UNIVERSITY OF CALIFORNIA, SAN DIEGO

Uncertainty Quantification in Environmental Flow and Transport Models

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Science with Specialization in Computational Science

by

Peng Wang

Committee in charge:

Daniel M. Tartakovsky, Chair
Konstantine P. Georgakakos
Theodore Groves
Jan Kleissl
Paul F. Linden
Tracy Nishikawa

2011

The dissertation of Peng Wang is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2011

DEDICATION

To “my parents, my adviser, my friends and all who have provided help in my life”

EPIGRAPH

夫君子之行　靜以修身　儉以養德
非澹泊無以明志　非寧靜無以致遠
夫學須靜也　才須學靜也
非學無以廣才　非志無以成學
淫慢則不能勵精　險躁則不能治性
年與時馳　意與日去　遂成枯落
多不接世　悲守窮廬　將復何及

"The practice of a cultivated man is to improve himself by quietude and develop virtue by frugality.

Without detachment, there is no clarification of the will; without serenity, there is no achievement of great ambition.

Learning requires research through concentration, talent requires cumulation through learning.

Without learning talent cannot be expanded; without determination learning cannot be consistent.

Laziness and indulgence does not encourage will; impulsion and audaciousness does not refine virtue.

Years elapse with hours, aspirations flee with days and one withers like a fallen leaf.

There is no contribution to others but only lament and regret in a shabby hut, for the lost prime that will never return.

–Zhuge liang

("letter to my son", 234 A.D.)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The printed pages of this dissertation hold far more than a culmination of four years of my graduate study. It is also a means through which I would like to express my deepest gratitude to many generous and inspiring people I have met since the beginning of my overseas studies nine years ago. The list is long, but I cherish each of your contributions to my development and enlightenment as a scholar and a man:

I am heartily thankful to my supervisor, Daniel M. Tartakovsky. Your enduring patience, encouragement, supervision and support enabled me to develop an understanding not only of my research subject but also of a grand picture of the physical world on a philosophical level. Your mere presence offered great comfort and relief to a graduate student who was delivering his first presentation in a conference of great minds. Your endless diligence has shown me the correct path to success and your wisdom have lifted me out of ignorance and impulsion. Whenever I had questions or doubts, of work or of life, you are always there to help and to provide light to steer me through the winding water. From you, I have understood the meaning of self-improvement and have caught a glimpse of *Tao*. You are not just an adviser or a mentor, but more of a friend and a father. All the words and phrases are inadequate to describe my indebtedness to you. Thank you very much for all you have done for me and my family, we will always carry this gratitude in our hearts.

親愛的爸爸媽媽，非常感謝您們長久以來對我無私的愛與支持。“父母在，不遠遊，遊必有方”。請您們原諒兒子在過去的九年里沒能在您們膝下盡孝。現在海外學習終於要告一段落了，回首這九個寒暑，我可以問心無愧地告訴您們兒子沒有虛度光陰。在朋友和老師們等幫助下，兒時成為科學家的夢想正在逐步實現。在以後的歲月里，我會努力提高自己的知識層面，完善自己的人格與操守，切實地為社會做出貢獻。

媽媽，我總是讓您勞累、操心。你的愛温暖著我，也促使著孩子在不斷地成長。請您放心，兒子在外會照顧好自己，處理好和大家的關係，向著自己的人生目標前進。

爸爸，雖然您不善於表達，但是您的言傳身教總是在默默地影響著孩子。以前很多時候無法理解您的教誨，隨著年齡的增長，我也漸漸發現那些簡單易懂的話語里包含著大智慧和您人格的魅力。

您們的愛，兒子是永世無法償還的。我能做到的只是讓您們安度晚年，把您們無私的愛傳承給下一代，給予我們身邊需要幫助的人。

Nick and Patrick, thank you so much for all your support throughout the years since we met in Durham. Every time I recalled our time together in England, Hungary, China and New Zealand, I appreciate more of this dear and unique friendship among us, you are like true my brothers. Best wishes for you two on the pursue of doctoral degrees.

Lastly, I offer my gratitude to all of those who supported and helped me in any respect during the completion of my research and development of my life. Thank you very much.

The text of this dissertation includes the reprints of the following papers, either accepted or submitted for consideration at the time of publication. The dissertation author was the primary investigator and author of these publications.

Chapter 2

Wang, P., Quinlan, P., Tartakovsky, D. M., (2009), 'Effects of spatio-temporal variability of precipitation on contaminant migration in vadose zone'. *Geophys. Res. Lett., vol. 36, pp. L12404, doi:10.1029/2009GL038347.*

Chapter 3

Wang, P., Tartakovsky, D. M., (2011), 'Probabilistic predictions of infiltration into heterogeneous media with uncertain hydraulic parameters'. *Int. J. Uncert. Quant., vol. 1, no. 1, pp. 35-47.*

Wang, P., Tartakovsky, D. M., (2011), 'Reduced complexity models for probabilistic forecasting of infiltration rates'. *Adv. Water Resour., vol. 34, pp. 375-382, doi:10.1016/j.advwatres.2010.12.007.*

Chapter 4

Wang, P., Tartakovsky, D. M., (2011), 'Uncertainty quantification in kinematic wave models'. *Submitted.*

Chapter 5

Wang, P., Tartakovsky, D. M., (2011), 'Uncertainty quantification of algae bloom in an enclosed aquatic system '. *Under preparation.*

VITA

2011	Ph.D. in Engineering Science with Specialization in Computational Science, University of California, San Diego.
2010	M.Sc. in Engineering Sciences with Mechanical Engineering, University of California, San Diego.
2007	M.Eng. in Mechanical Engineering (1*st* Class with Honours), University of Durham, Durham, United Kingdom.

JOURNAL PUBLICATIONS

Wang, P., Tartakovsky, D. M., (2011), 'Uncertainty quantification of algae bloom in an enclosed aquatic system '. *Under preparation.*

Wang, P., Tartakovsky, D. M., (2011), 'Uncertainty quantification in kinematic wave models'. *Submitted.*

Wang, P., Tartakovsky, D. M., (2011), 'Reduced complexity models for probabilistic forecasting of infiltration rates'. *Adv. Water Resour., vol. 34, pp. 375-382, doi:10.1016/j.advwatres.2010.12.007.*

Wang, P., Tartakovsky, D. M., (2011), 'Probabilistic predictions of infiltration into heterogeneous media with uncertain hydraulic parameters'. *Int. J. Uncert. Quant., vol. 1, no. 1, pp. 35-47.*

Wang, P., Quinlan, P., Tartakovsky, D. M., (2009), 'Effects of spatio-temporal variability of precipitation on contaminant migration in vadose zone'. *Geophys. Res. Lett., vol. 36, pp. L12404, doi:10.1029/2009GL038347.*

SELECT PRESENTATIONS

Wang, P., Tartakovsky, D. M., 'Probabilistic density function of runoff rate', (2010) *American Geophysical Union Fall Meeting, San Francisco CA, December 13-17.*

Wang, P., Tartakovsky, D. M., 'PDF solutions for infiltration into heterogeneous soils with uncertain parameters', (2010) *XVIII International Conference on Computational Methods in Water Resources, Barcellona, Spain, June 21-24.*

Wang, P., Tartakovsky, D. M., 'PDF solutions for infiltration into heterogeneous soils with uncertain parameters', (2009) *American Geophysical Union Fall Meeting, San Francisco CA, December 14-18.*

Wang, P., Tartakovsky, D. M., 'PDF solutions for infiltration into heterogeneous soils with uncertain parameters', (2009) *Worldwide Universities Network (WUN) Workshop on Critical Zone Research, University Park PA, September 7-11.*

AWARDS

Powell-Fochet Fellowship, School of Engineering, University of California, San Diego (2007-2010).

ABSTRACT OF THE DISSERTATION

Uncertainty Quantification in Environmental Flow and Transport Models

by

Peng Wang

Doctor of Philosophy in Engineering Science with Specialization in Computational Science

University of California, San Diego, 2011

Daniel M. Tartakovsky, Chair

This dissertation is a work on the development of mathematical tools for uncertainty quantification in environmental flow and transport models. In hydrology, data scarcity and insufficient site characterization are the two ubiquitous factors that render modeling of physical processes uncertain. Spatio-temporal variability (heterogeneity) poses significantly impact on predictions of system states. Standard practices are to compute (analytically or numerically) the first two statistical moments of system states, using their ensemble means as predictors of a system's behavior and variances (or standard deviations) as a measure of predictive uncertainty. However, such approaches become inadequate for risk assessment where one is typically interested in the probability of rare events. In other words, full

statistical descriptions of system states in terms of probabilistic density functions (PDFs) or cumulative density functions (CDFs), must be sought. This is challenging because not only parameters, forcings and initial and boundary conditions are uncertain, but the governing equations are also highly nonlinear. One way to circumvent these problems is to develop simple but realistic models that are easier to analyze. In chapter 3, we introduce such reduced-complexity approaches, based on Green-Ampt and Parlange infiltration models, to provide probabilistic forecasts of infiltration into heterogeneous media with uncertain hydraulic parameters. Another approach is to derive deterministic equations for the statistics of random system states. A general framework to obtain the cumulative density function (CDF) of channel-flow rate from a kinematic-wave equation is developed in the third part of this work. Superior to conventional probabilistic density function (PDF) procedure, the new CDFs method removes ambiguity in formulations of boundary conditions for the CDF equation. Having developed tools for uncertainty quantification of both subsurface and surface flows, we apply those results in final part of this dissertation to perform probabilistic forecasting of algae growth in an enclosed aquatic system.

Chapter 1

Introduction

知不知　尚矣
不知知　病矣
夫唯病病　是以不病
圣人不病　以其病病
是以不病

"Acknowledgement of ignorance is the strength of learning,

Pretension of knowledge is the sickness of learning.

Only understanding one's sickness can prevent disease.

Sage does not have any weakness because he sees his own weakness,

and hence he does not have any weakness."

-Lao-zi

"Dao De Jing" (chapter 71), 6th century B.C.

1.1 Uncertainty Quantification

Wisdom begins with the acknowledgement of one's ignorance - the limits of what we know. Such recognition of uncertainty is a classic empiricist philosopher's view on human learning. It gave rise to recognition that most physical systems are fundamentally stochastic [8, 9, 10, 11].

Uncertainty quantification (UQ) is a science of quantitative characterization and possible reduction of uncertainty in the prediction of system behavior [12]. Many scientific and engineering fields, for example climate and environmental modeling, energy generation, control and manufacturing, and process and system design, have fully embraced the importance of dealing with, and quantifying, predictive uncertainty.

Heterogeneity, insufficient site characterization and conceptual-mathematical limitations of modeling approaches render the predictions of groundwater flow and transport notoriously difficult. Continuous advancements in data acquisition are encouraging but will not provide us with full knowledge of system parameters. Measurements of hydraulic properties are prone to measurement and interpretive errors. This uncertainty is ubiquitous and cannot be solved with more computational resources. It makes hydrogeology ever more challenging and UQ is both necessary and increasingly deemed obligatory. This is the primary focus of this dissertation in which various probabilistic tools are developed to quantify uncertainty in models of environmental flow and transport.

In the remainder of this chapter, we present a brief review of common UQ methods in hydrogeology (section 1.2) and their shortcomings are also discussed. A considerable part of section 1.3 is devoted to probabilistic risk assessment (PRA) that prompts our study of new approaches to UQ. In section 1.4, those new methods and their applications in hydrogeology are outlined.

1.2 Current UQ methods

Many probabilistic analyses classify uncertainty into two categories: aleatory and epistemic. The former is the uncertainty "inherent in a nondeterministic (random, stochastic) phenomenon" that cannot be suppressed by more accurate measurements; while the latter is "attributable to incomplete knowledge about a phenomenon that affects our ability to model it" [13] - this is the primary concern of uncertainty quantification in hydrogeology.

From the earliest monograph [14] to the recent work [15], significant progress to address epistemic uncertainty in stochastic hydrogeology resulted in a number of competing approaches. Specifically, these approaches are developed to quantify structural (model) and parametric uncertainties associated with surface/subsurface process, which are mathematically represented through (stochastic) governing equations.

As outlined in Figure 1.1, structural uncertainty in hydrogeology often stems from one's limited knowledge of the hydrogeologic system makeup. It can be

quantified by random domain decompositions [16, 17]. Meanwhile, our inadequate understandings of subsurface flow and transport leads to various interpretations and mathematical representations of the corresponding physical (chemical or biological) processes. One can apply a Bayesian maximum entropy approach [18], maximum likelihood Bayesian averaging [19], etc. to quantify such uncertainty.

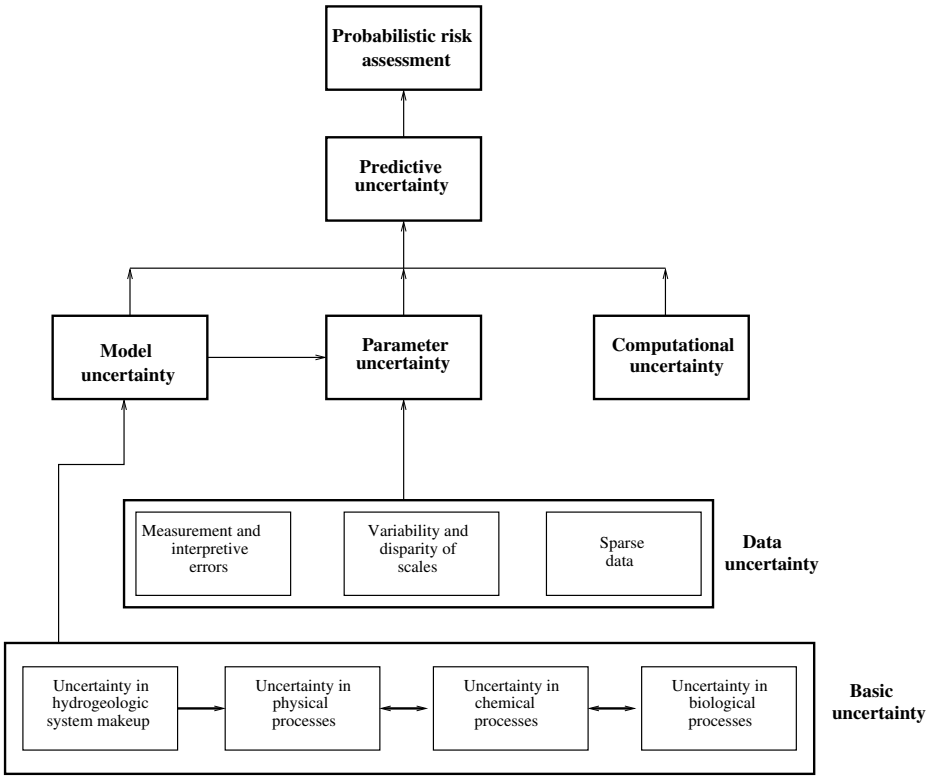

Figure 1.1: Sources of uncertainty in a distributed system.

Parametric uncertainty results from spatio-temporal heterogeneity coupled with scant and often noisy measurements of hydraulic and bio- geochemical properties. Routinely, researchers approach this type of uncertainty through a statistical

treatment of parameters. Specifically, a parameter $\mathcal{A}(\mathbf{x}, t, \omega)$ varies not only in the physical domain, $(\mathbf{x}, t) \in \mathcal{D}$, but also in the probability space $\omega \in \Omega$. Such probabilistic descriptions of parameters render the corresponding governing equations stochastic. Their solutions are given in terms of probability density functions (PDFs) or cumulative density functions (CDFs) of system states.Methodologies to address parametric uncertainty could be subdivided into, statistical methods and stochastic methods. Monte Carlo Simulations (MCS), moment differential equations (MDE) and generalized polynomial chaos expansions (PCEs) are the three most common practices and shall be discussed below.

Among statistical methods for UQ, MCS is commonly labeled as "brute force". It involves numerous simulations of deterministic codes to solve the governing equations and hence to obtain statistical distribution of target system states. In principle, the law of large numbers ensures that such approach becomes increasingly accurate as more simulations ($N \to \infty$) are performed. Yet in practice, we are confined to limited computing resources and MCS has a convergence rate of the order $1/\sqrt{N}$. Without any physical insight, the advancement in computing power is unlikely to match the growing size of models, i.e, the number of degrees of freedom (random variables).

MDE, on the other hand, as an example of "indirect" stochastic methods for UQ. By employing Reynolds decomposition, it represents a random parameter

$\mathcal{A}$ as the sum of its mean $\bar{\mathcal{A}}$ and a zero-mean fluctuation $\mathcal{A}'$

$$\mathcal{A} = \bar{\mathcal{A}} + \mathcal{A}', \qquad \langle \mathcal{A}'^2 \rangle = \sigma_{\mathcal{A}}^2 \tag{1.1}$$

where $\sigma_{\mathcal{A}}^2$ is the second statistical moment (variance) of $\mathcal{A}$. Ensemble averaging of stochastic governing equations leads to deterministic equations for the statistical moments of system states. Such MDEs typically require a closure approximation. Based on physical arguments, various closure approximations have been proposed. Often they assume mild heterogeneity (relatively small variance), $\bar{\mathcal{A}} \gg \mathcal{A}'$, which undermines the range of MDE applicability.

Grounded in rigorous mathematical theory [8], PCEs are an intrusive stochastic finite element method. As its name suggests, a second-order stochastic process (parameter) can be expanded:

$$\mathcal{A}(\mathbf{x}, t, \omega) \; = \sum_{i=1}^{\infty} a_i(\mathbf{x}, t)\, \Psi_i(\omega) \; \approx \sum_{i=1}^{N} a_i(\mathbf{x}, t)\, \Psi_i(\omega), \qquad \langle \Psi_i \Psi_j \rangle = \langle \Psi_i^2 \rangle\, \delta_{ij} \tag{1.2}$$

where the orthogonal Wiener-Askey polynomials $\Psi_i(\omega)$ correspond to the distribution of random variable ω and δ_{ij} denotes the Kronecker delta function. Knowledge of the deterministic polynomial chaos coefficients $a_i(\mathbf{x}, t)$ could fully characterize the random parameter $\mathcal{A}$. In theory, truncation at polynomial order N provides non-perturbative UQ methods, which could capture large fluctuations and highly non-Gaussian field. However, PCE also suffers drawbacks: it works well for unimodal distributions of random variables and large correlation length. Application of PCEs in subsurface hydrology can be found in [20].

1.3 Probabilistic Risk Assessment

The aforementioned approaches are primarily used to compute the first and second statistical moments of systems states. They are adequate to describe the mean system behavior whose prediction error can be quantified with the ensemble variance. However, when one is concerned with probabilistic risk assessment (PRA), i.e., the probability of rare events, those methods become insufficient and full statistical characterization of system states, i.e., PDFs or CDFs, is required.

As outlined in the report of National Research Council [21], the main focus of PRA is on uncertainty quantification to assess both the likelihood of an environmental hazard and the efficiency of alternative remediation efforts. Similar to most engineering applications, a comprehensive PRA should address the following three questions: "What can happen? How likely is it to happen? Given that it occurs, what are the consequences?" [22].

Consider, as an example, hazardous algae blooms (HABs) in an aquatic system. We say that "a system failed" at time t, if the population of a specific type of algae exceeds the Environmental Protection Agency mandated levels. By identifying the system's components, the general framework of PRA relates the occurrence of the system failure to the failures of its constitutive parts (basic events) through Boolean operators AND and OR. Typical basic events include "Natural attenuation fails", "Favorable Climate", "Remediation effort fails", "Accumulation of nutrients".

The procedure described above is a standard PRA framework whose output relies on the computation of probabilities of basic events. In contrast to complex artificial systems - for instance nuclear power plants and shuttles [22] - there are no reliability databases for natural (environmental) systems to quantify the uncertainty of individual components. Instead, one must solve stochastic partial differential equations that govern flow and transport to obtain the predictive uncertainty in Figure 1.1, i.e. PDFs or CDFs of system states. This is the goal of our study on the development of UQ methods and our results will be briefly introduced in the following section.

1.4 Research Outline

In addition to the uncertainty about parameters, forcing terms and initial and boundary conditions, nonlinearity of the governing equations complicates their analytical and numerical analysis. Such difficulty is elucidated in Chapter 2 where a series of flow and transport simulations are conducted to study the effects of spatio-temporal variability of precipitation on contaminant migration in the vadose zone. Specifically, we confirm that meteorological data based on different temporal averaging scales (annual and daily) could lead to large discrepancy of net infiltration and groundwater contamination level in arid and semi-arid regions. We further demonstrate that the accuracy of temporally averaged predictions is influenced by the degree of nonlinearity of the Richards equation describing flow

in partially saturated porous media. Additional errors are introduced when one ignores topographical and/or urban features that tend to focus and increase local infiltration rates.

Given the practical impossibility of obtaining accurate PDF solutions of the stochastic Richards equation in three dimensions, we develop novel reduced complexity models to circumvent this problem. Based on the Green-Ampt and Parlange models that provide alternative descriptions of flow in unsaturated porous media, we propose reduced complexity models for the probabilistic forecasting of infiltration rates in heterogeneous soils during surface runoff and/or flooding events. As presented in Chapter 3, these models yield closed-form semi-analytical expressions for the single- and multi-point PDFs of infiltration depth and corresponding infiltration rate, which quantify predictive uncertainty stemming from uncertainty in soil properties. We investigate the temporal evolution of these PDFs, the relative importance of uncertainty in various hydraulic parameters and their cross-correlation, and the impact of the choice of a functional form of the hydraulic function. Comparing to their counterparts obtained from a full infiltration model based on the Richards equation, the infiltration rate PDFs computed with the reduced complexity models are in close agreement at early times and provide conservative estimates of predictive uncertainty at all times.

Alternatively, one can attempt to derive deterministic equations for the PDFs/CDFs of system states from stochastic governing equations. In Chapter 4, this methodology is applied to kinematic wave theory in order to obtain a statisti-

cal description of channel flows with uncertain hydraulic parameters, input sources and boundary conditions. Introducing the new concept of fine-grained CDFs, our approach is equivalent to computing the PDFs of volumetric flow rate but provides easier formulation of boundary conditions for the new CDF equation. Since the nonlinearity of the governing equation is reduced without any approximation, we derive closed-form semi-analytical solutions for both raw and full CDFs of volumetric flow rate. The former facilitate their numerical simulations. We performed MCS to evaluate the accuracy of the CDF approach. Results show that the two approaches are in close agreement at all times.

Having developed various new tools for UQ, we conclude this dissertation by a stochastic model of algae bloom in Chapter 5. This is accomplished by computing the probabilistic distribution of algae population density using a Fokker-Planck equation. Final conclusions are drawn in Chapter 6.

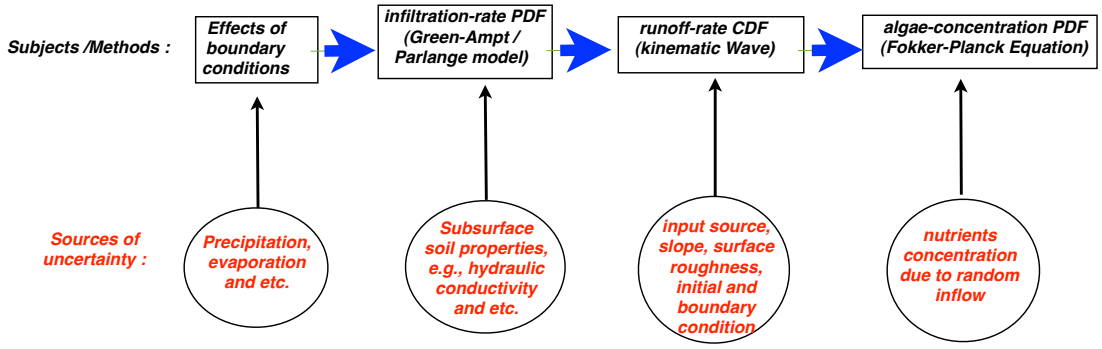

Figure 1.2: Workflow of Peng Wang's doctoral research.

Chapter 2

Effects of Boundary Conditions

2.1 Introduction

In this Chapter we investigate the effects of boundary condition uncertainty, i.e. spatio-temporal variability (heterogeneity and non-stationarity) of precipitation rate, on subsurface flow and transport in the vadose zone.

The vadose zone forms a major hydrologic link, and acts as a main conduit for anthropogenic contaminants, between land surface and groundwater aquifers. These two functions of the vadose zone are closely related, since the downward movement of water is the key mechanism of contaminant migration. Yet their respective analyses often require distinct methodologies and assumptions: even if large-scale averaged models are adequate to describe the effects of precipitation on groundwater recharge [23, 24], they might fail to ascertain the risk posed by contaminant spills to groundwater quality. This distinction becomes paramount

in regions with arid or semi-arid climate, which is defined by annual rainfall below $250 - 500$ mm ($10 - 20$ in). It is often argued that such low precipitation rates are not sufficient to drive contaminants from surface or near-surface spills to the water table, i.e., that contaminants released into soils in (semi-)arid regions pose no threat to groundwater quality.

This assertion rests on an implicit assumption that time and/or space averaged precipitation rates provide an adequate input for subsurface flow and transport models. The highly nonlinear nature of both the coupling between surface and subsurface processes and the Richards equation that is routinely used to describe flow in partially saturated porous media provides a clear indication that the superposition principle does not hold, i.e., that predictions based on averaged boundary conditions (infiltration rates) are at best an approximation. The adequacy of such approximations has been the subject of a handful of studies in the past few decades. Analyzing contaminant migration in homogeneous soils, [25] and [26] found breakthrough curves under time-varying and averaged boundary conditions to be similar, while [27] and [28] concluded that contaminant might travel significantly faster and further under time-variable infiltration than under its time-averaged counterpart. [29] attributed this discrepancy to the absence of root uptake from the latter studies and concluded that the use of time-averaged infiltration rates is justifiable. (Note that her analysis used infiltration rates typical of humid climates and calls for the use of modified soil parameters.) The more recent studies of [30] and [31] seem to support this finding, even though [30]

cautioned that it might become invalid under severe weather conditions and [31] added a few caveats discussed below.

Several crucial issues related to the adequacy of averaged precipitation rates as predictors of the risk of groundwater contamination remain unresolved. First, most of the studies mentioned above examined predictive errors stemming from the use of average infiltration rates rather than their precipitation counterparts. While the latter are readily available on an hourly basis, e.g., from meteorological stations, the former have to be estimated. Second, the impact of the degree of nonlinearity of the Richards equation, i.e., of the choice of particular constitutive laws, has not been investigated. Finally, surface's topography and/or local land use localize infiltration, enhancing a contaminant's downward migration in a manner that undermines the use of spatially averaged precipitation and infiltration rates. This chapter aims to elucidate the impact of spatio-temporal averaging of precipitation rates on flow and transport predictions. This question gains in significance now that global climate change is likely to result in more severe weather with stronger rainfall, greater runoff, and longer periods of drought even if resulting annual precipitation rates might remain unchanged [32].

2.2 Model Formulation

Flow in the vadose zone, i.e., distributions of volumetric flux $\mathbf{q}(\mathbf{x}, t)$, pressure head $\psi(\mathbf{x}, t)$ and water content $\theta(\mathbf{x}, t)$ at any point $\mathbf{x} = (x_1, x_2, x_3)^T$ and time

t, can be described by a combination of Darcy's law and the continuity equation,

$$\mathbf{q} = -K\nabla(\psi + x_3) \quad \text{and} \quad \frac{\partial\theta}{\partial t} = -\nabla \cdot \mathbf{q}, \tag{2.1}$$

respectively. A flow model is completed by specifying functional forms of unsaturated hydraulic conductivity $K = K(\theta)$ and retention curve $\psi = \psi(\theta)$. We assume that a porous medium is homogeneous (heterogeneity effects are discussed below). To be concrete, we set initial water content to its residual value, $\theta(\mathbf{x}, 0) = \theta_r$, and place the water table at $x_3 = -L = -150$ m. The boundary condition at the Earth surface $x_3 = 0$ is determined from atmospheric data and surface conditions. In the presence of ponding, $\psi = h_o$ where $h_0 = 0\,m$ denotes the height of standing water on a flat surface as in our case. In the absence of ponding, the boundary condition at $x_3 = 0$ is $q_3 = i$ if precipitation rate $P > 0$ and $q_3 = -e$ if $P = 0$. Infiltration rate i is computed from precipitation rate P, actual evaporation rate e, and runoff rate r over a unit area as $i = P - e - r$. The initial condition of the terrain also plays an important role in spatial-average effect, in this paper it is assumed to be dry.

In the current study, we employ a simple, linear relation between runoff and precipitation, $r = C_r P$ with $C_r = 0.1$, which is a reasonable approximation for (semi-)arid regions away from major rivers [33]. More complex relations would add another nonlinear feedback into the system. However, we must stress that in reality, runoff is a function of both soil water content θ and precipitation rate, whose motion can be approximated using kinematic wave model. A rigours general

framework is developed in Chapter 4 to address overland flow uncertainty and in a narrow scope, runoff uncertainty.

Actual evaporation rate $e = \max\{e_p, e_v\}$, where e_p is potential evaporative demand of atmosphere and e_v is a soil's ability to conduct water to the surface. The latter can be defined as $e_v = Kr_s(\psi_a - \psi_s)$ [1], where $\psi_a = RT/(M_w g) \ln H_r$ is the pressure potential of atmosphere, $\psi_s = \psi(x_3 = 0)$ is pressure potential at the surface and ψ_a is the pressure potential of atmosphere given by the Kelvin equation [34], in which R is universal gas constant, T is absolute air temperature ($^\circ K$), M_w is molecular weight of water and g is gravity acceleration, and H_r is relative humidity. In the absence of surface crust and vegetation, surface resistance r_s equals the reciprocal of the distance from the land surface.

Potential evaporation rate e_p is calculated from measurements of net solar radiation R_n, vapor pressure p_v, air temperature T_c ($^\circ C$), and mean wind speed at two meters above the ground U_w by using a modified Penman equation [35],

$$e_p = \frac{wR_n}{694.5(1 - 0.000946T_c)} + 24(1 - w)(p_s - p_v)f_w. \tag{2.2}$$

Here $p_s = 0.6108 \exp[17.27T_c/(T_c + 237.3)]$ is saturation vapor pressure; $f_w = 0.030 + 0.0576U_w$ is the wind function; and the weight function $w = \Delta/(\Delta + \gamma)$ where $\Delta = 4099p_s/(T_c + 237.3)^2$, and psychrometer constant $\gamma = 0.000646(1 + 0.000946T_c)(101.3 - 0.0115z + 5.44 \cdot 10^{-7}z^2)$ with z denoting the elevation of a weather station above the mean sea level.

Migration of a conservative contaminant with concentration $c(\mathbf{x}, t)$ is de-

scribed by advection-dispersion equation,

$$\partial c/\partial t + \nabla \cdot (\mathbf{u}c) = \nabla \cdot (\mathbf{D}\nabla c, \tag{2.3}$$

where $\mathbf{u} = \mathbf{q}/\omega$ is the mean macroscopic velocity, ω is the porosity, and $\mathbf{D}$ is the dispersion coefficient tensor whose longitudinal and transverse components are given by $D_L = D_m + \lambda_L|\mathbf{u}|$ and $D_T = D_m + \lambda_T|\mathbf{u}|$, respectively. Initially, the soil is contamination-free, $c(\mathbf{x}, 0) = 0$, except for the layer $x_3 \in [-0.5\,m, -0.4\,m]$ where the concentration is $c_0 = 100\,gm^{-3}$.

The raw meteorological data used in the simulations presented below come from a California Irrigation Management Information System (CIMIS) station located near Five Points, Fresno County, CA at the surface elevation $z = 86.9$ m. The data set, freely available on line, contains measurements of daily precipitation, air temperature, solar radiation, relative humidity, wind speed and vapor pressure collected from 1983 to 2008. The average annual precipitation rate during this time period was 0.2 m/year, which is representative of semi-arid regions.

Unless explicitly stated, the simulations reported below correspond to a sandy loam soil with porosity $\omega = 0.496$, residual water content $\theta_r = 0.15$, saturated hydraulic conductivity $K_s = 0.7$ m/day, and van Genuchten constitutive relations [1, Table 1]

$$K_r = (1 - |\alpha\psi|^{\beta-1} D^{-\gamma})^2/D^{\gamma/2}, \qquad D = 1 + |\alpha\psi|^{\beta}, \tag{2.4a}$$

$$\Theta = (1 + |\alpha\psi|^{\beta})^{-\gamma}, \qquad \Theta = (\theta - \theta_r)/(\omega - \theta_r), \tag{2.4b}$$

with parameters $\alpha = 0.847$, $\beta = 4.8$, $\gamma = 1 - 1/\beta$. In all transport simulations, we

set molecular diffusion to $D_m = 10^{-6}\, m^2/day$, longitudinal dispersivity to $\lambda_L = 0.1$ m, and transverse dispersivity to $\lambda_T = 0.01$ m.

Numerical code VS2DT [36] is used for daily numerical simulations over 25 years. Evaporation is simulated in VS2DT by a two-stage process, which requires three inputs: potential evaporation e_p, pressure potential of the atmosphere ψ_a and the surface resistance r_s. Since VS2DT can treat the surface as either a precipitation or evaporation boundary, but not both at the same time, we divide each daily recharge period into two periods: first taking effective (total) precipitation after runoff as infiltration, with rate modified so that total inflow mass will be the same; it is then followed by an evaporation period. It is noted that due to nonlinearity of Richards equation and constitutive relations (2.4), a reverse order of evaporation and precipitation periods would lead to different results.

2.3 Effects of Temporal Averaging

We start by analyzing the effects of temporal averaging of daily atmospheric data on the downward migration of moisture and contaminants. A one-dimensional soil column was discretized into 1500 cells, which puts the contamination source in the fifth cell from the surface. Fig. 2.1c compares temporal evolution of the wetting front (defined as the leading edge of a moisture plume wherein water content exceeds its initial value) computed with the daily meteorological data described above and its counterparts resulted from monthly and yearly averages of

these data. One can see that the yearly averages lead to predictions that are both quantitatively and qualitatively wrong, while monthly averages yield somewhat better predictions that still underestimate the extent of wetting. It is worthwhile recognizing that the use of daily data is in itself an approximation that averages instantaneous rainfall intensity albeit over shorter time intervals than monthly and yearly data do. Hence the actual errors introduced by the reliance of yearly meteorological data are even higher.

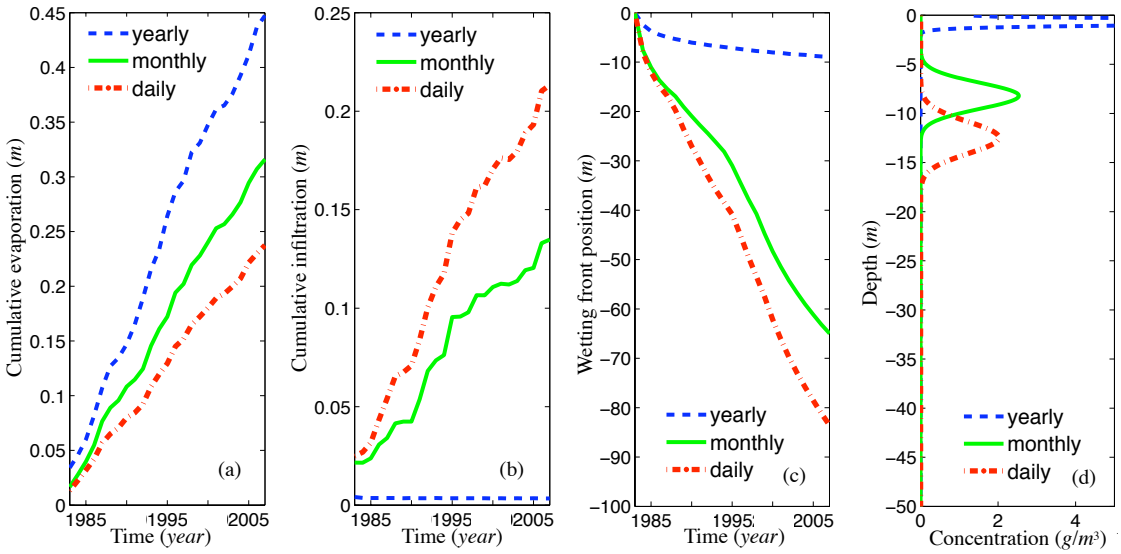

Figure 2.1: Cumulative evaporation (a), cumulative infiltration (b), temporal evolution of the wetting front (c), and final concentration profiles (d) predicted with daily meteorological data and their monthly and yearly averages.

The extent of subsurface contamination resulted from 25 years of infiltration is shown in Fig. 2.1d. The annual (averaged) precipitation data predict a contaminant that remains practically immobile in its initial position, which is consistent with routine claims made for (semi-)arid regions. This prediction is, however, at

variance with predictions obtained with daily and monthly meteorological data. One can see that contaminant does migrate downward with reduced concentration, which reflects the presence of more water in soil in daily simulations than in their monthly and especially yearly counterparts.

Figures 2.1a,b shed light on the cause of the apparent differences in modeling predictions based on daily, monthly, and yearly data. While yearly data result in zero infiltration, it is quite significant when computed from either monthly or daily data. Another important feature of these results is the increasing dichotomy between both infiltration and evaporation predicted from daily data and averaged data. At the end of 25 years, the actual evaporation computed from annual data is 0.447 m, which is almost twice the value of 0.2375 m computed from daily data. This finding is one of the reasons why Destouni [29], whose simulations spanned a one-year period, observed little differences between predictions based on daily and annual data. Another reason is that we are concerned with (semi-)arid climates that are characterized by severe precipitation patterns, while the simulations of Destouni [29] were conducted for humid conditions.

Both infiltration and contaminant migration in the vadose zone are influenced to a large degree by its hydraulic properties and heterogeneity. Figure 2.2 presents moisture profiles predicted with daily, monthly, and yearly averages for three homogeneous soil types: Fresno medium sand (saturated hydraulic conductivity $K_s = 400$ m/day), Columbia sandy loam ($K_s = 0.7$ m/day) and Yolo light clay ($K_s = 0.011$ m/day). Other hydraulic properties of these soils can be found

in [1, Table 1]. The predictive errors caused by the reliance on annual meteorological data can be quantified in terms of a relative error introduced by the temporal averaging of meteorological data, $\mathcal{E} = |z_f^{(d)} - z_f^{(y)}|/z_f^{(d)}$ where $z_f^{(d)}$ and $z_f^{(y)}$ are the wetting front's positions resulting from the use of daily and yearly data, respectively. The errors increase with hydraulic conductivity, as wetting fronts travel farther and faster without reaching the water table ($\mathcal{E} = 90\%$ and 76% for Columbia sandy loam and Yolo light clay, respectively). After the wetting front reaches the water table, as is the case reported in Figure 2.2 for Fresno medium sand at the end of 25 years, this error decreases slightly to $\mathcal{E} = 68.5\%$.

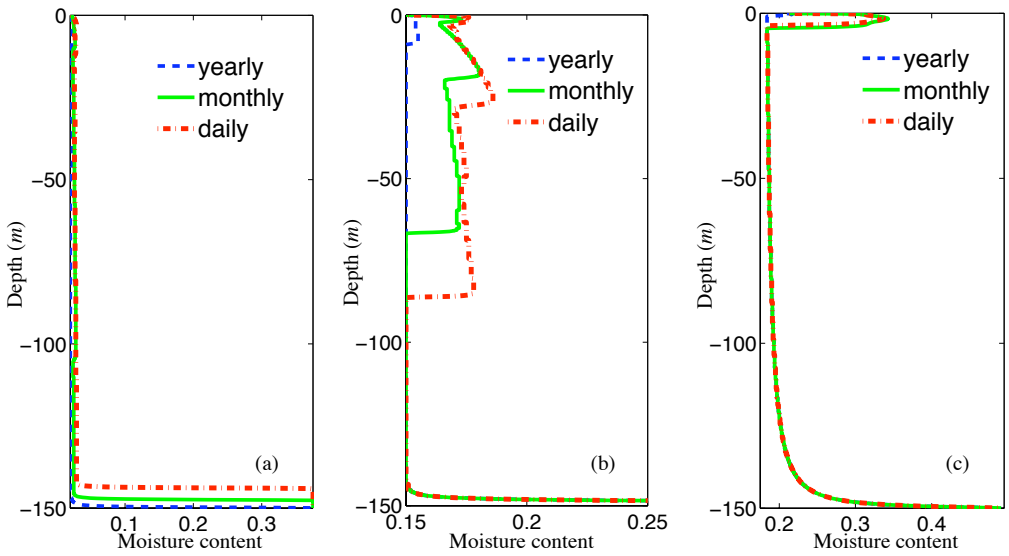

Figure 2.2: Moisture profiles at the end of 25 years of simulations for three soil types: (a) Fresno medium sand, $K_s = 400$ m/day; (b) Columbia sandy loam, $K_s = 0.7$ m/day; and (c) Yolo light clay, $K_s = 0.011$ m/day. These and other soil properties are taken from [1, Table 1].

Detailed investigation of the impact of soil's heterogeneity is carried out in the next chapter. Destouni [29] concluded that "Textural heterogeneity in the soil

profile, such as a clay layer in sandy loam, increases the discrepancy between the steady state and the transient flow model when root water uptake is neglected." Schoups et al. [31] seemingly contradict this conclusion by noting "Where time averaging does not give satisfactory results, it may still give adequate predictions of the spatial-ensemble distribution or statistical moments of the variable of interest." The veracity of such conclusions is hard to ascertain and is likely to be site specific.

2.4 Nonlinearity Effects

Nonlinearity of the Richards equation (2.1) stems from the dependence of both relative hydraulic conductivity K_r and pressure head ψ on water content θ. The choice of constitutive relations $K_r(\theta)$ and $\psi(\theta)$ is bound to influence the discrepancy between predictions based on daily and yearly averages of meteorological data. To investigate this phenomenon, we compare predictions based on the van Genuchten model (2.4) with those corresponding to the Brooks-Corey model

$$K_r = (\psi/\psi_b)^{-2+3\lambda}, \qquad \Theta = (\psi/\psi_b)^{-\lambda} \tag{2.5}$$

wherein $\psi_b = -0.85$, $\lambda = 1.6$ and $\theta_r = 0.11$, and the Haverkamp model

$$K_r = [1 + (\psi/\psi_a)^b]^{-1}, \qquad \Theta = [1 + (\psi/\alpha)]^{-\beta} \tag{2.6}$$

wherein $\psi_a = -0.9$, $b = 9.2$, $\alpha = -1.26$, $\beta = 4.6$, and $\theta_r = 0.16$. The parameters in these models are representative of sandy loam soils [1, Table 1], except for ψ_a and b which were obtained by fitting. An alternative way to parameterize (2.5) and

(2.6) is to ensure "hydraulic equivalency" between the three models (2.4), (2.5), and (2.6) [37, 38]. The latter choice is more rigorous but less frequently used by practitioners.

Relative errors in predictions of the wetting front penetration and cumulative infiltration after 25 years of simulations with van Genuchten, Haverkamp, and Brooks-Corey constitutive models are shown in Figure 2.3 for several values of saturated hydraulic conductivity. A relative error introduced by the temporal averaging of meteorological data is defined as $\mathcal{E} = |z_f^{(d)} - z_f^{(y)}|/z_f^{(d)}$, where $z_f^{(d)}$ and $z_f^{(y)}$ are the wetting front's positions resulting from the use of daily and yearly data, respectively.

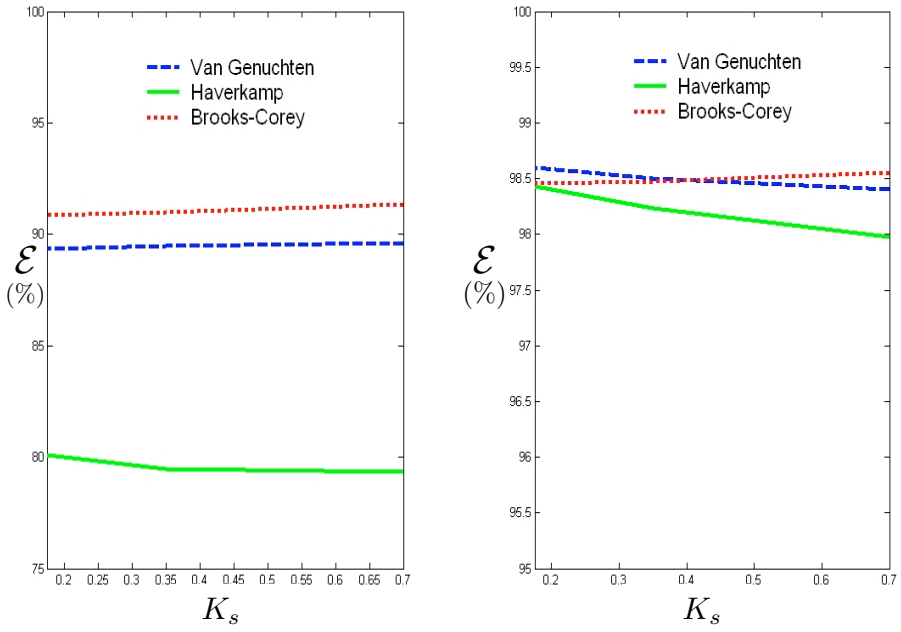

Figure 2.3: Relative errors in predictions of the wetting front penetration (a) and cumulative infiltration (b) after 25 years of simulations with van Genuchten, Haverkamp, and Brooks-Corey constitutive models, plotted as a function of saturated hydraulic conductivity $K_s = 0.7$ m/day, $= 0.35$ m/day, and $= 0.175$ m/day.

Our simulations revealed the choice of a constitutive model affects both the wetting front penetration and cumulative infiltration. The Brooks-Corey model leads to the largest $\mathcal{E}$, while the Haverkamp model results in the smallest error. After 25 years of infiltration, the relative error in wetting front predictions is $\mathcal{E} = 92\%$ for the Brooks-Corey model, 89% for Van Genuchten model and 79% for the Haverkamp model. We also found that the errors are largely insensitive to the value of saturated conductivity K_s.

2.5 Effects of Spatial Averaging

Topographic features and built environments often focus infiltration. Under such conditions, the use of large-scale meteorological data to predict contaminant transport amounts to spatial averaging, which is bound to introduce predictive errors due to nonlinearity of the governing equation. We analyze this phenomenon by comparing two-dimensional contaminant migration induced by uniform and localized infiltration regimes. Both regimes use the same annual meteorological data as before; the former assumes uniform infiltration and evaporation over the land surface, while the latter focuses them at a point.

Concentration isolines of $c = 0.01\, g/m^3$ at the end of 25 years of two-dimensional simulations with uniform and localized boundary conditions are shown in Figure 2.4. The spatial averaging of annual meteorological data underestimates the extent of downward contaminant migration from its initial location near the

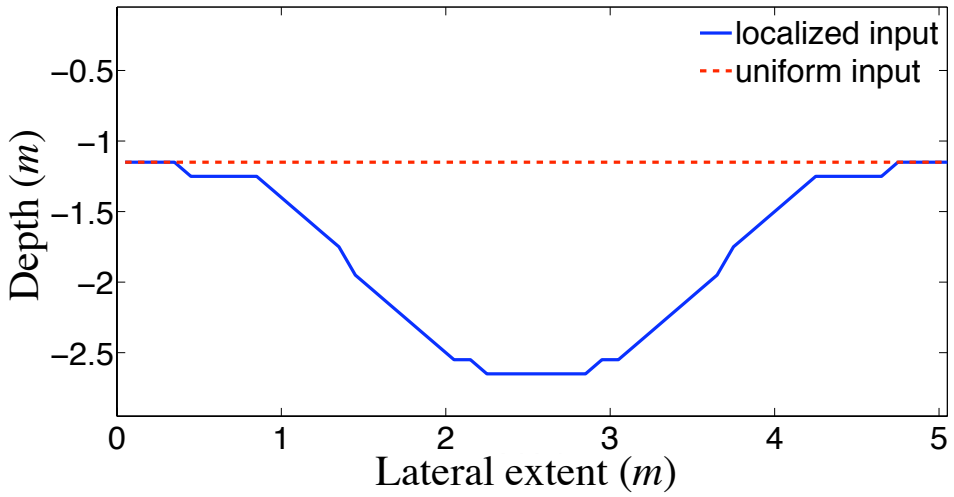

Figure 2.4: Extent of soil contamination after 25 years of simulations with spatially uniform and localized annual meteorological data.

Earth surface ($x_3 = -0.4\,m$) by the factor of two. The increased water content resulted from localized infiltration results in a drop of solute volumetric concentration and a V-shaped concentration profile.

2.6 Conclusions

We investigated the effects of relying on annual spatially averaged meteorological data to predict contaminant transport through the vadose zone in (semi-)arid regions. Our analysis leads to the following major conclusions.

1. Given high temporal variability of precipitation in (semi-)arid regions, the reliance on annual meteorological data might significantly underestimate the downward migration of contaminant.

2. Predictive errors stemming from the use of annual data increase with time

and are more pronounced in highly conductive soils.

3. Selection of constitutive models for the Richards equation, e.g., van Genuchten model versus Brooks-Corey model, influences the magnitude of predictive errors among various time periods of averaging input preciptation.

4. Surface topography and built environments further undermine the accuracy of predictions based on annual data by introducing errors associated with spatial averaging.

Having understood that boundary condition uncertainty plays an important role in the prediction of subsurface flow and transport, in the next chapter, we will develop two reduced-complexity models to obtain the probabilistic density function of wetting front location and infiltration rate, based on the Green-Ampt and Parlange infiltration models.

Wang, P., Quinlan, P., Tartakovsky, D. M., (2009), 'Effects of spatio-temporal variability of precipitation on contaminant migration in vadose zone'. Geophys. Res. Lett., vol. 36, pp. L12404, doi:10.1029/2009GL038347.

Chapter 3

Reduced Complexity Models for Probabilistic Forecasting of Infiltration Rates

3.1 Introduction

Flow in unsaturated porous media can be routinely described by Richards equation,

$$\frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla \psi) - \frac{\partial K}{\partial x_3}, \qquad K = K_s K_r(\theta), \qquad \theta = \theta(\psi), \tag{3.1}$$

Despite some reservations, e.g., [39, 16], it has become common to treat saturated hydraulic conductivity $K_s(\mathbf{x})$ as a multivariate log-normal random field whose ensemble statistics (e.g., mean, variance, and correlation length) can be inferred

from spatially distributed data by means of geostatistics. No such consensus exists about statistical distributions of various hydraulic parameters entering relative hydraulic conductivity $K_r(\theta)$ and retention curves $\theta(\psi)$. For example, various data analyses concluded that spatial variability of a soil parameter $\alpha_G(\mathbf{x})$ in the Gardner model of relative conductivity, which is often referred to as the reciprocal of the macroscopic capillary length, exhibits either a normal [40] or log-normal [41] distribution and is either correlated [42] or uncorrelated [40] with K_s. We defer a more detailed review of the statistical properties of both $\alpha_G(\mathbf{x})$ and parameters in the van Genuchten model of relative conductivity till section 3.2. Here, it suffices to say that any approach to uncertainty quantification for flow and transport in the vadose zone must be flexible enough to accommodate arbitrary statistical distributions of soil properties.

A large body of literature have attempted to solve the stochastic Richards equation includes [43, 44, 45, 46, 47, 48], to name just a few. With the exception of solutions based on the Kirchhoff transformation [49, 50, 51], such analyses require one to linearize constitutive relations in the Richards equation, introducing errors that are hard to quantify a priori. More important, none of these solutions can be used to estimate the probability of rare events, which is of crucial importance for uncertainty quantification and risk assessment [13]. To the best of our knowledge, no PDF solutions of the Richards equation are reported in the literature.

Given the practical impossibility of obtaining accurate PDF solutions of the stochastic Richards equation (3.1) in three spatial dimensions, we develop reduced

complexity models to compute PDFs of the rate of infiltration into heterogeneous soils with uncertain hydraulic parameters. Construction of such models starts with the selection of a simplified statistical model for soil properties. In the present analysis, we will make use of the Dagan-Bresler statistical parameterization [52], which reduces the spatial dimensionality of random parameter fields. For example, saturated hydraulic conductivity $K_s(\mathbf{x})$—the sole source of uncertainty in [52]—is treated as a two-dimensional *random field*, $K_s(x_1, x_2)$, i.e., a soil is treated as a collection of vertical tubes each of which is characterized by a different *random variable* K_s.

The Dagan-Bresler parameterization [52] enables one to model three-dimensional infiltration with a collection of one-dimensional (in the x_3 direction) solutions of either the Richards equation (3.1) or its approximations, such as the Green-Ampt model [53] and the three-parameter infiltration equation of Parlange et al. [54]. Stochastic analyses of the Richards equation with the Dagan-Bresler parameterization can be found in [55, 56]. Their counterparts based on the Green-Ampt equation were found to yield accurate predictions of infiltration into heterogeneous soils [57] and have been adopted in a number of subsequent investigations, e.g., [58, 59, 60, 61]. These and other similar analyses aimed to derive effective (ensemble averaged) infiltration equations, and some of them quantified predictive uncertainty by computing variances of system states.

Our goal here is to provide a full probabilistic description of infiltration into heterogeneous soils with uncertain parameters (i.e., to compute PDFs of rel-

evant dependent variables) by employing the reduced complexity models based on the Dagan-Bresler parameterization [52] and either the Green-Ampt [53] or the Parlange et al. [54] infiltration equations. From the outset, it is worthwhile emphasizing that the reliance on the Dagan-Bresler parameterization [52] formally limits our analysis to infiltration into top soils, and thus can be used to model surface response to rainfall events [59, 60] and transport phenomena in top soil [61]. Yet it was also used to derive effective properties of the whole vadose zone [62, 41]. Rubin and Or [55] provided an additional justification for the Dagan-Bresler parameterization by noting that "the determination of soil hydraulic properties through field methods...homogenize the properties vertically, thus eliminating the variability in the vertical direction in a practical sense."

Driven by the needs of probabilistic risk assessment, we focus on the derivation of single-point PDFs (rather than the first two moments) of system states describing infiltration into heterogeneous soils with uncertain hydraulic parameters. The two alternative reduced complexity models are formulated in section 3.3, which is preceded (section 3.2) by a brief summary of experimental evidence used to select statistical properties of saturated hydraulic conductivity K_s and fitting parameters in the constitutive laws $K_r = K_r(\theta)$ and $\theta = \theta(\psi)$. Section 3.4 presents analytical closed-form expressions for the infiltration depth and infiltration rate PDFs that can be used as input for probabilistic forecasting of surface runoff and flooding. In section 3.5, we first focus on Green-Ampt model and investigate the temporal evolution of the PDFs of a wetting front and corresponding infiltration

rate, the relative importance of uncertainty in various hydraulic parameters and their cross-correlation, and the impact of the choice of a functional form of K_r.

Similarly, in section 3.5.2, results from Parlange model is investigated to study the temporal evolution of the PDFs of infiltration rate, the relative importance of uncertainty in various hydraulic parameters, and the effects of their cross-correlation. A comparison of the PDFs of infiltration rate obtained from the reduced complexity models with those computed by means of Monte Carlo simulations of the Richards equation (3.1) is presented in section 3.5.3. Key findings of this analysis are summarized in section 3.6.

3.2 Problem Formulation

Consider infiltration into a heterogeneous soil with saturated hydraulic conductivity K_s, porosity ϕ, residual water content θ_r, and relative hydraulic conductivity $K_r(\psi;\alpha)$ that varies with pressure head ψ in accordance with a constitutive model and model parameters α. While the subsequent analysis can be applied to any constitutive relation, we will focus on the Gardner model [53, Table 2.1]

$$K_r = \mathrm{e}^{\alpha_G \psi} \tag{3.2}$$

and the van Genuchten model

$$K_r = \frac{[1-\psi_d^{mn}(1+\psi_d^n)^{-m}]^2}{(1+\psi_d^n)^{m/2}}, \quad \psi_d \equiv \alpha_{vG}|\psi|. \tag{3.3}$$

The model parameters α (α_G and $\{\alpha_{vG}, n, m = 1 - 1/n\}$ for the Gardner and van Genuchten models, respectively) and the rest of hydraulic properties mentioned above vary in space and are sparsely sampled. To quantify uncertainty about values of these properties at points $\mathbf{x} = (x_1, x_2, x_3)^T$ where measurements are unavailable, we treat them as random fields. Thus, a soil parameter $\mathcal{A}(\mathbf{x}, \omega)$ varies not only in the physical domain, $\mathbf{x} \in \mathcal{D}$, but also in the probability space $\omega \in \Omega$. A probability density function $p_{\mathcal{A}}$, which describes the latter variability, is inferred from measurements of $\mathcal{A}$ by invoking ergodicity. Experimental evidence for the selection of PDFs $p_{\mathcal{A}}$ for various soil parameters $\mathcal{A}$ is reviewed in the following section, and the Dagan-Bresler statistical model used in our analysis is formulated in section 3.3.1.

The overreaching aim of the present analysis is to quantify the impact of this parametric uncertainty on predictions of both the dynamics of wetting fronts and infiltration rates. Uncertainty in the former may significantly affect the accuracy and reliability of field-scale measurements of soil saturation [63], while uncertainty in the latter is of fundamental importance to flood forecasting [60].

3.2.1 Saturated hydraulic conductivity

In addition to the experimental studies reviewed in [49], the data analyses reported in [41, 59], etc. support our treatment of saturated hydraulic conductivity K_s as a lognormal random field.

3.2.2 Gardner's constitutive parameter

The (scarce) experimental evidence reviewed in [49] suggests that α_G, the reciprocal of the macroscopic capillary length, can be treated alternatively either as a Gaussian (normal) or as a log-normal random field. While the approach described below is capable of handling both distributions, in the subsequent computational examples we will treat α_G as a log-normal field, which is a model adopted in more recent computational investigations (e.g., [47, 41]).

3.2.3 van Genuchten's constitutive parameters

The van Genuchten hydraulic function (3.3) is a two-parameter model, obtained from its more general form by setting $m = 1 - 1/n$ and $l = 1/2$ (hence the power $m/2$ in the denominator). We employ this form because of its widespread use [53, Table 2.1], but the approach described below can be readily applied to quantify uncertainty in more general formulations with arbitrary m and l. The experimental evidence presented in [2], [41], and [64] shows that the coefficient of variation of α_{vG} is much larger than that of n. These data suggest that α_{vG} can be treated as a log-normal field and the shape factor n as a deterministic constant.

3.2.4 Correlations between hydraulic parameters

Experimental evidence on cross-correlation between K_s and α is inconclusive. Various data sets were used to conclude that these parameters are perfectly

correlated [42], uncorrelated [40], or anti-correlated [62]. The approach we present below is capable of handling an arbitrary degree of cross-correlation between K_s and α. Finally, the data reviewed in [49], as well as more recent data reported in [41], suggest that the coefficient of variation (CV) of K_s is generally much larger than α, i.e., that the former is much more variable than the latter.

Finally, since the difference between the full and residual saturations $\Delta\theta = \phi - \theta_r$ typically exhibits lower spatial variability than both K_s and α, we treat it as a deterministic constant to simplify the presentation. Our approach can be adopted to quantify uncertainty in $\Delta\theta$ and the shape factor n in the van Genuchten hydraulic function, as discussed in section 3.4.

3.3 Reduced Complexity Models

Construction of our reduced complexity models consists of two steps. First, the Dagan-Bresler statistical parameterization is used in section 3.3.1 to represent three-dimensional random fields $K_s(\mathbf{x})$ and $\alpha(\mathbf{x})$ as a collection of corresponding random variables K_s and α. Second, the Richards equation (3.1) is replaced with either the Green-Ampt [53] or the Parlange et al. [54] infiltration equations (3.3.2). To be specific, we consider infiltration under ponding, which is a prerequisite for overland flow [65]. Other infiltration regimes can be handled as well by modifying the Green-Ampt [53] and the Parlange et al. [54] infiltration equations accordingly. The accuracy of the infiltration-rate PDFs predicted with the reduced complexity

models is assessed via comparison with its counterpart obtained from MCS of the Richards equation (3.1) in section 3.5.3.

3.3.1 Statistical Model for Soil Parameters

Following [52], we restrict our analysis to infiltration depths that do not exceed vertical correlation lengths λ_v of (random) soil parameters $\mathcal{A}(\mathbf{x}, \omega)$. Then $\mathcal{A} = \mathcal{A}(x_1, x_2, \omega)$, so that a heterogeneous soil can be represented by a collection of one-dimensional (in the vertical direction x_3) homogeneous columns of length L_3, whose uncertain hydraulic properties are modeled as random variables (rather than random fields). The restriction $\lambda_v > L_3$ formally renders the Dagan-Bresler parameterization [52] suitable for heterogeneous top soils, and thus can be used to model surface response to rainfall events [60, 59] and transport phenomena in top soil [61]. Yet it was also used to derive effective properties of the whole vadose zone [62, 41].

Consider a three-dimensional flow domain $\Omega = \Omega_h \times [0, L_3]$, where Ω_h represents its horizontal extent. A discretization of Ω_h into N elements represents Ω by an assemblage of N columns of length L_3 and facilitates the complete description of a random field $\mathcal{A}(x_1, x_2, \omega)$—in the analysis below, $\mathcal{A}$ stands for K_s and α but can also include other hydraulic properties and the ponding pressure head ψ_0 at the soil surface $x_3 = 0$—with a joint probability function $p_{\boldsymbol{\mathcal{A}}}(A_1, \ldots, A_N)$. Probability density functions (PDFs) of hydraulic properties of the i-th column

are defined as marginal distributions,

$$p_{\mathcal{A}_i}(A_i) = \int p_{\boldsymbol{\mathcal{A}}}(A_1, \ldots, A_N) \mathrm{d}A_1 \ldots \mathrm{d}A_{i-1} \mathrm{d}A_{i+1} \ldots \mathrm{d}A_N. \tag{3.4}$$

Since statistical properties of soil parameters $\mathcal{A}$ are inferred from spatially distributed data by invoking ergodicity, the corresponding random fields (or their fluctuations obtained by data de-trending) must be spatially-stationary so that

$$p_{\mathcal{A}_i} = p_{\mathcal{A}} \quad \text{for} \quad i = 1, \ldots, N. \tag{3.5}$$

If soil parameters (e.g., K_s and α) are correlated, their statistical description requires the knowledge of a joint distribution. For multivariate Gaussian $Y_1 = \ln K_s$ and $Y_2 = \ln \alpha_G$ (or $Y_2 = \ln \alpha_{vG}$), their joint PDF is given by

$$p_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2\pi\sigma_{Y_1}\sigma_{Y_2}\sqrt{1-\rho^2}} \exp\left[-\frac{R}{2(1-\rho^2)}\right] \tag{3.6a}$$

where

$$R = \frac{(y_1 - \overline{Y}_1)^2}{\sigma_{Y_1}^2} - 2\rho\frac{y_1 - \overline{Y}_1}{\sigma_{Y_1}}\frac{y_2 - \overline{Y}_2}{\sigma_{Y_2}} + \frac{(y_2 - \overline{Y}_2)^2}{\sigma_{Y_2}^2}; \tag{3.6b}$$

$\overline{Y}_i$ and σ_{Y_i} denote the mean and standard deviation of Y_i $(i = 1, 2)$, respectively; and $-1 \leq \rho \leq 1$ is the linear correlation coefficient between Y_1 and Y_2. The lack of correlation between Y_1 and Y_2 corresponds to setting $\rho = 0$ in (3.6).

3.3.2 Simplified flow models

During infiltration into top soils, the Dagan-Bresler parameterization of soil heterogeneity can be supplemented with an assumption of vertical flow. The rationale for, and implications of, neglecting the horizontal component of flow velocity

can be found in [52, 55, 58] and other studies reviewed in the introduction. This assumption obviates the need to solve a three-dimensional flow problem, replacing the latter with a collection of N one-dimensional flow problems to be solved in homogeneous soil columns with random but constant hydraulic parameters.

The second step in the construction of our reduced complexity models for probabilistic estimation of infiltration rates $i(t)$ replaces the Richards equation (3.1) with either the Green-Ampt [53] or Parlange et al. [54] infiltration equations. The accuracy of these reduced complexity models is investigated in section 3.5.3 via comparison with Monte Carlo solutions of the two-dimensional Richards equation (3.1).

As mentioned above, we consider the Green-Ampt [53] and Parlange et al. [54] infiltration equations corresponding to ponding water of height ψ_0 at the soil surface $x_3 = 0$. Other infiltration scenarios can be handled in a similar manner by modifying these equations as discussed in the closing of this section.

Green-Ampt infiltration model

The Green-Ampt model provides an alternative description of flow in partially saturated porous media. The relative simplicity of the Green-Ampt formulation makes it easier to solve than the Richards equation, which explains its prevalence in large numerical codes—e.g., SCS developed by US EPA, DR3M developed by USGS, and HIRO2 developed by USDA—that are routinely used to predict overland and channel flows.

Let $I(t)$ denote (uncertain) cumulative infiltration due to ponding water of height ψ_0 at the soil surface $x_3 = 0$. The Green-Ampt model of infiltration approximates an S-shaped wetting front with a sharp interface (infiltration depth) $x_\mathrm{f}(t)$, which separates the uniformly "wet" ($\theta = \theta_\mathrm{wet}$) region behind the wetting front from a partially-saturated region with a uniform water content $\theta = \theta_\infty$ ahead of the front. To be specific, and without loss of generality, we set $\theta_\mathrm{wet} = \phi$ and $\theta_\infty = \theta_i$. If the x_3 coordinate is positive downward, Darcy's law defines macroscopic (Darcy's) flux q as (e.g., [53, Eq. 5-1])

$$q = -K_s \frac{\psi_\mathrm{f} - x_\mathrm{f} - \psi_0}{x_\mathrm{f}}. \tag{3.7}$$

Pressure head, ψ_f, at the infiltration depth $x_\mathrm{f}(t)$ is often set to a "capillary drive",

$$\psi_\mathrm{f} = -\int_{\psi_i}^{0} K_r(\psi)\,\mathrm{d}\psi, \tag{3.8}$$

where ψ_i is the pressure head corresponding to the water content θ_i. Theoretical derivations of this equation can be found in [66] and [67].

Mass conservation requires that the infiltration rate $i = q$, and that $i = \Delta\theta\,\mathrm{d}x_\mathrm{f}/\mathrm{d}t$ where $\Delta\theta = \phi - \theta_i$. Combined with (3.7), this leads to an implicit expression for the infiltration depth $x_\mathrm{f}(t)$,

$$x_\mathrm{f} - (\psi_0 - \psi_\mathrm{f}) \ln\left(1 + \frac{x_\mathrm{f}}{\psi_0 - \psi_\mathrm{f}}\right) = \frac{K_s}{\Delta\theta} t. \tag{3.9}$$

which is applicable to time intervals during which the height of ponding water, ψ_0, remains approximately constant. Substituting $\psi_\mathrm{f}(t)$ from (3.9) into (3.7) yields a Green-Ampt solution for the infiltration rate $i(t)$.

It is worthwhile emphasizing that several of the simplifying assumptions made above can be easily relaxed. First, since K_s and $\Delta\theta$ enter the stochastic solution (3.9) as the ratio $K_s^\star = K_s/\Delta\theta$ one can easily incorporate uncertainty in (randomness of) $\Delta\theta$ by replacing the PDF of K_s with the PDF of $K_s^\star$. Second, the implicit relation $F(x_\mathrm{f}, K_s/\Delta\theta, \alpha; t) = 0$ given by (3.9) and (3.8) allows one to express the PDF of x_f in terms of the PDFs of *any* number of hydraulic parameters by following the procedure described below. Our goal is to relate uncertainty in hydraulic parameters K_s and α to predictive uncertainty about the infiltration depth $x_\mathrm{f}(t)$ and the infiltration rate $i(t)$, i.e., to express the PDFs of the latter, $p_\mathrm{f}(x_\mathrm{f}; t)$ and $p_\mathrm{i}(i; t)$, in terms of the PDF of the former (3.6).

Parlange infiltration model

The Parlange et al. [54] infiltration model seeks to preserve a sigmoidal shape of infiltration fronts by postulating a functional form of the soil water diffusivity $D(\theta) \equiv K\,\mathrm{d}\psi/\mathrm{d}\theta$. Under ponded conditions, this equation takes the form [68],

$$I - K_i t = (\psi_0 + \psi_j)\frac{\Delta\theta K_s}{i - K_s} + \frac{S^2 - 2\psi_j K_s \Delta\theta}{2\Delta K}\ln\left(1 + \frac{\Delta K}{i - K_s}\right), \tag{3.10}$$

where cumulative infiltration rate $I(t)$ is related to infiltration rate $i(t)$ by $i = \mathrm{d}I/\mathrm{d}t$, and $\Delta K \equiv K_s - K_i$ with $K_i \equiv K(\theta_i)$. Following [69, 68], we approximate soil sorptivity S by

$$S^2 = \int_{\theta_i}^{\phi} (\phi + \theta - 2\theta_i) D(\theta)\mathrm{d}\theta. \tag{3.11}$$

Finally, the parameter ψ_j ($\psi_j < \psi$) represents a small pressure jump at saturation that is typically observed in soil-water characteristic curves. This soil parameter depends on the local pore structure, has limited range and effect on infiltration prediction, and remans "constant in time and independent of changing boundary conditions" [68]. Consequently, we treat ψ_j as a deterministic constant.

For constant ponding water heights ψ_0, solving (3.10) yields an implicit expression for the infiltration rate $i(t)$ [68],

$$\begin{aligned} t =& \frac{K_s(\psi_0 + \psi_j)\Delta\theta}{(i - K_s)\Delta K} - \frac{S^2 - 2\psi_j K_s \Delta\theta}{2\Delta K(i - K_i)} \\ &+ \frac{S^2 - 2K_s\Delta\theta(\psi_0 + 2\psi_j)}{2(\Delta K)^2} \ln\left(1 + \frac{\Delta K}{i - K_s}\right). \end{aligned} \tag{3.12}$$

For brevity, we will call this expression the Haverkamp solution for infiltration under ponded conditions, after the first author of [68]. Note that (3.12) reduces to (3.9) if one sets the soil water diffusivity $D(\theta)$ to be a delta function [53, pp. 159-161].

In conclusion, analytical solutions (3.9) and (3.12) correspond to ponded conditions with constant water heights ψ_0. Our reduced complexity models can handle other infiltration regimes by replacing (3.9) and (3.12) with their appropriate counterparts. For example, (3.12) can be replaced with the analytical solutions in [70] or [71] if infiltration is driven respectively by atmospheric pressure at the soil surface ($\psi_0 = 0$) or by temporally varying ponded water height $\psi_0(t)$. Likewise, infiltration under non-ponded conditions can be handled by replacing (3.9) with appropriately modified Green-Ampt solutions, many of which can be found

in [53]. What is important is that a properly chosen reduced complexity model provides a mapping $i = i(K_s, \alpha)$.

3.4 PDF Solutions

3.4.1 PDFs of infiltration depth

For small t, (3.9) can be approximated by an explicit relation [53, Eq. 5.12]

$$x_{\mathrm{f}} \approx \sqrt{\frac{2(\psi_0 - \psi_{\mathrm{f}})K_s t}{\Delta\theta}}. \tag{3.13}$$

For large t, flow becomes gravity dominated, $i \sim K_s$, and [53, p. 170]

$$x_{\mathrm{f}} \approx \frac{K_s}{\Delta\theta} t. \tag{3.14}$$

For intermediate t, various approximations, e.g., [72] and [53, p. 170], can be used to replace the implicit solution (3.9) with its explicit counterparts. We will use the implicit solution (3.9) to avoid unnecessary approximation errors.

Let $G_{\mathrm{f}}(x_{\mathrm{f}}^\star) = P(x_{\mathrm{f}} \leq x_{\mathrm{f}}^\star)$ denote the cumulative distribution function of x_{f}, i.e., the probability that the random position of the wetting front x_{f} takes on a value not larger than $x_{\mathrm{f}}^\star$. Since (3.9) provides an explicit dependence of random K_s on random x_{f} and α , i.e.,

$$K_s(x_{\mathrm{f}}, \alpha) = \frac{\Delta\theta}{t}\left[x_{\mathrm{f}} - (\psi_0 - \psi_{\mathrm{f}}) \ln\left(1 + \frac{x_{\mathrm{f}}}{\psi_0 - \psi_{\mathrm{f}}}\right)\right], \tag{3.15}$$

it follows from the definition of a cumulative distribution function that

$$G_{\mathrm{f}}(x_{\mathrm{f}}^\star) = \int_0^\infty \int_0^{K_s(x_{\mathrm{f}}^\star, \alpha)} p_{Y_1, Y_2}(K_s, \alpha) \frac{\mathrm{d}K_s \mathrm{d}\alpha}{K_s \alpha}. \tag{3.16}$$

The denominator in (3.16) reflects the transition from (3.6), the joint Gaussian PDF for Y_1 and Y_2, to the lognormal variables $K_s = \exp(Y_1)$ and $\alpha = \exp(Y_2)$.

The PDF of the random (uncertain) infiltration depth, $p_{\rm f}(x_{\rm f}^\star; t)$, can now be obtained as

$$p_{\rm f}(x_{\rm f}^\star; t) = \frac{{\rm d}G_{\rm f}(x_{\rm f}^\star; t)}{{\rm d}x_{\rm f}^\star}. \tag{3.17}$$

Using Leibnitz's rule to compute the derivative of the integral in (3.16) and (3.17), we obtain

$$p_{\rm f}(x_{\rm f}^\star; t) = \int_0^\infty \frac{p_{Y_1,Y_2}[K_s(x_{\rm f}^\star, \alpha), \alpha]}{\alpha K_s(x_{\rm f}^\star, \alpha)} \frac{\partial K_s(x_{\rm f}^\star, \alpha)}{\partial x_{\rm f}^\star} {\rm d}\alpha. \tag{3.18}$$

It is worthwhile emphasizing that (3.18) holds for an arbitrary implicit solution of the Green-Ampt equation, $F(x_{\rm f}, K_s/\Delta\theta, \alpha; t) = 0$, and, hence, the PDF solution (3.18) is applicable to a large class of infiltration regimes that are amenable to the Green-Ampt description. For the flow regime considered in the present analysis, $K_s(x_{\rm f}^\star, \alpha)$ is given by (3.15), and (3.18) takes the form

$$p_{\rm f}(x_{\rm f}^\star; t) = \frac{\Delta\theta}{t} \int_0^\infty \frac{p_{Y_1,Y_2}[K_s(x_{\rm f}^\star, \alpha), \alpha]}{\alpha K_s(x_{\rm f}^\star, \alpha)} \frac{x_{\rm f}^\star \, {\rm d}\alpha}{\psi_0 - \psi_{\rm f} + x_{\rm f}^\star}. \tag{3.19}$$

3.4.2 PDFs for Infiltration Rate

Let $G_{\rm i}(i^\star; t) = P[i \le i^\star]$ denote the cumulative distribution function of i at time t, i.e., the probability that the random infiltration rate i at time t does not exceed some value $i^\star$. Equations (3.7)–(3.9) and (3.12) define mappings $i = i(K_s, \alpha)$ for the two alternative reduced complexity models. These mappings enable

one to compute the cumulative distribution function $G_{\mathrm{i}}(i^\star;t)$ as

$$G_{\mathrm{i}}(i^\star;t)=\int_0^\infty\int_0^{\alpha(i^\star,K_s)} p_{Y_1,Y_2}(K_s,\alpha)\frac{\mathrm{d}\alpha\,\mathrm{d}K_s}{\alpha K_s}. \tag{3.20}$$

The PDF of the random (uncertain) infiltration rate, $p_{\mathrm{i}}=\mathrm{d}G_{\mathrm{i}}/\mathrm{d}i^\star$, is

$$p_{\mathrm{i}}(i^\star;t)=\int_0^\infty \frac{p_{Y_1,Y_2}[\alpha(i^\star,K_s),K_s]}{\alpha(i^\star,K_s)K_s}\frac{\partial\alpha(i^\star,K_s)}{\partial i^\star}\mathrm{d}K_s. \tag{3.21}$$

While the analysis above deals with two uncertain parameters, K_s and α, it can be readily generalized to account for uncertainty in other soil parameters, such as the van Genuchten parameter n. If M soil properties are uncertain then their statistics are characterized by a joint PDF, $p_{Y_1,\ldots,Y_M}$; the cumulative distribution function G_{i} in (3.20) is defined in terms of an M dimensional integral; and the subsequent derivation is modified accordingly.

Green-Ampt infiltration model

Computation of the infiltration-rate PDF, p_{i} is facilitated by the change of the integration variable in (3.21),

$$p_{\mathrm{i}}(i^\star;t)=\int_0^\infty \frac{p_{Y_1,Y_2}[K_s(i^\star,\alpha),\alpha]}{\alpha K_s(i^\star,\alpha)}\frac{\partial K_s(i^\star,\alpha)}{\partial i^\star}\mathrm{d}\alpha. \tag{3.22}$$

Here p_{Y_1,Y_2} and $K_s(i^\star,\alpha)$ are given by (3.6) and (3.7), respectively; and the derivative $\partial K_s/\partial i^\star$ is obtained from (3.7) as the reciprocal of

$$\frac{\partial i^\star}{\partial K_s}=1+\frac{\psi_0-\psi_{\mathrm{f}}}{x_{\mathrm{f}}}\left(1-\frac{K_s t}{\Delta\theta}\frac{x_{\mathrm{f}}-\psi_{\mathrm{f}}+\psi_0}{x_{\mathrm{f}}^2}\right). \tag{3.23}$$

Parlange infiltration model

For the van Genuchten constitutive relation (3.3), the soil sorptivity S in (3.11) takes the form

$$S^2 = \frac{K_s \Delta\theta}{\alpha_{vG}}(1-m)A(m), \tag{3.24a}$$

where m is the van Genuchten model shape parameter, $A(m)$ is given by

$$\begin{aligned} A = &\frac{\Gamma(1-m)\Gamma(3m/2-1)}{\Gamma(m/2)} - \frac{4}{3m-2} + \frac{\Gamma(m+1)\Gamma(3m/2-1)}{\Gamma(5m/2)} \\ &+ \frac{\Gamma(1-m)\Gamma(5m/2-1)}{\Gamma\left(3m/2\right)} - \frac{4}{5m-2} + \frac{\Gamma(m+1)\Gamma(5m/2-1)}{\Gamma(7m/2)}, \end{aligned} \tag{3.24b}$$

and $\Gamma(\cdot)$ is the complete Gamma function. For the sake of simplicity, and without loss of generality, we assume that the soil ahead of the wetting front is "dry", and set $\psi_i = -\infty$. (Other values of ψ_i can be handled as well by following the procedure outlined below.) Then $K_i = 0$ and substituting (3.24) into (3.12) yields an explicit relation between the three random variables $\alpha_{vG} = \alpha_{vG}(i, K_s)$,

$$\alpha_{vG}(i, K_s) = A\left[K_s - i\ln\left(\frac{i}{i-K_s}\right)\right]\frac{i-K_s}{2B(i,K_s)} \tag{3.25a}$$

where

$$\begin{aligned} B(i,K_s) = &(\psi_0+\psi_{str})iK_s - (\psi_0+2\psi_{str})i(i-K_s)\ln\left(\frac{i}{i-K_s}\right) \\ &+ K_s(i-K_s)(\psi_{str} - \frac{it}{\Delta\theta}). \end{aligned} \tag{3.25b}$$

Substituting (3.25) into (3.21) gives the infiltration-rate PDF,

$$p_{\mathrm{i}}(i^\star;t) = \frac{A}{2}\int_0^\infty \frac{K_s p_{Y_1,Y_2}[\alpha_{vG}(i^\star,K_s),K_s]}{\alpha_{vG}(i^\star,K_s)B(i^\star,K_s)^2}\Big\{\frac{K_s-i^\star}{\Delta\theta}K_s t \\ -(\psi_0+\psi_{str})\left[(2i^\star-K_s)\ln\left(\frac{i^\star}{i^\star-K_s}\right)-2K_s\right]\Big\}\mathrm{d}K_s. \tag{3.26}$$

Multi-point pdfs

As discussed in section 3.3.1, a complete description of the random infiltration rate $i(\mathbf{x},t)$ in the domain discretized into N elements requires the knowledge of an N-point PDF, $p_{\mathbf{i}}(i_1^\star,\ldots,i_N^\star;t)$, where $i_k^\star$ is a deterministic value (outcome) of the random infiltration rate i at the k-th column ($k=1,\ldots,N$). The reduced complexity models presented in section 3.3 allow one to compute such multi-point PDFs.

Consider a two-point PDF, $p_{\mathbf{i}}^{(2)}(i_1^\star,i_2^\star;t)$, which describes a joint distribution of infiltration rates $i(\mathbf{x}^k,t)$ ($k=1,2$) at points $\mathbf{x}^1=(x_1^1,x_2^1)^T$ and $\mathbf{x}^2=(x_1^2,x_2^2)^T$. Let $Y_{1,k}=\ln K_s(\mathbf{x}^k)$ and $Y_{2,k}=\ln\alpha(\mathbf{x}^k)$, with the joint two-point PDF $p_{Y_1,Y_2}^{(2)}(Y_{1,1}^\star,Y_{2,1}^\star;Y_{1,2}^\star,Y_{2,2}^\star)$. Recalling that (3.7)–(3.9) and (3.12) define the two alternative mappings $i=i(K_s,\alpha)$, we compute, in analogy with (3.20), the two-point cumulative distribution function $G_{\mathbf{i}}^{(2)}(i_1^\star,i_2^\star;t)$ as

$$G_{\mathbf{i}}^{(2)}(i_1^\star,i_2^\star;t) = \int_0^\infty\int_0^\infty\int_0^{\alpha_1(i_1^\star,K_{s1})}\int_0^{\alpha_2(i_2^\star,K_{s2})} p_{Y_1,Y_2}^{(2)}(K_{s1},\alpha_1;K_{s2},\alpha_2) \\ \times\frac{\mathrm{d}\alpha_1\mathrm{d}K_{s1}\mathrm{d}\alpha_2\mathrm{d}K_{s2}}{\alpha_1 K_{s1}\alpha_2 K_{s2}}. \tag{3.27}$$

The two-point PDF of the random (uncertain) infiltration rate, $p_{\mathbf{i}}^{(2)}(i_1^\star,i_2^\star;t)$, is

obtained as

$$p_{\mathbf{i}}^{(2)}(i_1^\star, i_2^\star; t) = \frac{\partial^2 G_{\mathbf{i}}^{(2)}}{\partial i_1^\star \partial i_2^\star}. \tag{3.28}$$

N-point PDFs, $p_{\mathbf{i}}^{(N)}$ with $N > 2$, can be computed in a similar manner.

N-point PDFs can be used both to predict (cross-)correlations of infiltration rates at multiple locations and to assimilate infiltration data via a straightforward Bayesian updating. We leave the latter aspect for future investigation.

3.5 Results and Discussion

3.5.1 Green-Ampt infiltration model

In this subsection, we explore the impact of various aspects of parametric uncertainty on the uncertainty in predictions of infiltration rate $i(t)$ and infiltration depth $x_{\mathrm{f}}(t)$. Specifically, we investigate the temporal evolution of the PDFs of the wetting front and the infiltration rate, the relative importance of uncertainty in K_s and α, and the effects of cross-correlation between them. This is done for the Gardner hydraulic function (3.2), in which case (3.8) results in the interfacial pressure head $\psi_{\mathrm{f}} = -\alpha_G^{-1}$. In conclusion, we explore how the choice of a functional form of the hydraulic function, i.e., the use of the van Genuchten model (3.3) instead of the Gardner relation (3.2), affects the predictive uncertainty.

Dimensionless form of PDFs

To facilitate an analysis of the effects of various sources of parametric uncertainty on the PDF, $p_{\rm f}(x_{\rm f}^{\star};t)$, of the uncertain (random) infiltration depth $x_{\rm f}(t)$, given by the analytical solution (3.19), we introduce the following dimensionless quantities. Let the averaged quantities $(\overline{\alpha})^{-1}$ and $\overline{K}_s$ represent a characteristic length scale and a characteristic value of saturated hydraulic conductivity, respectively. Then a characteristic time scale τ can be defined as

$$\tau = (\overline{\alpha}\overline{K}_s)^{-1}, \tag{3.29}$$

and the following dimensionless quantities can be introduced,

$$t' = \frac{t}{\tau}, \quad \psi' = \overline{\alpha}\psi, \quad \alpha' = \frac{\alpha}{\overline{\alpha}}, \quad K_s' = \frac{K_s}{\overline{K}_s}. \tag{3.30}$$

This leads to a PDF solution for the dimensionless infiltration depth $x_{\rm f}' = \overline{\alpha}x_{\rm f}$,

$$p_{\rm f}(x_{\rm f}';t') = \frac{\Delta\theta}{t'}\int_0^\infty \frac{p_{Y_1',Y_2'}[K_s'(x_{\rm f}',\alpha'),\alpha']}{\alpha' K_s'(x_{\rm f}',\alpha')}\frac{x_{\rm f}'\,{\rm d}\alpha'}{\psi_0' - \psi_{\rm f}' + x_{\rm f}'}. \tag{3.31}$$

Likewise, the PDF of the dimensionless infiltration rate $i' = i/\overline{K}_s$ takes the form

$$p_{\rm i}(i';t') = \int_0^\infty \frac{p_{Y_1',Y_2'}[K_s'(i',\alpha'),\alpha']}{\alpha' K_s'(i',\alpha')}\frac{\partial K_s'(i',\alpha')}{\partial i'}{\rm d}\alpha'. \tag{3.32}$$

In the following, we drop the primes to simplify the notation. Unless explicitly noted otherwise, the simulations reported below correspond to the dimensionless height of ponding water $\psi_0 = 0.1$, $\Delta\theta = 0.45$, the coefficients of variation $CV_{\ln K_s} \equiv \sigma_{Y_1}/\overline{Y}_1 = 3.0$ and $CV_{\ln \alpha_G} \equiv \sigma_{Y_2}/\overline{Y}_2 = 0.5$ with the means $\overline{Y}_1 = 0.25$ and $\overline{Y}_2 = 0.1$, and the cross-correlation coefficient $\rho = 0$. (The use of the soil data in Table 1 of

[2] in conjunction with these dimensionless parameters would result in the height of ponding water $\psi_0 = 0.6$ cm.)

PDF of wetting front

Since the initial position of the wetting front is assumed to be known, $x_{\mathrm{f}}(t = 0) = 0$, the PDF $p_{\mathrm{f}}(x_{\mathrm{f}}; 0) = \delta(x_{\mathrm{f}})$ where $\delta(\cdot)$ denotes the Dirac delta function. As the dimensionless time becomes large ($t \to \infty$), $p_{\mathrm{f}} \sim p_{K_s}$ in accordance with (3.14). The PDF $p_{\mathrm{f}}(x_{\mathrm{f}}; t)$ in (3.31) describes the temporal evolution of predictive uncertainty between these two asymptotes, with Figure 3.1 providing snapshots at dimensionless times $t = 0.01$, 0.1 and 0.5. (For the soil parameters reported in Table 1 of [2] this corresponds to dimensional times 1.5, 15 and 75 min, respectively). The uncertainty in predictions of infiltration depth increases rapidly, as witnessed by wider distributions with longer tails.

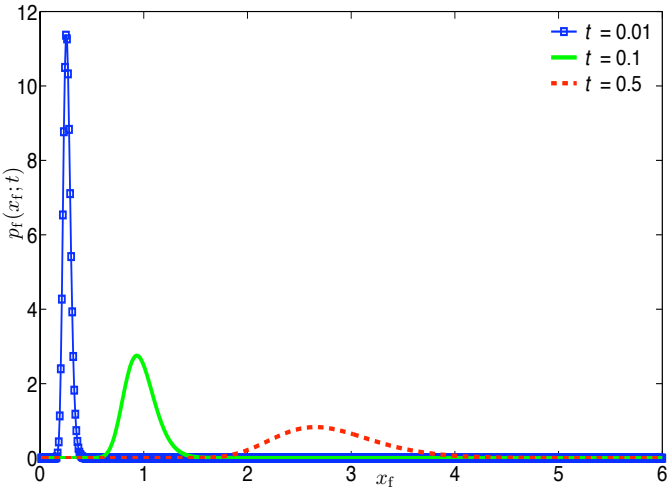

Figure 3.1: Temporal evolution of the PDFs of infiltration depth from Green-Ampt model, $p_{\mathrm{f}}(x_{\mathrm{f}}; t)$.

PDF of infiltration rate

Figure 3.2 provides snapshots, at dimensionless times $t = 0.01$, 0.1 and 0.5, of the temporal evolution of the PDF of infiltration rate, $p_{\mathrm{i}}(i;t)$, given by (3.32). Both the mean infiltration rate and the corresponding predictive uncertainty decrease with time. At later times (the dimensionless time $t = 5.0$, for the parameters used in these simulations), the PDF appears to become time invariant. This is to be expected on theoretical grounds, see (3.14), according to which $p_{\mathrm{i}}(i';t') \to p_K(K'_s)$ as $t' \to \infty$. The reduced χ^2 test confirmed this asymptotic behavior at dimensionless time $t = 100.0$.

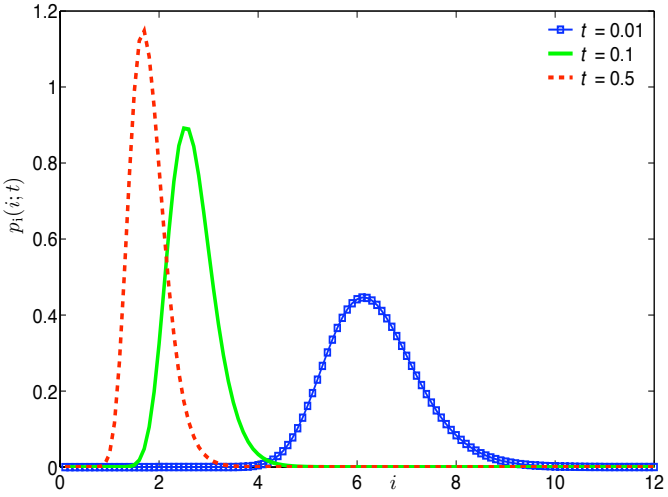

Figure 3.2: Temporal evolution of the PDF of the infiltration rate from Green-Ampt model, $p_{\mathrm{i}}(i;t)$.

Effects of parametric uncertainty

The degree of uncertainty in hydraulic parameters $\ln K_s$ and $\ln \alpha_G$ is encapsulated in their coefficients of variation $CV_{\ln K_s}$ and $CV_{\ln \alpha_G}$, respectively. Figure 3.3 demonstrates the relative effects of these two sources of uncertainty upon the predictive uncertainty, as quantified by the infiltration depth PDF $p_{\mathrm{f}}(x_{\mathrm{f}}; t)$, computed at $t = 0.1$. Uncertainty in saturated hydraulic conductivity K_s affects predictive uncertainty more than uncertainty in the Gardner parameter α_G does. Although not shown in Figure 3.3, we found similar behavior at later times $t = 0.5$ and 1.0. These findings are in agreement with those reported in [52, 73], wherein variances of state variables were used to conclude that uncertain saturated hydraulic conductivity K_s is the dominant factor affecting predictive uncertainty.

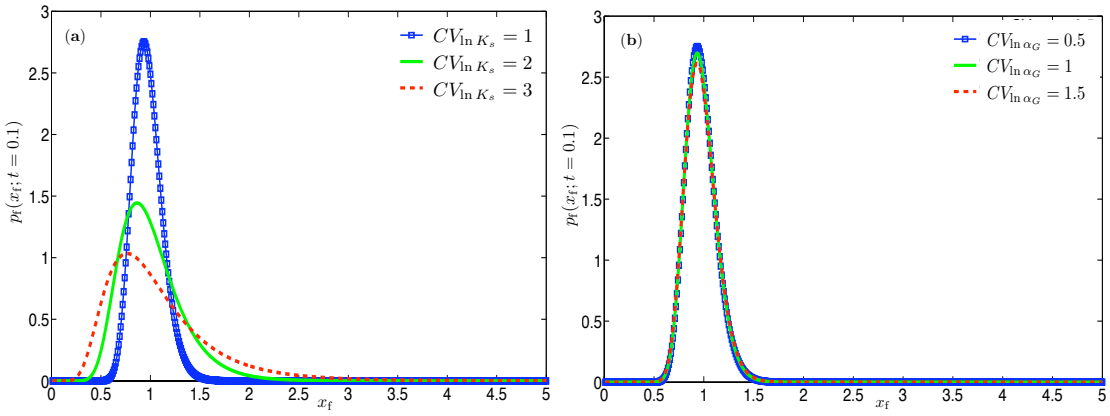

Figure 3.3: The infiltration depth PDF $p_{\mathrm{f}}(x_{\mathrm{f}}; t = 0.1)$ from Green-Ampt model for different levels of uncertainty in (a) saturated hydraulic conductivity K_s and (b) the Gardner parameter α_G.

Effects of cross-correlation

The question of whether various hydraulic parameters are correlated with each other remains open, with different data sets supporting opposite conclusions (see section 3.2). This suggests that the presence or absence of such cross-correlations is likely to be site-specific rather than universal. The general PDF solution (3.31) enables us to investigate the impact of cross-correlations between saturated hydraulic conductivity K_s and the Gardner parameter α_G on predictive uncertainty. This is done by exploring the dependence of the PDF of the wetting front, $p_{\mathrm{f}}(x_{\mathrm{f}}; t)$, on the correlation coefficient ρ. Figure 3.4 presents $p_{\mathrm{f}}(x_{\mathrm{f}}; t = 0.1)$ for $\rho = -0.99$, 0.0 and 0.99, which represent perfect anti-correlation, independence and perfect correlation between K_s and α_G, respectively. The perfect correlation between K_s and α_G ($\rho = 0.99$) results in the minimum predictive uncertainty (the width of the distribution), while the perfect anti-correlation ($\rho = -0.99$) leads to the maximum predictive uncertainty. Predictive uncertainty resulting from the lack of correlation between K_s and α_G ($\rho = 0.0$) falls amid these two limits. The impact of cross-correlation between soil hydraulic parameters (a value of ρ) decreases with time, falling from the maximum difference of about 21% at $t = 0.01$ to about 3% at $t = 0.1$.

Effects of selection of hydraulic function

Finally, we examine how the choice of a hydraulic function $K_r(\psi; \alpha)$ affects predictive uncertainty. Guided by the data analyses presented in section 3.2, we

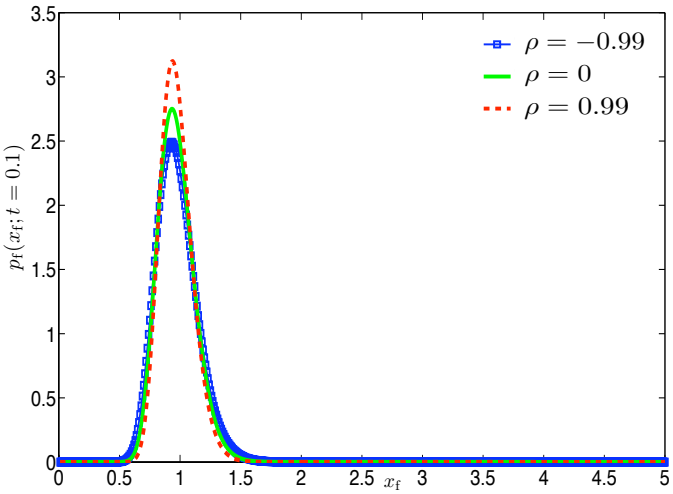

Figure 3.4: The infiltration depth PDF $p_{\mathrm{f}}(x_{\mathrm{f}}; t = 0.1)$ from Green-Ampt model for different levels of correlation ρ between hydraulic parameters K_s and α_G.

treat α_{vG} as the only uncertain parameter in the van Genuchten hydraulic function with $n = 1.5$. To make a meaningful comparison between predictions based on the Gardner (3.2) and van Genuchten (3.3) relations, we select statistics of their respective parameters α in a way that preserves the mean effective capillary drive defined by (3.8) [74, 38]. Specifically, we use the equivalence criteria to select the mean of $\ln \alpha_{vG}$ (-1.40, for the parameters used in these simulations) that maintains the same mean capillary drive as the Gardner model with $\overline{\ln \alpha_G} = 0.1$, and choose the variance of $\ln \alpha_{vG}$ as to maintain the original values of the coefficients of variation $CV_{\ln \alpha_{vG}} = CV_{\ln \alpha_G} = 0.5$. Figure 3.5 reveals that the choice between the van Genuchten and Gardner models has a significant effect on predictive uncertainty of the wetting front dynamics, although this influence diminishes with time. For example, the difference between the variances is 40% at $t = 0.01$ and 23% at $t = 0.1$.

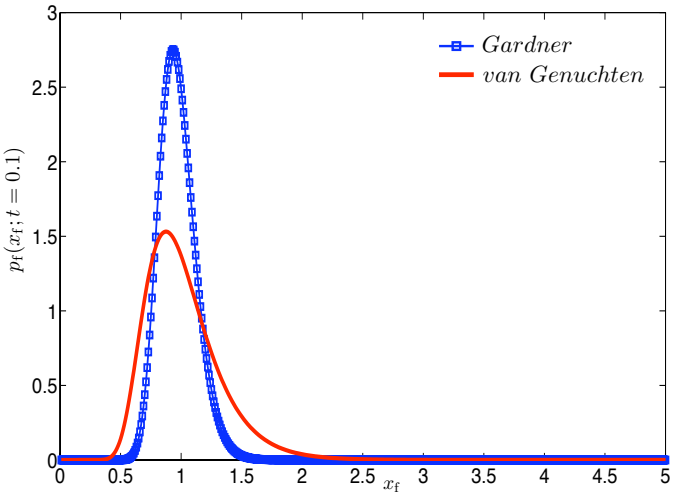

Figure 3.5: The infiltration depth PDF $p_f(x_f; t = 0.1)$ from Green-Ampt model, resulted from the use of the Gardner and van Genuchten hydraulic functions.

3.5.2 Parlange infiltration model

Here we carry out a similar analysis for the infiltration-rate PDF predicted with the Parlange model (Haverkamp solution). Unless stated otherwise, van Genuchten hydraulic functions are used. Specifically, we investigate temporal evolution of the infiltration-rate PDF, the relative importance of uncertainty in K_s and α_{vG} and the effects of cross-correlation between them.

To be concrete, we use the Bet-Dagan soil properties [2] reported in Table 3.1. Unless explicitly noted otherwise, the simulations reported below correspond to theponding water height $\psi_0 = 1$ cm, pressure jump $\psi_j = 2$ cm, and the cross-correlation coefficient $\rho = 0$.

Table 3.1: Hydraulic properties of the Bet-Dagan soil [2, Table 3].

	$\ln K_s$ (cm/min)	$\ln \alpha_{vG}$ (cm^{-1})	ϕ	θ_i	ψ_0 (cm)	ψ_j (cm)	van Genuchten n
mean	-3.58	-3.01	0.42	0.13	1	2	1.81
variance	0.89	0.63	-	-	-	-	-

Temporal evolution of infiltration rate PDFs

Figure 3.6 presents three snapshots of the temporal evolution of the infiltration-rate PDF, $p_{\text{i}}(i;t)$, at times $t = 5$, 50 and 100 min. Uncertainty associated with predictions of the infiltration rate under ponded conditions (i.e., the width of p_{i}) decreases with time. This is because, as time increases, top soil gradually saturates and the infiltration rate $i(t)$ approaches an (uncertain) value of the saturated hydraulic conductivity K_s in accordance with (3.12), i.e., $p_{\text{i}}(i^\star; t \to \infty) \to p_K(K_s^\star)$. It must be noted that at large times, the infiltration depth exceeds the vertical

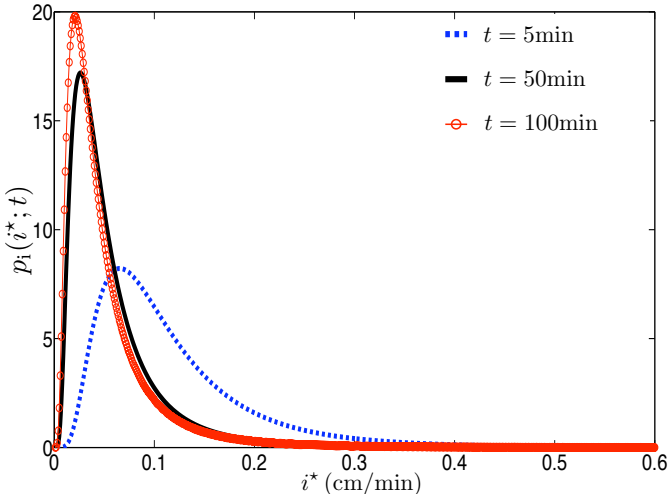

Figure 3.6: Temporal evolution of the infiltration-rate PDF $p_{\text{i}}(i^\star;t)$ given by (3.26).

correlation lengths of K_s and α_{vG}, which violates the conditions of validity of the reduced complexity models. Therefore, our analysis is formally limited to early infiltration times and ought to be used to compute the infiltration-rate PDFs that are necessary for probabilistic forecasting of surface runoff and flooding where uncertainty in infiltration rate predictions is highest (Fig. 3.6). Unless otherwise noted the subsequent figures correspond to $t = 5$ min.

The infiltration-rate PDFs exhibit long tails that superficially resemble those of lognormal distributions. To test whether a lognormal distribution $p_{\ln}(i^\star; t)$ can be used to approximate $p_{\mathrm{i}}(i^\star; t)$ in (3.26), we compute a relative error $\mathcal{E} \equiv 100\% \times |p_{\mathrm{i}} - p_{\ln}|/p_{\mathrm{i}}$. Both distributions have the same mean and variance. Figure 3.7 reveals a significant discrepancy between the tails of the two distributions (probabilities of rare events).

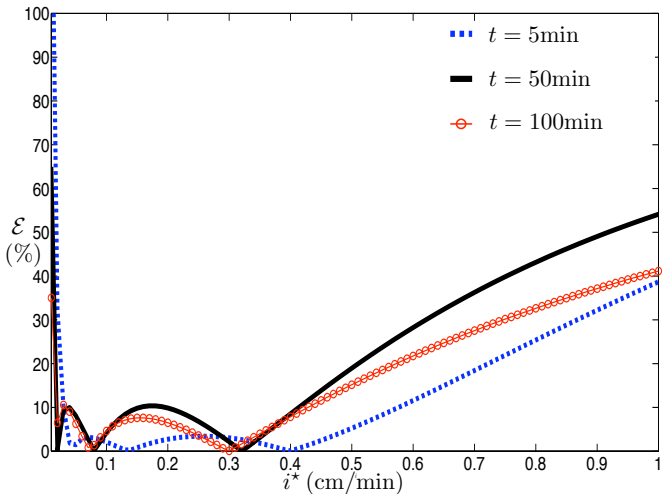

Figure 3.7: Relative error, $\mathcal{E}$, introduced by approximating the infiltration-rate PDF $p_{\mathrm{i}}(i^\star; t)$ with its lognormal counterpart $p_{\ln}(i^\star; t)$, from Parlange model.

Effects of parametric uncertainty

While the proposed approach can handle uncertainty in any number of hydraulic parameters, we focus on K_s and α_{vG} for the reasons discussed above. In this section, we investigate the relative importance of these two sources of parametric uncertainty. Uncertainty in both $\ln K_s$ and $\ln \alpha_{vG}$ is encapsulated in their respective coefficients of variation, $CV_{\ln K_s} \equiv \sigma_{Y_1}/\overline{Y}_1$ and $CV_{\ln \alpha_{vG}} \equiv \sigma_{Y_2}/\overline{Y}_2$. Figure 3.8 demonstrates their effects on predictive uncertainty (PDF of i at $t = 5$ min). The curves represent $p_{\mathrm{i}}(i^\star; t = 5 \text{ min})$ for the CV of one parameter set to 0.1, 0.3, 0.5 and the other parameter fixed at its value in Table 3.1. One can see that uncertainty in K_s has more pronounced effect on the predictive uncertainty than uncertainty in α_{vG} does. This finding is in accordance with previous observations [73, 52, 75].

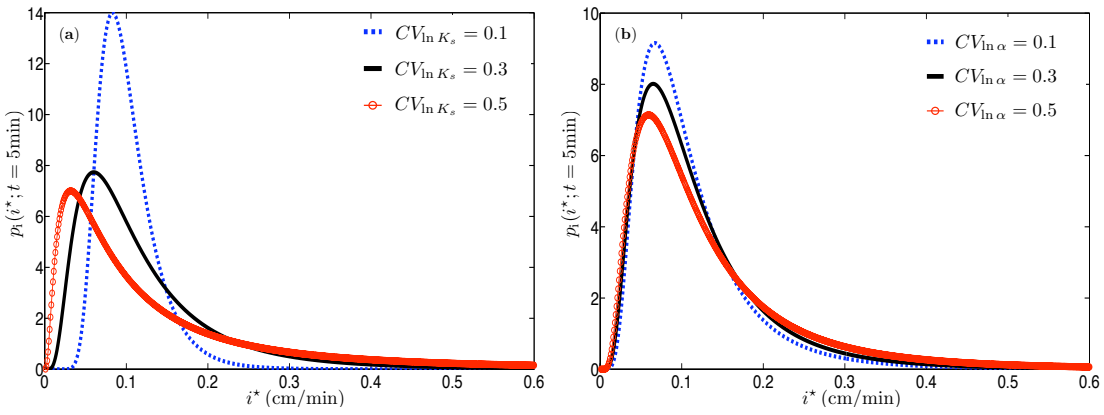

Figure 3.8: The infiltration-rate PDF $p_{\mathrm{i}}(i^\star; t = 5\text{min})$ from Parlange model for different levels of uncertainty in (a) saturated hydraulic conductivity K_s and (b) the van Genuchten parameter α_{vG}.

Effects of cross-correlation

The data reviewed in section 3.2 suggest that the presence, absence, or strength of cross-correlation between saturated hydraulic conductivity K_s and the van Genuchten parameter α_{vG} is site-specific rather than universal. Our reduced complexity models allow one to investigate the role of this cross-correlation on predictive uncertainty in infiltration rates $i(t)$. Figure 3.9 presents the infiltration-rate PDFs $p_{\mathrm{i}}(i^{\star};t)$ corresponding to K_s and α_{vG} that are anti-correlated ($\rho = -0.99$), uncorrelated ($\rho = 0.0$) and perfectly correlated ($\rho = 0.99$). The comparison of the three curves reveals that the perfect correlation between K_s and α_{vG} significantly reduces the predictive uncertainty.

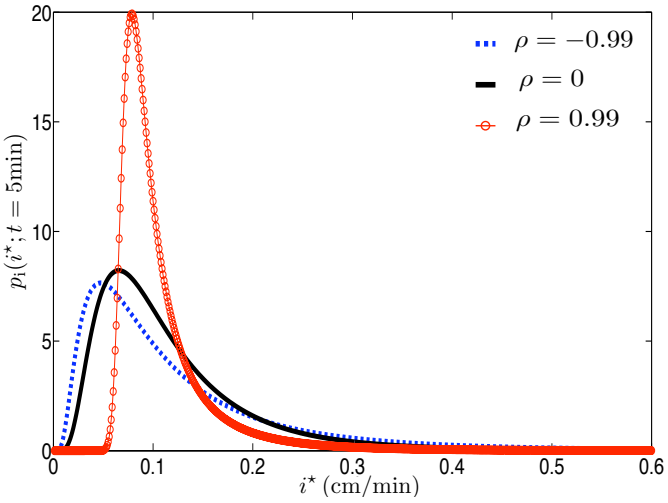

Figure 3.9: The infiltration-rate PDF $p_{\mathrm{i}}(i^{\star};t = 5$ min) from Parlange model for three degrees of correlation ρ between hydraulic parameters K_s and α_{vG}.

3.5.3 Comparison with Richards' equation

To validate our reduced complexity models, we compare their PDF solutions with that obtained by Monte Carlo Simulations (MCS) of the two-dimensional stochastic Richards equation (3.1). In these MCS, we used the geostatistical software library SGEMS to generate $N = 2000$ realizations of mutually-uncorrelated random fields $K_s(\mathbf{x})$ and $\alpha_{vG}(\mathbf{x})$ in a two-dimensional (15000cm$\times$200cm) domain discretized into 2500 nodes. For both parameters, we used an anisotropic exponential correlation function, with horizontal and vertical correlation lengths λ_h and λ_v, respectively. For each realization of $K_s(\mathbf{x})$ and $\alpha_{vG}(\mathbf{x})$, the Richards equation was solved with the USGS code VS2DT, and the infiltration rate $i(t)$ was determined at a surface midpoint. The results were used to compute the infiltration-rate PDF as $p_{\mathrm{i}}(i^\star;t) = (N\Delta_{bin})^{-1}\sum_{n=1}^{N} I(i_n \in \Delta_{bin}^{i^\star};t)$, where Δ_{bin} is a (uniform) bin size, $\Delta_{bin}^{i^\star}$ is the bin containing $i^\star$, and I is the indicator function.

The Dagan-Bresler statistical parameterization, which forms the foundation of our reduced complexity models, requires that $\lambda_v \ll \lambda_h$. This requirement was tested by setting $\lambda_v/\lambda_h = 18.75$ and 30.0 for the random fields $\ln K_s$ and $\ln \alpha_{vG}$, respectively. Figure 3.10 compares the infiltration-rate PDF computed via MCS of the Richards equation with those determined analytically from both the Green-Ampt and Haverkamp solutions. These analytical solutions used the same random values of K_s and α_{vG} at the surface midpoint as those used in the MCS. The PDFs computed with the two reduced complexity models are similar, with the Haverkamp

solution having a slight edge. Both agree with the PDF resulting from the Richards equation at early times (up to 40 min), but this agreement deteriorates with time. This is to be expected, since the conditions of validity of our reduced complexity models are violated as time becomes large enough for the wetting front to travel distances larger than the vertical correlation lengths of K_s and α_{vG}.

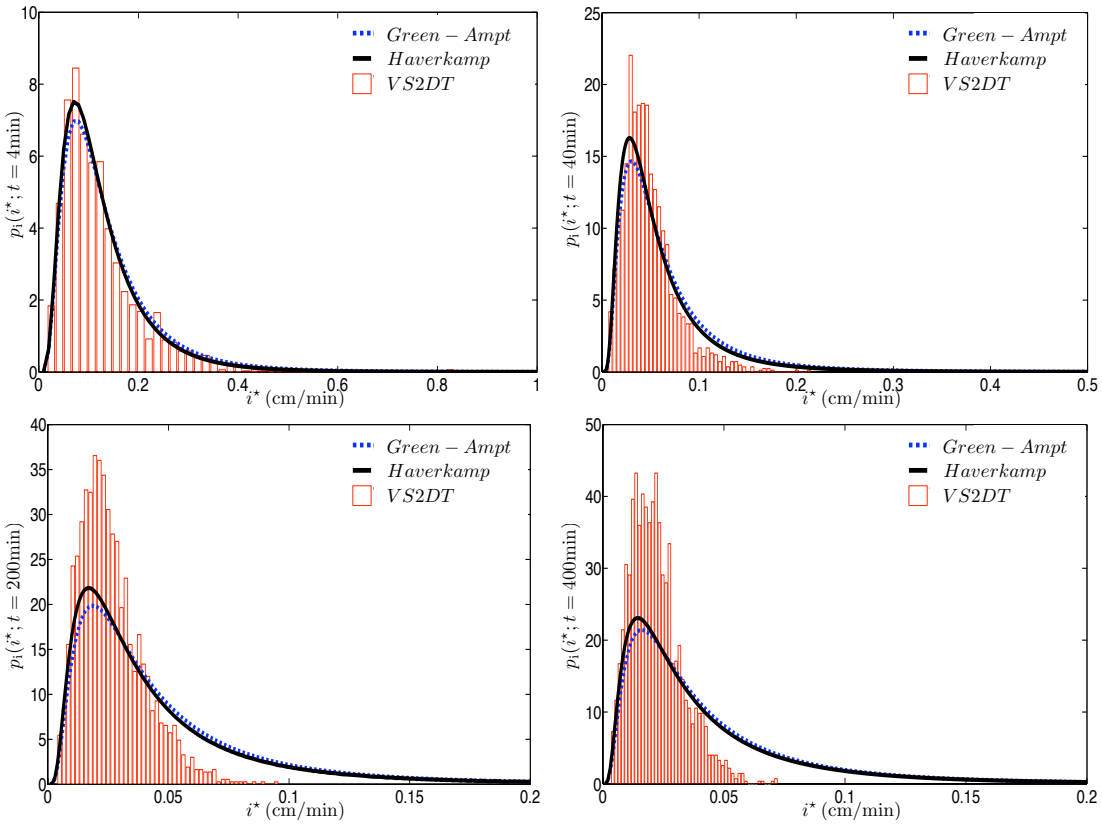

Figure 3.10: Temporal snapshots of the infiltration-rate PDFs computed with the two reduced complexity models (Green-Ampt and Haverkamp) and Monte Carlo simulations of the Richards equation (VS2DT). Ratios of the horizontal and vertical correlation lengths are $\lambda_v/\lambda_h = 18.75$ and 30.0 for $\ln K_s$ and $\ln \alpha_{vG}$, respectively.

Figure 3.11 provides a similar comparison for smaller ratios of λ_v/λ_h. The ratios $\lambda_v/\lambda_h = 4.0$ and 2.5 for $\ln K_s$ and $\ln \alpha_{vG}$ correspond to those observed in

the Bet-Dagan soil [2]. The reduced complexity models perform well at early times ($t = 4$ min) but their accuracy deteriorates faster (by $t = 40$ min), reflecting the increased importance of the lateral flow. At all times and for arbitrary correlation-length ratios, the reduced complexity models provide conservative estimates of predictive uncertainty.

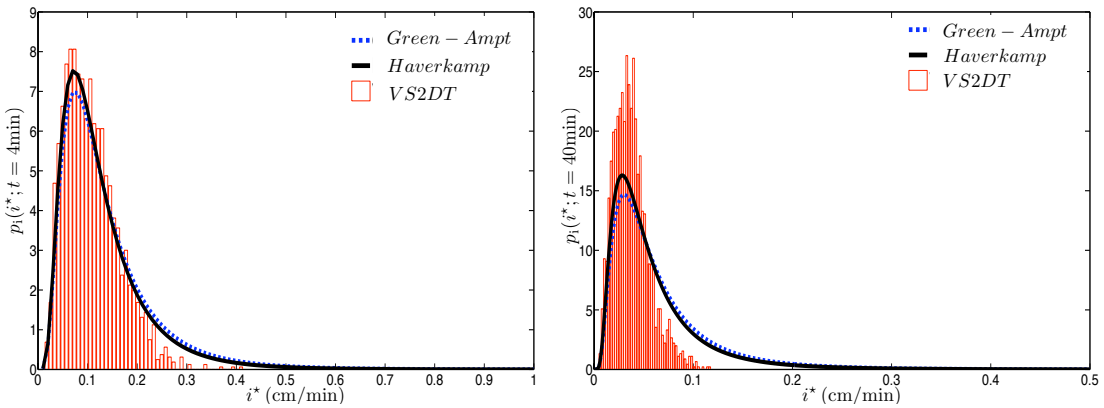

Figure 3.11: Temporal snapshots of the infiltration-rate PDFs computed with the two reduced complexity models (Green-Ampt and Haverkamp) and Monte Carlo simulations of the Richards equation (VS2DT). Ratios of the horizontal and vertical correlation lengths are $\lambda_v/\lambda_h = 4.0$ and 2.5 for $\ln K_s$ and $\ln \alpha_{vG}$, respectively.

3.6 Conclusions

We presented two reduced complexity models for the probabilistic forecasting of infiltration rates in heterogeneous soils during surface runoff and/or flooding events. The models are based alternatively on the Green-Ampt or Parlange models of infiltration under ponded conditions, both employing the Dagan-Bresler statistical parameterization. These models yield closed-form semi-analytical expressions

for the infiltration-rate PDFs (probability density functions), which quantify predictive uncertainty stemming from uncertainty in a soil's hydraulic parameters. Our analysis leads to the following major conclusions.

1. The infiltration-depth and infiltration-rate PDFs developed in this analysis allow one to evaluate probabilities of rare events, i.e., to estimate the probability of the infiltration rate exceeding a given value.

2. Predictive uncertainty (the infiltration-depth and infiltration-rate PDFs) is significantly more sensitive to the coefficient of variation of saturated hydraulic conductivity K_s than to that of the fitting parameters in the Gardner or van Genuchten hydraulic function.

3. The degree of cross-correlation between hydraulic parameters K_s and α has great influence on predictive uncertainty through joint PDF $p_{Y_1,Y_2}(K_s, \alpha)$ (3.6).

4. For Green-Ampt model, the choice of a functional form of the hydraulic function (e.g., the Gardner model vs. the van Genuchten model) has a significant effect on predictive uncertainty during early stages of infiltration. This effect diminishes with time.

5. The PDFs of infiltration rate computed with the two reduced complexity models are similar, with the Parlange model having a slight edge.

6. At early times the PDFs obtained from both models agree with their counterpart resulting from the Richards equation, but this agreement deteriorates

with time. The larger the ratio of vertical to horizontal correlation lengths of soil properties, the longer the reduced complexity models remain valid.

7. At all times and for arbitrary correlation-length ratios, the reduced complexity models provide conservative estimates of predictive uncertainty.

8. Nonlinear dependence of the infiltration rate on soil hydraulic parameters implies that the infiltration-rate PDF is in general not lognormal even if PDFs of the soil parameters are. Hence the nonlinear PDF mapping (3.21) should be used.

Reliance on the reduced complexity models of infiltration into heterogeneous soils with uncertain hydraulic parameters offers a number of advantages. Not only it allows one to compute single-point PDFs of the infiltration rate, it does so exactly, without introducing linearization errors that plague most stochastic analyses of the Richards equation. The reduced complexity models are capable of quantifying uncertainty in any number of hydraulic parameters and can be used with arbitrary constitutive laws (relative conductivity functions and retention curves). Finally, they make it possible to compute multi-point PDFs of infiltration rate. The latter can be used both to predict (cross-)correlations of infiltration rates at multiple locations and to assimilate infiltration data via a straightforward Bayesian updating.

The infiltration-rate PDFs presented here correspond to ponded conditions with constant water heights ψ_0. Other infiltration regimes can be handled in a

similar manner by replacing (3.9) and (3.12) with their appropriate counterparts. For example, (3.12) can be replaced with the analytical solutions in [70] or [71] if infiltration is driven respectively by atmospheric pressure at the soil surface ($\psi_0 = 0$) or by temporally varying ponded water height $\psi_0(t)$. Likewise, infiltration under non-ponded conditions can be handled by replacing (3.9) with appropriately modified Green-Ampt solutions, many of which can be found in [53]. What is important is that a properly chosen reduced complexity model provides a mapping $i = i(K_s, \alpha)$.

In the following chapter, we develop an alternative method to obtain the statistical distribution of volumetric channel flow rate. To be specific, an equation of fine-grained CDFs is derived from the stochastic kinematic wave equation that governs overland flow.

Wang, P., Tartakovsky, D. M., (2011), 'Probabilistic predictions of infiltration into heterogeneous media with uncertain hydraulic parameters'. Int. J. Uncert. Quant., vol. 1, no. 1, pp. 35-47.

Wang, P., Tartakovsky, D. M., (2011), 'Reduced complexity models for probabilistic forecasting of infiltration rates'. Adv. Water Resour., vol. 34, pp. 375-382, doi:10.1016/j.advwatres.2010.12.007.

Chapter 4

Uncertainty Quantification in Kinematic Wave Models

4.1 Introduction

In previous chapter, two reduced-complexity models for infiltration rate have been developed. Now a different UQ approach for overland flow based on kinematic wave theory (KWT) is proposed.

Since its development by Lighthill and Whitham [76, 77], KWT has been used to model a number of environmental phenomena, including overland flow, channel flow, mutliphase flow in porous media, erosion and sediment transport [78, 79]. It is routinely employed in analyses of urban storm-water drainage systems to route flood hydrographs [79].

The KWT theory postulates a functional relationship between a quantity

$k(\mathbf{x}, t)$ and its flux $\mathbf{q}(\mathbf{x}, t)$, $\mathbf{q} = \mathbf{q}(k)$, so that the phenomenon is described by the continuity equation

$$\frac{\partial k}{\partial t} + \nabla \cdot \mathbf{q} = S, \qquad \mathbf{q} = \mathbf{q}(k) \tag{4.1}$$

where $S(\mathbf{x}, t)$ is a source. This is in contrast with dynamic-wave models, which employ the conservation of momentum to establish a dynamic relation between $k(\mathbf{x}, t)$ and $\mathbf{q}(\mathbf{x}, t)$. For Froude numbers smaller than 1 (appropriate for flood waves), the dynamic waves (long gravity waves) do appear, but they attenuate rapidly and the main disturbance is carried downstream by kinematic waves only [76]. We use this application (overland flow in flood forecasting) to motivate the subsequent analysis.

When the KWT equation (4.1) is used to describe flow in long rivers, the functional relationship $\mathbf{q} = \mathbf{q}(k)$ is typically given by either Chézy or Manning formulae [79], which represent a balance between the friction at the bottom and the gravitational force. These constitutive relations are parameterized with a friction coefficient and a downward slope, both of which often exhibit high spatial variability and are usually underspecified by data. In addition to this parametric uncertainty, the source function S, which represents influx from tributaries and/or runoff from the ambient terrain, as well as initial and boundary conditions are subject to uncertainty. Although data acquisition continues to improve, ubiquitous data sparsity and measurement/interpretation errors render overland flow predictions inherently uncertain. This predictive uncertainty is routinely mentioned as

one of the fundamental challenges in flood forecasting [80].

A common approach to quantifying uncertainty in system parameters and driving forces is to treat them as random fields, whose statistics are inferred from available data. This renders the KWT equation (4.1) stochastic. Its solution is given in terms of probabilistic density functions (PDFs) of the system states $\mathbf{q}$ and k, and amounts to propagation of parametric uncertainty through the modeling process.

Early attempts to quantify uncertainty in modeling predictions based on the stochastic KWT equation (4.1) dealt with spatially-averaged quantities [81, 82, 83, 84, 85]. Spatially-distributed probabilistic predictions were obtained by solving the stochastic KWT equation (4.1) with Monte Carlo simulations (MCS) [86, 60] and stochastic finite elements [87, 88]. For transient nonlinear systems such as (4.1) these direct approaches are computationally expensive, and often prohibitively so, especially when the parameter fields have small correlation lengths and high variances. They are typically used to compute the first two ensemble moments of system states. Accurate estimates of the tails of system states' PDFs entail further computational costs.

We present an alternative approach to uncertainty quantification in flow models based on the stochastic KWT equation (4.1). The approach is based on the derivation of a deterministic differential equation for cumulative density functions (CDFs) of the system states $\mathbf{q}(\mathbf{x}, t)$ and $k(\mathbf{x}, t)$. Our framework is conceptually similar to the PDF equations approach used to describe the dynamics of (passive

or reactive) scalars in turbulent flows [89] and to quantify uncertainty in models of reactive transport in heterogeneous porous media [90]. Yet it offers a distinct advantage of removing the ambiguity in formulation of boundary conditions.

In section §4.2, we provide a shallow-water formulation of surface flow and identify the key sources of uncertainty. Section §4.3 contains the derivation of a CDF equation and corresponding boundary conditions. In section §4.4, this equation is solved analytically for two special cases describing flood dynamics in long rivers. We investigate the robustness and salient features of the CDF solutions in section §4.5, using MCS as a benchmark. The overall conclusions are drawn in section §4.6.

4.2 Problem Formulation

4.2.1 Governing equations

Motion of a homogeneous fluid whose horizontal extent is much larger than its vertical counterpart can be described by the shallow water equations. It is common to use their one-dimensional form, which is often referred to as the Saint-Venant equations,

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = S, \tag{4.2}$$

to model open-channel flow. In this application of the KWT equation (4.1), $k(x,t)$ [L^2] denotes the cross-sectional area of a channel occupied by the fluid at a point x

along the channel's length, $q(x,t)$ [$\mathrm{L}^3\,\mathrm{T}^{-1}$] is the volumetric flow rate, and $S(x,t)$ [$\mathrm{L}^2\,\mathrm{T}^{-1}$] denotes the lateral inflow rate. When kinematic waves in long rivers pass a junction with a tributary, the latter's effects on the flood movement are represented by S. The KWT equation (4.1) provides a good approximation of the flood dynamics if influence on the river upstream of the junction is neglected [76]. Since the kinematic wave approximation neglects backwater effects—the upstream propagation caused by local acceleration, convective acceleration, and pressure—the flow rate throughout the flow domain is non-negative, $q(x,t) \geq 0$ for all x and t.

For wide channels (i.e., channels whose hydraulic radius equals the depth of water), commonly used functional relations between k and q at any point x and time t (e.g., Darcy-Weisbach, Chézy, or Manning formulae) can be written as

$$q = \alpha\, k^{1/\beta}. \tag{4.3}$$

Here the parameter α represents the effects of surface slope and resistance, and the exponent β is a measure of turbulence that characterizes the flow regime as laminar, turbulent or transitional [79]. In general, both parameters can vary in space and time, $\alpha(x,t)$ and $\beta(x,t)$. Although the bed of an alluvial river varies with time [91], these changes occur on a time scale that is much larger than that of the flow, so that $\alpha = \alpha(x)$. While not strictly necessary, we assume that the exponent β is constant in order to simplify the presentation. Combining (4.2)

and (4.3) gives

$$\gamma(x)\frac{\partial q^{\beta}}{\partial t} + \frac{\partial q}{\partial x} = S(x,t), \qquad \gamma \equiv \alpha^{-\beta}. \tag{4.4}$$

The open-channel flow equation (4.4) is subject to the initial and boundary conditions

$$q(x=0,t) = q_0(t) \tag{4.5a}$$

$$q(x,t=0) = q_{\text{in}}(x). \tag{4.5b}$$

We allow the coefficient $\gamma(x)$, the source function $S(x,t)$, the inlet flow rate $q_0(t)$, and the initial flow rate $q_{\text{in}}(x)$ to be uncertain. The uncertainty is quantified by treating these functions as random fields. Within this probabilistic framework, a random quantity $\mathcal{A}(x,t;\omega)$ varies not only in the physical domain, $(x,t) \in (0,\infty)\times(0,\infty)$, but also in the probability space $\omega \in \Omega$. Our goal is to obtain a complete (single space-time point) probabilistic description of $q(x,t;\omega)$. In the following, the dependence of the random fields on ω is suppressed to simplify the notation.

4.2.2 Example of statistical parameterizations

Consider, as an example of the general relation (4.3), the Manning formula

$$q = \frac{\sqrt{s_0}}{n}k^{4/3}, \tag{4.6}$$

wherein $s_0(x)$ denotes the channel slope, and $n(x)$ (s/m$^{1/3}$) is the Manning's roughness coefficient. Both $s_0(x)$ and $n(x)$ are typically uncertain and often treated as

random (e.g., [92, 3, 4, 5] and the references therein). The data reviewed in these and other analyses suggest that no single distribution is capable of capturing their spatial variability at all sites, with the normal, lognormal, gamma, logistic or log-logistic PDFs found to fit various data sets best. The spatial correlations of $s_0(x)$ and $n(x)$, and their cross-correlation, are likewise site-specific. For the data analyzed in [92], the random field $s_0(x)$ was found to be spatially uncorrelated (white noise) and either weakly correlated or uncorrelated with other hydraulic parameters.

The relevant statistics of the parameter $\gamma(x) = (\sqrt{s_0}/n)^{-\beta}$ in (4.4) are related to those of $s_0(x)$ and $n(x)$ in Appendix A.

4.3 CDF equations

We start by introducing a "raw" (or "fine-grained") cumulative density function (CDF),

$$\Pi(Q; x, t) = \mathcal{H}[Q - q(x, t)], \tag{4.7}$$

where $\mathcal{H}$ is the Heaviside step function, and Q is a deterministic value (outcome) that the random flow rate q can take at a space-time point (x, t). Let $p_q(Q; x, t)$ denote a single-point probability density function (PDF) of q at the space-time point (x, t). Then taking the ensemble average (over random q) of (4.7) yields a

single-point CDF of q,

$$\overline{\Pi}(Q;x,t) \equiv \int_0^\infty \mathcal{H}(Q-q')p_{\mathrm{q}}(q';x,t)\mathrm{d}q' = F_q(Q;x,t). \tag{4.8}$$

For $q(x,t)$ in (4.4) and (4.5), its raw CDF satisfies a two-dimensional stochastic linear CWT equation (Appendix B)

$$\beta\gamma(x)Q^{\beta-1}\frac{\partial\Pi}{\partial t} + \frac{\partial\Pi}{\partial x} + S(x,t)\frac{\partial\Pi}{\partial Q} = 0 \tag{4.9}$$

subject to the initial and boundary conditions

$$\Pi(Q;x,t=0) = \Pi_{\mathrm{in}} = \mathcal{H}\left[Q - q_{\mathrm{in}}(x)\right], \tag{4.10a}$$

$$\Pi(Q;x=0,t) = \Pi_0 = \mathcal{H}\left[Q - q_0(t)\right], \tag{4.10b}$$

$$\Pi(0;x,t) = 0. \tag{4.10c}$$

The straightforward and unambiguous way in which the boundary condition (4.10b) is formulated provides the key advantage of our CDF method over commonly used PDF methods [89, 90]. The latter are formulated in terms of "raw" PDFs, $\Pi(Q,q;x,t) = \delta[Q-q(x,t)]$, whose value at $Q=0$ for any space-time point (x,t) is generally unknown.

The CDF formulation (4.9)–(4.10) offers a number of other advantages over direct solutions of the flow equations (4.4)–(4.5). First, one needs to compute (e.g., with MCS or stochastic finite elements) only the first ensemble moment of Π to obtain the full distribution of q. Second, linearity of the CDF equations (4.9)–(4.10) simplifies their theoretical and numerical analyses, enabling, for example, examination of the convergence and other properties of polynomial chaos methods [93].

More important for the subsequent analysis, one can take advantage of the large body of literature on stochastic averaging of linear advective transport in random velocity fields $\mathbf{v}(\mathbf{x}, t)$,

$$\frac{\partial \Pi}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} \Pi = 0. \tag{4.11}$$

In the context of (4.9)–(4.10),

$$\mathbf{x} = (x, Q)^T, \qquad \mathbf{v} = (v_x, v_Q)^T, \qquad v_x = \frac{Q^{1-\beta}}{\beta\gamma(x)}, \qquad v_Q = \frac{Q^{1-\beta} S(x,t)}{\beta\gamma(x)}. \tag{4.12}$$

Specifically, the ensemble averaging of (4.11) would yield an effective transport equation for the CDF of q,

$$\frac{\partial F_q}{\partial t} + \mathbf{v}_{\text{eff}} \cdot \nabla_{\mathbf{x}} F_q = \nabla_{\mathbf{x}} \cdot (\mathbf{D} \nabla_{\mathbf{x}} F_q), \tag{4.13}$$

where $\mathbf{v}_{\text{eff}}$ and $\mathbf{D}$ are the effective velocity and the eddy-diffusivity tensor, respectively. This equation is based on a closure approximation, but is asymptotically exact when F_q varies slowly with $\mathbf{x}$ and t relative to $\mathbf{v}$ [94].

In the present study, we consider two special cases of (4.9), $S = 0$ and $S = S(x)$, both of which enable one to obtain the CDFs F_q without resorting to closure approximations.

4.4 CDF Solutions

4.4.1 Flood propagation in the absence of lateral inflow

The open-channel flow equation (4.4) with $S \equiv 0$ provides a classical setting first analyzed by Lighthill and Whitham [76] to model flood propagation in long

rivers. The corresponding raw CDF problem (4.9)–(4.10) admits an analytical solution (Appendix C),

$$\Pi(Q;x,t) = \mathcal{H}(C-t)\mathcal{H}[Q - q_{\text{in}}(x^\star)] + \mathcal{H}(t-C)\mathcal{H}[Q - q_0(t-C)]. \tag{4.14a}$$

Here

$$C(x) = \int_0^x \beta Q^{\beta-1}\gamma(x')\mathrm{d}x'. \tag{4.14b}$$

and $x' = x^\star$ is a solution of the equation

$$C(x') = C(x) - t \tag{4.14c}$$

for a given Q, x and t.

For large times, $t > C$, the general solution (4.14) reduces to

$$\Pi(Q;x,t) = \mathcal{H}[Q - q_0(t-C)]. \tag{4.15}$$

4.4.2 Flood propagation under steady lateral inflow

In the open-channel flow equation (4.4), the source term $S = S(x)$ might represent either input from a river's tributaries (in which case S can be treated as a sum of delta functions) or runoff (in which case S is continuous) or their combination. The corresponding raw CDF problem (4.9)–(4.10) admits an analytical solution (Appendix D)

$$\Pi(Q;x,t) = \mathcal{H}(C-t)\mathcal{H}[Q - I(x,x^\star) - q_{\text{in}}(x^\star)] + \mathcal{H}(t-C)\mathcal{H}[Q - I(x,0) - q_0(t-C)]. \tag{4.16a}$$

Here

$$C = \int_{x_0}^{x} \beta[Q - I(x, x'')]^{\beta-1}\gamma \mathrm{d}x'', \qquad I(x, x') = \int_{x'}^{x} S\,\mathrm{d}x'', \tag{4.16b}$$

and $x' = x^\star$ is a solution of the equation

$$\int_{x_0}^{x'} \beta[Q - I(x, x'')]^{\beta-1}\gamma \mathrm{d}x'' = \int_{x_0}^{x} \beta[Q - I(x, x'')]^{\beta-1}\gamma \mathrm{d}x'' - t \tag{4.16c}$$

with

$$x_0 = \begin{cases} 0 & Q \geq I(x, 0) \\ \\ \eta & Q < I(x, 0) \end{cases}, \qquad I(\eta, 0) = I(x, 0) - Q. \tag{4.16d}$$

If $S(x) \equiv 0$, (4.16) reduces to (4.14).

4.4.3 CDF solutions

Expressions (4.14) and (4.16) map the random system parameter $\gamma(x)$ and driving forces $q_0(t)$, $S(x)$, and $q_{\text{in}}(x)$ onto the raw CDF Π. To simplify the presentation, we take q_{in} to be deterministic, and analyze in detail flow in the absence tributaries ($S = 0$). This setting captures the salient features of the CDF method, and its extension to more complicated flow scenarios is relatively straightforward.

The parametric uncertainty can now be quantified by p_{γ,q_0}, a joint PDF of random inputs $\gamma(x)$ and $q_0(t)$. Since $\gamma(x)$ and $q_0(t)$ represent two different physical phenomena, they can be treated as independent, so that $p_{\gamma,q_0} = p_\gamma p_{q_0}$ and (4.8) gives rise to a CDF solution

$$F_q(Q; x, t) = \int\int \Pi(\Gamma, Q_0; x, t) p_\gamma(\Gamma) p_{q_0}(Q_0) \mathrm{d}\Gamma \mathrm{d}Q_0. \tag{4.17}$$

The (non-Gaussian, correlated) random field $\gamma(x)$ enters (4.14) only as an integrand in

$$I_\gamma(x) = \int_0^x \gamma(x')\mathrm{d}x'. \tag{4.18}$$

Therefore, (4.17) can be replaced with

$$F_q(Q;x,t) = \int\int \Pi(I,Q_0;x,t)p_{I_\gamma}(I)p_{q_0}(Q_0)\mathrm{d}I\mathrm{d}Q_0. \tag{4.19}$$

It remains to compute $p_{I_\gamma}(I;x)$, the PDF of $I_\gamma(x)$.

Let λ_γ denote the correlation length of $\gamma(x)$. For $x \ll \lambda_\gamma$, $\gamma(x')$ on the interval $[0,x]$ is approximately constant, $I_\gamma(x) \approx x\gamma$ and $C(x)$ in (4.14) can be approximated by

$$C(x) \approx \beta Q^{\beta-1}x\gamma(x), \qquad x \ll \lambda_\gamma. \tag{4.20}$$

For $x \gg \lambda_\gamma$, $I_\gamma(x)$ becomes Gaussian with mean $x\overline{\gamma}$ and variance $2x\sigma_\gamma^2$, and $C(x)$ becomes

$$C(x) \approx \beta Q^{\beta-1}N(x\overline{\gamma}, 2x\sigma_\gamma^2), \qquad x \gg \lambda_\gamma. \tag{4.21}$$

If $\gamma(x)$ lacks spatial correlation [92], this expression becomes exact. For intermediate x, we approximate the statistics of $I_\gamma(x)$ with the central limit theorem (CLT)-based approach [95] (see Appendix E).

Below we use a computational example to investigate the accuracy and robustness of the alternative approximations of $p_{I_\gamma}(I;x)$, and their effects on the flow-rate CDF F_q, via comparison with Monte Carlo simulations.

4.4.4 Computational example

We set the initial flow rate to $q_{\text{in}} = 0.5\,\text{m}^3/\text{s}$ and the flow rate at the inlet $x = 0$ to

$$q_0(t) = \overline{q}_0 \left| \sin\left(\frac{\pi t}{P}\right) \right| [1 + q_0'(t)]. \tag{4.22}$$

The mean flow rate $\overline{q}_0 = 1\,\text{m}^3/\text{s}$ satisfies the subcritical flow condition required for the kinematic wave approximation to be valid, and P denotes the period. The fluctuating term $q_0'(t)$ is white noise. Its statistics, as well as those of the random channel slope $s_0(x)$ and Manning coefficient $n(x)$ are summarized in Table 4.1, wherein CV denotes the coefficient of variation (absolute value of the ratio of the standard deviation to the mean), and $\rho(r)$ and λ are the correlation function and correlation length, respectively. The size of the flow domain (e.g., the length of a river downstream from $x = 0$) is $L = 20\,\text{km}$, while the correlation length is $\lambda = 200\,\text{m}$.

Table 4.1: Statistics of the uncertain (random) parameters. These values are representative of data in [3, 4, 5, 6].

Parameter	PDF	$\rho(r)$	Mean	CV	λ
$q_0'(t)$	normal	$\delta(r)$	0	0.1	–
$s_0(x)$	lognormal	$\exp(r/\lambda)$	0.01	0.25	200
$n(x)$	lognormal	$\exp(r/\lambda)$	0.037	0.25	200

4.5 Results and discussion

The subsequent results are presented in terms of the dimensionless quantities defined as follows. Let the deterministic quantities λ and $\overline{q}_0$ represent a

characteristic length scale and a characteristic volumetric flux, respectively. Their ratio define a characteristic time scale τ,

$$\tau = \lambda^3/\overline{q}_0. \tag{4.23}$$

We introduce dimensionless quantities

$$\tilde{x} = \frac{x}{\lambda}, \quad \tilde{t} = \frac{t}{\tau}, \quad \tilde{Q} = \frac{Q}{\overline{q}_0}, \quad \tilde{q}_{\text{in}} = \frac{q_{\text{in}}}{\overline{q}_0}, \quad \tilde{\gamma} = \frac{\gamma \overline{q}_0^{3/4}}{\lambda^2}, \quad \tilde{P} = \frac{P}{\tau}. \tag{4.24}$$

In the simulations reported below, we set $\tilde{P} = 1$.

MCS consist of 2000 realizations. The geostatistical software SGEMS was used to generate mutually-uncorrelated random fields of $s_0(x)$ and $n(x)$ on 500 nodes evenly distributed on the interval $[0, L]$. These were used in (4.14) to compute realizations of the raw CDF, Π, and its mean, the flow rate CDF F_q.

Close to the inlet, $\tilde{x} \ll 1$, the approximation of the integral $C(x_d)$ with (4.20) is expected to lead to an accurate solution for the CDF $F_q(\tilde{Q}; \tilde{x}, \tilde{t})$. Figure 4.1 shows three temporal snapshots of $F_q(\tilde{Q}; \tilde{x}, \tilde{t})$ computed at $\tilde{x} = 0.4$ with MCS, the constant γ approximation (4.20) (Const), and the CLT-based approximation (CLT). At early times ($\tilde{t} = 1 \times 10^{-6}$), F_q maintains its deterministic initial state (4.10a) of a step function. As time increases, both parametric uncertainty and uncertainty in boundary conditions propagate downstream and lead to a rising predictive uncertainty at $t = 1 \times 10^{-5}$. At later times, predictive uncertainty at a given point is increasingly dominated by the boundary fluctuation while the effects of parametric (s_0 and n) uncertainty become negligible. The late-time solution (4.15) and the snapshot at $t = 0.005$ illustrate this behavior.

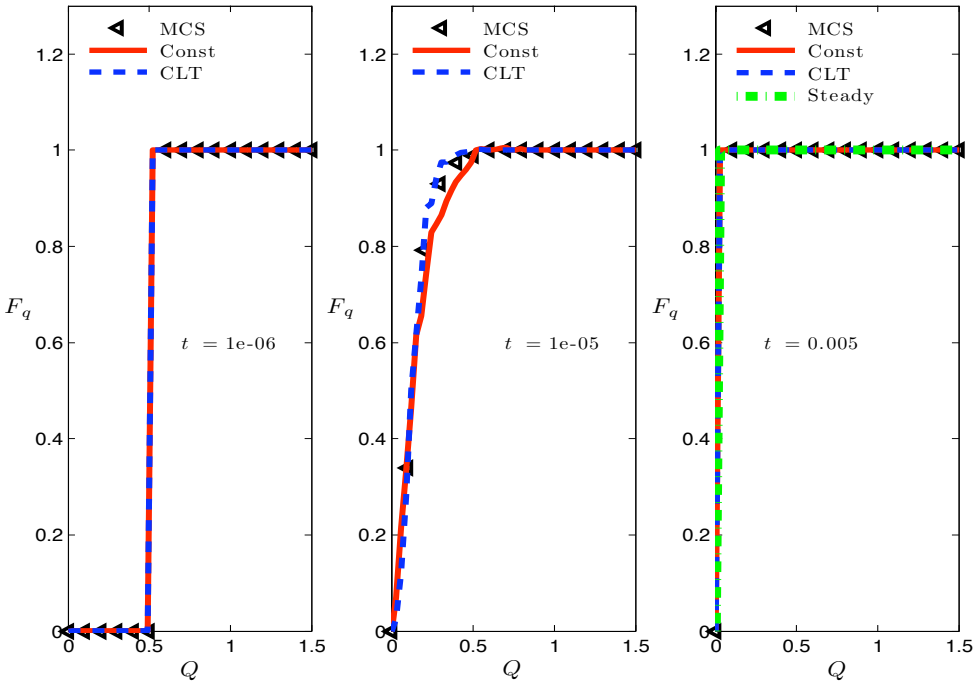

Figure 4.1: Temporal evolution of the flow rate CDF, $F_q(\tilde{Q}; \tilde{x} = 0.4, \tilde{t})$, computed with MCS, the constant γ approximation (4.20) (Const), and the CLT-based approximation (CLT).

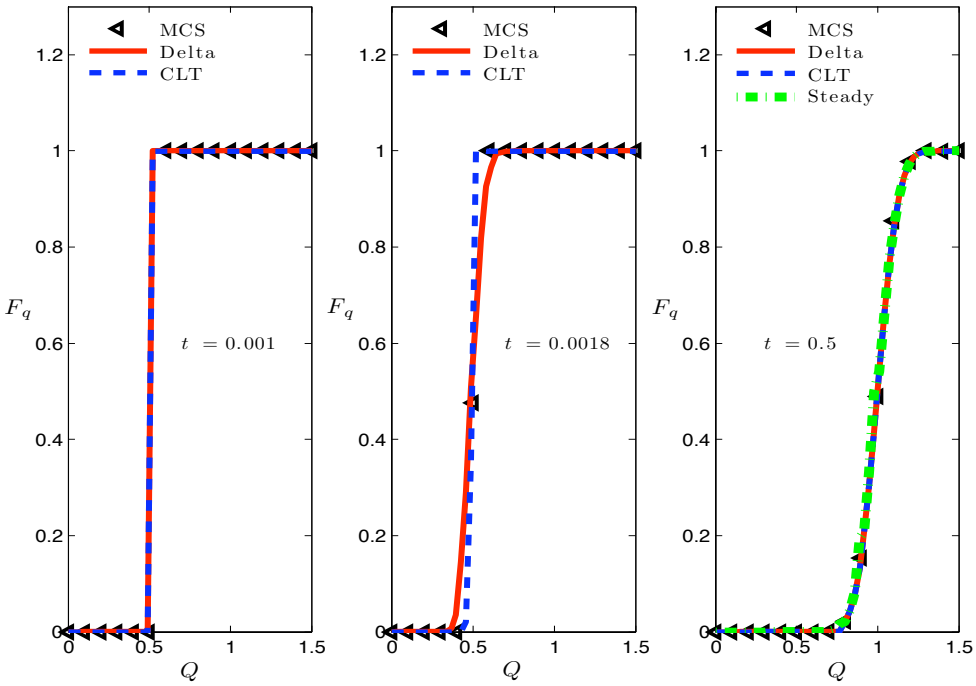

Figure 4.2: Temporal evolution of the flow rate CDF, $F_q(\tilde{Q}; \tilde{x} = 100, \tilde{t})$, computed with MCS, the white noise γ approximation (4.21) (Delta), and the CLT-based approximation (CLT). Also shown for $\tilde{t} = 0.5$ is the large-time solution (4.15).

Far away from the inlet, $\tilde{x} \gg 1$, the approximation (4.21) is expected to be accurate. Figure 4.2 demonstrates that this indeed is the case for $F_q(\tilde{Q}; \tilde{x} = 100, \tilde{t})$, especially at early and large times. The CLT-based approximation is in perfect agreement with MCS at all times. The initial state predominates the early time solution ($t = 1 \times 10^{-3}$). At intermediate times ($\tilde{t} = 1.8 \times 10^{-3}$), fluctuations from both the hydraulic parameters and the boundary condition affect the flow rate's predictions. With increasing time, the latter uncertainty exerts ever stronger influence, and $F_q(\tilde{Q}; \tilde{x}, \tilde{t})$ is dominated by the white-noise $q_0'(t)$. There is good match with the large-time solution (4.15) at $t = 0.5$.

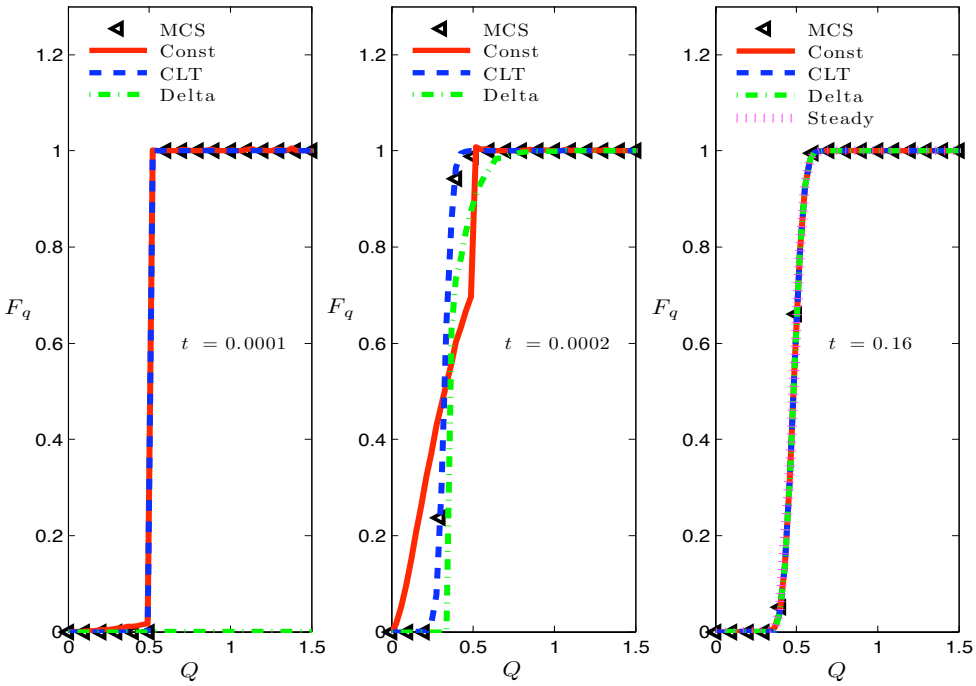

Figure 4.3: Temporal evolution of the flow rate CDF, $F_q(\tilde{Q}; \tilde{x} = 10, \tilde{t})$, computed with MCS, the constant γ approximation (4.20) (Const), the CLT-based approximation (CLT), and the white noise γ approximation (4.21) (Delta). Also shown for $\tilde{t} = 0.16$ is the large-time solution (4.15).

At intermediate distances ($\tilde{x} = 10$), both the constant γ and white noise γ approximations are expected to fail. Figure 4.3 reveals that the failure occurs at intermediate times ($\tilde{t} = 2 \times 10^{-4}$) when predictive uncertainty is highest. At all times, the CLT-based approximation is in excellent agreement with MCS. The large-time solution is appropriate at $t = 0.16$.

4.6 Conclusion

We developed a probabilistic approach to quantify parametric uncertainty in first-order hyperbolic conservation laws (kinematic wave equations). The approach relies on the derivation of a deterministic equation for the cumulative density function (CDF) of the system state, in which probabilistic descriptions (probability density functions or PDFs) of the system parameters and/or initial and boundary conditions serve as inputs. The accuracy and robustness of solutions of the CDF equation for one such system, the Saint-Venant equations of river flows, were investigated via comparison with Monte Carlo simulations. Our analysis leads to the following major conclusions.

1. CDF equations, and their (semi-)analytical solutions, provide a computationally efficient alternative to the existing methods for uncertainty quantification, such as Monte Carlo simulations and stochastic finite element methods (polynomial chaos expansions, stochastic collocation methods, etc.).

2. CDF equations are ideally suited for handling input parameters and/or initial and boundary conditions that exhibit small correlation lengths. This is in contrast with stochastic finite element methods and other numerical approaches that rely on the Karhunen-Loève representation of random parameter fields.

3. CDF equations offer an operational advantage over PDF equations that are often used in other contexts, e.g., to analyze transport of passive tracers and reactive species in turbulent (randomly fluctuating) velocity fields. This is

because CDF equations allow for straightforward and unambiguous determination of boundary conditions with respect to sample variables.

In the next chapter, we will propose a general framework to quantify the uncertainty of multiple algae-groups' densities through a Langevin equation.

Wang, P., Tartakovsky, D. M., (2011), 'Uncertainty quantification in kinematic wave models'. Submitted.

Chapter 5

Stochastic Forecasting of Algae Bloom in Lakes

5.1 Introduction

In the previous two chapters, two UQ methods to obtain full statistical description of random surface/subsurface flows have been developed. The reduced-complexity model and CDF method not only enable one to assess the probability of rare events but also aid stochastic analysis in ecological systems. This chapter presents a stochastic model that governs dynamic growth of several phytoplankton groups in an enclosed aquatic system, whose nutrients are primarily brought by random runoff.

Cyanobacteria, also known as blue-green algae, is one of the most important phytonplankton groups. It has attracted a lot of attentions over the years from

both negative and positive prospectives. Widespread in a diverse range of water bodies, cynobacteria consumes nutrients (nitrogen and phosphorous) as part of its growing process. Increasing human activity, especially discharge of wastewater, has significantly accelerated eutrophication of many aquatic systems worldwide [96]. As a result, there is an explosion of harmful algae blooms (HABs) that pose serious risks for human and animal health and ecosystem sustainability. A conservative estimate of the economic cost of HABs and eutrophication in U.S. alone amounts to $2.2-4.6 billion annually [97]. Ironically, on the other end of its effect spectrum, recent research has suggested various applications of algal biomass, such as bio-diesel, animal feed, heating, electricity, and even pharmaceutical and cosmetic products.

Comparing to many engineering systems, it is more challenging to study dynamics in ecosystems typically governed by numerous parameters and complex interactions between different bio-groups. Its analysis is further complicated by the uncertainty exhibited in many large-scale applications. In recent years, a growing number of works [7, 98, 99] dealt with stochastic analyses of HABs systems. Yet their UQ is either relatively simple to account for more general cases or are not physically based and fail to quantify the probability tails required for PRA.

In this chapter, an alternative framework is developed to address both underlying dynamics and uncertainty of algae growth in a heterogeneous ecosystem. The problem is formulated in section 5.2 with a brief review of associated parameters and their uncertainty. Section 5.3 contains a derivation of a set of deterministic

equations, i.e. Fokker-Planck equations, whose solution are the PDFs of density of various algae groups. Section 5.4 presents an example for the cyanobacteria-concentration PDFs. Final conclusions are drawned in section 5.5.

5.2 Problem Formulation

HABs occur most often when there are abundant nutrients (nitrogen and phosphorous), warm water ($> 20^oC$), sunlight and stagnant or quiescent water [100]. It is often assumed that an aquatic system is well mixed or at least there is no spatial dependence for the top layer of water. As a result, the growing process is usually modeled by an ordinary differential equation.

The present study adopts a model [7] that focuses on four algae groups: Diatoms, Chrysophycea, nitrogen-fixing Cyanobacteria, and minor species. Without loss of generality, we extend this model to a number of N algae groups, whose growth rate of biomass concentration C_{A_i} ([mg m^{-3}]) in a lake is described by

$$\frac{dC_{A_i}}{dt} = \left(\tilde{\mu}_i - \frac{\tilde{\sigma}_i}{h} - \frac{Q_{out}}{V} - p_i C_Z\right) C_{A_i}, \qquad i = 1, 2, ..., N \tag{5.1}$$

The underlying dynamic of (5.1) requires a set of parameters relating growth to gain (natural growth rate $\tilde{\mu}_i$) and loss (non-predatory loss rate $\tilde{\sigma}_i$, outflow rate Q_{out} and zooplankton predator rate $p_i C_Z$). V (m^3) and h (m) represent the lakes' volume and mean depth, respectively.

Natural growth rate $\tilde{\mu}_i$ and non-predatory loss rate $\tilde{\sigma}_i$ are defined as

$$\tilde{\mu}_i = \mu_i \theta_i^{T-T_{ref}} \frac{I}{K_{I_i} + I} \frac{P}{K_{P_i} + P} \frac{N}{K_{N_i} + N} \tag{5.2}$$

$$\tilde{\sigma}_i = \sigma_i \theta_\sigma^{T-T_{ref}} \tag{5.3}$$

where the rate coefficients K's with various subscripts are defined in Table 5.1. θ_i and θ_σ denote temperature coefficients for growth and non-predatory loss rate, T is the average temperature in the lake, and P and N are the average concentrations of phosphorus and nitrogen in the aquatic system. Inside phytoplankton the latter concentrations are given by

$$P = P_{tot} - \sum_{i=1}^{N} \alpha_i C_{A_i}, \qquad N = N_{tot} - \sum_{i=1}^{N} \beta_i C_{A_i}, \tag{5.4}$$

where P_{tot} and N_{tot} are the total average concentrations of phosphorus and nitrogen in the lake, respectively. Nutrient content of each algae group for phosphorus and nitrogen is denoted by the deterministic constants α_i and β_i, respectively.

The Monod-form of algae growth rate (5.2) depends almost linearly on irradiance I, phosphorous P and nitrate N when these quantities are small. It is bounded by a maximum value of 1 for large intake of light and nutrients.

It is common [7, 98, 101] to identify parameters in the governing equation (5.1) through either measurements of Markov chain Monte Carlo (MCMC). The data reported in [7, 98, 101] suggest that temperature T, global irradiance I, outflow rate Q_{out} and predatory loss $p_i C_Z$ typically exhibit much smaller variation than nutrients concentrations over the summer. Consequently, we treat the concentrations of phosphorous P_{tot} and nitrogen N_{tot} as two random time-dependent

parameters, while assuming that the remaining parameters are deterministic. Our goal here is to derive a mapping scheme that translates these parametric uncertainty into the statistics of system state $W(\{\boldsymbol{\xi}\}, t)$, where W is the joint PDF of various phytoplankton-specices concentrations and $\{\boldsymbol{\xi}\} = \xi_1, \xi_2, ..., \xi_N$ denote the deterministic values of population concentrations in probability place (i.e., outcomes).

Table 5.1: Notations and units for the model parameters, data variables and constants from Table 3 in [7]

μ_i	(day^{-1})	maximum growth rate at 20^oC
σ_i	(day^{-1})	maximum non-predatory loss rate at 20^oC
θ_i	(–)	temperature coefficients for growth rate
θ_σ	(–)	temperature coefficients for non-predatory loss rate
K_{I_i}	(W m^{-2})	global irradiance half-saturation coefficient
K_{P_i}	(mg m^{-3})	phosphorus half-saturation coefficient
K_{N_i}	(mg m^{-3})	nitrogen half-saturation coefficient
$p_i C_Z$	(day $^{-1}$)	zooplankton rate
α_i	(–)	relative phosphorus content of algae
β_i	(–)	relative nitrogen content of algae
P	(mg m^{-3})	total phosphorus concentration available for the algae
P_{tot}	(mg m^{-3})	total phosphorus concentration
N	(mg m^{-3})	total nitrogen concentration available for the algae
N_{tot}	(mg m^{-3})	total nitrogen concentration
T, T_{ref}	(^oC)	temperature, the reference temperature (20^oC)
Q	(m^3 day^{-1})	outflow
I	(W m^{-2})	global irradiance
V	(m^3)	volume of lake
h	(m)	depth of lake

5.3 Stochastic Models

The supply of nutrients by surface runoff and discharge of waste water is the primary factor leading to eutrophication in many water bodies. Hence we identify fluctuation of runoff rate $q(t)$ as the common cause of the uncertainty of P_{tot} and N_{tot}, which can be modeled as

$$P_{tot} = P_0 + \frac{c_P Q}{V}, \qquad N_{tot} = N_0 + \frac{c_N Q}{V}, \tag{5.5}$$

where P_0 and N_0 (mg m^{-3}) are the initial concentration in the lake, c_P and c_N ([mg m^{-3}]) represent the concentration of phosphorus and nitrogen in the runoff, respectively. Statistics of random cumulative runoff $Q(t) = \int_0^t q(t')\mathrm{d}t'$ can be derived from our CDF analysis of overland flow (Chapter 4).

Taking a Taylor expansion around the mean runoff volume $\bar{Q}$ (m^3) and denoting $Q' = Q - \bar{Q}$, we express random growth rates as

$$\tilde{\mu}_i \;=\; \tilde{\mu}_i(\bar{Q}) \;+\; \frac{\partial \tilde{\mu}_i}{\partial Q}(\bar{Q})\, Q' \;+\; \frac{\partial^2 \tilde{\mu}_i}{\partial Q^2}(\bar{Q})\, Q'^2 \;+\; \mathcal{O}(Q'^3). \tag{5.6}$$

Substituting the first two terms in this Taylor expansion into (5.1) gives

$$\begin{aligned} \frac{dC_{A_i}}{dt} &= \left(\tilde{\mu}_i(\bar{Q}) - \frac{\tilde{\sigma}_i}{h} - \frac{Q_{out}}{V} - p_i C_Z\right) C_{A_i} + \frac{\partial \tilde{\mu}_i}{\partial Q}(\bar{Q})\, C_{A_i}\, Q' \\ &= h_i(\{C_A\}, t) + g_{i1}(\{C_A\}, t)\, Q'(t), \qquad \{C_A\} = C_{A_1}, C_{A_2}, ..., C_{A_N} \end{aligned} \tag{5.7}$$

This is a system of nonlinear Langevin equations with multiplicative noise $Q'(t)$.

Its Kramers-Moyal coefficients are given [102]

$$D_i(\{\boldsymbol{\xi}\},t) \;=\; h_i(\{\boldsymbol{\xi}\},t) \;+\; g_{kj}\frac{\partial}{\partial\xi_k}g_{ij}(\{\boldsymbol{\xi}\},t), \tag{5.8}$$

$$D_{ij}(\{\boldsymbol{\xi}\},t) \;=\; g_{ik}(\{\boldsymbol{\xi}\},t)\,g_{jk}(\{\boldsymbol{\xi}\},t), \tag{5.9}$$

$$D^{(v)}_{i_1\ldots i_v}(\{\boldsymbol{\xi}\},t) \;=\; \frac{1}{v!}\lim_{\tau\to 0}\frac{1}{\tau}\langle[C_{A_{i_1}}(t+\tau)-\xi_{i_1}]\ldots[C_{A_{i_v}}(t+\tau)-\xi_{i_v}]\rangle. \tag{5.10}$$

For a white-noise $Q'(t)$, such that

$$\langle Q'(t)\rangle = 0, \qquad \langle Q'(t)\,Q'(t')\rangle = 2\delta(t-t'), \tag{5.11}$$

all the Kramers-Moyal coefficients ($n \geq 3$) vanish [102], and we obtain a Fokker-Planck equation for the joint probability density functions of N algae groups' concentrations $W(\{\boldsymbol{\xi}\},t)$

$$\frac{\partial W}{\partial t} \;=\; \left[-\sum_{i=1}^{N}\frac{\partial}{\partial\xi_i}D_i(\{\boldsymbol{\xi}\},t) \;+\; \sum_{i,j=1}^{N}\frac{\partial^2}{\partial\xi_i\,\partial\xi_j}D_{ij}(\{\boldsymbol{\xi}\},t)\right]W. \tag{5.12}$$

There is a rich literature on solutions of the Fokker-Planck equations of several variables [102]. In the following section, we will consider the blue-green algae groups only.

5.4 Results and discussion

For a single algae group, e.g., Cynanobacteria, (5.12) reduced to:

$$\frac{\partial W_c}{\partial t} \;=\; \left[-\frac{\partial}{\partial\xi}D_1(\xi,t) \;+\; \frac{\partial^2}{\partial\xi^2}D_2(\xi,t)\right]W_c, \tag{5.13}$$

with the drift (D_1) and diffusion (D_2) coefficients

$$D_1 = \left(\tilde{\mu}_1(\bar{Q}) - \frac{\tilde{\sigma}_1}{h} - \frac{Q_{out}}{V} - p_1 C_Z\right)\xi + \left(\frac{\partial^2\tilde{\mu}_1}{\partial Q\partial\xi}(\bar{Q})\,\xi + \frac{\partial\tilde{\mu}_1}{\partial Q}(\bar{Q})\right)\frac{\partial\tilde{\mu}_1}{\partial Q}(\bar{Q})\,\xi, \tag{5.14a}$$

$$D_2 = \left(\frac{\partial\tilde{\mu}_1}{\partial Q}(\bar{Q})\,\xi\right)^2. \tag{5.14b}$$

Rearranging (5.13), we obtain a one-dimensional advection-diffusion equation

$$\frac{\partial W_c}{\partial t} = -\frac{\partial}{\partial\xi}\left(D_1 - \frac{\partial D_2}{\partial\xi}\right)W_c + \frac{\partial}{\partial\xi}\left(D_2\frac{\partial W_c}{\partial\xi}\right). \tag{5.15}$$

For the Monod dependence of the growth rate (5.2), this gives

$$\tilde{\mu}_1 = \mu_1\theta_1^{T-T_{ref}}\frac{I}{K_{I_1}+I}\,\frac{\bar{P}}{K_{P_1}+\bar{P}}\,\frac{\bar{N}}{K_{N_1}+\bar{N}} \tag{5.16a}$$

$$\frac{\partial\tilde{\mu}_1}{\partial\xi} = -\mu_1\theta_1^{T-T_{ref}}\frac{I}{K_{I_1}+I}\left[\frac{\alpha_1 K_{P_1}}{\left(K_{P_1}+\bar{P}\right)^2}\frac{\bar{N}}{K_{N_1}+\bar{N}} + \frac{\bar{P}}{K_{P_1}+\bar{P}}\frac{\beta_1 K_{N_1}}{\left(K_{N_1}+\bar{N}\right)^2}\right] \tag{5.16b}$$

$$\frac{\partial\tilde{\mu}_1}{\partial Q} = \frac{\mu_1}{V}\,\theta_1^{T-T_{ref}}\frac{I}{K_{I_1}+I}\left[\frac{K_{P_1}c_P}{\left(K_{P_1}+\bar{P}\right)^2}\,\frac{\bar{N}}{K_{N_1}+\bar{N}} + \frac{\bar{P}}{K_{P_1}+\bar{P}}\,\frac{K_{N_1}c_N}{\left(K_{N_1}+\bar{N}\right)^2}\right] \tag{5.16c}$$

$$\frac{\partial^2\tilde{\mu}_1}{\partial Q\partial\xi} = \frac{\mu_1}{V}\,\theta_1^{T-T_{ref}}\frac{I}{K_{I_1}+I}\,\left[\frac{2\alpha_1 K_{P_1}c_P}{\left(K_{P_1}+\bar{P}\right)^3}\,\frac{\bar{N}}{K_{N_1}+\bar{N}} + \frac{\bar{P}}{K_{P_1}+\bar{P}}\,\frac{2\beta_1 K_{N_1}c_N}{\left(K_{N_1}+\bar{N}\right)^3}\right.$$
$$\left. - \frac{K_{P_1}K_{N_1}(\beta_1 c_P + \alpha_1 c_N)}{\left(K_{P_1}+\bar{P}\right)^2\left(K_{N_1}+\bar{N}\right)^2}\right] \tag{5.16d}$$

It is noted that the mean concentrations of nutrients in the lake are

$$\bar{P} = P_0 + \frac{c_P\bar{Q}}{V} - \alpha_1\xi, \qquad \bar{N} = N_0 + \frac{c_N\bar{Q}}{V} - \beta_1\xi \tag{5.17}$$

Numerical simulations are performed to obtain a PDF solution to (5.15) with data from previous investigations [7, 98]. A Gaussian distribution $\mathcal{N}(5,1)$

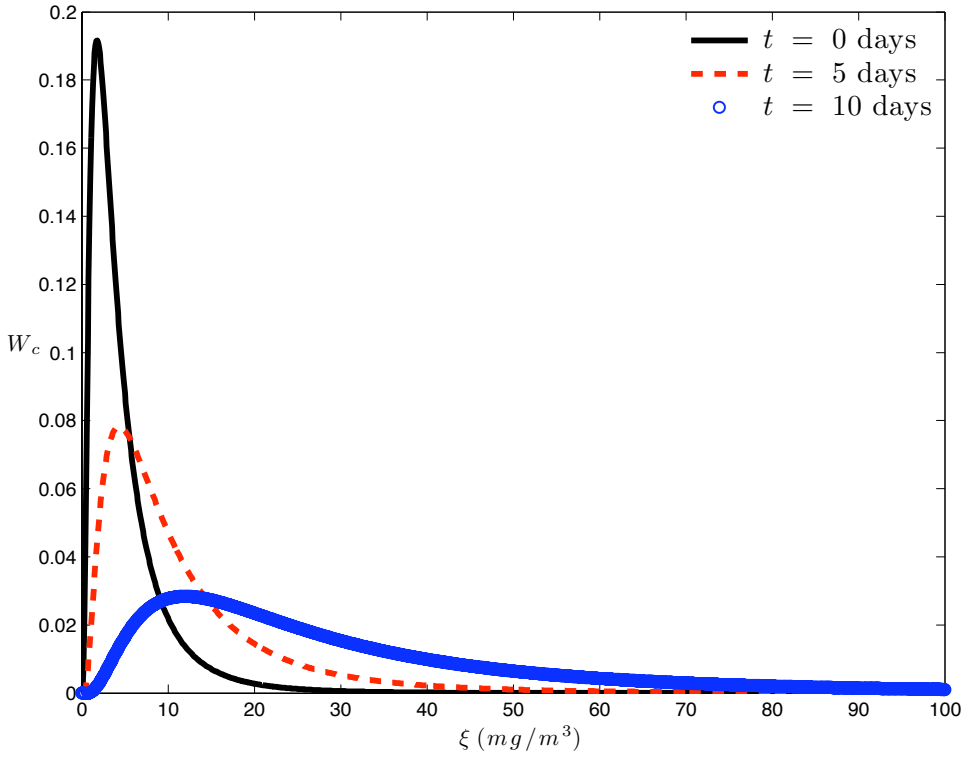

Figure 5.1: Temporal evolution of the cynanobacteria concentration PDF $W_c(\xi, t)$ at (a) $t = 0$ days, (b) $t = 5$ days and (c) $t = 10$ days.

is prescribed for the initial concentration PDF. Figure 5.1 demonstrates the temporal evolution of cynanobacteria population density PDF $W_c(\xi, t)$ through three timeframes: (a) $t = 0$ days, (b) $t = 5$ days and (c) $t = 10$ days. With continuous nutrients' inflow, blue-green algae exhibits rapid growth from its initial mean concentration of 5 mg m^{-3} to 33 mg m^{-3} over a week. At the same time, widening distributions (increasing variance) indicate a rising uncertainty in the prediction of algae concentration. Overall, the shape of W_c gradually diffuses and propagates with time, as expected from the advection-diffusion equation (5.15).

In Figure 5.2 the effect of average runoff volume $\bar{Q}$ on algae bloom is in-

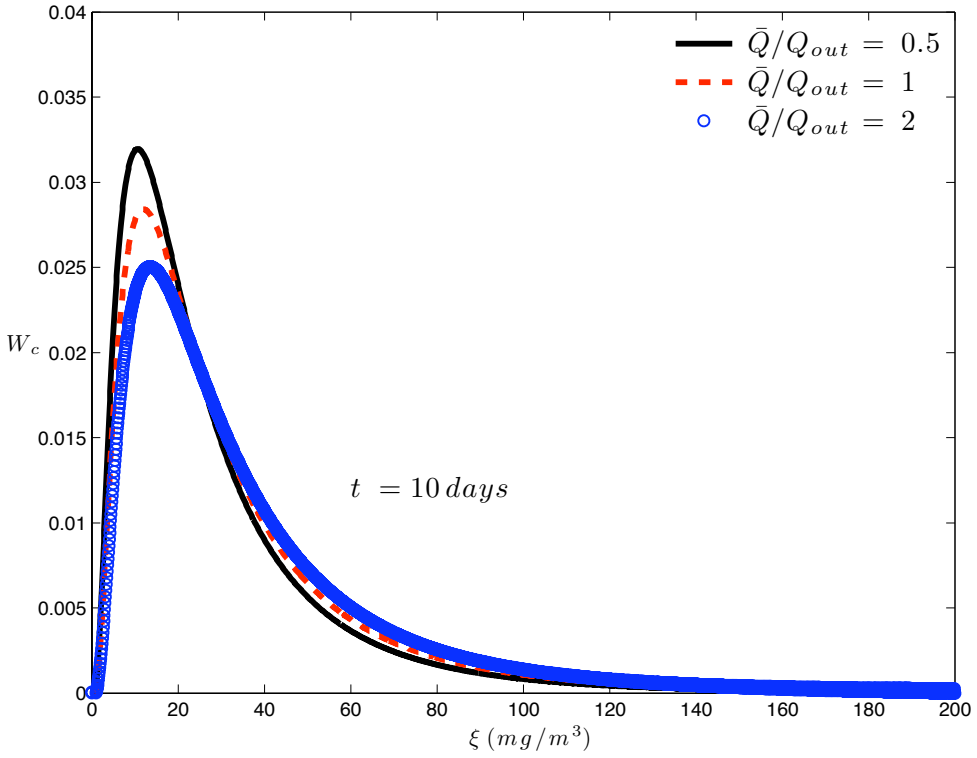

Figure 5.2: Effects of runoff $\bar{Q}$ on the cynanobacteria concentration PDF $W_c(\xi, t)$ at $t = 10$ days.

vestigated with respect to the lake outflow rate Q_{out}. After $t = 10$ days, greater inflows ($\bar{Q}/Q_{out} = 2$) have introduced more uncertainty to the predictions of the cynanobacteria population density, as indicated by larger variance. This is expected, because inflow nutrients are the primary factors contributing to algae bloom in lakes and provide the major source of system uncertainty. A reduction of its average value by higher outflow rate ($\bar{Q}/Q_{out} = 0.5$) would lead to smaller predictive uncertainty. However, the overall impact is limited due to its small volume relative to the volume of the lake.

The initial concentration exerts great influence on the development of algae

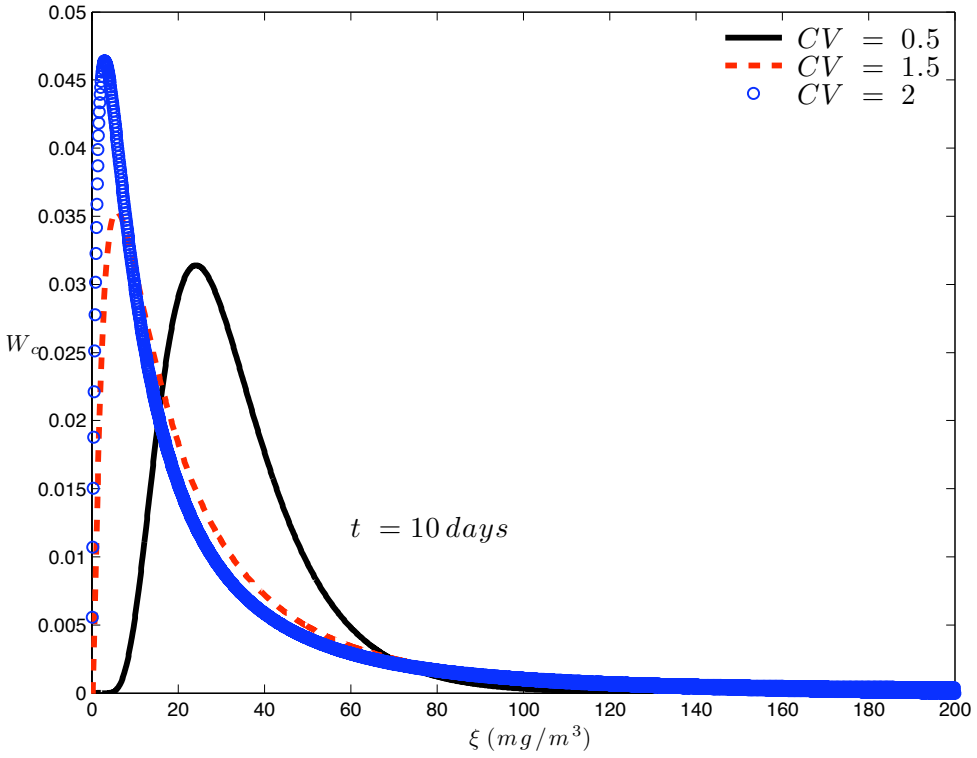

Figure 5.3: The cynanobacteria concentration PDF $W_c(\xi, t)$ at $t = 10$ days for different levels of uncertainty at initial concentration.

bloom in lakes. This is illustrated in Figure 5.3 for $W_c(\xi, t)$ at $t = 10$ days with three CV (standard deviation divided by mean value) levels of initial distribution. Larger initial fluctuations lead to greater uncertainty (longer distribution tails) at later stages of algae growth.

5.5 Conclusion

We present a general framework to quantify the uncertainty in nutrients' inflow via a full statistical description (PDF) of the concentration of several phy-

toplankton groups. Based on a physical model routinely used for algae population dynamics in a lake, a Fokker-Planck equation is derived. Its solution yields the joint PDF of multiple-algae-species concentrations.

1. The new framework enables one to obtain probabilistic density functions (PDFs) of several phytoplankton groups' concentrations. Its results also facilitate probabilistic risk assessments, which rely on probabilities of rare events.
2. Average runoff inflow has limited impact on algae growth due to its relative small volume compared to the volume of the lake.
3. Uncertainty in initial concentration significantly affects overall predictive uncertainty.

Chapter 6

Conclusions

This dissertation leads to the following major conclusions:

1. Given high temporal variability of precipitation in (semi-)arid regions, the reliance on annual meteorological data might significantly underestimate the downward migration of contaminant through the vadose zone in (semi-)arid regions. Specially, predictive errors stemming from such data increase with time and are more pronounced in highly conductive soils. It is also influenced by the selection of constitutive models for the Richards equation, e.g., van Genuchten model versus Brooks-Corey model. Additional errors associated with spatial averaging further undermine the accuracy of predictions based on annual data when one takes account of surface topography and built environments.

2. The two reduced complexity models provide accurate results of PDFs of

infiltration rate compared to their counterpart resulting from the Richards equation. Though this agreement deteriorates with time, the reduced complexity models provide conservative estimates of predictive uncertainty at all times. Reliance on the reduced complexity models of infiltration into heterogeneous soils with uncertain hydraulic parameters offers a number of advantages. Not only it allows one to compute closed-form semi-analytical expressions for single-point PDFs of the infiltration rate, it does so exactly, without introducing linearization errors that plague most stochastic analyses of the Richards equation. It also allows one evaluate probabilities of rare events. The reduced complexity models are capable of quantifying uncertainty in any number of hydraulic parameters and can be used with arbitrary constitutive laws (relative conductivity functions and retention curves). It has been shown that predictive uncertainty (the infiltration-depth and infiltration-rate PDFs) is significantly more sensitive to the coefficient of variation of saturated hydraulic conductivity K_s than to that of the fitting parameters in the Gardner or van Genuchten hydraulic function, while it is also greatly influenced by the degree of cross-correlation between hydraulic parameters K_s and α_{VG}. For Green-Ampt model, the choice of a functional form of the hydraulic function (e.g., the Gardner model vs. the van Genuchten model) has a significant effect on predictive uncertainty during early stages of infiltration - this effect diminishes with time. Finally, reduced complexity models make it possible to compute multi-point PDFs of infiltration rate. The latter can be used both

to predict (cross-)correlations of infiltration rates at multiple locations and to assimilate infiltration data via a straightforward Bayesian updating.

3. We developed a probabilistic approach to quantify parametric uncertainty in first-order hyperbolic conservation laws (kinematic wave equations). The approach relies on the derivation of a deterministic equation for the cumulative density function (CDF) of the system state, in which probabilistic descriptions (probability density functions or PDFs) of the system parameters and/or initial and boundary conditions serve as inputs. The accuracy and robustness of solutions of the CDF equation for one such system, the Saint-Venant equations of river flows, were investigated via comparison with Monte Carlo simulations and yielded satisfactory results. CDF equations, and their (semi-)analytical solutions, provide a computationally efficient alternative to the existing methods for uncertainty quantification, such as Monte Carlo simulations and stochastic finite element methods (polynomial chaos expansions, stochastic collocation methods, etc.). In addition, CDF equations are ideally suited for handling input parameters and/or initial and boundary conditions that exhibit small correlation lengths. This is in contrast with stochastic finite element methods and other numerical approaches that rely on the Karhunen-Loève representation of random parameter fields. Lastly, CDF equations offer an operational advantage over PDF equations that are often used in other contexts, e.g., to analyze transport of passive tracers and reactive species in turbulent (randomly fluctuating) velocity fields. This is because CDF equa-

tions allow for straightforward and unambiguous determination of boundary conditions with respect to sample variables.

4. We propose a general framework to obtain probabilistic density functions (PDFs) of several phytoplankton groups' concentrations with uncertain nutrients' inflow. Its results also facilitate probabilistic risk assessment which typically concerns probabilities of rare events. Specifically we find that uncertainty of initial concentration significantly affects overall predictive uncertainty, while average runoff inflow has limited impact on algae growth due to its relative small volume compared to the volume of the lake.

Appendix A

Statistical properties of γ

The random coefficient $\gamma = \left(\sqrt{s_0}/n\right)^{-\beta}$ is defined in terms of the two random parameters, $s_0(x)$ and $n(x)$. Its single-point PDF $p_\gamma(\Gamma; x)$ can be expressed in terms of the PDFs of $s_0(x)$ and $n(x)$ as follows. Let $G_\gamma(\Gamma) = P(\gamma \leq \Gamma)$ denote the cumulative density function of γ, i.e., the probability that the random coefficient γ at point x takes on a value not larger than Γ. By definition,

$$G_\gamma(\Gamma) = \int_0^\infty \int_0^{N(\Gamma,S_0)} p_{n,s_0}(N, S_0)\mathrm{d}N\mathrm{d}S_0 \tag{A.1}$$

where $p_{n,s_0}(N, S_0)$ is the joint PDF of the channel slope s_0 and the Manning coefficient n at point x. The PDF p_γ can now be obtained as

$$p_\gamma(\Gamma) = \frac{\mathrm{d}G_\gamma}{\mathrm{d}\Gamma} = \int_0^\infty p_{n,s_0}\left[N(\Gamma, S_0), S_0\right] \frac{\partial N(\Gamma, S_0)}{\partial \Gamma}\mathrm{d}S_0. \tag{A.2}$$

If s_0 and n are mutually independent, (A.2) reduces to

$$p_\gamma(\Gamma) = \frac{1}{\beta}\Gamma^{1/\beta-1} \int_0^\infty p_n[N(\Gamma, S_0)] p_{s_0}(S_0)\sqrt{S_0}\mathrm{d}S_0. \tag{A.3}$$

If $Y_1(x) = \ln s_0(x)$ and $Y_2(x) = \ln n(x)$ are mutually uncorrelated multivariate Gaussian stationary (statistically homogeneous) random fields, their two-point PDFs are given by

$$p_{_2Y_i}(\xi_1, \xi_2) = \frac{1}{2\pi\sigma^2_{Y_i}\sqrt{1-\rho^2_{Y_i}}} \exp\left[-\frac{R_i(\xi_1,\xi_2)}{2\sigma^2_{Y_i}(1-\rho^2_{Y_i})}\right], \qquad i = 1, 2, \tag{A.4a}$$

where

$$R_i = (\xi_1 - \overline{Y}_1)^2 - 2\rho_{Y_i}(\xi_1 - \overline{Y}_i)(\xi_2 - \overline{Y}_i) + (\xi_2 - \overline{Y}_i)^2 \tag{A.4b}$$

and $\rho_{Y_i}(x_1, x_2)$ denotes the linear correlation function between $Y_i(x_1)$ and $Y_i(x_2)$. The two-point covariance of $\gamma(x)$, $C_\gamma(x_1, x_2) = \langle\gamma'(x_1)\gamma'(x_2)\rangle$, is defined by

$$C_\gamma(x_1, x_2) = \langle\gamma(x_1)\gamma(x_2)\rangle - \overline{\gamma}^2 = C_1(x_1, x_2)C_2(x_1, x_2) - \overline{\gamma}^2. \tag{A.5}$$

The covariances $C_1(x_1, x_2) = \langle[(s_0(x_1)s_0(x_2)]^{-\beta/2}\rangle$ and $C_2(x_1, x_2) = \langle[n(x_1)n(x_2)]^\beta\rangle$ can be expressed in terms of the statistics of $s_0(x)$ and $n(x)$,

$$C_1(x_1, x_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{\beta}{2}(S_1+S_2)} p_{_2Y_1}(S_1, S_2)\mathrm{d}S_1\mathrm{d}S_2 = \mathrm{e}^{-\beta\overline{Y}_1+\beta^2(1+\rho_{Y_1})\sigma^2_{Y_1}/4} \tag{A.6}$$

and

$$C_2(x_1, x_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{\beta(N_1+N_2)} p_{_2Y_2}(N_1, N_2)\mathrm{d}N_1\mathrm{d}N_2 = \mathrm{e}^{2\beta\overline{Y}_2+\beta^2(1+\rho_{Y_2})\sigma^2_{Y_2}}. \tag{A.7}$$

Appendix B

Derivation of Raw CDF Equation

It follows from the definition of Π in (4.7) that

$$\frac{\partial \Pi}{\partial x} = \frac{\partial \mathcal{H}[q(x,t) - Q]}{\partial x} = \frac{\partial \Pi}{\partial q}\,\frac{\partial q}{\partial x} = -\frac{\partial \Pi}{\partial Q}\frac{\partial q}{\partial x}. \tag{B.1}$$

Multiplying (4.4) with $\partial \Pi / \partial Q$ and making use of (B.1) yields

$$\beta\gamma q^{\beta-1}\frac{\partial \Pi}{\partial Q}\frac{\partial q}{\partial t} + \frac{\partial \Pi}{\partial Q}\frac{\partial q}{\partial x} = \frac{\partial \Pi}{\partial Q}S. \tag{B.2}$$

Since $\partial \Pi / \partial Q = \delta(Q - q)$ where $\delta(\cdot)$ is the Dirac delta function, and since for any test function $g(\cdot)$ the following relation holds $g(q)\delta(Q - q) = g(Q)\delta(Q - q)$, one can rewrite (B.2) as

$$\beta\gamma Q^{\beta-1}\frac{\partial \Pi}{\partial Q}\frac{\partial q}{\partial t} - \frac{\partial \Pi}{\partial x} = \frac{\partial \Pi}{\partial Q}S. \tag{B.3}$$

Finally, substituting the relation

$$\frac{\partial \Pi}{\partial t} = \frac{\partial \mathcal{H}[q(x,t) - Q]}{\partial t} = \frac{\partial \Pi}{\partial q}\,\frac{\partial q}{\partial t} = -\frac{\partial \Pi}{\partial Q}\frac{\partial q}{\partial t} \tag{B.4}$$

into (B.3) yields an equation for the raw CDF (4.9).

Appendix C

Solution for $S = 0$

Taking the Laplace transformation of (4.9) with $S \equiv 0$ yields

$$\frac{\mathrm{d}\hat{\Pi}}{\mathrm{d}x} + \beta Q^{\beta-1}\gamma s\hat{\Pi} = \beta Q^{\beta-1}\gamma\Pi_{\mathrm{in}}, \tag{C.1}$$

where $\hat{\Pi}(Q; x, s)$ is the Laplace transform of $\Pi(Q; x, t)$. This equation is subject to the boundary condition obtained from (4.10b),

$$\hat{\Pi}(Q, x = 0, s) = \hat{\Pi}_0 = \int_0^\infty \mathcal{H}\left[Q - q_0(t)\right]\, \mathrm{e}^{-st}\, \mathrm{d}t. \tag{C.2}$$

A solution of (C.1) and (C.2) is

$$\hat{\Pi} = \int_0^x \mathrm{e}^{-s[C(x)-C(x')]}B(x')\mathrm{d}x' + \hat{\Pi}_0\, \mathrm{e}^{-sC(x)}, \tag{C.3}$$

where

$$B(x) = \beta Q^{\beta-1}\gamma(x)\mathcal{H}[Q - q_{\mathrm{in}}(x)], \qquad C(x) = \int_0^x \beta Q^{\beta-1}\gamma(x')\mathrm{d}x'. \tag{C.4}$$

The inverse Laplace transform of (C.3)–(C.4) is given by

$$\Pi = \int_0^x \delta[t - C(x) + C(x')]B(x')\mathrm{d}x' + \mathcal{H}(t - C)\mathcal{H}[Q - q_0(t - C)]. \tag{C.5}$$

Evaluating the quadrature, while recalling the definitions of B and C in (C.4), yields

$$\Pi = \mathcal{H}[Q - q_{\text{in}}(x^\star)] + \mathcal{H}(t - C)\mathcal{H}[Q - q_0(t - C)]. \tag{C.6}$$

Here $x' = x^\star$ is a solution of the equation

$$C(x') = C(x) - t \tag{C.7}$$

for a given x and t. It follows from (C.4) that $C(x' = x^\star) \geq 0$ for all x and t. This imposes the constraint $C(x) \geq t$ on the parameter space of (C.7), which translates into the Heaviside function $\mathcal{H}(C - t)$ in (4.14).

Appendix D

Solution for $S = S(x)$

Taking the Laplace transformation of (4.9)–(4.10) with $S = S(x)$ yields

$$\frac{\partial \hat{\Pi}}{\partial x} + S(x)\frac{\partial \hat{\Pi}}{\partial Q} = -\beta Q^{\beta-1}\gamma(x)(s\hat{\Pi} - \Pi_{\text{in}}) \tag{D.1}$$

subject to the boundary conditions

$$\hat{\Pi}(Q; 0, s) = \hat{\Pi}_0(Q, s), \qquad \hat{\Pi}(0; x, s) = 0. \tag{D.2}$$

A family of characteristics $Q = Q(x; \xi)$ is defined by

$$\frac{\mathrm{d}Q}{\mathrm{d}x} = S(x), \qquad Q(x = 0) = \xi, \tag{D.3}$$

which yields an equation for characteristics

$$Q = \int_0^x S \,\mathrm{d}x' + \xi. \tag{D.4}$$

The "label" ξ defines the origin of each characteristic line, such that (see Figure D.1)

1. for $\xi \geq 0$, characteristics originate from the Q-axis ($x = 0$) and the solution is determined by the boundary condition on x;

2. for $\xi < 0$, characteristics originate from the x-axis ($x = \eta$) and the solution is determined by the boundary condition on Q. The constant η is a solution of $\int_0^\eta S \, \mathrm{d}x' = -\xi$.

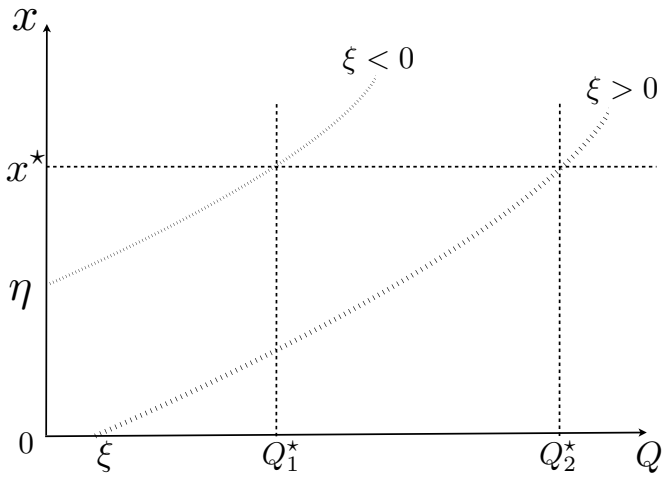

Figure D.1: Characteristic curves in the (x, Q) plane for $\hat{\Pi}(Q; x, s)$.

Along the characteristics (D.4), the equation (D.1) takes the form

$$\frac{\mathrm{d}\hat{\Pi}}{\mathrm{d}x} = -\beta Q^{\beta-1}\gamma(x)(s\hat{\Pi} - \Pi_{\mathrm{in}}). \tag{D.5}$$

The two boundary conditions in (D.2) give rise to the boundary condition for (D.5),

$$\hat{\Pi}(Q; x_0, s) = \hat{\Pi}_0(\xi, s), \qquad x_0 = \begin{cases} 0 & \xi \geq 0 \\ & \\ \eta & \xi < 0 \end{cases}, \qquad \hat{\Pi}_0 = \int_0^\infty \mathcal{H}\left[\xi - q_0(t)\right] \mathrm{e}^{-st} \, \mathrm{d}t. \tag{D.6}$$

Substituting (D.4) into (D.5), solving the resulting ODE, and eliminating ξ in favor of x and Q in the solution, yields

$$\hat{\Pi} = \int_{x_0}^{x} \mathrm{e}^{-s(C-A)} B \mathrm{d}x' + \hat{\Pi}_0(x_0, s)\mathrm{e}^{-sC} \tag{D.7}$$

where

$$A = \int_{x_0}^{x'} \beta[Q - I(x, x'')]^{\beta-1}\gamma \mathrm{d}x'' \tag{D.8a}$$

$$B = \beta[Q - I(x, x')]^{\beta-1}\gamma \mathcal{H}[Q - I(x, x') - q_{\mathrm{in}}(x')] \tag{D.8b}$$

$$C = \int_{x_0}^{x} \beta[Q - I(x, x'')]^{\beta-1}\gamma \mathrm{d}x'', \qquad I(x, x') = \int_{x'}^{x} S \, \mathrm{d}x''. \tag{D.8c}$$

The inverse Laplace transform of (D.7)–(D.8) is given by

$$\Pi = \int_{x_0}^{x} \delta(t - C + A) B \mathrm{d}x' + \mathcal{H}(t - C)\mathcal{H}[Q - I(x, 0) - q_0(t - C)]. \tag{D.9}$$

Evaluating the quadrature, while recalling the definitions of A, B and C in (D.8), yields

$$\Pi = \mathcal{H}[Q - I(x, x^\star) - q_{\mathrm{in}}(x^\star)] + \mathcal{H}(t - C)\mathcal{H}[Q - I(x, 0) - q_0(t - C)]. \tag{D.10}$$

Here $x' = x^\star$ is a solution of the equation

$$A(x, x') = C(x, x') - t \tag{D.11}$$

for a given x and t. It follows from (D.8) that $A(x, x^\star) \geq 0$ for all x and t. This imposes the constraint $C(x, x') \geq t$ on the parameter space of (D.11), which translates into the Heaviside function $\mathcal{H}(C - t)$ in (4.16).

Appendix E

Integration of correlated random fields

For intermediate x, we follow the approach presented in [95] to compute the statistics of the integral $I_\gamma(x)$ in (4.18). It is briefly reviewed here for completeness. Let us subdivide the integration interval $[0, x]$ into $2N$ intervals of length $\Delta = x/(2N)$. Then (4.18) can be rewritten as

$$I_\gamma(x) = \sum_{i=1}^{N}(I_i + J_i), \qquad I_i = \int_{(2i-1)\Delta}^{2i\Delta} \gamma(x')\mathrm{d}x', \qquad J_i = \int_{(2i-2)\Delta}^{(2i-1)\Delta} \gamma(x')\mathrm{d}x'. \tag{E.1}$$

Since $\gamma(x)$ is stationary, the integrals I_i and J_i $(i = 1, \ldots, N)$ have the same mean $\overline{I} = \overline{\gamma}\Delta$ and variance

$$\sigma_I^2 = \sigma_\gamma^2 \int_{(2i-1)\Delta}^{2i\Delta}\int_{(2i-1)\Delta}^{2i\Delta} \rho_\gamma(x' - x'')\mathrm{d}x'\mathrm{d}x'' = 2\sigma_\gamma^2 \int_0^\Delta (\Delta - x)\rho_\gamma(y)\mathrm{d}y. \tag{E.2}$$

The correlation coefficient between the two sums is given by

$$\rho_N\left(\sum_{i=1}^{N} I_i, \sum_{j=1}^{N} I_j\right) = \frac{2N-1}{2N} \frac{\int_0^\Delta \rho_\gamma(y)\mathrm{d}y}{\int_0^\Delta (\Delta - y)\rho_\gamma(y)\mathrm{d}y}. \tag{E.3}$$

According to the central limit theorem for dependent processes, $I_\gamma(x) = \sum_{i=1}^{N}(I_i + J_i)$ is asymptotically (as $N \to \infty$) Gaussian with mean $2N\overline{I}$ and variance $2N(1+\rho_N)\sigma_I^2$.

Bibliography

[1] E.G. Lapalla, R.W. Healy, and E.P. Weeks, "Documentation of computer program vs2d to solve the equations of fluid flow in variably saturated porous media," Tech. Rep. 89-4099, U.S. Geological Survey, 1987.

[2] D. Russo and M. Bouton, "Statistical analysis of spatial variability in unsaturated flow parameters," *Water Resour. Res.*, vol. 28, no. 7, pp. 1911–1925, 1992.

[3] T. K. Gates and M. Al-Zahrani, "Spatiotemporal stochastic open-channel flow. I: Model and its parameter data," *J. Hydrol. Engrg.*, vol. 122, no. 11, pp. 641–651, 1996.

[4] T. K. Gates and M. Al-Zahrani, "Spatiotemporal stochastic open-channel flow. II: Simulation experiments," *J. Hydrol. Engrg.*, vol. 122, no. 11, pp. 652–661, 1996.

[5] T. Moramarco and V. P. Singh, "A practical method for analysis of river waves and for kinematic wave routing in natural channel networks," *Hydrol. Process.*, vol. 14, pp. 51–62, 2000.

[6] L. Liang and M. L. Kavvas, "Modeling of solute transport and macrodispersion by unsteady stream flow under uncertain conditions," *J. Hydrol. Engrg.*, vol. 13, no. 6, pp. 510–520, 2008.

[7] Olli Malve, Marko Laine, Heikki Haario, Teija Kirkkala, and Jouko Sarvala, "Bayesian modelling of algal mass occurrences–using adaptive mcmc methods with a lake water quality model," *Environmental Modelling and Software*, vol. 22, no. 7, pp. 966–977, 2007.

[8] Norbert Wiener, "The homogeneous chaos," *American Journal of Mathematics*, vol. 60, no. 4, pp. 897–936, 10 1938.

[9] U. Frisch, *Probabilistic methods in applied mathematics*, vol. 1, Academic, New York, 1968.

[10] G. C. Papanicolaou, D. McLaughlin, and R. Burridge, "A stochastic gaussian beam," *Journal of Mathematical Physics*, vol. 14, no. 1, pp. 84–89, 1973.

[11] N. G. Van Kampen, "Stochastic differential equations," *Physics Reports*, vol. 24, no. 3, pp. 171–228, 3 1976.

[12] C. Tong, "Uncertainty quantification," `https://computation.llnl.gov/casc/uncertainty_quantification/`, May 2005.

[13] D. M. Tartakovsky, "Probabilistic risk analysis in subsurface hydrology," *Geophys. Res. Lett.*, vol. 34, pp. L05404, 2007.

[14] M. I Shvidler, *Filtration Flows in Heterogeneous Media: A Statistical Approach*, Springer, New York, 1964.

[15] Y. Rubin, *Applied Stochastic Hydrogeology*, Oxford University Press, New York, 2003.

[16] C. L. Winter and D. M. Tartakovsky, "Groundwater flow in heterogeneous composite aquifers," *Water Resour. Res.*, vol. 38, no. 8, 2002.

[17] L. Guadagnini, A. Guadagnini, and D. M. Tartakovsky, "Probabilistic reconstruction of geologic facies," *J. Hydrol.*, vol. 294, pp. 57–67, 2004.

[18] G. Christakos, "A bayesian maximum entropy view to the spatial estimation problem," *Math. Geol.*, vol. 22, no. 7, pp. 763–777, 1990.

[19] S. P. Neuman, "Maximum likelihood bayesian averaging of uncertain model predictions," *Stochastic Environ. Res. Risk. Assess*, vol. 17, no. 5, pp. 291–305, 2003.

[20] R. Ghanem, "Scales of fluctuation and the propagation of uncertainty in random porous media,," *Water Resour. Res.*, vol. 34, no. 9, pp. 2123–2136, 1998.

[21] "Understanding risk: Informing decisions in a democratic society," Tech. Rep., National Research Council (NRC), Washington, D.C: National Academy Press, 1996.

[22] T. Bedford and R. Cooke, *Probabilistic Risk Analysis: Foundations and Methods*, Cambridge Univ. Press, 2003.

[23] J. A. Izbicki, J. Radyk, and R. L. Michel, "Water movement through a thick unsaturated zone underlying an intermittent stream in the western Mojave Desert, southern California, USA," *J. Hydrol.*, vol. 238, no. 3-4, pp. 194–217, 2000.

[24] J. Izbicki, J. Radyk, and R. Michel, "Movement of water through the thick unsaturated zone underlying Oro Grande and Sheep Creek Washes in the western Mojave Desert, USA," *Hydrogeol. J.*, vol. 10, no. 3, pp. 409–427, 2002.

[25] P. J. Wierenga, "Solute distribution profiles computed with Steady-State and transient water movement models," *Soil Sci. Soc. Am. J.*, vol. 41, no. 6, pp. 1050–1055, 1977.

[26] F. Beese and P. J. Wierenga, "Solute transport through soil with adsorption and root water uptake computed with a transient and a Constant-Flux model," *Soil Sci.*, vol. 129, no. 4, pp. 245–252, 1980.

[27] E. Bresler and G. Dagan, "Unsaturated flow in spatially variable fields 3. Solute transport models and their application to two fields," *Water Resour. Res.*, vol. 19, no. 2, pp. 429–435, 1982.

[28] D. Russo, W. A. Jury, and G. L. Butters, "Numerical analysis of solute transport during transient irrigation, 1. The effect of hysteresis and profile heterogeneity," *Water Resour. Res.*, vol. 25, no. 10, pp. 2109–2118, 1989.

[29] G. Destouni, "Applicability of the steady state flow assumption for solute advection in field soils," *Water Resour. Res.*, vol. 27, no. 8, pp. 2129–2140, 1991.

[30] J. D. Marshall, B. Wilson Shimada, and P. R. Jaffe, "Effect of temporal variability in infiltration on contaminant transport in the unsaturated zone," *J. Contam. Hydrol.*, vol. 46, no. 1-2, pp. 151–161, 2000.

[31] G. Schoups, J. W. Hopmans, and K. K. Tanji, "Evaluation of model complexity and space-time resolution on the prediction of long-term soil salinity dynamics, western san joaquin valley, california," *Hydrolog. Process.*, vol. 20, no. 13, pp. 2647–2668, 2006.

[32] F. H. Lambert, A. R. Stine, N. Y. Krakauer, and J. C. H. Chiang, "How much will precipitation increase with global warming?," *EOS Trans., AGU*, vol. 89, no. 12, pp. 193–194, 2008.

[33] D. Ackerman and K. Schiff, "Modeling storm water mass emissions to the Southern California bight," *J. Envir. Engrg.*, vol. 129, no. 4, pp. 308–317, 2003.

[34] N.E. Edelfson and A.B.C. Anderson, "Thermodynamics of soil moisture," *Hilgardia*, vol. 15, no. 2, pp. 31–298, 1943.

[35] W. O. Pruitt and J. Doorenbos, "Empirical calibration, a requisite for evapotranspiration formulae based on daily or longer mean climatic data?," in *Proceedings of the International Round Table Conference on "Evapotranspiration"*. International Commission on Irrigation and Drainage, 1977, p. 22.

[36] R. W. Healy, "Simulation of solute transport in variably saturated porous media with supplemental information on modifications to the USGS's computer program VS2DT," Tech. Rep. 90-4025, USGS, 1990.

[37] H. J. Morel-Seytoux, Ph. D. Meyer, M. Nachabe, Jaoudat Tourna, M. T. van Genuchten, and R. J. Lenhard, "Parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics: Preserving the effective capillary drive," *Water Resour. Res.*, vol. 32, no. 5, pp. 1251–1258, 1996.

[38] D. M. Tartakovsky, A. Guadagnini, and M. Riva, "Stochastic averaging of nonlinear flows in heterogeneous porous media," *J. Fluid Mech.*, vol. 492, pp. 47–62, 2003.

[39] J. J. Gómez-Hernández and X.-H. Wen, "To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology," *Adv. Water Resour.*, vol. 21, no. 1, pp. 47–61, 1998.

[40] Y. Mualem, "A new model for predicting the hydraulic conductivity of unsaturated porous media," *Water Resour. Res.*, vol. 12, no. 3, pp. 513–522, 1976.

[41] J. Zhu, M. H. Young, and M. Th. van Genuchten, "Upscaling schemes and relationships for the Gardner and van Genuchten hydraulic functions for heterogeneous soils," *Vadose Zone J.*, vol. 6, no. 1, pp. 186–195, 2007.

[42] D. Russo, I. Russo, and A. Laufer, "On the spatial variability of parameters of the unsaturated hydraulic conductivity," *Water Resour. Res.*, vol. 33, no. 5, pp. 947–956, 1997.

[43] T.-C. J. Yeh, L. W. Gelhar, and A. L. Gutjahr, "Stochastic analysis of unsaturated flow in heterogeneous soils. 2. Statistically anisotropic media with variable α," *Water Resour. Res.*, vol. 21, no. 4, pp. 457–464, 1985.

[44] A. Mantoglou and L. W. Gelhar, "Stochastic modeling of large-scale transient unsaturated flow system," *Water Resour. Res.*, vol. 23, no. 1, pp. 37 – 46, 1987.

[45] X. Hu and J. Cushman, "Nonequilibrium statistical mechanical derivation of a nonlocal Darcy's Law for unsaturated/saturated flow," *Stoch. Hydrol. Hydraul.*, vol. 8, no. 2, pp. 109–116, 1994.

[46] D. Russo, "Stochastic analysis of the velocity covariance and the displacement covariance tensors in partially saturated heterogeneous anisotropic porous formations," *Water Resour. Res.*, vol. 31, no. 7, pp. 1647–1658, 1995.

[47] G. Severino and A. Santini, "On the effective hydraulic conductivity in mean vertical unsaturated steady flows," *Adv. Water Resour.*, vol. 28, no. 9, pp. 964–974, 2005.

[48] D. Russo and A. Fiori, "Stochastic analysis of transport in a combined heterogeneous vadose zone–groundwater flow system," *Water Resour. Res.*, vol. 45, pp. 10.1029/2008WR007157, 2009.

[49] D. M. Tartakovsky, S. P. Neuman, and Z. Lu, "Conditional stochastic averaging of steady state unsaturated flow by means of Kirchhoff transformation," *Water Resour. Res.*, vol. 35, no. 3, pp. 731–745, 1999.

[50] A. M. Tartakovsky, L. Garcia-Naranjo, and D. M. Tartakovsky, "Transient flow in a heterogeneous vadose zone with uncertain parameters," *Vadose Zone J.*, vol. 3, no. 1, pp. 154–163, 2004.

[51] Z. Lu, S. P. Neuman, A. Guadagnini, and D. M. Tartakovsky, "Conditional moment analysis of steady state unsaturated flow in bounded, randomly heterogeneous soils," *Water Resour. Res.*, vol. 38, no. 4, pp. 10.1029/2001WR000278, 2002.

[52] G. Dagan and E. Bresler, "Unsaturated flow in spatially variable fields 1. Derivation of models of infiltration and redistribution," *Water Resour. Res.*, vol. 19, no. 2, pp. 413–420, 1983.

[53] Arthur W. Warrick, *Soil Water Dynamics*, Oxford University Press, 2003.

[54] J-Y Parlange, I. Lisel, R. D. Braddock, and R. E Smith, "The three-parameter infiltration equation," *Soil Sci*, vol. 133, no. 6, pp. 337–341, 1982.

[55] Y. Rubin and D. Or, "Stochastic modeling of unsaturated flow in heterogeneous soils with water uptake by plant roots: The parallel columns model," *Water Resour. Res.*, vol. 29, no. 3, pp. 619–631, 1993.

[56] J. Zhu and B. P. Mohanty, "Soil hydraulic parameter upscaling for steady-state flow with root water uptake," *Vadose Zone J.*, vol. 3, no. 4, pp. 1464–1470, 2004.

[57] E. Bresler and G. Dagan, "Unsaturated flow in spatially variable fields 2. Application of water flow models to various fields," *Water Resour. Res.*, vol. 19, no. 2, pp. 421–428, 1983.

[58] P. Indelman, I. Touber-Yasur, B. Yaron, and G. Dagan, "Stochastic analysis of water flow and pesticides transport in a field experiment," *J. Contam. Hydrol.*, vol. 32, no. 1-2, pp. 77–97, 1998.

[59] H. Meng, T. R. Green, J. D. Salas, and L. R. Ahuja, "Development and testing of a terrain-based hydrologic model for spatial Hortonian infiltration and runoff/on," *Environ. Model. Soft.*, vol. 23, no. 6, pp. 794–812, 2008.

[60] R. Morbidelli, C. Corradini, and R. S. Govindaraju, "A simplified model for estimating field-scale surface runoff hydrographs," *Hydrol. Process.*, vol. 21, no. 13, pp. 1772–1779, 2007.

[61] K. F. Zeller and N. T. Nikolov, "Quantifying simultaneous fluxes of ozone, carbon dioxide and water vapor above a subalpine forest ecosystem," *Environ. Pollut.*, vol. 107, no. 1, pp. 1–20, 2000.

[62] J. Zhu and B. P. Mohanty, "Spatial averaging of van Genuchten hydraulic parameters for steady-state flow in heterogeneous soils: A numerical study," *Vadose Zone J.*, vol. 1, no. 2, pp. 261–272, 2002.

[63] S. Gómez, G. Severino, L. Randazzo, G. Toraldo, and J. M. Otero, "Identification of the hydraulic conductivity using a global optimization method," *Agric. Water Manage.*, vol. 96, pp. 504–510, 2009.

[64] H. Saito, K. Seki, and J. Simunek, "An alternative deterministic method for the spatial interpolation of water retention parameters," *Hydrol. Earth Syst. Sci.*, vol. 13, pp. 453–465, 2009.

[65] R. Allan Freeze, "A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope," *Water Resour. Res.*, vol. 16, no. 2, pp. 391–408, 1980.

[66] S. P. Neuman, "Wetting front pressure head in the infiltration model of green and ampt," *Water Resour. Res.*, vol. 12, no. 3, pp. 564–566, 1976.

[67] D. A. Barry, J.-Y. Parlange, G. C. Sander, , and M. Sivaplan, "A class of exact solutions for richards' equation.," *Journal of Hydrology*, vol. 142, no. 1-4, pp. 29–46, 1993.

[68] R. Haverkamp, J.-Y. Parlange, J. L. Starr, G. Schmitz, and C. Fuentes, "Infiltration under ponded conditions: 3. a predictive equation based on physical parameters," *Soil Sci*, vol. 149, no. 5, 1990.

[69] J.-Y. Parlange, "On solving the flow equation n unsaturated soils by optimization: horizontal infiltration," *Soil Sci. Soc. Am. Proc.*, , no. 39, pp. 415–418, 1975.

[70] J. Y. Parlange and J.T.R. Haverkamp, "Infiltration under ponded conditions: 1. optimal analytical solution and comparison with experimental observations," *Soil Sci*, vol. 139, pp. 205–311, 1985.

[71] D. A. Barry, J.-Y. Parlange, R. Haverkamp, and P. J. Ross, "Infiltration under ponded conditions: 4. an explicit predictive infiltration formula," *Soil Sci*, vol. 160, no. 1, 1995.

[72] S. E. Serrano, "Improved decomposition solution to Green and Ampt equation," *J. Hydrol. Engrg.*, vol. 8, no. 3, pp. 158–160, 2003.

[73] A. Coppola, A. Basile, A. Comegna, and N. Lamaddalena, "Monte Carlo analysis of field water flow comparing uni- and bimodal effective hydraulic parameters for structured soil," *J. Contam. Hydrol.*, vol. 104, no. 1-4, pp. 153–165, 2009.

[74] H. J. Morel-Seytoux, Ph. D. Meyer, M. Nachabe, J. Tourna, M. T. van Genuchten, and R. J. Lenhard, "Parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics: Preserving the effective capillary drive," *Water Resour. Res.*, vol. 32, no. 5, pp. 1251–1258, 1996.

[75] X.Foussereau, W. Graham, and P.Rao, "Stochastic analysis of transient flow in unsaturated heterogeneous soils," *Water Resour. Res.*, vol. 36, no. 4, pp. 891–910, 2000.

[76] M. J. Lighthill and G. B. Whitham, "On kinematic waves. i. flood movement in long rivers," *Proc. R. Soc. London, Ser. A*, vol. 229, no. 1178, pp. 281–316, 1955.

[77] M. J. Lighthill and G. B. Whitham, "On kinematic waves. ii. a theory of traffic flow on long crowded raods," *Proc. R. Soc. London, Ser. A*, vol. 229, no. 1178, pp. 317–345, 1955.

[78] V. P. Singh, *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*, Wiley: New York, 1996.

[79] V. P. Singh, "Kinematic wave modeling in water resources: a historical perspective," *Hydrol. Process.*, vol. 15, pp. 671–706, 2001.

[80] R. J. Moore, S. J. Cole, V. A. Bell, and D. A. Jones, "Issues in flood forecasting: ungauged basins, extreme floods and uncertainty," in *Frontiers in flood research*, I. Tchiguirinskaia, K. N. N. Thein, and P. Hubert, Eds., vol. IAHS Publ. 305, pp. 103–122. Int. Assoc. Hydrol. Sci., 2006.

[81] P. S. Eagleson, "Dynamics of flood frequency," *Water Resour. Res.*, vol. 8, no. 4, pp. 878–898, 1972.

[82] M. Kavvas and R. S. Govindaraju, "Stochastic overland flows, part 1: Physics-based evolutionary probability distributions," *Stoch. Hydrol. Hydraul.*, vol. 5, no. 2, pp. 89–104, 1991.

[83] R. S. Govindaraju and M. Kavvas, "Stochastic overland flows, part 2: Numerical solutions evolutionary probability density functions," *Stoch. Hydrol. Hydraul.*, vol. 5, no. 2, pp. 105–124, 1991.

[84] M. Kavvas, "Nonlinear hydrologic processes: Conservation equations for determining their means and probability distributions," *J. Hydrol. Engrg.*, vol. 8, no. 2, pp. 44–53, 2003.

[85] Jaeyoung Yoon and M. Levent Kavvas, "Probabilistic solution to stochastic overland flow equation," *J. Hydrol. Engrg.*, vol. 8, no. 2, pp. 54–63, 2003.

[86] Luc Séguis, Bernard Cappelaere, Christophe Peugeot, and Baxter Vieux, "Impact on sahelian runoff of stochastic and elevation-induced spatial distributions of soil parameters," *Hydrol. Process.*, vol. 16, no. 2, pp. 313 – 332, 2002.

[87] G. Lin, C.-H. Su, and G. E. Karniadakis, "Predicting shock dynamics in the presence of uncertainties," *J. Comput. Phys.*, vol. 217, no. 1, pp. 260V276, 2006.

[88] J. Tryoen, O. Le Maître, M. Ndjing, and A. Ern, "Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems," *J. Comp. Phys.*, vol. 229, no. 18, pp. 6485–6511, 2010.

[89] S. B. Pope, *Turbulent Flows*, Cambridge University Press, 2000.

[90] D. M. Tartakovsky and S. Broyda, "Pdf equations for advective-reactive transport in heterogeneous porous media with uncertain properties," *J. Contam. Hydrol.*, vol. 120-121, pp. 129–140, 2011.

[91] J. A. Seddon, "River hydraulics," *Trans. Amer. Soc. Civ. Engrs.*, vol. 43, pp. 179–243, 1900.

[92] D. L. Buhman, T. K. Gates, and C. C. Watson, "Stochastic variability of fluvial hydraulic geometry: Mississippi and Red rivers," *J. Hydr. Engrg.*, vol. 128, pp. 426–437, 2002.

[93] David Gottlieb and Dongbin Xiu, "Galerkin method for wave equations with uncertain coefficients," *Commun. Comput. Phys.*, vol. 3, no. 2, pp. 505–518, 2008.

[94] R. H. Kraichnan, "Eddy viscosity and diffusivity: Exact formulas and approximations," *Complex Systems*, vol. 1, pp. 805–820, 1987.

[95] O. Ditlevsen, G. Mohr, and P. Hoffmeyer, "Integration of non-Gaussian fields," *Probab. Eng. Mech.*, vol. 11, pp. 15–23, 1996.

[96] Wayne Carmichael, *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*, vol. 619, chapter 4: A world overview - One-hundred-twenty-seven years of research on toxic cyanobacteria - Where do we go from here?, pp. 105–125, Srpinger New York, 2008.

[97] W.K. Dodds, W.W. Bouska, J.L. Eitzmann, T.J. Pilger, K.L. Pitts, A.J. Riley, J.T. Schloesser, and D.J. Thornbrugh, "Eutrophication of u.s, freshwaters: analysis of potential economic damages," *Environ. Sci. Technol.*, vol. 43, pp. 12–19, 2009.

[98] Heikki Haario, Leonid Kalachev, , and Marko Laine, "Reduced models for algae growth," *Bulletin of Mathematical Biology*, vol. 71, no. 7, pp. 1626–1648, 2007.

[99] Dong-Wei Huang, Hong-Li Wang, Jian-Feng Feng, and Zhi-Wen Zhu, "Modelling algal densities in harmful algal blooms (hab) with stochastic dynamics," *Applied Mathematical Modelling*, vol. 32, no. 7, pp. 1318–1326, 2008.

[100] H.K. Hudnell, C. Jones, B. Labisi, V. Lucero, D. R. Hill, and J. Eilers, "Freshwater harmful algal bloom (fhab) suppression with solar powered circulation (spc)," *Harmful Algae*, vol. 9, no. 2, pp. 208–217, 2010.

[101] Marko Laine, *Adaptive MCMC methods with applications in environmental and geophysical models*, Ph.D. thesis, Lappeenranta university of tehcnology, March 2008.

[102] H. Risken, *The Fokker-Planck Equation: Methods of Solutions and Applications*, Springer, 2nd edition, 1989.