# MODIFIED FINITE DIFFERENCE METHOD USING RANDOM SAMPLING FOR NONLINEAR EPIDEMIC MODELS

**MAHA ABDULJABBAR MOHAMMED**

**FACULTY OF SCIENCE**
**UNIVERSITY OF MALAYA**
**KUALA LUMPUR**

**2018**

# MODIFIED FINITE DIFFERENCE METHOD USING RANDOM SAMPLING FOR NONLINEAR EPIDEMIC MODELS

**MAHA ABDULJABBAR MOHAMMED**

**THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**INSTITUTE OF MATHEMATICAL SCIENCES**
**FACULTY OF SCIENCE**
**UNIVERSITY OF MALAYA**
**KUALA LUMPUR**

**2018**

**UNIVERSITY OF MALAYA**

**ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: **Maha Abduljabbar Mohammed**

Matric No: **SHB130006**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Project Paper/Research Report/Dissertation/Thesis:

("**MODIFIED FINITE DIFFERENCE METHOD USING RANDOM SAMPLING FOR NONLINEAR EPIDEMIC MODELS**"):

Field of Study: **Numerical Simulation for Ordinary Differential Equations**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation

**UNIVERSITI MALAYA**

**PERAKUAN KEASLIAN PENULISAN**

Nama: Maha Abduljabbar Mohammed

No. Matrik: **SHB130006**

Nama Ijazah: **DOKTOR FALSAFAH**

Tajuk Kertas Projek/Laporan Penyelidikan/Disertasi/Tesis:

("**KAEDAH BEZA TERHINGGA TERUBAHSUAI MENGGUNAKAN PENSAMPELAN RAWAK BAGI MODEL EPIDEMIK TAK LINEAR**")

Bidang Penyelidikan: **Simulasi Berangka untuk Persamaan Pembezaan Biasa**

Saya dengan sesungguhnya dan sebenarnya mengaku bahawa:

(1) Saya adalah satu-satunya pengarang/penulis Hasil Kerja ini;
(2) Hasil Kerja ini adalah asli;
(3) Apa-apa penggunaan mana-mana hasil kerja yang mengandungi hakcipta telah dilakukan secara urusan yang wajar dan bagi maksud yang dibenarkan dan apa-apa petikan, ekstrak, rujukan atau pengeluaran semula daripada atau kepada mana-mana hasil kerja yang mengandungi hakcipta telah dinyatakan dengan sejelasnya dan secukupnya dan satu pengiktirafan tajuk hasil kerja tersebut dan pengarang/penulisnya telah dilakukan di dalam Hasil Kerja ini;
(4) Saya tidak mempunyai apa-apa pengetahuan sebenar atau patut semunasabahnya tahu bahawa penghasilan Hasil Kerja ini melanggar suatu hakcipta hasil kerja yang lain;
(5) Saya dengan ini menyerahkan kesemua dan tiap-tiap hak yang terkandung di dalam hakcipta Hasil Kerja ini kepada Universiti Malaya ("UM") yang seterusnya mula dari sekarang adalah tuan punya kepada hakcipta di dalam Hasil Kerja ini dan apa-apa pengeluaran semula atau penggunaan dalam apa jua bentuk atau dengan apa juga cara sekalipun adalah dilarang tanpa terlebih dahulu mendapat kebenaran bertulis dari UM;
(6) Saya sedar sepenuhnya sekiranya dalam masa penghasilan Hasil Kerja ini saya telah melanggar suatu hakcipta hasil kerja yang lain sama ada dengan niat atau sebaliknya, saya boleh dikenakan tindakan undang-undang atau apa-apa tindakan lain sebagaimana yang diputuskan oleh UM.

Tandatangan Calon Tarikh:

Diperbuat dan sesungguhnya diakui di hadapan,

Tandatangan Saksi Tarikh:

Nama:

Jawatan:

**MODIFIED FINITE DIFFERENCE METHOD USING RANDOM SAMPLING FOR NONLINEAR EPIDEMIC MODELS**

## ABSTRACT

In this thesis, new modified numerical simulation processes are proposed to solve social epidemic models in the form of nonlinear initial value problems (IVP) of ordinary differential equations with multiple random variable parameters. The variables of the systems are dependent on time $t$. The utilization of Monte Carlo (MC) simulation with central divided difference formula is repeated $n$ times to simulate values of the variable parameters of the Spain weight reduction model as random sampling instead being limited as real values with respect to time. The mean of the $n$ final solutions via this integrated technique, named in short as Mean Monte Carlo Finite Difference (MMCFD) method, represents the final solution of the system. The numerical outputs are tabulated, graphed and compared with previous statistical estimations for 2013, 2015 and 2030 respectively. The solutions of FD and MMCFD are found to be in good agreement with small standard deviation of mean and small measure of difference. In the social epidemic of cocaine abuse in Spain, the FD numerical method is integrated with Latin hypercube sampling (LHS) technique in every simulation to simulate random variable parameters for the stochastic-deterministic model. The mean of final solutions of the FD iterations is known as Mean Latin Hypercube Finite Difference (MLHFD) solutions. The results obtained are compared with deterministic solutions of classical FD and homotopy analysis methods as relative to the previous statistical estimations from 1995 to 2015. Good agreement between the two is perceived with small errors. The MLHFD results are tabulated, graphed and discussed pertaining to the model's expected behavior until 2045. MMCFD and MLHFD are proposed for the first time in this thesis to calculate and to predict future behavior of the epidemic models considered. The results show the range for random distribution for the present numerical solutions obtained are in good agreement and approximation as compared to the existing randomized statistical estimations.

**Keywords**

Ordinary differential equation, Finite difference method, Random distribution, Random sampling, Monte Carlo method, Latin hypercube sampling, Prediction interval, Epidemiology.

**KAEDAH BEZA TERHINGGA TERUBAHSUAI MENGGUNAKAN PENSAMPELAN RAWAK BAGI MODEL EPIDEMIK TAK LINEAR**

## ABSTRAK

Dalam tesis ini, proses simulasi berangka terubahsuai yang baharu telah dicadangkan untuk menyelesaikan model wabak sosial dalam bentuk masalah nilai awal (MNA) tak linear bagi persamaan pembezaan biasa dengan parameter pemboleh ubah rawak berganda. Pemboleh ubah sistem bersandar kepada masa *t*. Penggunaan simulasi Monte Carlo (MC) dengan formula terbahagi berpusat diulang sebanyak *n* kali untuk mensimulasikan nilai parameter pemboleh ubah model pengurangan berat badan Spain sebagai pensampelan rawak dan bukan hanya terhad sebagai nilai sebenar bersandar kepada masa. Purata bagi *n* penyelesaian akhir melalui teknik bersepadu ini dinamakan sebagai kaedah Min Monte Carlo Beza Terhingga (MMCBT) yang mewakili penyelesaian akhir bagi sistem. Keputusan berangka dijadualkan, digrafkan dan dibandingkan dengan anggaran statistik terdahulu pada tahun 2013, 2015 dan 2030. Penyelesaian BT dan MMCBT didapati konsisten dengan purata sisihan piawai kecil dan ukuran perbezaan kecil. Dalam wabak sosial penyalahgunaan kokain di Spain, kaedah berangka BT disepadukan dengan teknik pensampelan hiperkubus Latin (PHL) dalam setiap simulasi untuk mensimulasikan parameter pemboleh ubah rawak bagi model stokastik-deterministik tersebut. Purata penyelesaian akhir bagi lelaran BT dikenali sebagai penyelesaian Min Latin Hiperkubus Beza Terhingga (MLHBT). Keputusan yang diperoleh telah dibandingkan dengan penyelesaian kaedah deterministik BT klasik dan analisis homotopi sejajar dengan anggaran statistik terdahulu iaitu dari tahun 1995 hingga 2015. Keputusan di antara keduanya didapati amat jitu dengan ralat kecil. Hasil keputusan MLHBT dijadualkan, digrafkan dan dibincangkan berhubung dengan jangkaan tingkah laku model sehingga 2045. MMCBT dan MLHBT telah dicadangkan buat pertama kalinya di dalam tesis ini untuk menghitung dan meramal tingkah laku masa depan bagi model-model wabak yang dipertimbangkan. Keputusan menunjukkan bahawa julat taburan rawak untuk penyelesaian berangka terkini yang diperolehi adalah konsisten dan mempunyai penghampiran yang baik jika dibandingkan dengan anggaran statistik rawak sedia ada.

**Kata kunci**

Persamaan pembezaan biasa,  Kaedah beza terhingga, Taburan rawak, Pensampelan rawak, Kaedah Monte Carlo, Pensampelan hiperkubus Latin, Selang ramalan, Epidemiologi.

**ACKNOWLEDGEMENTS**

All words cannot describe my infinite gratitude to Allah S.W.T who helped me in each step of the doctoral work and writing of this thesis during the years of my Ph.D. study. I can never complete this thesis without His assistance.

My greatest appreciation to my respected supervisor Dr. Noor Fadiya Mohd Noor who always gave me assistance and guidance to do the best research work. I owe it to her a deeply thank and I will never forget what she did to help me to complete this thesis from the beginning of my study until the last step of this work. I would like to give my greatest thank to my co. supervisors Dr. Zailan Siri and Dr. Adriana Irawati Nur Ibrahim. I really appreciate their guidance to upgrade this research and they have helped me to complete this study. I thank all my supervisors again; really I cannot complete my study without all the helps from my supervisors. I cannot give them what they deserve whatever I did.

I would like to give a sincere gratitude to my parents who always support me to face all the hard problems in my life during my study interval, especially when I was outside of my country. My mother and my father gave me the encouragement and they advised me to consider the study the most important thing in my life. I dedicate my work reward for this thesis to the spirit of my father Abduljabbar Mohammed (may Allah have mercy on him).

I would like to express my greatest gratitude to all my family members, especially my second mother, my sisters, my brothers, my uncles and my aunts. They always provided me their continuous encouragement to be patient in the hard life and study.

I am indebted to my father Abduljabbar, my lecturer Prof. Abbas Najim Salman and my uncle Mr. Nabeel. They have supported me financially through my study as well as they have supported me by their continuous encouragement.

My particular appreciation goes to my lecturer Prof. Abbas Najim Salman who provided me assistance, guidance and encouragement to complete the doctoral study. I will never forget what he did to help me during my study.

Special acknowledgement to the Government of Iraq who provided me with the education scholarship. I am really grateful to Institute of Mathematical Sciences, Faculty of Science, University of Malaya (UM) specially, Madame Budiyah Binti Yeop, Madame Fadila and Ms. Najwa. Malaysia that allowed me to complete my study among their students. Greatest thanks to all lecturers from the Mathematical department, Ibn-Al-Haitham College of Education, University of Baghdad, who gave me the encouragement to finish this study, especially assistance given by Prof. Mohammed Wahdan and Prof. Dr. Luma Naji. Special thanks to all my Ph.D. friends who had taught me what can benefit me in my study; Waleeda Suidan, Tammar Hussain, Lubna Ghalib, Nadia Faiq and to my dear friends Niran Sabah and Alaa Majed.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF SYMBOLS AND ABBREVIATIONS

| | |
|---|---|
| BT | Finite difference method in the Malay language |
| CDF | Cumulative probability distribution |
| CI | Confidence interval |
| CM-LHS | Combined Latin hypercube sampling method |
| DE | Differential equation |
| DEs | Differential equations |
| DLHS | Discrete Latin hypercube sampling method |
| Eq. | Equation |
| Eq.s | Equations |
| FD | Finite difference method |
| FE | Finite element method |
| FV | Finite volume method |
| HAM | Homotopy analysis method |
| HAM-Pade | Homotopy-Pade technique |
| IVP | Initial value problem |
| LHS | Latin hypercube sampling method |
| LNR | Lymph nodes removed |
| MC | Monte Carlo simulation technique |
| MLHBT | Mean Latin Hypercube Finite Difference method in the Malay language |
| MLHFD | Mean Latin Hypercube Finite Difference method |
| MMCBT | Mean Monte Carlo Finite Difference method in the Malay language |
| MMCFD | Mean Monte Carlo Finite Difference method |
| MNA | Initial value problem in the Malay language |
| NFD | Nonstandard Finite difference method |
| ODE | Ordinary differential equation |
| ODEs | Ordinary differential equations |
| PDE | Partial differential equations |
| PHL | Latin hypercube sampling method in the Malay language |
| RAND | Generator function for random variables in MATLAB |
| RK | Runge-Kutta method |
| RK2 | Runge-Kutta method of order 2 |
| RK4 | Runge-Kutta method of order 4 |
| RK45 | Runge-Kutta method of order 4 and 5 |
| RK78 | Runge-Kutta method of order 7 and 8 |
| RV | Random variable |

| | |
|---|---|
| SEIAR | Susceptible-Exposed-Infectious-Infectious without symptoms-Recovered model type |
| SEIR | Susceptible-Exposed-Infectious-Recovered model type |
| SIR | Susceptible-Infectious-Recovered model type |
| SIS | Susceptible-Infectious-Susceptible model type |
| SIVRS | Susceptible-Infectious-Variant-Recovered-Susceptible-model type |
| $A$ | Infectious without symptoms population |
| $\frac{dy}{dt}$ | Measure of difference |
| $E$ | Exposed population |
| $\|E\|$ | Difference error in Chapter 4 |
| $\|E_a\|$ | Absolute approximate error in Chapter 3 |
| $F$ | Cumulative probability distribution |
| $F^{-1}$ | The inverse of cumulative probability distribution |
| $f$ | Function |
| $h$ | Step size in Chapter 3 |
| $h_0$ | Initial value of habitual users of cocaine abuse model |
| $h(t)$ | Habitual users of cocaine abuse model |
| $I$ | Infectious population |
| $l$ | counter for variables in a model |
| $i$ | counter for iterations of a numerical method |
| $j$ | counter for simulations of a random sampling |
| $k$ | The maximum number of variables in a model |
| $m$ | The maximum number of iterations for a numerical method |
| $n$ | The maximum number of simulations of a random sampling |
| $n_0$ | Initial value of non-users of cocaine abuse model |
| $N(t)$ | Subpopulation of individuals with normal weight |
| $n(t)$ | Non-users of cocaine abuse model |
| $o_0$ | Initial value of occasional users of cocaine abuse model |
| $O(h)$ | Truncation error |
| $O(h^m)$ | Truncation error of order $m$ |
| $O(1/n)$ | LHS's sampling error |
| $O(1/\sqrt{n})$ | Monte Carlo's sampling error |
| $O(t)$ | Subpopulation of individuals with obesity |
| $o(t)$ | Occasional users of cocaine abuse model |

| | |
|---|---|
| $p$ | Average overweight or obese individual needs to reduce a mean of 7 kg weight to transit to normal weight or overweight in reduce weight model |
| $p^{th}$ | The percentile value of $th\%$ |
| $R$ | Recovered population |
| $\mathbb{R}$ | Real space |
| $R_0$ | Basic reproduction number |
| $r_0$ | Initial value of regular users of cocaine abuse model |
| $r(t)$ | Regular users of cocaine abuse model |
| $1/p$ | Average time an individual needs to return to S(t) from O(t) or N(t) from S(t) by diet and physical activity of reduce weight model |
| $S$ | Susceptible population |
| $S(t)$ | Subpopulation of individuals with overweight |
| $sol_{i,j}(t)$ | Monte Carlo finite difference of a system for $i$ numerical iteration and for $j$ simulation |
| $sol_{m,j}(t)$ | Monte Carlo finite difference final solutions of the last iteration $m$ and $n$ simulations |
| $sol_{m,j}(t)_l$ | Monte Carlo finite difference final solutions of the last iteration $m$ for a variable $l$ in a model with $n$ simulations |
| $\mu_{sol}$ | MMCFD solution |
| $\mu_{sol,l}$ | Mean of Monte Carlo finite difference final solutions for each variable $l$ |
| $t$ | Time |
| $t_0$ | Initial of time |
| $V$ | Variant |
| $x_{sol}$ | Deterministic or numerical simulation value |
| $w$ | Step size in Chapter 4 |
| $y$ | A variable of a system |
| $y'$ | Derivative of a system variable y |
| $Z$ | Random variable |
| $\Delta t$ | Time change |
| $\beta_1$ | Transmission rate that depends on unhealthy lifestyles due to social pressure in Chapter 3 |
| $\beta_2$ | Transmission rate of cocaine abuse due to social pressure in cocaine abuse model in Chapter 4 |
| $\gamma_1$ | Rate at which 24-65 years old overweight adults become obese due to their unhealthy lifestyles of reduce weight model in Chapter 3 |

| | |
|---|---|
| $\gamma_2$ | The rate at which an occasional user becomes a regular user in cocaine abuse model in Chapter 4 |
| $\delta$ | Value in an interval |
| $\varepsilon$ | Rate at which obese individual become overweight in reduce weight model in Chapter 3 |
| $\varepsilon_1$ | 0.4% proportion of obese population who do physical activity in order to reduce weight in reduce weight model |
| $\varepsilon_2$ | 2.4% proportion of obese population who improve diet to reduce weight |
| $\varepsilon_c$ | The rate at which a habitual user leaves cocaine due to therapy course in cocaine abuse model in Chapter 4 |
| $\mu_1$ | Average stay time in system of 24-65 years old adults in reduce weight model in Chapter 3 |
| $\mu_2$ | Birth /death rate of cocaine abuse model in Chapter 4 |
| $\rho$ | Rate that overweight individuals turn into normal weight in reduce weight model |
| $\rho_1$ | 2% proportion of overweight population who do physical activity to reduce weight in reduce weight model |
| $\rho_2$ | 4.2% proportion of overweight population who improve diet to reduce weight in reduce weight model |
| $\sigma$ | Standard deviation |
| $\sigma_c$ | The rate at which a regular user becomes a habitual user of cocaine abuse model in Chapter 4 |
| $\sigma_n$ | Standard deviation of MC |

## LIST OF APPENDICES

# CHAPTER 1: INTRODUCTION

## 1.1 Introduction

Ordinary differential equation (ODE) is an equation for unknown functions of dependent variable and its derivatives. If the independent variable of ODE is in terms of a variety of time and does not appear explicitly, then the system is called an autonomous system or sometimes, a time-invariant system. The highest derivative is named the order of the differential equation (Teschl, 2012). Nonlinear autonomous systems of first-order ordinary differential equations for initial value problems are considered in this thesis for the public campaign effects on the weight loss and the cocaine abuse model in Spain. Other applications of the autonomous system are abundant in the behavioral studies of dynamical systems (Hirsch et al., 2012).

A mathematical model is a description of a natural phenomenon, either stochastic model or deterministic model. A stochastic model provides various outcomes which include one or more random variables solved by a stochastic method such as Monte Carlo simulation (Taylor & Karlin, 2014; Wilkinson, 2012; Ullah & Wolkenhauer, 2011). On the other hand, a deterministic model does not has a random variable, thus the solutions obtained are only unique for a given specific time. These models are solved by deterministic methods such as Runge-Kutta, finite-difference, finite volume, finite-element etc. (Lambert, 1991). The third kind of method is hybrid stochastic/ deterministic modelling approach which represents a stochastic-deterministic model. Stochastic-deterministic modelling is a mixed of deterministic and stochastic terms shown in one system as in Chemical Master Equation (CME) application (Mazza & Benaim, 2014; Menz, 2013). In the present thesis, the stochastic-deterministic models are considered to predict the overall behavior of the real epidemic models. The epidemic models are the wide application of nonlinear autonomous stochastic-deterministic models (Allman & Rhodes, 2004; Diekmann & Heesterbeek, 2008; Hethcote, 2000) which considered in this study.

Most epidemic models can be represented by systems of ordinary differential equations depending on independent time $t$, where some of these systems are solved practically by using simulation approaches for randomness distribution of the parameters involved in such models (Brauer & Castillo-Chavez, 2001). Similarly, two modified approaches are firstly proposed in the present thesis namely, Mean Monte Carlo Finite Difference (MMCFD) and Mean Latin Hypercube Finite Difference (MLHFD) methods. These modified methods are the integration products between statistical simulation processes for random sampling with classical numerical iteration approach of finite difference (FD) formula. We can explain these concepts in a simple way, starting with a random variable which is a variable that carries a value determined by a chance event. In other words, it is a possible value of numerical outcomes of a random phenomenon (Tucker, 1998). Hence random sampling is a method of selecting a sample randomly from a statistical population in order to study the entire population when the random sample represents the population (Perros, 2009). There are many kinds of random sampling and the most common processes are random, systematic, stratified and cluster samplings.

In the current study, Monte Carlo (MC) simulation technique and Latin Hypercube Sampling (LHS) procedure are used to simulate parameters of the selected stochastic-deterministic models for random sampling distribution. Monte Carlo (MC) is a form of sampling that refers to the traditional technique for using random numbers to sample from a probability distribution. It is considered as a simple type of random sampling (Dagpunar, 2007). The second random sampling method is Latin hypercube sampling (LHS) which is a statistical method for generating a random sample of parameter values from a multidimensional distribution. It is considered as an extension of stratified sampling that divides a set or a population into several strata where the selection of a random sample is done from each stratum. LHS is faster and saves effort with the time (De Veaux et al., 2012).

On the other hand, finite difference (FD) method is one an established approximate method using to solve the differential equations numerically. In the present thesis, the first-order autonomous initial value problems representing nonlinear systems of ordinary differential equations (ODEs) are solved numerically by integrating the FD method with the statistical simulation procedures of Monte Carlo (MC) and Latin hypercube sampling (LHS) to obtain random distributed solutions for the deterministic models with random parameters.

## 1.2 Problem Statement

Individual bad habit can be transmitted as a social epidemic within a big society due to consistent social contact and pressure. Such epidemics can be worsened without appropriate precautions and remedies. In order to control the spread of these epidemics, researchers embarked to represent them as mathematical models. These models can be written in the form of nonlinear systems of ordinary differential equations. The parameters of such models have random distribution in nature. Therefore the classical methods may not be always appropriate to solve these systems for future estimation due to insufficient data of deterministic approach. The importance of this study comes from the fact that some real models of nonlinear systems of ODEs consist of random variables that require solutions pertaining to randomness property. Since parameter estimation in random sampling distribution is considered, modified statistical-numerical methods can be explored especially when the high-dimensional parameter space is taken into consideration. On the other hand, these modified methods support the prediction interval for the solutions obtained whereas such concept is not appropriately accepted in conventional numerical deterministic approaches.

The statistical simulation technique itself may be appropriate for some reasons; through statistical simulation, we can get better understanding from detailed observation of the system and, to analyze the phenomenal changes and the effects of information under study. Sometimes, simulations can design an experiment for the complicated system or a new

system. The simulations can also be used to analyze a dynamic system with their real time (Rubinstein, 1981, p. 9-10 ).

## 1.3 Research Objectives

This research embarks to fulfill the following objectives:

- To introduce modified methods which combined statistical simulation techniques with an iterative numerical method that is finite difference method.
- To apply the new modified methods for solving selected social epidemic models in the form of nonlinear systems of ordinary differential equations with random sampling distribution of the parameters.
- To compare the numerical simulation results obtained from the new modified methods with possible and existing deterministic and non-deterministic solutions.
- To analyze the numerical simulation results obtained graphically and tabularly towards the solutions of the epidemic models.
- To estimate prediction intervals for random distribution of the numerical simulation solutions obtained.

## 1.4 Scope of Research

Social epidemic models with random parameters are considered in our study. These models are normally treated as deterministic problems with a probability process that can be programmed in computers in order to save cost and time (Rubinstein, 1981). Random sampling process is used to predict unknown parameters in such models. Two existing social epidemic models in Spain are selected from literatures in order to demonstrate the applications of the newly proposed modified finite difference methods; weight loss due to public campaign for Mean Monte Carlo Finite Difference (MMCFD) method and cocaine abuse for Mean Latin Hypercube Finite Difference (MLHFD) method. Further stability

analysis on the epidemic models considered in this thesis is beyond the scope of the present study.

## 1.5 Outline of Thesis

This research embarks on finding the alternative modified methods of statistical-numerical approaches in order to supply numerical simulation solutions for some real deterministic nonlinear epidemic systems as well as to give prediction ranges of these solutions. This thesis is divided into four parts; introduction of the research, literature review and modified numerical simulation processes, analysis and solutions of selected social epidemic models using the modified numerical simulation processes and finally, the overall conclusion and suggestion of the thesis.

In Chapter 1, the preliminaries of this research are outlined briefly in the subsections of introduction, research objectives, problem statement and scope of research. Chapter 2 provides a brief literature review ideas and concepts of ordinary differential equations, stochastic-deterministic models and their applications, epidemic models, numerical iteration methods with special emphasis on finite difference (FD) method, statistical simulation techniques of Monte Carlo (MC), Latin hypercube sampling (LHS) and the modified numerical simulation processes.

Next in Chapter 3, a new modified approach between Monte Carlo simulation and finite difference method, namely Mean Monte Carlo Finite Difference (MMCFD) method explained in Chapter 2, is applied to solve weight reduction model due to public campaign in Spain. The MMCFD results obtained are compared with present FD solutions and the existing statistical estimation from the literature. On the other hand, another modified approach between Latin hypercube sampling (LHS) simulation and finite difference (FD) method, namely Mean Latin Hypercube Finite Difference (MLHFD) method newly proposed and explained in Chapter 2, is applied in Chapter 4 to solve the cocaine abuse

model in Spain. The MLHFD results obtained are compared with present FD solutions and the results of HAM-Pade and statistical estimation by other researchers.

Finally, the overall findings and conclusion of the research are provided in Chapter 5. Recommendations to improve and to expand the research concerns for future works on this subject are also suggested in this last chapter.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Introduction

"Differential equation (DE) is an equation that describes how a state $y(t)$ changes." (Logan, 2006, p. 2). The order of DE is the highest derivative in this equation (Logan, 2006). DE has an important role in discrete mathematics that characterized various time $t$ continuous over a particular interval. The time change is denoted by $\Delta t$ and time $t$ in the current study is an independent variable. This time may happen in the past $t - \Delta t$ (delay time), the current time $t$ (present time) or the future time $t + \Delta t$ (Giordano et al., 2003). "An ordinary differential equation (ODE) is an equation that involves one or more derivatives of an unknown function. A solution of the ODE is a specific function that satisfied the equation" (Cheney & Kincaid, 1999, p. 370).

The general form of an ordinary differential equation (ODE) with initial value can be written as follows:

$$y' = f(t, y), \qquad a \le t \le b, \qquad y(a) = a_0, \tag{2.1}$$

where $y$ is a scalar function in $d$-dimensional space of $\mathbb{R}$, $t$ is time while $a$, $b$ and $a_0$ are positive real constant values, (Lambert, 1991). Each system variable depends on independent time $t$ and the parameters of the system are considered as random parameters depending on time $t$ in the current study.

The focus of our study is on the initial value problem (IVP) for nonlinear autonomous system of first order differential equations. Autonomous system means the independent variable time $t$ does not appear explicitly as follows (Logan, 2006):

$$y' = f(y). \tag{2.2}$$

The initial value problem (IVP) depends on initial conditions in solving the autonomous system (Logan, 2006). The first-order means the highest order of the equations is one. If

first order differential equations have exact solutions, then they are mostly solved by separation of variables (Vuldin, 2008). We can define the linear differential equations that satisfy the following concept. Suppose that $y$, $y_1$ and $y_2$ are scalars (vectors or functions), $f$ is a linear function if it satisfies the following conditions (Vuldin, 2008):

1. Additivity;

$$f(y_1 + y_2) = f(y_1) + f(y_2), \tag{2.3}$$

2. Homogeneity;

$$f(ay) = af(y), \tag{2.4}$$

where $a$ is a constant value. Otherwise, the function $f$ is nonlinear.

## 2.2 Mathematical Modelling

A model is considered as "an object or concept that is used to represent something else." (Meyer, 1984, p. 2). The mathematical model is sometimes defined as a collection of equations that describes the phenomenon in different fields such as science, biology, engineering, etc. (Logan, 2006). The idea of mathematical models is that people are trying to describe some things in their real life or real problems in the world using mathematical formulae. The attempts of formatting the real problems are expressed in mathematical descriptions of functions, inequalities, differential equations, constants, variables, etc. Naturally, the areas of building mathematical models are in nature, mathematical biology, social sciences, physical, mechanics and so on (Meyer, 1984). The modelling of a system of differential equations starts from making formulations for the equations, then giving initial points and expressions so that the changes in time may be accommodated.

In general, to make a mathematical model, first, we assume simple assumptions about realistic problems in our life then we try to test these assumptions and to predict the model statistically. On the other hand, we try to formulate these hypotheses as a system of

mathematical equations and before solving these equations or a system of the equations approximately or numerically (Fulford et al., 1997).

Formulating real problems with mathematical language is called mathematical modelling. One benefit of mathematical modelling: it has the idealization for selecting suitable variables, parameters, formulas and other relevant assumptions and theories. Mathematics covers the area of theorems, analysis and processes for doing logical concepts and equations in real life. With mathematics, we can analyze the real problems numerically via the modelling by computer. In short, the mathematical models require understanding of all facets of the problem to obtain a good model for the problem. Then we develop and test the model. The simple process of the mathematical model in the differential equations can be written as below (Nagle et al., 2008):

$$\frac{dy}{dt} = input\ rate - output\ rate. \tag{2.5}$$

The first order differential equations have real applications in many aspects of the life such as fluid flow and clinical medicine (Nagle et al., 2008).

A mathematical modelling has either probabilistic or deterministic behavior. The model that has changing elements is a probabilistic model while if it has fixed elements, it is a deterministic model. The deterministic methods can give approximations for either deterministic or probabilistic behavior. The probabilistic methods such as Monte Carlo (MC) can also give an approximation for either probabilistic or deterministic behavior (Giordano et al., 2003, p. 219).

The classification of the mathematical modelling can be described in different ways: static and dynamic or deterministic and stochastic models. If there is a variation in time that appears clearly in a model, it means that the model is a dynamic model, like the changes that happen in Newton's law of motion, predator-prey model, etc. If the variation in time does not appear in the model it means that the model is static. The model that has fixed solutions

is a deterministic model, for instance, the models that are solved by numerical methods. However, the presence of randomness in at least one variable in a model indicates that it is a stochastic model, for example the epidemic models (Rubinstein, 1981, p. 4).

In some of the real complicated problems, the mathematical modelling cannot explain the observed behavior analytically and therefore the prediction of behavior becomes necessary for this situation. The modelling for the empirical model construction is called simulation modelling. The behavior of the model may be simulated directly. Then the alternative behavior by simulation may be investigated. A simulation modelling tries to verify the relationship between two or more variables by creating some ranges which are confidence intervals for particular distributions or empirical confidence intervals (prediction intervals). The widely simulation method is the Monte Carlo method (Giordano et al., 1997) which is considered in this study. Peng (2016) discussed some medical problems on lymph nodes of breast cancer using alternative statistical techniques and studied the effect of lymph nodes removed (LNR) on breast cancer patients.

Sometimes the linear differential equations are not consistent with the observed data in some real problems; in this case, the nonlinear models may be better than the real linear models. Some examples of nonlinear models are the logistic equation, restricted growth that becomes a nonlinear population growth model when the population has unbounded rising growth such as measles epidemics that transfer by spreading the virus quickly between individuals during one week (Fulford et al., 1997). Lee (2008) generated random samples from mixed Poisson distributions using an acceptance-rejection algorithm to solve the nonlinear system.

A stochastic model has various values in each time and has a population distribution for the parameters and variables of the mathematical models. Therefore it needs assumptions, theory and analytic study. The important part of the stochastic model is that it has a sample distribution which is an experimental structure and considered as an approximation of the

population distribution (Mooney & Swift, 1999). Kypraios et al. (2017) analyzed the stochastic epidemic models with approximate Bayesian Computation method. Stockdale et al. (2017) predicted stochastic epidemic model and used Markov chain Monte Carlo method.

A behavior of deterministic mathematical model that deals with discrete time is called discrete dynamical systems (Mooney & Swift, 1999). In our study, we are interested with discrete stochastic-deterministic dynamical systems for autonomous nonlinear real models. There are many areas of biological mathematical modelling such as:

- Biochemistry: It deals with biochemical problems by applying for the computer programs in order to solve and to analyze the biochemical information. Therefore, it is also named the "computational biochemistry" (Elnashaie & Uhlig, 2007; Tsai, 2002).

- Biology: The design of the growing or declining of a population over time by mathematical language is known as biological study. This modelling is based on the population of animals, humans, disease, epidemic, microbiology, biomedical catalytic etc., for instance logistic equation in population growth, predator-prey model, epidemic models, infectious disease model (Allman & Rhodes, 2004; Brauer et al., 2001; De Vries et al., 2006; Ingalls, 2013; Murray, 2003; Shonkwiler & Herod, 2009; Murray, 2002).

- Biomechanical: This study deals with dynamic modelling of animals and human population using experimental data processing. The goal of such models is to study the motion that has started the population long time ago (Zinkovsky et al., 1996).

- Biostatistics: This approach focuses on the application of statistics in biological areas to understand the statistical methods that can be applied in the medical and

health areas, such as medical and clinical research, health services research, biology lab experiments and many other biological matters (Lachin, 2011; Lee, 2010; Newman, 2001; Van Belle et al., 2004).

- Epidemiology: This study focuses on biomedical sciences, epidemics, diseases, bacterial or viral growth, etc., to understand the dynamics of such models. Some of these models are stochastic models that have a stochastic term, consisting of differential equations with random model variables or with both random variables and random parameters (Chowell et al., 2009; Mazza & Benaim, 2014; Taylor & Karlin, 2014; Ullah & Wolkenhauer, 2011; Wilkinson, 2012). Other epidemic models are either deterministic models that consist of differential equations with missing randomness or stochastic-deterministic epidemic models that consist of differential equations with randomness in the model parameters. (Allen et al., 2008; Brauer et al., 2001; Chowell et al., 2009; Frauenthal, 1980; Krämer et al., 2010; Ma et al., 2009; Newman, 2001).

## 2.3 Stability and Linearization

A nonlinear system can be defined as a problem that has “variables cannot be written as a linear sum of independent components" (Vuldin, 2008, p.1). In general, nonlinear systems are difficult and sometimes it is impossible to solve them because most of such problems are not more understood than the linear problems especially because the real-life physical problems are nonlinear problems that are difficult and highly complicated in format and modelling (Vuldin, 2008).

Most of the time, the nonlinear dynamics of the system of differential equations cannot be solved analytically because of the change of time in these models although with the aid of existing numerical methods. In this case, the graphical analysis can describe the behaviors of nonlinear system variables. The solutions of such systems are called solution

curve, path, trajectory or orbit of the system. The plane of the system is called a phase plane. The value which satisfies the system and makes the functions on the right side is equal to zero and then the derivatives of the system are equal to zero is called equilibrium point or steady state. The study of trajectory behavior if the trajectories converge to the equilibrium point; this connotation is the idea of stability. In the other words, the stability is the analysis of the trajectories if they are near to the equilibrium points.

The classifications of equilibrium points are local asymptotically stable, global asymptotically stable or instability. The study of stability depends on how far is the distance between the near trajectories and the equilibrium points with small perturbations (Logan, 2006, p. 233). The equilibrium point is stable if any trajectory starts to approach a point and still approaches the same point in the future. The equilibrium point is asymptotically stable if any trajectory stable and starts approaching an equilibrium point and time $t$ close to infinity, (Giordano et al., 2003). In other words, the system is local asymptotically stable means "the system will return to the original point overtime" (Logan, 2006, p. 34). While the global stability means "the system returns to the state for all perturbations" (Logan, 2006, p. 34). An equilibrium point is not stable, if it is unstable or called semi-stable, which is not asymptotically stable. It means the orbits are moving away from the equilibrium point (Giordano et al., 2003; Logan, 2006; Nagle et al., 2008; La Salle, 1976). The study that treats the behavior of a nonlinear system of differential equations is called dynamical system study.

The important subject in the nonlinear dynamic system is the stability and how to find an approximation of linear system to nonlinear system models. This linear transformation in this process is called the linearization (Logan, 2006, p. 234). Linearization is a study of the nonlinear system behavior near equilibrium points by replacing the curve or surface of the nonlinear equations of the tangent line or the tangent plane (Mooney & Swift, 1999, p. 285). In order to understand the relationship between stability and linearization, the stability idea depends on studying the behavior of the trajectories near the equilibrium points. This

analysis is considered as a mirror to the nonlinear system by a linear system with small perturbations. It simulates the behavior of the nonlinear system using eigenvalues of the system in differential equations. That means the analysis of the linear system behavior with its properties corresponds to the analysis of the nonlinear system behavior with properties. The linearization of DE can be done by the eigenvalues of the linear system; this process is called locally stability (Logan, 2006, p. 234). Wang et al. (2017) determined the equilibrium points of the knowledge transmission model and evaluated the global asymptotic stability dealing with complex networks.

## 2.4 Mathematical Model of Epidemic

A model which deals with an epidemic that spreads rapidly in a large size population is called an epidemic model. The epidemic models are stochastic-deterministic models that format as a system of the first order differential equations. The aim of the epidemic modelling study is to analyze the epidemic behavior either decays, grows or remains in the population with the time (Diekmann & Heesterbeek, 2008). The important tool for testing the stability of epidemic model's behavior is the basic reproduction number ($R_0$). $R_0$ is a threshold quantity and considered as a tool to determine whether an epidemic occurs or the disease simply dies out. This value determines the probability of the transmission of disease (Diekmann & Heesterbeek, 2008). If $R_0$ is less than one, it means that the disease may not become an infection during the infectious period, therefore the infection will fade away in the future. In case of $R_0$ is more than one, there is an epidemic in the population. If $R_0$ is equal to one, means that the disease becomes endemic, that is the disease remains in the population at a consistent rate, such that each infected individual transmits the disease to other susceptible individual (Allen et al., 2008; Brauer et al., 2001; Chowell et al., 2009; Diekmann & Heesterbeek, 2008; Frauenthal, 1980; Krämer et al., 2010; Ma et al., 2009; Newman, 2001).

Recently, Chowell (2017) estimated epidemic model parameters using least-square fitting. Zarebski et al. (2017) analyzed a mathematical model of epidemics of seasonal influenza in Australia using the likelihood-based method. Kim et al. (2017) studied the optimal control strategies of influenza epidemic model in Korea. Levy et al. (2017) studied how to reduce the effect of Ebola virus outbreaks in Sudan through the health campaigns in the community. Kumar and Srivastava (2017) solved the problem of infectious disease by vaccination and treatment using optimal control approach. There are other researchers who analyzed the behavior of some mathematical epidemic models recently. Champredon et al. (2017) discussed Ebola synthetic epidemics. Gouvêa Jr (2017) used a new simulation method to estimate the mathematical model of Dengue for mosquito. Chowell (2017) discussed dynamic of epidemic outbreaks and estimated the parameters using fitting approach. The fitting data of a deterministic compartmental model for leprosy in Brazil have been evaluated (Blok et al., 2017). Song et al. (2017) analyzed the dynamics of infected diseases with time delay using nonlinear incidence rate where the latent period of disease stages was studied. Wang et al. (2017) established the relation between epidemic spread and the information efficiency for this epidemic using mobility patterns in complex networks and they noticed that the credibility of information can reduce infectious disease outbreak.

The epidemic populations are divided into separate classes according to the humans' vulnerability towards the disease; Susceptible ($S$), Exposed ($E$), Infectious ($I$), Recovered ($R$). Susceptible is the group of people with possibility of being infected, Exposed is the group of individuals who are infected but not yet infectious, Infectious are people who capable to spread the disease while Recovered are the individuals who have immunity from the disease and cannot infect others. The profile of a disease that can be represented by Susceptible-Exposed-Infectious-Recovered type is known as epidemic SEIR model. There are also others simple types of the disease models such as SIS (Susceptible-Infectious-Susceptible) type and SIR (Susceptible-Infectious-Recovered) type. The preliminaries of SIS, SIR and SEIR dynamic models are outlined by Diekmann and Heesterbeek (2008) and

Logan (2006). Recently, Azizi et al (2017) analyzed SIS model of the spread of infectious diseases with mixing distributions. Zaman et al. (2017) suggested a SIR epidemic model for the optimal vaccination and treatment. Wang et al. (2017) analyzed the behavior of the SEIR model with random networks. Lee and Chowell (2017) examined the optimal control strategies for the flu-like epidemics with SIR model using various parameters.

A social epidemic can be understood as an outbreak of a bad habit that happens under social pressure or custom. Social epidemiological models have been analyzed by many researchers to better understand the dynamics of the complicated phenomena. Santonja et al. (2010) predicted the future behavior of alcohol consumption in the Spanish population by estimating the parameters of the model and by fitting the model to real data. Sánchez et al. (2011) studied the evolution of cocaine abuse in Spain and predicted a decrease of cocaine abuse in the future by using the sensitivity analysis approach. Guerrero et al. (2011) studied the effect of the smoke-free law on the evolution of smoking habits in Spain before and after applying this law during 2006 to 2009 and predicted the effect of this law on the growth of the smoking habit in the Spanish population. Santonja et al. (2012) analyzed the effects of public health campaigns in Spain to change the people lifestyle behavior, nutritional behavior and to promote physical exercises in order to reduce their excess weight. Maha et al. (2015) solved the initial value problem of ordinary differential equations representing the social epidemic model of excess weight loss in Valencia numerically using Runge-Kutta methods RK2, RK4, RK45 and RK78 methods. With respect to the stability study of the social epidemic model, the global asymptotical stability of the basic reproduction number of an alcohol model was derived and the numerical simulation results were discussed by Zhu and Zhu (2017).

The stability of the epidemic models was also evaluated in the recent year. An SIR model of nonlinear autonomous system of delay ODE was discussed by Akimenko (2017). The reproduction number was derived from the cholera model (Berge et al., 2017), the stability of the SIR of deterministic model for HIV/AIDS was analyzed, the random

parameters of the model were simulated using LHS by Simpson and Gumel (2017), the stability of the SIRS model for infectious diseases were determined (Agaba et al., 2017), the global asymptotic stability of nonlinear dynamics of epidemic models were evaluated by Liu et al. (2017). Moreover, the basic reproduction number of SEIAR malaria model was computed (Cai et al., 2017) while the SIVRS epidemic model with complex networks was proposed by Xu et al. (2017).

## 2.5 Data Source

In this thesis, we consider secondary data as reported by Santonja et al. (2012) who studied the effects of public health campaigns towards the weight loss to demonstrate the application of our Mean Monte Carlo Finite Difference (MMCFD) method in Chapter 3. We also consider secondary data in Chapter 4 as reported by Sánchez et al. (2011) who studied cocaine abuse model in Spain to demonstrate application of our proposed Mean Latin Hypercube Finite Difference (MLHFD) method. We do not have access to the primary data source as captured by Santonja et al. (2012) and Sánchez et al. (2011) and based on APA citation style (DeCleene & Fogo, 2012) accepted by the University of Malaya Library, we may mention the primary data source only by giving citation to the secondary data source as available to us.

In Chapter 3 of this thesis, Santonja et al. (2012) studied the effects of public health campaigns towards weight loss as experienced by a Valencia community in Spain (24-65 years old) from 2000 to 2005. The primary data needed to conduct this study was reported to be obtained from Valencian Department of Health (2000 and 2005a) by Santonja et al. (2012). The public health campaigns include healthy nutritional habits aided with physical activity as well as strategies to reduce weight among the overweight and obese subpopulations in the Spanish population based on the Valencian Health Plan (2005-2009) which is supplied by Valencian Department of Health (2005b) as cited by Santonja et al.

(2012). Santonja et al. (2012) also reported that their model parameters were estimated by using the sources by Valencian Department of Health (2000) and (2005a) based on the health survey for the Valencian community in the year of 2000 and 2005.

In Chapter 4 of this thesis, Sánchez et al. (2011) studied the cocaine abuse problem in Spain during a period of ten years (1995 - 2005). Sánchez et al. (2011) predicted the profiles of cocaine abuse subpopulations up to 2015. The primary data source was accessed by Sánchez et al. (2011) from the Spanish Health Ministry, National Drug Strategy 2000–2008 and Official State Gazette in 2009. The classification of the subpopulations for the cocaine abuse problem was described by the Spanish Health Ministry in 2008 as cited by Sánchez et al. (2011). Spanish birth and death data were collected from Spanish Statistics Institute in 2008 as reported by Sánchez et al. (2011). Moreover, the real data to reduce the cocaine abuse were supplied by the Drug National Observatory Reports in 2000 and 2007 as reported by Sánchez et al. (2011).

## 2.6 Methodology

Most of the deterministic epidemic models can be represented as systems of ordinary differential equations depending on various times $t$. Some of these systems can be solved by using a simulation approach to introduce randomness in the parameters of such models. Commonly, Monte Carlo (MC) simulation technique and Latin hypercube sampling (LHS) procedure can be used to simulate the values of the parameters. In this study, two simulation methods with random sampling are integrated with a numerical deterministic approach. to solve the social epidemics problems; statistical simulation processes. We explain all these approaches in this section.

### 2.6.1 Random Number and Random Sampling

Random variable (RV) is a random value that has a probability distribution. It can be extended to a random sampling computerization. This process is called a simulation or a resampling and sometimes it is called Monte Carlo (Gentle, 2003). The simulation term originates from the Latin word „„*simulare*"" and it generally means "the application of a model with the objective to derive strategies that help in solving a problem or to answer a question pertaining to a system." (Velten, 2009, P. 7). The most common application of simulation is in stochastic processes. A simulation becomes necessary when difficulty exists in population resampling and when the mathematical analysis becomes intractable (Gentle, 2003). Other reasons to do simulation are discussed by Rubinstein (1981, p. 9-10). In short, we need to apply simulation approach in order to obtain a random distribution of a variable studied.

### 2.6.2 Monte Carlo (MC) Simulation Process

MC sampling refers to the traditional technique of using random or pseudorandom numbers to sample from a probability distribution (Rubinstein, 1981). MC method becomes a universal technique after the development of computers. Monte Carlo name is derived from the city in Principality of Monaco (Sobol, 1994). Using computer software, MC algorithm generates random numbers that have a probability density function that equal to one if the numbers are between 0 and 1, and equal to 0 elsewhere. These numbers are considered as random variables distributed uniformly on $(0,1)$. The random quality that is obtained by a simulation process is pseudo-random in nature (Rubinstein, 1981). These random numbers are called random number generators, pseudorandom or quasi-random (Rubinstein, 1981, p. 20).

Naturally the computer produces independent random variables that have uniform distribution on $(0,1)$ (Rubinstein, 1981). Direct sampling methods produce independent random variables (samples) that have other probability distributions. The usual direct sampling methods are the Box Muller algorithm for normal distribution and the mapping method (Graham & Talay, 2013). In the current study, we use direct sampling methods to produce the uniform distribution with a particular interval. We can generate pseudo random numbers by using MATLAB function RAND to generate random variables that have uniform distribution (Cheney & Kincaid, 1999).

MC simulation process generates uniform random numbers. The following MC procedure is used in our study (Carsey & Harden, 2014): Firstly, we generate random numbers on the interval (0,1) such that each one of these random numbers is considered as a random variable that has uniform distribution on the interval (0,1) (standard uniform distribution). Then the inverse transform method is used to transform the random variables which have the standard uniform distribution, into random variables that have specific distribution (Cheney & Kincaid, 1999). The inverse transform method (inversion method) has the following formula:

Let $Z$ be a random variable that has a specific distribution and cumulative distribution function $F$. Suppose that $F^{-1}$ is the inverse of function $F$ and suppose that ς is a continuous random variable that distributes uniformly on interval $(0,1)$, such that (Cheney & Kincaid, 1999):

$$\mathbb{P}(F^{-1}(\varsigma) \leq z) = \mathbb{P}\big(\varsigma \leq F(z)\big) = F(z). \tag{2.6}$$

Due to ς is a continuous random variable then $\mathbb{P}\big(\varsigma = F(z)\big) = 0$ and by taking the $F^{-1}$ for $\varsigma = F(z)$, the random variable $Z$ is equal to $F^{-1}(\varsigma)$ and it has been written as:

$$Z = F^{-1}(\varsigma). \tag{2.7}$$

Box-Muller transformation is an example of a method that uses the inverse transform to convert two uniform random variables into normally distributed random variables. In this thesis, $Z$ is considered as a RV which has a uniform distribution on $(a, b)$ where $a$ and $b$ are integer's numbers and represent lower and upper bounds.

The probability density function of uniform distribution on $(a, b)$ is

$$f(z) = \begin{cases} \dfrac{1}{b-a}, & for \quad a \le z \le b, \\ 0 & otherwise. \end{cases}, \tag{2.8}$$

and the cumulative distribution function $F$ of uniform distribution on $(a, b)$ is

$$F(z) = \begin{cases} 0 & for \quad z < a, \\ \dfrac{z-a}{b-a} & for \quad a \le z \le b, \\ 1 & for \quad z > b. \end{cases} \tag{2.9}$$

Then the inverse of the uniform cumulative distribution function $F^{-1}$ has the following formula:

$$z = F^{-1}(\varsigma) = a + (b-a)\varsigma, \tag{2.10}$$

where $\varsigma$ has the standard uniform distribution (Rubinstein , 1981, p. 65).

Recently, Ndanguza et al. (2017) estimated the parameters of the Ebola epidemic model of a stochastic differential equation using Monte Carlo. Zhao et al. (2017) also simulated the unknown parameters by using the Monte Carlo process. Enduri and Jolad (2017) estimated the reproduction number of dengue epidemic and simulated the model parameters through Monte Carlo sampling method. A mathematical model from the stochastic differential equations are modelled by Yalim et al. (2017) where the authors discussed the noise of the model and used Monte Carlo simulations.

### 2.6.3 Latin Hypercube Sampling (LHS) Technique

LHS is one of the most popular statistical methods to generate random samples. It is an extension of stratified sampling in multiple dimensions to generate better samples with uniform distribution written in the form of a matrix such that each column positions as a variable (parameter in our study) and each row as a sample. Prior to that, the sample that has a uniform distibution on $(0,1)$ must be generated while the input random numbers, the number of dimensions (number of parameters in our study) and the sample size (number of simulations in the current study) must be determined (Dagpunar, 2007). In the current work, the range (0,1) of each dimension (parameter) is divided into $n$-equal disjoint levels (sometimes called the stratum, stratification or bins) such that the number of parameter is equal to the number of dimension. Then each one of these parameters has $n$-equal intervals (strata) with the form $((i-1)/n\,, i/n]$ where $i = 1,2,\dots,n$ and $n$ is the number of simulation's repetitions. That means, each sample from each stratification has uniform probability distribution on $1/n$. Next, only one sample is selected randomly from each stratum (Liu & Yang et al., 2015). The basic idea of LHS random sampling is to transfer the generated random samples which have the $(0,1)$ uniform distribution to strata of the cumulative probability distribution (CDF); by dividing the CDF into equal partitions then selecting just one random sample value from each partition of the CDF. Then the inverse of the CDF obtained from the inversion method is used to find a random variable that is distributed uniformly on the created interval in the current work. Finally, each parameter is considered as a random vector of the random variables which belong to the uniform distribution on the created interval (Dagpunar, 2007).

In general, to simulate model parameters using LHS technique, most samples used are those with uniform distribution. Recently Li et al. (2017) simulated the epidemiological parameters of Ebola outbreak in West Africa uses LHS. Li et al. (2017) proposed a method to solve nonlinear dynamic systems with uncertain parameters using regression and

Chebyshev polynomials and simulated the model parameters by LHS. In the current study, we apply LHS technique in order to simulate the model parameters.

### 2.6.4 The Difference between MC and LHS

Both MC and LHS techniques are unbiased estimation with some differences. LHS is an extension of stratified sampling for the multidimensional data distribution. It is a variance-reduction technique and one of the most popular statistical methods for generating random samples of parameters' values from multiple dimensions.

Most of the sampling methods are varied in terms of how the simulation results are obtained and how long the simulation time is consumed. Monte Carlo (MC) simulation process refers to the traditional sampling technique using random numbers to sample from a particular probability distribution. Since MC is a pure randomness process in essence, the sample values may be found anywhere within the range of the input distribution and it can be inefficient. Apart from MC, LHS is a randomized stratified sampling for the input probability distributions where it evenly spreads the sample points into smaller domain and sample size of equal probability. Furthermore, the difference between MC and LHS is attributed to LHS experimental design which deals with characteristic of multilevel sampling written in the form of a matrix. MC is memoryless in approach such that the new values are generated without containing the previous account's sample. On the other hand, LHS is a memory system in approach such that it remembers in which row and column the selected sample was taken. The important point is that the LHS approach can give better simulation results than MC with the same number of repetition because it is faster to reach good representation of probability distribution than MC sampling (Gentle, 2003).

Furthermore, the difference between MC and LHS is attributed to LHS experimental design which deals with a characteristic of multilevel sampling written in the form of a matrix. MC is memoryless in approach such that the new values are generated without

containing the previous account's sample. On the other hand, LHS is a memory system in approach such that it remembers in which row and column the selected sample was taken. The important point is that the LHS approach can give better simulation results than MC with the same number of repetition because it is faster to reach good representation of probability distribution than MC sampling (Gentle, 2003). The Monte Carlo's sampling error is $O(1/\sqrt{n})$ whereas the LHS sampling error is $O(1/n)$, in other words, with multidimensional systems, LHS converges faster than MC (Mainik, 2015).

The simulation methods are different in the number of simulations (length of the time), the quality and accuracy of results. Because MC is a random process (the sample can come from in anywhere in the range of the input distribution), therefore, MC needs to use a large number of repetitions to give more accurate simulation results. On the other hands, only a few numbers of repetitions for LHS is needed to obtain accurate results. That means LHS gives a good accuracy with less repetitions compare to MC, for example if LHS get an accuracy with $n$ simulations whereas MC need the $n^2$ to get the same accuracy. Another comparison; with MC, random samples may be selected not in the close region to a model under study. While LHS is spread the sample values until the range of the simulated parameter is covered (Mainik, 2015).

### 2.6.5 Numerical Finite Difference (FD) Method

The system of differential equations which have time-dependent coefficients can be solved numerically by finite difference method (FD). Sometimes it is called a method of lines that can be considered as a discretization method (Gustafsson, 2011). FD is a numerical method to solve initial value problem. The results of FD represent discrete numerical values that approximate the exact solution, (Logan, 2006). FD is an iteration process to solve differential equations (Fulford et al., 1997). It considered as an approximation of the derivative of the differential equation.

There are three types of finite approximation methods; finite difference, finite volume and finite element. Finite difference (FD) method is the oldest numerical method to solve differential equations. It approximates the derivatives of differential equations and deals with the points where the solution domain is treated as a grid system (Ferziger et al., 2002, P. 35). Finite volume (FV) method deals with the integral form and approximates surface and volume integrals when the solution domain is subdivided into a limited number of neighboring volumes. The surface and volume integrals are approximated by appropriate quadrature formulae (Ferziger et al., 2002, p. 36). Finite element (FE) method has the most properties of the FV method. The domain is divided into a collection of discrete volumes of finite elements. The solution is approximated by a linear function (Ferziger et al., 2002, p. 36).

Yusof (2011) used the FD method to solve partial differential equations and applied it to the lithium-ion concentrations. Simulation technique for parameters was used to analyze the results of boundary value problem using *t*-test.

The general form of FD for an ordinary differential equation (ODE) with initial value Eq. (2.1) can be written as follows:

FD discrete the time $t$ in $(a, b)$ into $m$ subintervals which is equal to endpoints $t_i = a + ih$, for $i = 1,2, \dots, m$, where $m$ is the maximum number of iterations and $h = \frac{b}{m}$ is a step size.

The step size $h$ = 1 (per day, week or year) is chosen here since we are solving the real social epidemic model estimated on a time basis. Therefore, Eq. (2.1) becomes:

$$y'(t_i) = f\big(t_i, y(t_i)\big). \tag{2.11}$$

By using the central difference formula of FD, Eq. (2.2) becomes (Ferziger et al., 2002):

$$\frac{dy}{dt} \approx \frac{y_{i+1} - y_{i-1}}{2h}, \quad for\ each \quad i = 1,2, \dots m. \tag{2.12}$$

### 2.6.6 Derivative and Properties of Finite Difference (FD) Method

Any numerical method involves error. There are many kinds of numerical errors; the most common errors are the absolute error, relative error, and truncation error. Suppose that $a$ and $b$ are two values, one the exact value ($a$) and the other approximate value ($b$). The formula of absolute error is $|a - b|$, and relative error is $\frac{|a-b|}{|a|}$ (Cheney & Kincaid, 1999, p. 4).

Truncation error comes from cutting a part of the Taylor series approximation of a function. That means, the cutting terms of Taylor series approximation that approaches zero, is symbolized $O(h^m)$ of order $m$ and named truncation error when $h$ is the step size and $m$ is a positive integer value. The order error is affecting the accuracy. That means the numerical method that has higher error order is more accurate than the lower error order. Such as the best length of Euler (0.1), modified Euler (0.01) and Runge-Kutta (0.0001) methods. We can expect a truncation error during computer perform with determined length (Cheney & Kincaid, 1999; Logan, 2006; Giordano et al., 2003). Suppose that a step size *h* is a fixed positive number. Some of higher powers for approximations of a function using Taylor series, approach to zero. The error term of central FD is (Giordano et al., 2003):

$$O(h^2) = -\frac{1}{6}h^2 y'''(t). \tag{2.13}$$

We can explain the deriving of the central finite difference formula by using Taylor theorem as follows:

The expression of Taylor series of order $m$ at $h$ is

$$y(t+h) = y(t) + hy'(t) + \frac{1}{2}\, h^2\, y''(t) + \frac{1}{6}h^3 y'''(t) + \cdots + \frac{1}{m!}h^m y^{(m)}(t) + \cdots \quad (2.14)$$

The approximate functions at $(t+h)$ and $(t-h)$ lead to

$$y(t+h) = y(t) + hy'(t) + \frac{1}{2}\, h^2\, y''(t) + \frac{1}{6}h^3 y'''(t) + \cdots \quad (2.15)$$

$$y(t-h) = y(t) - hy'(t) + \frac{1}{2}\, h^2\, y''(t) - \frac{1}{6}h^3 y'''(t) + \cdots \quad (2.16)$$

We obtained by subtraction

$$y(t+h) - y(t-h) = 2hy'(t) + \frac{2}{3!}\, h^3\, y'''(t) + \frac{2}{5!}h^5 y''''(t) + \cdots \quad (2.17)$$

The central finite difference formula with its truncation error can be written as

$$y'(t) = \frac{1}{2h}[y(t+h) - y(t-h)] - \frac{1}{6}h^2 y'''(\delta), \quad (2.18)$$

where $\delta$ is a value in the interval of a function $y$ and $-\frac{1}{6}h^2 y'''(\delta)$ is the error term (Cheney & Kincaid, 1999, p. 173).

To construct a formula of finite difference methods to solve differential equations, let first mention that in our study considers the function $f$ in Eq. (2.1) is a nonlinear of $y$, $h$ is the step size, $m$ is an integer and $t = 1,2,3, \ldots, m$. Suppose that

$$t \approx t_m, y(t) \approx y_m, f(y) \approx f_m \text{ and } f_m \approx f(y_m). \quad (2.19)$$

The first derivative form in calculus is given for the forward, backward and central difference schemes is given as

$$\frac{dy}{dt} = \lim_{h\to 0} \begin{cases} \dfrac{y(t+h)-y(t)}{h}, \\ \dfrac{y(t)-y(t-1)}{h}, \\ \dfrac{y(t+h)-y(t-h)}{2h}. \end{cases} \rightarrow \begin{cases} \dfrac{y_{m+1}-y_m}{h}, \\ \dfrac{y_m-y_{m-1}}{h}, \\ \dfrac{y_{m+1}-y_{m-1}}{2h}. \end{cases} \tag{2.20}$$

Substitute the corresponding Equations (2.19) and (2.20) in Eq. (2.1) to obtain the finite difference schemes of Eq. (2.1);

$$\frac{y_{m+1}-y_m}{h} = f_m, \tag{2.21}$$

$$\frac{y_m-y_{m-1}}{h} = f_m, \tag{2.22}$$

$$\frac{y_{m+1}-y_{m-1}}{2h} = f_m. \tag{2.23}$$

The Equations expressions (2.21), (2.22) and (2.23) are called the respectively forward Euler, backward Euler and central finite difference schemes (Mickens, 1994, p. 2-3).

The numerical method is named implicit if the next step term like $y_{m+1}$ appears on two sides of the formula (Nagle et al., 2008). The backward Euler method is an implicit scheme because the next step $y_m$ appear on two sides and the scheme must be solved for $y_m$, (Mickens, 1994). We can derive from FD some iteration numerical methods to estimate the next step $i+1$ of FD such as Euler of order $O(h)$ and Runge-Kutta RK4 method of order $O(h^4)$.

Numerical instabilities of the discrete dynamic system of nonlinear differential equations can happen when the finite difference equations have not corresponding solutions

of the differential equations in the discrete model. In the other words, this case is an indicator that the difference schemes in the discrete model of differential equations do not have the ability to get a good mathematical modelling to the problem under study. The main reason of numerical instabilities is that the parameters of the model have space larger than the corresponding differential equations. In general, the numerical instabilities occur insure for all step sizes when the order of finite difference equations becomes larger than the order of differential equations because the higher order of the difference equations has general solutions more than the differential equations that corresponding them. Therefore the central finite difference scheme to solve first order differential equations has numerical instabilities because it is of order two (with two linearly independent solutions) that more than the order of differential equations. The second reason of numerical instabilities is if there are no restrictions on the step size that the discrete model requires, for example the forward Euler and Runge-Kutta methods have this case of numerical instabilities. As well the central finite difference scheme to solve first order autonomous differential equations of the discrete model may be having the numerical instabilities (Mickens, 1994, p.60-61).

### 2.6.7 Numerical Simulation Methods

Numerical simulation methods obtain results by solving the system of differential equations using a numerical method and simulation processes. Previously, an inverse problem for nonlinear parabolic was solved by finite difference scheme jointed with Monte Carlo algorithm and the unknown diffusion coefficient was estimated using polynomial format (Farnoosh & Ebrahimi, 2010). Hosseini and Shahabian (2011) used a hybrid numerical method that consists of Galerkin finite element with Newmark finite difference methods and discussed the stochastic analysis. Finite difference method was integrated with Monte Carlo simulation process in order to predict the behavior of the dam (Rohaninejad & Zarghami, 2012), the random variables which were generated by Monte Carlo method in

this problem have a Gaussian distribution. In recent year, the elliptical partial differentials are analyzed by the stochastic finite element method (Pryse & Adhikari, 2017). Nonlinear random differential equations were solved by generalized polynomial chaos method (Cortés et al., 2017). Faes & Moens (2017) studied the uncertainty of the parameter in a spatial interval on a physical model using LHS and finite element.

#### 2.6.7.1 Mean Monte Carlo Finite Difference (MMCFD) Method

The Monte Carlo simulation with finite difference (FD) method is repeated $n$ times to simulate values of the variable parameters as random sample. The mean of the $n$ final solutions via this integrated technique, named in short as mean Monte Carlo finite difference (MMCFD) method, represents the final solution of the system. This method is proposed for the first time to calculate the numerical solution obtained for each subpopulation as a vector distribution.

In Chapter 3 there are three variables considered in the weight reduction model which are $N(t)$, $S(t)$ and $O(t)$, while in Chapter 4 the variables for the cocaine abuse model are $n(t)$, $o(t)$, $r(t)$ and $h(t)$. Therefore in Chapter 3, $k = 3$ and in Chapter 4, $k = 4$ where $k$ denotes the number of variables in a model. Hence the Monte Carlo finite difference solutions for one variable ($k = 1$) may appear as follows:

$$\left(sol_{i,j}(t)\right)_{k=1} = \begin{pmatrix} sol_{11} & \cdots & sol_{1n} \\ \vdots & \ddots & \vdots \\ sol_{m1} & \cdots & sol_{mn} \end{pmatrix}_{m\times n}, i = 1,2,\dots,m,\ j = 1,2,\dots,n, \qquad (2.24)$$

where $m$ is the maximum number of iterations as well as it is the number of weeks (in Chapter 3) or years (in Chapter 4) in each considered interval. $n$ is the number of random sample space for model parameters and it also represents the maximum number of simulations. The numerical simulation results ($sol_{i,j}(t)$ of Monte Carlo finite difference)

obtained for one variable of a system are presented in the above Eq. (2.24) for $n$ simulations and $m$ numerical iterations.

Then, the numerical simulated solutions of Monte Carlo finite difference of a system for $i$ numerical iteration and for $j$ simulation can also be written as:

$$sol_{i,j}(t) = \begin{pmatrix} Solution\ of\ variable\_1 \\ Solution\ of\ variable\_2 \\ \vdots \\ Solution\ of\ variable\_k \end{pmatrix} = \begin{pmatrix} (sol_{i,j}(t))_1 \\ (sol_{i,j}(t))_2 \\ \vdots \\ (sol_{i,j}(t))_k \end{pmatrix}. \tag{2.25}$$

The Monte Carlo finite difference final solutions which correspond to the last numerical iteration $m$ for $k$ variables in a model with $n$ simulations as represented in the last row of Eq. (2.24), can be written as:

$$sol_{m,n}(t) = \begin{pmatrix} \overbrace{sol_{m,1}(t)_1}^{\text{Final solutions of variable\_1 } l=1} & \cdots & \overbrace{sol_{m,1}(t)_k}^{\text{Final solutions of variable\_}k\ l=k} \\ \vdots & \ddots & \vdots \\ sol_{m,n}(t)_1 & \cdots & sol_{m,,n}(t)_k \end{pmatrix}, \qquad l = 1,2,\dots,k. \tag{2.26}$$

The mean of each column of the final solution in Eq. (2.26) represents solution of the system for each variable $l$ that is displayed in the following terms:

$$\mu_{sol,l} = \sum_{j=1}^{n} \frac{sol_{m,j}(t)_l}{n}, \qquad l = 1,2,\dots,k, \tag{2.27}$$

such that $sol_{m,j}(t)_l$ are elements in Eq. (2.26), where $m$ is the last numerical iteration, $j$ is the number of simulations and $k$ is the total number of variables in a model. Hence, the mean of Monte Carlo finite difference solutions (MMCFD solution) of a system, $\mu_{sol}$ is given as:

$$\mu_{sol} = (\mu_{sol,1}, \mu_{sol,2}, \dots, \mu_{sol,k}), \tag{2.28}$$

where $k$ is the total number of variables in a model and $\mu_{sol}$ is an estimated solution of a system (the MMCFD solution of a system). The MMCFD method is fully executed using MATLAB software where the flow chart is presented in Figure 2.1 while the algorithm is given in Appendix A.

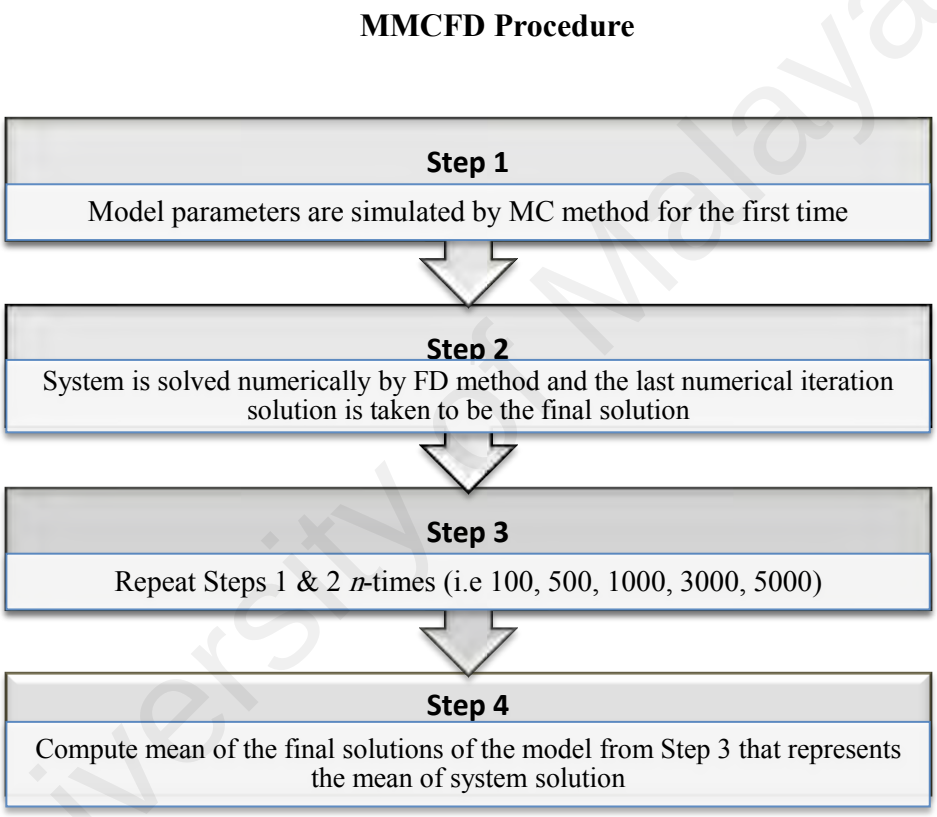


**Figure 2.1:** Flow chart of MMCFD method

### 2.6.7.2 Mean Latin Hypercube Finite Difference (MLHFD) Method

Mean Latin Hypercube Finite Difference (MLHFD) method is a new modified numerical simulation technique proposed by integrating two methods of different natures together in this study a statistical simulation process with random sampling, LHS and a numerical deterministic approach, FD. The deterministic system under study depends on a continuous time that appears in the variables and parameters. MLHFD benefits from the

deterministic system due to its random distribution properties in various model variables/parameters. Furthermore, the MLHFD numerical simulation process is more flexible to predict the range of solutions than the classic numerical methods which depend on a fixed time. The MLHFD method is fully executed using MATLAB software and the algorithm is presented as a flow chart in Figure 2.2 and the program in Appendix B.

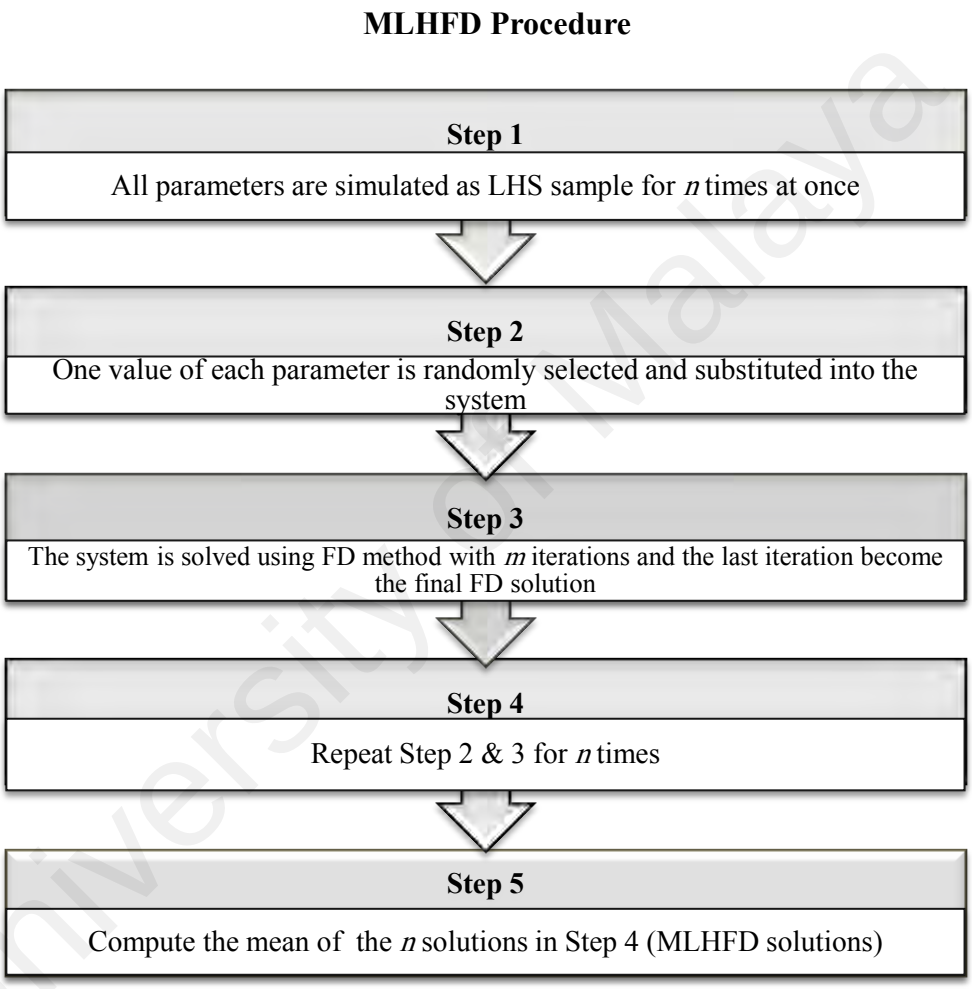


**Figure 2.2:** Flow chart of MLHFD method

### 2.6.8 Statistical Indicators

This study focuses on certain statistical indicators such as measure of difference $|E|$, standard deviation of MC ($\sigma_n$) and percentile of distribution in order to compare between MMCFD or MLHFD results and the predicted results. For the purpose of comparison

between numerical and statistical methods, we define a measure of difference $|E|$, written as (Cheney & Kincaid, 1999):

$$|E| = \left|x_\rho - x_{sol}\right|, \tag{2.29}$$

where $x_{sol}$ is a deterministic or numerical simulation value while $x_p$ is a predicted value. A small value of $|E|$ would indicate in the present work.

A standard deviation $\sigma$ is calculated as a special measuring of an error for the MMCFD method. The expression $\sigma/\sqrt{n}$ reduces the error with the square root of the sample size $n$, approximating to the standard deviation of the means, $\sigma_n$ which has the following formula:

$$\sigma_n \approx \frac{\sigma}{\sqrt{n}}, \tag{2.30}$$

where $n$ represents the number of simulations in the present computation and the sample size of the model random variables. Since the value of the standard deviation $\sigma$ of the population is not given, the variance of $\sigma_n$ defined as:

$$\sigma_n^2 = \langle I^2 \rangle - \langle I \rangle^2, \tag{2.31}$$

is used in the current work with

$$\langle I \rangle = \frac{1}{n}\sum_{j=1}^{n} sol_{m,j}(t)_l, \tag{2.32}$$

and

$$\langle I^2 \rangle = \frac{1}{n}\sum\nolimits_{j=1}^{n} \left(sol_{m,j}(t)_l\right)^2, \tag{2.33}$$

such that $sol_{m,j}(t)_l$ are elements in Eq. (2.26), $m$ is the last numerical iteration, $j = 1,2,\dots,n$ number of simulations, $l = 1,2,\dots,k$ and $k$ is the total number of variables in the model under study. For more details on $\sigma_n$, please refer to Jacoboni and Lugli (1989).

In general, the best result is obtained when the standard deviation of the means $\sigma_n$ is the smallest.

### 2.6.9 Range and Prediction Interval

Another statistical indicator to describe data (solution) is prediction interval. It is considered as description of the numerical simulation solutions obtained in present study. The proposed modified method is useful to determine prediction interval when random distribution of the numerical solution is necessary for estimation of real epidemiological models. Prediction interval is an interval contains upper and lower bounds for predicted values of the distribution for each subpopulation of a model. It can be obtained by using the $p^{th}$ percentiles ($p\%$) to give upper and lower limits. A percentile is a value inside a distribution that can divide ordered predicted values into two or more parts by drawing lines between these values. These predicted values belong to the vector distribution of random variables. As a consequence, the $p^{th}$ percentile of the predicted values is inside a population. This percentile value is equal to or less than the number that required to calculate it when $0 < p < 100$. Suppose that $n$ is the total number of predicted values in a distribution (random sample size). The index becomes $((n) \times (p \div 100))$ and represents the place for the $p^{th}$ percentile value within the population distribution. All the predicted values are arranged from the smallest time value until the largest time value in the distribution, see (Wilcox, 2006; Patel, 1989).

In Chapter 3, for example 0.581503 is the upper limit ($95^{th}$ percentile of $N(t)$ at the end of 2015 for 5000 simulations) of predicted values in the population. It means 95% of predicted values for the population are equal to or less than 0.581503. The lower limit of this prediction interval is the $5^{th}$ percentile. The median here is the $50^{th}$ percentile of the predicted values. In other words, The $50^{th}$ percentile is a data that splits the entire data distribution into two pieces such that 50% of the values lie in one piece and the other values

lie in the second piece, for more clarity, see (Wilcox, 2006). In summary, the predicted values are arranged from the smallest time value to the largest time value of the distribution. These percentiles are dependent on the size of the prediction interval. The prediction interval obtained takes into account the empirical 5% and 95% percentiles. (Wilcox, 2006).

# CHAPTER 3: NUMERICAL SIMULATION TECHNIQUE FOR WEIGHT REDUCTION MODEL DUE TO PUBLIC HEALTH CAMPAIGNS IN SPAIN

## 3.1 Introduction

Differential equation arises in formulation of many real life and industrial based dynamical relations that describe growth, transformation and change of specific dependent and independent mechanisms that may affect the whole system of inquisition. Generally, differential equations are categorized as ordinary differential equations (ODEs) or partial differential equations (PDEs). ODEs are differential equations with only one independent variable that reflected in the denominator expression, while PDEs have two or more independent variables in the system. There are three types of solutions for differential equations, which are exact, analytical, and numerical solutions.

Many researchers have conducted study on differential equations using vast choices of method. Since exact solutions do not exist for many complex models, closed form and analytical solutions are preferred as second choice alternative depending on the versatility, practicality and mathematical severity of the problems. Mohyud-Din et al. (2009) reviewed the applications of the analytical methods of He‟s and modified variational iterations, homotopy perturbation, parameter expansion and exp-function to solve a wide class of nonlinear initial and boundary value problems. Further discussion on such methods is also studied in Ali et al. (2016), Mohyud-Din et al. (2012), and Ul Hassan and Mohyud Din (2016). On the other hand, numerical methods are favored by numerical scientists and engineers due to simpler algorithm and computer-handling operations on complex solutions, less time consumption, and when the exact or analytical solutions of certain models are not yet found.

Finite difference (FD) method is one of established numerical methods to solve differential equations approximately. This method is widely applied in practical mathematical and physical modelling (Karris, 2007), numerical analysis/simulation and discrete domain problems throughout many fields of science, engineering, social science, medicine and economics. Witnessing continuous demand and progress in these fields, many authors employed FD in their works including on ODEs. Rao and Chakravarthy (2013) presented FD method for singularly perturbed linear second-order differential difference equations of convection–diffusion type with a shift parameter when the shift is smaller or greater than a perturbation parameter. Moghaddam and Mostaghim (2013) used a new numerical method that depends on FD to provide feasible solution for boundary problem of delayed differential equations of fractional order. In 2014, linear second-order differential difference equation of convection–diffusion type with singularly perturbed boundary value problem is discussed, where a FD method is introduced with fitting parameter from the theory of singular perturbation (Rao & Chakravarthy, 2014). Others works on the FD method can be found in Liu and Chen (2014), Su et al. (2013) and Zhao (2013).

Monte Carlo (MC) method is an eminent statistical modelling method in applied mathematics to develop models of phenomena with unknown probabilistic property. Because of its simplicity and good statistical results, the MC method has become a widespread technique used for doubt quantity. Due to increasing demand of statistical modelling and fast development of technology, MC method is employed with multiprocessor computing system in order to simulate many independent statistical experiments at a time. This progress enables difficult problems of probabilistic nature to be solved using the MC method by direct simulation of the natural probability model as well as to solve deterministic problems that are represented as constructed probability processes. These processes are modelled, and the numerical solutions are framed in the form of statistical estimators of deterministic equations with either random coefficients or random boundary conditions or right sides random (Rubinstein, 1981). The MC method can also be

applied to solve problems in finance, science, engineering, economics, and actual mathematics. It can be used to generate random data in behavioral studies as well as to estimate numerical quantities via repeated sampling. Moreover, mathematical algorithms can be randomized using computers to carry out statistical sampling in optimization problems via the MC method.

There are many researchers who organized research in parameters estimation of ODEs with MC. Spigler and Zerbetto (2013) studied the oscillatory behavior of tension leg platform offshore structure using MC method, while De Souza et al. (2014) provided MC solution of ODEs from a first law analysis and calculated the stochastic moments. Licea et al. (2013) dealt with Riccati equation that has random coefficients and compared the solutions with MC method. Meanwhile Buezas et al. (2013) solved a nonlinear ODE using finite element method with MC simulation technique. Later, Leander et al. (2014) presented an article related to estimated parameters of ODEs that have experimental data for discrete time. It is demonstrated that transforming an ODE to a stochastic differential equation as an objective function eliminates problem of local minima. Rohaninejad and Zarghami (2012) predicted the behavior of impounded dams by employing MC with FD method, where the statistical test Kolmogorov–Smirnov is used to predict the dam behavior. Other applications of MC method are given in Abdulle and Blumenthal (2013), Dereich and Heidenreich (2011) and Kovtanyuk et al. (2012).

Obesity and overweight are excessive fats in one‟s body measured using a simple formula of body mass index (BMI). These abnormal body weights have been medically proven as major risks to modern population with known linkage to many death cases around the world due to cardiovascular diseases, such as stroke and heart disease, diabetes, musculoskeletal disorders especially osteoarthritis, and also due to colon, endometrial, and breast cancers. Based on statistics from the World Health Organization, the world obesity rate has increased by 200 percent since 1980 with 42-million children under five years old are obese or overweight in 2013 alone. Mainly caused by imbalance of daily nutritional

intake and lack of physical activities to burn extra calories taken, the excess weight continues to become a serious health problem, especially in the low- and middle-income countries exasperated by insufficient prevention strategies, law policies, and power enforcement by the corresponded governments. As much as the rate of excess weight is rising at worrying stage, the research and studies devoted to this issue also keep rising.

Santonja et al. (2012) studied the effects of a public campaign to reduce excess weight in a Spanish region within a community of Valencia that is caused by unhealthy lifestyles. A mathematical model is presented to track on the progression of excess weight loss influenced by prevention strategy of public health campaigns. The campaigns focus on overweight and/or obese individuals in order to reduce their weight. In Snyder (2007), the author reached to a conclusion that nutritional campaigns should be able to change feeding behavior. Christakis and Fowler (2007) investigated similar issue by assuming the excess weight (obesity and overweight) as a socially transmitted epidemic disease. Further studies on excess weight from diverse perspectives are conducted in Livingston et al. (2001); Navarro-Barrientos et al. (2011); and Thomas et al. (2014).

In the present study, FD method integrated with MC technique is considered to simulate coefficients of time $t$ (week) of a nonlinear system of equations representing three Valencia community subpopulations based on BMI (Santonja et al. 2012). The system consists of three nonlinear ODEs with multiple coefficients/parameters that are described as random variables with probability distributions. The integrated method we named as mean Monte Carlo finite difference (MMCFD) method is used to simulate the parameters of the model (Santonja et al. 2012) and secondly to calculate the prediction interval for the numerical solution of every dependent variable of the system when the real-valued coefficients are provided previously.

Because of random sampling and rounding errors, the solution must be supported by prediction interval based on percentile of variables distribution. In particular, we are using the $5^{th}$ and $95^{th}$ percentiles in the present work. This interval of numerical results is later compared with the interval of statistical predictions (Santonja et al. 2012) derived by Latin hypercube sampling (LHS) technique. In order to obtain empirical prediction intervals, this article leads to set up interval by using the percentiles from the simulated values. This prediction interval represents empirical prediction intervals. The MMCFD solutions are also compared with the results generated by the FD method via central divided difference formula. The MMCFD method shows better accuracy and convergence to the given predicted results than the FD method. Thus, it is expected that the MMCFD method has a potential in future estimations with probabilistic model.

### 3.2 Mathematical Model

Consider the epidemiological model of a group of people in the Valencia community of Spain studied by Santonja et al. (2012) on the effects of public health campaigns towards the community weight loss. The selected group of adults at the age of 23 years old is divided into three subpopulations according to their BMI, which is calculated as BMI = weight/height$^2$ (Santonja et al., 2012):

- $N$ – Subpopulation of individuals with normal weight (BMI < 25)
- $S$ – Subpopulation of individuals with overweight (25 ≤ BMI < 30)
- $O$ – Subpopulation of individuals with obesity (BMI ≥ 30)

The transition of individual weight is observed at the age of 24 years old in all subpopulations of $N$, $S$ and $O$ which is a year prior to the public health campaign organized in 2000. The initial subpopulations of individuals at the age of 23 years old with normal weight, overweight and obesity are denoted by $N(t=0) = N_0 = 0.704$, $S(t=0) = S_0 = 0.250$ and $O(t=0) = O_0 = 0.046$ (Santonja et al., 2012) respectively where $t$ is a time

unit in week with $t = 0$ is the last week of the year 2000. The average subpopulation proportions are represented by the terms $\mu_1 N$, $\mu_1 S$ and $\mu_1 O$ respectively. It is further assumed that when the individuals reach 65 years old, they will leave the system in all subpopulations. Due to unhealthy social lifestyle, the rate of transformation from normal weight subpopulations to overweight and obese subpopulations is nonlinearly modelled by $\beta_1 N\,(S + O)$ while the rate of transformation of individuals from overweight to obese subpopulation is a linear term modelled by $\gamma_1 S$. Moreover, individuals in $S$ transit to $N$ at the rate of $\rho = \rho_1 \times \rho_2 \times p$ while the individuals in $O$ transit to $S$ at the rate of $\varepsilon = \varepsilon_1 \times \varepsilon_2 \times p$. The size of the $S$ and $O$ subpopulations are modelled by $\rho S$ and $O$ respectively. Hence the transition model is presented as a nonlinear system of ODEs with respect to independent time variable, $t$ (week) as follows (Santonja et al., 2012):

$$N'(t) = \mu_1 N_0 - \mu_1 N(t) - \beta_1 N(t)\big(S(t) + O(t)\big) + (\rho_1 \times \rho_2 \times p)S(t), \tag{3.1}$$

$$S'(t) = \mu_1 S_0 + \beta_1 N(t)\big(S(t) + O(t)\big) - (\mu_1 + \gamma_1 + \rho_1 \times \rho_2 \times p)S(t) + (\varepsilon_1 \times \varepsilon_2 \times p)O(t), \tag{3.2}$$

$$O'(t) = \mu_1 O_0 + \gamma_1 S(t) - (\mu_1 + \varepsilon_1 \times \varepsilon_2 \times p)O(t). \tag{3.3}$$

Other parameters of the model (Santonja et al., 2012) are listed in Table 3.1 with further estimation details. This model is a system of ODEs with random variable solutions depending on time $t$. The model parameters are also random variables with estimated values. This is where the MMCFD method can be very useful to solve the stochastic model numerically. More details for data source are mentioned in section 2.5.

**Table 3.1:** Model parameters provided by Santonja et.al (2012)

| Parameters | Remarks | Parameters | Remarks |
|---|---|---|---|
| $N_0 = 0.704$ | Proportion of normal weight at 23-year old in 2000. | $p$ =1/7 (per week) | Average overweight or obese individual needs to reduce a mean of 7 kg weight to transit to normal weight or overweight. |
| $S_0 = 0.25$ | Proportion of overweight at 23-year old in 2000. | $\rho$ | Rate that overweight individuals turn into normal weight. |
| $O_0 = 0.046$ | Proportion of obese at 23-year old in 2000 | $\rho_1 = 0.020$ | 2% proportion of overweight population who do physical activity to reduce weight. |
| $\mu_1$ = 0.0004578 | Average stay time in system of 24-65 years old adults; $\mu_1$=1/2184 $\text{week}^{-1}$ spent time of individuals in system (42 years=2184 weeks). | $\rho_2 = 0.042$ | 4.2% proportion of overweight population who improve diet to reduce weight. |
| $\beta_1$= 0.00112 | Transmission rate that depends on unhealthy lifestyles due to social pressure. | $\varepsilon = \varepsilon_1 \times \varepsilon_2 \times p$ | Rate at which obese individual become overweight. |
| $\gamma_1$ = 0.0003226 | Rate at which 24-65 years old overweight adults become obese due to their unhealthy lifestyles. | $\varepsilon_1$ = 0.004 | 0.4% proportion of obese population who do physical activity in order to reduce weight. |
| $1/p$ | Average time an individual needs to return to $S(t)$ from $O(t)$ or $N(t)$from $S(t)$ by diet and physical activity. | $\varepsilon_2$ = 0.024 | 2.4% proportion of obese population who improve diet to reduce weight. |

### 3.3 Solution Approach

Upon substitution of the Eq. (2.12) into the nonlinear system of ODEs (3.1)-(3.3) gives:

$$\frac{N_{i+1} - N_{i-1}}{2h} = \mu_1 N_0 - \mu_1 N_i - \beta_1 N_i (S_i + O_i) + (\rho_1 \times \rho_2 \times p) S_i, \tag{3.4}$$

$$\frac{S_{i+1} - S_{i-1}}{2h} = \mu_1 S_0 + \beta_1 N_i (S_i + O_i) - (\mu_1 + \gamma_1 + p_1 \times \rho_2 \times p) S_i + (\varepsilon_1 \times \varepsilon_2 \times p) O_i, \tag{3.5}$$

$$\frac{O_{i+1} - O_{i-1}}{2h} = \mu_1 O_0 + \gamma_1 S_i - (\mu_1 + \varepsilon_1 \times \varepsilon_2 \times p) O_i, \tag{3.6}$$

Where $i = 1,2, \dots, m$ and $m$ is the number of numerical iterations.

All model parameters (Santonja et al., 2012) are simulated by MC method prior to numerical solutions of the system (3.4)-(3.6) by using FD method in MMCFD approach with the last numerical iteration represents the final solutions of the model corresponding to specific time interval (see Table 3.2).

After that, the parameters are re-simulated before they are used to solve the system by FD method again. This process is applied for 100, 500, 1000, 3000 and 5000 times. As a result, the system has 100-5000 final solutions that are used to calculate the average of them. This average is called in this thesis as "the MMCFD final solutions" (the mean of Monte Carlo finite difference final solutions) that represents the system solution. These mean values correspond to the average of the last iteration of Monte Carlo finite difference solutions in each repeated simulation cum the average of the last component of final solutions vector (the end values in the ordered random sample for $N(t)$, $S(t)$ and $O(t)$) are listed in Table 3.2. The flow chart of MMCFD procedure is further presented in Figure 2.1.

We compute the time $t$ as a unit in week by using Wolfram Alpha program where the number of weeks from the beginning of 2001 to the end of 2013 is calculated as 678 weeks, from the beginning of 2001 to the end of 2015 is 782 weeks, and from the beginning of 2001 to the end of 2030 is 1565 weeks. The number of weeks in each interval is the same number of iterations. Random predictions in 2013 and over the next few years in 2015 and 2030 can be performed by assuming all model parameters to have uniform probability distributions on given prediction interval since these parameters are random variable sampling too. In this case, the outcome results are called as "numerical simulated solutions" by means of MMCFD method. Since we have previous results for this model that are estimated by a statistical method with estimated parameters (Santonja et al., 2012), then comparison between the present Monte Carlo finite difference final solutions with the predicted results (Santonja et al., 2012) can be conducted.

All MMCFD computations are done in MATLAB environment. Some graphical results (for 5000 simulations and the number of numerical iterations is 678 weeks in 2013, 782 weeks in 2015, 1565 weeks in 2030) are plotted using MagicPlot software while the box plots are drawn by using S-plus software. In recapitulation, MMCFD is beneficial to solve linear and nonlinear ODEs system for actual stochastic model that has multiple parameters as random variables and provides the system numerical simulated solutions whenever preliminary data of the model are given previously. All the statistical indicators explained in Chapter 2 and further discussed in the next section are showing good results with the present numerical simulated solutions of MMCFD method.

### 3.4 Results and Discussion

In this section, existence of unique numerical solution by using MMCFD method is presented for the system (3.1)-(3.3) when constant parameters are considered (Santonja et al., 2012). Because of random variables are included as parameters in this model, the MMCFD method is suggested to generate new numerical simulated solutions for this model so that random variables sampling can be estimated by MC method through a stochastic process. Good inference is obtained when the results of all numerical simulated solutions of MMCFD (the mean of Monte Carlo finite difference final solutions) are proved to be closer results to statistical values (Santonja et al., 2012) than the FD numerical results with difference recurrence of simulations (100, 500, 1000, 3000 and 5000 simulations) in 2013, 2015 and 2030 (see Table 3.2). Since different approaches (numerical and statistical) are employed and due to source of errors, the MMCFD numerical simulated solutions of the system (3.1)-(3.3) are expected to have a percentage of difference with statistical and FD results especially at the end of 2013, 2015 and 2030 for the subpopulations $N(t)$, $S(t)$ and $O(t)$ as clearly supported by Table 3.2, Table 3.3 and Figures 3.1- 3.3 respectively.

**Table 3.2:** Mean of MMCFD final solutions and measure of difference for the subpopulations

| Subpopulation | Method | | 2013 | | 2015 | | 2030 | |
|---|---|---|---|---|---|---|---|---|
| | | | Solutions | $\|E\|$ | Solutions | $\|E\|$ | Solutions | $\|E\|$ |
| $N(t)$ | Santonja et al. in 2012 | | 0.4350 | - | 0.4244 | - | 0.3658 | - |
| | FD | | 0.573849 | 0.138849 | 0.555952 | 0.131552 | 0.445840 | 0.080040 |
| | Present results MMCFD | 100 Sim. | 0.564445 | 0.129445 | 0.543390 | 0.118990 | 0.402395 | 0.036595 |
| | | 500 Sim. | 0.562819 | 0.127819 | 0.541705 | 0.117305 | 0.401929 | 0.036129 |
| | | 1000 Sim. | 0.563993 | 0.128993 | 0.543033 | 0.118633 | 0.403929 | 0.038129 |
| | | 3000 Sim. | 0.564385 | 0.129385 | 0.543536 | 0.119136 | 0.405434 | 0.039634 |
| | | 5000 Sim. | 0.564039 | 0.129039 | 0.543108 | 0.118708 | 0.404293 | 0.038493 |
| $S(t)$ | Santonja et al. in 2012 | | 0.3970 | - | 0.4004 | - | 0.4132 | - |
| | FD | | 0.325999 | 0.071001 | 0.335716 | 0.064684 | 0.388068 | 0.025132 |
| | Present results MMCFD | 100 Sim. | 0.329961 | 0.067039 | 0.341145 | 0.059255 | 0.404574 | 0.008626 |
| | | 500 Sim. | 0.331786 | 0.065214 | 0.343118 | 0.057282 | 0.406629 | 0.006571 |
| | | 1000 Sim. | 0.330428 | 0.066572 | 0.341593 | 0.058807 | 0.404404 | 0.008796 |
| | | 3000 Sim. | 0.329959 | 0.067041 | 0.341030 | 0.059370 | 0.403192 | 0.010008 |
| | | 5000 Sim. | 0.330336 | 0.066664 | 0.341476 | 0.058924 | 0.404109 | 0.009091 |
| $O(t)$ | Santonja et al. in 2012 | | 0.1680 | - | 0.1752 | - | 0.2210 | - |
| | FD | | 0.100152 | 0.067848 | 0.108332 | 0.066868 | 0.166092 | 0.054908 |
| | Present results MMCFD | 100 Sim. | 0.104405 | 0.063595 | 0.114111 | 0.061089 | 0.190559 | 0.030441 |
| | | 500 Sim. | 0.104078 | 0.063922 | 0.113681 | 0.061519 | 0.188752 | 0.032248 |
| | | 1000 Sim. | 0.104245 | 0.063755 | 0.113858 | 0.061342 | 0.188956 | 0.032044 |
| | | 3000 Sim. | 0.104368 | 0.063632 | 0.113973 | 0.061227 | 0.188750 | 0.032250 |
| | | 5000 Sim. | 0.104218 | 0.063782 | 0.113818 | 0.061382 | 0.188732 | 0.032268 |

(Sim.~simulation)

In Table 3.2, solutions of $N(t)$, $S(t)$ and $O(t)$ are tabulated in comparison with results from (Santonja et al., 2012), FD method and MMCFD method. The corresponding measure of difference between FD or MMCFD methods with respect to statistical predictions (Santonja et al., 2012), are also listed. In general, normal weight subpopulation, $N(t)$ of the Valencia community is expected to reduce as the year increases from 2000 (see Table 3.1; $N_0 = 0.704$, $S_0 = 0.250$, $O_0 = 0.046$) to 2030 while the subpopulations of overweight $S(t)$ and obese $O(t)$ are growing over the same period. It is observed that the measure of difference of the MMCFD method are consistently smaller than the measure of difference of the FD method in all years of 2013, 2015 and 2030 under various repeated simulations (100, 500, 1000, 3000 and 5000 simulations). Although the current results depend on 5000 simulations, in order to fit with estimations by Santonja et al. (2012), different numbers of repeated simulations are tested in Table 3.2 to find the best number of simulation that produces the closest results to the predicted values of Santonja et al. (2012). In general, the

statistical behavior is random with MMCFD method and it is noted that MMCFD results are extremely convergent with regard to different number of simulations. It is observed that the mean of final solutions of both normal weight and overweight subpopulations for MMCFD is the most convergent to the predicted values (Santonja et al., 2012) in 500 simulations in all years of 2013, 2015 and 2030. However the MMCFD results of obese subpopulation in these years are the closest to the statistical values (Santonja et al., 2012) by using only 100 simulations.

On the other hand, the absolute growth rate percentage at the end of 2013, 2015 and 2030 as well as the absolute yearly average growth rate percentage as compared to the initial conditions of each subpopulation in the last week of 2000 (Santonja et al., 2012) are presented in Table 3.3. It is found that the normal weight and overweight subpopulations ($N(t)$ and $S(t)$) are decreasing at average rates of 1.6% ($N(t)$) and 2.2% ($S(t)$) per year by statistical prediction (Santonja et al., 2012), at 1.2% ($N(t)$) and 1.8% ($S(t)$) by FD method and at 1.4% ($N(t)$) and 2.1% ($S(t)$) by MMCFD method from the last week of 2000 until the end of 2030. Oppositely, the growth rate of obese subpopulation $O(t)$ is increasing over the 30 years. The average yearly increasing rate for the obese subpopulation is estimated at 12.7%, 8.7% and 10.3% from (Santonja et al., 2012), FD and MMCFD solutions respectively. The statistical predictions by Santonja et al. (2012) produce the highest growth rates for all subpopulations and years; the FD method gives the minimum rates while the MMCFD method provides the intermediate rates between the two different approaches of statistical and numerical methods.

**Table 3.3:** Absolute growth rates and absolute average yearly growth rates of the subpopulations

| Subpop. | 2000 Initial | Method | \|Growth\| (%) | | | \|Average Growth\| (%) | | |
|---|---|---|---|---|---|---|---|---|
| | | | 2013 | 2015 | 2030 | 2013 | 2015 | 2030 |
| $N(t)$ | 0.704 | Santonja et al. in 2012 | 38.2 | 39.7 | 48.0 | 2.9 | 2.6 | 1.6 |
| | | FD | 18.5 | 21.0 | 36.7 | 1.4 | 1.4 | 1.2 |
| | | MMCFD* 5000 Sim. | 19.9 | 22.9 | 42.6 | 1.5 | 1.5 | 1.4 |
| $S(t)$ | 0.250 | Santonja et al. in 2012 | 58.8 | 60.2 | 65.3 | 4.5 | 4.0 | 2.2 |
| | | FD | 30.4 | 34.3 | 55.2 | 2.3 | 2.3 | 1.8 |
| | | MMCFD* 5000 Sim. | 32.1 | 36.6 | 61.6 | 2.5 | 2.4 | 2.1 |
| $O(t)$ | 0.046 | Santonja et al. in 2012 | 265.2 | 280.9 | 380.4 | 20.4 | 18.7 | 12.7 |
| | | FD | 117.7 | 135.5 | 261.1 | 9.1 | 9.0 | 8.7 |
| | | MMCFD* 5000 Sim. | 126. 6 | 147.4 | 310.3 | 9.7 | 9.8 | 10.3 |

Note that * are the present results.

According to Figure 3.1, normal weight curves for both FD and MMCFD methods are gradually declining. The decrement of $N(t)$ is slightly higher in 2015 than in 2013. It is also noted that the normal weight subpopulations continue to fall until it approaches the predicted value (Santonja et al., 2012) at the end of 2030. This convinces the expectation that $N(t)$ will decrease in future. In general, the normal weight curves for MMCFD are decreasing significantly than the normal weight curves of FD from 2001 to 2030.

It is clear that, MMCFD of normal weight subpopulation curves converge to the predicted values of (Santonja et al., 2012) faster than the FD curves for all 30 years especially at the end of 2030, (see Figure 3.1).

Based on Figure 3.2, there is a gradual growing of the overweight curves for both FD and MMCFD methods from 2001 until the end of 2030. The curves of $S(t)$ are higher in 2015 than in 2013. Moreover, the overweight subpopulation curves significantly increase before they approach the predicted value (Santonja et al., 2012) at the end of 2030. In other words, $S(t)$ is expected to increase in the next years with the overweight curves of MMCFD method are higher rising than the FD curves from 2001 to the end of 2030. Obviously, the

overweight subpopulation curves of MMCFD method are closer to the predicted values (Santonja et al., 2012) than the FD curves in all years.

According to Figure 3.3, the obese curves are rising progressively in 2013. Also, the progress is more evidenced in 2015 as compared to 2013. Moreover, $O(t)$ is approaching the predicted value (Santonja et al., 2012) closely at the end of 2030. In conclusion, $O(t)$ is expected to increase in the next few years with the obese curves of MMCFD are consistently higher than the FD curves from 2001 to the end of 2030. Apparently, the obese subpopulation curves of MMCFD are closer to the predicted solutions (Santonja et al., 2012) than the FD curves in all years. Even though all curves of normal weight, overweight and obese starting at the same initial points in the beginning of 2001, however only MMCFD solutions for all $N(t)$, $S(t)$ and $O(t)$ are closer to the given predicted points (Santonja et al., 2012) at the end of 2013, 2015 and 2030 as compared to the FD solutions.

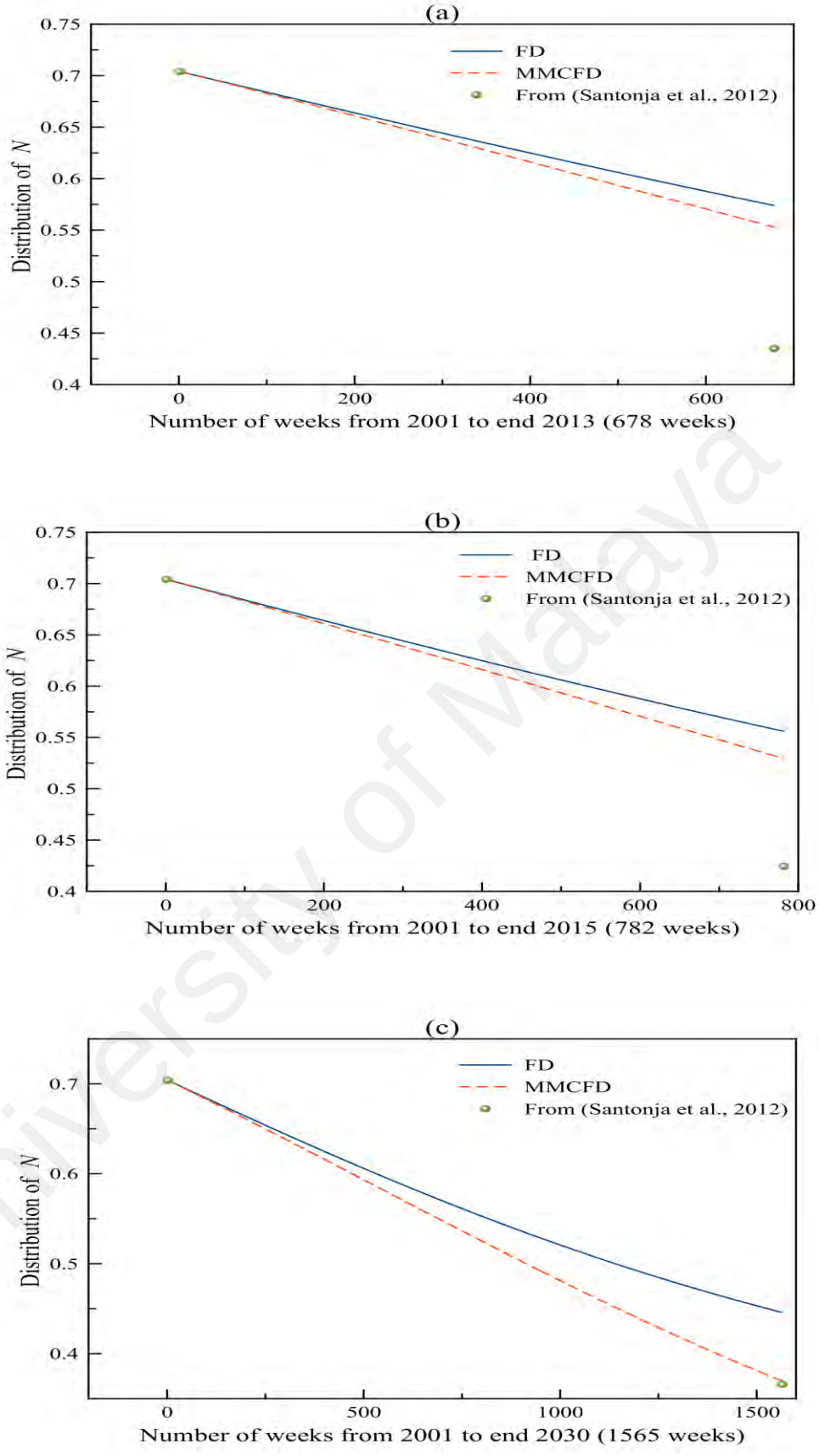


**Figure 3.1:** Numerical solutions using FD and a sample solution of MMCFD of $N(t)$ with 5000 simulations from 2001 to the end of: (a) 2013 (b) 2015 (c) 2030

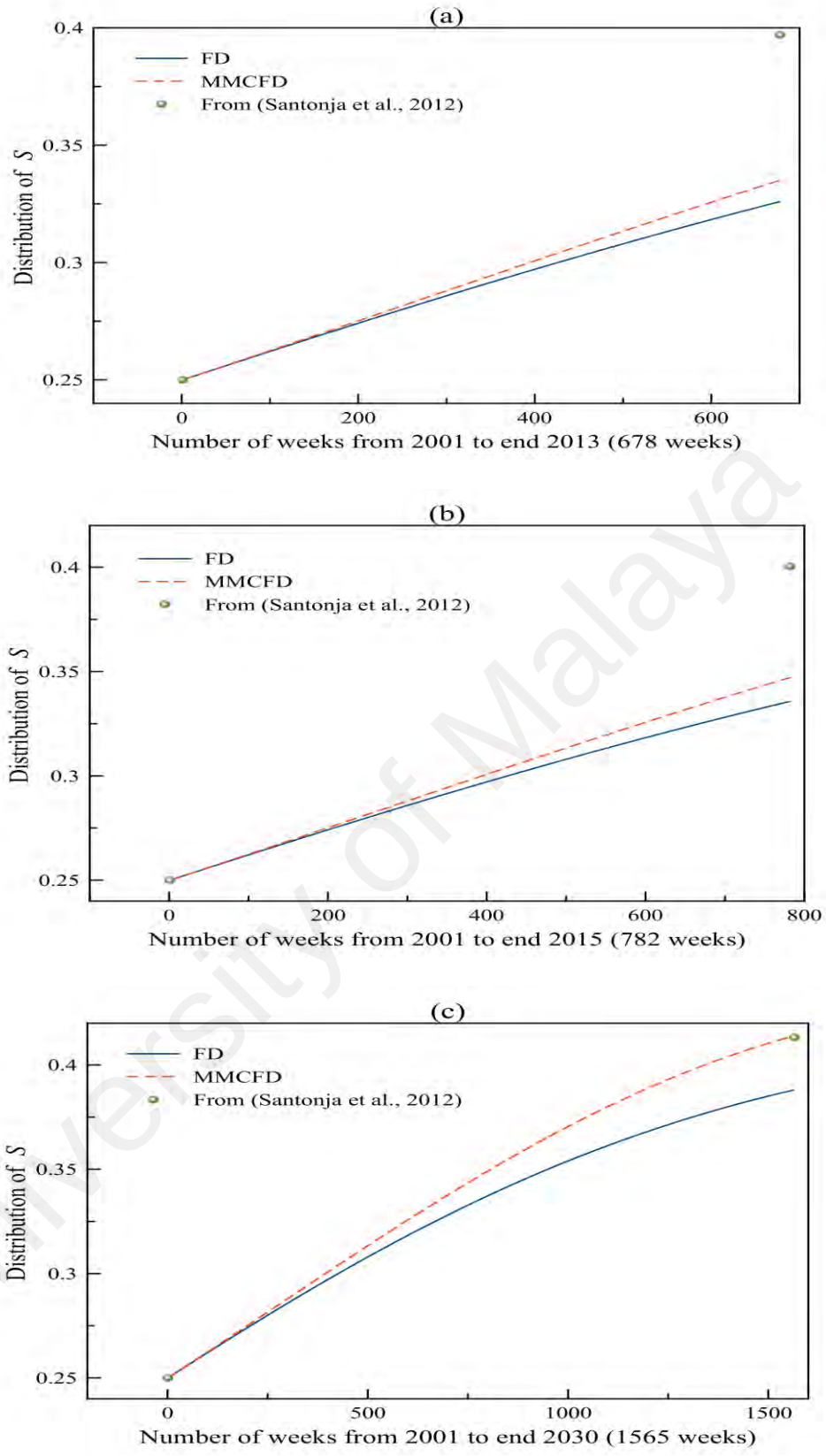


**Figure 3.2:** Numerical solutions using FD and a sample solution of MMCFD of $S(t)$ with 5000 simulations from 2001 to the end of: (a) 2013 (b) 2015 (c) 2030

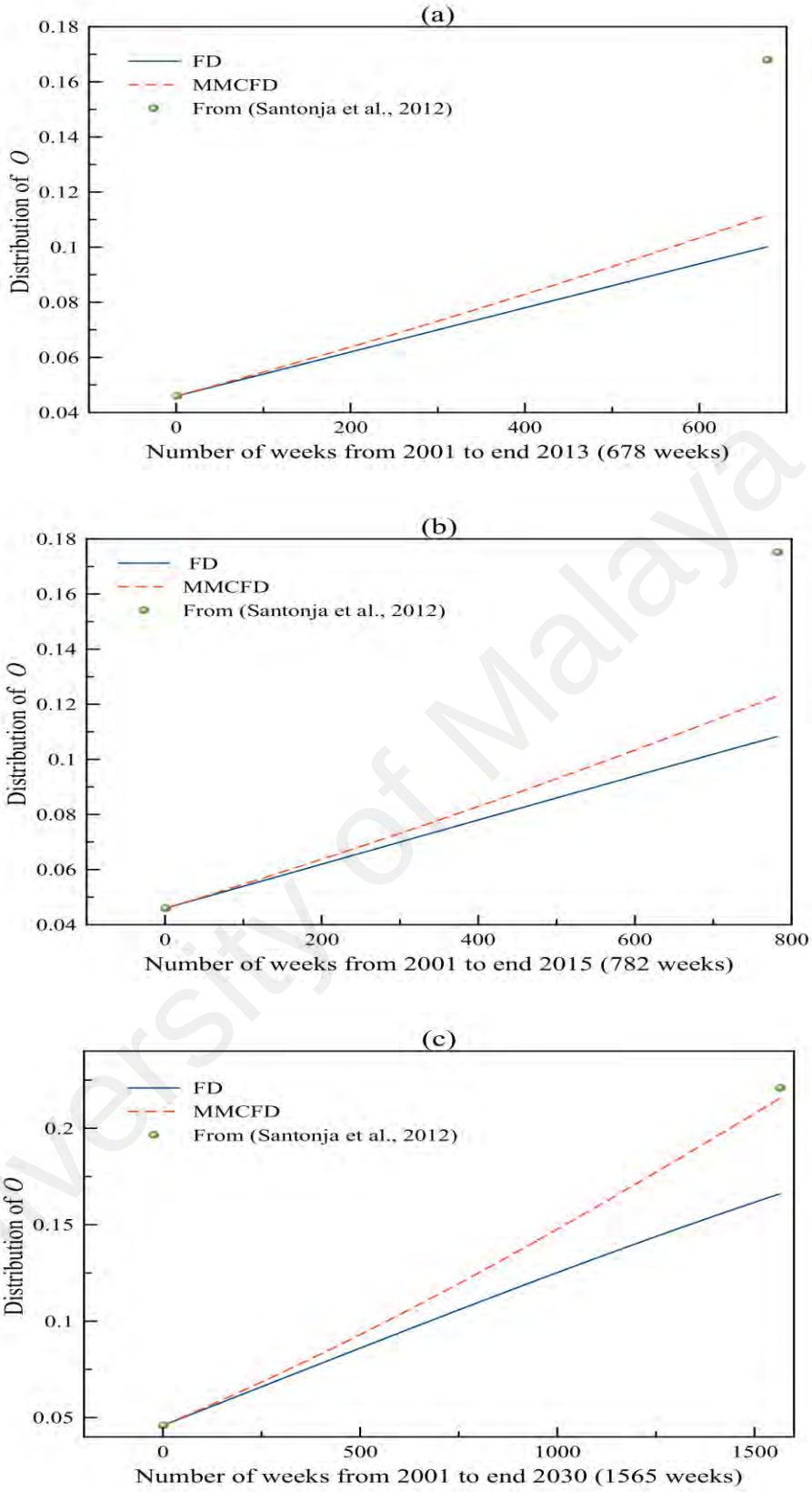


**Figure 3.3:** Numerical solutions using FD and a sample solution of MMCFD of $O(t)$ with 5000 simulations from 2001 to the end of: (a) 2013 (b) 2015 (c) 2030

It can be concluded that the random error from simulation causes difference between FD and MMCFD results and the error increases with more iterations in the future as showed in Figures 3.1-3.3. On the other hand, the convergence between the present numerical simulated results and the given statistical predictions computed as in Eq.(2.30) corresponds to our expectation and agrees with the existing research (Santonja et al., 2012) for decreasing $N(t)$ and increasing $S(t)$ and $O(t)$ in the future. As a consequence, MMCFD numerical simulated solutions of the model are better approximation than FD numerical solutions towards the statistical predictions in all time intervals considered.

Based on Table 3.4, standard deviation of mean, $\sigma_n$ in equations (2.31)-(2.34) of MMCFD for all subpopulations until 2030 is so small. Therefore, this indicator signifies an acceptable error (see (Jacoboni & Lugli, 1989)) for present work. From 2001 until the end of 2030 at 5000 simulations, standard deviation of error for normal weight subpopulation is not more than 0.006730, while the standard deviation of error for the overweight individual drops to 0.005075. At the same time, the standard deviation of error for obese falls to 0.003191. The important point to be addressed here is the MMCFD standard deviation of mean, $\sigma_n$ has a very low percentage of error as can be seen in Table 3.4. Finally, $\sigma_n$ increases with years as a result of statistical simulation errors.

**Table 3.4:** Standard deviation of mean of MMCFD using 5000 simulation for each subpopulation

| Subpopulation | Standard Deviation of Mean | 2013 | 2015 | 2030 |
|---|---|---|---|---|
| $N(t)$ | $\sigma_n$ | 0.002987 | 0.003470 | 0.006730 |
| $S(t)$ | $\sigma_n$ | 0.002789 | 0.003171 | 0.005075 |
| $O(t)$ | $\sigma_n$ | 0.001008 | 0.001204 | 0.003191 |

The prediction intervals of MMCFD solutions $N(t)$, $S(t)$ and $O(t)$ are approaching the range of intervals by Santonja et al. (2012) with a percentage of error. This is due to the

difference in the sample size of model solutions, the method used to solve the differential equations and the difference in simulation technique to estimate parameters (see Table 4 (Santonja et al., 2012) and Tables 3.5).

In general, many factors can cause significant errors in the present study. Firstly, modelling error might happen in the system (3.1)-(3.3) as given in (Santonja et al., 2012) where this system represents the epidemiological model on the effectiveness of public health campaigns towards body weight loss. This error comes from prediction and actual data. Secondly, the sampling and rounding errors come from simulation method to estimate model parameters that are later being transported to predictions. The third source of the error is due to different simulation techniques to estimate model parameters. In the present work, the parameters are simulated by using the classical Monte Carlo process while Santonja et al. (2012) previously simulated the parameters by using Latin Hypercube Sampling (LHS) technique that is another type of Monte Carlo sampling. LHS is applied by Santonja et al. (2012) to generate 5000 different values of each model at the same time. Finally, the error comes from the type of distribution (uniform distribution) for parameters and solutions of the model (Santonja et al., 2012). Also, the numerical method is subjected to two common numerical errors; the round off error which losing of precision due to computer rounding of decimal quantities and data propagation error. Therefore, the error is not restricted and extends in all interval areas of the distribution.

**Table 3.5:** Prediction interval ($5^{th}$ percentile, $95^{th}$ percentile) for MMCFD solutions of $N(t)$, $S(t)$ and $O(t)$ using 5000 simulation at the end of 2013, 2015 and 2030

| Subpopulation | 2013 | 2015 | 2030 |
|---|---|---|---|
| $\boldsymbol{N(t)}$ | (0.5297, 0.5970) | (0.5030, 0.5815) | (0.3226, 0.4806) |
| $\boldsymbol{S(t)}$ | (0.2994, 0.3625) | (0.3063, 0.3781) | (0.3475, 0.4635) |
| $\boldsymbol{O(t)}$ | (0.0931, 0.1162) | (0.1004, 0.1283) | (0.1542, 0.2291) |

Because of random errors (sampling and rounding errors) in the particular model, the solutions of $N(t)$, $S(t)$ and $O(t)$ can be given by prediction interval come from niform distributions in terms of random variables $N(t)$, $S(t)$ and $O(t)$. The prediction interval consists of percentiles of final solutions for each model variable $N(t)$, $S(t)$ and $O(t)$ running at 5000 evaluations, empirically at 5% and 95% percentiles. These percentiles are used to account 90% confidence interval that implies 90% of the predicted values inside this confidence interval. Consider prediction intervals of distributions for $N(t)$, $S(t)$ and $O(t)$ contain a lower limit ($5^{th}$ percentile) and an upper limit ($95^{th}$ percentile). According to MMCFD method with 5000 simulations, the $95^{th}$ percentile for $N(t)$ predicted values inside the population is variously distributed for different years (see Table 3.5 and Figures 3.4-3.6). In recapitulation, we have chosen the $5^{th}$ and $95^{th}$ percentile because we want to get a 90% confidence interval for our numerical simulated solutions as an alternative to the 90% confidence interval of statistical predictions given by Santonja et al. (2012). Although the $5^{th}$ and $95^{th}$ percentiles are selected in the present work, we can also use other percentiles depending on the size of the confidence interval.

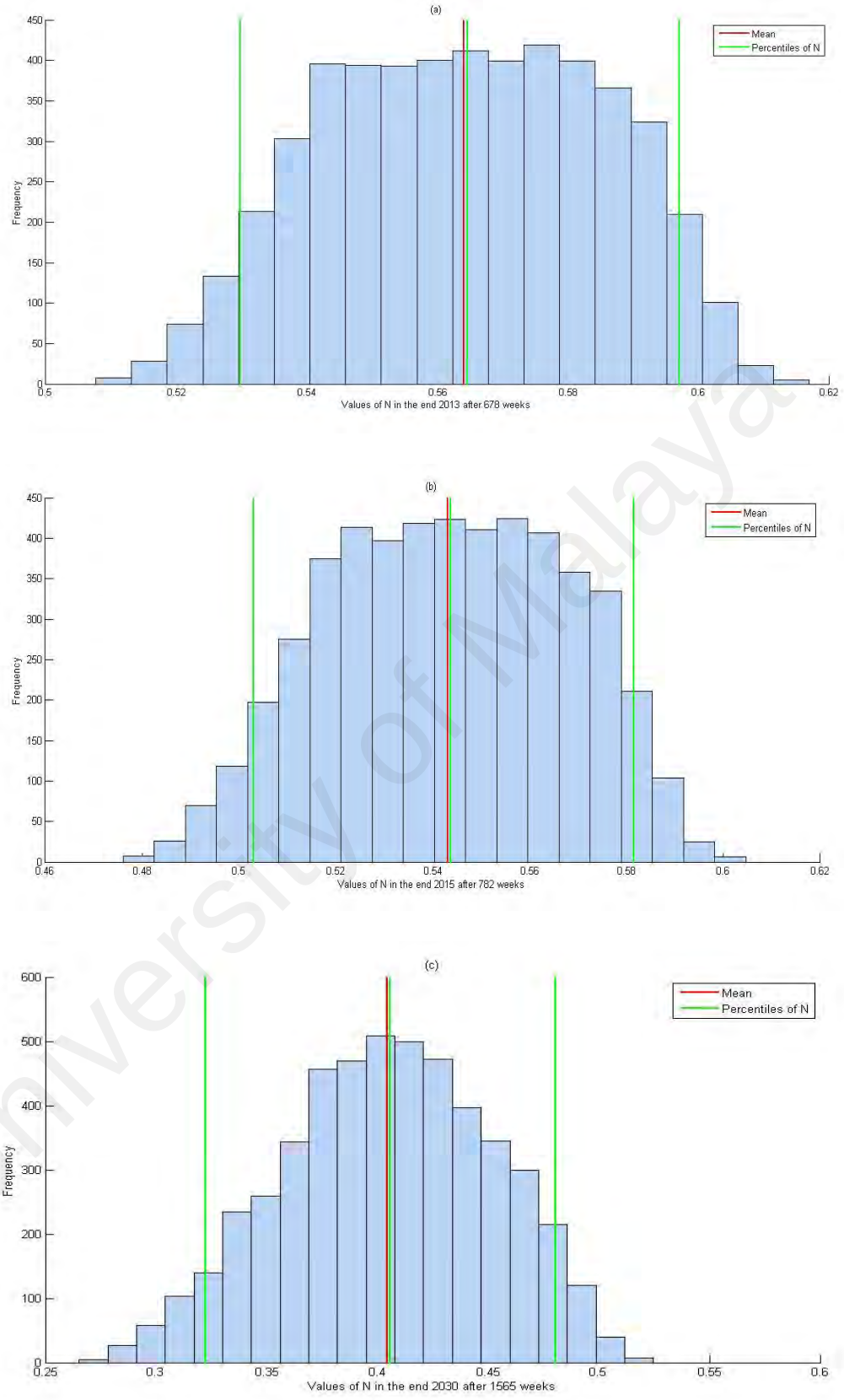


**Figure 3.4:** The 5th, 50th and 95th percentiles for MMCFD solutions of $N(t)$ with 5000 simulations from 2001 to the end of : (a) 2013 (b) 2015 (c) 2030

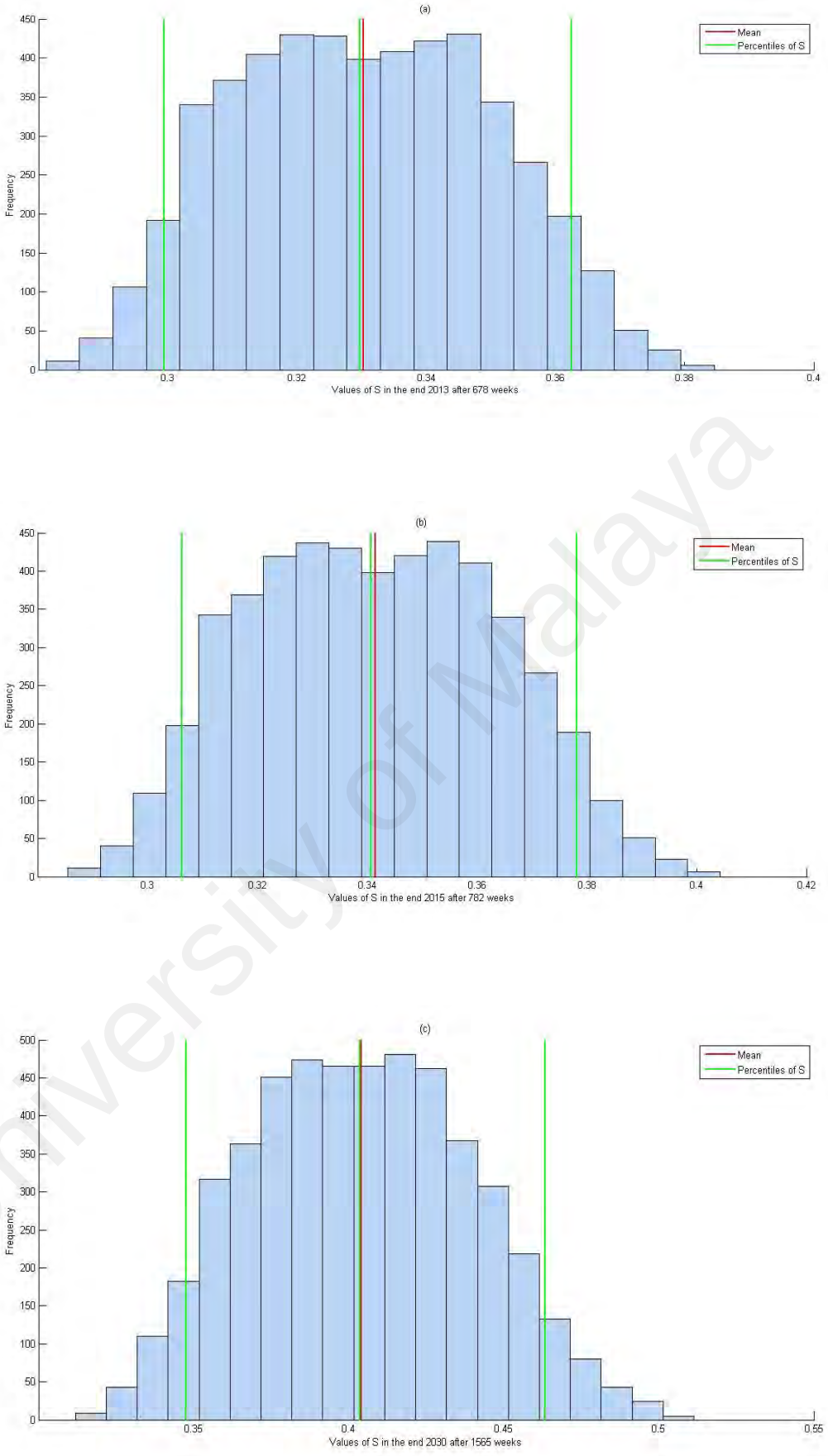


**Figure 3.5:** The 5th, 50th and 95th percentiles for MMCFD solutions of $S(t)$ with 5000 simulations from 2001 to the end of : (a) 2013 (b) 2015 (c) 2030

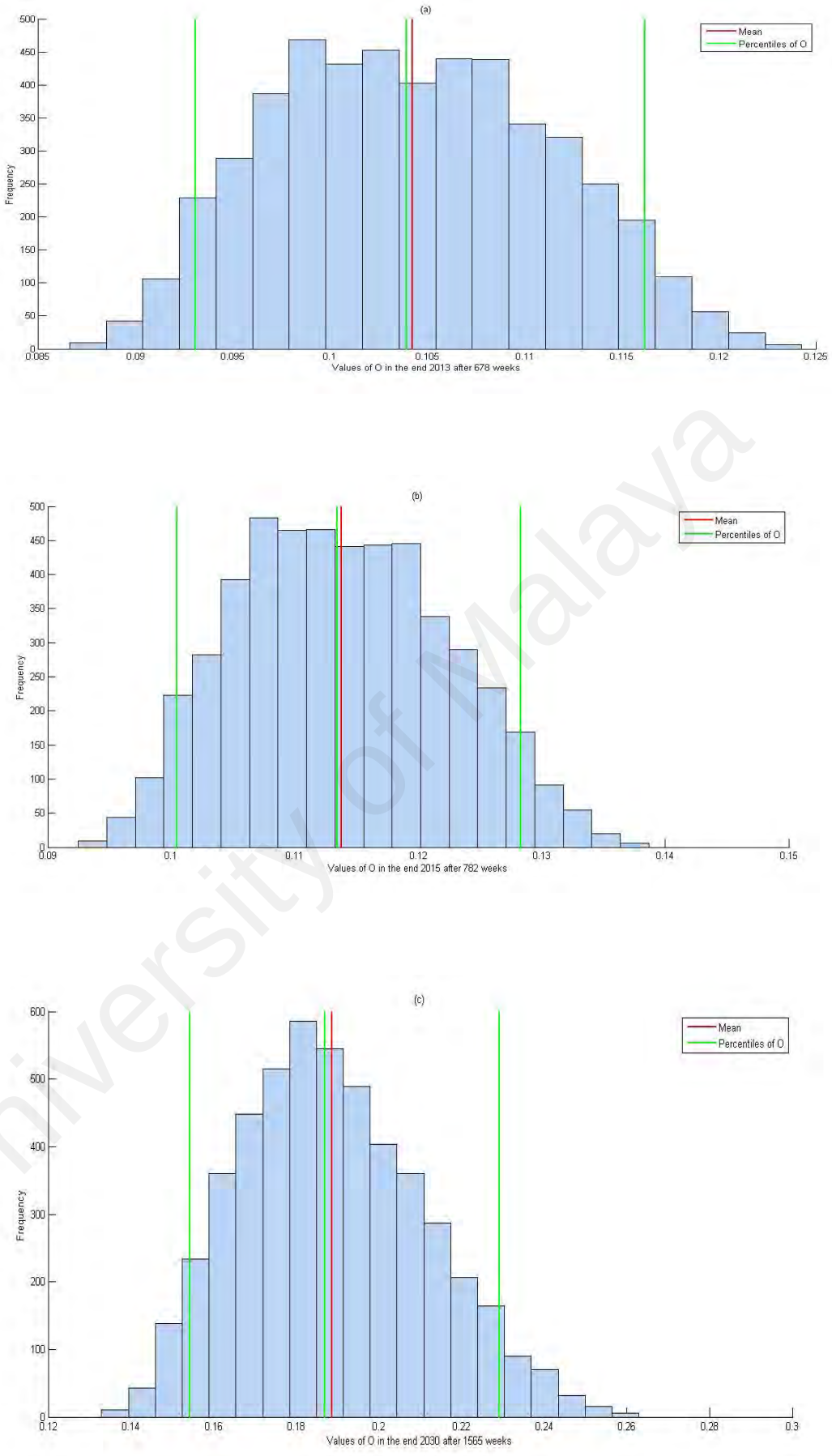


**Figure 3.6:** The 5th, 50th and 95th percentiles for MMCFD solutions of $O(t)$ with 5000simulations from 2001 to the end of : (a) 2013 (b) 2015 (c) 2030

Comparing between the present MMCFD results and the statistical prediction intervals (Santonja et al., 2012), Santonja et al. (2012) presented 90% confidence intervals for $N(t)$, $S(t)$ and $O(t)$ as listed in Table 4 (Santonja et al., 2012, p. 37) which are taken from the empirical at 5th and 95th percentiles. The current MMCFD-prediction interval consists of percentiles of $N(t)$, $S(t)$ and $O(t)$ empirically at 5th and 95th percentiles too. Both prediction intervals of solutions to this model are obtained by using 5000 simulations of model parameters. It is clear that all MMCFD solutions in Table 3.2 and 3.6 with 5000 simulations are inside the prediction intervals of percentile as presented in Table 3.5. In addition, the 50th percentile as the median for MMCFD solution distributions is close to the mean of these solutions as shown in all Figures 3.4 to 3.6 from 2001 to the end of 2013, 2015, and 2030, respectively.

**Table 3.6:** 50th percentile (median) of MMCFD results using 5000 simulations with the mean

| Subpopulation | Percentile | 2013 | 2015 | 2030 |
|---|---|---|---|---|
| $N(t)$ | **50th, Median** | 0.564596 | 0.543749 | 0.405707 |
| | **Mean** | 0.564039 | 0.543108 | 0.404293 |
| $S(t)$ | **50th, Median** | 0.329789 | 0.340697 | 0.403530 |
| | **Mean** | 0.330336 | 0.341476 | 0.404109 |
| $O(t)$ | **50th, Median** | 0.103911 | 0.113443 | 0.186972 |
| | **Mean** | 0.104218 | 0.113818 | 0.188732 |

Box plots analyzing the results of $N(t)$, $S(t)$ and $O(t)$ from 2001 to the end of 2030 are provided graphically in Figure 3.7. They show the central point (median), the extreme values (maximum and minimum values) and the outliers which are the points outside the boundaries. The box plots present the pattern of the numerical simulated solutions of the ODE system $N(t)$, $S(t)$ and $O(t)$ throughout the 30 years.

Using 5000 simulations, the boxplot of $N(t)$ lays above in 2013, it declines slightly in 2015, but it declines significantly in 2030 (with one outlier point at bottom). That means the subpopulation values of $N(t)$ take smaller values in the end of 2030 as compared to 2013

and 2015 (see median in Table 3.6). Therefore, $N(t)$ is a decreasing function in the future (see Figure 3.7 (a)).

Inversely, the boxplots of $S(t)$ and $O(t)$ are located at the bottom in 2013, then they rise slightly in 2015. There is a major hike for $S(t)$ and $O(t)$ in 2030 more than 2015. There is only one outlier value above the box for $S(t)$ in 2030 but more outliers are at the top part of the box for $O(t)$ in 2030. That means, the subpopulation values of $S(t)$ and $O(t)$ have greater values at the end of 2030 than in 2013 and 2015 (see median in Table 3.6). Therefore, $S(t)$ and $O(t)$ are both increasing functions in the future (see Figures 3.7 (b) and (c)). The statistical behaviors shown in Figures 3.1-3.3 and in Figure 3.7 confirmed that $N(t)$ will decrease while $S(t)$ and $O(t)$ will increase in the future.

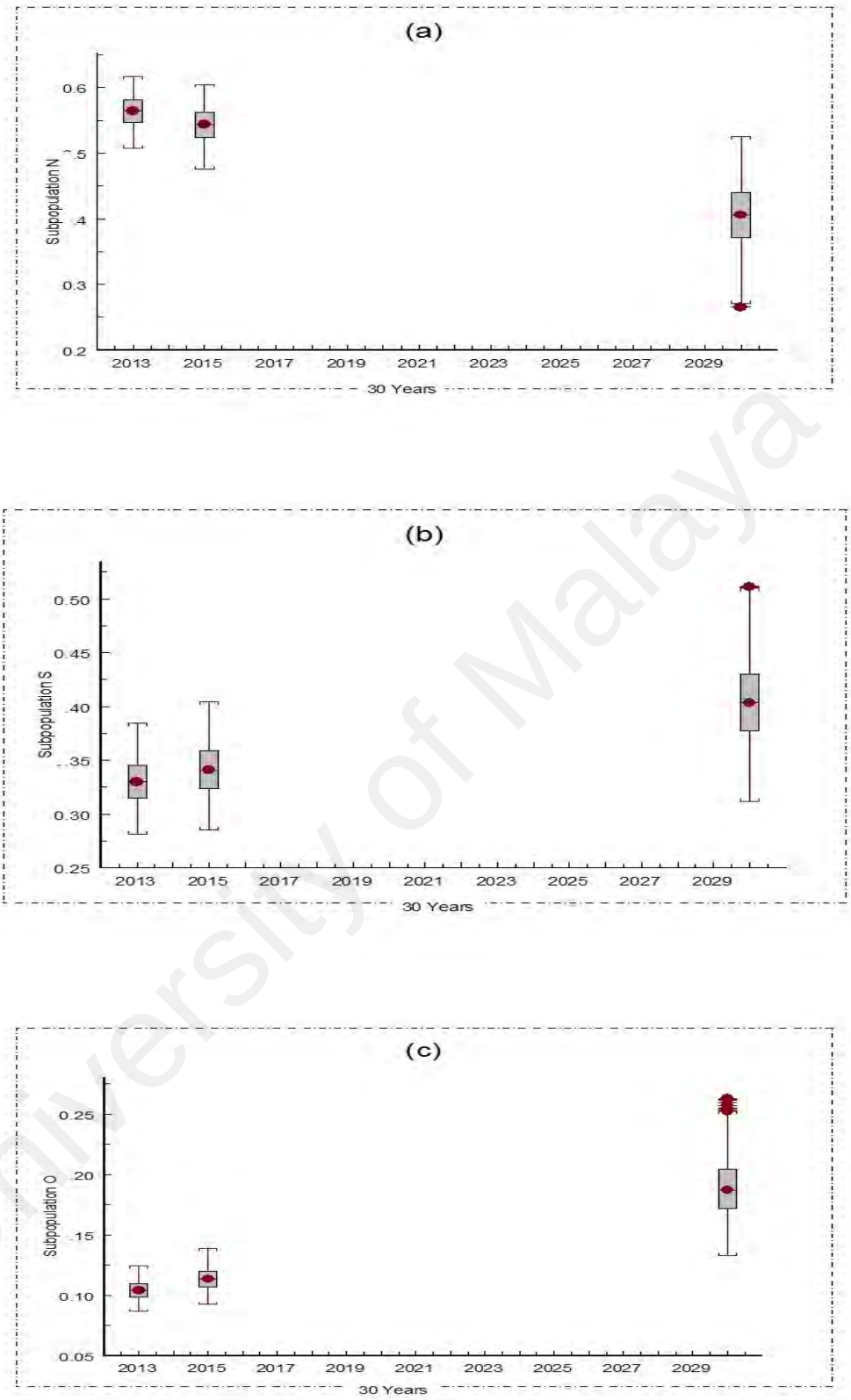


**Figure 3.7:** Box plot along 30 years (2001- 2030) for MMCFD solutions with 5000 simulations of: (a) $N(t)$ (b) $S(t)$ (c) $O(t)$

### 3.6 Conclusion

In this study, the Mean Monte Carlo Finite Difference (MMCFD) method is proposed for the first time to solve an epidemiology model explicating the effects of public health campaign on body weight loss in Spain. The MMCFD method is suggested to generate new numerical simulated solutions for this model so that random variables sampling can be estimated by MC method through a stochastic process.

Due to the similar nature of the MMCFD method and the statistical estimation method, it is found that the MMCFD method produced numerical simulated solutions of the model with better approximation towards previous statistical predictions as compared to the FD method in all time intervals considered. Through this work, it is shown that the MMCFD method is beneficial to generate random coefficient sampling for the nonlinear stochastic epidemic system. Besides, the MMCFD method is also promising to create agreement and estimation balance between a statistical method and a classical numerical method.

# CHAPTER 4: MEAN LATIN HYPERCUBE FINITE SOLUTIONS FOR COCAINE ABUSE IN SPAIN

## 4.1 Introduction

Mathematical epidemic model is necessary to analyze the contributing mechanisms behind the nature of an infectious disease in order to control its spread and to predict probability for it to happen in future. An epidemic can be defined as a sudden outbreak of a disease that happens firstly to those who were exposed to the disease. The epidemic populations are divided into Susceptible ($S$), Exposed ($E$), Infectious ($I$), Recovered ($R$) as the mentioned in Chapter 2, section 2.4. The recent study of SEIR disease dynamics is examined by Jiang et al. (2016). Many types of epidemic models have been developed such as SIS, SIR and SEIR, with or without time delays, of which some of them have been discussed by Shu et al. (2012), Safi and Garba (2012), Guerrero et al. (2013), Witbooi (2013), Wang et al. (2014), De la Sen et al. (2015), Bai (2015) and Xu et al. (2015). Many researchers have provided numerical solutions of such time-dependent models. One of the common methods used is finite difference (FD) method (Gustafsson, 2011) while one of the common software applications is MATLAB (Shampine et al. 2003).

Latin hypercube sampling (LHS) is an extension of stratified sampling for the multidimensional data distribution. It is a variance-reduction technique statistical methods for generating random samples, (Gentle, 2003). The main approach of LHS that use in this Chapter can be outlined here briefly. Firstly the curve of the cumulative probability distribution (CDF) is divided into equal intervals which are uniform domains. Next, a random value is selected from each interval of the probability distribution. The difference between MC and LHS is mentioned in Chapter 2, section 2.5.4. Some preliminary concepts of simulation are provided by Dagpunar (2007).

Nakayama (2011) estimated confidence interval (CI) using FD formula by updating the simulation using LHS. Nakayama (2012) also developed the LHS to increase statistical

efficiency after the CI is simulated using FD method. Budgaga et al. (2016) modified the model simulation parameters by employing dimensionality reduction technique using LHS to forecast the simulation outcomes in real time. On the other hand, Pedro et al. (2016) employed the technique of Latin hypercube sampling to evaluate simulation of the basic reproduction number, $R_0$ with 5000 sets of sample parameters for a mosquito population.

Sánchez et al. (2011) used the LHS stratified random sampling technique to estimate model parameters and to create the prediction interval for each parameter in the Spain cocaine abuse model. Later Guerrero and Vazquez-Leal (2014) solved this model by using the homotopy analysis method (HAM). Then, the model is further tackled analytically using HAM-Pade and multi-stage HAM-Pade approximations which are applied on the year intervals of [0, 50], [50, 150] and $[150, \infty]$. A network epidemiological mathematical model is studied by Santonja et al. (2010) for short-term evolution of the cocaine individual consumers. The sensitivity of the model parameters is examined. The increasing trends of cocaine abuse in Spain from majority of subpopulations are observed. Santonja et al. (2010) used LHS to generate variation of 5000 values for the parameters which have uniform probability distribution on the intervals [0, 0.085], [0.06, 0.2], [0.32, 0.72], [0, 4] and [0, 0.19] respectively. The 90% confidence interval which is the mean value of 5000 simulations is achieved for the cocaine abuse model. On the other hand, Monzó (2015) used LHS technique to predict the profile of each subpopulation of the cocaine consumers by 95% confidence interval for few years until 2020.

Cocaine abuse is considered as a socially epidemiological model. Therefore, the growth of the cocaine abuse population must be monitored and analyzed closely. The importance of this study is due to real evolution of the prevalence of cocaine consumed and because the harmful effects of cocaine to one's health happen quickly. An individual who takes cocaine will eventually become more dependent on the cocaine over the time if no precaution is taken. Therefore, the study of the model is crucial to avoid the spread of the epidemic that leads to sickening addiction. The purpose of the current study is to predict the evolution of a

social cocaine abuse habit as an epidemic relative to the previous predictions by Sánchez et al. (2011). The numerical simulation solutions for the cocaine abuse in Spain are obtained by using a modified mean process that combines the numerical deterministic method FD and the statistical random sampling technique LHS. This modified approach simulates the model parameters by LHS technique before integration with FD iterations. The last iteration of the FD numerical result selected from each repetition is called as the final solution of FD. The mean of the FD final solutions of the current process is considered as the solution for this system. This approach is named in short as Mean Latin Hypercube Finite Difference (MLHFD) method. It is introduced to solve a set of nonlinear first order ordinary differential equations (ODEs) by using random sampling of equal probability distribution.

The current technique is an upgrade of the work by Mohammed et al. (2016) by modifying the Mean Monte Carlo Finite Difference (MMCFD) method to benefit from LHS. The stratified Latin hypercube sampling process allows faster random variation of the numerical simulation results to be obtained. That means, the simulated results produced are closer to the statistical values via fewer number of numerical FD iterations as compared to the previous MMCFD approach by Mohammed et al. (2016). Moreover, the advantage of the proposed MLHFD technique over the deterministic methods such as FD and HAM is that it has properties that can provide probability distribution of the randomized solutions for the nonlinear ODE system. The MLHFD process can be used in solving the epidemiological models to describe dynamics of disease and to explain future behavior of the system under search when the data needed for a particular interval of time are not sufficient to derive accurate deterministic solutions. This chapter is organized as follows: In Section 4.2, the application of SEIR epidemic model of the cocaine abuse problem in Spain is discussed. Subsequently in Section 4.3, the methodology approaches; finite difference (FD) method of the cocaine abuse model is presented. In the same section, the MLHFD method is explained. Next, the analysis and discussion of the numerical simulation results

obtained for the cocaine abuse model are done in Section 4.4. Finally, the overall conclusion of the study is provided in Section 4.5.

## 4.2 SEIR Model of Cocaine Abuse

Cocaine abuse is considered as a socially transmitted disease. The probability of cocaine epidemic transmission depends on social contact or social peer pressure that usually influences the spread of such disease. This model was presented by Sánchez et al. (2011) with real data of the cocaine abuse in Spain in a period of time of ten years (1995 - 2005). It was applicated successfully for short term prediction up to 2015. The mathematical model of the cocaine abuse consists of four subpopulations represent the proportions of the total cocaine abuse population. It is described by the following nonlinear system of ordinary differential equations (Sánchez et al., 2011):

$$n'(t) = \mu_2\big(1 - n(t)\big) - \beta_2 n(t)\big(o(t) + r(t) + h(t)\big) + \varepsilon_c h(t), \tag{4.1}$$

$$o'(t) = \beta_2 n(t)\big(o(t) + r(t) + h(t)\big) - \gamma_2 o(t) - \mu_2 o(t), \tag{4.2}$$

$$r'(t) = \gamma_2 o(t) - \sigma_c r(t) - \mu_2 r(t), \tag{4.3}$$

$$h'(t) = \sigma_c r(t) - \mu_2 h(t) - \varepsilon_c h(t), \tag{4.4}$$

where $n(t)$, $o(t)$, $r(t)$ and $h(t)$ are model variables depending on time $t$.

Further descriptions and initial values of the cocaine abuse subpopulations are given in Table 4.1. Since the total size of constant population has been normalized to unity, the variables satisfy (Diekmann & Heesterbeek, 2008):

$$n(t) + o(t) + r(t) + h(t) = 1, \tag{4.5}$$

such that the equations (4.1)-(4.4) have the following region (Diekmann & Heesterbeek, 2008):

$$\varphi = \{(n, o, r, h) \in R_+^4 : n > 0, o \geq 0, r \geq 0, h \geq 0, n + o + r + h = 1\}. \quad (4.6)$$

The population assumes a constant; the birth rate which is equal to the death rate. Other model parameters that influence the spread of the cocaine abuse in Spain are also described in Table 4.2. The same values of the parameters by Sánchez et al. (2011) are used in the present work in order to imitate the behavior of the cocaine abuse but by using the MLHFD method up to 50 years prediction (1995 – 2045). More details for data source of the cocaine abuse model are mentioned in section 2.5.

Consider the cocaine abuse model as the SEIR model which has a latent duration and suppose the latent duration reflects the time before cocaine harmfully affects its consumers' health or before the consumers are causing harm to themselves due to excessive cocaine abuse etc. This latent duration can be shown in the subpopulations of Exposed ($E$) and Infectious ($I$). Similarly in the standard SEIR model, $n(t)$ (non-users) represents the proportion of individuals who are susceptible to cocaine abuse but do not use cocaine. $o(t)$ (occasional users) represents the proportion of individuals who are taking cocaine every now and then but they may become cocaine addicts in the future if mixed with regular users or habitual users. $r(t)$ (regular users) represents the proportion of individuals who accepted usual cocaine abuse in their lives. At the same time, they are the same individuals who are infectious to $n(t)$ or $o(t)$, yet this class does not suffer from damage to their health (this case is not addictive). $h(t)$ (habitual users) represents the proportion of individuals who are excessively addicted to cocaine and therefore they suffer from damage to their health. This class is considered as extremely cocaine-abused and typically they die or in some cases they transit back to the regular users, occasional users or non-users classes.

**Table 4.1:** Basic information on variables and real initial proportions values for the subpopulations of the cocaine abuse model in Spain at the beginning of the year 1995 (Sánchez et al., 2011)

| Variables of model | Classification variables | Variables description | Initial variables | Initial values |
|---|---|---|---|---|
| $\boldsymbol{n(t)}$ | Non-users | the proportion of individuals who have never taken cocaine. | $n_0$ | 0.944 |
| $\boldsymbol{o(t)}$ | occasional users | the proportion of individuals who take cocaine sometimes in their lives. | $o_0$ | 0.034 |
| $\boldsymbol{r(t)}$ | regular users | the proportion of individuals who usually abuse cocaine in their lives. | $r_0$ | 0.018 |
| $\boldsymbol{h(t)}$ | habitual users | the proportion of individuals who extremely abuse cocaine until they become addicted. | $h_0$ | 0.004 |

**Table 4.2:** Parameters' descriptions and estimated proportions values (Sánchez et al., 2011)

| Parameters of model | Parameter values ($\mathbf{year^{-1}}$) | Parameter description |
|---|---|---|
| $\boldsymbol{\mu_2}$ | 0.01 | birth /death rate |
| $\boldsymbol{\beta_2}$ | 0.09614 | transmission rate of cocaine abuse due to social pressure |
| $\boldsymbol{\gamma_2}$ | 0.0596 | the rate at which an occasional user becomes a regular user |
| $\boldsymbol{\sigma_c}$ | 0.0579 | the rate at which a regular user becomes a habitual user |
| $\boldsymbol{\varepsilon_c}$ | 0.0000456 | the rate at which a habitual user leaves cocaine due to therapy course |

## 4.3 Methodology

Mean Latin Hypercube Finite Difference (MLHFD) method is a new modified numerical simulation technique proposed by integrating two methods of different natures together; a statistical simulation process with random sampling, LHS and a numerical deterministic approach, FD. The mean of the solutions resulting from the methods' integration process is taken as the alternative solution for the cocaine abuse model with

random variables/parameters. All these approaches are discussed in Chapter 2 while the application of MLHFD method on the cocaine abuse model is explained in this section.

LHS process can be applied to simulate all parameters of the cocaine abuse model simultaneously. Each parameter is considered as a random vector of time-dependent random variables which have uniform distribution on a created interval that has the form $(a - 0.2a, a + 0.2a)$ where $a$ is the previously predicted value of the parameter (Sánchez et al., 2011) with 20% variation range. More details on this are given by Santonja et al. (2012). In our model, we consider five parameters that must be simulated by LHS. Finally, the random sample of each parameter has the number of LHS simulation values which are created by MATLAB software. LHS is presented with different sample sizes (100, 1000 and 5000 repetitions) to estimate parameters for the initial value problem of the nonlinear real cocaine abuse model and to further solve it. In this chapter, the MLHFD method is fully executed using MATLAB software and the algorithm is presented as a flow chart in Figure 2.2. Previously, the cocaine abuse model was solved by Santonja et al. (2010), Sánchez et al. (2011) and Monzó (2015) using the mathematical built-in code NDSolve[ ].

Applying FD method to the cocaine abuse model (4.1)-(4.4), the following expressions are firstly derived by using backward finite difference formula:

$$\frac{n_1 - n_0}{w} = \mu_2(1 - n_0) - \beta_2 n_0(o_0 + r_0 + h_0) + \varepsilon_c h_0, \tag{4.6}$$

$$\frac{o_1 - o_0}{w} = \beta_2 n_0(o_0 + r_0 + h_0) - \gamma_2 o_0 + \mu_2 o_0, \tag{4.7}$$

$$\frac{r_1 - r_0}{w} = \gamma_2\, o_0 - \sigma_c r_0 - \mu_2 r_0, \tag{4.8}$$

$$\frac{h_1 - h_0}{w} = \sigma_c r_0 + \mu_2 h_0 - \varepsilon_c h_0. \tag{4.9}$$

After the first solutions $n_1$, $o_1$, $r_1$ and $h_1$ are obtained, the central difference formula as in Eq. (2.12) is employed to generate:

$$\frac{n_{i+1} - n_{i-1}}{2w} = \mu_2(1 - n_i) - \beta_2 n_i(o_i + r_i + h_i) + \varepsilon_c h_i, \tag{4.10}$$

$$\frac{o_{i+1} - o_{i-1}}{2w} = \beta_2 n_i(o_i + r_i + h_i) - \gamma_2 o_i + \mu_2 o_i, \tag{4.11}$$

$$\frac{r_{i+1} - r_{i-1}}{2w} = \gamma_2 o_i - \sigma_c r_i - \mu_2 r_i, \tag{4.12}$$

$$\frac{h_{i+1} - h_{i-1}}{2w} = \sigma_c r_i + \mu_2 h_i - \varepsilon_c h_i, \qquad \text{for} \quad i = 1,2, \dots m, \tag{4.13}$$

where $m$ is the number of FD numerical iterations.

## 4.4 Results and Discussion

In this section, the predicted values are arranged from the smallest time value (in a year) until the largest time value of the distribution. In this study, the 5$^{th}$ and 95$^{th}$ percentiles are computed to get a 90% prediction interval for the MLHFD numerical simulated solutions. These percentiles are dependent on the size of the prediction interval. The prediction interval obtained takes into account the empirical 5% and 95% percentiles. The predicted values inside the prediction interval of all the subpopulations of non-users $n(t)$, occasional users $o(t)$, regular users $r(t)$ and habitual users $h(t)$ are obtained for 5000 repetitions. This is to match with the previous works by Santonja et al. (2010), Sánchez et al. (2011) and Monzó (2015) who estimated the cocaine abuse model parameters using stratified random sampling LHS technique for 5000 runs, while considering these parameters as unknown. Similarly these parameters are treated as random variables which have uniform distribution on the interval in the present work.

**Table 4.3:** Prediction intervals (5th percentile, 95th percentile) for MLHFD solutions of $n(t), o(t), r(t)$ and $h(t)$ with the step size, $w$ = 1 (yearly)

| Subpopulation | (100 repetitions) | (1000 repetitions) | (5000 repetitions) |
|---|---|---|---|
| **From 1995 to 2015 ($t \leq 20$)** | | | |
| $n(t)$ | (0.7120625, 0.8247000) | (0.7106284, 0.8240017) | (0.7100279, 0.8237870) |
| $o(t)$ | (0.0941251, 0.1709375) | (0.0948229, 0.1765881) | (0.0940196, 0.1772913) |
| $r(t)$ | (0.0439347, 0.0788158) | (0.0443243, 0.0764917) | (0.0447349, 0.0762423) |
| $h(t)$ | (0.0314621, 0.0461554) | (0.0308476, 0.0456079) | (0.0307725, 0.0452352) |
| **Subpopulation** | **(100 repetitions)** | **(1000 repetitions)** | **(5000 repetitions)** |
| **From 1995 to 2045 ($t \leq 50$)** | | | |
| $n(t)$ | (0.1698301,0.4226823) | (0.1658055,0.4196748) | (0.1649787,0.4192537) |
| $o(t)$ | (0.2306565,0.3329301) | (0.2362721,0.3331346) | (0.2333103,0.3366635) |
| $r(t)$ | (0.1539181,0.2669994) | (0.1531605,0.2591247) | (0.1527628,0.2589150) |
| $h(t)$ | (0.1672605,0.3048938) | (0.1667958,0.3051621) | (0.1698970,0.3054480) |

The present work is different from the previous research because each parameter of the model depends on the given initial values with 20% variation. In other words, the current work generates different interval for each parameter such that each MLHFD estimated parameter (as a random variable) has a uniform distribution on the generated interval in the form of $(a - 0.2a, a + 0.2a)$, where $a$ is the predicted value of the parameter from Santonja et al. (2010), Sánchez et al. (2011) and Monzó (2015). They discussed sensitivity analysis of the cocaine abuse model using 90% confidence interval with 5000 mean realizations of the model estimation for each studied year from 1995 until 2013 and 2015. On the other hand, the present work considers the Mean Latin Hypercube Finite Difference (MLHFD) results with varying simulations of 100, 1000 and 5000 respectively as the solutions of the epidemic model. Prediction intervals (5th percentile, 95th percentile) of the obtained random distribution of the MLHFD results are computed as cocaine abuse prediction at 90% empirical confidence intervals from 1995 until 2045 yearly as shown in Table 4.3 where the mean and the 90% predicted values are inside the prediction intervals.

In this section, the MLHFD solutions for the cocaine abuse model in Spain are analyzed and discussed. The previous solutions obtained via statistical estimation (Sánchez et al., 2011) and HAM-Pade approximation (Guerrero & Vazquez-Leal, 2014) are listed in Table 4.4 in comparison with the present FD and MLHFD results. Based on Table 4.4, the MLHFD results for all subpopulations non-users $n(t)$, occasional users $o(t)$, regular users $r(t)$ and habitual users $h(t)$ are the closest to the predicted values by Sánchez et al. (2011) at the end of 2015 for the year interval (0, 20) with the real step size $w$ = 1 (yearly) as compared to $w$ = 0.25 (quarter yearly), $w$ = 0.5 (half yearly) and $w$ = 2 (every 2 years). It is found that by incorporating LHS simulation with numerical FD method gives advantage to MLHFD in reducing number of classical FD iterations for solution of this model.

**Table 4.4:** Solutions for the cocaine abuse model at the end of 2015

| Model Variables | Predicted Values (Sánchez et al., 2011) | HAM-Pade (Guerrero & Vazquez-Leal, 2014) | Step Size, $w$ (year) | FD | Present MLHFD Results | | |
|---|---|---|---|---|---|---|---|
| | | | | | 100 repetitions | 1000 repetitions | 5000 repetitions |
| $n(t)$ | 0.785 | 0.7564470 | 2 | 0.3949467 | 0.4336910 | 0.4338025 | 0.4338027 |
| | | | 1 | 0.7568123 | 0.7698149 | 0.7698045 | 0.7698022 |
| | | | 0.5 | 0.8779570 | 0.8820085 | 0.8820078 | 0.8820073 |
| | | | 0.25 | 0.9166647 | 0.9181969 | 0.9181976 | 0.9181975 |
| $o(t)$ | 0.125 | 0.1397084 | 2 | 0.2820587 | 0.2700598 | 0.2703740 | 0.2705877 |
| | | | 1 | 0.1395866 | 0.1331838 | 0.1333245 | 0.1333881 |
| | | | 0.5 | 0.0728612 | 0.0706692 | 0.0706977 | 0.0707115 |
| | | | 0.25 | 0.0502798 | 0.0494117 | 0.0494171 | 0.0494203 |
| $r(t)$ | 0.055 | 0.0629923 | 2 | 0.1693121 | 0.1583501 | 0.1579137 | 0.1577963 |
| | | | 1 | 0.0629128 | 0.0594605 | 0.0593218 | 0.0592785 |
| | | | 0.5 | 0.0322767 | 0.0312826 | 0.0312524 | 0.0312427 |
| | | | 0.25 | 0.0234631 | 0.0231212 | 0.0231149 | 0.0231126 |
| $h(t)$ | 0.035 | 0.0408056 | 2 | 0.1536825 | 0.1378991 | 0.1379099 | 0.1378132 |
| | | | 1 | 0.0406883 | 0.0375408 | 0.0375492 | 0.0375312 |
| | | | 0.5 | 0.0169051 | 0.0160398 | 0.0160421 | 0.0160385 |
| | | | 0.25 | 0.0095924 | 0.0092703 | 0.0092704 | 0.0092696 |

For comparison purpose, the corresponding absolute approximate errors ($|E_a|$) of FD and MLHFD solutions are shown numerically in Table 4.5. $|E_a|$ represents the absolute value of the difference between FD and MLHFD solutions with the predicted values by Sánchez et al. (2011) in the year interval (0, 20) from 1995 to 2015. The errors of MLHFD

solutions for all subpopulations with various simulation numbers (100, 1000 and 5000) are smaller than the errors of FD results across various step sizes ($w$ = 0.25 (quarter yearly), $w$ = 0.5 (half yearly), $w$ = 1 (yearly) and $w$ = 2 (every 2 years)). Based on Tables 4.4-4.5, MLHFD results for non-users $n(t)$, occasional users $o(t)$, regular users $r(t)$ and habitual users $h(t)$ are closer to the predicted values by Sánchez et al. (2011) than the approximate HAM-Pade and numerical FD results for the system during the 20 years. From Figures 4.1(a)-(c), all MLHFD curves of the non-users $n(t)$, occasional users $o(t)$ and regular users $r(t)$ with 1000 and 5000 repetitions converge to the predicted values (Sánchez et al., 2011) faster than the HAM-Pade and FD curves from 1995 to 2015.

**Table 4.5:** Absolute approximate error, $|E_a|$ for FD and MLHFD solutions as relative the predicted values (Sánchez et al., 2011) at the end of 2015

| **Model variables** | **Step size, $w$ (year)** | **FD ($\|E_a\|$)** | **Present MLHFD Results ($\|E_a\|$)** | | |
|---|---|---|---|---|---|
| | | | **100 repetitions** | **1000 repetitions** | **5000 repetitions** |
| $n(t)$ | 2 | 0.3900533 | 0.3513090 | 0.3511975 | 0.3511973 |
| | 1 | 0.0281877 | 0.0151851 | 0.0151955 | 0.0151978 |
| | 0.5 | 0.0929570 | 0.0970085 | 0.0970078 | 0.0970073 |
| | 0.25 | 0.1316647 | 0.1331969 | 0.1331976 | 0.1331975 |
| $o(t)$ | 2 | 0.1570587 | 0.1450598 | 0.1453740 | 0.1455877 |
| | 1 | 0.0145866 | 0.0081838 | 0.0083245 | 0.0083881 |
| | 0.5 | 0.0521388 | 0.0543308 | 0.0543023 | 0.0542885 |
| | 0.25 | 0.0747202 | 0.0755883 | 0.0755829 | 0.0755797 |
| $r(t)$ | 2 | 0.1143121 | 0.1033501 | 0.1029137 | 0.1027963 |
| | 1 | 0.0079128 | 0.0044605 | 0.0043218 | 0.0042785 |
| | 0.5 | 0.0227233 | 0.0237174 | 0.0237476 | 0.0237573 |
| | 0.25 | 0.0315370 | 0.0318788 | 0.0318851 | 0.0318874 |
| $h(t)$ | 2 | 0.1186825 | 0.1028991 | 0.1029099 | 0.1028132 |
| | 1 | 0.0056883 | 0.0025408 | 0.0025492 | 0.0025312 |
| | 0.5 | 0.0180949 | 0.0189602 | 0.0189579 | 0.0189615 |
| | 0.25 | 0.0254076 | 0.0257298 | 0.0257296 | 0.0257304 |

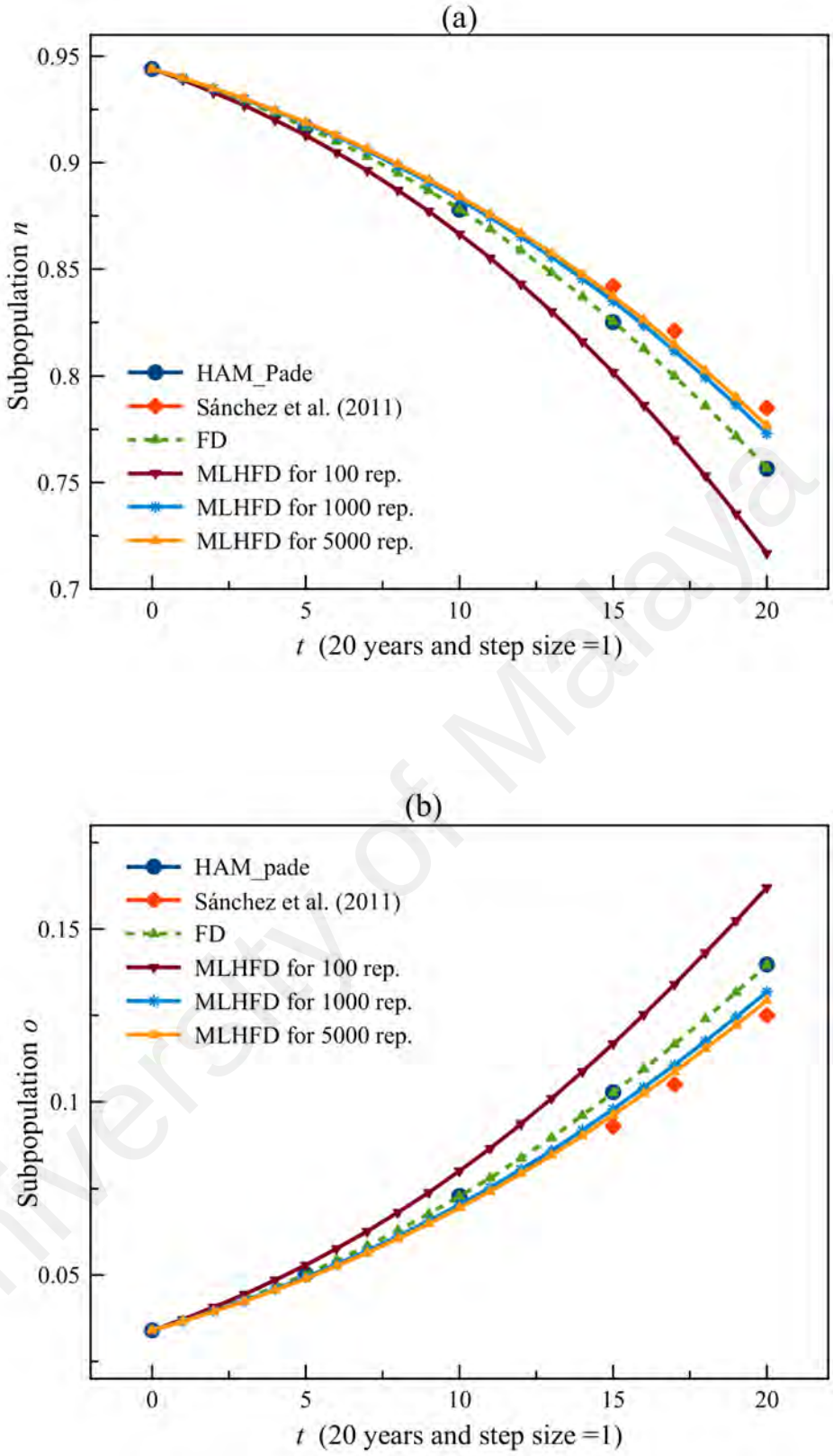


**Figure 4.1:** Variation of solutions between statistical predictions (Sánchez et al., 2011), HAM-Pade (Guerrero & Vazquez-Leal, 2014), FD and MLHFD (100, 1000 and 5000 repetitions) from 1995 to 2015 yearly

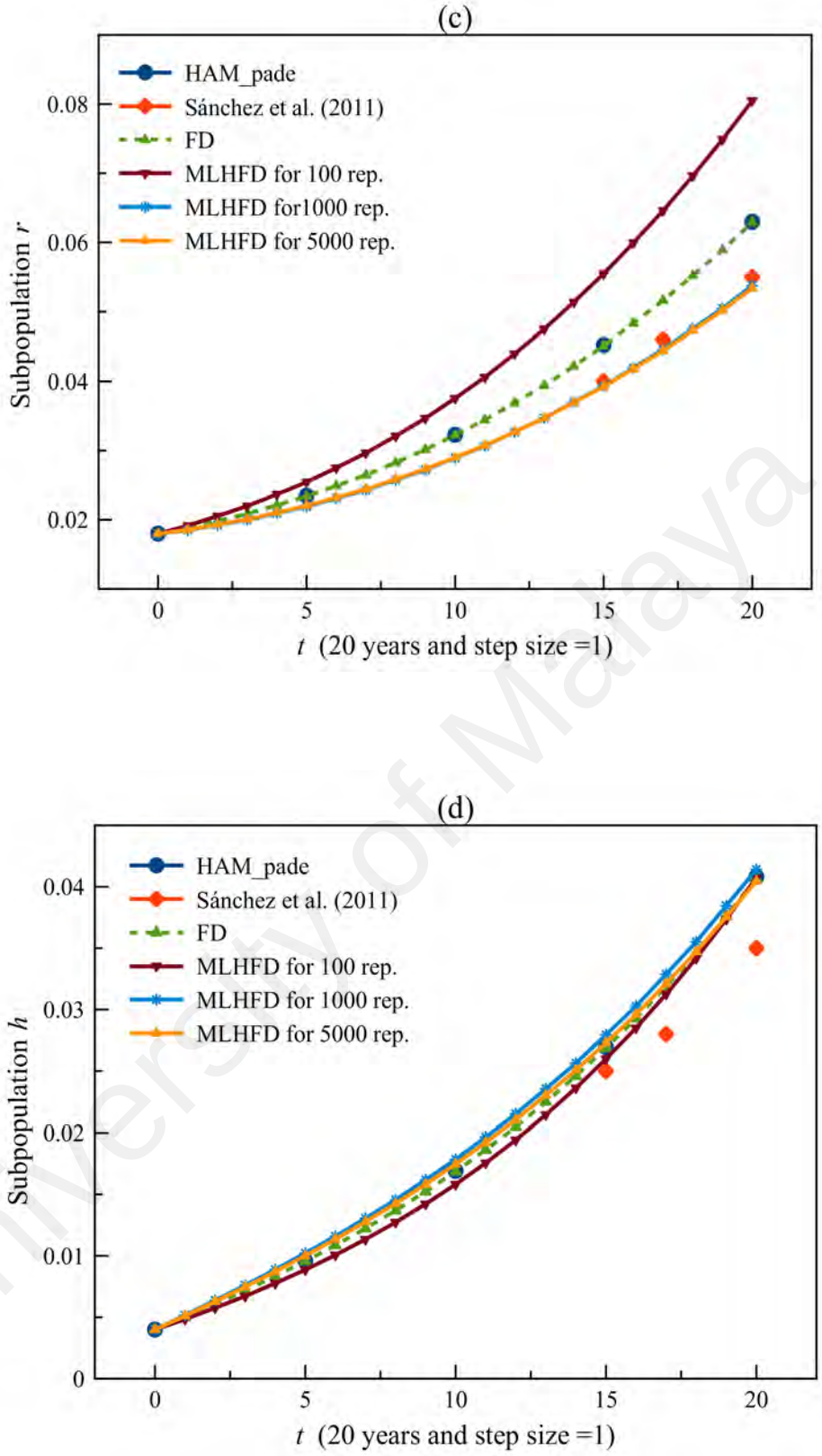


**Figure 4.1, continued** Variation of solutions between statistical predictions (Sánchez et al., 2011), HAM-Pade (Guerrero & Vazquez-Leal, 2014), FD and MLHFD (100, 1000 and 5000 repetitions) from 1995 to 2015 yearly

The MLHFD results of the cocaine abuse model using 100, 1000 and 5000 repetitions with the step size, $w$ = 1 (yearly) are demonstrated in Figure 4.1 for 20 years and in Figure 4.2 for 50 years respectively. It appears that all non-users $n(t)$ curves for HAM-Pade, FD and MLHFD methods are gradually declining in both Figures 4.1 and 4.2. Oppositely, the curves of occasional users $o(t)$, regular users $r(t)$ and habitual users $h(t)$ for all methods are significantly rising in these figures. Except in Figure 4.1(d), all HAM-Pade and FD curves lie in between MLHFD graphs from 1995 to 2015 or 2045.

FD method is used to find the deterministic solution of the epidemic system numerically when the model parameters are constants while MLHFD method is used when random distribution in the model becomes necessary such that the parameters are treated as random variables. The advantage of MLHFD method over FD method is that it can reduce the number of numerical iterations of FD. Previous statistical estimations by Sánchez et al. (2011) were conducted from the year 1995 to 2015 while HAM-Pade deterministic solutions by Guerrero & Vazquez-Leal (2014) described the cocaine abuse model from 1995 to 2045.

Hence, the present study also predicted the behavior of the epidemic for 50 years as demonstrated in Figure 4.2. These present results of MLHFD provide alternative estimations to the model when random distribution of the parameters is taken into account. Additionally, the obtained MLHFD numerical simulation results at the end of 2045 are listed together with other deterministic solutions in Table 4.6.

**Table 4.6** Solutions for the cocaine abuse model at the end of 2045

| Model variables | HAM-Pade (Guerrero & Vazquez-Leal, 2014) | Step size, *w* (year) | FD | Present MLHFD Results | | |
|---|---|---|---|---|---|---|
| | | | | 100 repetitions | 1000 repetitions | 5000 repetitions |
| $\boldsymbol{n(t)}$ | 0.2547923 | 2 | 0.0997666 | 0.1226348 | 0.1229011 | 0.1229287 |
| | | 1 | 0.2549121 | 0.2773194 | 0.2774777 | 0.2774819 |
| | | 0.5 | 0.6732753 | 0.6796014 | 0.6795964 | 0.6795933 |
| | | 0.25 | 0.8533917 | 0.8552158 | 0.8552127 | 0.8552119 |
| $\boldsymbol{o(t)}$ | 0.2877409 | 2 | 0.1792473 | 0.1007323 | 0.1006916 | 0.1004725 |
| | | 1 | 0.2881075 | 0.2828697 | 0.2830872 | 0.2832937 |
| | | 0.5 | 0.1813977 | 0.1787472 | 0.1789779 | 0.1790881 |
| | | 0.25 | 0.0868917 | 0.0860935 | 0.0861456 | 0.0861698 |
| $\boldsymbol{r(t)}$ | 0.2111729 | 2 | 0.0912135 | 0.2559868 | 0.2580461 | 0.2592917 |
| | | 1 | 0.2111649 | 0.2055668 | 0.2051807 | 0.2051177 |
| | | 0.5 | 0.0856571 | 0.0841173 | 0.0838788 | 0.0838054 |
| | | 0.25 | 0.0381966 | 0.0377096 | 0.0376565 | 0.0376397 |
| $\boldsymbol{h(t)}$ | 0.2461309 | 2 | 0.6297725 | 0.5206461 | 0.5183611 | 0.5173071 |
| | | 1 | 0.2458156 | 0.2342441 | 0.2342544 | 0.2341068 |
| | | 0.5 | 0.0596706 | 0.0575342 | 0.0575469 | 0.0575132 |
| | | 0.25 | 0.0215200 | 0.0209811 | 0.0209852 | 0.0209787 |

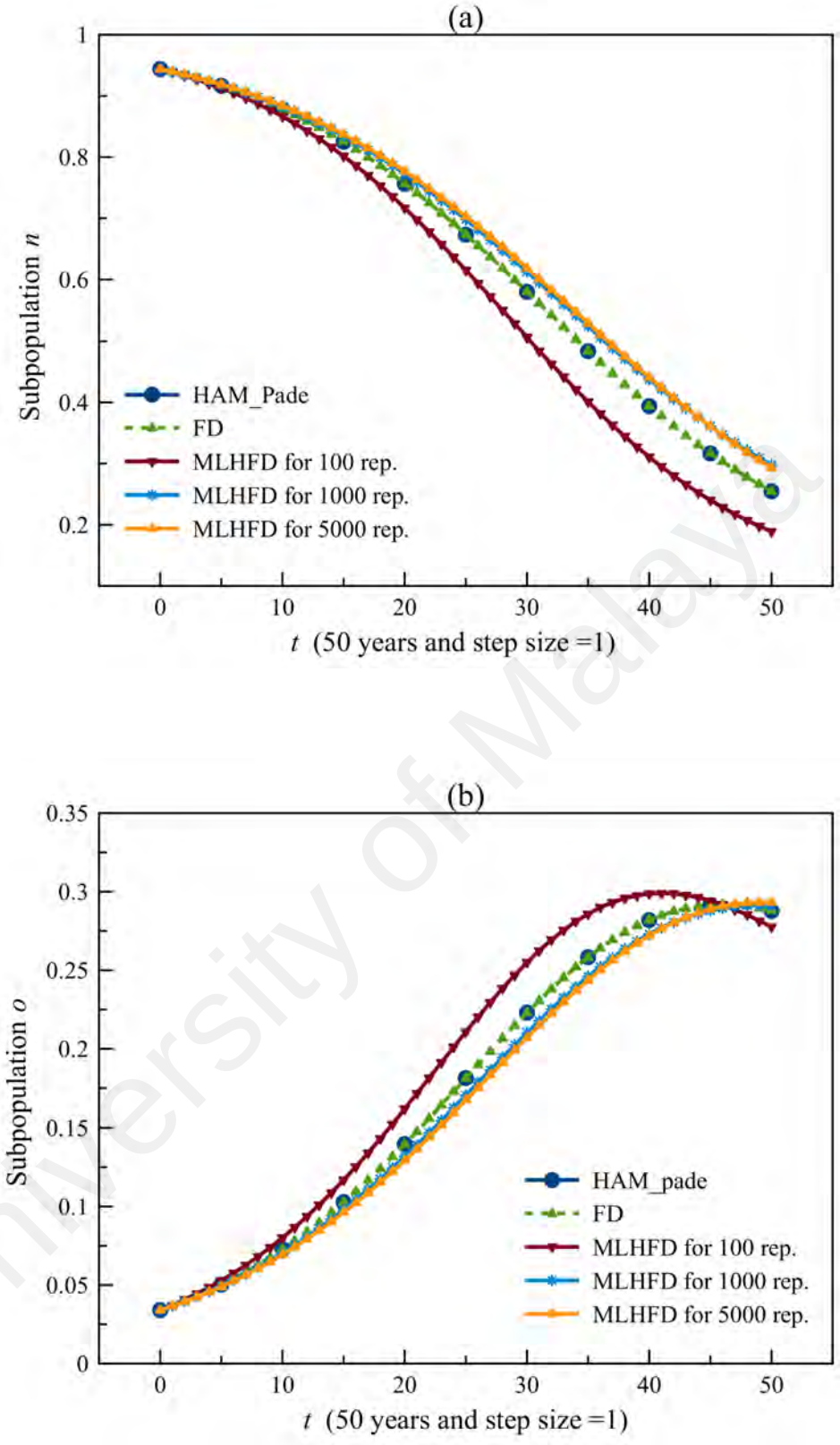


**Figure 4.2:** Variation of solutions between HAM-Pade (Guerrero & Vazquez-Leal, 2014), FD and MLHFD (100, 1000 and 5000 repetitions) from 1995 to 2045 yearly

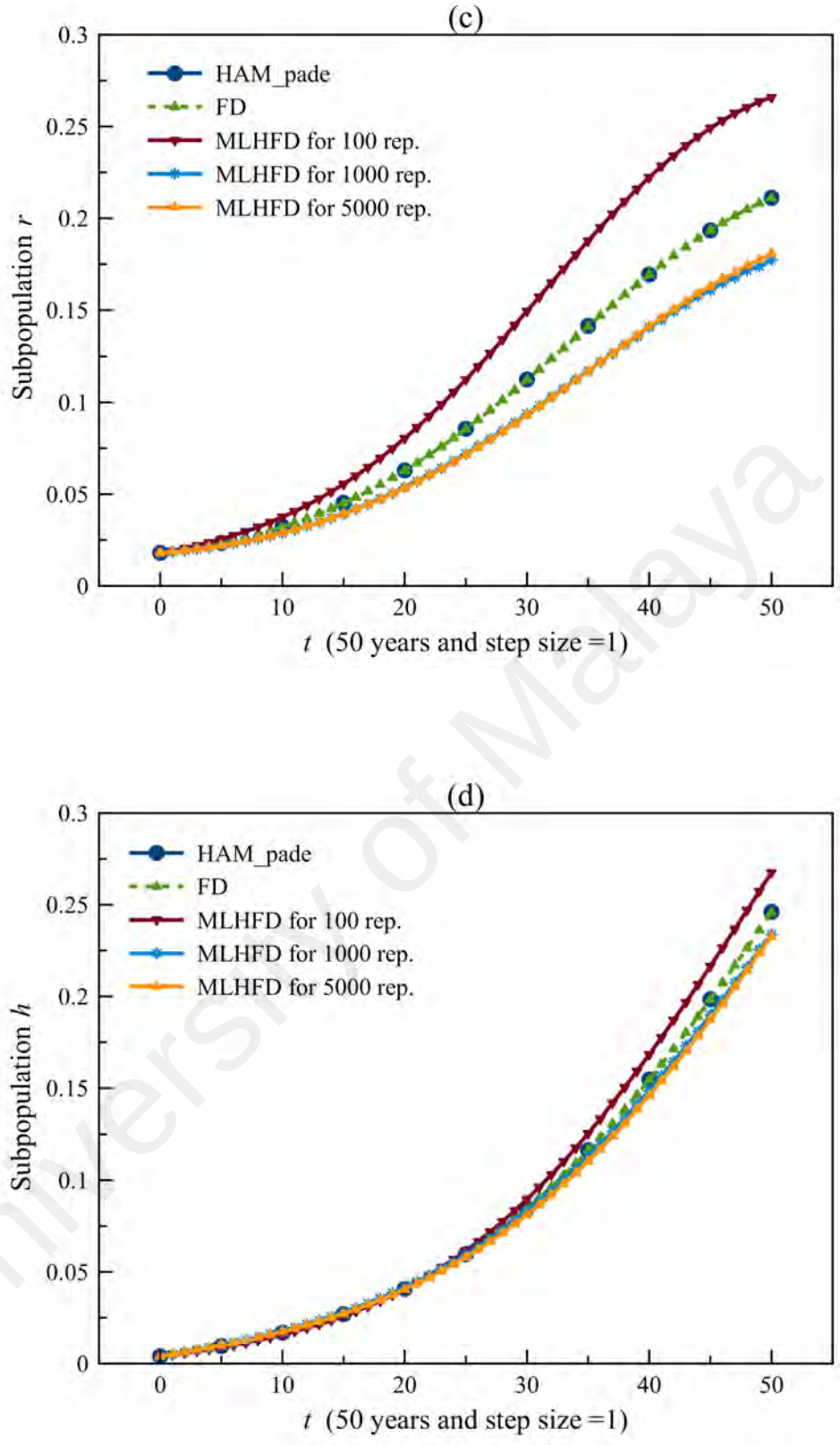


**Figure 4.2, continued:** Variation of solutions between HAM-Pade (Guerrer & VazquezLeal, 2014) FD and MLHFD (100, 1000 and 5000 repetitions) from 1995 to 2045 yearly

The overall MLHFD results for the cocaine abuse model are summarized graphically in Figure 4.3 in the year interval (0, 50) from 1995 with the step size $w$ = 1 (yearly) for all subpopulations; non-users $n(t)$, occasional users $o(t)$, regular users $r(t)$ and habitual users $h(t)$ using 5000 repetitions. The proportion of non-user individuals $n(t)$ is expected to decrease tremendously suggesting that healthy population in Spain will decay in future due to cocaine abuse epidemic. On the other hand, the habitual users $h(t)$ who abuse extra cocaine until they become addicted, are expected to increase yearly during the 50 years. Similar increasing trends are expected to happen to the profiles of occasional users $o(t)$ and regular users $r(t)$. The proposed MLHFD numerical simulation method is suitable to predict the range of solutions of such real model with random variables/parameters than the classic numerical FD and approximate HAM-Pade methods which are deterministic in nature.

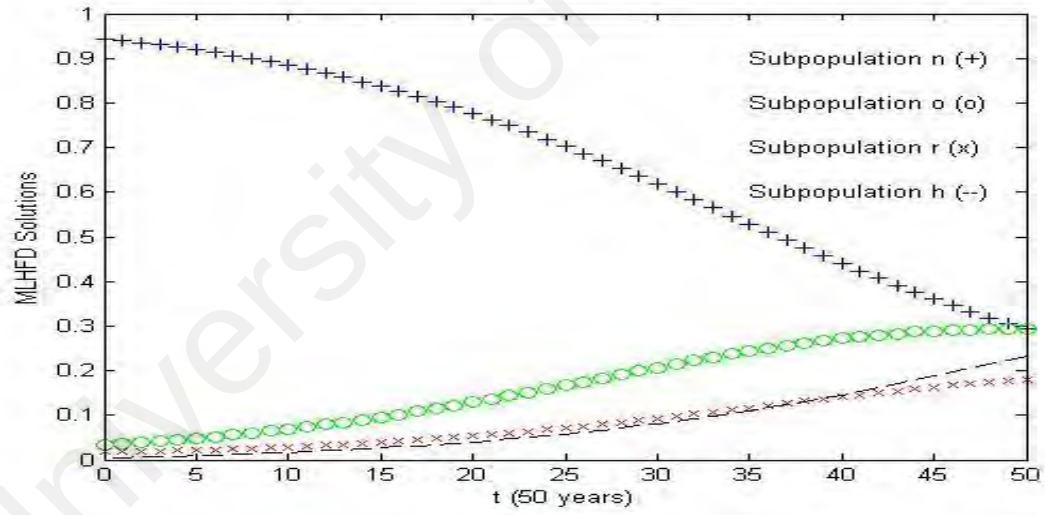


**Figure 4.3:** MLHFD numerical simulation results using 5000 repetitions from 1995 to 2045 with the step size, $w$ = 1 (yearly)

## 4.5 Conclusion

Cocaine abuse model has been studied in order to understand and to analyze the behavior of the social epidemic dynamics. A modified method that combined the statistical Latin hypercube sampling (LHS) technique with the numerical finite difference (FD) method is proposed to solve the nonlinear system of ordinary differential equations. Named

as Mean Latin Hypercube Finite Difference (MLHFD), this method is an upgrade of Mean Monte Carlo Finite Difference (MMCFD) method proposed by the same authors last year. These methods improved the existing statistical/numerical process by integrating each statistical simulation together with the last finite difference iteration in each simulation and finally taking the mean solution for the selected number of simulations/iterations as the final solution of the method. The advantage of MLHFD over MMCFD is that it produces faster simulation via LHS random group sampling distribution. In this work, the MLHFD results are tabulated together with the results of FD, HAM-Pade and the previous statistical estimations for the cocaine abuse model. The advantage of MLHFD over FD and HAM-Pade deterministic solutions is due to the random distribution provided by LHS for the real epidemic model. Using random distribution properties, the upper and lower boundaries of the predicted range for the cocaine abuse solutions can be determined by MLHFD method. It is found that FD, HAM-Pade and previous statistical estimation lie within the 90% prediction interval of the MLHFD solutions. Since FD and HAM-Pade are deterministic in nature while the previous statistical estimation is random, therefore the MLHFD solutions are always closer to the previous statistical predictions as compared to other deterministic solutions. The optimum solutions of MLHFD are achieved when the step size is based on annual unit ($w = 1$) at 5000 simulations for 20 years prediction as relative to the previous statistical estimations. In all computations done, it is found that the rate of cocaine non-users will decline while the number of occasional, regular and habitual users of cocaine in Spain will gradually increase in the future. Furthermore, the statistical estimations and the deterministic results by the previous researchers are found to be within the MLHFD prediction intervals for all the years and for all the subpopulations considered.

# CHAPTER 5: CONCLUSION AND SUGGESTION

## 5.1 Conclusions

In this thesis, modified finite difference methods for random sampling of social epidemic models are proposed. The importance of the current study and the research objectives have been clarified in Chapter 1. The problem statement which highlighted why this study is necessary has been provided. The scope of the research and the outline of thesis have also been displayed. The important role of simulation technique has been explained for solving the deterministic epidemic models with random parameters. Several preliminary concepts have been explained on ordinary differential equations, mathematical models, random sampling approach, simple MC statistical simulation process, stratified simulation technique of LHS, numerical FD method and the current proposed modified numerical simulation techniques of Mean Monte Carlo Finite Difference (MMCFD) and Mean Latin Hypercube Finite Difference (MLHFD) have been reviewed in Chapter 2.

The first modified finite difference method is proposed to solve an epidemiological model explicating the effects of public health campaign on body weight loss in Spain. The application of this method is carefully demonstrated in Chapter 3. The modified statistical-numerical method for solving the social epidemic model discussed earlier by Santonja et al. (2012) is named as Mean Monte Carlo Finite Difference (MMCFD) method. The proposed MMCFD method is a modified numerical simulation method that combined the statistical Monte Carlo (MC) technique with the numerical finite difference (FD) method. It is found that MMCFD method produced closer approximated solutions towards previous statistical prediction by Santonja et al. (2012) as compared to the deterministic FD method in all time intervals considered. Therefore MMCFD method is promising to create alternative estimation value between statistical and numerical methods. The MMCFD is also suggested as an alternative modified numerical simulation method to create prediction interval for

random distributions of the social epidemic solutions for the weight reduction model due to public campaign in Spain. From the results in Chapter 3, the curves of the normal weight subpopulation, $N(t)$ obtained from both FD and MMCFD methods are gradually declining. In general, the $N(t)$ curves for MMCFD are decreasing significantly than the $N(t)$ curves of FD from 2001 to 2030. Oppositely the curves for overweight and obesity subpopulations ($S(t)$ and $O(t)$) using MMCFD method are rising higher than the FD curves throughout the 30 years. Moreover, the overweight and the obesity curves are expected to increase in the future. MMCFD method produced faster and closer approximate solutions towards the previous statistical prediction by Santonja et al. (2012) as compared to FD method in all time intervals and for all subpopulations considered. The Mean Monte Carlo Finite Difference (MMCFD) method is promising to create alternative estimation value between the Monte Carlo (MC) statistical technique and the numerical finite difference (FD) method.

The second modified finite difference method is proposed in Chapter 4. This time, this method is applied to solve a cocaine abuse problem in Spain. This modified statistical-numerical method for solving the cocaine abuse model is named as Mean Latin Hypercube Sampling Finite Difference (MLHFD) method. The proposed MLHFD method is a modified numerical simulation method that combined the statistical Latin hypercube sampling (LHS) technique with the numerical finite difference (FD) method. From the results obtained in Chapter 4, MLHFD method produced closer approximated solutions towards previous statistical prediction (Sánchez et al., 2011) as compared to FD and HAM-Pade (Guerrero & Vazquez-Leal, 2014) methods in all time intervals considered due to randomization in the results variation which enables MLHFD to predict the range of results distribution of the model. It can be concluded that MLHFD method is promising to create alternative estimation value between statistical and numerical methods. Furthermore MLHFD is suggested as alternative modified numerical simulation methods to create prediction interval for random distribution of the cocaine abuse model solutions.

In overall conclusion, MMCFD and MLHFD techniques are modified approaches between statistical simulation and numerical finite difference approach. MMCFD and MLHFD are considered as better numerical simulation processes than numerical FD method with respect to the closest results to the existing statistical estimations under study. MLHFD method is fast and saves effort with time when dealing with the real models that have statistical probability. MMCFD and MLHFD processes are flexible to solve the nonlinear systems of the ODEs with random parameters. From this thesis, it can be shown that these techniques are the best compliment techniques between statistical and numerical approaches to solve such real stochastic-deterministic models. Hence, the applications of the present modified finite difference methods on the selected social epidemic models have been demonstrated sufficiently. The growth of social habits as epidemics can also be analyzed closely in the studied models of weight loss and cocaine abuse problems in Spain. The proposed modified methods of Mean Monte Carlo Finite Difference (MMCFD) and Mean Latin Hypercube Finite Difference (MLHFD) are beneficial to predict ranges for obtaining numerical simulation results when random distribution of the numerical solutions are necessary for estimation of real epidemiological models.

## 5.2 Suggestions

The current study deals with modification of the standard finite difference (FD) method. These methods can be applied on solving several differential equations such as decay equations, logistic differential equation, harmonic oscillator, Burgers' equation, unidirectional wave equation, etc. One of the future progress from the current study is the proposal of modified nonstandard finite difference (NFD) method on these previously mentioned equations (Martín-Vaquero et al., 2017; Mickens, 1994). We can study the properties of numerical standard finite difference method by using nonstandard finite difference method. The NFD method is related to numerical instabilities and it can be any discrete scheme of a system of DEs based on some rules (Mickens, 2005, p. 4). The numerical instabilities can happen when the order of the finite difference rule is greater than

the order of the DE (Mickens, 1994, p. 60). The general idea of numerical instabilities is to construct a discrete model that has a solution such that this solution has the same qualitative properties correspond to the differential equations for all step sizes by deleting the elementary numerical instabilities (Mickens, 1994, p. 14). Recently, some researchers solved the DEs by using the NFD method (Arenas et al., 2008; Jódar et al., 2008; Garba et al., 2011; Anguelov et al., 2014; Garba et al., 2015; Arenas et al., 2016; Martín-Vaquero et al., 2017; Wood et al., 2017; Xu et al., 2017).

MMCFD and MLHFD methods are suggested to solve more nonlinear IVP systems of ODEs representing other social epidemic models such as alcohol consumption (Santonja et al., 2010; Zhu & Zhu, 2017), anti-drug (Fierro et al., 2015; Goldstein et al., 2016) and smoking habit (Sikander et al., 2017; Haq et al., 2017). The MMCFD and MLHFD methods can also be applied to solve autonomous system of delay differential equations. The time delay in the latent interval can reduce the infected diseases. In a recent study, the latent period of disease steps was studied with time delay (Song et al., 2017). Another suggestion that can be discussed with the MMCFD and MLHFD methods includes the application of these techniques on higher order IVP for nonlinear autonomous system of ODEs, partial DEs and fractional ODEs that have random variables (Haq et al., 2017).

Stability analysis can be done to discuss the epidemic model behavior by deriving the basic reproduction number $R_0$ of the epidemic models. $R_0$ is conventionally used to test the stability of the epidemic models, whether the disease becomes endemic, decays, grows or remains in the population. The recent study of stability for epidemic models is given by Enduri and Jolad (2017). Other direction of study than can be explored is related to the information that can reduce infectious disease outbreak and deals with complicated networks. Wang et al. (2017) established such relation between epidemic spread and the information of complex networks using mobility patterns.

Apart from that, RK numerical iteration methods with different orders such as RK2, RK4, RK45 and RK78 can be considered to be merged with MC or LHS simulation techniques as new modified finite difference methods to solve deterministic models with random parameters (Mohammed et al., 2015). Additionally, other stratified simulation processes can be used to improve the MMCFD or MLHFD methods such as Discrete Latin Hypercube Sampling (DLHS) (Maschio & Schiozer, 2016) and Combined Multiple-LHS (CM-LHS) (Nakayama, 2011). Finally, the idea of optimization to select the best number of statistical simulation repetition can be incorporated with numerical method so that the generated numerical simulation results from the modified finite difference methods may produce closer solutions to existing statistical estimation considered under the future study.

# REFERENCES


Abdulle, A., & Blumenthal, A. (2013). Stabilized multilevel Monte Carlo method for stiff stochastic differential equations. *Journal of Computational Physics, 251*, 445-460.

Agaba, G. O., Kyrychko, Y. N., & Blyuss, K. B. (2017). Mathematical model for the impact of awareness on the dynamics of infectious diseases. *Mathematical Biosciences, 286*, 22-30.

Akimenko, V. (2017). An age-structured SIR epidemic model with fixed incubation period of infection. *Computers & Mathematics with Applications, 73*(7), 1485-1504.

Ali, A., Iqbal, M. A., Ul-Hassan, Q. M., Ahmad, J., & Mohyud-Din, S. T. (2016). An efficient technique for higher order fractional differential equation. *SpringerPlus, 5*(1), 281.

Allen, L. J., Brauer, F., Van den Driessche, P., & Wu, J. (2008). *Mathematical epidemiology*. German: Springer.

Allman, E. S., & Rhodes, J. A. (2004). *Mathematical models in biology: an introduction*. USA: Cambridge University Press.

Anguelov, R., Dumont, Y., Lubuma, J. M. S., & Shillor, M. (2014). Dynamically consistent nonstandard finite difference schemes for epidemiological models. *Journal of Computational and Applied Mathematics, 255*, 161-182.

Arenas, A. J., González-Parra, G., & Chen-Charpentier, B. M. (2016). Construction of nonstandard finite difference schemes for the SI and SIR epidemic models of fractional order. *Mathematics and Computers in Simulation, 121*, 48-63.

Arenas, A. J., Moraño, J. A., & Cortés, J. C. (2008). Non-standard numerical method for a mathematical model of RSV epidemiological transmission. *Computers & Mathematics with Applications, 56*(3), 670-678.

Azizi, A., Ríos-Soto, K., Mubayi, A., & Hyman, J. M. (2017). A risk-based model for predicting the impact of using condoms on the spread of sexually transmitted infections. *Infectious Disease Modelling, 2*(1), 100-112.

Bai, Z. (2015). Threshold dynamics of a time-delayed SEIRS model with pulse vaccination. *Mathematical Biosciences, 269*, 178-185.

Berge, T., Bowong, S., & Lubuma, J. M. S. (2017). Global stability of a two-patch cholera model with fast and slow transmissions. *Mathematics and Computers in Simulation, 133*, 142-164.

Blok, D. J., Crump, R. E., Sundaresh, R., Ndeffo-Mbah, M., Galvani, A. P., Porco, T. C., . . . Richardus, J. H. (2017). Forecasting the new case detection rate of leprosy in four states of Brazil: A comparison of modelling approaches. *Epidemics, 18*, 92-100.

Brauer, F., & Castillo-Chavez, C. (2001). *Mathematical models in population biology and epidemiology* (2nd ed.). London: Springer.

Budgaga, W., Malensek, M., Pallickara, S., Harvey, N., Breidt, F. J., & Pallickara, S. (2016). Predictive analytics using statistical, learning, and ensemble methods to support real-time exploration of discrete event simulations. *Future Generation Computer Systems, 56*, 360-374.

Buezas, F. S., Rosales, M. B., & Sampaio, R. (2013). Propagation of uncertainties and multimodality in the impact problem of two elastic bodies. *International Journal of Mechanical Sciences, 75*, 145-155.

Cai, L., Li, X., Tuncer, N., Martcheva, M., & Lashari, A. A. (2017). Optimal control of a malaria model with asymptomatic class and superinfection. *Mathematical Biosciences, 288*, 94-108.

Carsey, T. M., & Harden, J. J. (2014). *Monte Carlo Simulation and Resampling Methods for Social Science*. California: SAGE Publications.

Champredon, D., Li, M., Bolker, B. M., & Dushoff, J. (2017). Two approaches to forecast Ebola synthetic epidemics. *Epidemics*. doi:10.1016/j.epidem.2017.02. 011

Cheney, W., & Kincaid, D. (1999). *Numerical mathematics and computing* (4th ed.). USA: Brooks/Cole Publishing Company.

Chowell, G. (2017). Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling, 2*(3), 379-398.

Chowell, G., Hyman, J. M., Bettencourt, L. M., & Castillo-Chavez, C. (2009). *Mathematical and statistical estimation approaches in epidemiology*. New York: Springer.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine, 2007*(357), 370-379.

Cortés, J. C., Romero, J. V., Roselló, M. D., & Villanueva, R. J. (2017). Improving adaptive generalized polynomial chaos method to solve nonlinear random differential equations by the random variable transformation technique. *Communications in Nonlinear Science and Numerical Simulation, 50*, 1-15.

Dagpunar, J. S. (2007). *Simulation and Monte Carlo: With applications in finance and MCMC*. England: John Wiley & Sons Ltd.

De la Sen, M., Alonso-Quesada, S., & Ibeas, A. (2015). On the stability of an SEIR epidemic model with distributed time-delay and a general class of feedback vaccination rules. *Applied Mathematics and Computation, 270*, 953-976.

De Souza, M. V. C., Colaço, M. J., & Leiroz, A. J. K. (2014). Application of the generalized Polynomial Chaos expansion to the simulation of an internal combustion engine with uncertainties. *Fuel, 134*, 358-367.

De Veaux, R. D., Velleman, P. F., & Bock, D. E. (2012). *Intro Stats:* (3rd ed.). USA: Pearson Education, Inc.

De Vries, G., Hillen, T., Lewis, M., Müller, J., & Schönfisch, B. (2006). *A course in mathematical biology: quantitative modeling with mathematical and computational methods*.USA: Siam.

DeCleene, K. E., & Fogo, J. (2012, 23 Mar 2012). Publication Manual of the American Psychological Association. *Occupational therapy in health care, 26,* 90-92.

Dereich, S., & Heidenreich, F. (2011). A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. *Stochastic Processes and their Applications, 121*(7), 1565-1587.

Diamantopoulos, A., & Schlegelmilch, B. B. (2000). *Taking the fear out of data analysis: a step-by-step approach*. Singapore: Cengage Learning EMEA.

Diekmann, O., & Heesterbeek, J. A. P. (2008). *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. German: Springer-Verlag Berlin Heidelberg.

Elnashaie, S. S., & Uhlig, F. (2007). *Numerical techniques for chemical and biological engineers using MATLAB®: A simple bifurcation approach*. USA: Springer Science & Business Media.

Enduri, M. K., & Jolad, S. (2017). Estimation of reproduction number and non stationary spectral analysis of dengue epidemic. *Mathematical Biosciences, 288*, 140-148.

Faes, M., & Moens, D. (2017). Identification and quantification of spatial interval uncertainty in numerical models. *Computers & Structures, 192*, 16-33.

Farnoosh, R., & Ebrahimi, M. (2010). Monte Carlo simulation via a numerical algorithm for solving a nonlinear inverse problem. *Communications in Nonlinear Science and Numerical Simulation, 15*(9), 2436-2444.

Ferziger, J. H., Peric, M., & Leonard, A. (2002). *Computational methods for fluid dynamics* (3rd ed.). Germany: Springer-Verlag Berlin Heidelberg.

Fierro, I., González-Luque, J. C., Seguí-Gómez, M., & Álvarez, F. J. (2015). Alcohol and drug use by Spanish drivers: Comparison of two cross-sectional road-side surveys (2008–9/2013). *International Journal of Drug Policy, 26*(8), 794-797.

Frauenthal, J. C. (1980). *Mathematical modeling in epidemiology*. New York: Springer Science & Business Media.

Fulford, G., Forrester, P., & Jones, A. (1997). *Modelling with differential and difference equations* . UK: Cambridge University Press.

Garba, S. M., Gumel, A. B., Hassan, A. S., & Lubuma, J. M. S. (2015). Switching from exact scheme to nonstandard finite difference scheme for linear delay differential equation. *Applied Mathematics and Computation, 258*, 388-403.

Garba, S. M., Gumel, A. B., & Lubuma, J. M. S. (2011). Dynamically-consistent non-standard finite difference method for an epidemic model. *Mathematical and Computer Modelling, 53*(1), 131-150.

Gentle, J. E. (2003). *Random number generation and Monte Carlo methods* (2nd ed.). USA: Springer Science Business Media, Inc.

Giordano, F. R., Weir, M. D., & Fox, W. P. (2003). *A first course in mathematical modeling, Brooks* (3rd ed.). USA: Brooks/Cole, Thomson Learning

Goldstein, N. D., Burstyn, I., LeVasseur, M. T., & Welles, S. L. (2016). Drug use among men by sexual behaviour, race and ethnicity: Prevalence estimates from a nationally representative US sample. *International Journal of Drug Policy, 36*, 148-150.

Gouvêa Jr, M. M. (2017). Time-spatial model on the dynamics of the proliferation of Aedes aegypti. *Communications in Nonlinear Science and Numerical Simulation, 44*, 130-143.

Graham, C., & Talay, D. (2013). *Stochastic Simulation and Monte Carlo Methods: Mathematical Foundations of Stochastic Simulation*. London New York: Springer Berlin Heidelberg.

Guerrero, F., Santonja, F., & Villanueva, R. (2013). Solving a model for the evolution of smoking habit in Spain with homotopy analysis method. *Nonlinear Analysis: Real World Applications, 14*(1), 549-558.

Guerrero, F., Santonja, F.-J., & Villanueva, R.-J. (2011). Analysing the Spanish smoke-free legislation of 2006: a new method to quantify its impact using a dynamic model. *International Journal of Drug Policy, 22*(4), 247-251.

Guerrero, F., & Vazquez-Leal, H. (2014). Application of multi-stage HAM-Padé to solve a model for the evolution of cocaine consumption in Spain. *TWMS Journal of Pure and Applied Mathematics, 5*(2), 241-255.

Gustafsson, B. (2011). *Fundamentals of scientific computing*. London New York: Springer-Verlag Berlin Heidelberg.

Hahn, G. J., & Meeker, W. Q. (2011). *Statistical Intervals: A Guide for Practitioners*. Canada: John Wiley & Sons.

Haq, F., Shah, K., ur Rahman, G., & Shahzad, M. (2017). Numerical solution of fractional order smoking model via Laplace Adomian decomposition method. *Alexandria Engineering Journal,*. doi:10.1016/j.aej.2017.02.015

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM review, 42*(4), 599-653.

Hirsch, M. W., Smale, S., & Devaney, R. L. (2012). *Differential equations, dynamical systems, and an introduction to chaos* (2nd ed.). USA: Academic press in Elsevier.

Hosseini, S. M., & Shahabian, F. (2011). Transient analysis of thermo-elastic waves in thick hollow cylinders using a stochastic hybrid numerical method, considering Gaussian mechanical properties. *Applied Mathematical Modelling, 35*(10), 4697-4714.

Ingalls, B. P. (2013). *Mathematical modeling in systems biology: an introduction*. London: England MIT press.

Jacoboni, C., & Lugli, P. (1989). *The Monte Carlo method for semiconductor device simulation*. New York: Springer Science & Business Media.

Jiang, Z., Ma, W., & Wei, J. (2016). Global Hopf bifurcation and permanence of a delayed SEIRS epidemic model. *Mathematics and Computers in Simulation, 122*, 35-54.

Jódar, L., Villanueva, R. J., Arenas, A. J., & González, G. C. (2008). Nonstandard numerical methods for a mathematical model for influenza disease. *Mathematics and Computers in Simulation, 79*(3), 622-633.

Karris, S. T. (2007). *Numerical analysis using Matlab and Exce* (3rd.). USA: Orchard Publications.

Kim, S., Lee, J., & Jung, E. (2017). Mathematical model of transmission dynamics and optimal control strategies for 2009 A/H1N1 influenza in the Republic of Korea. *Journal of Theoretical Biology, 412*, 74-85.

Kovtanyuk, A. E., Botkin, N. D., & Hoffmann, K.-H. (2012). Numerical simulations of a coupled radiative–conductive heat transfer model using a modified Monte Carlo method. *International Journal of Heat and Mass Transfer, 55*(4), 649-654.

Krämer, A., Kretzschmar, M., & Krickeberg, K. (2010). *Modern infectious disease epidemiology: Concepts, methods, mathematical models, and public health*. USA: Springer.

Kumar, A., & Srivastava, P. K. (2017). Vaccination and treatment as control interventions in an infectious disease model with their cost optimization. *Communications in Nonlinear Science and Numerical Simulation, 44*, 334-343.

Kypraios, T., Neal, P., & Prangle, D. (2017). A tutorial introduction to Bayesian inference for stochastic epidemic models using Approximate Bayesian Computation. *Mathematical Biosciences, 287*, 42-53.

La Salle, J. P. (1976). *The stability of dynamical systems*: SIAM. doi: 10.1137/ 1.9781611970 432 .bm

Lachin, J. M. (2011). *Biostatistical methods: the assessment of relative risks* (2nd ed.). USA: John Wiley & Sons.

Lambert, J. D. (1991). *Numerical methods for ordinary differential systems: the initial value problem*. USA: John Wiley & Sons, Inc.

Leander, J., Lundh, T., & Jirstrand, M. (2014). Stochastic differential equations as a tool to regularize the parameter estimation problem for continuous time dynamical systems given discrete time measurements. *Mathematical Biosciences, 251*, 54-62.

Lee, J. K. (2010). *Statistical bioinformatics: for biomedical and life science researchers.* Canada: John Wiley & Sons.

Lee, S., & Chowell, G. (2017). Exploring optimal control strategies in seasonally varying flu-like epidemics. *Journal of Theoretical Biology, 412*, 36-47.

Lee, W. J. (2008). *Computer Generation and Computation of Mixed Poisson and Related Distributions.* (Doctoral Dissertation), University of Malaya, Kuala Lumpur, Malaysia.

Levy, B., Edholm, C., Gaoue, O., Kaondera-Shava, R., Kgosimore, M., Lenhart, S., . . . Nyabadza, F. (2017). Modeling the role of public health education in Ebola virus disease outbreaks in Sudan. *Infectious Disease Modelling, 2*(3), 323-340.

Li, C., Chen, B., Peng, H., & Zhang, S. (2017). Sparse regression Chebyshev polynomial interval method for nonlinear dynamic systems under uncertainty. *Applied Mathematical Modelling, 51*, 505-525.

Li, Q., Lu, F., Dai, C., Fan, M., Wang, W., & Wang, K. (2017). Simulating the potential role of media coverage and infected bats in the 2014 Ebola outbreak. *Journal of Theoretical Biology, 412*, 123-129.

Licea, J., Villafuerte, L., & Chen-Charpentier, B. M. (2013). Analytic and numerical solutions of a Riccati differential equation with random coefficients. *Journal of Computational and Applied Mathematics, 239*, 208-219.

Liu, L.-B., & Chen, Y. (2014). Maximum norm a posteriori error estimates for a singularly perturbed differential difference equation with small delay. *Applied Mathematics and Computation, 227*, 801-810.

Liu, S., Ruan, S., & Zhang, X. (2017). Nonlinear dynamics of avian influenza epidemic models. *Mathematical Biosciences, 283*, 118-135.

Livingston, E. H., Sebastian, J. L., Huerta, S., Yip, I., & Heber, D. (2001). Biexponential model for predicting weight loss after gastric surgery for obesity. *Journal of Surgical Research, 101*(2), 216-224.

Logan, J. D. (2006). *A first course in differential equations*. USA: Springer.

Ma, Z., Zhou, Y., & Wu, J. (2009). *Modeling and dynamics of infectious diseases.* China: World Scientific.

Mainik, G. (2015). Risk aggregation with empirical margins: Latin hypercubes, empirical copulas, and convergence of sum distributions. *Journal of Multivariate Analysis, 141*, 197-216.

Martín-Vaquero, J., Martín del Rey, A., Encinas, A. H., Hernández Guillén, J. D., Queiruga-Dios, A., & Rodríguez Sánchez, G. (2017). Higher-order nonstandard finite difference schemes for a MSEIR model for a malware propagation. *Journal of Computational and Applied Mathematics, 317*, 146-156.

Maschio, C., & Schiozer, D. J. (2016). Probabilistic history matching using discrete Latin Hypercube sampling and nonparametric density estimation. *Journal of Petroleum Science and Engineering, 147*, 98-115.

Mazza, C., & Benaim, M. (2014). *Stochastic dynamics for systems biology*. U.S: CRC Press of Taylor & Francis Group.

Menz, S. (2013). *Hybrid stochastic-deterministic approaches for simulation and analysis of biochemical reaction networks.* (Doctoral Dissertation), Freie Universität, Berlin.

Meyer, W. J. (1984). *Concepts of mathematical modeling*. USA: McGraw-Hill, Inc.

Mickens, R. E. (1994). *Nonstandard finite difference models of differential equations*. USA: world scientific.

Mickens, R. E. (2005). *Advances in the applications of nonstandard finite difference schemes*. Singapore: World Scientific.

Moghaddam, B. P., & Mostaghim, Z. S. (2013). A numerical method based on finite difference for solving fractional delay differential equations. *Journal of Taibah University for Science*, 7(3), 120-127.

Mohammed, M. A., Noor, N. F. M., Siri, Z., & Ibrahim, A. I. N. (2015). Numerical solution for weight reduction model due to health campaigns in Spain [AIP Conference Proceeding]. At Hotel Grand BlueWave, Shah Alam, Kuala Lumpur, Malaysia: AIP Publishing, 1682, 020005-1-020005-6.

Mohammed, M. A., Ibrahim, A. I. N., Siri, Z., & Noor, N. F. M. (2016). Mean Monte Carlo Finite Difference Method for Random Sampling of a Nonlinear Epidemic System. *Sociological Methods & Research*, 1-28. doi: 10.1177/00491241166 72683

Mohyud-Din, S., Yıldırım, A., & Gulkanat, Y. (2012). Approximate analysis of population dynamics with density-dependent migrations and the Allee effects. *International Journal of Numerical Methods for Heat & Fluid Flow, 22*(2), 243-250.

Mohyud-Din, S. T., Noor, M. A., & Noor, K. I. (2009). Some relatively new techniques for nonlinear problems. *Mathematical Problems in Engineering, 2009*, 1-25. doi: 10.1155/2009/234849

Monzó, M. R. (2015). *Uncertainty quantification in dynamical models. An application to cocaine consumption in Spain.* (Doctoral Dissertation), Politècnica de València, Valencia, Spain.

Mooney, D. D., & Swift, R. J. (1999). *A course in mathematical modeling*. USA: The Mathematical Association of America.

Murray, J. (2002). *Mathematical Biology: I. An introduction* (3rd ed.). USA: Springer.

Murray, J. D. (2003). *Mathematical Biology. II Spatial Models and Biomedical Applications* (3rd ed.) USA :Springer-Verlag New York Incorporated.

Nagle, R. K., Saff, E. B., & Snider, A. D. (2008). *Fundamentals of differential equations* (7th ed.): USA: Pearson Education Inc. /Addison Wesley.

Nakayama, M. K. (2011). Asymptotically valid confidence intervals for quantiles and values-at-risk when applying Latin hypercube sampling. *International Journal on Advances in Systems and Measurements, 4*, 86-94.

Nakayama, M. K. (2012, 28-31 October). *Confidence intervals for quantiles when applying replicated Latin hypercube sampling and sectioning.* Paper presented at the Proceedings of the 2012 Autumn Simulation Multiconference California, USA.

Navarro-Barrientos, J.-E., Rivera, D. E., & Collins, L. M. (2011). A dynamical model for describing behavioural interventions for weight loss and body composition change. *Mathematical and computer modelling of dynamical systems, 17*(2), 183-203.

Newman, S. C. (2001). *Biostatistical methods in epidemiology*. USA: John Wiley & Sons.

Patel, J. (1989). Prediction intervals-A review. *Communications in Statistics-Theory and Methods, 18*(7), 2393-2465.

Pedro, S., Tonnang, H. E., & Abelman, S. (2016). Uncertainty and sensitivity analysis of a Rift Valley fever model. *Applied Mathematics and Computation, 279*, 170-186.

Peng, L. F. (2016). *Statistical appraisal in solving some medical problems* (Doctoral Disserttion), University of Malaya, Kuala Lumpur, Malaysia.

Perros, H. G. (2009). *Computer Simulation Techniques: The definitive introduction*!: Retrieved from https://repository.lib.ncsu.edu/bitstream/handle/1840.2/2542/simulation.pdf?sequence=1

Pryse, S. E., & Adhikari, S. (2017). Stochastic finite element response analysis using random eigenfunction expansion. *Computers & Structures, 192*, 1-15.

Rao, R. N., & Chakravarthy, P. P. (2013). A finite difference method for singularly perturbed differential-difference equations arising from a model of neuronal variability. *Journal of Taibah University for Science, 7*(3), 128-136.

Rao, R. N., & Chakravarthy, P. P. (2014). An exponentially fitted tridiagonal finite difference method for singularly perturbed differential-difference equations with small shift. *Ain Shams Engineering Journal, 5*(4), 1351-1360.

Rohaninejad, M., & Zarghami, M. (2012). Combining Monte Carlo and finite difference methods for effective simulation of dam behavior. *Advances in Engineering Software, 45*(1), 197-202.

Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo method*. Canada: John Wiley & Sons.

Safi, M. A., & Garba, S. M. (2012). Global stability analysis of SEIR model with holling type II incidence function. *Computational and mathematical methods in medicine, 2012*, 1-8. doi:10.1155/2012/826052

Sánchez, E., Villanueva, R.-J., Santonja, F.-J., & Rubio, M. (2011). Predicting cocaine consumption in Spain: A mathematical modelling approach. *Drugs: education, prevention and policy, 18*(2), 108-115.

Santonja, F.-J., Lombana, I.-C., Rubio, M., Sánchez, E., & Villanueva, J. (2010). A network model for the short-term prediction of the evolution of cocaine consumption in Spain. *Mathematical and Computer Modelling, 52*(7), 1023-1029.

Santonja, F.-J., Morales, A., Villanueva, R.-J., & Cortés, J.-C. (2012). Analysing the effect of public health campaigns on reducing excess weight: A modelling approach for the Spanish Autonomous Region of the Community of Valencia. *Evaluation and program planning, 35*(1), 34-39.

Santonja, F.-J., Sánchez, E., Rubio, M., & Morera, J.-L. (2010). Alcohol consumption in Spain and its economic cost: a mathematical modeling approach. *Mathematical and Computer Modelling, 52*(7), 999-1003.

Shampine, L. F., Gladwell, I., & Thompson, S. (2003). *Solving ODEs with MATLAB*. USA: Cambridge University Press.

Shu, H., Fan, D., & Wei, J. (2012). Global stability of multi-group SEIR epidemic models with distributed delays and nonlinear transmission. *Nonlinear Analysis: Real World Applications, 13*(4), 1581-1592.

Sikander, W., Khan, U., Ahmed, N., & Mohyud-Din, S. T. (2017). Optimal Solutions for a Bio Mathematical Model for the Evolution of Smoking Habit. *Results in Physics*, 7,510-517.

Simpson, L., & Gumel, A. B. (2017). Mathematical assessment of the role of pre-exposure prophylaxis on HIV transmission dynamics. *Applied Mathematics and Computation, 293*, 168-193.

Snyder, L. B. (2007). Health communication campaigns and their impact on behavior. *Journal of nutrition education and behavior, 39*(2), S32-S40.

Sobol, I. M. (1994). *A Primer for the Monte Carlo Method*. USA: Taylor & Francis.

Song, L.-P., Zhang , R.-P., Feng , L.-P., & Shi, Q. (2017). Pattern dynamics of a spatial epidemic model with time delay. *Applied Mathematics and Computation, 292*, 390-399.

Spigler, R., & Zerbetto, R. (2013). Random perturbations effects on the stability of Tension Leg Platforms in offshore engineering and of other large structures. *Applied Mathematical Modelling, 37*(5), 2881-2899.

Stockdale, J. E., Kypraios, T., & O'Neill, P. D. (2017). Modelling and Bayesian analysis of the Abakaliki smallpox data. *Epidemics, 19*, 13-23.

Su, H., Li, W., & Ding, X. (2013). Numerical dynamics of a nonstandard finite difference method for a class of delay differential equations. *Journal of Mathematical Analysis and Applications, 400*(1), 25-34.

Taylor, H. M., & Karlin, S. (2014). *An introduction to stochastic modeling* (3rd ed.). USA: Academic press.

Teschl, G. (2012). *Ordinary differential equations and dynamical systems*. USA: American Mathematical Society Providence.

Thomas, D. M., Weedermann, M., Fuemmeler, B. F., Martin, C. K., Dhurandhar, N. V., Bredlau, C., . . . Bouchard, C. (2014). Dynamic model predicting overweight, obesity, and extreme obesity prevalence trends. *Obesity, 22*(2), 590-597.

Tsai, C. S. (2002). *An introduction to computational biochemistry.* USA: John Wiley & Sons.

Tucker, H. G. (1998). *Mathematical Methods in Sample Surveys*. Singapore: World scientific.

Ul Hassan, Q. M., & Mohyud-Din, S. T. (2016). Investigating biological population model using exp-function method. *International Journal of Biomathematics, 9*(02), 1650026.

Ullah, M., & Wolkenhauer, O. (2011). *Stochastic approaches for systems biology*. USA: Springer Science & Business Media.

Van Belle, G., Fisher, L. D., Heagerty, P. J., & Lumley, T. (2004). *Biostatistics: a methodology for the health sciences* (2nd ed.). Canada: John Wiley & Sons.

Velten, K. (2009). *Mathematical Modeling and Simulation: Introduction for Scientists and Engineers*. Germany: Wiley.

Vuldin, R. (2008). *Non-linear differential equations* (1st ed.). India: IVY Publishing House.

Wang, B., Han, Y., & Tanaka, G. (2017). Interplay between epidemic spread and information propagation on metapopulation networks. *Journal of Theoretical Biology, 420*, 18-25.

Wang, H., Wang, J., Ding, L., & Wei, W. (2017). Knowledge transmission model with consideration of self-learning mechanism in complex networks. *Applied Mathematics and Computation, 304*, 83-92.

Wang, X., Wei, L., & Zhang, J. (2014). Dynamical analysis and perturbation solution of an SEIR epidemic model. *Applied Mathematics and Computation, 232*, 479-486.

Wang, Y., Cao, J., Alsaedi, A., & Ahmad, B. (2017). Edge-based SEIR dynamics with or without infectious force in latent period on random networks. *Communications in Nonlinear Science and Numerical Simulation, 45*, 35-54.

Wilcox, R. R. (2006). Confidence intervals for prediction intervals. *Journal of Applied Statistics, 33*(3), 317-326.

Wilkinson, D. J. (2012). *Stochastic modelling for systems biology* (2nd ed.). UK: CRC press.

Witbooi, P. J. (2013). Stability of an SEIR epidemic model with independent stochastic perturbations. *Physica A: Statistical Mechanics and its Applications, 392*(20), 4928-4936.

Wood, D. T., Kojouharov, H. V., & Dimitrov, D. T. (2017). Universal approaches to approximate biological systems with nonstandard finite difference methods. *Mathematics and Computers in Simulation, 133*, 337-350.

Xu, D., Xu, X., Xie, Y., & Yang, C. (2017). Optimal control of an SIVRS epidemic spreading model with virus variation based on complex networks. *Communications in Nonlinear Science and Numerical Simulation, 48*, 200-210.

Xu, J., Geng, Y., & Hou, J. (2017). A non-standard finite difference scheme for a delayed and diffusive viral infection model with general nonlinear incidence rate. *Computers & Mathematics with Appl*ications. doi:10.1016/j.camwa.2017 .06.041

Xu, R., Wang, Z., & Zhang, F. (2015). Global stability and Hopf bifurcations of an SEIR epidemiological model with logistic growth and time delay. *Applied Mathematics and Computation, 269*, 332-342.

Yalim, J., Welfert, B. D., & Lopez, J. M. (2017). Evaluation of closure strategies for a periodically-forced Duffing oscillator with slowly modulated frequency subject to Gaussian white noise. *Communications in Nonlinear Science and Numerical Simulation, 44*, 144-158.

Yusof, N. B. (2011). *Simulation of the Lithium-Ion Performance Using Finite Difference Method.* (Doctoral Dissertation), University of Malaya, Kuala Lumpur, Malaysia.

Zaman, G., Kang, Y. H., Cho, G., & Jung, I. H. (2017). Optimal strategy of vaccination & treatment in an SIR epidemic model. *Mathematics and Computers in Simulation, 136*, 63-77.

Zarebski, A. E., Dawson, P., McCaw, J. M., & Moss, R. (2017). Model selection for seasonal influenza forecasting. *Infectious Disease Modelling, 2*(1), 56-70.

Zhao, H., Gao, Z., Gao, Y., & Wang, C. (2017). Effective robust design of high lift NLF airfoil under multi-parameter uncertainty. *Aerospace Science and Technology, 68*, 530-542.

Zhao, J. (2013). Compact finite difference methods for high order integro-differential equations. *Applied Mathematics and Computation, 221*, 66-78.

Zhi-zhao Liu, W. L., and Ming Yang. (2015). Two General Extension Algorithms of Latin Hypercube Sampling. *Mathematical Problems in Engineering, 2015*, 1-9. doi:10.1155/2015/450492

Zhu, C.-C., & Zhu, J. (2017). Stability of a reaction–diffusion alcohol model with the impact of tax policy. *Computers & Mathematics with Applications, 74*(4), 613-633.

Zinkovsky, A., Sholuha, V., & Ivanov, A. (1996). *Mathematical modelling and computer simulation of biomechanical systems*. USA: World Scientific.

# LIST OF PUBLICATIONS AND SUBMITED PAPERS

**List of publications**

1. Mohammed, M. A., Noor, N. F. M., Siri, Z., & Ibrahim, A. I. N. (2015). Numerical solution for weight reduction model due to health campaigns in Spain [*AIP Conference Proceeding*]. At Hotel Grand BlueWave, Shah Alam, Kuala Lumpur, Malaysia: AIP Publishing, 1682, (020005-1-020005-6).

2. Mohammed, M.A., Ibrahim, A.I.N, Siri, Z., & Noor, N.F.M. (2016). Mean Monte Carlo Finite Difference Method for Random Sampling of a Nonlinear Epidemic System. *Sociological Methods & Research*, 1-28. doi: 10.1177/ 0049124116672683