be overestimated by the LEFPE methodology. Three reasons can explain the overestimation of this statistical parameter. (1) the second-order cumulant expansion used to derive the LEFPE, (2) the numerical diffusion inherent to any numerical model, and (3) the representation of deterministic conditions by means of Delta pulses.

The developed methodology can also include variability in the storativity and boundary conditions. Subsequently, this stochastic framework shows great promise in dealing with high-dimensional problems and including correlated structures in the random field variables.

Bibliography

Ablowitz, M. J., Herbst, B. M., & Schober, C. (1996). On the numerical solution of the sine–gordon equation: I. integrable discretizations and homoclinic manifolds. *Journal of Computational Physics*, *126*, 299–314. https://doi.org/10.1006/JCPH.1996.0139

Ahmed, S., & Marsily, G. D. (1987). Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. *Water Resources Research*, *23*, 1717–1737. https://doi.org/10.1029/WR023i009p01717

Alecsa, C. D., Boros, I., Frank, F., Knabner, P., Nechita, M., Prechtel, A., Rupp, A., & Suciu, N. (2020). Numerical benchmark study for flow in highly heterogeneous aquifers. *Advances in Water Resources*, *138*, 103558. https://doi.org/10.1016/j.advwatres.2020.103558

Anderson, M. P., & Woessner, W. W. (2015). *Applied groundwater modelling: Simulation of flow andadvective transport* (M. P. Anderson, W. W. Woessner, & R. J. Hunt, Eds.; Second Edi). https://doi.org/https://doi.org/10.1016/B978-0-08-091638-5.00019-5

Ascher, U. M., Ruuth, S. J., & Wetton, B. T. R. (1995). Implicit-explicit methods for time-dependent partial differential equations. *Source: SIAM Journal on Numerical Analysis*, *32*, 797–823.

Ascher, U. M., Ruuth, S. J., & Wetton, B. T. R. (2006). Implicit-explicit methods for time-dependent partial differential equations. *http://dx.doi.org/10.1137/0732037*, *32*, 797–823. https://doi.org/10.1137/0732037

Baalousha, H. (2008). Fundamentals of groundwater modelling. *Groundwater: Modelling, Management and Contamination. New York: Nova Science Publishers, Inc.*

Bakshevskaya, V. A., & Pozdnyakov, S. P. (2013). Methods of modeling hydraulic heterogeneity of sedimentary formations. *Water Resources*, *40*, 767–775. https://doi.org/10.1134/S0097807813070026

Baye, A. Y., Razack, M., Ayenew, T., & Zemedagegnehu, E. (2013). Estimating transmissivity using empirical and geostatistical methods in the volcanic aquifers of upper awash basin, central ethiopia. *Environmental Earth Sciences*, *69*, 1791–1802. https://doi.org/10.1007/s12665-012-2011-6

Bear, J., & Verruijt, A. (1987). Modeling groundwater flow and pollution. https://doi.org/10.1007/978-94-009-3379-8

Bohling, G. C., Liu, G., Dietrich, P., & Butler, J. J. (2016). Reassessing the made direct-push hydraulic conductivity data using a revised calibration procedure. *Water Resources Research*, *52*, 8970–8985. https://doi.org/10.1002/2016WR019008

Bredehoeft, J. D. (2003). From models to performance assessment: The conceptualization problem. *Ground Water*, *41*, 571–577. https://doi.org/10.1111/j.1745-6584.2003.tb02395.x

Buet, C., & Dellacherie, S. (2010). On the chang and cooper scheme applied to a linear fokker-planck equation. *Communications in Mathematical Sciences*, *8*, 1079–1090. https://doi.org/10.4310/CMS.2010.V8.N4.A15

Buet, C., & Thanh, K.-C. L. (2007). Positive, conservative, equilibrium state preserving and implicit difference schemes for the isotropic fokker-planck-landau equation. https:

//hal.archives-ouvertes.fr/hal-00142408%20https://hal.archives-ouvertes.fr/hal-00142408/document

Bürger, R., Gavilán, E., Inzunza, D., Mulet, P., & Villada, L. M. (2020). Implicit-explicit methods for a convection-diffusion-reaction model of the propagation of forest fires. *Mathematics 2020, Vol. 8, Page 1034*, *8*, 1034. https://doi.org/10.3390/MATH8061034

Butt, M. M. (2021). Two-level difference scheme for the two-dimensional fokker–planck equation. *Mathematics and Computers in Simulation*, *180*, 276–288. https://doi.org/10.1016/J.MATCOM.2020.09.001

Cahyono, M. (1993). *Three-dimensional numerical modelling of sediment transport processes in non-stratified estuarine and coastal waters*. The University of Bradford. The University of Bradford.

Carle, S. F., & Fogg, G. E. (1996). Transition probability-based indicator geostatistics. *Mathematical Geology*, *28*, 453–476. https://doi.org/10.1007/bf02083656

Castagna, M., Becker, M. W., & Bellin, A. (2011). Joint estimation of transmissivity and storativity in a bedrock fracture. *Water Resources Research*, *47*, 9504. https://doi.org/10.1029/2010WR009262

Casulli, V. (1990). Semi-implicit finite difference methods for the two-dimensional shallow water equations. *JCoPh*, *86*, 56–74. https://doi.org/10.1016/0021-9991(90)90091-E

Cayar, M., & Kavvas, M. L. (2009). Ensemble average and ensemble variance behavior of unsteady, one-dimensional groundwater flow in unconfined, heterogeneous aquifers: An exact second-order model. *Stochastic Environmental Research and Risk Assessment*, *23*, 947–956. https://doi.org/10.1007/s00477-008-0263-1

Ceyhan, M. S., & Kavvas, M. L. (2018). Ensemble modeling of the theis equation under uncertain parameter conditions. *Journal of Hydrologic Engineering*, *23*, 04018011. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001632

Chandra, R., Dagum, L., Kohr, D., Menon, R., Maydan, D., & McDonald, J. (2001). *Parallel programming in openmp*. Morgan kaufmann.

Chang, J. S., & Cooper, G. (1970). A practical difference scheme for fokker-planck equations. *Journal of Computational Physics*, *6*, 1–16. https://doi.org/https://doi.org/10.1016/0021-9991(70)90001-X

Chaudhry, J. H., Estep, D., Ginting, V., Shadid, J. N., & Tavener, S. (2015). A posteriori error analysis of imex multi-step time integration methods for advection–diffusion–reaction equations. *Computer Methods in Applied Mechanics and Engineering*, *285*, 730–751. https://doi.org/10.1016/J.CMA.2014.11.015

Connell, L. D. (1995). An analysis of perturbation based methods for the treatment of parameter uncertainty in numerical groundwater models. *Transport in Porous Media*, *21*, 225–240. https://doi.org/10.1007/BF00617407

Dagan, G. (1982). Stochastic modeling of groundwater flow by unconditional and conditional probabilities: 1. conditional simulation and the direct problem. *Water Resources Research*, *18*, 813–833. https://doi.org/10.1029/WR018i004p00813

Dagan, G. (1986). Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale. *Water Resources Research*, *22*, 120S–134S. https://doi.org/10.1029/WR022i09Sp0120S

Dagan, G. (2002). An overview of stochastic modeling of groundwater flow and transport: From theory to applications. *Eos, Transactions American Geophysical Union*, *83*, 621. https://doi.org/10.1029/2002EO000421

Dagan, G., & Rubin, Y. (1988). Stochastic identification of recharge, transmissivity, and storativity in aquifer transient flow' a quasi-steady approach. *WATER RESOURCES RESEARCH*, *24*.

Dawson, C. (1995). High resolution upwind-mixed finite element methods for advection-diffusion equations with variable time-stepping. *Numerical Methods for Partial Differential Equations*, *11*, 525–538. https://doi.org/10.1002/NUM.1690110508

de Brito Fontenele, S., Mendonça, L. A. R., de Araújo, J. C., Santiago, M. M. F., & de Brito Gonçalves, J. Y. (2014). Relationship between hydrogeological parameters for data-scarce regions: The case of the araripe sedimentary basin, brazil. *Environmental Earth Sciences*, *71*, 885–894. https://doi.org/10.1007/s12665-013-2491-z

Dehghan, M. (2004). Weighted finite difference techniques for the one-dimensional advection–diffusion equation. *Applied Mathematics and Computation*, *147*, 307–319. https://doi.org/10.1016/S0096-3003(02)00667-7

Diaz, A. (2019). *Ensemble modeling of unsteady, one-dimensional subsurface stormflow with uncertain hillslope parameters*. University of California, Davis.

Dib, A., & Kavvas, M. L. (2018). Ensemble modeling of stochastic unsteady open-channel flow in terms of its time–space evolutionary probability distribution – part 2: Numerical application. *Hydrology and Earth System Sciences*, *22*, 2007–2021. https://doi.org/10.5194/hess-22-2007-2018

Dogrul, E. C., Kavvas, M. L., & Chen, Z. (1998). Prediction of subsurface stormflow in heterogeneous sloping aquifers. *Journal of Hydrologic Engineering*, *3*, 258–267. https://doi.org/10.1061/(ASCE)1084-0699(1998)3:4(258)

Domenico, P. A., & Schwartz, F. W. (1998). *Physical and chemical hydrogeology, 2nd edn john wiley sons*.

Ercan, A., & Kavvas, M. L. (2012). Ensemble modeling of hydrologic and hydraulic processes at one shot: Application to kinematic open-channel flow under uncertain channel properties by the stochastic method of characteristics. *Journal of Hydrologic Engineering*, *17*, 168–181. https://doi.org/10.1061/(asce)he.1943-5584.0000425

Fleckenstein, J. H., Niswonger, R. G., & Fogg, G. E. (2006). River-aquifer interactions, geologic heterogeneity, and low-flow management. *Ground Water*, *44*, 837–852. https://doi.org/10.1111/j.1745-6584.2006.00190.x

Fok, J., Guo, B., & Tang, T. (2002). Combined hermite spectral-finite difference method for the fokker-planck equation. *Mathematics of Computation*, *71*, 1497–1528. https://doi.org/10.1090/S0025-5718-01-01365-5

Freeze, R. A. (1975). A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research*, *11*, 725–741. https://doi.org/10.1029/WR011i005p00725

Galán, R. F., Ermentrout, G. B., & Urban, N. N. (2007). Stochastic dynamics of uncoupled neural oscillators: Fokker-planck studies with the finite element method. *Physical Review E*, *76*, 056110. https://doi.org/10.1103/PhysRevE.76.056110

Gardiner. (2009). *Stochastic methods* (Vol. 4). Springer Berlin.

Gelhar, L. W. (1986). Stochastic subsurface hydrology from theory to applications. *Water Resources Research*, *22*, 135S–145S. https://doi.org/10.1029/WR022i09Sp0135S

Gelhar, L. W., & Axness, C. L. (1983). Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, *19*, 161–180. https://doi.org/10.1029/WR019i001p00161

Gotovac, H., Cvetković, V., & Andričević, R. (2009). Adaptive fup multi-resolution approach to flow and advective transport in highly heterogeneous porous media: Methodology, accuracy and convergence. *Advances in Water Resources*, *32*, 885–905. https://doi.org/10.1016/j.advwatres.2009.02.013

Gray, W. G. (1984). Comparison of finite difference and finite element methods. In J. Bear & M. Y. Corapcioglu (Eds.), *Fundamentals of transport phenomena in porous media* (pp. 899–952). Springer Netherlands. https://doi.org/10.1007/978-94-009-6175-3_18

Harris, E., Falconer, R. A., Kay, D., & Stapleton, C. (2002). Development of a modelling tool to quantify faecal indicator levels in cardiff bay. *Proceedings of the Institution of Civil Engineers: Water and Maritime Engineering*, *154*, 129–135. https://doi.org/10.1680/WAME.2002.154.2.129

Hassan, A. E., Cushman, J. H., & Delleur, J. W. (1998). A monte carlo assessment of eulerian flow and transport perturbation models. *Water Resources Research*, *34*, 1143–1163. https://doi.org/10.1029/98WR00011

Hayley, K. (2017). The present state and future application of cloud computing for numerical groundwater modeling. *Groundwater*, *55*, 678–682. https://doi.org/10.1111/GWAT.12555

Hughes, J. D., Langevin, C. D., & Banta, E. R. (2017). *Documentation for the modflow 6 framework: U.s. geological survey techniques and methods*. https://doi.org/10.3133/TM6A57

Jafari, F., Javadi, S., Golmohammadi, G., Mohammadi, K., Khodadadi, A., & Mohammadzadeh, M. (2016). Groundwater risk mapping prediction using mathematical modeling and the monte carlo technique. *Environmental Earth Sciences*, *75*, 1–11. https://doi.org/10.1007/S12665-016-5335-9/FIGURES/13

Jeremić, B., Sett, K., & Kavvas, M. L. (2007). Probabilistic elasto-plasticity: Formulation in 1d. *Acta Geotechnica 2007 2:3*, *2*, 197–210. https://doi.org/10.1007/S11440-007-0036-X

Johnson, E. A., Wojtkiewicz, S. F., Bergman, L. A., & Spencer, B. F. (1997). Observations with regard to massively parallel computation for monte carlo simulation of stochastic dynamical systems. *International Journal of Non-Linear Mechanics*, *32*, 721–734. https://doi.org/10.1016/S0020-7462(96)00097-2

Kampen, N. G. V. (1974). A cumulant expansion for stochastic linear differential equations. i. *Physica*, *74*, 215–238. https://doi.org/10.1016/0031-8914(74)90121-9

Karapiperis, K., Sett, K., Kavvas, M. L., & Jeremić, B. (2016). Fokker–planck linearization for non-gaussian stochastic elastoplastic finite elements. *Computer Methods in Applied Mechanics and Engineering*, *307*, 451–469. https://doi.org/10.1016/J.CMA.2016.05.001

Karatzas, G. P. (2017). Developments on modeling of groundwater flow and contaminant transport. *Water Resources Management 2017 31:10*, *31*, 3235–3244. https://doi.org/10.1007/S11269-017-1729-Z

Kashefipour, S. M., & Zahiri, J. (2010). Comparison of empirical equations' application in the advection-dispersion equation (ade) on sediment transport modelling. *World Applied Sciences Journal*, *11*, 1015–1024.

Kavvas, M. L., & Karakas, A. (1996). On the stochastic theory of solute transport by unsteady and steady groundwater flow in heterogeneous aquifers. *Journal of Hydrology*, *179*, 321–351. https://doi.org/https://doi.org/10.1016/0022-1694(95)02835-8

Kavvas, M. L., & Wu, J.-L. (2002). Conservation equations for solute transport by unsteady and steady flows in heterogeneous aquifers: The cumulant expansion/lie operator methodology. *Stochastic Methods in Subsurface Contaminant Hydrology*, 281–306. https://doi.org/10.1061/9780784405321.CH07

Keese, A. (2003). A review of recent developments in the numerical solution of stochastic partial differential equations (stochastic finite elements). https://doi.org/10.24355/DBBS.084-200511080100-583

Kim, S., Kavvas, M. L., & Yoon, J. (2005). Upscaling of vertical unsaturated flow model under infiltration condition. *Journal of Hydrologic Engineering*, *10*, 151–159. https://doi.org/10.1061/(asce)1084-0699(2005)10:2(151)

Kitanidis, P. K. (1997). Kitanidis, p. k. 1997. introduction to geostatistics. applications in hydrogeology. xx 249 pp. cambridge, new york, port chester, melbourne, sydney: Cambridge university press. price £55.00, 74.95($hardcovers$), 19.95, US29.95 ⋆◎♀ ⋆♪ o 5. *Geological Magazine*, *135*, 723–732.

Král, R., & Náprstek, J. (2017). Theoretical background and implementation of the finite element method for multi-dimensional fokker–planck equation analysis. *Advances in Engineering Software*, *113*, 54–75. https://doi.org/10.1016/J.ADVENGSOFT.2017.02.013

Lai, J., Luo, Y., & Ren, L. (2012). Numerical evaluation of depth effects of double-ring infiltrometers on soil saturated hydraulic conductivity measurements. *Soil Science Society of America Journal*, *76*, 867–875. https://doi.org/10.2136/sssaj2011.0048

Larsen, E. W., Levermore, C. D., Pomraning, G. C., & Sanderson, J. G. (1985). Discretization methods for one-dimensional fokker-planck operators. *Journal of Computational Physics*, *61*, 359–390. https://doi.org/10.1016/0021-9991(85)90070-1

Leonard, B. (1979). A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Computer Methods in Applied Mechanics and Engineering*, *19*, 59–98. https://doi.org/10.1016/0045-7825(79)90034-3

Leonard, B. (1991). The ultimate conservative difference scheme applied to unsteady one-dimensional advection. *Computer Methods in Applied Mechanics and Engineering*, *88*, 17–74. https://doi.org/10.1016/0045-7825(91)90232-U

Li, S., McLaughlin, D., & Liao, H.-S. (2003). A computationally practical method for stochastic groundwater modeling. *Advances in Water Resources*, *26*, 1137–1148. https://doi.org/https://doi.org/10.1016/j.advwatres.2003.08.003

Liang, L., & Kavvas, M. L. (2008). Modeling of solute transport and macrodispersion by unsteady stream flow under uncertain conditions. *Journal of Hydrologic Engineering*, *13*, 510–520. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:6(510)

Lin, B., & Falconer, R. A. (1997). Tidal flow and transport modeling using ultimate quickest scheme. *Journal of Hydraulic Engineering*, *123*, 303–314. https://doi.org/10.1061/(ASCE)0733-9429(1997)123:4(303)

Lin, Chen, Y.-W., Chang, L.-C., Yeh, M.-S., Huang, G.-H., & Petway, J. (2017). Groundwater simulations and uncertainty analysis using modflow and geostatistical approach with conditioning multi-aquifer spatial covariance. *Water*, *9*, 164. https://doi.org/10.3390/w9030164

Lu, Z., & Zhang, D. (2005). Analytical solutions to statistical moments for transient flow in two-dimensional, bounded, randomly heterogeneous media. *Water Resources Research*, *41*, 1–13. https://doi.org/10.1029/2004WR003389

Ma, X., Li, S., & Zhu, W. (2009). A new method in groundwater flow modeling. *Journal of Hydrodynamics*, *21*, 245–254. https://doi.org/10.1016/S1001-6058(08)60142-0

Mantoglou, A., & Gelhar, L. (1987). Capillary tension head variance, mean soil moisture content, and effective specific soil moisture capacity of transient unsaturated flow in stratified soils. *Water Resources Research*, *23*, 47–56. https://doi.org/10.1029/WR023I001P00047

Masud, A., & Bergman, L. (2005). Solution of the four dimensional fokker-planck equation: Still a challenge.

Meerschaert, M. M., Dogan, M., Dam, R. L. V., Hyndman, D. W., & Benson, D. A. (2013). Hydraulic conductivity fields: Gaussian or not? *Water Resources Research*, *49*, 4730–4737. https://doi.org/10.1002/wrcr.20376

Meier, P. M., Carrera, J., & Sánchez-Vila, X. (1998). An evaluation of jacob's method for the interpretation of pumping tests in heterogeneous formations. *Water Resources Research*, *34*, 1011–1025. https://doi.org/10.1029/98WR00008

Mil'shtejn, G. N. (1975). Approximate integration of stochastic differential equations. *Theory of Probability Its Applications*, *19*, 557–562. https://doi.org/10.1137/1119062

Morbidelli, R., Saltalippi, C., Flammini, A., Cifrodelli, M., Picciafuoco, T., Corradini, C., & Govindaraju, R. S. (2017). In situ measurements of soil saturated hydraulic conductivity: Assessment of reliability through rainfall-runoff experiments. *Hydrological Processes*, *31*, 3084–3094. https://doi.org/10.1002/hyp.11247

Mu, Y., Zhu, L., qing Shen, T., Zhang, M., & yuan Zha, Y. (2020). Influence of correlation scale errors on aquifer hydraulic conductivity inversion precision. *Water Science and Engineering*, *13*, 243–252. https://doi.org/10.1016/j.wse.2020.09.004

Neuman, S. P. (1990). Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water Resources Research*, *26*, 1749–1758. https://doi.org/10.1029/WR026i008p01749

Neumann, L. E., Unek, J. S., & Cook, F. J. (2011). Implementation of quadratic upstream interpolation schemes for solute transport into hydrus-1d. *Environmental Modelling and Software*, *26*, 1298–1308. https://doi.org/10.1016/j.envsoft.2011.05.010

Ohara, N., Kavvas, M. L., & Chen, Z. Q. (2008). Stochastic upscaling for snow accumulation and melt processes with pdf approach. *Journal of Hydrologic Engineering*, *13*, 1103–1118. https://doi.org/10.1061/(asce)1084-0699(2008)13:12(1103)

Ohara, N. (2003). *Numerical and stochastic upscaling of snowmelt process*. Universit of California, Davis.

Park, B., & Petrosian, V. (1996). Fokker-planck equations of stochastic acceleration: A study of numerical methods. *apjs*, *103*, 255. https://doi.org/10.1086/192278

Patriarche, D., Castro, M. C., & Goovaerts, P. (2005). Estimating regional hydraulic conductivity fields - a comparative study of geostatistical methods. *Mathematical Geology*, *37*, 587–613. https://doi.org/10.1007/s11004-005-7308-5

Pletcher, R., Tannehill, J., & Anderson, D. (1997). *Computational fluid mechanics and heat transfer, second edition*. Taylor & Francis. https://books.google.com/books?id=ZJPbtHeilCgC

Qian, Y., Wang, Z., & Zhou, S. (2019). A conservative, free energy dissipating, and positivity preserving finite difference scheme for multi-dimensional nonlocal fokker–planck equation. *Journal of Computational Physics*, *386*, 22–36. https://doi.org/10.1016/J.JCP.2019.02.028

Rahman, A., Tsai, F. T.-C., White, C. D., & Willson, C. S. (2008). Coupled semivariogram uncertainty of hydrogeological and geophysical data on capture zone uncertainty analysis. *Journal of Hydrologic Engineering*, *13*, 915–925. https://doi.org/10.1061/(asce)1084-0699(2008)13:10(915)

Rajaram, H. (2016). *Debates—stochastic subsurface hydrology from theory to practice: Introduction*. https://doi.org/10.1002/2016WR020066

Razack, M., & Lasm, T. (2006). Geostatistical estimation of the transmissivity in a highly fractured metamorphic and crystalline aquifer (man-danane region, western ivory coast). *Journal of Hydrology*, *325*, 164–178. https://doi.org/10.1016/j.jhydrol.2005.10.014

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

Renard, P. (2007). Stochastic hydrogeology: What professionals really need? *Ground Water*, *45*, 531–541. https://doi.org/10.1111/j.1745-6584.2007.00340.x

Rojas, R., Feyen, L., & Dassargues, A. (2008). Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and bayesian model averaging. *Water Resources Research*, *44*. https://doi.org/10.1029/2008WR006908

Rotzoll, K., & El-Kadi, A. I. (2008). Estimating hydraulic conductivity from specific capacity for hawaii aquifers, usa. *Hydrogeology Journal*, *16*, 969–979. https://doi.org/10.1007/s10040-007-0271-0

Rubin, Y., Chang, C.-F., Chen, J., Cucchi, K., Harken, B., Heße, F., & Savoy, H. (2018). Stochastic hydrogeology's biggest hurdles analyzed and its big blind spot. *Hydrol. Earth Syst. Sci*, *22*, 5675–5695. https://doi.org/10.5194/hess-22-5675-2018

Rubin, Y., & Dagan, G. (1989). Stochastic analysis of boundaries effects on head spatial variability in heterogeneous aquifers: 2. impervious boundary. *Water Resources Research*, *25*, 707–712. https://doi.org/10.1029/WR025I004P00707

Scharffenberg, W. A., & Kavvas, M. L. (2011). Uncertainty in flood wave routing in a lateral-inflow-dominated stream. *Journal of Hydrologic Engineering*, *16*, 165–175. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000298

Sepehrian, B., & Radpoor, M. K. (2015). Numerical solution of non-linear fokker–planck equation using finite differences method and the cubic spline functions. *Applied Mathematics and Computation*, *262*, 187–190. https://doi.org/10.1016/J.AMC.2015.03.062

Sett, K., Jeremić, B., & Kavvas, M. L. (2007). Probabilistic elasto-plasticity: Solution and verification in 1d. *Acta Geotechnica 2007 2:3*, *2*, 211–220. https://doi.org/10.1007/S11440-007-0037-9

Shu, C.-W., & LeVeque, R. J. (1991). Numerical methods for conservation laws. *Mathematics of Computation*, *57*, 875. https://doi.org/10.2307/2938728

Shukla, H. S., & Tamsir, M. (2018). An exponential cubic b-spline algorithm for multi-dimensional convection-diffusion equations. *Alexandria Engineering Journal*, *57*, 1999–2006. https://doi.org/https://doi.org/10.1016/j.aej.2017.04.011

Simpson, M. J., & Clement, T. P. (2003). Comparison of finite difference and finite element solutions to the variably saturated flow equation. *Journal of Hydrology*, *270*, 49–64. https://doi.org/10.1016/S0022-1694(02)00294-9

Smith, L., & Freeze, R. A. (1979). Stochastic analysis of steady state groundwater flow in a bounded domain: 2. two-dimensional simulations. *Water Resources Research*, *15*, 1543–1559. https://doi.org/10.1029/WR015i006p01543

Sudicky, E. A. (1986). A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resources Research*, *22*, 2069–2082. https://doi.org/10.1029/WR022i013p02069

Tayfur, G., & Kavvas, M. L. (1994). Spatially averaged conservation equations for interacting rillinterrill area overland flows. *Journal of Hydraulic Engineering*, *120*, 1426–1448. https://doi.org/10.1061/(ASCE)0733-9429(1994)120:12(1426)

Theis, C. (1935). The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage. *Transactions, American Geophysical Union*, *16*, 519–524. https://doi.org/10.1029/TR016i002p00519

Thiem, G. (1906). Hydrologische methoden. *Gebhardt, Leipzig.*

Todd, D. K. (1980). *Groundwater hydrology.* Wiley.

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

Townley, L. R., & Wilson, J. L. (1985). Computationally efficient algorithms for parameter estimation and uncertainty propagation in numerical models of groundwater flow. *Water Resources Research*, *21*, 1851–1860. https://doi.org/10.1029/WR021i012p01851

Tu, T., Ercan, A., & Kavvas, M. L. (2019). One-dimensional solute transport in open channel flow from a stochastic systematic perspective. *Stochastic Environmental Research and Risk Assessment*, *33*, 1403–1418. https://doi.org/10.1007/s00477-019-01699-7

Tu, T., Ercan, A., & Kavvas, M. L. (2020a). Modeling one-dimensional nonreactive solute transport in open channel flows under uncertain flow and solute loading conditions. *Journal of Hydrologic Engineering*, *25*, 04020035. https://doi.org/10.1061/(asce)he.1943-5584.0001957

Tu, T., Ercan, A., & Kavvas, M. L. (2020b). Probabilistic solution to two-dimensional stochastic solute transport model by the fokker-planck equation approach. *Journal of Hydrology*, *580*, 124250. https://doi.org/10.1016/j.jhydrol.2019.124250

Wang, H., Shu, C.-W., & Zhang, Q. (2015). Stability and error estimates of local discontinuous galerkin methods with implicit-explicit time-marching for advection-diffusion problems. *http://dx.doi.org/10.1137/140956750*, *53*, 206–227. https://doi.org/10.1137/140956750

Wei, G. W. (2000). A unified approach for the solution of the fokker-planck equation. *Journal of Physics A: Mathematical and General*, *33*, 4935. https://doi.org/10.1088/0305-4470/33/27/311

Wood, B. D., & Kavvas, M. L. (1999). Ensemble-averaged equations for reactive transport in porous media under unsteady flow conditions. *Water Resources Research*, *35*, 2053–2068. https://doi.org/10.1029/1999WR900113

Woodbury, A. D., & Sudicky, E. A. (1991). The geostatistical characteristics of the borden aquifer. *Water Resources Research*, *27*, 533–546. https://doi.org/10.1029/90WR02545

Wu, J. C., & Zeng, X. K. (2013). *Review of the uncertainty analysis of groundwater numerical simulation*. https://doi.org/10.1007/s11434-013-5950-8

Xia, C.-A., Guadagnini, A., Hu, B. X., Riva, M., & Ackerer, P. (2019). Grid convergence for numerical solutions of stochastic moment equations of groundwater flow. *Stochastic Environmental Research and Risk Assessment 2019 33:8*, *33*, 1565–1579. https://doi.org/10.1007/S00477-019-01719-6

Yeh, Chang, J.-R. .-., & Short, T. E. (1992). An exact peak capturing and oscillation-free scheme to solve advection-dispersion transport equations. *Water Resources Research*, *28*, 2937–2951. https://doi.org/10.1029/92WR01751

Yeh, T. C. (1992). Stochastic modelling of groundwater flow and solute transport in aquifers. *Hydrological Processes*, *6*, 369–395. https://doi.org/10.1002/hyp.3360060402

Yoshimura, H., & Fujita, I. (2019). Investigation of free-surface dynamics in an open-channel flow. *https://doi.org/10.1080/00221686.2018.1561531*, *58*, 231–247. https://doi.org/10.1080/00221686.2018.1561531

Zárate-Miñano, R., & Milano, F. (2016). Construction of sde-based wind speed models with exponentially decaying autocorrelation. *Renewable Energy*, *94*, 186–196. https://doi.org/10.1016/j.renene.2016.03.026

Zheng, C., & Bennett, G. D. (2002). Applied contaminant transport modeling / chunmiao zheng, gordon d. bennett. http://search.ebscohost.com.ebsco-eds.wbg2.bg.agh.edu.pl/login.aspx?direct=true&db=edshlc&AN=edshlc.009544465-3&lang=pl&site=eds-live&scope=cite

Zhu, L., Gong, H., Chen, Y., Li, X., Chang, X., & Cui, Y. (2016). Improved estimation of hydraulic conductivity by combining stochastically simulated hydrofacies with geophysical data. *Scientific Reports*, *6*. https://doi.org/10.1038/srep22224