Cairo University

THE USE OF CONNECTIVITY CLUSTERS AND PERCOLATION CONCEPTS IN STOCHASTIC MODELING OF HIGHLY HETEROGENEOUS POROUS MEDIA

By

Mohamed Abdullah Awad Taha

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

In

IRRIGATION AND HYDRAULICS ENGINEERING

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2018

THE USE OF CONNECTIVITY CLUSTERS AND PERCOLATION CONCEPTS IN STOCHASTIC MODELING OF HIGHLY HETEROGENEOUS POROUS MEDIA

By

Mohamed Abdullah Awad Taha

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

In

IRRIGATION AND HYDRAULICS ENGINEERING

Under the Supervision of

Prof. Ahmed Emam Ahmed Hassan	**Prof. Hesham Bekhit Mohamed**
……………………..	…………………………
Professor of Hydrogeology Irrigation and Hydraulics Department Faculty of Engineering, Cairo University	Professor of Water Resources Irrigation and Hydraulics Department Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2018

THE USE OF CONNECTIVITY CLUSTERS AND PERCOLATION CONCEPTS IN STOCHASTIC MODELING OF HIGHLY HETEROGENEOUS POROUS MEDIA

By

Mohamed Abdullah Awad Taha

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

In

IRRIGATION AND HYDRAULICS ENGINEERING

Approved by the Examining Committee

____

Prof. **Ahmed Emam Ahmed Hassan**, Thesis Main Advisor
Professor of Hydrogeology
Faculty of Engineering-Cairo University

____

Prof. **Hesham Bekhit Mohamed**, Advisor
Professor of Water Resources
Faculty of Engineering-Cairo University

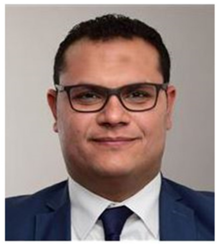
____

Prof. **Abdulwahab Mohamed Amer**, Internal Examiner
Professor of Hydraulics
Faculty of Engineering-Cairo University

____

Prof. **Ahmed Ali Hassan**, External Examiner
Professor of Hydrogeology
Faculty of Engineering-Ain Shams University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2018



Engineer's Name: Mohamed Abdullah Awad Taha
Date of Birth: 28/09/1983
Nationality: Egyptian
E-mail: Mohamed.Awad@dar.com
Phone: +201068059097
Address: 124 K Hadayak Al Ahram
Registration Date: 01/03/2012
Awarding Date: 2018
Degree: Doctor of Philosophy
Department: Irrigation and Hydraulics Engineering

Supervisors:

Prof. Ahmed Emam Hassan
Prof. Hesham Bekhit Mohamed

Examiners:

Prof. Ahmed Ali Hassan (External Examiner)
Professor of Hydrogeology
Faculty of Engineering-Ain Shams University
Prof. Abdulwahab Mohamed Amer (Internal Examiner)
Prof. Ahmed Emam Ahmed Hassan (Thesis Main Advisor)
Prof. Hesham Bekhit Mohamed (Advisor)

Title of Thesis:

The Use of Connectivity Clusters and Percolation Concepts in Stochastic Modeling of Highly Heterogeneous Porous Media

Key Words:
Stochastic; Monte Carlo; percolation; connectivity; clustering

Summary:
The main objective of this research is to present a sampling framework to reduce the number of the required Monte Carlo realizations utilizing the connectivity properties of the hydraulic conductivity distributions in three dimensional domains. This objective is achieved through studying the influence of the connectivity and preferential flow paths on transport modeling in highly heterogonous media and testing different sampling techniques to select a compact yet representative sample. Applying different sampling techniques together with several indicators suggested that a compact sample representing only 10% of the total number of realizations can be used to produce results close to the results of the full set of realizations. In addition to that, Artificial Neural Network (ANN) is utilized to predict the transport outputs using the clusters connectivity indicators. The results are in good match with the traditional transport modelling using MT3DMS. The developed sampling frame work is applied to the famous MADE-2 site experiment. The selected compact sample succeeded to match the observed plume concentrations along the simulated domain.

Acknowledgments

The author wants to express a sincere acknowledgement to his advisors, Prof. Dr. Ahmed Emam Hassan and Prof. Dr. Hesham Bekhit Mohamed at the Irrigation and Hydraulics Department, Faculty of Engineering, Cairo University for their motivation, time, materials, precious remarks and endless support throughout the course of this research. Moreover, the author highly appreciates the contributions of Prof. Dr. Abdelwahab Amer and Prof. Dr. Ahmed Aly Hassan. The author is deeply grateful to Dar's staff and management for their guidance, time, materials and support. Finally, the author deeply thanks his family, beloved mother and his beloved wife and children Nour and Yehia for their support, motivation, and love.

Table of Contents

List of Tables

List of Figures

Nomenclature

ADE	Advection Dispersion Equation
ANNs	Artificial Neural Networks
BTCs	Breakthrough Curves
CP	Control Plane
CPU	Central Processing Unit
GMEAN	Geometric Mean
GSLIB	Geostatistical Software Library
K	Hydraulic Conductivity
LHS	Latin Hypercube Sampling
MADE	Macrodispersion Experiment Site in Mississippi
NF	Net Fractions
RGF	Random Field Generator
SGSIM	Sequential Gaussian Simulation
TCF	Total Connected Fractions
TOR	Tortuosity

Abstract

Stochastic geostatistical techniques are essential tools for groundwater flow and transport modeling in highly heterogeneous media. Typically, these techniques require extremely large numbers of realizations to accurately simulate the high variability and account for uncertainty. These large numbers of realizations impose several constraints on the stochastic techniques (e.g., increasing the computational effort, model size, grid spacing, and time step, and stability issues).

Understanding connectivity of subsurface layers gives an opportunity to overcome these constraints. This research presents a sampling framework to reduce number of required Monte Carlo realizations utilizing connectivity properties of the hydraulic conductivity distributions in three dimensional domains.

Different geostatistical distributions are tested in this study including exponential and spherical distributions. It was found that the total connected fraction of the largest clusters and their tortuosity are highly correlated with the percentage of mass arrival and first arrival quantiles at different control planes. Applying different sampling techniques together with several indicators suggested that a compact sample representing only 10% of the total number of required realizations can be used to produce results that are close to the results of the full set of realizations.

Also, the proposed sampling techniques specially utilizing low conductivity clustering show very promising results in terms of matching the full range of realizations. In addition, the size of selected clusters relative to domain size significantly affects transport characteristics and connectivity indicators.

Artificial Neural Networks (ANNs) are utilized to predict transport outputs using clusters connectivity indicators. The results are compared with the traditional transport modelling using MT3DMS and good matching was observed. The developed sampling framework is applied to the famous MADE-2 site experiment which is characterized by highly heterogeneous subsurface conditions. The selected compact sample succeeded to match the observed plume concentrations along with the simulated domain.

Chapter 1 : Introduction

1.1. General

Modeling flow and transport in highly heterogeneous media has been an active field of research for several decades (e.g. [1-5]). Accurate flow and transport modeling in highly heterogeneous aquifers is a complicated task due to the uncertainty in subsurface geology as well as difficulty to select an appropriate modeling approach. Simulation of contaminant transport in heterogeneous media is a very challenging task. It is important to understand the relation between the spatial distribution of a contaminant migrating in groundwater and variation of subsurface geologic conditions and its associated hydraulic properties. This understanding is a prerequisite for appropriate choice of a modeling approach that can accurately describe the movement of contaminants in highly heterogeneous aquifers.

The characteristics of geologic media cannot be determined with certainty. Some types of uncertainty can be reduced (e.g., via collecting more data) while others cannot be reduced (such as uncertainty stemming from spatial variability). Even if there are unlimited financial and human resources to characterize geologic media, there is a limit to characterization effort and number of exploratory boreholes that can be drilled at a site. Therefore, considering the uncertainty in both conceptualization of the model as well as parametric uncertainty of the hydraulic properties mandates the use of stochastic modeling approaches (e.g. [4, 6, 7]).

This irreducible uncertainty motivated many researchers to cast groundwater problems in a stochastic framework. Implementation of the stochastic and numerical simulations within a Monte Carlo framework has thus become common practice in the past two decades.

Stochastic geostatistical techniques are usually used to generate alternative fine-scale three-dimensional realizations of subsurface parameters that are consistent with the available data [8]. Assessment of aquifer response uncertainty is provided by processing a large number of fine-scale realizations through groundwater modeling programs. The issue of the number of realizations needed to achieve convergence for the statistics of concern becomes crucial when medium heterogeneity increases. Studies have shown that second order moments (variances) for contaminant concentration or mass flux require a much larger number of realizations to converge as compared to the first order moments (mean or expected values), even for relatively moderate heterogeneity [9]. This requirement, in addition to numerical constraints such as domain size, grid resolution, time step, and convergence issues, increases the computational effort involved in these simulations. Therefore, the application of Monte Carlo techniques become prohibitive from both time and cost perspectives, even with today's advanced computing resources.

Three dimensional studies usually employ a small number of realizations ($\approx$1000 - 2000) (e.g., [10-14]), which strongly influences accuracy and convergence of the computed statistics. Even when computational resources allow for few thousand realizations of *3-D* models, higher order moments and other outputs of concern may not be produced accurately. Computationally efficient numerical methods are, therefore, needed to simulate groundwater flow and transport processes in highly heterogeneous geologic media, while providing accurate estimates for the expected values, uncertainties, and higher order moments of the outputs of concern.

In the past fifteen years the rule of connectivity in highly heterogeneous aquifers has been addressed along with its effect on the selected modeling techniques [15-21]. It has been argued that the presence of network of connected highly permeable lenses will allow a portion of the plume to spread quickly. In the same time, the rest of the plume will remain concentrated around the source and result in long tail similar to the case of transport in highly heterogeneous aquifers. Study of subsurface geologic formation connectivity has also been addressed in the context of the percolation theory.

Deutsch [22] developed some indicators to quantify and assess connectivity in highly heterogeneous aquifers. He introduced three indicators based on measuring the number and the size of connected bodies in a *3-D* Cartesian grid. More recently, Knudby and Carrera [17] proposed and evaluated nine different indicators of connectivity in order to assess the possibility of predicting flow connectivity from statistical connectivity and, consequently, transport connectivity from flow connectivity [21].

The rule of connectivity indicators and its relation to flow and transport modeling were evaluated by Zinn and Harvey [15], Binachi et al. [21] and Tyukhova [23]. They concluded that distribution of preferential flow paths and contaminant exit locations are clearly influenced by the presence of the connected zones of high hydraulic conductivity.

1.2. Problem Statement

Flow and transport modeling in highly heterogeneous media is constantly faced with ambiguity and variability which always lead to uncertainty. Monte Carlo simulations provide a mean for examining all possible results and allowing for better decision making under uncertainty. Monte Carlo simulations relay on building models of all possible outcomes using a range of values for a certain parameter(s) that has inherent uncertainty. This technique runs the models and obtains results for a large number of realizations each time using a different set of random values from the probability functions for the parameter of concern. However, Monte Carlo simulation requires thousands of realizations to converge to a reliable solution based on degree of variability and the data limitation. These numerous number of Monte Carlo realizations require huge computational efforts and resources to accurately capture the uncertainty in low and transport results.

Connectivity characteristics of hydraulic conductivity may play a significant role in enhancing the efficiency of Monte Carlo simulations and reducing computational efforts. Traditionally, Monte Carlo approach implies developing thousands of