Incorporating Particle Filtering And System Dynamic Modelling in Infection Transmission of Measles and Pertussis

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Computer Science
University of Saskatchewan
Saskatoon

By
Xiaoyan Li

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
176 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5C9

Or
Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building
110 Science Place
Saskatoon, Saskatchewan
Canada
S7N 5C9

Abstract

Childhood viral and bacterial infections remain an important public problem, and research into their dynamics has broader scientific implications for understanding both dynamical systems and associated methodologies at the population level. Measles and pertussis are two important childhood infectious diseases. Measles is a highly transmissible disease and is one of the leading causes of death among young children under 5 globally. Pertussis (whooping cough) is another common childhood infectious disease, which is most harmful for babies and young children and can be deadly.

While the use of ongoing surveillance data and – recently – dynamic models offer insight on measles (or pertussis) dynamics, both suffer notable shortcomings when applied to measles (or pertussis) outbreak prediction. In this thesis, I apply the Sequential Monte Carlo approach of particle filtering, incorporating reported measles and pertussis incidence for Saskatchewan during the pre-vaccination era, using an adaptation of a previously contributed measles and pertussis compartmental models. To secure further insight, I also perform particle filtering on age structured adaptations of the models. For some models, I further consider two different methods of configuring the contact matrix.

The results indicate that, when used with a suitable dynamic model, particle filtering can offer high predictive capacity for measles and pertussis dynamics and outbreak occurrence in a low vaccination context. Based on the most competitive model as evaluated by predictive accuracy, I have performed prediction and outbreak classification analysis. The prediction results demonstrated that the most competitive models could predict the measles (or pertussis) outbreak patterns and classify whether there will be an outbreak or not in the next month (Area under the ROC Curve of measles is 0.89, while pertussis is 0.91).

I conclude that anticipating the outbreak dynamics of measles and pertussis in low vaccination regions by applying particle filtering with simple measles and pertussis transmission models, and incorporating time series of reported case counts, is a valuable technique to assist public health authorities in estimating risk and magnitude of measles and pertussis outbreaks. Such approach offers particularly strong value proposition for other pathogens with little-known dynamics, important latent drivers, and in the context of the growing number of high-velocity electronic data sources. Strong additional benefits are also likely to be realized from extending the application of this technique to highly vaccinated populations.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Nathaniel D. Osgood for the continuous support of my Master study and related research, for his patience, motivation, and immense knowledge. I could not have imagined having a better advisor for my Master study. During the two years, I have learned a lot from Prof. Osgood – not only in the academic level, but also in his high ethical standards. His optimism, kindness, care for others and passion for research are influenced me deeply, which will always remind me to become a better person in my rest life. Moreover, I take it as one of my most lucky things in my life to study and work with Prof. Osgood.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Christopher Dutchyn, Prof. Kevin Stanley, and Prof. Vineet Saini, for their insightful comments and encouragement. And also for the questions which incented me to widen my research from various perspectives.

I thank my fellow lab mates in for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last two years. Especially, I would like to thank Weicheng Qian and Anahita Safarishahrbijari for the suggestions in building the models; thank Lujie Duan for pointing a bug in my previous version models; thank Bryce Keeler for correcting the English of Chapter 1 of this thesis. I would also like to thank the stuffs in the computer science department, espetially Christine, Sophie and Gwen for their help during my master study period.

Last but not the least, I would like to thank my family: my parents, my husband, my daughters, my aunt and most importantly – my grandmother. I thank my parents (Wanqing Li and Liying Zhao) and my aunt (Jinling Li) for helping me to take care of my daughters during these two years. I would like to thank my husband and also my best friend – Gang Lyu. Your love, support and encouragement are always my impetus to move forward. I also thank my daughters – Emily Lyu and Alice Lyu for taking care of yourselves when Mom can not be there for you. Finally but most importantly, I thank my grandmother – Junying Zheng. Although you could not be there with me anymore, I know that my merits – brave and strong – if I have, are all learned and inherited from you. You will always be my hero and live in my deep heart forever.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

1.1.1 General burden of communicable diseases

Infectious diseases have significantly influenced the health status of peoples throughout the world. Historically, the number of deaths caused by infectious diseases is staggering. For example, in the fourteenth-century century in Europe, 25 million people (out of a total population of approximately 100 million) died due to the bubonic plague alone [2]. In the past several decades, infectious diseases still cause deaths among people worldwide, even in the vaccination era. For example, the human immunodeficiency virus (HIV), the etiological agent of acquired immunodeficiency syndrome (AIDS), has a significant impact on the mortality pattern of people both in developing and developed countries [2]. Measles is still one of the leading cause of death globally among children under 5 years old [80]. Recently, reflecting the growth of global worldwide, and closer contacts between human populations and animal populations, some new infectious diseases have emerged and have caused mass mortality, such aforementioned HIV, as the severe acute respiratory syndrome (SARS) [85] outbreak in China in 2002, pandemic H1N1 [77] in 2009 worldwide, and the recent Ebola [9] virus outbreak in West Africa among 2014-2016.

Understanding the outbreak patterns of infectious diseases can aid in forecasting and help public health agencies design intervention strategies to prevent and control such diseases, such as by setting outbreak response measures, setting vaccination targets, and allocating financial and human resources, etc.

Childhood communicable diseases are a significant part of communicable diseases. Although vaccinations are used worldwide, most children will have at least 6 to 8 respiratory infections (i.e., infections in the lungs and breathing tubes) each year [46]. Childhood viral and bacterial infections remain an important public problem, and research into their dynamics has broader scientific implications for understanding both dynamical systems and associated methodologies at the population level [92, 30, 13]. In this thesis, the dynamics of measles and pertussis are investigated to estimate and predict their transmission patterns.

Measles is a serious childhood infectious disease, which often leads to complications and death. Before the vaccination era, almost every child had been infected by measles [48]. Even in the vaccination era, measles is still commonly transmitted worldwide, including across countries in Europe, Asia, the Pacific and Africa [33].

It remains a leading cause of vaccine-preventable infant mortality [33]. In recent years, several countries been able to declare measles elimination (absence of continuous measles transmission for greater than 12 months [33]) by implementing a highly effective vaccination program and other control measures, like the United States and Canada. However, measles can still be brought in by unvaccinated travelers from other countries where measles still persists. Such transmission can result in outbreaks which are dangerous and costly to control [33]. Thus, research into measles dynamics remains essential for prevention and control. Moreover, the dynamics of measles have played a key role in theoretical epidemiology [11, 13, 97, 76, 2, 43, 44, 18, 49, 48]. Additionally, measles presents a unique opportunity because it exhibits both endemic and episodic dynamics [13]. In large cities, the dynamics of measles exhibits endemic cycles, while in small towns it exhibits recurrent outbreaks with periods of local extinction [13]. Before mass vaccination began in the 1960s, the outbreak of measles presented both regular (annual, biennial and triennial cycles) and irregular dynamics (with some exhibiting low dimensional chaos) [30, 76].

Pertussis (whooping cough) is another common childhood infectious disease. It is a serious respiratory infection caused by the pertussis bacteria, which is most harmful for babies and young children and can be deadly [37]. Before the vaccination era, about 8000 people died each year because of pertussis in the United States. In recent years, although vaccinations are widely spread in the United States, between 15,000 and 50,000 cases of pertussis are reported anually, and fewer than 20 people die each year [37]. With regards to natural history and transmission dynamics, pertussis is different from measles, in that it is more complex. For example, pertussis immunity acquired from infection and vaccination wanes over time, whereas immunity from measles is typically permanent. Moreover, more than one type of infected persons must be considered in pertussis dynamics. The reason is that the infectivity of an infected person who has a history of previous infection is lower than that of a newly infected person [48, 50].

Finally, in this thesis, the outbreak patterns of both measles and pertussis are studied by the use of models incorporating particle filtering algorithms, traditional mathematical dynamic models and noisy surveillance data.

1.1.2 Shortcomings of the previous methods

Traditionally, many public health agencies seek to anticipate future transmission of infectious diseases based upon time series of reported incident case counts. In recent years, simulation models have increasingly been used to predict the spread of infectious diseases within the population, and the potential impact of interventions on that spread. Although both these methods have their own benefits, and have played an important role in estimating and predicting measles and pertussis outbreak patterns, each suffers from notable limitations.

In many jurisdictions worldwide, the surveillance reports of some significant epidemiology diseases are collected on a regular basis, such as for measles, pertussis, chickenpox, etc. These data – often summarized as epidemiological curves – have offered great value for public health agencies seeking to understand trends

in the incidence and anticipate future evolution. For example, identifying and confirming suspected cases through surveillance allows early detection of outbreaks, and supports comparing recent activity to reference periods. However, the reported time series data are often quite noisy, particularly for contexts marked by smaller population or low incidence or diagnosis rates [18], such as are characteristic for measles and pertussis. Such reports are also delayed – sometimes by months, and are frequently woefully incomplete, due to the fact that many families of those infected elect not to seek care, especially in case that the infection is not serious. When used unassisted, it is difficult to secure quantitatively rigorous insights from time series into the future evolution of the currently observed patterns – evolution that will be shaped by a complex combination of local circumstances, including birth rates, contact patterns, and the size of the local pool of susceptible as affected by vaccination rates and recent history of other outbreaks, etc. Perhaps most importantly when considered as a tool for planning, use of time series alone cannot itself be used to investigate counterfactuals, such as how future incidence is likely to be affected by an outbreak response immunization campaign, enhanced contact tracing, advisories, or social distancing measures.

Dynamic modeling has played a significant role in providing insight into infectious disease outbreak dynamics [11, 13, 97, 76, 43, 44, 18, 49, 48]. Most such contributions employing models seek to incorporate some aspects of local epidemiology, and often draw on surveillance data. While they serve as powerful tools for investigating counterfactuals, such models also suffer from an essential set of shortcomings. Firstly, while dynamic models are commonly calibrated to empirical data, this process is typically undertaken on a one-time basis, and with significant human involvement. While calibration can allow for estimation of model parameters, it provides weak support for ongoing estimation of the latent state needed to keep the state of a model aligned with observations on an ongoing basis. It is rare for a dynamic model to incorporate ongoing arriving ground data; and while systems doing so can be found for other infectious diseases [20], the author is not aware of any such support for pertussis and measles. More profoundly, a dynamic model of necessity represents a simplified characterization of processes in the real world. Inevitably, such models often omit, simplify and mis-estimate some factors. These drawbacks and the infeasibility of anticipating the realized outcome of factors represented as stochastic in the model will inevitably lead the model to diverge from the unfolding epidemiological situation.

1.2 Research goals of this thesis

Particle filtering is a machine learning algorithm based on the idea of Monte Carlo approximation and for recursive Bayesian inference [74]. The particle filtering algorithm is normally implemented in the state space model which represents the dynamic system, including both a state transition model and measurement model.

In this thesis, the particle filtering algorithm can link the aggregate system dynamics (compartmental) models of measles and pertussis to the empirical datasets (noisy surveillance data) to compensate for weaknesses of both. Specifically, the state transition models are employed by the compartmental models of measles

and pertussis in a way that incorporates stochastic processes as system noise, while the surveillance data of measles and pertussis are used as the empirical data in the particle filtering algorithm. Thus, the stochastic processes are incorporated in the system dynamics models and the particle filtering algorithm allows for on-line learning. Moreover, the noisy surveillance data can be incorporated in the particle filtering models to inform and provide a "survival of the fittest" among particles, where each such particle posits some state for the entire compartmental model at each point in time. In this regard, it is to be underscored that particle filtering supports estimation of the entire state of the dynamic model (both latent or observable) for each point in time.

This thesis seeks to support more accurate estimation and prediction dynamics for pertussis and measles by applying a computational statistics technique that combines the best features of insights from ongoing (although noisy) empirical data and dynamic models (although fraught by systematic errors, omissions, and stochastic divergence over time) while mitigating important weaknesses of each. The use of sequential Monte Carlo methods in the form of particle filtering [84, 95, 27, 78, 102, 107, 73, 77, 91, 64, 101] has provided an effective and versatile approach to solving this problem in other infectious diseases, such as influenza. This thesis investigates the combination of particle filtering methods with the compartmental models (e.g., SEIR model) of measles and pertussis to recurrently estimate the latent state of the population with respect to the natural history of infection, to anticipate measles and pertussis evolution and outbreak transitions in the pre-vaccination era.

1.3 Thesis statement

I hypothesize that when applied to measles or pertussis transmission models, particle filtering will allow for localized estimation of model state and improve predictive accuracy of the models.

I propose to judge whether the model supports localized estimation of model state by examining if the high probability density region of the model's prediction of empirical data lies near those data. For the sake of the thesis statement, I will further consider an acceptable improvement to predictive accuracy as being achieved if the sampled discrepancy of model's predictions vs. observed data (as defined by the discrepancy in chapter 3) is reduced.

1.4 Thesis organization

The balance of this thesis is organized as follows. Chapter 2 provides the background of this research, including a literature review – related to mathematical epidemiology modelling and the machine learning method of particle filtering, two mathematical methodologies used in this thesis – the detailed mathematical introduction of the particle filtering algorithm and the deduction of the age-structured demographic model, and a brief introduction of the empirical datasets employed in this thesis. Chapter 3 introduces the particle filtering model of measles, with the aggregate demographic model. Chapter 4 introduces the particle filtering

model of measles, with age-structured demographic models (with two age groups). Chapter 5 introduces the particle filtering models of pertussis with 4 different models, including an aggregate model, an age-structured model with two age groups, and two age-structured model with 32 age groups (employing different methods in calculating the contact matrix). Chapter 6 introduces a method for classifying outbreak occurrence with the prediction results of the particle filtering models of both measles and pertussis. Chapter 7 provides a summary and conclusion of this research, and discuss the contribution of this thesis.

Chapter 2

Background

2.1 Literature review

There is a notable literature in two fields relevant to the research purpose of this thesis (incorporating the machine learning method of particle filtering with system dynamic models of pertussis and measles, to estimate and predict transmission patterns): infectious disease transmission modelling and particle filtering. The infectious disease transmission modelling literature mainly covers the mathematical characterization and simulation of communicable diseases, especially in measles and pertussis. The particle filtering component of the literature discussed here primarily covers this machine learning method, especially as it relates to infectious transmission simulation.

2.1.1 Mathematical epidemiology modelling

Mathematical models are important tools in characterizing and analyzing the transmission pattern of infectious diseases. The associated mathematical equations are formulated by several assumptions that simplify aspects the real world (e.g., contact process, infectious individuals) to variables, parameters, and formulas [2, 48]. The joint application of mathematical models and computer simulation to the study of infectious disease transmission provides a helpful approach for supporting research insight into the spread of disease within the population. This approach could also contribute to transmission prediction, determine the sensitivity of parameters, provide conceptional results (e.g., the basic reproductive number threshold R_0), evaluate intervene policies, etc.

A central idea of mathematical epidemiology models is dividing the total population into different subgroups [45, 93, 61, 2]. Thus, the mathematical epidemiology models are also named compartmental models. Hethcote (2000) [48] summarized them to a general form of $MSEIR$ model. M indicates the infants that still retain passive immunity. After losing these maternal antibodies, they will flow into the susceptible class (S). In some situations, the newborns lack passive immunity. They would correspondingly immediately enter the susceptible class S. Individuals recently infected by infectious individuals and are yet not in the infectious state will be labeled to be exposed class (E). Then, if they enter the infectious state, they will be labeled as the infectious class (I). Finally, after the end of the infectious period, these individuals will flow into the recovered class (R) with permanent or temporary infection-acquired immunity.

The first mathematical epidemiology model was formulated by Daniel Bernoulli (1760) [12], to evaluate the effectiveness of variolation against smallpox. Then after a prolonged gap, deterministic mathematical epidemiological models currently widely applied emerged in the early twentieth centuries.

Hammer (1906) [45] introduced one of the most important concepts in epidemiological mathematics, which is called the "mass action" principle by analogizing transmission contacts to chemical reaction kinetics. This principle obtains under the assumption that the population is homogeneously mixed. It posits that the number of new infectious individuals is related proportionally both to the density of susceptible individuals and density of infectious individuals per unit time per unit spatial area.

Ross (1911) [93] extended the discrete-time mathematical model of Hammer (1906) [45] to ordinary differential equations (continuous-time model), in the prevention of malaria.

Kermack and McKerdrick (1927) [61] contributed to the establishment of threshold theory. They found that a threshold density of susceptible population existed in general. The epidemic will not occur if the susceptible population density is below this threshold, even if there are a few of infectious individuals entering the community. Both this threshold theory and the "mass" action principle are the cornerstones of modern theoretical epidemiology [2].

Soper (1929) [99] employed Ross' mathematical framework of ordinary differential equations to analyze the underlying mechanisms of periodically occurring infectious diseases.

After these periods, the literature related to mathematical epidemiology models grew rapidly [8, 2, 51, 49, 50, 48, 72, 22, 63]. However, there existed two approaches for calculating the homogeneous mixing process. This situation persisted until 1995, when de Jong *et al.* [60] interpreted these two approaches and identified the correct one.

To understand this interpretation, it bears emphasis that the "force of infection" (denote as λ) characterizes the hazard rate (probability density) with which infection spreads to susceptibles in epidemiology mathematical models, and plays a central role in mathematical epidemiology. De Jong *et al.* (1995) [60] characterized two traditional approaches for characterizing the force of infection as "true mass action" and "pseudo mass action". In the approach of "pseudo mass action" [1], the force of infection was calculated by $\lambda = \beta SI$, where S is the total numbers of the susceptible individuals, I is the total number of the infectious individuals, and β is the transmission coefficient. De Jong *et al.* [60] noted that this "pseudo mass action" would violate the "mass action" proposed by Hammer (1906) [45] in the situation that the population density kept constant, while the total population size changed. However, the "pseudo mass action" fitted the situation where the population density and population size would not change much. Then, De Jong *et al.* [60] argued that "true mass action" represented force of infection as $\lambda = \beta SI/N$, where N was the total population. In short, it simply let the "force of infection" depend on the total population of N, in contrast with the "pseudo mass action" method. Finally, De Jong *et al.* [60] compared the results with these two approaches to prove the correctness of the "true mass action". They further discussed that the reason the "true mass action" did not perform well is that this homogeneous mix theory is too simple to simulate some

complex situations (e.g., considering heterogeneous contacts). However, the "true mass action" theory offers a greater robustness than does "pseudo mass action".

It is notable that the compartments included in the $MSEIR$ model [48] were not all contained in each mathematical models. Whether one or some of these classes were contained or not was determined by the characteristic of the specific disease, and also by the purpose of the models. For example, among the models that simulated measles, Bjornstad *et al.* (2002) [13] only considered the compartments of S, I and R , while Earn *et al.* (2000) [30] and Bolker and Grenfell (1995) [15] considered the classes of S, E, I and R. For those diseases lacking lifelong natural immunity, recovered individuals (the class R) will flow back to S, after their immunity disappears. The classic models for this kind of diseases (e.g., influenza) are SIRS [29, 59, 50]. During the vaccine era in which immunization is widely used to protect the susceptible individuals, a new compartment was imported into the mathematical structure, which was the vaccinated class (denoted as V). Susceptible individuals obtaining immunity by vaccination will entering into this class. Many mathematical models have been formulated considering this compartment in the vaccination era [5, 71, 50].

Although the classic mathematical epidemiology framework with an aggregated population demonstrated exceptional versatility in researching of infectious diseases, some assumptions were too simple to simulate some complex situations, such as the homogeneous mixing between susceptible and the infectious individuals. Several approaches have been employed to let the mathematical models characterize and simulate the real world more accurately.

Several infectious diseases exhibit strong seasonal fluctuation in specific areas or throughout the world, such as influenza [29], measles [30], etc. Normally, one or more specific parameters would be imported or modified to capture such seasonal fluctuation. Bolker and Grenfell (1995) [15] and Dushoff *et al.* (2004) [29] modified the transmission coefficient (β) from constant in the classic "mass action" to vary sinusoidally with time to capture the seasonal fluctuation ($\beta = \beta_0(1 + \beta_1 cos 2\pi t)$). Earn *et al.* (2000) [30] let the transmission coefficient vary with school terms (high during school terms, low otherwise). Meanwhile, the values of the transmission coefficient also changed based on the fluctuation patterns of measles in different areas (e.g., annually, biennially, etc.) – the transmission coefficients were becoming smaller, while the fluctuation intervals were bigger.

Age-structured models [50, 97, 15, 30, 48, 105] were also widely employed to simulate heterogeneous mixing in the real world, instead of the simplified homogeneous mixing. The age-structured models, initially applied in sexual transmitted diseases, were employed for other infectious diseases, especially for the childhood infectious disease, like measles [15, 97] and pertussis [50].

Lotka (1922) [68] proposed a demographic model with a stable population. This demographic model consisted of an initial-boundary problem with a partial differential equation for age-dependent population growth. Kermack and McKendrick (1927) [61] combined this demographic model with epidemic models. This mathematical epidemiology model sometimes is called the Lotka-McKendrick model or the McKendrick-von Foerster model [48].

Schenzle (1984) [97] introduced an age-structured model of pre- and post-vaccination to measles transmission. It indicated the results of this age-structured model could match the empirical data (annual numbers of reported measles cases in England and Wales) better than the global mass-action model.

Hethcote (1997) [50] introduced an age-structured pertussis model with steady-state population age structure. It combined Lotka's [68] demographic model with the pertussis epidemiology model. In this age-structured model, the population was split into 32 age groups. It consisted of 384 ordinary differential equations in total, while in each age group, there were 12 ordinary differential equations related to the total compartments.

With the growth of computational power, agent-based modelling and simulation is becoming a relatively new approach to simulate systems consisting of autonomous and interacting agents [69, 70, 90, 7, 38, 86], compared with deterministic mathematical models. Agent-based simulation can simulate heterogeneity on the level of each individual, compared with age-structured mathematical modelling, which is on the level of refined population groups. Agent-based models could also capture the complexity of the human behaviors, particularly regarding to mobility patterns and social networks [38].

Compared with the deterministic mathematical models with differential equations, the agent-based models could consider stochastic processes in the real world more easily. Osgood (2004) [81] compared the difference of the computational resource and error scaling among three methods of representing the heterogeneity – using stocks disaggregated by attribute value, agent-based disaggregation and using co-flows. This research concluded that the structured stock disaggregation models were best suited to problems with low number of attribute dimensions or a large population size, while the agent-based models suited for problems exhibit important heterogeneity with respect to medium or high numbers of attributes dimensions and population sizes. Osgood (2007) [82] provided an overview of trade-offs between the two modeling traditions – agent-based modeling and system dynamics, and argued that the most important difference was between individual-based and aggregate models, rather than the formalism used to characterize change. Osgood (2009) [83] focused on representations of multiple co-morbidities by both aggregate and individual representations. It concluded each of the methods had their own advantages and disadvantages. The individual representation exhibited high resolution and transparent representation. However, they have significant drawbacks that slow the modeling. While the aggregate representations of co-morbidities are possible, they suffer from a variety of shortcomings, such as low fidelity.

Perez and Dragicevic (2009) [86] proposed a GIS-agent based model. This model simulated the outbreak of measles in an urban area where different activities were taking place in the daily routine of each. The results of this model indicated that it could successfully generate different scenarios of outbreaks in complex and real geographic urban settings, by considering the movement and interaction of each individual in the model. Subsequently, Liu *et al.* (2015) [66] built an agent-based model to study the effects of the vaccination coverage, clustering of immunity, and contact investigations in preventing uncontrolled measles outbreaks.

Doroshenko *et al.* (2016) [24] employed agent-based modelling to study the effects of an outbreak response

immunization focused on young individuals in averting pertussis cases. The important elements of the structure of this model were drawn from the research of Hethcote [50] with a deterministic aggregate model employing differential equations. Sanstead *et al.* (2015) [96] built an age-based model to study age-related trends in the incidence of pertussis. They fitted three years of outbreak data (2004, 2008 and 2012) to the model, to identify the factors that have contributed to them. Finally, they found no single factor accounted for these observed trends, and that the agent-based modelling was a useful tool for future research.

Although the agent-based modelling approach has its own advantages compared with traditional aggregate mathematical models based on differential equations, it requires more computational and cognitive costs, which may limit the sensitivity analysis and model scope. Due to the limited resources at hand, the cost and benefit of each approach guide the choice [90].

To apply mathematical modelling to simulate infectious diseases, several parameters or processes need to be simplified, compared to their real world complexity. For example, the contact process among individuals in the real world is stochastic. However, the contact process is normally simplified to be a represented by constant contact rate parameter in deterministic models and some agent-based models. Given that simplification effort is required to quantify these parameters through several approaches. These approaches include (in approximately the decreasing order of the preference) "the use of recorded measurements, inference from related data, logic, educated guesswork, or adjustments needed to provide a better simulated fit to history" [53]. In a situation lacking recourse to detailed measurement of a parameter, while it is too important to be ignored in the model, we could following a calibration approach to estimate the value of these parameters, to adjust the model to fit available historical data. Normally, the calibration method used in the infectious diseases is performed offline [24, 96, 104, 95]. And it is notable that the off-line calibration process is typically time-consuming and inflexible in terms of handling on-going empirical data.

Models represented by differential equations could also capture the stochastic processes in the level of population groups, by combining the mathematical models (differential equations) with statistical epidemiology [71, 63, 8, 13, 44], including through the use of stochastic differential equations.

Bjornstad *et al.* (2002) [13] worked on a model that paralleled several approaches and factors in ecological population dynamics, including "intrinsic density dependence vs. extrinsic forcing, deterministic vs. stochastic dynamics, spatial coupling and synchronize, etc". The model was particularly notable for combining the SEIR deterministic model with the statistical knowledge. The model was named TSIR (Time-series Susceptible-Infected-Recovered) model, and could capture both endemic cycles and episodic outbreaks of measles. Finally, they analyzed this model with short-term empirical data on measles epidemiology drawn from 60 cities. It indicated that this model could describe and regenerate the quantitative and qualitative properties of these different types of dynamics. Grenfell *et al.* (2002) [44] as a companion paper showed how the stochastic TSIR model could capture the long-term dynamical patterns of measles. Finally, it also discussed the balance between the noises (stochastic factors) and determinism scales in the context of the highly nonlinear ecological system of measles.

2.1.2 Particle filtering method

Particle filtering is a Monte Carlo or simulation-based algorithm for recursive Bayesian posterior inference [74]. In data science, normally we need to estimate some unknown quantities from some given observed data from the real world. Bayes' theorem provides an approach to calculate the posterior distribution of the unknown quantities given observed data. In the Particle Filtering context, these observed data are most commonly time series – data arriving over time. The Bayesian method of particle filtering has the ability to incorporate such observed data on-line [25]. The particle filtering method is widely used in several areas, including "tracking, time series forecasting, online parameter learning, etc." [74].

Markov models (or Markov chains) play an important role in the analysis of time series [10]. In latent Markov models, the states in the state space are normally unobservable in the real world. Typically, the latent states are considered as being characterized by a first-order Markov chain and in a discrete time context. The chain indicates that the latent states at next timepoint are only dependent on the state at the previous timepointd the observed data at each timepoint is only dependent on the latent state at that timepoint [74, 10, 25]. The HMM (Hidden Markov Model) is a widely used algorithm for latent Markov models, in the situation where the hidden states are discrete and finite sets [89, 74].

The Kalman filter is a well known and widespread algorithm which contributes to the Bayesian inference of continuous latent state state space models, in linear and Gaussian dynamic systems [52, 74]. Prior to the introduction of the particle filtering algorithm, the extended Kalman filter (EKF) is the most popular method to perform on nonlinear estimation [58, 67]. It is notable that the extended Kalman filter still assumes that the posterior follows the Gaussian distribution [58, 67, 74]. Thus, for more general non-linear non-Gaussian state space models, the extended Kalman filter algorithm may be biased and finally lead to filter divergence [67].

Particle filtering is referred to under many different names, including Bayesian filtering, optimal (non-linear) filtering, stochastic filtering, bootstrap filtering, the condensation algorithm, interacting particle approximations and on-line inference and learning [25, 6]. The particle filtering method was advanced to be applied to a more general situation, with non-linear and non-Gaussian models. Several excellent tutorials [6, 26, 17] on particle filtering have been published in the following years. They have given systematically introductions and examples of the set of particle filtering algorithms.

Gordon *et al.* (1993) [42] is commonly regarded as the first instance of the particle filtering algorithm. The key idea of particle filtering is to represent the required posterior probability density function (PDF) of system state by a numerical method of a set of weighted samples, in contrast to the analytic method of Kalman filter, which assumes a Gaussian distribution around a maximum likelihood estimate. If the total number of samples is large enough, we consider that these weighted samples could collectively represent the PDF. With Bayesian recursive estimation, the PDF could be calculated by two recurrence equations theoretically. Finally, by incorporating sequential importance sampling / resampling [94], the particle filtering is an algorithm for propagating and updating these samples (particles) to obtain an estimate of the value of

system state (including latent states) at each time step. When considering their weighting, the distribution of these samples approximates the PDF at the corresponding time steps.

After the particle filtering algorithm was introduced by Gordon *et al.* (1993) [42], it has become a popular approach for solving estimation problems numerically on-line. Several variants of the particle filtering algorithm have been proposed under the generic framework of sequential importance sampling.

The sampling importance re-sampling filter is the one introduced by Gordon *et al.* (1993) [42]. It has the advantages that both the importance weight and the importance density could be easily evaluated and sampled. However, it also suffers from a notable weakness. As the state space variables are explored without the consideration of the observed data, this filtering algorithm could be inefficient and is sensitive to outliers [6, 87]. Moreover, it may result in a rapid loss of diversity of particles, because of the re-sampling undertaken in each iteration.

The auxiliary sampling importance re-sampling filter was introduced by Pitt and Shephard (1999) [87]. This algorithm could overcome the weakness of sampling importance re-sampling filter in terms of sensitivity to outliers. Specifically, the auxiliary sampling importance re-sampling filter was derived from the framework of sequential importance sampling, by introducing a new importance density which related the probability of the current latent states and the index given the observed data. Then, the pair of the current state variables and the index of particles of the last time step could be sampled from this importance density. Finally, a set of samples from the current state could be obtained by omitting the index in the pairs from the marginal density of current states given the observed data. Because the indexes were considered as the auxiliary variables, this algorithm then was named auxiliary sampling importance re-sampling filter.

Musso *et al.* (2001) [75] proposed the regularized particle filtering algorithm, to overcome the problem of particle impoverishment (loss of the diversity of particles), where most of the particles occupy the same point of the sate space in some specific time steps, offering a poor representation of the posterior density. The regularized particle filtering is similar to the method of sampling importance re-sampling introduced in Gordon *et al.* (1993) [42], except for the re-sampling step. Specifically, the regularized particle filtering samples from the continuous kernel density [98] of the posterior density of the current states given the observed data in the re-sampling step. However, the sampling importance re-sampling algorithm re-samples from the discrete posterior density represented by the particles.

The property of recursively calculating the density of posterior and updating weights allows the particle filtering algorithm to incorporate ongoing arrival of observed data online. It is powerful in some application fields, such as tracking. However, in the off-line situation, it may lose information if the density of the current states are only calculated based on the observations received up to the current time. The observed data received after the current time in the off-line situation also contribute to the estimation of the density of the current states (especially in the situation of state space models with high-dimensional states). Then, the approach of calculating the density of current states based on the observations in the later time steps is known as smoothing. Several previous research has been conducted related to smoothing particle filtering

algorithms. Kitagawa and Sato (2001) [62] worked on the fixed-lag approximation. It is also the simplest smoothing approach. Some more complex smoothing algorithms were also proposed, like forward filtering-backward smoothing, the generalized two-filter formula, etc. [26]. However, the implementation of the smoothing algorithm is a high cost of the computer resources, compared with the normal particle filtering algorithm without smoothing, both in memory and computational time.

The resample-move algorithm [41] was normally applied to address the degeneracy problem. Like the technique of Markov Chain Monte Carlo (MCMC) [3]. The key idea of the resample-move is to design an MCMC transition kernel (such as in Metropolis-Hastings sampler or Gibbs sampler scheme), and then apply this transition kernel to each particle, before or after the re-sample steps. By the application of the resample-move algorithm, the diversity of particles would be increased. Finally, the degeneracy problem would be reduced. This indicates that the MCMC resample-move algorithm is simple and could easily to be incorporated in any invariant of particle filtering algorithms. Cappe *et al.* (2007) [17] indicated that the MCMC schemes were found to be particularly effective in fixed-lag smoothing approaches and in static parameter estimation. Similar to the smoothing algorithm, the MCMC resample-move algorithm also requires high computational loads of the computer resources.

Although several variants of the particle filtering algorithm have been proposed, the limitations have not been eliminated. For example, the variance of posterior density would be high, if the variance of the likelihood function is high. In such cases, frequent re-sampling would be performed. This would taken to make the results of the particle filtering unreliable [26]. In high-dimensional systems with complex patterns of probability density functions, the variance of the density functions may be high, because it is difficult to be represent by simple probability density functions. Loss of the diversity of particles is a problem for a long sequence of time steps.

Andrieu *et al.* (2010) [4] proposed a particle Markov chain Monte Carlo (PMCMC) method by combining the two powerful approaches in sampling from high dimensional probability distributions – the particle filtering algorithm to sample from trajectories of states and MCMC to sample from parameter values, yielding a joint distribution of parameter values and model state trajectories. In this new method, the particle filtering was employed to generate high dimensional proposal distributions for MCMC algorithms. While this PMCMC algorithm also requires high computational loads of the computer resources, several research has been conducted to address this problem, such as parallel implementation in recent years [47, 65, 100].

2.1.3 Particle filtering in epidemiology modelling

The particle filtering algorithm has been widely used in system dynamic models since it was proposed, including in the epidemiology modelling [84, 95, 27, 55, 78, 102, 107, 73, 77, 91, 64, 101].

Ong *et al.* (2010) [77] deployed a real-time influenza epidemic surveillance system by incorporating the particle filtering to a standard and stochastic compartmental SEIR model, with the surveillance data arriving each day. Process data and forecasts were then uploaded to a public website. This research indicated that

the real-time surveillance system was capable of showing the progress of the epidemic and indicating the time of peaks. Moreover, this particle filtering model was capable of being used for the prediction when the epidemic wave ended and when a second wave appeared. Also, the longer the data fed in along the epidemic cycle, the greater the accuracy of the prediction of the particle filtering model.

Dukic *et al.* (2012) [27] incorporated Google Flu Trends data with the particle filtering method and state-space methodology to track the evolution of the epidemic process. Specifically, the research extended a classical mathematical epidemiology model (a seasonal susceptible-exposed-infected-recovered SEIR model) as the state-space model. The particle filtering model incorporated with this state-space model could track the influenza outbreaks, based on U.S. Google Flu Trends.

Dureau *et al.* (2013) [28] also extended the mathematical SEIR framework to yield a state space model and incorporated this state space model into the particle filtering and particle MCMC algorithm. And the final machine learning models were fed by the observed data of 2009 H1N1 in England. Moreover, this research used stochastic extensions (Brownian Motion) to capture the character of the time-varying parameters in the state space models. Finally, a diffusion-driven mathematical model with age structure (divided into two age groups – children and adults) was implemented in the machine learning algorithm.

Osgood and Liu (2014) [84] explored the robustness, tolerance and correctness of the particle filtering model implementing a SEIR mathematical framework and the epidemiological and measurement processes to the particle filtering algorithm with on-going ground truth data, which were generated by an agent-based models (ABM). Finally, this research compared the ground truth situation within the ABM to the results of particle filtering in the context of an aggregate, incomplete and systematically biased compartmental model, and indicated that the particle filtering approach can significantly improve model accuracy despite systematic bias, and offers the potential to be widely applied.

Yang *et al.* (2014) [107] compared six state-of-the-art filter methods in the estimation and prediction of influenza, by incorporating mathematical models with the SIRS framework and Google Flu Trends data. These six models included three particle filter methods – basic particle filter (PF) with re-sampling and regularization, maximum likelihood estimation via iterated filtering (MIF), and particle Markov chain Monte Carlo (PMCMC) – and three ensemble filters – the ensemble Kalman filter (EnKF), the ensemble adjustment Kalman filter (EAKF), and the rank histogram filter (RHF). The results showed that in the context of influenza and among the specific six filter methods, the basic particle filter method and the three ensemble methods performed better than the particle Markov chain Monte Carlo and the maximum likelihood estimation via iterated filtering.

In the previous literature, the state space models implemented in the particle filtering algorithm were all extended from the deterministic mathematical models (compartmental models). However, there is also some research explored by applying agent-base modeling as the state space model in the particle filtering algorithm.

Kreuger and Osgood (2015) [64] initially implemented an agent based model for use with the particle

filtering algorithm. The research employed three different network types in the agent-based models – random, small world, distance-based – and three particle weight update rules – no weight updating, weight updating every one time step, weight updating every five time steps. The results indicated that although the particle filtering model did not show benefits in general, the particle filtering models have the capability to compensate for mismatches between the model characterization of the dynamics of the system and the actual dynamics as captured in the ground-truth model. The authors proposed several possible hypotheses for the limitations observed in this research.

Tabataba *et al.* (2014) [101] implemented a high resolution agent-based model using realistic social contact networks with a synthetic dataset for Ebola for use with a variant of the particle filtering algorithm – smart beam particle filtering. The results showed that the smart beam particle filtering approach could find near-optimal results with a lower number of particles.

The literature related to particle filtering applied in the fields of infectious disease modelling have indicated that this filtering method could benefit the estimation of system state variables, the inference of model parameters and the prediction of transmission. Two research works related to the combination of variants of Kalman filter and epidemiology models are further reviewed as supplementary.

Qian *et al.* (2014) [88] applied an extended Kalman filter approach in an SIR aggregate mathematical epidemiology framework, by evaluating this system with a two level system – a ground truth data generated from an agent-based model featuring an empirical dynamic contact network as gathered via wireless proximity sensing and a noisy dataset of measurements. This research concluded that although the Kalman filter algorithm suffered some limitations, it still offered capacity to improve outbreak peak estimation and could compensate the inaccuracies in aggregate model structure and parameter estimates. Chen *et al.* (2012) [18] extended the SIR mathematical framework for a state space model, and then linked this aggregate state space model with an extended Kalman filter algorithm. Measles surveillance data with annual time series from 1980 to 2007 of four cities was fed into the model. The authors could employ this model to incorporate surveillance data to estimate parameters and predict unobserved elements in the model. In addition, a particle filter framework has also been applied in this paper for comparison with the results of the extended Kalman filter model. Although the extended Kalman filter has was shown to have several limitations, the results generated from the extended Kalman filter model and particle filtering models of parameter estimation were demonstrated to be similar.

2.2 Introduction of two mathematical methodologies

2.2.1 Particle filtering algorithm

Particle filtering is a general Monte Carlo sampling method for performing inference in state space models. The approach estimates time-evolving internal states of dynamic systems where random perturbations are present, and information about the state is obtained via noisy measurements made at each time step. Particle

filtering is a popular numerical method for filtering problems in general state space methods, especially in non-linear non-Gaussian scenarios [6, 74, 26]. In a general discrete-time state-space model in a particle filtering context, the states of a system evolve between measurement points k according to equations specifying how the current states are related to previous one [6, 74]:

$$x_k^N = g_k(x_{k-1}^N, \vartheta_{k-1}) \tag{2.1}$$

where x_k^N is the vector representing the states of the system at time k, N indicates the length of the vector x. ϑ_{k-1} is the vector representing state process noise. g_k is a possibly non-linear and time-dependent function describing the evolution of the state vector between measurement points. In cases of non-linear state equations (such as obtain for communicable disease models), g_k is in general not expressible in a closed form, with the mapping between successive observation points depending instead on numeric integration. However, in this project, we are dealing with a continuous-time state space system – the system of ordinary differential equations (ODEs). The state model of particle filtering in a continuous-time dynamic system can be characterized as follows:

$$\frac{dx^N}{dt} = g(x^N, \vartheta) \tag{2.2}$$

Such ODEs can be solved by numerical integration over discrete time intervals. In such case, the function g_k in Equation (2.1) represents the numerical integration of function g in Equation (2.2), that is:

$$x_k^N = g_k\left(x_{k-1}^N, \vartheta_{k-1}\right) = \int_{k-1}^{k} g(x^N(t), \vartheta(t))dt + x_{k-1}^N \tag{2.3}$$

Within this equation, it bears emphasis that g_k is in general not specifiable in closed form and the state vector x_k^N is the integration of the ODE system in the time intervals [*k-1*, *k*]. x_k^N is assumed in general to be latent and unobservable, and information concerning x_k^N is obtained from a noisy observation vector y_k^M. The measurement model describes how noisy measurements are related to the states [6]:

$$y_k^M = h_k(x_k^N, n_k) \tag{2.4}$$

Where h_k is also a possibly non-linear time-dependent function describing the measurement process and n_k is the vector representing measurement noise. M indicates the length of the observation vector y_k. This thesis focuses on cases in which the observed data are measured at evenly spaced time points, with different integer spacing optionally being allowed for different particular measurands (e.g., the total measles cases can be measured monthly, while the measles cases of each age group are measured yearly). In this dynamic system, the state model is represented by continuous-time functions, which are then sampled with Equation (2.3) in uniformly distributed discrete time to build the measurement model – Equation (2.4). On the basis of Equations (2.2)-(2.4), we apply the approach for particle filtering of discrete-time systems of Equation (2.1) to our continuous-time system without loss of generality, simply by re-scaling time. Moreover, the noise in the state model of the continuous-time system can also be characterized in numerical integration of a stochastic differential equation, which will be discussed in greater detail in Equation (2.3).

Making the simplifying assumption for exposition that all measurements are uniformly distributed and to occur at integer times, the intention of the particle filter model is to derive or sample from the sequential marginal distribution of $\left\{p(x_{0:n}^N|y_{0:n}^M)\right\}_{n=0:T}$ for all the particle numbers n up to time T, where x_0 indicates the initial value of the state vector, and there is no measurement at y_0. The technique for achieving this is a recursive formulation. To characterize the means by which the particle filter method operates, we introduce below the Recursive Bayes Filter [6], general Sequential Importance Sampling [6, 74, 84], and Resampling [6, 74, 84].

Recursive Bayes filter

Suppose we seek filtered estimates of x_k^N at time k based on the observed data $y_{1:k}^M$. From the Bayesian perspective, the tracking problem is to recursively calculate the probability of being in state x_k^N, given the empirical data $y_{1:k}^M$. It is notable that the initial probability density function (PDF) of $p(x_0^N|y_0^M) = p(x_0^N)$ is known as the prior. Thus, $p(x_k^N|y_{1:k}^M)$ could be obtained recursively by two stages: prediction and update [6].

The prediction stage: Suppose the required $\mathrm{p}(x_{k-1}^N|y_{1:k-1}^M)$is available at time *k-1*. The prediction stage calculates the prior pdf $\mathrm{p}(x_k^N|y_{1:k-1}^M)$ of the time k using state update equation g_k (see Equation (2.1) and (2.3)) to calculate state vector x_k^N based on a posited state vector x_{k-1}^N. For the current case, $\mathrm{g_k}$ is given by a numerical integration of the compartmental model. According to the Chapman–Kolmogorov equation:

$$p\left(x_k^N|y_{1:k-1}^M\right) = \int_{-\infty}^{+\infty} p\left(x_k^N, x_{k-1}^N \ |y_{1:k-1}^M\right) dx_{k-1}^N \tag{2.5}$$

By the probabilistic chain rule, in general: $p\left(AB|C\right) = \frac{p(ABC)}{P(C)} = \frac{p(A|BC)P(BC)}{P(C)} = p(A|BC)P(B|C)$

Thus Equation (2.5) can be rewritten as:

$$p\left(x_k^N|y_{1:k-1}^M\right) = \int_{-\infty}^{+\infty} p\left(x_k^N \ \middle| x_{k-1}^N, \ y_{1:k-1}^M\right) p(x_{k-1}^N|y_{1:k-1}^M) dx_{k-1}^N \tag{2.6}$$

Because for a first-order Markov-Chain, $p\left(x_k^N \ \middle| x_{k-1}^N, \ y_{1:k-1}^M\right) = \ p\left(x_k^N|x_{k-1}^N\right)$, the first term in the integral above can be rewritten, yielding the key relation:

$$p\left(x_k^N|y_{1:k-1}^M\right) = \int_{-\infty}^{+\infty} p\left(x_k^N|x_{k-1}^N\right) p(x_{k-1}^N|y_{1:k-1}^M) dx_{k-1}^N \tag{2.7}$$

It bears emphasis that Equation (2.7) can be used to sample from $p\left(x_k^N|y_{1:k-1}^M\right)$ via Monte Carlo integration.

The update stage: Given the capacity to sample from $p\left(x_k^N|y_{1:k-1}^M\right)$ provided by the prediction stage, the update stage provides a means of calculating $p\left(x_k^N|y_{1:k}^M\right)$. By the application of Bayes' rule, we can obtain:

$$p\left(x_k^N|y_{1:k}^M\right) = \ \frac{p(y_{1:k}^M|x_k^N)p(x_k^N)}{p(y_{1:k}^M)} \tag{2.8}$$

Recognizing that $p\left(y_{1:k}^M\right) = p(y_k^M,\ y_{1:k-1}^M)$, we can rewrite this as:

$$p\left(x_k^N \middle| y_{1:k}^M\right) = \frac{p(y_k^M,\ y_{1:k-1}^M | x_k^N) p(x_k^N)}{p(y_k^M,\ y_{1:k-1}^M)} \tag{2.9}$$

By application of Bayes' rule and the probabilistic chain rule as above, we obtain:

$$\begin{aligned} p\left(y_k^M,\ y_{1:k-1}^M \middle| x_k^N\right) &= p\left(y_{1:k-1}^M, y_k^M \middle| x_k^N\right) \\ &= p\left(y_{1:k-1}^M \middle| y_k^M, x_k^N\right) p\left(y_k^M \middle| x_k^N\right) \\ &= p\left(y_k^M \middle| x_k^N\right) \frac{p(y_k^M, x_k^N | y_{1:k-1}^M) p(y_{1:k-1}^M)}{p(y_k^M, x_k^N)} \end{aligned} \tag{2.10}$$

by the definition of a joint distribution: $p\left(y_{1:k-1}^M, y_k^M \middle| x_k^N\right) = p(y_{1:k-1}^M |\ y_k^M, x_k^N) p(y_k^M | x_k^N)$.

Similarly, we could get:

$$p\left(y_k^M, y_{1:k-1}^M\right) = p(y_k^M |\ y_{1:k-1}^M) p(y_{1:k-1}^M) \tag{2.11}$$

Substituting Equation (2.10) and Equation (2.11) into Equation (2.9), and cancelling the term involving $p(x_k^N)$, we obtain:

$$\begin{aligned} p\left(x_k^N \middle| y_{1:k}^M\right) &= \frac{p(y_k^M | x_k^N) p(y_k^M, x_k^N | y_{1:k-1}^M) \cancel{p(y_{1:k-1}^M)} p(x_k^N)}{p(y_k^M |\ y_{1:k-1}^M) \cancel{p(y_{1:k-1}^M)} p(y_k^M, x_k^N)} \\ &= \frac{p(y_k^M | x_k^N) p(x_k^N | y_{1:k-1}^M) p(y_k^M | x_k^N, y_{1:k-1}^M) p(x_k^N)}{p(y_k^M |\ y_{1:k-1}^M) p(y_k^M | x_k^N) p(x_k^N)} \end{aligned} \tag{2.12}$$

By the properties of a first-order Markov chain, we know that $p\left(y_k^M \middle| x_k^N, y_{1:k-1}^M\right) = p\left(y_k^M \middle| x_k^N\right)$. Cancelling the $p\left(y_k^M \middle| x_k^N\right)$ term in both numerator and denominator, we obtain:

$$p\left(x_k^N \middle| y_{1:k}^M\right) = \frac{p(y_k^M | x_k^N) p(x_k^N | y_{1:k-1}^M)}{p(y_k^M |\ y_{1:k-1}^M)} \tag{2.13}$$

where the second term in the numerator is as given in Equation (2.7) and the denominator is a normalization term also depending on that term. That normalization term is given by:

$$p(y_k^M |\ y_{1:k-1}^M) = \int_{-\infty}^{+\infty} p(y_k^M | X_k^N) p(X_k^N | y_{1:k-1}^M) dX_k^N \tag{2.14}$$

Because Equations (2.14) does not depend on x_k^N, it can be treated as a constant when sampling from $p\left(x_k^N \middle| y_{1:k-1}^M\right)$. In Equation (2.13), the density $p(y_k^M | x_k^N)$ is defined by the measurement model (2.4). Reflecting the application of Bayes' rule in Equation (2.13), it is fruitful to think of each observation as leading to Bayesian update. Within this update, the observed data y_k^M is used to update the prior density $p(x_k^N | y_{1:k-1}^M)$ by multiplication of that prior by the density $p(y_k^M | x_k^N)$ so as to obtain the posterior density $p(x_k^N | y_{1:k}^M)$ of the current state x_k^N given all previous data $y_{1:k}^M$.

The prediction stage (as given by Equation (2.7)) depends on $p(x_{k-1}^N | y_{1:k-1}^M)$ in order to determine $p\left(x_k^N \middle| y_{1:k-1}^M\right)$. By contrast, the update stage takes the output from the prediction stage and specifies (via Equation (2.13)) how to calculate $p\left(x_{k-1}^N \middle| y_{1:k-1}^M\right)$ on from it – exactly what is needed to continue on to

the next iteration of the prediction stage. Recursively calculating Equation (2.7) and (2.13) with alternations of prediction and Bayesian stages for progressively larger values of k can finally yield the solution of $\left\{p(x_{0:n}^N|y_{0:n}^M)\right\}_{n=0:T}$. For a Gaussian and linear system, these equations can be solved analytically via the well-known Kalman filter algorithm. However, in general, these two functions cannot be carried out analytically, such as for the model considered here, which exhibits both non-Gaussian and non-linear characteristics. As result, we require the approximation method of Monte Carlo Sampling.

Sequential importance sampling

Sequential importance sampling (SIS) is the most basic Monte Carlo method used to sample when the predict and update equation of the recursive Bayes filter are not analytically tractable [6]. The key idea of SIS is to estimate the posterior distribution $p(x_{0:k-1}^N|y_{1:k-1}^M)$ at a given time – for example, *k-1* – with a weighted set of samples $\left\{x_{0:k-1}^{N(i)},\ w_{k-1}^{(i)}\right\}_{i=1}^{N_s}$ (N_s is the total number of particles). It bears emphasis that $x_0^{N(i)}$ indicates the initial value of the state vector of sample *i*. The weight $\{w_{k-1}^{(i)}\}_{i=1}^{N_s}$ of the initial state is normally set to a uniformly value, and the weight are normalized at each time step, and thus $\sum_{i=1}^{N_s} w_k^{(i)} = 1$. Then recursively update these samples to obtain samples approximating the posterior distribution at the next time step $p\left(x_{0:k}^N|y_{1:k}^M\right)$. These samples are also named particles [6].

Before introducing SIS, the simpler concept of importance sampling will be demonstrated. In importance sampling, we seek to sample (draw values) from a target distribution $p(x)$. In many cases, this can be difficult – for example, $p(x)$ may not be analytically expressible, with the value of $p(x)$ for a given x requiring computation by numerical methods. Sampling via traditional methods (e.g., by computing a cumulative distribution function) can thus pose problems; in other cases, x is drawn from a space with moderate to high dimensionality, and construction of a cumulative distribution is cumbersome and inefficient. Instead, sampling from $p(x)$ can be approximated by drawing samples from a different "proposal" distribution of $q(x)$ supporting easy sampling. Suppose that we sample $\{x^{(i)}\}_{i=1}^n$ from $q(x)$ which can represent the key features of $p(x)$. Then the target density can be approximated as [6, 74, 84]:

$$\widehat{p}\,(x) = \sum_{i=1}^{n} w(x^{(i)})\delta_{X^{(i)}} \tag{2.15}$$

where $w(x^{(i)}) = \frac{p(x^{(i)})}{q(x^{(i)})}$, and $\delta_{X^{(i)}}(x)$ is a Dirac - delta mass centered at $x^{(i)}$. This "importance sampling" emphasizes (via higher weights) samples which are more heavily weighted in $p(x)$, but requires just the ability to evaluate $\frac{p(x)}{q(x)}$ for those samples drawn from $q(x)$, rather than sampling from $p(x)$ (e.g., by computing the cumulative distribution and sampling using it).

We now return to SIS – which extends the principle of importance sampling to successive times. If applied to the posterior distribution at time *k-1*, the importance sampling yields:

$$p\left(x_{0:k-1}^{N(i)}\middle|y_{1:k-1}^M\right) \approx \sum_{i=1}^{n} w_{k-1}^{(i)}\delta_{x_{0:k-1}^{N(i)}} \tag{2.16}$$

Then, if a sample $x_{0:k-1}^{N(i)}$is drawn from a proposal distribution $q(x_{0:k-1}^{N(i)}|y_{1:k-1}^{M})$, according to importance sampling, we could get:

$$w_{k-1}^{(i)} = \frac{p\left(x_{0:k-1}^{N(i)}\middle|y_{1:k-1}^{M}\right)}{q\left(x_{0:k-1}^{N(i)}\middle|y_{1:k-1}^{M}\right)} \tag{2.17}$$

The key idea of SIS is to update samples $x_{0:k-1}^{N(i)}$ and their weights $w_{k-1}^{(i)}$. Then they can approximate the posterior distribution at the next time step $p(x_{0:k}^{N(i)}|y_{1:k}^{M})$. To do this, we first assume that the value of the proposal distribution at time k can be factored by the definition of joint distribution as follows ([6]):

$$\begin{aligned} q(x_{0:k}^{N}|y_{1:k}^{M}) &= q(x_{k}^{N}|x_{0:k-1}^{N}, y_{1:k}^{M})q(x_{0:k-1}^{N}|y_{1:k}^{M}) \\ &= q(x_{k}^{N}|x_{0:k-1}^{N}, y_{1:k}^{M})q(x_{0:k-1}^{N}|y_{1:k-1}^{M}) \end{aligned} \tag{2.18}$$

where the final term on the right is simply that being calculated but at the previous time step. This recurrence relation specifies a formulation that expresses the value of the proposal distribution at time k in terms of its value at the previous time $q(x_{0:k-1}^{N}|y_{1:k-1}^{M})$ and the probability (density) of the updated state taking into account the new measurement at time k, which is $q(x_{k}^{N}|x_{0:k-1}^{N}, y_{1:k}^{M})$. If a sample (particle) is drawn with a weight $w_{k-1}^{(i)}$ from the importance-sampling approximated distribution $q(x_{0:k-1}^{N}|y_{1:k-1}^{M})$ at time k-1, then the task is to formulate an equation to update the weight of that particle to draw from the target distribution $q(x_{0:k}^{N}|y_{1:k}^{M})$ for time k.

According to Equation (2.13) we could get:

$$\begin{aligned} p\left(x_{0:k}^{N}|y_{1:k}^{M}\right) &= \frac{p(y_{k}^{M}|x_{0:k}^{N})p(x_{0:k}^{N}|y_{1:k-1}^{M})}{p(y_{k}^{M}|\ y_{1:k-1}^{M})} \\ &= \frac{p(y_{k}^{M}|x_{0:k}^{N})p(x_{k}^{N},\ x_{0:k-1}^{N}|y_{1:k-1}^{M})}{p(y_{k}^{M}|\ y_{1:k-1}^{M})} \\ &= \frac{p\left(y_{k}^{M}\middle|x_{k}^{N}, x_{0:k-1}^{N}\right)p\left(x_{k}^{N}\middle|\ x_{0:k-1}^{N}, y_{1:k-1}^{M}\right)p(x_{0:k-1}^{N}|y_{1:k-1}^{M})}{p(y_{k}^{M}|\ y_{1:k-1}^{M})} \end{aligned} \tag{2.19}$$

Because the process is assumed to be well characterized by a first-order Markov chain, we know that:

$$p\left(y_{k}^{M}\middle|x_{k}^{N}, x_{0:k-1}^{N}\right) = p\left(y_{k}^{M}\middle|x_{k}^{N}\right) \tag{2.20}$$

$$p\left(x_{k}^{N}\middle|\ x_{0:k-1}^{N}, y_{1:k-1}^{M}\right) = p\left(x_{k}^{N}\middle|\ x_{k-1}^{N}\right) \tag{2.21}$$

Substituting Equation (2.20) and (2.21) into Equation (2.19), we obtain:

$$p(x_{0:k}^{N}|y_{1:k}^{M}) = \frac{p(y_{k}^{M}|x_{k}^{N})p(x_{k}^{N}|x_{k-1}^{N})p(x_{0:k-1}^{N}|y_{1:k-1}^{M})}{p(y_{k}^{M}|y_{1:k-1}^{M})} \tag{2.22}$$

According to Equation (2.14), we know that the denominator of Equation (2.22) is constant. Then, by substituting Equation (2.22) and (2.18) into (2.17), we could get:

$$w_{k}^{(i)} \propto \frac{p\left(y_{k}^{M}\middle|x_{k}^{N(i)}\right)p\left(x_{k}^{N(i)}\middle|x_{k-1}^{N(i)}\right)p\left(x_{0:k-1}^{N(i)}\middle|y_{1:k-1}^{M}\right)}{q\left(x_{k}^{N(i)}\middle|x_{0:k-1}^{N(i)}, y_{1:k}^{M}\right)q\left(x_{0:k-1}^{N(i)}\middle|y_{1:k-1}^{M}\right)} \tag{2.23}$$

Finally, recognizing that the final terms in the numerator and denominator are as given for $w_{k-1}^{(i)}$ in Equation (2.17), we obtain ([74], [6]):

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p\left(y_k^M \middle| x_k^{N(i)}\right) p\left(x_k^{N(i)} \middle| x_{k-1}^{N(i)}\right)}{q\left(x_k^{N(i)} \middle| x_{0:k-1}^{N(i)}, y_{1:k}^M\right)} \tag{2.24}$$

Because process g_k is Markovian, , then we can get$q\left(x_k^{N(i)} \middle| x_{0:k-1}^{N(i)}, y_{1:k}^M\right) = q\left(x_k^{N(i)} \middle| x_{k-1}^{N(i)}, y_k^M\right)$.

Then, in this case, the recursively updated equations simplify to:

$$x_k^{N(i)} \sim q\left(x_k^{N(i)} \middle| x_{k-1}^{N(i)}, y_k^M\right) \tag{2.25}$$

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p\left(y_k^M \middle| x_k^{N(i)}\right) p\left(x_k^{N(i)} \middle| x_{k-1}^{N(i)}\right)}{q\left(x_k^{N(i)} \middle| x_{k-1}^{N(i)}, y_k^M\right)} \tag{2.26}$$

Finally, the general algorithm of Sequential Importance Sampling as applied to particle filtering can be summarized as follows. Let $\mathrm{N_s}$ be the number of particles, $X_k^{N(i)}$ the state vector of particle i at time k ([74], [84], [6], [26]).

Algorithm 1: SIS particle filter

1. At time k=0:

(1) Sample $X_0^{N(i)}$ from $q_0(x_0^N)$;

(2) Compute a *weight* for each particle $w_0^{(i)} = \frac{1}{N_s}$. It indicates that the weight at initial time step follows uniform distribution.

2. At time $k \geq 1$, perform a recursive update as follows:

(1) Advance the sampled state by sampling $X_k^{N(i)} \sim q_k(x_k^N | y_k, X_{0:k-1}^{N(i)})$ and set $X_{0:k}^{N(i)} = (X_{0:k-1}^{N(i)}, X_k^{N(i)})$;

(2) Update the weights to reflect the probabilistic and state update models as follows:

$$w_k^{(i)} = W_{k-1}^{(i)} \frac{p(y_k^M | X_k^{N(i)}) p(X_k^{N(i)} | X_{k-1}^{N(i)})}{q(X_k^{N(i)} | X_{k-1}^{N(i)}, y_k^M)}.$$

Normalize the weights $W_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_s} w_k^{(i)}}$

The degeneracy problem and resampling

The SIS particle filter commonly suffers from a strong degeneracy problem – as the algorithm continues, many – and eventually most – particles will develop a negligible weight, by iterating Equation (2.25) and (2.26). This occurs because we are sampling in a high dimensional space, using a myopic proposal distribution [74].

Thus, we follow many other implementations by seeking quantify the degree of degeneracy using the effective sample size [74], so as to determine the acuteness of the problem. We estimate the degree of degeneracy at time k using the following measure [6, 74, 84]:

$$S_{\text{eff}} = \frac{1}{\sum_{i=1}^{N_s} \left(w_k^{(i)}\right)^2} \tag{2.27}$$

Where a smaller S_{eff} means a larger variance of the weights, hence more degeneracy. There are two main solutions to the degeneracy problem: resampling and choosing a good proposal distribution. Below we introduce resampling, which was used in this research.

The key idea underlying resampling is a form of "survival of the fittest". To achieve this, the resampling step will monitor the effective sample size. Then, whenever it drops below a threshold, we will create new particles in which the ones with higher weight tend to reproduce, and those the particles with low weight tend to die out. The new particles inherit their parent's values but carry uniform and normalized weight. Specifically, to perform resampling for N_s particles, we sample the new particles from the set of the old particles, where the probability that a sample is drawn from the old particles occurs with a probability proportional to the weight of that particle. Following sampling, the weight of each of the new particles is set to $\frac{1}{N_s}$.

Finally, we can get the generic particle filter algorithm by combining algorithm 1 and resampling step as follows:

Algorithm 2: Generic particle filter

1. At time k=0:

(1) Sample $X_0^{N(i)}$ from $q_0(x_0^N)$;

(2) Compute a *weight* for each particle $w_0^{(i)} = \frac{1}{N_s}$. It indicates that the weight at initial time step follows uniform distribution.

2. At time $k \geq 1$, perform a recursive update as follows:

(1) Advance the sampled state by sampling $X_k^{N(i)} \sim q_k(x_k^N | y_k, X_{0:k-1}^{N(i)})$ and set $X_{0:k}^{N(i)} = (X_{0:k-1}^{N(i)}, X_k^{N(i)})$;

(2) Update the weights to reflect the probabilistic and state update models as follows:

$$w_k^{(i)} = W_{k-1}^{(i)} \frac{p(y_k^M | X_k^{N(i)}) p(X_k^{N(i)} | X_{k-1}^{N(i)})}{q(X_k^{N(i)} | X_{k-1}^{N(i)}, y_k^M)}.$$

Normalize the weights $W_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_s} w_k^{(i)}}$

(3) Calculate the S_{eff} : $\frac{1}{\sum_{i=1}^{N_s} \left(w_k^{(i)}\right)^2}$

(4) If $S_{\text{eff}} < S_T$ (S_T is the minimum effective sample size – the threshold of resampling), perform resampling to get a new set of $X_k^{N(i)}$. And set the weight of the new particles as $\frac{1}{N_s}$.

2.2.2 The demographic model

The generic exponential growth demographic model

** The contents in this section primarily follow the derivation and exposition from the research work of Hethcote (1997) [50]**

A widely used initial-boundary-value problem of the age-structured model [68, 50, 54] is employed in this thesis. Suppose $A(a,t)$ is the age distribution function, with variable a related to age, and variable t related to time. Then, the partial differential equation for the population is as follows[54, 50]:

$$\frac{\partial A(a,t)}{\partial a} + \frac{\partial A(a,t)}{\partial t} = -\mu(a)A(a,t) \tag{2.28}$$

where $\mu(a)$ is the death rate of age a. Suppose $v(a)$ is the birth rate function of age a. It is notable that the birth rate $v(a)$ is not dependent on time t, that is, the birth rate of any age a is invariant over time. Then, the birth population at time t is:

$$A(0,t) = \int_0^\infty v(a)A(a,t)da \tag{2.29}$$

Equation (2.29) is the boundary condition of the partial equation (2.28). And the initial condition of (2.28) can be given by $A(a,0) = N(a)$, where $N(a)$ is the population distribution of age a at time $t = 0$.

The idea of building the age-structured demographic model is that we split the model via age into n intervals $[a_{i-1}, a_i), i \in [1,n]$, where $a_0 = 0$, and $a_n = \infty$. Then, in each interval, we could get an ordinary differential equation over time t by integrating $A(a,t)$ at the age interval $[a_{i-1}, a_i)$.

In this research, the population is assumed to be subject to exponential growth, with the equation of the age distribution with time being given by:

$$A(a,t) = e^{qt}N(a) \tag{2.30}$$

Under these assumptions, the partial differential equation (2.28) could be transformed into n ordinary differential equations in the age interval $[a_{i-1}, a_i), i \in [1,n]$:

$$N_i(t) = \int_{a_{i-1}}^{a_i} A(a,t)da = \int_{a_{i-1}}^{a_i} e^{qt}N(a)da = e^{qt}N_i(0) \tag{2.31}$$

where $N_i(t)$ is the total population of each age group $[a_{i-1}, a_i), i \in [1,n]$ at time t, while $N_i(0)$ is the initial value of $N_i(t)$ at time $t = 0$. Similarly, we assume that the death rate $\mu_i = \int_{a_{i-1}}^{a_i} \mu(a)da$, and the birth rate $v_i = \int_{a_{i-1}}^{a_i} v(a)da$.

Via substitution of Equation (2.30) into (2.28), it can be recognized that:

$$\frac{dN(a)}{da} = -(q + \mu(a))N(a) \tag{2.32}$$

Solving the ordinary differential Equation (2.32) in the interval of $[a_{i-1}, a_i)$ and assuming as the initial condition the population in age a_{i-1}, which is $N(a_{i-1})$ yields the following:

$$N(a) = N(a_{i-1})e^{-(q+u_i)(a-a_{i-1})} \tag{2.33}$$

Integrating Equation (2.33) in interval $[a_{i-1}, a_i)$, the total population $N_i(t)$ in the age interval $[a_{i-1}, a_i)$ can be demonstrated to be:

$$N_i(0) = \int_{a_{i-1}}^{a_i} N(a)da = \int_{a_{i-1}}^{a_i} A(a,t)da = \frac{N(a_{i-1})(1 - e^{-(q+\mu_i)(a_i - a_{i-1})})}{q + \mu_i} \tag{2.34}$$

Because there are n age groups in the model, we could define a constant variable of aging rate c_i for each age group i ($i \in [1, n]$), where $N(a_i) = c_i N_i(0)$. Then, according to Equation (2.33) and (2.34), it can be recognized that:

$$c_i = \frac{N(a_i)}{N_i(0)} = \frac{q + \mu_i}{e^{(q+\mu_i)(a_i - a_{i-1})} - 1} \tag{2.35}$$

Consider the Taylor expansion of $e^{(q+\mu_i)(a_i - a_{i-1})}$, which is $e^{(q+\mu_i)(a_i - a_{i-1})} = 1 + (q + \mu_i)(a_i - a_{i-1}) + \frac{(q+\mu_i)(a_i - a_{i-1})^2}{2!} + ...$ When $a_i - a_{i-1}$ or $q + \mu_i$ is a small value, then we can calculate c_i via the following approximation:

$$c_i \approx \frac{1}{a_i - a_{i-1}} \tag{2.36}$$

It is notable that the rate constant for the final age group $c_n \approx \frac{1}{a_n - a_{n-1}} = \frac{1}{\infty - a_{n-1}} \approx 0$.

Taking the above into account, if we integrate Equation (2.28) over the age interval of $[a_{i-1}, a_i)$, we get:

$$\int_{a_{i-1}}^{a_i} \frac{\partial A(a,t)}{\partial a}da + \int_{a_{i-1}}^{a_i} \frac{\partial A(a,t)}{\partial t}da = \int_{a_{i-1}}^{a_i} (-\mu(a)A(a,t))da \tag{2.37}$$

according to Equation (2.30), the first term of (2.37) could be:

$$\int_{a_{i-1}}^{a_i} \frac{\partial A(a,t)}{\partial a}da = \int_{a_{i-1}}^{a_i} \frac{\partial e^{qt}N(a)}{\partial a}da = e^{qt}[N(a_i) - N(a_i - 1)] \tag{2.38}$$

substituting $N(a_i) = c_i N_i(0)$ to Equation (2.30) yields:

$$\int_{a_{i-1}}^{a_i} \frac{\partial A(a,t)}{\partial a}da = e^{qt}[c_i N_i(0) - c_i N_{i-1}(0)] \tag{2.39}$$

substituting Equation (2.31) to Equation (2.39), we obtain:

$$\int_{a_{i-1}}^{a_i} \frac{\partial A(a,t)}{\partial a}da = c_i N_i(t) - c_i N_{i-1}(t) \tag{2.40}$$

according to Equation (2.31), the second term of Equation (2.37) can be written as:

$$\int_{a_{i-1}}^{a_i} \frac{\partial A(a,t)}{\partial t}da = \int_{a_{i-1}}^{a_i} \frac{\partial e^{qt}N(a)}{\partial t}da = \frac{de^{qt}}{dt}\int_{a_{i-1}}^{a_i} N(a)da = \frac{de^{qt}}{dt}N_i(0) = \frac{dN_i(t)}{dt} \tag{2.41}$$

Substituting Equation (2.40) and (2.41) into Equation (2.37), we obtain:

$$\frac{dN_i(t)}{dt} = c_i N_{i-1}(t) - (c_i + \mu_i)N_i(t) \tag{2.42}$$

Recall that the boundary condition of Equation (2.42) is given by the total inflow of newborns: $\sum_{j=1}^{n} v_j N_j(t)$. Finally, the ordinary differential equations of the demographic model are as follows:

$$\begin{aligned} \frac{dN_1(t)}{dt} &= \sum_{j=1}^{n} v_j N_j(t) - (c_1 + \mu_1)N_1(t) \\ \frac{dN_i(t)}{dt} &= c_{i-1}N_{i-1}(t) - (c_i + \mu_i)N_i(t), \quad i \geq 2 \end{aligned} \tag{2.43}$$

The equilibrium demographic model

In this thesis, I assume that the population is in equilibrium because the empirical data of population (population in Saskatchewan from 1921 to 1956) does not change dramatically (Figure 2.3). This approximation assumes that the total population $N_i(t)$ of age group i will stay stable (remain invariant) with time t. Thus, the rate constant q is 0. And according to Equation (2.43), for this simplified context, we obtain:

$$\begin{aligned} \frac{dN_1(t)}{dt} &= \sum_{j=1}^{n} v_j N_j(t) - (c_1 + \mu_1) N_1(t) = 0 \\ \frac{dN_i(t)}{dt} &= c_{i-1} N_{i-1}(t) - (c_i + \mu_i) N_i(t) = 0, \quad i \geq 2 \end{aligned} \tag{2.44}$$

Then, the death rate u_i could be calculated as follows:

$$\begin{aligned} \mu_1 &= \frac{\sum_{j=1}^{n} v_j N_j(t) - c_1 N_1(t)}{N_1(t)} \\ \mu_i &= \frac{c_{i-1} N_{i-1}(t) - c_i N_i(t)}{N_i(t)}, \quad i \geq 2 \end{aligned} \tag{2.45}$$

2.3 Introduction of empirical data

2.3.1 The surveillance data

This thesis benefits from the fact that both pertussis and measles are formally classified as notifiable illnesses for the province of Saskatchewan (Canada). In this thesis, measles and pertussis reporting data for Saskatchewan are used as empirical data for the particle filtering models. These data were public aggregate data obtained from the Government of Saskatcewan's "Public Health Annual Report of Saskatchewan [21]. This thesis employs four datasets drawn from that report – for each of measles and pertussis, datasets are used for monthly reported cases aggregated across the total population, and yearly reported cases in each age group. In the yearly empirical datasets, the annual reported cases are split into different age groups. Within this dataset, the age stratification is inconsistent. As a result, the splitting in some years fails to precisely match the stratification of the age groups in the models. For these cases, I proportionally split the yearly empirical reported cases into overlapping age groups within the model.

This study employs measles and pertussis reported cases in Saskatchewan specifically during the pre-vaccination era. The time of the monthly empirical data extends from Jan. 1921 to Dec. 1956, with the dataset offering a total of 432 records for each pathogen. The yearly age-specific data extend from 1925 to 1956, reflecting the fact that reporting of age-specific data was only started in 1925. Every record contains four features – Date, reported cases and population size. To make them consistent with the population size of the dynamic model – the average population from 1921 to 1956 (863,545) – the reported cases are normalized to the same population size as the model, thus yielding estimated incidence rates rather than incident case counts. The monthly reported cases of measles and pertussis are shown in Figure 2.1 and Figure

2.2, relatively; it can be readily appreciated that the time series demonstrate the classic patterns of waxing and waning incorporating both stochastic and regular features characteristic of many childhood infectious diseases in the pre-vaccination era.

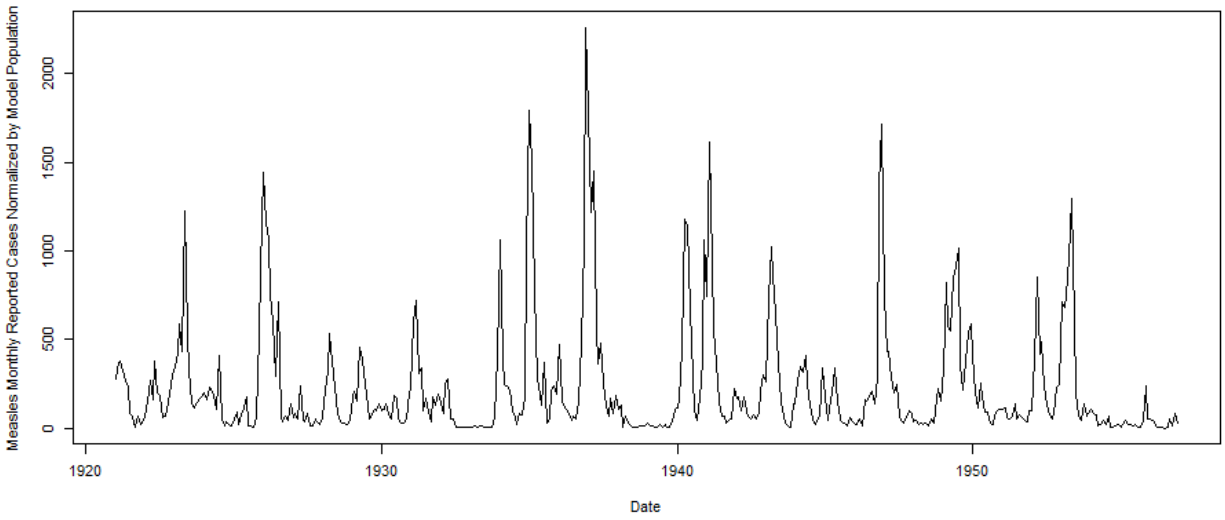


Figure 2.1: The monthly reported measles cases in Saskatchewan from 1921 to 1956 normalized by the population employed in the model (863,545)

2.3.2 The demographic data

In this thesis, the demographic parameters play a significant role in the models, and especially in the age structured models used. The parameters related to the population are abstracted from the empirical population of Saskatchewan from 1921 to 1956 [16]. The empirical demographic data indicate that the total population of Saskatchewan does not show drastic fluctuation [16] over the year range from 1921 to 1956. During these years, the empirical population lie in the interval from 757,000 to 932,000 (2.3). For this thesis, I let the model population constantly stay in 863,545, which is the average population among 1921 to 1956. It is notable that I employ the equilibrium population model – the total population and population among each age group (in the age-structured models) stay the same, rather than change. Similarly, the model retains as invariant the value of the population in each age group, according to the average population among 1921 to 1956 in within the Saskatchewan age pyramid [16]. The population of Saskatchewan from 1921 to 1956 of each age is depicted in Figure 2.3.

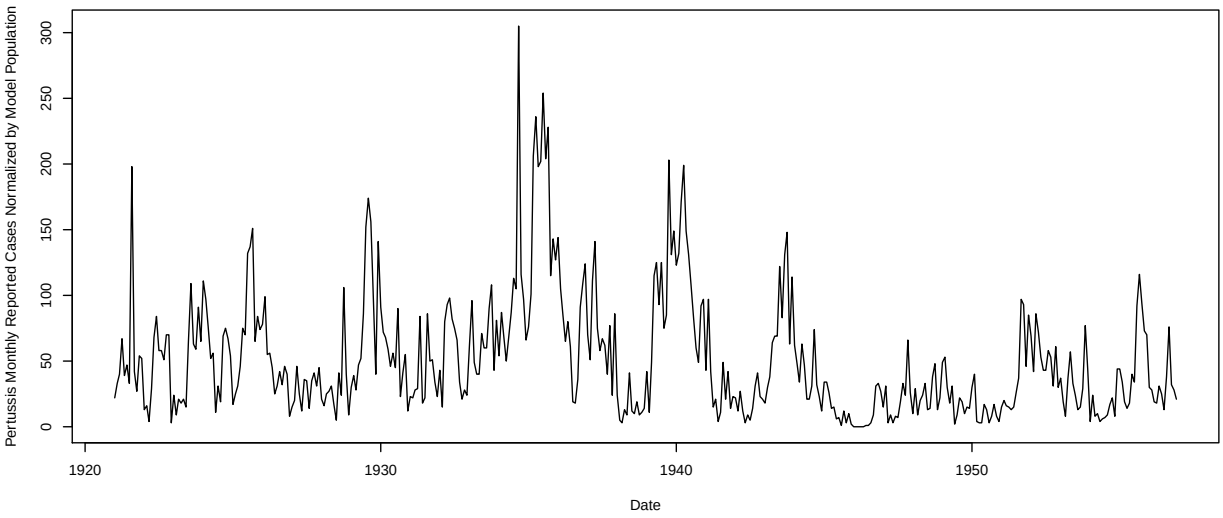


Figure 2.2: The monthly reported pertussis cases in Saskatchewan from 1921 to 1956 normalized by the population employed in the model (863,545)

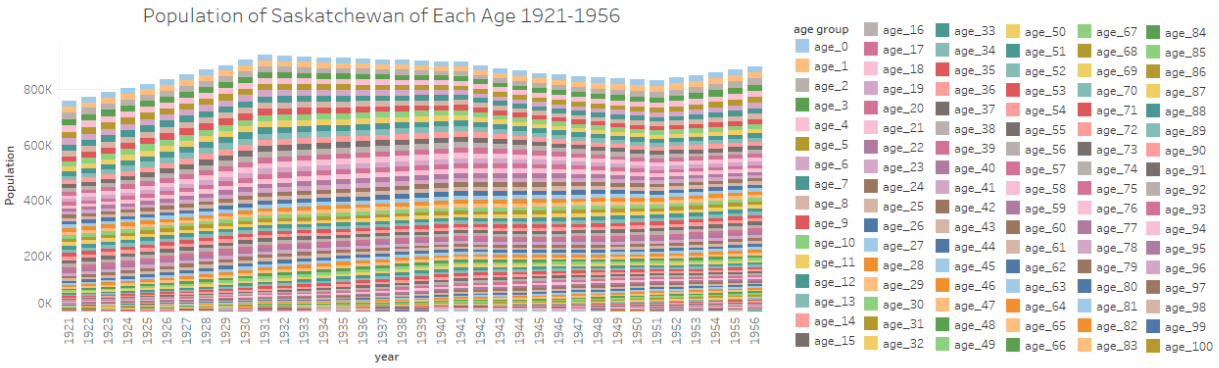


Figure 2.3: The total population in Saskatchewan from 1921 to 1956 of each age.

Chapter 3

Measles aggregate population model of particle filtering

3.1 Introduction

Measles is a highly contagious viral disease. It remains one of the leading causes of death among young children globally, and has imposed a significant morbidity and mortality burden where vaccination coverage is low [80]. About 30% of all cases of measles have one or more complications including diarrhea, otitis media, pneumonia or encephalitis. Measles mortality was estimated to be 0.2% in the United States between 1985 and 1992 [106]. Prior to widespread vaccination, measles caused an estimated 2.6 million deaths each year [80]. In 2016, approximately 89,780 people died from measles - mostly children under the age of 5 [80]. The Americas has become the first region in the world to have eliminated measles [103]. However, the importation of cases from other regions leads to outbreaks in unimmunized and under-immunized populations. Understanding measles outbreak patterns can aid in forecasting and help public health agencies to design intervention strategies to prevent and control it, such as setting outbreak response measures, setting vaccination targets, and allocating financial and human resources, etc.

Measles is a highly contagious virus. The virus of measles lives in the nose and throat mucus of an infected person, which can spread to the other individuals through coughing and sneezing [36]. It can normally live for up to two hours in the air or on objects and surfaces [106]. The measles transmission is primarily from person to person by large respiratory droplets [106]. The CDC Reports that "if other people breathe the contaminated air or touch the infected surface, then touch their eyes, noses, or mouths, they can become infected. Measles is so contagious that if one person has it, 90% of the people close to that person who are not immune will also become infected" [36]. After 7 to 14 days, the symptoms of measles generally appear [35]. The symptoms of measles normally include high fever, cough, runny nose and red and watery eyes [35]. Three to five days after symptoms begin, a rash (flat red spots that appear on the face, neck, trunk, arms, legs and feet) usually breaks out [35]. Then the infected individuals can spread measles to others from four days before through four days after a characteristic rash appears [36]. In temperate areas, measles shows a seasonal pattern that it occurs primarily in late winter and spring [106]. Finally, it is notable that measles is a disease of humans and the virus of measles is not spread by any other animal species [36].

Surveillance data and simulation models are two of the most prominent contemporary methods used to estimate and predict outbreak patterns of measles. However, both of these methods have notable shortcomings. Surveillance data are normally very noisy. And it is difficult to secure quantitatively rigorous insights into the future evolution of the outbreak pattern by only employing the surveillance data. Finally, surveillance data can not be used to investigate counterfactuals. Although dynamic modeling is increasingly widely used, it also exhibits notable disadvantages. Firstly, a calibration process for a dynamic model is typically undertaken manually on a one-time basis. The laborious nature of this process renders it difficult for a dynamic model to incorporate ongoing arriving data. Secondly, a dynamic model represents a simplified characterization of processes in the real world. It is inevitable that a dynamic model often omits, simplifies and mis-estimates some factors. As characterized in the introduction of particle filtering in previous chapters, particle filtering, the machine learning algorithm based on the idea of Monte Carlo approximation and for recursive Bayesian inference [74], can link the system dynamics models of measles and pertussis to the empirical datasets (noisy surveillance data) to compensate for weaknesses of both.

This chapter seeks to support more accurate estimation and prediction of measles dynamics by applying a computational statistics technique (particle filtering) that combines the best features of insights from ongoing (although noisy) empirical data and dynamic models (although fraught by systematic errors, omissions, and stochastic divergence over time) while mitigating important weaknesses of each. Specifically, this chapter investigates the combination of particle filtering methods with a compartmental model (SEIR model) with aggregate population of measles to recurrently estimate the latent states of the population with respect to the natural history of infection with measles, to anticipate measles evolution and outbreak transitions in pre-vaccination era. Finally, a particle filtering model supports the estimation of the entire state in the system dynamics model across all the entire simulation time horizon by incorporating incoming empirical data. This asset supports researchers by providing research insight related to the latent states (such as the historical Susceptible, Infectious, Exposed, Recovered population, etc. in each time point during incorporating the empirical data) in the dynamics system of measles.

3.2 The mathematical dynamic model

This project employs a measles SEIR model [30] as the disease transmission component of the state-equation model in particle filtering. A time unit of months is used, so as to be consistent with the empirical data [21]. This model, which can be found in [30], contains 4 state variables: (S-Susceptible, E-Exposed, I-Infectious, R-Recovered). The original model [30] makes use of a formulation in which each state variable is of unit dimension, representing a fraction of the entire population. However, for the sake of comparison against empirical data, the model in this project is represented in a re-dimensionalized fashion, with the state variables representing counts of persons. In the first step, I re-dimensionalized the original model. The

resulting aggregate compartmental equations are as follows:

$$\begin{aligned}
\frac{dS}{dt} &= Nv - (\beta\frac{I}{N} + \mu)S \\
\frac{dE}{dt} &= \beta\frac{I}{N}S - (\sigma + \mu)\,E \\
\frac{dI}{dt} &= \sigma E - (\gamma + \mu)\,I \\
\frac{dR}{dt} &= \gamma I - \mu R
\end{aligned} \tag{3.1}$$

The meaning of the states and parameters are as follows: S, E, I and R are the count of Susceptible, Exposed, Infectious and Recovered people in the population, respectively. N is the total number of people, and N=S+E+I+R. v is the birth rate (of unit 1/Month) and μ is the death rate (also of unit 1/Month). σ^{-1} and γ^{-1} are the mean incubation and infectious periods (in months) of the disease, respectively. β is the rate of effective contact between individuals, and reflects both the contact rate and transmission probability (β = contact rate $\times$ transmission probability), and is thus of unit 1/Month. The population size N was obtained from Saskatchewan during the years from 1921 to 1956. I take the mean of the population across these years to be the value of parameter N. This is associated with the empirical dataset [21]. As noted below, while β was treated in [30] as a constant, I take advantage of particle filtering by allowing it to vary over the course of the timeframe explored. The other values of parameters are obtained from [30].

3.3 Particle filter implementation

3.3.1 The state space model

The SEIR model is employed as the governing equations underlying the state space model (denoted as g_k) of particle filtering, which is introduced in Equation (3.1). Then each particle at time k, noted as $X_k^{N(i)}$, represents a complete copy of the system states at that point of time. Except for the basic states in the SEIR model (S, E, I, R), models of infection transmission are often related to more complex dynamics – such as parameters evolving according to stochastic processes.

In the aggregate model (i.e., the model not stratified by age), three essential stochastic processes are considered. Firstly, I consider changes in the transmissible contact rate linking infectious and susceptible persons, which is represented by the parameter β. A second area in which I consider parameter evolution is with respect to the disease reporting process. Specifically, to simulate this process, a parameter – representing the probability that a given measles infectious case is reported C_r, and a state I_m – calculating the accumulative measles infectious cases per unit time (per Month in this project) – are implemented. Finally, the model includes a stochastic process (specifically, a Poisson process) associated with incidence of infection. This process reflects the small number of cases that occur over each small unit of time (Δt). The stochastics associated with these factors represents a composite of two factors. Firstly, there is expected to both be stochastic variability in the measles infection processes and some evolution in the underlying transmission

dynamics in terms of changes in reporting rate, and changes in mixing. Secondly, such stochastic variability allows characterization of uncertainty associated with respect to model dynamics – reflecting the fact that both the observations and the model dynamics share a high degree of fallability. Given an otherwise deterministic simulation model such as that considered here, there is a particular need to incorporate added stochastic variability in parameters and flows capture the uncertainty in simulation results.

To estimate the changing values of these two stochastic parameters (β and C_r) and to investigate the capacity of the particle filter to adapt to parameters whose effective values evolve over simulation, the state associated with each particle includes the contact transmission rate β and reported rate C_r. Thus, each particle in this project is associated with a state vector $x = [S, E, I, R, \beta, C_r, I_m]^T$.

Reflecting the fact that the transmissible contact rate β varies over the entire range of positive real numbers, I treat the natural logarithm of the transmissible contact rate (denoted by β) as undergoing a random walk according to a Wiener Process (Brownian Motion) [56, 28]. The stochastic differential equation of transmissible contact rate can thus be described according to Stratonovich notation as:

$$dln(\beta) = s_\beta dW_t \tag{3.2}$$

where dW_t is a standard Wiener process following the normal distribution with 0 of mean and unit rate of variance. Then, $dln(\beta)$ follows the normal distribution with 0 of mean and variance ${s_\beta}^2$. In this project, I selected an initial value of β following a uniform distribution in the interval [40, 160) across all particles. Unless otherwise noted, the constant value of the diffusion coefficient s_β used is 0.8.

Over the multi-decadal model time horizon, and particularly on account of shifting risk perception, there can be notable evolution in the degree to which infected individuals or their guardians seek care. To capture this evolution, I consider evolution in the fraction of underlying measles cases that are reported (denoted by C_r). Reflective of the fact that C_r varies over the range [0,1], I characterize the logit of C_r as also undergoing Brownian Motion according to Stratonovich notation as:

$$d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t \tag{3.3}$$

where the initial value of C_r among particles follows a uniform distribution in the interval [0.11, 0.15). The diffusion coefficient (s_r) associated with the perturbations to the logit of C_r over dt is selected to be a constant 0.03 across all particles in this project.

To incorporate the empirical data, I further implement a convenience state I_m with respect to the cumulative count of infectious cases per unit time (Month). The state I_m is different from the state I. Specifically, the state of an cumulative count of infectious cases per unit time I_m only considers all the inflows to the infectious state and without all the outflows, to simulate the same process of reporting the measles cases in the real world. The cumulative infectious cases I_m of measles at time k is:

$$I_{mk} = \int_{k-1}^{k} (\sigma E) dt \tag{3.4}$$

Then, the reported cases at time k calculated by the model would be:

$$I_{rmk} = I_{mk} C_r \tag{3.5}$$

In total, the compartmental model without age stratification is the combination of the classic SEIR model and these three stochastic processes:

$$\begin{aligned}
&\frac{dS}{dt} = Nv - A_1 - \mu S \\
&\frac{dE}{dt} = A_1 - (\sigma + \mu) E \\
&\frac{dI}{dt} = \sigma E - (\gamma + \mu) I \\
&\frac{dR}{dt} = \gamma I - \mu R \\
&d(ln\beta) = s_\beta dW_t \\
&d(ln(\frac{C_r}{1 - C_r}) = s_r dW_t \\
&I_{mk} = \int_{k-1}^{k} (\sigma E) dt \\
&I_{rmk} = I_{mk} C_r \\
&A_1 = \frac{Poisson(\beta \frac{I}{N} S \Delta t)}{\Delta t}
\end{aligned} \tag{3.6}$$

To solve the system above, I made use of Euler integration with a time step of 0.01 Month (i.e., Δt= 0.01 Month in Equation (3.6)).

3.3.2 The measurement model

As introduced in particle filtering tutorials [6, 26], the measurement model characterizes the relationship between the measured data and the model. In this project, I denote the measurement vector as y_k^M, where M indicates the length of the measurement vector.

In the aggregate model, the measurement vector y_k^M is one-dimensional (representing empirical dataset of monthly reported measles infected cases), that is, M equals 1. I denote the value of empirically reported cases as I_{em}. Then, at time k, the measurement model can be represented as:

$$I_{emk} = I_{rmk} + n_{mk} \tag{3.7}$$

where I_{rmk} is calculated by the states of I_m and C_r in the space transition model of Equation (3.6), and n_{mk} is the measurement noise related to the monthly reported cases.

It is notable that the entire state ($[S, E, I, R, \beta, C_r, I_m]^T$) in the state transition model (Equation (3.6)) is latent. That is, such state variables are not directly observable in the real world. However, on the basis of this hidden state, the particle filtering framework supports characterization of model analogues for quantities in the empirical dataset. Finally, the particle filtering model can be grounded by the noisy observed data,

while the entire state in the particle filtering model can be estimated across the whole time horizon in which empirical data is incorporated.

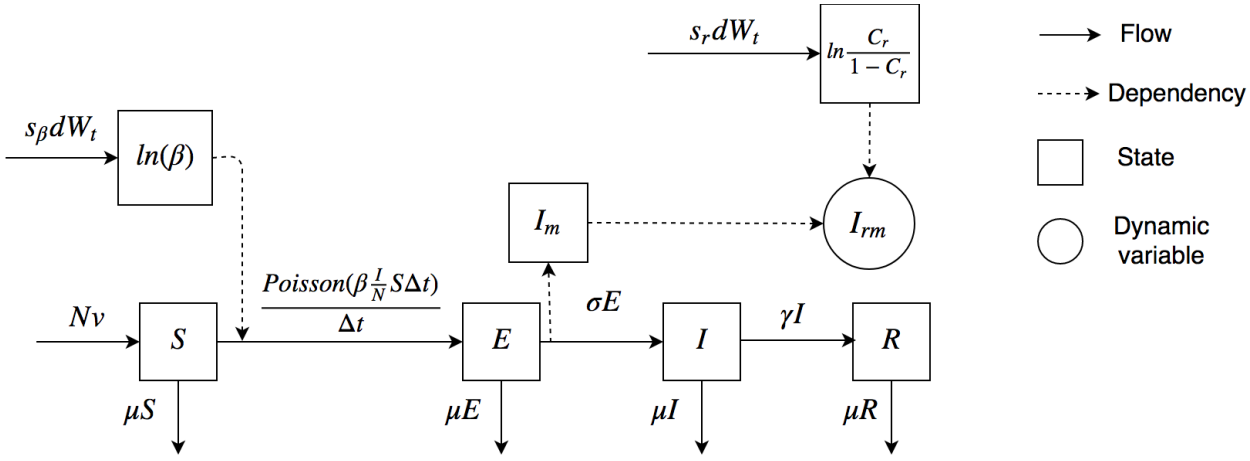


Figure 3.1: The mathematical structure of the particle filtering aggregate model of measles.

Figure 3.1 represents the mathematical structure of the measles particle filtering model with aggregate population.

3.3.3 Parameters and initial values

The important fixed parameters in the models are γ^{-1}, σ^{-1}, v, μ. The values of birth rate is obtained from the Annual Report of the Saskatchewan Department of Public Health [21]. To let the total population (N) keep stable (un-changed) across the whole timeframe of the model, the value of the death rate (μ) is similar to the birth rate. The values of parameters of γ^{-1} and σ^{-1} are as given by [30]. Finally, all the compartmental parameters are specified at table 3.1. The initial value of all stocks in the particle filtering models are given in table 3.2.

Table 3.1: Table showing the value of parameters in the measles aggregate particle filtering model.

Parameter	**Value**	**Units**
γ^{-1}	5	Day
v	0.03	1/Year
μ	0.03	1/Year
N	863,545	Person
σ^{-1}	8	Day

Table 3.2: Table showing initial values of the S, E, I, R stocks in the measles aggregate particle filtering model. The units are all Person.

Parameter	Value
S_0	96354 - I_0
E_0	0
I_0	Uniform[0, 9635)
R_0	767191

3.3.4 The proposal distribution

The Condensation Algorithm [14, 74] is applied in this project to implement the particle filter model. It is the simplest and most widely used proposal distribution, and consists simply of a proposal distribution that is the same as the prior . If we suppose [74, 6]:

$$q(x_k^{N(i)}|x_{k-1}^{N(i)}, y_k^M) = p(x_k^{N(i)}|x_{k-1}^{N(i)}) \tag{3.8}$$

Then the SIS recursively update Equations (2.25) and (2.26) reduce to:

$$x_k^{N(i)} \sim p(x_k^{N(i)}|x_{k-1}^{N(i)}) \tag{3.9}$$

$$w_k^{(i)} \propto w_{k-1}^{(i)} p(y_k^M|x_k^{N(i)}) \tag{3.10}$$

It is notable that sampling from the term $p(x_k^{N(i)}|x_{k-1}^{N(i)})$ in Equation (3.9) represents simply advancing the model from time *k-1* to time k, while the second term $p(y_k^M|x_k^{N(i)})$ in Equation (3.10) is simply the likelihood function. This strategy represents a "generate and test" approach: for a given particle holding the value $x_{k-1}^{N(i)}$in the posterior at the last observation point (at time *k-1*), we sample values by simply advancing that particle in the dynamic model to the time of the next observation, with an induced probability of an outcome $x_k^{N(i)}$ of $p(x_k^{N(i)}|x_{k-1}^{N(i)})$, and then evaluate how consistent the empirical evidence is with the predictions implied by the resulting state of that particle by evaluating the likelihood function $p(y_k^M|x_k^{N(i)})$ from the probabilistic model, multiplying the value of the weight associated with that particle by the value of that likelihood function. This has the effect of lowering the weight associated with particles that predict observations inconsistent with the empirical observation (and thus have lower values of the likelihood function), and elevating the weight for those that imply observations that are consistent with the empirical observation (and thus are associated with higher values of that likelihood function).

3.3.5 Likelihood function

In this project, the observed data is the monthly reported incidence case count of measles. As previously introduced, the measured data is the reported cases of measles in this project. The likelihood function

$p(y_k^M|x_k^N)$ describes such a reporting process, and specifies the probability that a given measles case in the dynamic model will be measured. One way of characterizing the reporting process of measles would be analogous to a coin flip, with each case being subject to an independent distributed probability of reporting. Such a treatment of reporting would imply a likelihood function characterized by the probability mass function of a binomial distribution, with a count of trials equal to the count of underlying incident cases posited by a particle.

However, if we choose a binomial distribution as the basis for a likelihood function in this project, it imposes a high risk of causing a problem of singularity during weight re-normalization [84]. This can be caused by situations where all particles in the model are associated with an infectious state smaller than the empirical data observed. Because the number of trials for each such particles will be less than the value of the empirical incident case count, the probability of a binomial draw yielding the observed data is 0 for each particle and thus the weights of each particle would become zero.

Finally, I followed the past contributions [23, 84, 78, 95] in selecting the negative binomial distribution as the basis for the likelihood function, which allows for greater robustness than the classic binomial distribution.

The equation associated with the likelihood function is thus as follows:

$$p(y_k|x_k) = \binom{y_k + r - 1}{y_k} p^{y_k}(1-p)^r \tag{3.11}$$

where y_k is the empirical data (reported measles cases) at time k, $p = \frac{x_k}{x_k+r}$ representing the probability that a given measured reported case is in fact a true reported incident case, and x_k is the (integer rounded) incident cases resulting from the dynamic model at time k. r is the dispersion parameter associated with the negative binomial distribution. In all scenarios reported in this project, the value of r is chosen to be 10.

In the aggregate model, the measured data is a one-dimensional vector consisting of the monthly reported cases. It indicates that the weight update rule (likelihood function) of the aggregate model could simply achieved by calculating the value of $p(y_{mk}|I_{rmk})$, where (y_{mk} is the empirical data as given by the monthly reported measles cases at time k, and I_{rmk} is the reported cases calculated by the dynamic model.

3.3.6 Evaluating particle filter performance

To assess the accuracy of particle filtering for robust estimation of model states, it is essential to evaluate the estimation and predictive capacity of the particle filtered models. In this project, I therefore sought to calculate the discrepancy at each observation time (Month in this project) [84] between the model generated data and empirical data, using a linear measure. Typically, there will be thousands of particles included in each model run. To calculate the discrepancy of particle filter results by incorporating empirical data across all time points, I sample n particles by importance sampling for each such time. The monthly discrepancy of each time is simply the Root Mean Squared error (RMSE) between the monthly empirical data at that time and the related data calculated by the particle filtering model [84, 107].

Moreover, to investigate the predictive capability and efficiency of particle filter model, I defined a variable

"Prediction Start Time", denoted by T^*; it indicates the time $1 \leq k < T^*$ up to which the weights of particles are updated based on the observed data. When $k \geq T^*$, the particle filtering ceases – the weights of particles are no longer updated and no further re-sampling occurs. During that period (i.e., following the Prediction Start Time), particles simply continue to evolve according to the state space model shown in Equation (3.6). For such a Prediction Start Time T^*, the model calculated a prediction discrepancy using a simple variant of the strategy of the discrepancy used in considering all the time frame, but limited to considering only times T^* and larger.

Finally, in this project, the discrepancy of the particle filtering model is compared to the discrepancy of the normal compartmental model and calibration model. The discrepancies of the normal compartmental model and calibration model are calculated simply by the root mean across the whole timeframe between the result of models and empirical dataset.

3.3.7 Model characterization

To research on the performance of incorporating particle filtering into the compartmental model, I have built 3 models – the normal deterministic model, the calibration model and the aggregate particle filtering model. In the particle filtering model, the number of particles in the particle filtering algorithm is 5000. The description of each model is listed as follows:

(1) *Pure*. It is simply the deterministic SEIR model with aggregate population, see Equation (3.1). The value of the initial infectious population is 90, the initial susceptible population is 89910, the initial exposed population is 0, and the initial recovered population is 773545. The values of β and reporting rate C_r are 50 and 0.11, respectively.

(2) *Calibrated*. It is the calibration model of the SEIR model with aggregate population. Then, the relatively uncertain and stochastic parameters, including initial infectious population, initial susceptible population, the parameter β and reporting rate C_r are obtained by calibration. Finally, they are respectively 930, 89070, 49.598 and 0.119. The initial exposed population is 0, and the initial recovered population is 773545.

(3) *PF*. The particle filtering model with homogeneous mixing of all population, see Equation (3.6).

By comparing the discrepancy of these models, we sought to identify the performance of the particle filtering model. The model with the smallest discrepancy would be the most reliable in simulating the measles outbreak pattern in the province of Saskatchewan during the pre-vaccination era from 1921 to 1956. To assess model results, the stochastic model – particle filtering model is run for 5 realizations, each with random seeds generated from the same set. I then calculate the average and 95% confidence intervals of the mean discrepancy.

3.4 Results

Table 3.3: Comparison of the average discrepancy of all models by incorporating empirical data across all observation points.

Model	Monthly Discrepancy
Pure	249.0
Calibrated	207.5
PF	104.6 (99.4, 109.9)

The five particle filtering model is run for 5 realizations, each with different random seeds. Shown here are the average and 95% confidence intervals (in parentheses) of the mean discrepancy for each model variant.

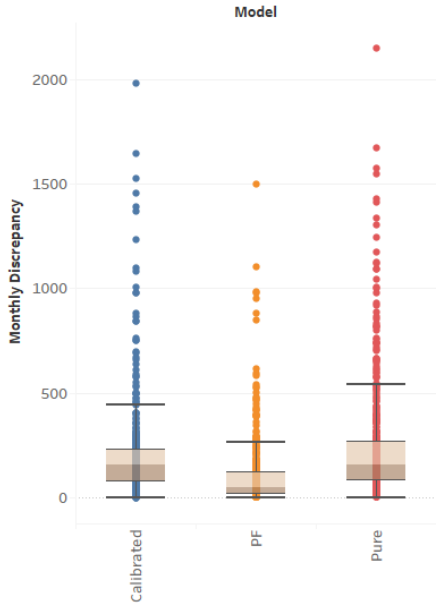


Figure 3.2: Box plots of monthly discrepancy of all models by incorporating empirical data across all observation points.

3.4.1 Results of incorporating the empirical dataset across all timeframe

To depict the particle filter results, 2D histograms of values calculated by the particles in the particle filtering model are plotted. To let the 2D histogram plots characterize the model's calculated data with proper weighting in accordance with the principles of importance sampling, I plot the results of the particles sampled by weights. The resulting plot thus represents an approximation to the probability distribution of the values characterized by the particle filtering model. It is notable that the number of particles performed, the number of particles sampled in 2D histograms and the number of particles sampled (also by weight) in calculating

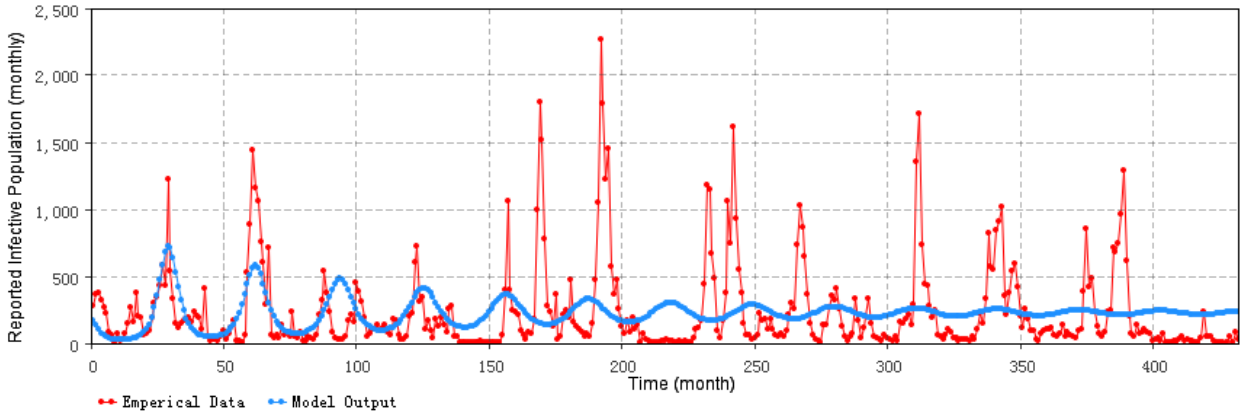


Figure 3.3: Comparison of the reported measles cases of the calibration model and the empirical data (monthly).

the discrepancy in the particle filtering model in this chapter are all 5000.

Table 3.3 specifies the discrepancies of the models implemented in this chapter – the deterministic dynamic model with parameters values obtained from literature and estimated (denoted as *Pure*), the deterministic dynamic model with parameters obtained from a calibration process incorporating matching of the model against empirical data (denoted as *Calibrated*) and the particle filtering model (denoted as PF) of aggregate population. It is notable that both the model of *Pure* and *Calibrated* are deterministic models. Thus, each model can simply calculates a certain fixed discrepancy. By contrast, the particle filtering model PF includes several stochastic processes. Thus, 5 runs of the particle filtering model have been calculated, with the average (across realizations) value and the 95% confidence interval (in parentheses) of the mean discrepancy for each model variant are listed. Table 3.3 demonstrates that the particle filtering model strongly decreases the model discrepancy by approximately a factor of 2.0 compared with the calibrated model. And the particle filtering model exhibits the highest accuracy among these three models, by virtue of the fact that it offers the smallest discrepancy. This reduction in discrepancy proves the hypothesis in the thesis statement that incorporating particle filtering algorithm could help to enhance model accuracy.

Figure 3.2 shows the box plot of the discrepancy among the three models, where each data point represents one month. It is notable that the dataset providing the discrepancy of the box plot in the particle filtering model is calculated by the average value among five runs for each month (the particle filtering model is run for 5 realizations, each with different random seeds and then the average monthly discrepancy among these five runs at each time point is plotted). Figure 3.2 also indicates that the particle filtering model has the smallest median discrepancy. Moreover, the dataset of the discrepancy of the particle filtering model exhibits narrower distribution between monthly discrepancies than the other two deterministic models. Thus, the particle filtering model exhibits the greatest and most consistent accuracy.

Figure 3.3 displays the comparison of the empirical data (on the one hand) and the reported measles cases generated by the deterministic model where the parameters are calibrated with reference to the empirical

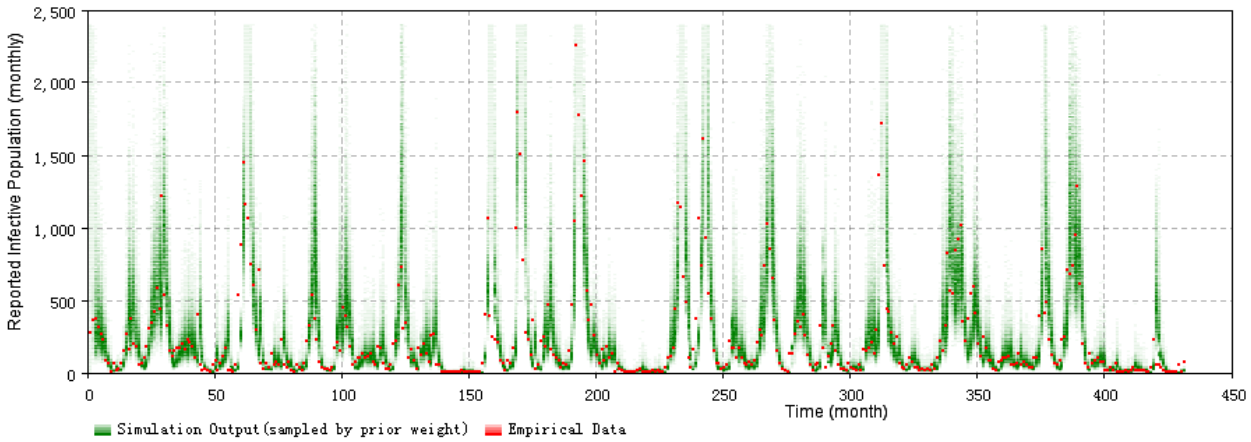


Figure 3.4: 2D histogram prior result of total timeframe of the particle filtering model (monthly).

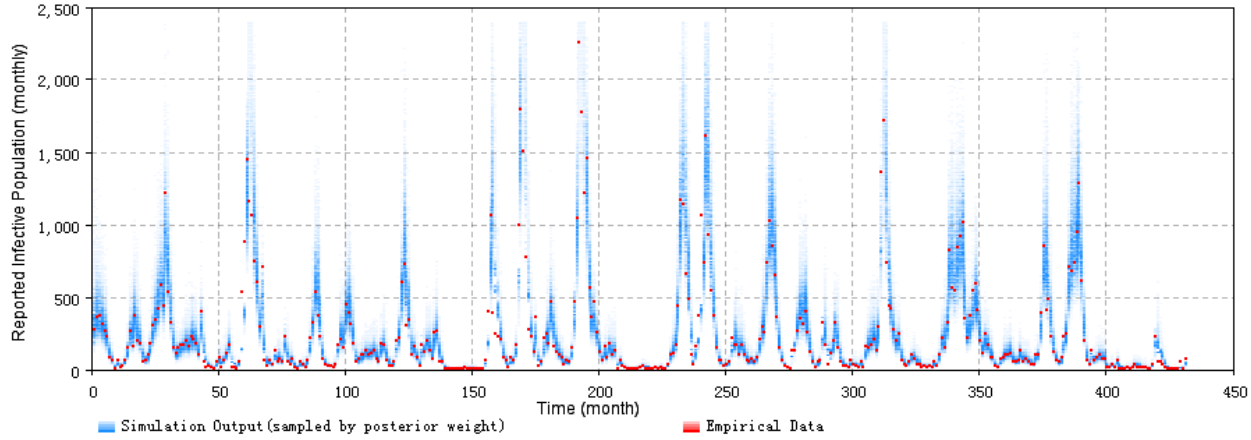


Figure 3.5: 2D histogram posterior result of total timeframe of the particle filtering model (monthly).

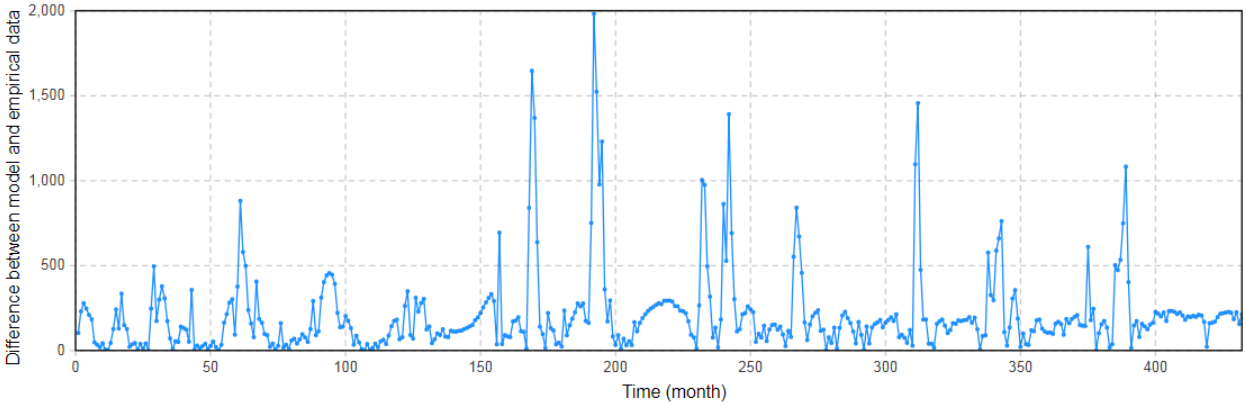


Figure 3.6: Difference between the results of the calibrated model and the empirical data for measles (monthly).

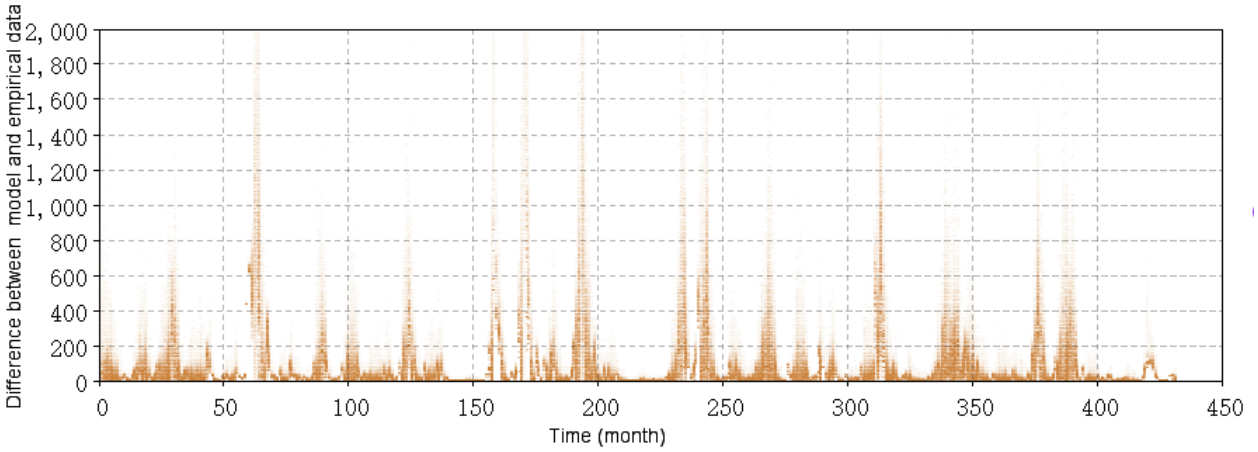


Figure 3.7: 2D histogram plot of the difference between samples from the monthly posterior distributions of the particle filtering model and the empirical data for measles (monthly).

data (on the other). Figure 3.3 indicates that the deterministic model with calibrated parameters exhibits difficulty in tracking the oscillations in the outbreaks of measles in the middle and later times. The reason is that the deterministic model of measles employed in this chapter ($SEIR$ model) approaches a stable equilibrium.

Figure 3.4 displays samples from the monthly prior distributions of the particle filtering model for the entire timeframe. For this case, the results of particle filtering model are sampled according to the weights of all particles before the weights are updated by incorporating the empirical data of the current time point. The values of empirical data points are shown in red, while the distribution of the sampled prior results of particle filtering model is shown in green at each time.

Figure 3.5 presents the samples from the monthly posterior distributions of the particle filtering model for the entire timeframe. For this case, the results of particle filtering model are sampled according to the weights of all particles after the weights are updated by incorporating the empirical data of the current time point. The values of empirical data points are shown in red, while the distribution of the sampled posterior results of particle filtering model is shown in blue at each time.

Figure 3.6 shows the difference for each time point between the empirical data and the results of the deterministic model with calibrated parameters.

Figure 3.7 represents the 2D histogram plot between the empirical data and samples from the monthly posterior distributions of the particle filtering model by incorporating the empirical data across the whole timeframe. It is notable that the particles sampled in figure 3.5 are the same as the particles samples in figure 3.7.

Both figure 3.4 and figure 3.5 indicate that most of the empirical data are located at the range where the particles exhibit high posterior probability. This reflects the fact that the particles could suitably track the oscillation of the outbreak pattern of measles, given the combination of model prediction and observation-

based updating that forms the basis for the particle filter. It bears emphasis that the results of the particle filtering model sampled in figure 3.4 are before the weight update process in each step, while the results in figure 3.5 are after the weight update process in each step. The value of sampled particles of figure 3.4 spread in a wider range, compared with figure 3.5. This difference in dispersion indicates that the weight update process of particle filtering algorithm has the capability to combine the empirical data to the particle filtering model to constrain the particles in a tighter range as suggested by the empirical data.

From the results above, we can see that the particle filtering model not only can decrease the discrepancy between the model results and the empirical data (see table 3.3 and the comparison between figure 3.6 and figure 3.7), but also can track the oscillation of the outbreaks. This contrast is particularly evident when comparing the results of particle filtering model (figure 3.5) and deterministic model with calibrated parameters (figure 3.3). All of these results indicate that incorporating particle filtering in the compartmental model of measles could help to improve the simulation accuracy and track the outbreaks.

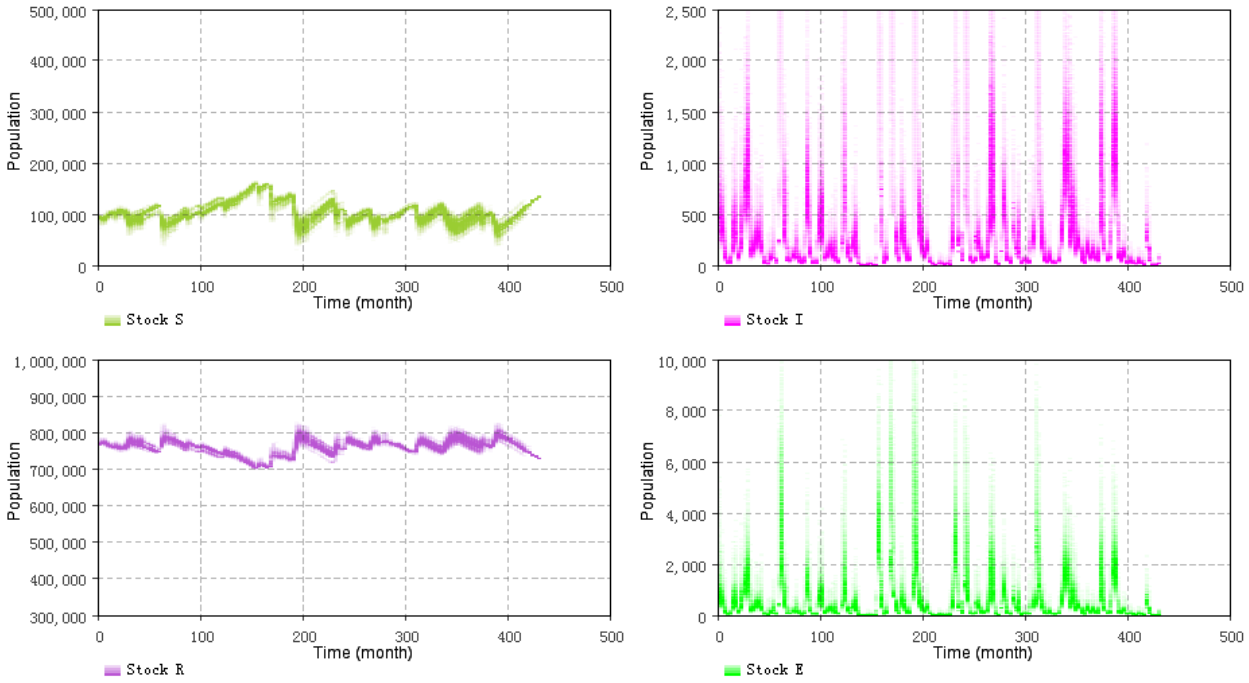


Figure 3.8: 2D histogram results for the latent S, E, I, R stocks estimated by the particle filtering model.

Particle filtering models can contribute to the estimation of model states and aid in estimating dynamic model parameters. It is notable that, as is widely the case in dynamic models, the states in the compartmental models are latent (e.g., Susceptible (S), Exposed (E), Infectious (I) and Recovered (R) stocks in the SEIR model in Equation (3.1) of measles). What can be empirically observed is the noisy reported measles cases related to the Infectious (I) state. However, the methodology of particle filter provides an approach to estimate (via sampling from) the distribution of values of these latent states. This ability to estimate the value of latent states such as the reservoir of susceptible people can aid researchers and public health agencies to in terms

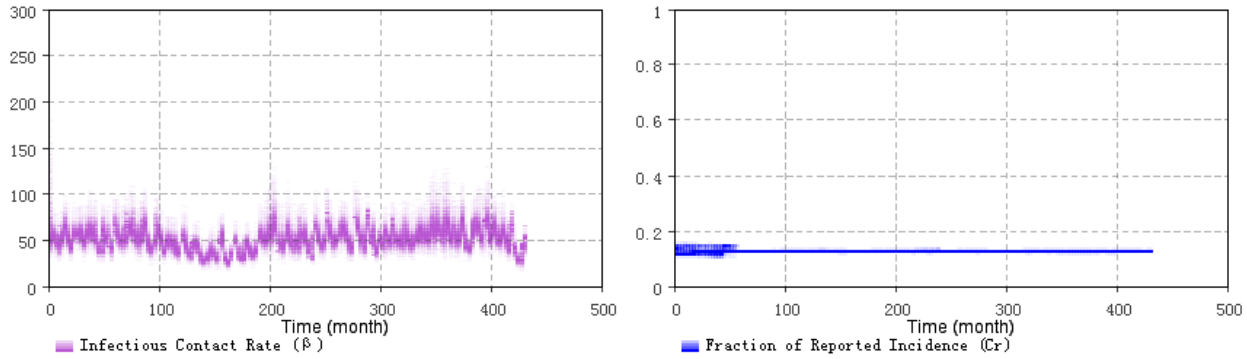


Figure 3.9: 2D histogram results for the dynamic parameters estimated by the particle filtering model.

of understanding the underlying epidemiological situation from multiple lines of evidence, as constrained by understanding of the structure of the system, as characterized by a dynamic model. To illustrate this, I employ a similar method to the above to plot the 2D histogram of the latent states of Susceptible, Infectious, Exposed and Recovered sampled according to importance sampling principles. figure 3.8 shows the results of the 2D histogram plots of the latent states. These plots indicate that most of the population are located in the Recovered (R) stock. And there are clear relationship under the four stocks. After an outbreak and before the next outbreak of measles, the population in the stock of Susceptible (S) is piled up, with the newborns come in, who are all susceptible individuals. Then during the next outbreak, the population in stock of the Susceptible (S) is consumed by the infectious, and eventually flows to the stock of Recovered (R). This lies in accordance with the expectations for measles transmission in the real world and builds confidence in the capacity of the model to meaningfully estimate latent state.

Figure 3.9 displays the estimation of the stochastic parameters of the particle filtering model. It is notable that although the stochastic parameters – the infectious contact rate of measles in Saskatchewan during the pre-vaccination era (denoted as β) and the report rate of measles (denoted as C_r) are simply static parameters in the traditional deterministic $SEIR$ model of measles, they are treated as latent states in the particle filtering models. The particle filtering models are capable of estimate all the latent states of the dynamic systems. In the project, the particle filtering model think that the infectious contact rate of measles in Saskatchewan during the pre-vaccination era is around 50, and the report rate of measles is around 0.12 to 0.13.

3.4.2 Prediction results of the minimal discrepancy model

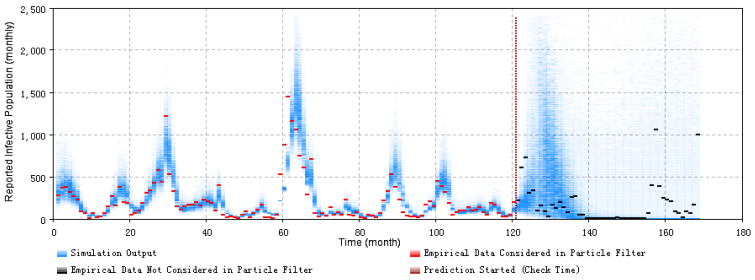
In this section, I assess the predictive capacity of the particle filtering model introduced in the previous section. By changing different Prediction Start Time of T^*, we have performed the prediction from different archetypal situations. These situations are listed as follows:

(1) Prediction started from the first or second time points of an outbreak.

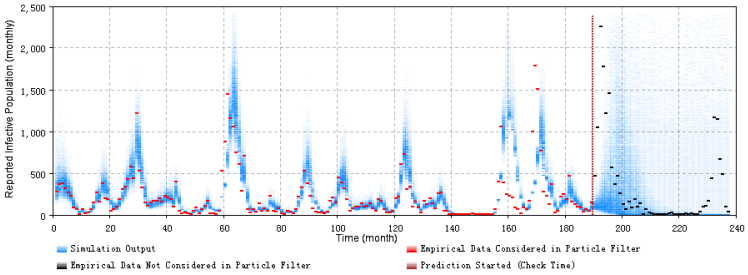
(2) Prediction started before the next outbreak.

(3) Prediction started from the peak of an outbreak.

(4) Prediction started from the end of an outbreak.



(a)

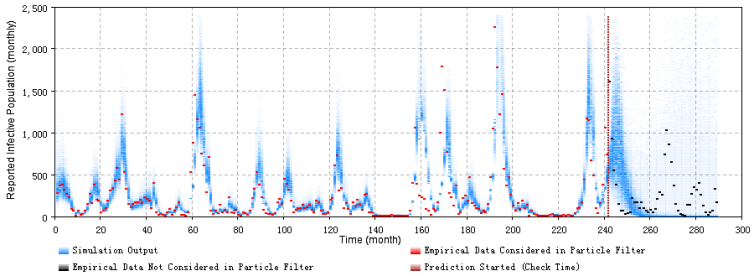


(b)

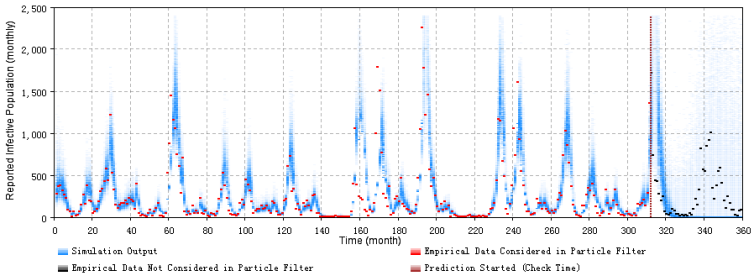
Figure 3.10: 2D histogram of predicting from the first or second time points of an outbreak of the minimum discrepancy model. (a) predicted from the 121 months. (b) predicted from the 190 months.

Figure 3.10 displays the prediction results of the particle filtering model in situations in which the prediction started from the first or second time points of an outbreak. This is illustrated with two experiments from where prediction started from month 121 ($T^* = 121$) – with β is 0.4, and the monthly prediction discrepancy is 337.6 – and 190 ($T^* = 190$) – with β is 0.4, and the monthly prediction discrepancy is 367.7 – respectively. In the prediction process of the particle filtering model, the weights of particles will stop being updated at the time given by "Prediction Started Time" (T^*) by incorporating the empirical data. From that point forward, all the particles run without new empirical data being considered. In this chapter, all the prediction experiments are run 4 years from the time of "Prediction Start Time".In the 2D histogram plot of figure 3.10, the empirical data considered in the particle filtering process (i.e., incorporated in training the models) are shown in red, while the empirical data that was not considered in the particle filtering process (i.e., data points only displayed to compared with the results of models) are shown in black. The vertical straight line labels the "Prediction Start Time" of T^* of each experiment.

Figure 3.11 displays the prediction results of the particle filtering model in situations in which the prediction started from a peak of an outbreak. As above, this is illustrated with two experiments, one in which

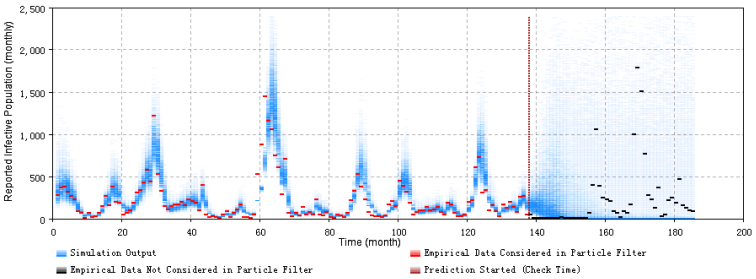


(a)

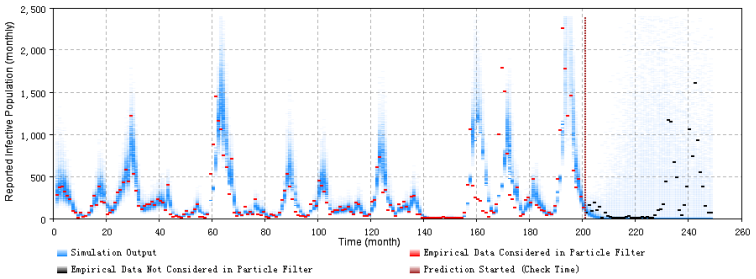


(b)

Figure 3.11: 2D histogram of predicting from the peak of an outbreak of the minimum discrepancy model. (a) predicted from the 242 months. (b) predicted from the 312 months.

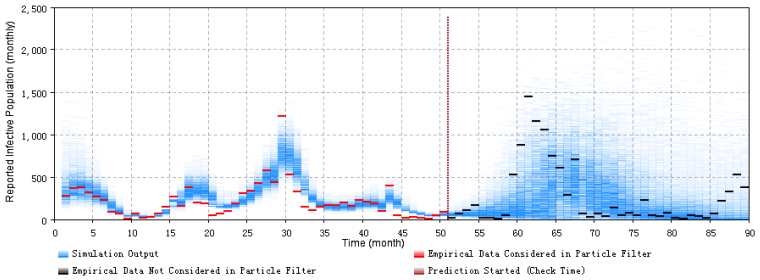


(a)

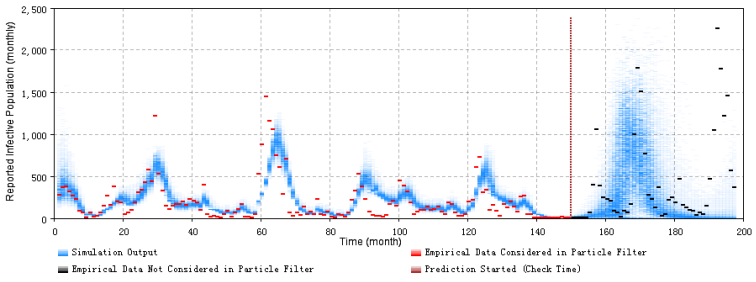


(b)

Figure 3.12: 2D histogram of predicting from the end of an outbreak of the minimum discrepancy model. (a) predicted from the 138 months. (b) predicted from the 201 months.



(a)



(b)

Figure 3.13: 2D histogram of predicting before the next outbreak of the minimum discrepancy model. (a) predicted from the 51 months. (b) predicted from the 150 months.

prediction started from month 242 ($T^* = 242$) – with β is 0.4, and the monthly prediction discrepancy is 352.5 – and another from month 312 ($T^* = 312$) – with β is 0.4, and the monthly prediction discrepancy is 350.4 – respectively. The layout of the 2D histogram plot of figure 3.11 is the same as figure 3.10.

Figure 3.12 displays the predictive results of the particle filtering model in situations in which the prediction initiated at the end of an outbreak with two experiments, prediction started from month 138 ($T^* = 138$) – with β is 0.4, and the monthly prediction discrepancy is 347.7 – and month 201 ($T^* = 201$) – with β is 0.4, and the monthly prediction discrepancy is 300.8 – respectively. The layout of the 2D histogram plot of figure 3.12 is the same as figure 3.10.

Figure 3.13 displays the predictive results of the particle filtering model in situations predicting before the next outbreak with two experiments, prediction started from month 51 ($T^* = 51$) – with β is 0.2, and the monthly prediction discrepancy is 356.2 – and month 150 ($T^* = 150$) – with β is 0.1, and the monthly prediction discrepancy is 362.6 – respectively. The layout of the 2D histogram plot of figure 3.13 is the same as for figure 3.10.

These prediction results suggest that the particle filter model offers the capacity to probabilistically anticipate measles dynamics with a fair degree of accuracy. From the 2D histogram plots, empirical data lying after Prediction Start Time – and thus not considered by the particle filtering machinery – mostly lie within the high-density range of the particles. Notably, in such examples, the particle filter model appears to be able to accurately anticipate a high likelihood of a coming outbreak and non-outbreak. Such an ability could offers substantial value for informing the public health agencies with accurate predictions of the

anticipated evolution of measles over coming months.

3.5 Discussion and conclusion

This chapter presents a new method for tracking the patterns of measles outbreaks in low vaccination regions by applying particle filtering with measles transmission models, and incorporating time series of measles incidence.

Particle filtering has mitigated significant weaknesses and simplifications associated with the aggregate System Dynamics transmission model and noisy empirical data. By incorporating ongoing arriving empirical data, the particle filter model has demonstrated the capacity to correct for the distortions that accompany the System Dynamics model aggregation, such as assumptions of random mixing and homogeneity. In this dataset, particle filter offered strong performance in estimating the population state underlying the outbreak pattern of measles and predicting future trends.

It is to be emphasized that particle filtering samples estimates of the entire state of the dynamic model (both latent or observable) for each point in time incorporated within the empirical dataset. In the particle filtering model of measles with aggregate population in this chapter, the latent states related to the Susceptible population, the Infectious population, the Exposed population, the Recovered population, the infectious contact rate and the report rate of measles in each time point. The final of these supports calculation of a derivative value that can be compared with the empirical data. This capacity to sample from the full state of the system – both latent and observable – can help researchers to perform related research, including better understanding historical circumstances.

It is notable that the impact of the stochastic processes added in the state space model has distinct consequences during the estimation period and prediction period. In the estimation period where the empirical datasets are incorporated in the particle filtering algorithm, if these stochastic processes are more noisy (e.g., the parameters of s_β and s_r are larger), the particle filtering models generally benefit from an easier ability to trace the empirical datasets if equipped with a sufficiently large particle count. This reflects the fact that, given sufficiently large stochastics, some particles are likely to be consistent with empirical datapoints close to that observed; such particles can then be favoured by the resulting weight updates, and allow for effective adaptation to the new empirical datapoint. However, during the prediction period, it is particularly desirable that the stochastic processes not be too noisy. This reflects the fact that, absent new empirical data, in the presence of a noisy state space model, the particle filtering model will grow rapidly in its uncertainty regarding the current state of the system. Thus, the noise of the stochastic processes in the state space model should be controlled in a proper range – if the model is more confident (exhibits continued tight support in the distribution of estimated current state as the time since the last data point increase), it is more helpful during the prediction period, while it will make the estimation period more difficult, as the model will tend to be less able to readily respond to new datapoints outside of the range of particle predictions. However,

if the model is more noisy, it is more helpful during the estimation period by allowing for a greater overlap between the particle distribution and the empirical data and henceforth a faster adaptation to incoming data; by contrast, the presence of pronounced stochastics in the prediction period will lead the model to rapidly exhibit broad uncertainty regarding system state as time increases since the last datapoint.

Another merit of the particle filtering model is that the particle filtering model could manage the estimation of the initial values of each stocks easily. In traditional deterministic models, the initial values is difficult to estimate. Normally, the models may need to be run in a long period of burn-in time to estimate the initial values. However, in the particle filtering models, the initial values of the stocks could be assigned in a proper range, such as following a uniform distribution in the particle filtering model introduced in this chapter. Then, the initial values of each stock can be estimated by the particle filtering algorithm, by allowing for a "survival of the fittest" that rewards particles positing an initial state of the model that is consistent with the evidence. It is notable that the estimation of the initial state of the measles aggregate particle filtering models in this chapter is calculated by the particle filtering algorithm automatically. Thus, no human time is needed in estimating the initial values of the stocks in the particle filtering models with aggregate population structure.

This chapter proves the hypotheses proposed in the thesis statements for measles. Firstly, comparing the empirical data and the 2D histogram plots with the particle filtering models' output data (the monthly reported cases of measles) indicates that the high probability density region of the model's posterior prediction of incident cases lies near the empirical data, especially when compared with the results of the output of the traditional calibration models. Secondly, the discrepancy between the measles particle filtering model's (posterior) predictions vs. observed data is reduced by a factor of 2.0. This indicates that the particle filtering algorithm is capable of supporting greater measles model accuracy when compared with the traditional calibration process. It is notable that with the deterministic model with calibrated parameters is difficult to track the fluctuations of the outbreak pattern of measles, while the particle filtering model is capable of tracking oscillations in the outbreak of measles.

Chapter 4

Measles age-stratified model of particle filtering

4.1 Introduction

Measles has a severe impact on children's health and is one of the leading causes of death among young children globally. Measles transmission pattern may be different among different age groups [97]. For example, the age composition of daily contacts may be different for different age groups; children may spend more time with other children and their caregivers, rather than with other adults. Moreover, the rates of contacts sufficiently close to transmitting infection can differ between age groups, such as due to hygienic disparities.

To capture such differences, beyond the aggregate model introduced in chapter 3, where the susceptible individuals contact with the infectious individuals homogeneously, this chapter investigates the predictive performance of particle filtering of age-structured measles models. Specifically, within such models, the total population are split into two age groups – children and adults. The susceptible persons heterogeneously contact with the infectious persons between different age groups, while the susceptible individuals contact with the infectious individuals homogeneously in the same age group.

Moreover, I have investigated the performance of the particle filtering models by incorporating multiple empirical datasets. There are two categories of measles empirical datasets – monthly measles reported cases across all population in Saskatchewan from 1921 to 1956, and yearly measles reported cases of different age groups from 1925 to 1956. Such datasets are introduced in detail in chapter 2. Two categories of measles particle filtering age-structured models are built in this chapter – only incorporating a single empirical dataset (the measles monthly reported cases) and (by contrast) incorporating 3 empirical datasets (the measles monthly reported cases, and the yearly measles reported cases of the two age groups). Then, by comparing the discrepancy of the aggregate particle filtering model and age-structured particle filtering models, and between the age-structured particle filtering model incorporated with only one empirical dataset and (separately) 3 empirical datasets, I have researched on the performance among these different measles particle filtering models in this chapter.

Finally, this chapter explores the capability of particle filtering model simulating intervention experiments simply by modifying parameters in the particle filtering models.

4.2 The age-structured mathematical model

In this variant of the mathematical model, I use subscripts "c" and "a" for a quantity to denote the child- and adult-specific values of that quantity, respectively. I further assume in the demographic model, that the population of each age group (N_c, N_a) does not change (whose formulation and derivation are introduced in the chapter 2). Similar to the parameter of total population (N) in the aggregate model, the mean of the population of each age group across the timeframe in the age pyramid of Saskatchewan [16] is employed as the value of N_c and N_a (where the sum of N_c and N_a equals N). Before introduce the measles age-structured mathematical models, I will introduced the general deduction of the age structured epidemiology model first.

4.2.1 The mathematical deduction of the age structured epidemiology model

Similar to the equilibrium demographic model introduced in chapter 2 [50], the initial-boundary-value problem of the SEIR model with age structure could be listed as follows:

$$\begin{aligned}
&\frac{\partial S}{\partial t} + \frac{\partial S}{\partial a} = -\lambda(a,t)S - \mu(a)S \\
&\frac{\partial E}{\partial t} + \frac{\partial E}{\partial a} = \lambda(a,t)S - \sigma E - \mu(a)E \\
&\frac{\partial I}{\partial t} + \frac{\partial I}{\partial a} = \sigma E - \gamma I - \mu(a)I \\
&\frac{\partial R}{\partial t} + \frac{\partial R}{\partial a} = \gamma I - \mu(a)R \\
&\lambda(a,t) = \int_0^\infty \beta(a,a')\frac{I(a',t)}{N(a',t)}da'
\end{aligned} \tag{4.1}$$

where $\beta(a,a')$ is the infectious contact rate between age group a and age group a'.

The boundary condition at age 0 are all 0, except all the birth population going to $S(0,t)$:

$$S(0,t) = \int_0^\infty v(a)N(a,t)da \tag{4.2}$$

The initial conditions are the total population of each age group at time 0.

Similarly to the demographic model, I split the total population to n age groups. Then, for each age group i in the age interval $[a_{i-1}, a_i)$, I could get four ordinary differential equations of S_i, E_i, I_i, R_i. The definition of S_i, E_i, I_i, R_i are listed as follows:

$$\begin{aligned}
S_i(t) &= \int_{a_{i-1}}^{a_i} S(a,t)da \\
E_i(t) &= \int_{a_{i-1}}^{a_i} E(a,t)da \\
I_i(t) &= \int_{a_{i-1}}^{a_i} I(a,t)da \\
R_i(t) &= \int_{a_{i-1}}^{a_i} R(a,t)da
\end{aligned} \tag{4.3}$$

And for each age group i, I have:

$$N_i(t) = S_i(t) + E_i(t) + I_i(t) + R_i(t) \tag{4.4}$$

Thus, I could get totally $4n$ ordinary differential equations. The infectious contact rate $\beta(a, a')$ between any two age groups (a and a') are also assumed to be constant. Then, I could have $\beta(a, a') = \beta_{ij}$, where age group i locates in the age interval $[a_{i-1}, a_i)$, and age group j locates in the age interval $[a_{j-1}, a_j)$. Similarly to the demographic model in [50], I have $S(a_i, t) = c_i S_i(t)$, $E(a_i, t) = c_i E_i(t)$, $I(a_i, t) = c_i I_i(t)$, $R(a_i, t) = c_i R_i(t)$. Finally, if I integrate the partial differential equations in the age interval $[a_{i-1}, a_i)$, I could get the final epidemiology model with $4n$ ordinary differential equations:

$$\begin{aligned}
\frac{dS_1}{dt} &= \sum_{j=1}^{n} v_j N_j(t) - \lambda_1 S_1 - \mu_1 S_1 - c_1 S_1 \\
\frac{dE_1}{dt} &= \lambda_1 S_1 - \sigma_1 E_1 - \mu_1 E_1 - c_1 E_1 \\
\frac{dI_1}{dt} &= \sigma_1 E_1 - \gamma_1 I_1 - \mu_1 I_1 - c_1 I_1 \\
\frac{dR_1}{dt} &= \gamma_1 I_1 - \mu_1 R_1 - c_1 R_1 \\
\frac{dS_i}{dt} &= c_{i-1} S_{i-1} - \lambda_i S_i - \mu_i S_i - c_i S_i \quad i \geq 2 \\
\frac{dE_i}{dt} &= \lambda_i S_i + c_{i-1} E_{i-1} - \sigma_i E_i - \mu_i E_i - c_i E_i \quad i \geq 2 \\
\frac{dI_i}{dt} &= \sigma_i E_i + c_{i-1} I_{i-1} - \gamma_i I_i - \mu_i I_i - c_i I_i \quad i \geq 2 \\
\frac{dR_i}{dt} &= \gamma_i I_i + c_{i-1} R_{i-1} - \mu_i R_i - c_i R_i \quad i \geq 2 \\
\lambda_i &= \sum_{j=1}^{n} \beta_{ij} \frac{I_j}{N_j}
\end{aligned} \tag{4.5}$$

Specifically, the contact matrix in this age structured epidemiology model is: $\begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{bmatrix}$.

4.2.2 The re-dimensionalized age structured model

By applying the method of mathematical deduction of the age structured epidemiology model to the re-dimensionalized mathematical structure of measles, the resulting age-structured SEIR model is as follows:

$$\begin{aligned}
\begin{bmatrix} \frac{dS_c}{dt} \\ \frac{dS_a}{dt} \end{bmatrix} &= \begin{bmatrix} N_a v_a \\ 0 \end{bmatrix} + \begin{bmatrix} -\omega S_c \\ \omega S_c \end{bmatrix} - \begin{bmatrix} \beta_c S_c \\ \beta_a S_a \end{bmatrix} \circ \left(\begin{bmatrix} f_{cc} & f_{ca} \\ f_{ac} & f_{aa} \end{bmatrix} \times \begin{bmatrix} \frac{I_c}{N_c} \\ \frac{I_a}{N_a} \end{bmatrix} \right) - \begin{bmatrix} \mu_c S_c \\ \mu_a S_a \end{bmatrix} \\
\begin{bmatrix} \frac{dE_c}{dt} \\ \frac{dE_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega E_c \\ \omega E_c \end{bmatrix} + \begin{bmatrix} \beta_c S_c \\ \beta_a S_a \end{bmatrix} \circ \left(\begin{bmatrix} f_{cc} & f_{ca} \\ f_{ac} & f_{aa} \end{bmatrix} \times \begin{bmatrix} \frac{I_c}{N_c} \\ \frac{I_a}{N_a} \end{bmatrix} \right) - \sigma \begin{bmatrix} E_c \\ E_a \end{bmatrix} - \begin{bmatrix} \mu_c E_c \\ \mu_a E_a \end{bmatrix} \\
\begin{bmatrix} \frac{dI_c}{dt} \\ \frac{dI_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega I_c \\ \omega I_c \end{bmatrix} + \sigma \begin{bmatrix} E_c \\ E_a \end{bmatrix} - \gamma \begin{bmatrix} I_c \\ I_a \end{bmatrix} - \begin{bmatrix} \mu_c I_c \\ \mu_a I_a \end{bmatrix} \\
\begin{bmatrix} \frac{dR_c}{dt} \\ \frac{dR_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega R_c \\ \omega R_c \end{bmatrix} + \gamma \begin{bmatrix} I_c \\ I_a \end{bmatrix} - \begin{bmatrix} \mu_c R_c \\ \mu_a R_a \end{bmatrix}
\end{aligned} \tag{4.6}$$

where $\circ$ indicates the Hadamard (element-wise) product; $\times$ indicates matrix multiplication; $\begin{bmatrix} \beta_c f_{cc} & \beta_c f_{ca} \\ \beta_a f_{ac} & \beta_a f_{aa} \end{bmatrix}$ is the contact matrix: f_{cc} indicates the fraction of children's infectious contacts that occur with other children; similarly f_{ca} indicates the fraction of children's infectious contacts that occur with adults, and f_{ca}=1-f_{cc}; f_{ac} indicates the fraction of adult's infectious contacts that occur with children; $f_{aa} = 1 - f_{ac}$ indicates the fraction of adult's infectious contacts that occur with other adults; ω is the aging rate out of the age group of children (which carries the same meaning with c_1 in the demographic model in [50]). v_a is the birth rate for adults, for children, the birth rate is 0. The other parameters hold the same role and values as in the age-aggregated model.

4.2.3 The contact matrix model

Among the age groups, the number of contacts from the child age group to the adult age group in a given interval of time must equal the number of contacts from adults age group to the child age group in that same interval of time. This leads to the following equalities:

$$\begin{aligned}
N_a C_a f_{ac} &= N_c C_c f_{ca} \\
f_{cc} + f_{ca} &= 1 \\
f_{ac} + f_{aa} &= 1
\end{aligned} \tag{4.7}$$

where C_a is the contact rate of the adults age group, while C_c is the contact rate of the child age group. Suppose the transmission probability is β_p, then I could have $\beta_c = C_c \beta_p$ and $\beta_a = C_a \beta_p$. It indicates that $\frac{C_c}{C_a} = \frac{\beta_c}{\beta_a}$.

Finally, by solving Equation (4.7), I arrive at Equation (4.8) below:

$$\begin{aligned} f_{ca} &= 1 - f_{cc} \\ f_{ac} &= \begin{cases} \frac{N_c\beta_c}{N_a\beta_a}\left(1 - f_{cc}\right), & \text{if } \left[\frac{N_c\beta_c}{N_a\beta_a}\left(1 - f_{cc}\right)\right] < 1.0 \\ 1.0, & \text{if } \left[\frac{N_c\beta_c}{N_a\beta_a}\left(1 - f_{cc}\right)\right] \geq 1.0 \end{cases} \\ f_{aa} &= 1 - f_{ac} \end{aligned} \tag{4.8}$$

In the contact matrix $\begin{bmatrix} \beta_c f_{cc} & \beta_c f_{ca} \\ \beta_a f_{ac} & \beta_a f_{aa} \end{bmatrix}$, as covered more fully in section below, the parameters β_c, β_a and f_{cc} are treated as varying across the model time horizon (like the parameter of β in the measles aggregate particle filtering model – Equation (3.1) in chapter 3). Based upon their values, the other parameters in contact matrix (f_{ca}, f_{ac}, f_{aa}) can be calculated as in Equation (4.8).

4.2.4 The equilibrium demographic model

The population model is listed as follows, introduced in chapter 2 [50]:

$$\begin{aligned} \frac{dN_c}{dt} &= N_a v_a - \mu_c N_c - \omega N_c \\ \frac{dN_a}{dt} &= \omega N_c - \mu_a N_a \end{aligned} \tag{4.9}$$

Where N_c is the population of the child age group; N_a is the population of the adult age group; v_a is the birth rate (applying only to adults); μ_c is the death rate of child age group; μ_a is the death rate of the adult age group.

While measles infection can be lethal, for simplicity, the death rates of all states in the models of this paper are the same. The death rate of the "infectious (I)" state should, in theory, be higher than the other states in the pre-vaccination era. According to measles history in CDC (Centers for Disease Control and Prevention) [34], there were 400 to 500 deaths reported among 3-4 million measles annually before 1963 (the vaccination starting year) in the United States. Thus, the measles causes death rate of the infectious state is about 0.125% to 0.167% yearly among the total population. Moreover, measles infection was nearly universal during childhood [34]. Thus, the measles attributable death rate of children should be higher than for adults. At the same time, the death rate among the total population of Saskatchewan during 1914 to 1921 is around 0.5% to 1.4%, for example, the death rate in 1921 of Saskatchewan is 0.81% [21]. However, the death rates of all states in the models of this paper are the same, to make the models simpler to be implemented.

As a result of the assumption of an invariant population size, it follows that death rates for children and adults are as follows: the detailed mathematical derivation is introduced in chapter 2 in Equation (2.44) and Equation (2.45).

$$\begin{aligned} \mu_c &= \frac{N_a}{N_c} v_a - \omega \\ \mu_a &= \frac{N_c}{N_a} \omega \end{aligned} \tag{4.10}$$

4.3 Particle filter implementation

4.3.1 The state space model

In the age stratified model, four stochastic parameters and two extra states are considered dynamically, compared with the mathematical model (Equations (4.6)–(4.10)). The stochastic process related to incidence of infection is also considered in the age structured state space model, as in the aggregated model in chapter 3 (Equation (3.6)). The first stochastic parameter is the same as in the aggregate group model: the disease reporting process parameter (C_r), whose dynamics are characterized according to Equation (3.3). The second is the rate of transmissible contacts between infectious persons and susceptible persons of child age group, which is represented by the parameter (β_c) in the age structured model – Equation (4.6). The equation and chosen values of (β_c) are the same with Equation (3.2). The third stochastic parameter (M_a) represents the ratio of the adult age group's transmissible contact rate (β_a) to that of the child age group (β_c). And I have $\beta_a = M_a\beta_c$. Reflecting the fact that this parameter represents a non-negative real number, similar to the rate of transmissible contacts (β) in the aggregate state space model, I treat the natural logarithm of M_a as undergoing a random walk according to a Brownian motion:

$$d(lnM_a) = s_{M_a} dW_t \tag{4.11}$$

It is a widespread perception that because of limited hygienic awareness and other factors, children are subject to higher rates of transmissible contact than adults. This would suggest that the value of the multiplier M_a normally less than 1.0. Thus, I elected to impose an initial value of M_a across all particles as drawn from a uniform distribution with support [0.2, 1). And, the diffusion coefficient (s_{M_a}) associated with the evolution of $d(lnM_a)$ is chosen to be a constant value of 0.5 among all particles.

The fourth stochastic parameter in the stratified model is the fraction of the contact of children that occurs with other children, denoted as f_{cc}. This parameter appears in the contact matrix, and varies over the range from 0 to 1. As a result, the dynamic process for f_{cc} is similar to the disease report rating C_r with the Equation (3.3), specifically:

$$d(ln(\frac{f_{cc}}{1-f_{cc}})) = s_{cc} dW_t \tag{4.12}$$

The initial value of f_{cc} I employed follows a uniform distribution in the interval [0.2, 1.0). I assumed a constant value of 0.2 as the diffusion coefficient (s_{cc}) associated with the logit of f_{cc}.

Finally, the new state of cumulative infectious count per unit time is implemented similar to aggregate model (Equation (3.4)), except for its division into two distinct states according to stratification into two age groups (I_{mc} and I_{ma}). The discrete time equations of I_{mc} and I_{ma} and reported infectious count per unit time in model I_{rmc} and I_{rma} at time k are as follows:

$$
\begin{aligned}
I_{mck} &= \int_{k-1}^{k} (\sigma E_c) dt \\
I_{mak} &= \int_{k-1}^{k} (\sigma E_a) dt \\
\begin{bmatrix} I_{rmck} \\ I_{rmak} \end{bmatrix} &= C_r \begin{bmatrix} I_{mck} \\ I_{mak} \end{bmatrix}
\end{aligned}
\tag{4.13}
$$

The state vector x^N is $[S_c, S_a, E_c, E_a, I_c, I_a, R_c, R_a, \beta_c, M_a,\ f_{cc},\ C_r, I_{mc}, I_{ma}]^T$ in the age stratified model, and N equals 14. The complete set of state equation for the age-stratified model is given in Equation (4.14):

$$
\begin{aligned}
\begin{bmatrix} \frac{dS_c}{dt} \\ \frac{dS_a}{dt} \end{bmatrix} &= \begin{bmatrix} N_a v_a \\ 0 \end{bmatrix} + \begin{bmatrix} -\omega S_c \\ \omega S_c \end{bmatrix} - \begin{bmatrix} \mu_c S_c \\ \mu_a S_a \end{bmatrix} - A_2 \\
\begin{bmatrix} \frac{dE_c}{dt} \\ \frac{dE_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega E_c \\ \omega E_c \end{bmatrix} - \sigma \begin{bmatrix} E_c \\ E_a \end{bmatrix} - \begin{bmatrix} \mu_c E_c \\ \mu_a E_a \end{bmatrix} + A_2 \\
\begin{bmatrix} \frac{dI_c}{dt} \\ \frac{dI_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega I_c \\ \omega I_c \end{bmatrix} + \sigma \begin{bmatrix} E_c \\ E_a \end{bmatrix} - \gamma \begin{bmatrix} I_c \\ I_a \end{bmatrix} - \begin{bmatrix} \mu_c I_c \\ \mu_a I_a \end{bmatrix} \\
\begin{bmatrix} \frac{dR_c}{dt} \\ \frac{dR_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega R_c \\ \omega R_c \end{bmatrix} + \gamma \begin{bmatrix} I_c \\ I_a \end{bmatrix} - \begin{bmatrix} \mu_c R_c \\ \mu_a R_a \end{bmatrix} \\
d(ln\beta_c) &= s_{\beta_c} dW_t \\
d(ln(\frac{f_{cc}}{1-f_{cc}})) &= s_{cc} dW_t \\
d(lnM_a) &= s_{M_a} dW_t \\
\beta_a &= M_a \beta_c \\
d(ln(\frac{C_r}{1-C_r})) &= s_r dW_t \\
f_{ca} &= 1 - f_{cc} \\
f_{ac} &= \begin{cases} \frac{N_c \beta_c}{N_a \beta_a} (1 - f_{cc}), & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a} (1 - f_{cc})\right] < 1.0 \\ 1.0, & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a} (1 - f_{cc})\right] \geq 1.0 \end{cases} \\
f_{aa} &= 1 - f_{ac} \\
\mu_c &= \frac{N_a}{N_c} v_a - \omega \\
\mu_a &= \frac{N_c}{N_a} \omega \\
I_{mck} &= \int_{k-1}^{k} (\sigma E_c) dt \\
I_{mak} &= \int_{k-1}^{k} (\sigma E_a) dt \\
\begin{bmatrix} I_{rmck} \\ I_{rmak} \end{bmatrix} &= C_r \begin{bmatrix} I_{mck} \\ I_{mak} \end{bmatrix}
\end{aligned}
\tag{4.14}
$$

$$A_2 = \frac{Poisson\left(\begin{bmatrix}\beta_c S_c \Delta t \\ \beta_a S_a \Delta t\end{bmatrix} \circ \left(\begin{bmatrix}f_{cc} & f_{ca} \\ f_{ac} & f_{aa}\end{bmatrix} \times \begin{bmatrix}\frac{I_c}{N_c} \\ \frac{I_a}{N_a}\end{bmatrix}\right)\right)}{\Delta t}$$

Reflective of the structure of the age group stratification in available data, it is notable that I have considered two different age group configurations in this paper: one where the child age group includes those up to 5 years old (M_{age_5}) and another where it includes those up to 15 years old (M_{age_15}).

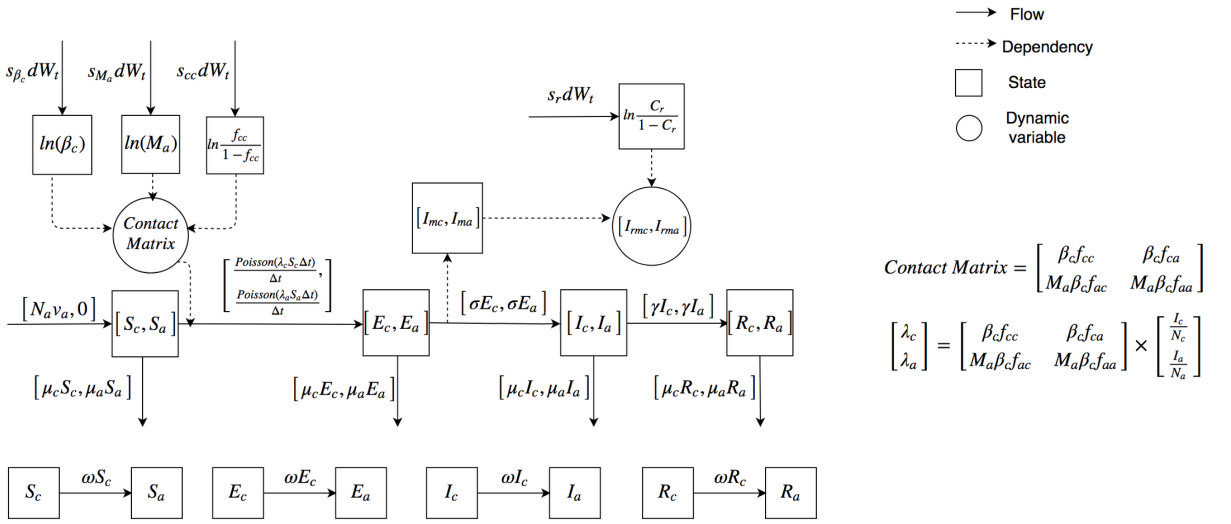


Figure 4.1: The mathematical structure of the particle filtering age stratified model of measles.

4.3.2 Parameters and initial values

The important fixed parameters in the age structured models are γ^{-1}, σ^{-1}, v_a, N, N_{c15}, N_{a15}, N_{c5}, N_{a5}. The values of birth rate were also obtained from the Annual Report of the Saskatchewan Department of Public Health [21]. The values of parameters of γ^{-1} and σ^{-1} are as given by [30]. Moreover, to compare the results, I have built two types of age structured models – one where the lower age group consists of children below 5 years of age (with population in child age group denoted as N_{c5}, and population of adult age group denoted as N_{a5}), and another in which children consist of individuals below 15 years of age (population of age groups denoted as N_{c15} and N_{a15}, respectively). Thus, the birth rates are different among these two types of models (denoted as v_{a5} and v_{a15} respectively), to let all the models have a similar birth population per unit time. Finally, all the compartmental parameters and initial values are specified at Table 4.1 and Table 4.2.

4.3.3 The measurement model

Model only incorporating with the monthly empirical dataset

In the age-structured particle filtering models only incorporating with the monthly empirical dataset, the measurement model is similar with the aggregate particle filering model, which is introduced in chapter 3

Table 4.1: Table showing the value of parameters of the measles age-structured particle filtering models.

Parameter	Value	Units
γ^{-1}	5	Day
σ^{-1}	8	Day
v_{a15}	0.045	1/Year
v_{a5}	0.034	1/Year
N_{c5}	98,743	Person
N_{a5}	764,802	Person
N_{c15}	286,537	Person
N_{a15}	577,008	Person

Table 4.2: Table showing initial values of the stocks in the measles age-structured particle filtering models. The units are all person.

Parameter	Value
S_{0c5}	43177 - I_{0c5}
E_{0c5}	0
I_{0c5}	Uniform[0, 4318)
R_{0c5}	55566
S_{0a5}	43177 - I_{0a5}
E_{0a5}	0
I_{0a5}	Uniform[0, 4318)
R_{0a5}	721625
S_{0c15}	67454 - I_{0c15}
E_{0c15}	0
I_{0c15}	Uniform[0, 6745)
R_{0c15}	219083
S_{0a15}	28900 - I_{0a15}
E_{0a15}	0
I_{0a15}	Uniform[0, 2890)
R_{0a15}	548108

of Equation (3.7). The measurement model of age stratified model only incorporating with the monthly empirical dataset is:

$$I_{emk} = (I_{rmak} + I_{rmck}) + n_{mk} \tag{4.15}$$

where I_{emk} is monthly measles reported cases, the same as in the aggregate model in Equation (3.7); I_{rmak} and I_{rmck} are calculated in the state space model of the age stratified model in Equation (4.14); n_{mk} is the measurement noise associated with the montly empirical dataset.

Model incorporating with both of the monthly and yearly empirical datasets

In the age stratified particle filtering model incorporating with both of the monthly and yearly empirical datasets, empirical observations include three components. Thus, in the measurement vector y_k^M, M is 3 in age stratified model. In addition to the empirical data associated with monthly reported cases, empirical data further include annual reported cases for each of the two age groups (children and adults). The measurement model of age stratified model incorporating with both of the monthly and yearly empirical datasets can thus be represented as:

$$\begin{aligned} I_{emk} &= (I_{rmak} + I_{rmck}) + n_{mk} \\ I_{eyck_y} &= I_{yck_y} + n_{yck_y} \\ I_{eyak_y} &= I_{yak_y} + n_{yak_y} \end{aligned} \tag{4.16}$$

where the parameters I_{emk}, n_{mk}, I_{rmak} and I_{rmck} are the same as in Equation (4.15); I_{eyck_y} consists of the annual measured cases of child age group, while I_{eyak_y} represents the annual measured cases of adult age group; I_{yck_y} and I_{yak_y} are the annual reported cases calculated by the state space model of Equation (4.14) of children and adult age group, respectively; n_{yck_y} and n_{yak_y} are the measurement noise associated with these two age groups.

It is notable that the subscript k_y indicates annual time points, while the unit of time in the models in this paper is month. Thus I_{yck_y} and I_{yak_y} could be obtained by the sum of I_{rmak} and I_{rmck} in the model each year.

4.3.4 The proposal distribution

The Condensation Algorithm [14, 74] is also applied in this project to implement the age-structured particle filter model, same as the aggregate particle filtering model introduced in the chapter 3.

4.3.5 Likelihood function

Model only incorporating with the monthly empirical dataset

In the particle filtering age strucutred model only incorporating with the monthly empirical dataset, because the model lacks the capacity to distinguish between individuals within different age groups as necessary to compare to the yearly age-stratified reported values, the measured data is a one-dimensional vector consisting of the monthly reported cases. The likelihood function is similar to the measles aggregate particle filtering model introduced in chapter 3 in Equation (3.11), by using the negative binomial distribution. Thus, the weight update rule (likelihood function) of the age structured model only incorporating with the monthly empirical dataset could also be achieved by calculating the value of $p(y_{mk}|(I_{rmak}+I_{rmck}))$, where ($y_{mk}$ is the empirical data as given by the monthly reported measles cases at time k, and $(I_{rmak}+I_{rmck})$ is the monthly reported cases across the total population calculated by the dynamic model.

Model incorporating with both of the monthly and yearly empirical datasets

In the age stratified model incorporating with both of the monthly and yearly empirical datasets, the weight update rule is similar to the model only incorporating with the monthly empirical dataset, except for the update associated with the close of each year. Specifically, the weights of particles associated with the age stratified model from January to November are only updated by the monthly empirical data – monthly measles reported cases at each time (using the likelihood function given in Equation (3.11)). However, the weight at the end of the last month (December) of each year is updated by the combination of three parts. The likelihood formulation of age stratified model is listed as follows:

$$
\begin{aligned}
&L_{AgeStructuredModel} = L_{month} * L_{yearlyChild} * L_{yearlyAdult} \\
&L_{month} = p(y_{mk}|I_{rmk}) \\
&L_{yearlyChild} = \begin{cases} 1, & \text{if } (k \mod 12) \neq 0 \\ p(y_{yck}|I_{ryck}), & \text{if } (k \mod 12) = 0 \end{cases} \\
&L_{yearlyAdult} = \begin{cases} 1, & \text{if } (k \mod 12) \neq 0 \\ p(y_{yak}|I_{ryak}), & \text{if } (k \mod 12) = 0 \end{cases}
\end{aligned}
\tag{4.17}
$$

where L_{month} is the likelihood function based on the monthly empirical data for the total population. The other two likelihood functions reflect the fact that annual totals are available on an age-specific basis at year end. $L_{yearlyChild}$ is the likelihood function based on the yearly empirical data for the child age group. $L_{yearlyAdult}$ is the likelihood function based on the yearly empirical data of the adult age group.

4.3.6 Evaluating particle filter performance

The monthly discrepancy of each time is simply the Root Mean Squared error (RMSE) between the monthly empirical data at that time and the related data calculated by the particle filtering model [84, 107]. To get the yearly discrepancy of each time (here, successive Months), the RMSE was calculated for each age group of each year (similary to monthly discrepancy). Then the yearly discrepancy is the sum across all age groups of the yearly RMSE over 12 (to convert the unit to Month). For the discrepancy calculated in the prediction models, with a Prediction Start Time T^*, the model calculated a prediction discrepancy using a simple variant of the strategy of the discrepancy used in considering all the time frame, but limited to considering only times T^* and larger.

4.3.7 Empirical dataset

In the yearly empirical dataset, these yearly reported cases are split into different age groups. In a small minority of years (from 1926 to 1941), the age categories present in the reported data do not correspond neatly to the age group categories in the models (considering children as being those within their first 5 years or first 15 years). For these cases, I split them into the age categories of the models proportionally.

The yearly empirical data related to multiple age categories are available from year 1925 to 1956. During the process in preparing the yearly empirical data match the two age groups in the two age group particle filtering models (children age group is for those people up to 5 or up to 15 years), I need to split some age categories due to two reasons. The first reason is because the division of the age group in empirical dataset does not match the two age groups in particle filtering models. Specifically, from year 1926 to 1941, I need to split the counts of reported measles cases in age category "1-6 years" in age 5 proportionally (four fifths goes to the child age group, and one fifth goes to adult age group). This problem only related to the age group model of child group up to 5 years old. The second reason is because there is a category in the empirical yearly dataset of "age not stated". Thus, I need to split the counts in this category to corresponding age groups with the two age group models – the persons in the child age group up to 5 or 15 years old proportionally (based on the proportion calculated by the age categories has labeled age clearly).

4.3.8 Model characterization

To research on the performance of incorporating particle filtering into the age structured compartmental model, I have built 4 age-structured particle filtering models. In the particle filtering models, the number of particles in the particle filtering algorithm is 5000. These models respectively listed as follows:

(1) $PF_{age\_5\_monthly}$. The age structure model where the child age group includes those less than 5 years old, and only incorporated with the monthly reported empirical data.

(2) $PF_{age\_5\_both}$. The age structure model where the child age group includes those less than 5 years old,

and incorporated with both the monthly reported and yearly reported age group empirical data.

(3) $PF_{age\_15\_monthly}$. The age structure model where the child age group includes those less than 15 years old, and only incorporated with the monthly reported empirical data.

(4) $PF_{age\_15\_both}$. The age structure model where the child age group includes those less than 15 years old, and incorporated with both the monthly reported and yearly reported age group empirical data.

To compare the age structured particle filtering model with the aggregate particle filtering model, the result of the aggregate particle filtering model introduced in chapter 3 is also included in this chapter, denoted as $PF_{aggregate}$.

4.4 Results

4.4.1 Results of models incorporating empirical datasets across all timeframe

Table 4.3: Comparison of the average discrepancy of all seven models by incorporating empirical data across all observation points.

Model	Monthly	Yearly in Month	Total
$PF_{aggregate}$	104.6 (99.4, 109.9)	NONE	NONE
$PF_{age\_5\_monthly}$	96.1 (91.5, 100.7)	179.5 (160.6, 198.3)	275.5 (260.2, 290.9)
$PF_{age\_5\_both}$	97.8 (94.1, 101.5)	144.8 (112.5, 177.1)	242.6 (210.2, 275.1)
$PF_{age\_15\_monthly}$	95.7 (89.8, 101.6)	45.9 (39.9, 51.9)	141.6 (133.8, 149.4)
$PF_{age\_15\_both}$	96.1 (91.6, 100.5)	39.9 (34.3, 45.6)	136.0 (127.6, 144.5)

Each of the particle filtering models is run for 5 realizations (each with a distinct random seed drawn from the same set). Shown here are the average and 95% confidence intervals (in parentheses) of the mean discrepancy for each model variant.

Table 4.3 lists the discrepancies of the measles particle filtering models, including the model with an aggregate population (denoted as $PF_{aggregate}$) and the other four models featuring age structured populations. Each particle filtering model is run for 5 realizations, and the average (across realizations) value and the 95% confidence interval (in parentheses) of the mean discrepancy for each model variant are listed. It is notable that among the four age structured particle filtering models, two situations are considered. Firstly, the population is divided according to an age threshold (either age 5 and age 15). Secondly, for each age structured models divided in age 5 and age 15, I further compared the impact of incorporating only the monthly empirical dataset with the whole population and (by contrast) incorporating both the monthly empirical dataset with the whole population and the yearly empirical datasets for each age group, respectively. It is notable that the yearly discrepancy is not available for the aggregated population model, in light of

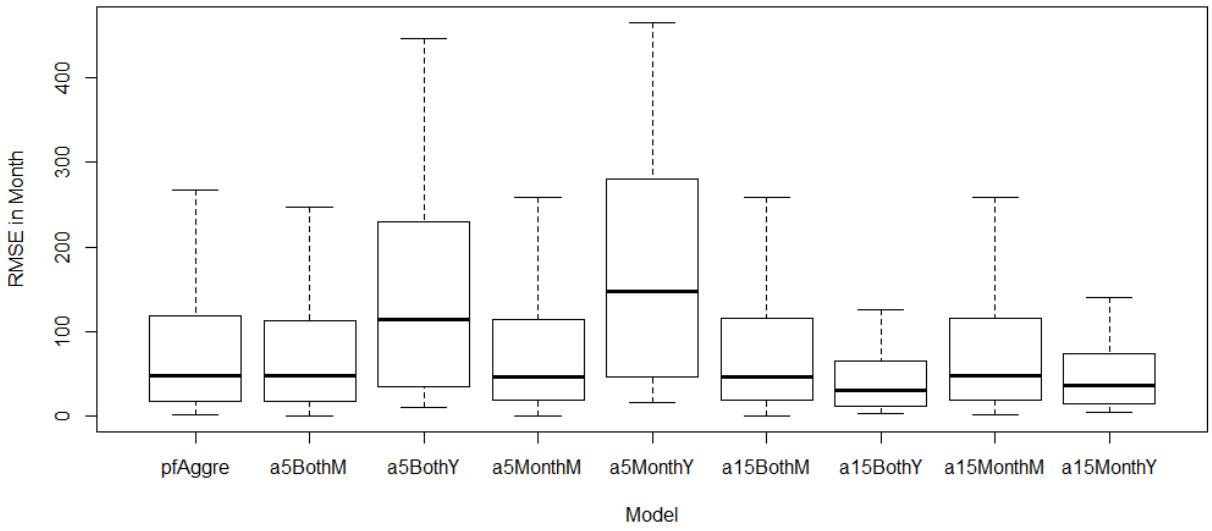


Figure 4.2: Box plots of monthly and yearly discrepancy of all measles models by incorporating empirical data across all observation points.

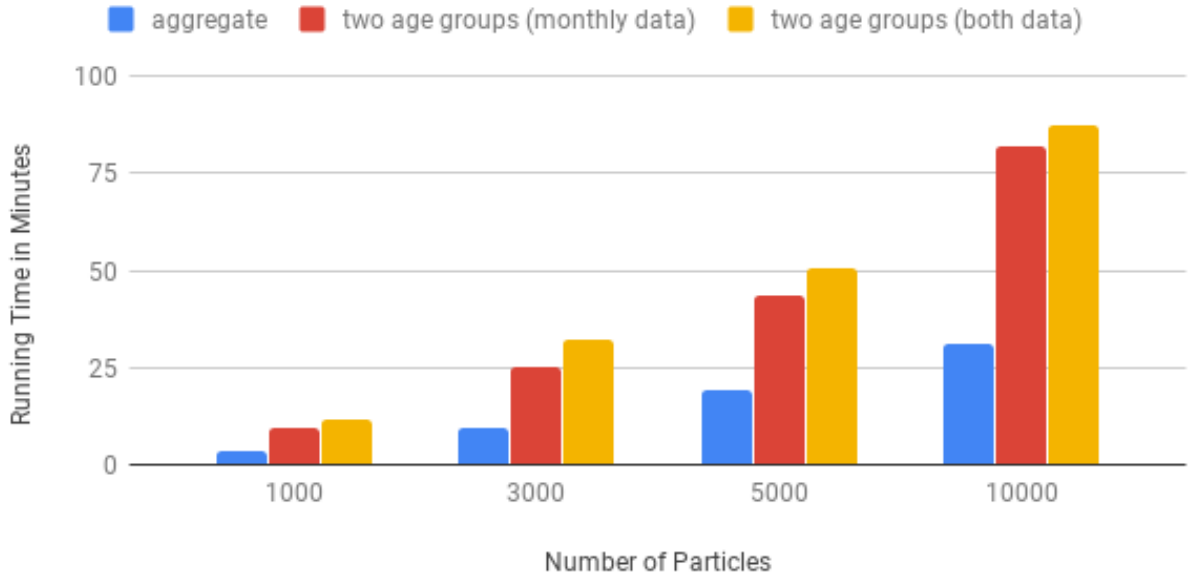


Figure 4.3: Running time of the measles particle filtering models.

the absence of age-specific population categories against which the age-specific incident case counts can be compared. The results of Table 4.3 suggests that the age-structured particle filtering models perform better (as measured by discrepancy) than the aggregated population model, because the monthly discrepancy of all the four stratified age group models are smaller than that for the aggregated population model. Secondly, results indicate that an appropriate splitting of the age groups is also important in improving the simulation results of the models. Table 4.3 indicate that the discrepancy of the stratified age group models splitting the age group at age 15 years are consistently smaller than the models splitting the age group at age 5 years. Finally, results suggest that incorporating both the monthly empirical reported cases and yearly empirical data of each age group may be also helpful in improving the simulation accuracy of the models (when compared to only considering the monthly results), but further realizations are required to confirm these results. Finally, the results suggest that the model $PF_{age\_15\_both}$ offers the minimum discrepancy. It is notable that while aggregate population models cannot be compared directly against the other models in terms of total discrepancy, such models suffer from the least favorable score in terms of the metric by which comparisons can be made (the monthly discrepancy).

Figure 4.2 shows the box plot of the discrepancy among the five particle filtering models, in which each of the data points being summarized represents the discrepancy for a single month. It is notable that the dataset of the discrepancy of the box plot in the particle filtering model is calculated by the average value among five realizations at each time point. Both the monthly and yearly distribution of the discrepancies of each age structured models are plotted in figure 4.2. This box plot also indicates that the model $PF_{age\_15\_both}$ has the smallest median discrepancy. Moreover, the datasets of the discrepancy of the model $PF_{age\_15\_both}$ has a narrow distribution, especially for the dataset of the yearly discrepancy.

Figure 4.3 shows the running time of the measles particle filtering models with three categories – the aggregate model, the two-age-group model only incorporating monthly empirical data, and the two-age-group model incorporating both monthly and yearly empirical data. It is notable that all of the models are built using Anylogic 8.1.0 software, a Java-based simulation platform whose flexibility facilitates incorporation of the particle filtering mechanisms. Figure 4.3 indicates that the running time grows linearly with an increase in the number of particles. However, the running time scales superlinearly (rather than linearly) with the number of the age groups in the model. The author suspects that the reason is that compared with the aggregate model (only one age group), the two-age-group model requires additional calculation of the contact matrix between these two age groups. And the two-age-group model incorporated both monthly and yearly data requires added running time compared with the two-age-group model only incorporating the monthly data. It is notable that the running time of the deterministic model with aggregate population structure – and lacking any particle filtering machinery – is only 0.056 minutes, which is far faster than the particle filtering models. Finally, all the measles particle filtering models were run on the author's personal computer, with an Intel i7 2.7 GHZ processor, and memory size of 8GB DRAM).

According to the results shown by table 4.3 and figure 4.2, in the models examined here, the choice of

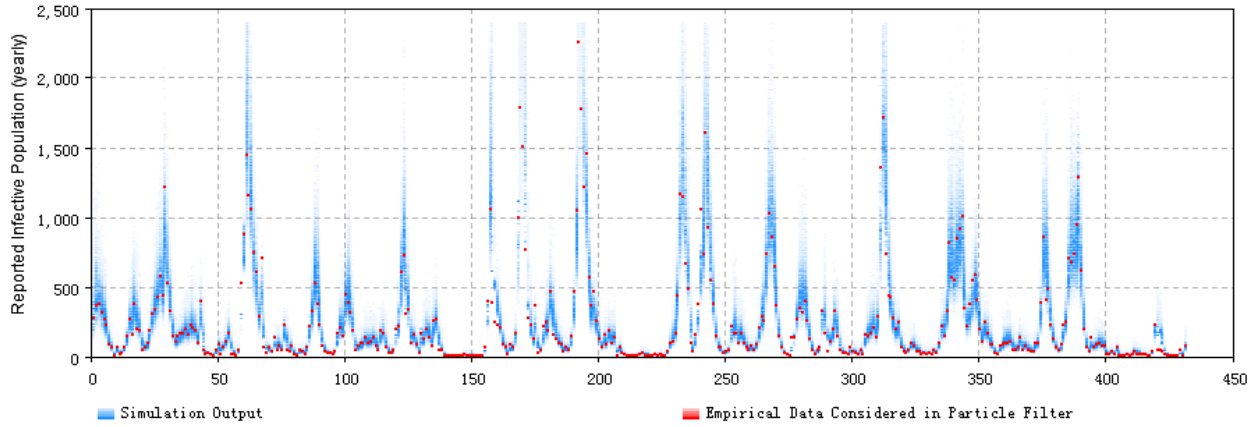
whether to divide the population in age 5 or 15 influences the discrepancies of the age structured models to a greater degree than does the choice of whether to incorporate the monthly empirical dataset or incorporating both the monthly and yearly empirical datasets. Thus, the 2D histograms sampled monthly posterior distributions of the monthly reported measles cases across all population and yearly reported cases of measles of each age groups of the age structured particle filtering models are plotted, taking the two filtering models incorporating both the monthly and yearly empirical datasets as an example (see figure 4.4 and figure 4.5).

Figure 4.4 displays 2D histogram plots of sampled monthly posterior distributions from both the monthly reported measles cases of all population and the yearly reported measles cases of each age group of the measles particle filtering model by dividing the population at the age of 5 and incorporating both the monthly and yearly empirical datasets (denoted as $PF_{age\_5\_both}$). The empirical data are also shown in the 2D histogram by the red dash present for each month. Figure 4.4 indicates that although the monthly empirical data are mostly located in the high density range of the results of model, the yearly empirical data of both age groups are mostly located out the range of the results of the model.

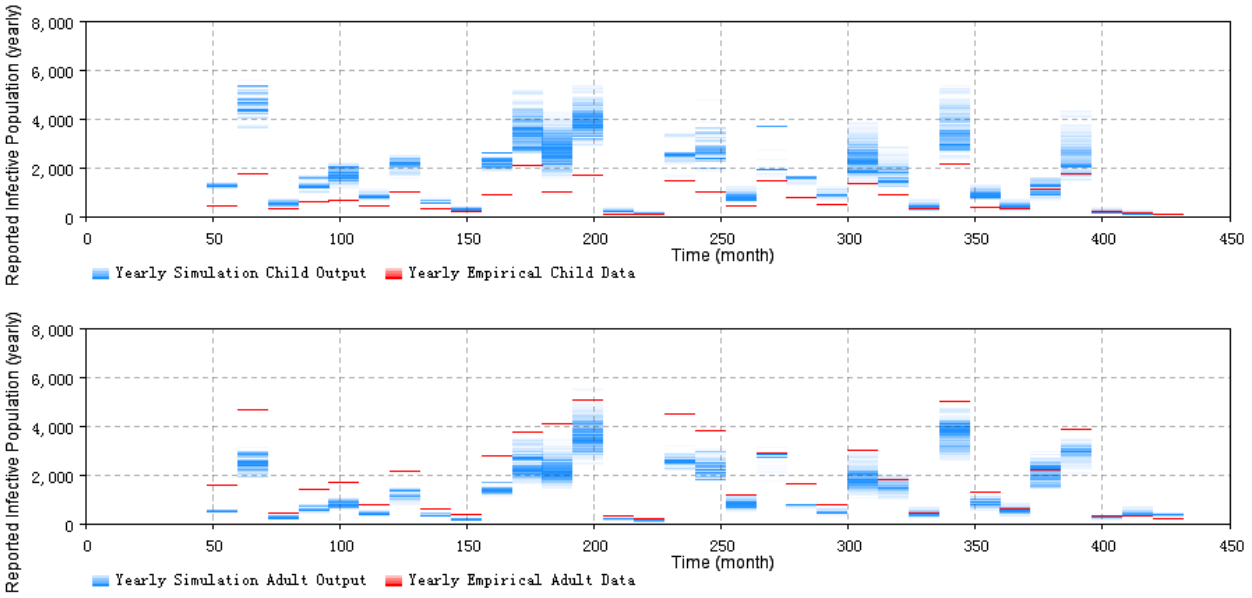
Figure 4.5 displays the 2D histogram plots of sampled monthly posterior distributions from both the monthly reported measles cases of all population and the yearly reported measles cases of each age group of the measles particle filtering model by dividing the population at the age of 15 and incorporating both the monthly and yearly empirical datasets (denoted as $PF_{age\_15\_both}$). The empirical data are also shown in the 2D histogram by the red line in each time. Figure 4.5 indicates that both the monthly empirical data and yearly empirical data with each age group are mostly located in the high density range of the results of model.

By comparing figure 4.4 and figure 4.5, we can see that for the models examined here, the particle filtering models where the individuals in the child age group are up to 15 years of age can estimate the yearly reported measles cases of each age group more accurately than the models particle filtering models where the individuals in the child age group are up to year 5.

Particle filtering models can contribute to the estimation of model states. Although there are more latent states in the age structured particle filtering models with two age groups compared than for the aggregate population particle filtering model discussed in chapter 3, the age structured particle filtering models in this chapter can also estimate the model states effectively. Similar to the measles particle filtering model with an aggregate population discussed in chapter 3, the states in the age-structured models are latent. What can be empirically observed is the noisy reported measles cases related to the Infectious states of both age groups (I_a and I_c). However, the methodology of particle filtering provides an approach to estimate (via sampling from) the distribution of values of these latent states. This ability to estimate the value of latent states such as the reservoir of susceptible people can aid researchers and public health agencies to in terms of understanding the underlying epidemiological situation from multiple lines of evidence, as constrained by understanding of the dynamics emerging from the structure of the underlying system, as characterized by a dynamic model. Figure 4.6 shows the 2D histogram plots of the latent stocks in the age structured model

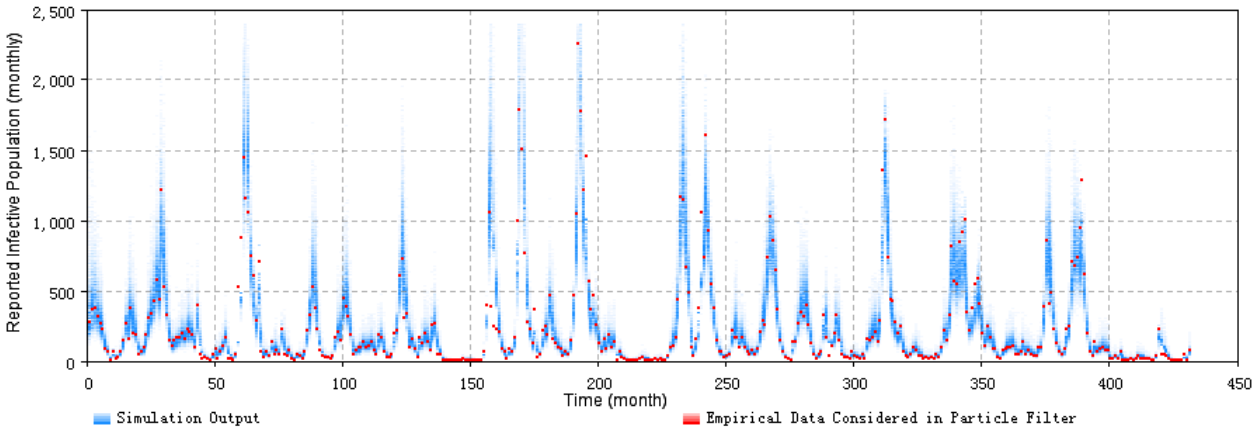


(a)

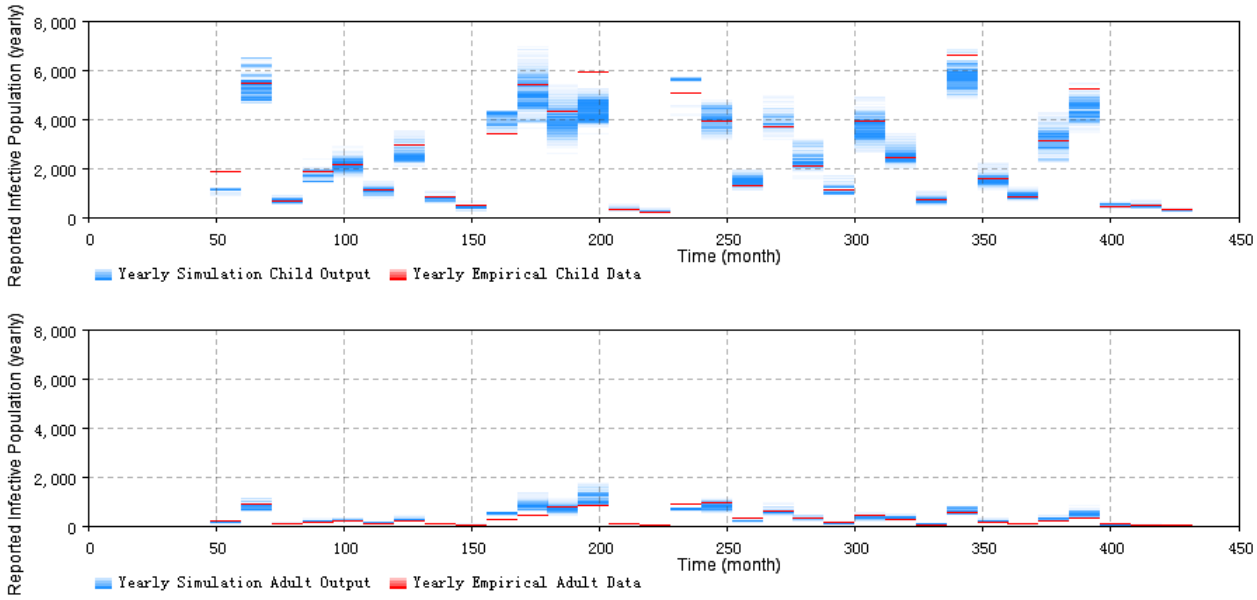


(b)

Figure 4.4: 2D histogram posterior result of total timeframe of the model incorporating both monthly and yearly empirical datasets split at year 5. (a) the monthly particle filtering result across all population. (b) the yearly particle filtering result of the children and adult age groups.

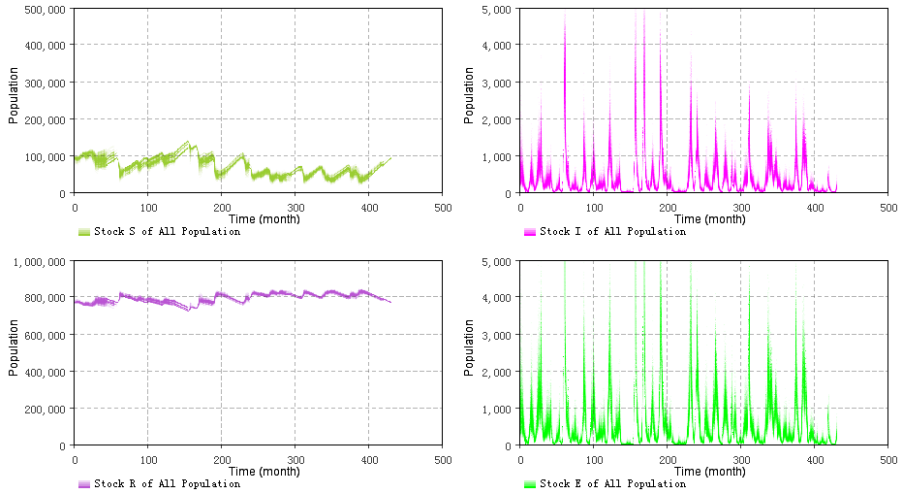


(a)

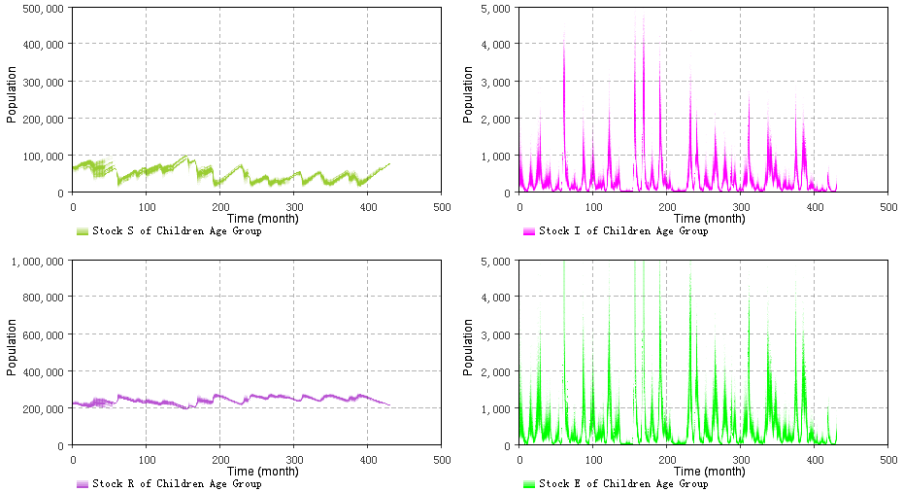


(b)

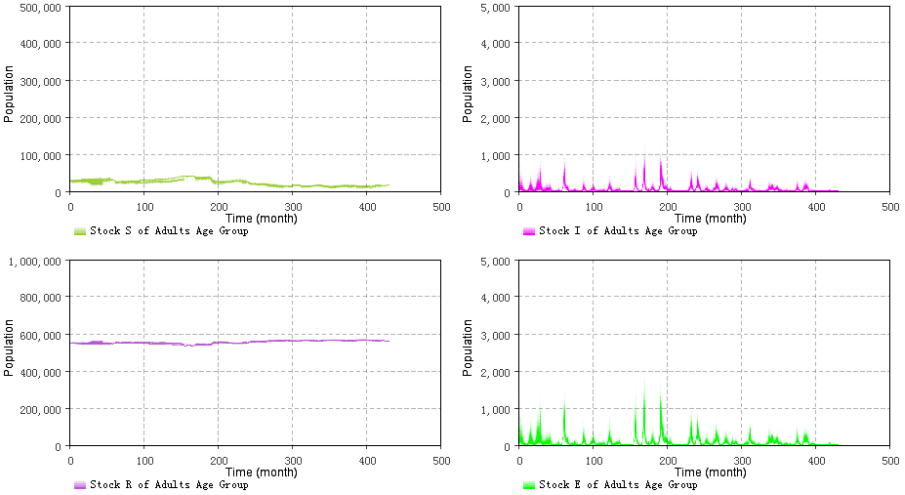
Figure 4.5: 2D histogram posterior result of total timeframe of the model incorporating both monthly and yearly empirical datasets split at year 15. (a) the monthly particle filtering result across all population. (b) the yearly particle filtering result of the children and adult age groups.



(a)



(b)



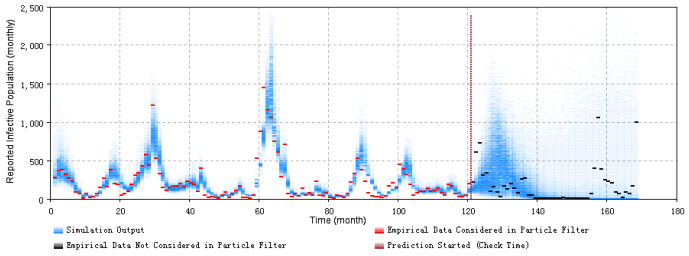
(c)

Figure 4.6: 2D histogram results for the S, E, I, R stocks with different age groups of the minimum discrepancy model incorporating the empirical data across all timeframe. (a) across all population. (b) the child age group (those within their first 15 years of life). (c) the adult age group (years 15 and up).

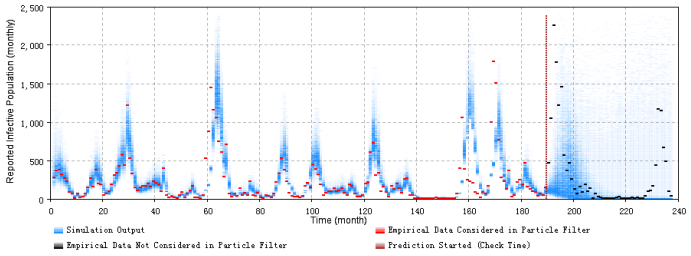
with the minimum discrepancy (the child age group up to age 15 and incorporating both the monthly and yearly empirical datasets across all timeframe). Figure 4.6 indicates that most of the susceptible, exposed and infectious people are located in the child (less than 15 years) age group, while most of the recovered population are located in the adult (equal and greater than 15 years) age group. This lies in accordance with the expectations for measles transmission in the real world, and builds confidence in the capacity of the model to meaningfully estimate latent state. As noted below, estimation of latent state can be an important enabler for understanding of the effects of interventions.

4.4.2 Prediction results of the minimal discrepancy model

In this section, I assess the predictive capacity of the minimal discrepancy model identified in the previous section. This minimal discrepancy model is still the strongest performing one among the group of models $PF_{age\_15\_both}$. By changing different Prediction Start Time of T^*, I have performed prediction from different archetypal situations, the same situations as the ones introduced in the chapter 3:



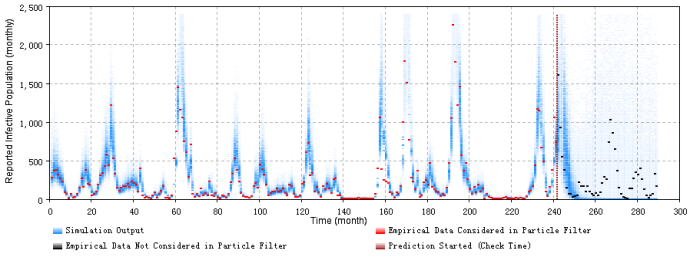
(a)



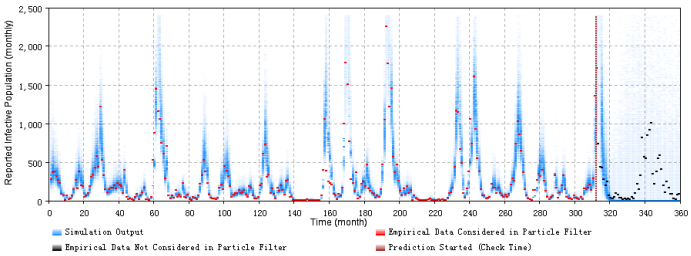
(b)

Figure 4.7: 2D histogram of predicting from the first or second time points of an outbreak of the minimum discrepancy model (a) predicted from the month 121. (b) predicted from the month 190.

Figure 4.7 displays the prediction results for the minimum discrepancy model in situations in which the prediction started from the first or second time points of an outbreak. It does so with two experiments which are prediction started from month 121 ($T^* = 121$) – with monthly prediction discrepancy 306.0, the sum of yearly prediction discrepancy of all age groups per month is 246.7 – and month 190 ($T^* = 190$) – with monthly prediction discrepancy 320.4, the sum of yearly prediction discrepancy of all age groups per month is 237.2 – respectively. In the prediction process for the particle filtering model, the weights of particles will

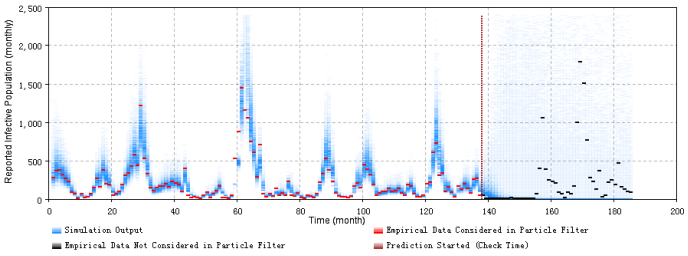


(a)

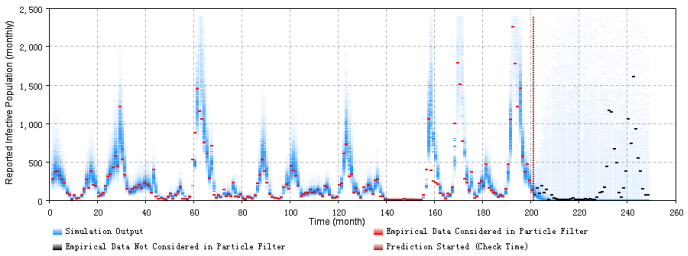


(b)

Figure 4.8: 2D histogram of predicting from the peak of an outbreak of the minimum discrepancy model. (a) predicted from the month 242. (b) predicted from the month 312.

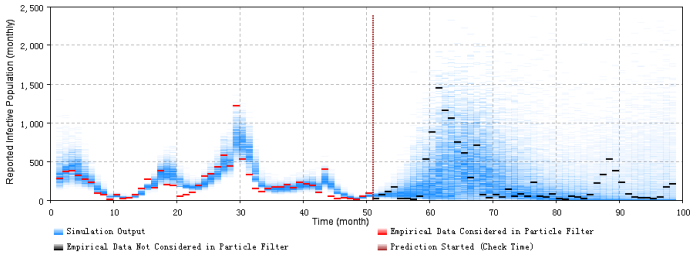


(a)

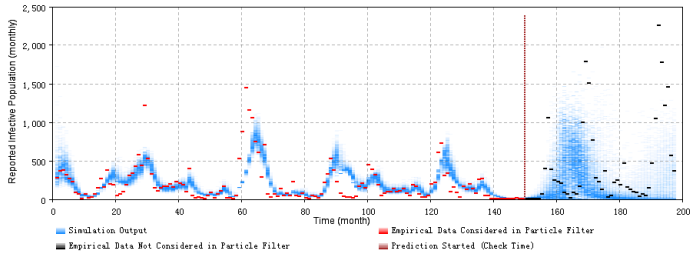


(b)

Figure 4.9: 2D histogram of predicting from the end of an outbreak of the minimum discrepancy model. (a) predicted from the month 138. (b) predicted from the month 201.



(a)



(b)

Figure 4.10: 2D histogram of predicting before the next outbreak of the minimum discrepancy model. (a) predicted from the month 51. (b) predicted from the month 150.

cease updating (as a result of incorporating the empirical data) at the "Prediction Started Time". From that point forwards, all the particles run without new empirical data being considered. In this chapter, all the prediction experiments are run for 4 years following the "Prediction Start Time" time. In the 2D histogram plot of figure 4.7, the empirical data considered in the particle filtering process (i.e., incorporated in training the models) are shown in red, while the empirical data that was not considered in the particle filtering process (i.e., data points only displayed to compared with the results of models) are shown in black. The vertical straight line labels the "Prediction Start Time" of T^* of each experiment.

Figure 4.8 displays the prediction results of the particle filtering model in the situation in which the prediction started from the peak of an outbreak. As above, it also does so with two experiments, prediction started from month 242 ($T^* = 242$) – with monthly prediction discrepancy 305.7, the sum of yearly prediction discrepancy of all age groups per month is 205.2 – and month 312 ($T^* = 312$) – with monthly prediction discrepancy 306.9, the sum of yearly prediction discrepancy of all age groups per month is 201.6 – respectively. The layout of the 2D histogram plot of figure 4.8 is the same as that for figure 4.7.

Figure 4.9 displays the prediction results of the particle filtering model in the situations in which the prediction started from the end of an outbreak. As above, it accomplishes this with two experiments , prediction started from month 138 ($T^* = 138$) – with monthly prediction discrepancy 302.6, the sum of yearly prediction discrepancy of all age groups per month is 248.0 – and month 201 ($T^* = 201$) – with monthly prediction discrepancy 316.7, the sum of yearly prediction discrepancy of all age groups per month is 217.3 – respectively. The layout of the 2D histogram plot of figure 4.9 is the same as that for figure 4.7.

Figure 4.10 displays the prediction results of the particle filtering model in situations in which we predict

the next outbreak using two experiments, prediction started from month 51 ($T^* = 51$) – with monthly prediction discrepancy 324.9, the sum of yearly prediction discrepancy of all age groups per month is 198.8 – and month 150 ($T^* = 150$) – with monthly prediction discrepancy 353.0, the sum of yearly prediction discrepancy of all age groups per month is 268.5 – respectively. The layout of the 2D histogram plot of figure 4.10 is the same as that for figure 4.7.

These prediction results display the prediction results of these situations with the monthly 2D histogram of minimum discrepancy particle filtering model with age structure. During several prediction time points following the "Prediction Start Time" of the experiments shown in figures 4.7 – 4.10, the measles minimum discrepancy model offers the capacity to probabilistically anticipate measles dynamics with a fair degree of accuracy. Moreover, by comparing with the prediction discrepancy of figures 4.7 – 4.10 to figures 3.10–3.13 in chapter 3 of the aggregate model, the age structured particle filtering model (the one with the minimum discrepancy) offers higher accuracy with smaller prediction discrepancy.

4.4.3 Intervention with the minimum discrepancy model

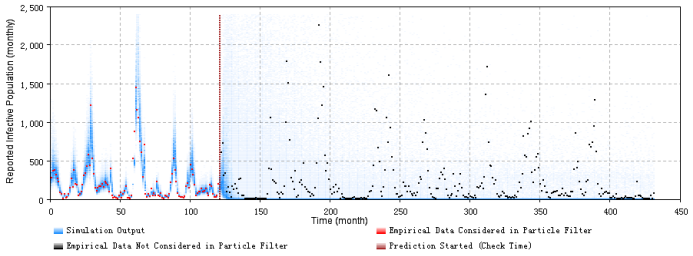


Figure 4.11: 2D histogram of simulating an outbreak-response quarantine intervention. This is realized by decreasing the contact rate to be 20% less than the value before intervention

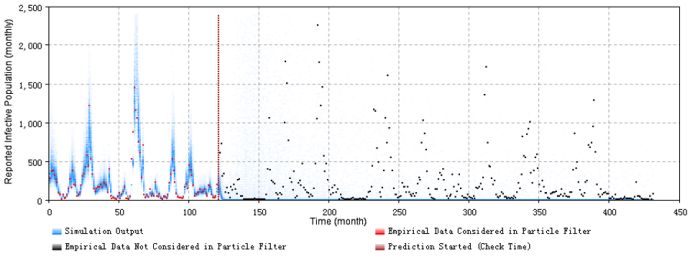


Figure 4.12: 2D histogram of simulating an outbreak-response quarantine intervention. This is realized by decreasing the contact rate to be 50% less than the value before intervention

Based on the minimum discrepancy particle filtering pertussis model discussed in the previous section, I have implemented several intervention experiments to simulate in a stylized fashion public health intervention policies to control pertussis outbreaks. Specifically, two intervention policies have been explored – quarantining infected individuals and vaccinating the susceptible population.

Within this investigation, the intervention strategies are normally performed before or at the very be-

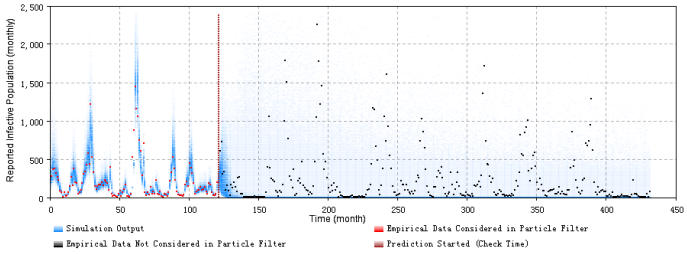


Figure 4.13: 2D histogram of simulating an outbreak-response vaccination intervention. This is realized by setting the value of the vaccination parameter to 0.2.

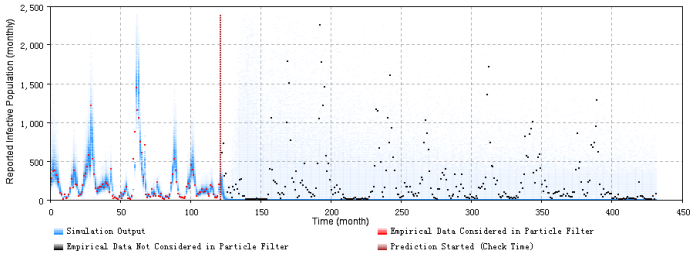


Figure 4.14: 2D histogram of simulating an outbreak-response vaccination intervention. This is realized by setting the value of the vaccination parameter to 0.5.

ginning of an outbreak. Moreover, to support ready comparing with the baseline prediction result of the minimum discrepancy model absent any interventions, all of the intervention strategies are simulated as initiating at the start month of an outbreak (month 121) in this project. The normal prediction result of the minimum discrepancy model absent any interventions is shown in figure 4.7 (a).

Figure 4.11 and figure 4.12 display the simulation results of the quarantine intervention strategies by decreasing the contact rate parameter to be 20% and 50% less than the value before intervention, respectively. The 2D histogram plot of the prediction result absent any intervention shown in figure 4.7 (a). The red dots represent the empirical data incorporated in the particle filtering model, while the black dots represent the empirical data not incorporated in the model, but included for comparison. It is notable that the empirical data shown in black correspond to a situation of without any intervention (in contrast to the counter-factual character of the intervention scenario). By comparing the quarantine intervention results (see figure 4.11 and figure 4.12) with the model result without intervention shown in figure 4.7 (a) and the empirical data during the intervention period (not incorporated in the particle filtering model), we can see that, although the interventions are implemented in a stylized fashion, by virtue of its ability to estimate the underlying epidemiological state, the measles particle filtering model is capable of evaluating public health intervention policies to reduce or even avoid the outbreak of measles.

To simulate an outbreak-response intervention of vaccination (immunization), a vaccination parameter was added to the model to move a fraction of the population from the Susceptible stock (prior to vaccination) to the Recovered stock (following vaccination). Figure 4.13 and figure 4.14 show the results of the vaccination intervention. The layout of the 2D histogram plots of the vaccination intervention shown in figure 4.13 and

figure 4.14 are similar those for the plots of quarantine interventions shown in figure 4.11 and figure 4.12. While the interventions are implemented in a stylized fashion, these results of interventions demonstrate that, by virtue of its ability to estimate the underlying epidemiological state, the measles particle filtering model is capable of evaluating public health intervention policies to reduce or even avoid the outbreak of measles.

4.5 Discussion and conclusion

This chapter explores the age-structured particle filtering models by dividing the population into two age groups – children and adults – to capture the transmission difference between those groups. Moreover, four variants of age-structured measles particle filtering models are implemented by considering all combinations of two distinct and dichotomous choices – dividing the population at the year 5 or 15, and incorporating only the monthly empirical dataset or incorporating both the monthly and yearly empirical datasets into the age-structured particle filtering models, respectively. Then, by comparing the results (including the result of the aggregate particle filtering model introduced in chapter 3), the strongest predictive performance emerged from the age stratified model whose child age group is defined as including those up to 15 years old, and considering both monthly empirical data regarding the total population and yearly reported cases of each age group. However, testing the models over additional realizations is required in order to validate these results. As discussed in chapter 3, the particle filtering algorithm can mitigate significant weaknesses and simplifications associated with aggregate compartmental models and noisy empirical data. It is notable that while age-structured compartmental models capture heterogeneous mixing among the different age groups, the individuals within the same age group are engaged in homogeneous mixing.

Although the total number of the states in the age-structured particle filtering model is higher than for the aggregate population model, the particle filtering supports estimation of the entire state of the dynamic model with a relatively higher dimensional state space (both latent or observable) for each point in time, by virtue of incorporating with the empirical dataset. In the age-structured particle filtering model of measles with two age groups in this chapter, the latent states related to the Susceptible population, the Infectious population, the Exposed population, the Recovered population of each age group, the infectious contact rate of each age group, and also the fraction of total contacts of the children age group with the infected individuals in the children age group. The latter of those supports derivation of a quantity that can be compared to the reported rate of measles incidence in each time point within the empirical data. This capacity to estimate the complete state of the system can help provide insight into the current epidemiological context, and support researchers in performing related research.

Similar to the measles particle filtering model with aggregate population structure introduced in the chapter 3, noise in the stochastic processes in the state space model impacts the particle filtering model differently during the estimation and prediction periods. As a result, the noise in the particle filtering models in this chapter should also be controlled within a proper range, by tuning the parameters of diffusion

coefficients in the stochastic processes related to Brownian motion.

Although particle filtering models with age-structured population also benefit from the ability to estimate the initial values of the stocks in the system dynamics models, estimation of the initial values in the age-structured population models are more complex than for the aggregate population particle filtering models introduced in chapter 3. The reason is that, in the age-structured population particle filtering models, the population distribution both among different stocks with each age group and between different age groups need to be considered. In this project, the population distribution among the different age groups are tuned manually, while the population distribution among different stocks within an age group is estimated by the particle filtering algorithm, by imposing initial distributions of values across different particles, and allowing the distribution of particles to then evolve according to the particle filtering algorithm.

For the age-structured measles models considered here, the hypotheses proposed in the thesis statements are proven. Firstly, the 2D histogram plots with the particle filtering models' output data (as sampled from the posterior distribution of the monthly and yearly reported cases of measles) and the empirical data indicate that the high probability density region of the model's prediction of empirical data lies near those data. Secondly, the discrepancy of the age-structured particle filtering model is even smaller than the aggregate population particle filtering model. And chapter 3 has demonstrated how the aggregate population particle filtering model has been proven to be significantly higher accuracy prediction of such case counts than are supported within a traditional calibrated model.

Finally, the experiments regarding interventions indicate that measles particle filtering models are capable of evaluating interventions by examining the outcome of potential intervention strategies using model results via exogenously altering the value of model parameters at pre-specified points in time, such as the contact rate and the total number of individuals in each stocks.

Chapter 5

Pertussis models of particle filtering

5.1 Introduction

Pertussis is another common childhood epidemiology disease, which is a highly contagious disease of the respiratory tract that caused by the bacteria – Bordetella pertussis [79]. It is most dangerous for infants. In the infants, the paroxysms maybe followed by a period of apnoea [79]. It could also cause severe complications. And the most relatively common complication is pneumonia, while seizures and encephalopathy occur more rarely [79].

In the pre-vaccination era, pertussis was one of the most common childhood infectious diseases and a major cause of the childhood mortality. In 1860, the mortality rate of all-age pertussis in Demark is 0.015% [19]. Moreover, pertussis is a childhood epidemiology disease which is more prevalent among infants and children. The researches of historical mortality rate from pertussis indicate that the death rate in infancy is higher than other groups [19].

Like measles, pertussis is also a very contagious disease only found in humans, and spreading from person to person [31]. It is a highly communicable disease, and evidence show a secondary attack rate of 80% among susceptible household contacts [32]. Pertussis can spread from one person to another person by coughing, sneezing, and spending a lot of time together [31]. In contrast to measles – in which natural exposure confers life-long immunity – immunity to pertussis is widely believed to wane relatively rapidly, leading to significant risks of infection even in adults who have been previously infected. It is notable that babies can be infected by adults, such as parents, older siblings, and caregivers who might not even know they have already contracted this disease [31]. Also in contrast to measles, pertussis shows no distinct seasonal pattern. However, it may increase in the summer and fall [32].

In recent years, there are an estimated 24.1 million cases of pertussis and about 160,700 deaths per year [37]. Since the 1980s, there is an increased trend in the reported cases of pertussis in the United States [37]. The most recent peak year of the reported cases of pertussis in the United States is 2012, when the Centers for Disease Control and Prevention (CDC) reported 48,277 cases, but many more go undiagnosed and unreported [37]. Thus, researches aimed at estimating and predicting the transmission dynamics of pertussis could help the public agencies to control it, such as performing intervention before the predicted next outbreak.

In this chapter, I will explore the performance of combining the machine learning algorithm – particle filtering, with a widely used compartmental model [50] and empirical datasets of reported pertussis cases in Saskatchewan during the pre-vaccination era (from 1921 to 1956). Moreover, pertussis is a high contagious childhood epidemiology disease and most serious for babies. To capture the heterogeneous characteristics of contacts between the infectious individuals and the other individuals, age-structured models are also performed. Specifically, I have performed two categories of age-structured particle filtering models – with 2 age groups and with 32 age groups. Moreover, I have explored and deduced three methods in calculating the contact matrix, to reduce the freedom (unknown) variables related to the calculating of the contact matrix. Finally, I compare the results get from all the particle filtering models by incorporating the empirical data across the whole timeframe and conduct research on prediction and interventions with the minimum discrepancy particle filtering model.

5.2 The mathematical models

The infectious dynamics of pertussis is more complex than the infectious disease of measles. The infection of measles – an infectious disease caused by Rubeola virus, confers lifelong immunity [40]. However, the pertussis is transmitted by the bacteria of Bordetella pertussis. The immunity acquired by the infection of pertussis is temporary. And, after the most recent pertussis infection increases, the immunity of a person decreases [48]. Thus, people with lower immunity have higher infectivity, and the individuals have fully susceptible are generally have the highest probability to get infected when they have exposed to the pertussis infections.

In this research, I have employed the structure of the pertussis mathematical model in [50]. To capture the characteristics of pertussis in immunity waning and the different level of infectious and recovered, the compartmental model in [50] further divided the infectious population into three groups: weak infectious (I_w), medium infectious (I_m), and fully infectious (I), and divided the recovered population into four groups: R_1, R_2, R_3 and R_4. The immunity of these recovered persons is increased from R_1 to R_4. Moreover, to correspondence with the empirical data (pertussis reported cases in Saskatchewan from 1921 to 1956 in the pre-vaccination era), the stocks related to vaccination (V_1, V_2, V_3 and V_4) in the original compartmental model [50] are not included in this research.

Figure 5.1 shows the mathematical structure [50] of the compartmental pertussis model. In this compartmental model of pertussis in the pre-vaccination era (without considering vaccination), the total population is divided into 8 distinct epidemiological classes. The newborns enter directly to the class S of the susceptible individuals. If a susceptible individual contact with an infective individual and is infected with pertussis successfully, this previous susceptible person becomes infectious and enters the class I of the full infective. Individuals in class I of infective have full cases of pertussis with all the usual symptoms. When individuals recovered from the class I of infective, they have the immunity and enter the class of R_4.

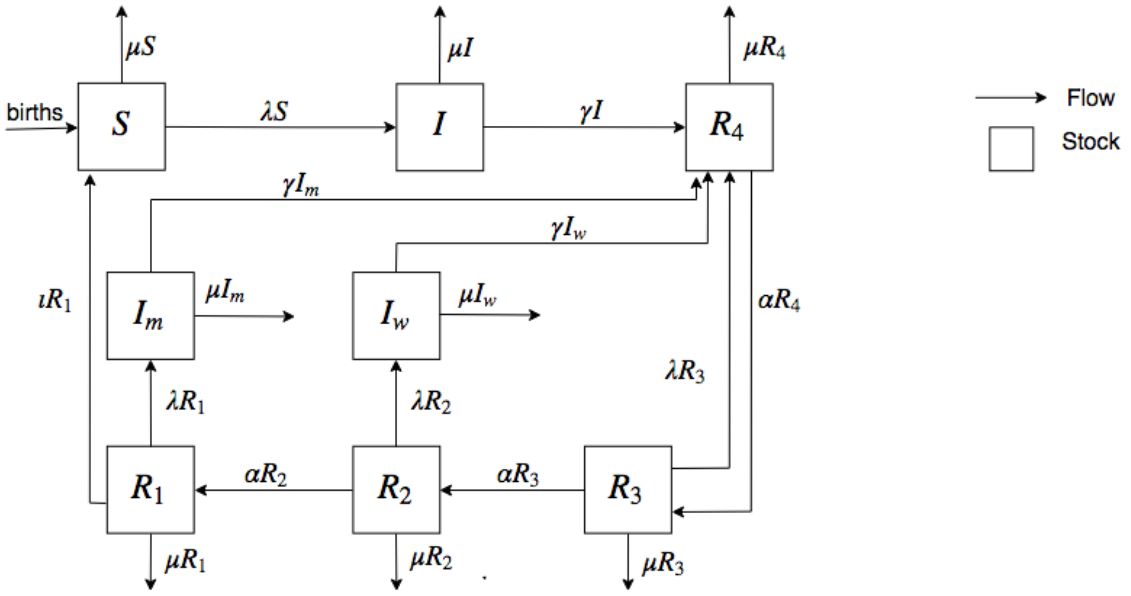


Figure 5.1: The transfer diagram for the pertussis model without vaccination. re-produced from [50]

Individuals in the class of R_4 have the strongest immunity. Thus, they are protected and can not be infected by pertussis. However, as time goes by, their immunity fades and they enter to a less strong immunity class of R_3. When individuals in class R_3 expose to an infective, they move back to the highest immunity class of R_4. If they are not re-exposed to an infective again, their immunity keeps fading and they enter to the relatively lower immunity class of R_2. When a person in the class of R_2 re-exposed to an infective sufficiently for transmission to occur, this person enters the class of I_w with weak infectivity. Individuals in the class of I_w have the weakest infective capability to infect a susceptible. After they recovered, the individuals in class of I_w get the highest immunity and enter the class of R_4 again. If people in the class of R_2 are not re-exposed to the infectives, they keep losing immunity and enter the lowest immunity class of R_1. Similarly, if a person in class of R_1 re-exposed to an infective, this person gets infected with mild infectivity and enters the class of I_m. Individuals in the class of I_m have a higher infectious capability, compared with those in the class of the weak infective (I_w), and have a lower infectious capability, compared with the individuals in the class of the full infective (I). When recovered, the individuals in the class of I_m enter the class of R_4 again. If the individuals in the class of R_1 are not re-exposed, they eventually lose all of their immunity and move back to the class of S with susceptible, finally.

It is notable that the model in the paper of Hethcote (1997) [50] makes use of a formulation in which each state variable is of unit dimension, representing a fraction of the population in different age groups of the same class. However, for the sake of comparison against empirical data, the model in this thesis is represented in a re-dimensionalized fashion, with the state variables representing counts of persons. Based on this structure in figure 5.1, four models are considered in this research (suppose n is the total number of age

groups in the models) – aggregate population ($n = 1$), two-age-groups model ($n = 2$), 32-age-groups model with the contact matrix introduced in the paper of Hethcote (1997) [50] and 32-age-groups model with the re-balanced contact matrix. These models are introduced as follows, respectively.

5.2.1 The aggregate model

In the aggregate model, the contacts of the individuals between the infectious (including the persons in stocks of I, I_m and I_w) and the others (including the persons in the other stocks – S, R_1, R_2, R_3 and R_4) are mixed homogeneously. Based on the mathematical structure referred to Hethcote (1997) [50], the equations of the aggregate compartmental model of pertussis are listed as follows:

$$\begin{aligned}
\frac{dS}{dt} &= Nv - (\lambda + \mu)S + \iota R_1 \\
\frac{dI}{dt} &= \lambda S - (\gamma + \mu)\, I \\
\frac{dI_m}{dt} &= \lambda R_1 - (\gamma + \mu)\, I_m \\
\frac{dI_w}{dt} &= \lambda R_2 - (\gamma + \mu)\, I_w \\
\frac{dR_1}{dt} &= \alpha R_2 - (\lambda + \mu + \iota) R_1 \\
\frac{dR_2}{dt} &= \alpha R_3 - (\lambda + \mu + \alpha) R_2 \\
\frac{dR_3}{dt} &= \alpha R_4 - (\lambda + \mu + \alpha) R_3 \\
\frac{dR_4}{dt} &= \gamma(I + I_m + I_w) + \lambda R_3 - (\alpha + \mu) R_4 \\
\lambda &= \frac{\beta(I + \rho_m I_m + \rho_w I_w)}{N} \\
N &= S + I + I_m + I_w + R_1 + R_2 + R_3 + R_4
\end{aligned} \tag{5.1}$$

where N is the number of the total population; v is the birth rate of the total population, while μ is the death rate. It is notable that the value of death rate (μ) across all stocks are the same, to let the model to be simpler to be implemented, although the death rates in the stocks of the infectives (I, I_m, I_w) are theoretically a little higher than the other stocks, because of the extra deaths caused by the pertussis. The mean period of the immunity lose from the stock of R_1 to S is $\frac{1}{\iota}$; The mean period of the infectives (I, I_m, I_w) recovered is $\frac{1}{\gamma}$; the mean period of the temporary immunity for losing in the stocks of R_1, R_2 and R_3 is $\frac{1}{\alpha}$; ρ_m and ρ_w are the relative infectivities of the individuals in the mild-disease and weak-disease infectivity stocks I_m and I_w.

5.2.2 The age-structured model of 2 age groups

Similar to the two-age-groups age-structured model of measles introduced in chapter 4, where the total population is divided into two age groups – child and adult (denoted as "c" and "a"). The methods of

calculating the contact matrix and the death rates are also the same with the two-age-groups particle filtering model of measles. The age-structured model with two age groups of pertussis is listed as follows:

$$
\begin{aligned}
\begin{bmatrix} \frac{dS_c}{dt} \\ \frac{dS_a}{dt} \end{bmatrix} &= \begin{bmatrix} N_a v_a \\ 0 \end{bmatrix} + \begin{bmatrix} -\omega S_c \\ \omega S_c \end{bmatrix} - \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} S_c \\ S_a \end{bmatrix} - \begin{bmatrix} \mu_c S_c \\ \mu_a S_a \end{bmatrix} + \begin{bmatrix} \iota_c R_{1c} \\ \iota_a R_{1a} \end{bmatrix} \\
\begin{bmatrix} \frac{dI_c}{dt} \\ \frac{dI_a}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega I_c \\ \omega I_c \end{bmatrix} + \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} S_c \\ S_a \end{bmatrix} - \gamma \begin{bmatrix} I_c \\ I_a \end{bmatrix} - \begin{bmatrix} \mu_c I_c \\ \mu_a I_a \end{bmatrix} \\
\begin{bmatrix} \frac{dI_{mc}}{dt} \\ \frac{dI_{ma}}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega I_{mc} \\ \omega I_{mc} \end{bmatrix} + \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{1c} \\ R_{1a} \end{bmatrix} - \gamma \begin{bmatrix} I_{mc} \\ I_{ma} \end{bmatrix} - \begin{bmatrix} \mu_c I_{mc} \\ \mu_a I_{ma} \end{bmatrix} \\
\begin{bmatrix} \frac{dI_{wc}}{dt} \\ \frac{dI_{wa}}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega I_{wc} \\ \omega I_{wc} \end{bmatrix} + \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} - \gamma \begin{bmatrix} I_{wc} \\ I_{wa} \end{bmatrix} - \begin{bmatrix} \mu_c I_{wc} \\ \mu_a I_{wa} \end{bmatrix} \\
\begin{bmatrix} \frac{dR_{1c}}{dt} \\ \frac{dR_{1a}}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega R_{1c} \\ \omega R_{1c} \end{bmatrix} + \alpha \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} - \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{1c} \\ R_{1a} \end{bmatrix} - \iota \begin{bmatrix} R_{1c} \\ R_{1a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{1c} \\ \mu_a R_{1a} \end{bmatrix} \\
\begin{bmatrix} \frac{dR_{2c}}{dt} \\ \frac{dR_{2a}}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega R_{2c} \\ \omega R_{2c} \end{bmatrix} + \alpha \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} - \iota \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{2c} \\ \mu_a R_{2a} \end{bmatrix} \\
\begin{bmatrix} \frac{dR_{3c}}{dt} \\ \frac{dR_{3a}}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega R_{3c} \\ \omega R_{3c} \end{bmatrix} + \alpha \begin{bmatrix} R_{4c} \\ R_{4a} \end{bmatrix} - \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \iota \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{3c} \\ \mu_a R_{3a} \end{bmatrix} \\
\begin{bmatrix} \frac{dR_{4c}}{dt} \\ \frac{dR_{4a}}{dt} \end{bmatrix} &= \begin{bmatrix} -\omega R_{4c} \\ \omega R_{4c} \end{bmatrix} + \gamma \begin{bmatrix} I_c + I_{mc} + I_{wc} \\ I_a + I_{ma} + I_{wa} \end{bmatrix} + \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \alpha \begin{bmatrix} R_{4c} \\ R_{4a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{4c} \\ \mu_a R_{4a} \end{bmatrix} \\
\begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} &= \begin{bmatrix} \beta_c f_{cc} & \beta_c f_{ca} \\ \beta_a f_{ac} & \beta_a f_{aa} \end{bmatrix} \times \begin{bmatrix} \frac{I_c + \rho_m I_{mc} + \rho_w I_{wc}}{N_c} \\ \frac{I_a + \rho_m I_{ma} + \rho_w I_{wa}}{N_a} \end{bmatrix} \\
\begin{bmatrix} N_c \\ N_a \end{bmatrix} &= \begin{bmatrix} S_c \\ S_a \end{bmatrix} + \begin{bmatrix} I_c \\ I_a \end{bmatrix} + \begin{bmatrix} I_{mc} \\ I_{ma} \end{bmatrix} + \begin{bmatrix} I_{wc} \\ I_{wa} \end{bmatrix} + \begin{bmatrix} R_{1c} \\ R_{1a} \end{bmatrix} + \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} + \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} + \begin{bmatrix} R_{4c} \\ R_{4a} \end{bmatrix} \\
\mu_c &= \frac{N_a}{N_c} v_a - \omega \\
\mu_a &= \frac{N_c}{N_a} \omega
\end{aligned}
\tag{5.2}
$$

The meaning of the parameters in this model is the same as the ones in the pertussis aggregate model – Equations (5.1) and in the two-age-groups particle filtering model of measles in chapter 4.

It is notable that according to Equation (2.36), the more number of age groups divided in the age-structured model, the more accurate of the estimation of the value of the aging parameter c_i, which is denoted as ω in this model – Equations (5.2). Another advantage of considering more age groups in the age-structured model is that the model could capture the heterogeneous mixing between the individuals in the infective stocks (I, I_m and I_w) and the individuals in the other stocks (S, R_1, R_2, R_3 and R_4) more precisely. Thus, I have explored other two pertussis age-structured models with 32 age groups.

The key part needs to be studied in expanding the age-structured model from 2 age groups to 32 age groups is the calculation of the contact matrix. The advantage of the method in calculating the contact matrix in the age-structured model with 2 age groups in this thesis is that the contacts between any two age groups (e.g. i and j) is balanced – the number of total contacts of an age group i to group j equals to the

number of total contacts of the age group j to group i. However, this method has a notable disadvantage that the unknown parameters in calculating the contact matrix grow super-linearly (n-squared) with the total number of age groups (denoted as n) in the model. This disadvantage makes the model difficult to be expanded to a large number of age groups. Thus I have explored two other methods to calculate the contact matrix sub-linearly or linearly with the total number of age groups – the method of obtaining an un-balanced contact matrix in the research of Hethcote (1997) with a constant number of unknown parameters [50] and the method could calculate a re-balanced contact matrix with the number of unknown parameters grows linearly with the total number of age groups.

The prove of the super-linearly grows of the total numbers of unknown parameters in the contact matrix with the total number of age groups is shown in Appendix A.

5.2.3 The general age-structured model with more than two age groups

The basic mathematical model of the general age-structured model with more than two age groups is listed as follows. It is similar to Equation (5.2). However, the following equations are written in multiple equations, instead of matrix. It is notable that in this thesis, I considered the total population of each age group remains stable (constant). Thus, the death rate of each age group μ_i could be calculated by Equation (2.45), which is introduced in chapter 2. And the calculation of the parameters c_i are also introduced in Equation (2.36) in chapter 2.

$$
\begin{aligned}
\frac{dS_1}{dt} &= \sum_{j=1}^{n} v_j N_j + \iota R_{11} - (c_1 + \lambda_1 + \mu_1) S_1 \\
\frac{dS_i}{dt} &= c_{i-1} S_{i-1} + \iota R_{1i} - (c_i + \lambda_i + \mu_i) S_i \quad 2 \leq i \leq n \\
\frac{dI_1}{dt} &= \lambda_1 S_1 - (c_1 + \gamma + \mu_1) I_1 \\
\frac{dI_i}{dt} &= c_{i-1} I_{i-1} + \lambda_i S_i - (c_i + \gamma + \mu_i) I_i \quad 2 \leq i \leq n \\
\frac{dI_{m1}}{dt} &= \lambda_1 R_{11} - (c_1 + \gamma + \mu_1) I_{m1} \\
\frac{dI_{mi}}{dt} &= c_{i-1} I_{m,i-1} + \lambda_i R_{1i} - (c_i + \gamma + \mu_i) I_{mi} \quad 2 \leq i \leq n \\
\frac{dI_{w1}}{dt} &= \lambda_1 R_{21} - (c_1 + \gamma + \mu_1) I_{w1} \\
\frac{dI_{wi}}{dt} &= c_{i-1} I_{w,i-1} + \lambda_i R_{2i} - (c_i + \gamma + \mu_i) I_{wi} \quad 2 \leq i \leq n \\
\frac{dR_{11}}{dt} &= \alpha R_{21} - (\lambda_1 + \iota + c_1 + \mu_1) R_{11} \\
\frac{dR_{1i}}{dt} &= c_{i-1} R_{1,i-1} + \alpha R_{2i} - (\lambda_i + \iota + c_i + \mu_i) R_{1i} \quad 2 \leq i \leq n \\
\frac{dR_{21}}{dt} &= \alpha R_{31} - (\lambda_1 + \iota + c_1 + \mu_1) R_{21} \\
\frac{dR_{2i}}{dt} &= c_{i-1} R_{2,i-1} + \alpha R_{3i} - (\lambda_i + \iota + c_i + \mu_i) R_{2i} \quad 2 \leq i \leq n \\
\frac{dR_{31}}{dt} &= \alpha R_{41} - (\lambda_1 + \iota + c_1 + \mu_1) R_{31}
\end{aligned}
\tag{5.3}
$$

$$\frac{dR_{3i}}{dt} = c_{i-1}R_{3,i-1} + \alpha R_{4i} - (\lambda_i + \iota + c_i + \mu_i)R_{3i} \quad 2 \leq i \leq n$$
$$\frac{dR_{41}}{dt} = \gamma(I_1 + I_{m1} + I_{w1}) + \lambda_1 R_{31} - (\alpha + c_1 + \mu_1)R_{41}$$
$$\frac{dR_{4i}}{dt} = c_{i-1}R_{4,i-1} + \gamma(I_i + I_{mi} + I_{wi}) + \lambda_i R_{3i} - (\alpha + c_i + \mu_i)R_{4i} \quad 2 \leq i \leq n$$
$$N_i = S_i + I_i + I_{mi} + I_{wi} + R_{1i} + R_{2i} + R_{3i} + R_{4i} \quad 1 \leq i \leq n$$
$$\mu_1 = \frac{\sum_{j=1}^{n} v_j N_j - c_1 N_1}{N_1}$$
$$\mu_i = \frac{c_{i-1}N_{i-1} - c_i N_i}{N_i} \quad 2 \leq i \leq n$$

It is notable that the index of the age group in the model is denoted by i. Two methods of calculating the contact matrix with the number of unknown parameters grow less than or equal linearly with the number of the total age groups are introduced as follows.

The un-balanced contact matrix

This method is introduced in the research of Hethcote (1997) [50]. In this method, there are no unknown parameters. Suppose l_k is the contact rate – the average number of persons contacted by a person of age group k per unit time. And only adequate contacts are sufficient to transmit the disease. This method based on a simple proportionate mixing assumption that the number of total persons contacted by one person in the age group k (l_k) is distributed among the population in the age group j in proportion to the fractions l_j/D of all contacts per unit time received by people in the age group j, where $D = \sum_{k=1}^{n} l_k N_k / \sum_{k=1}^{n} N_k$. Thus, the element in the contact matrix is $l_j l_k / D$.

Finally, the force of infection (λ) used in Equation (5.3) in [50] by re-dimensionalization could be calculated as follows:

$$\lambda_j = \sum_{k=1}^{n} \frac{l_k l_j}{D} \frac{I_k + \rho_m I_{mk} + \rho_w I_{wk}}{\sum_{k=1}^{n} N_k} \tag{5.4}$$

The advantage of this method is that there are no unknown parameters in calculating the contact matrix. And it is straightforward in calculating the contact matrix as long as the parameters of the contact rate of each age groups are known. However, this method of calculating the contact has a notable disadvantage, which is the value of the total contacts of the age group i to the age group j is not equal to the value of the total contacts of the age group j to the age group i.

The unbalance of the contact matrix in the paper of Hethcote (1997) [50] could be proved. The number of total contacts of the age group j to age group k is $N_j l_j l_k / D$, while the number of total contacts of the age group k to age group j is $N_k l_j l_k / D$. Generally, the value of $N_j l_j l_k / D$ is not equal to the value of $N_k l_j l_k / D$. Thus this contact matrix is unbalanced.

Then, I explored another method to calculate the contact matrix with the balanced feature and the total number of the unknown parameters grows linearly with the number of the age groups.

The re-balanced contact matrix

To calculate the balanced contact matrix, I have employed the method introduced in the research by Garnett and Bowden (2000) [39].

In the beginning, I introduce the method of calculating the basic contact matrix which is balanced already and with one unknown parameters. Before introduced, I import a mixing parameter, denoted as ϵ. The mixing parameter ϵ determines where mixing occurs on a scale from fully associative – persons only contact with the individuals in the same age group (e.g. $\epsilon = 0$) and random mixing – the contact among the total population is homogeneous (e.g. $\epsilon = 1.0$). Then, the fraction of the average persons that an individual in age group i that contact with the persons in the age group of j, which is the parameters of f_{ij} in the contact matrix are calculated as follows:

$$f_{ij} = (1.0 - \epsilon)\delta_{ij} + \epsilon \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j} \right) \tag{5.5}$$

where δ_{ij} is the identity matrix. And the elements in the contact matrix is $l_i f_{ij}$.

The total contacts of age group i to age group j ($N_i l_i f_{ij}$) equal the total contacts of age group j to age group i ($N_j l_j f_{ji}$), in this basic contact matrix. And the only unknown parameter is ϵ. However, in general, the mixing parameter related to each age group should be different. For example, the mixing parameter of ϵ of young children in school age maybe lower than the ϵ of the little baby, because the children in the school age contacts more to their peers in the school than the other groups, while the little baby contacts more with their parents or care-taker than the other babies. Thus, in the next step, I expend the mixing parameter ϵ to a vector, where each element represents the mixing parameter of each related age group ϵ_i.

Then, in this method of calculating the contact matrix with a vector of mixing parameters, the equation of f_{ij} is listed as follows:

$$f_{ij} = (1.0 - \epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j} \right) \tag{5.6}$$

Similarly, the elements in this contact matrix with a vector of mixing parameters are $l_i f_{ij}$. It is notable that the total contacts between any two age groups calculated based on this contact matrix are unbalanced. Specifically, the number of total contacts of age group i to age group j is $N_i l_i \left[(1.0 - \epsilon)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j} \right) \right]$, while the number of the total contact of age group j to age group i is $N_j l_j \left[(1.0 - \epsilon)\delta_{ji} + \epsilon_j \left(\frac{N_i l_i}{\sum_{j=1}^{n} N_j l_j} \right) \right]$. In general, the mixing parameters of any two age groups are not the same. Thus, the total numbers of contacts calculated by this contact matrix between any two age groups are not always the same.

To make the contact matrix balanced, I have employed the method introduced in [39] to re-balance the contact matrix. A parameter, denoted as Δ_{ij}, is imported to represent the ratio of the number of total contacts between any two different age groups ($i \neq j$) (for the same age group, the total number of contacts are always the same). Then, the equation of Δ_{ij} is:

$$\Delta_{ij} = \frac{N_i l_i f_{ij}}{N_j l_j f_{ji}} = \frac{\epsilon_i N_i l_i N_j l_j}{\epsilon_j N_i l_i N_j l_j} = \frac{\epsilon_i}{\epsilon_j}, \quad i \neq j \tag{5.7}$$

Then, the main idea of re-balancing the contact matrix is to extend the vector of contact rates (the elements of the contact rates are denoted as l_i) to a new matrix of contact rates l_{ij}. The elements in the matrix of contact rates l_{ij} represent the number of persons in the age groups j that a person in the age group i could contact in average. Then, according to [39], the equations of l_{ij} and l_{ji} could be defined separately:

$$\begin{aligned} l_{ij} &= l_i \Delta_{ij}^{\theta} = l_i \left(\frac{\epsilon_j}{\epsilon_i}\right)^{\theta} \\ l_{ji} &= l_j \Delta_{ij}^{-(1-\theta)} = l_j \left(\frac{\epsilon_j}{\epsilon_i}\right)^{-(1-\theta)} \end{aligned} \tag{5.8}$$

where θ is the re-balanced parameter.

Because both l_{ij} and l_{ji} represent the same matrix, a relationship could be generated, which is $l_{ij} = l_{ji}$. Then, I could get the value of the parameter of θ ($\theta = 0.5$). Substitute the value of θ ($\theta = 0.5$) to Equation (5.8), the matrix of contact rate – l_{ij} could be generated as follows:

$$l_{ij} = l_i \left(\frac{\epsilon_j}{\epsilon_i}\right)^{0.5} \tag{5.9}$$

Finally, the element of contact matrix $l_{ij} f_{ij}$ and force of infection λ_i are:

$$\begin{aligned} l_{ij} f_{ij} &= l_i \left(\frac{\epsilon_j}{\epsilon_i}\right)^{0.5} \left[(1.0 - \epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j}\right)\right] \\ \lambda_i &= p_i \sum_{j=1}^{n} \frac{l_{ij} f_{ij} (I_j + \rho_m I_{mj} + \rho_w I_{wj})}{N_j} \end{aligned} \tag{5.10}$$

where p_i is the transmission probability of age group i.

Finally, except the aggregate and age–structured model with two age groups, another two age-structured models with 32 age groups – the model introduced in the [50] and the model with re-balanced contact matrix are also employed as the base model to build the state-space models of the particle filtering algorithm. It is notable that the division of the 32 age groups is the same in the paper of Hethcote (1997) [50]. These age groups are 0–1 month, 2–3 month, 4–5 month, 6–11 month, 1 year, 2 year, 3 year, 4 year, 5 year, 6 year, 7 year, 8 year, 9 year, 10 year, 11 year, 12 year, 13 year, 14 year, 15 year, 16 year, 17 year, 18 year, 19 year, 20–24 year, 25–29 year, 30–39 year, 40–49 year, 50–59 year, 60–69 year, 70–79 year, 80–89 year, 90 year and plus. And the two age groups in the age structured model with two age groups are 0–4 year and 5 years and plus.

5.3 Particle filter implementation

The implementation of the particle filtering models of pertussis is similar to the implementation of the particle filtering models of measles, which have been introduced in the previous chapters, except for the difference of the state space models and the total number of the age groups. Specifically, the condensation algorithm is employed as the proposal distribution. The distribution employed in the likelihood function is the negative

binomial distribution, with the dispersion parameter r is 10. The measurement model is also similar to the particle filtering models of measles. However, in the age-structured models with 32 age groups of pertussis, the yearly empirical data are divided into 6 age groups, to training the particle filtering models. This fine division can influence the implementation of the likelihood function and calculation of discrepancy. Finally, I have built four pertussis particle filtering models, which are the aggregate model, age-structured model of 2 age groups, age-structured models of 32 age groups with the calculation method of contact matrix introduced in the paper of Hethcote (1997) [50], and age-structured model of 32 age groups with the re-balanced contact matrix – introduced in the paper of Garnett and Bowden (2000) [39]. I will discuss the implementation of the particle filtering algorithm with pertussis models as follows.

5.3.1 The state space model

The method of implementation of the state space models of pertussis is similar to the particle filtering of measles – considering several dynamic processes in the state space model, based on the compartmental models – pure Ordinary Differential Equations (ODEs). The state space models employed in this thesis to build the pertussis particle filtering models are discussed as follows.

The aggregate model ($n = 1$)

In the aggregate model, the individuals contact with the infectious (in the stocks of I_w, I_m and I) homogeneously. The pure ODEs model is introduced in Equations (5.1). Similar to the implementation of the aggregate particle filtering model of measles in chapter 3, three stochastic processes are considered. The first is the changes in the transmissible contact rate linking infectious and susceptible persons, which is represented by the parameter β. The second is also with respect to the disease reporting process. Specifically, a parameter – representing the probability that a given pertussis infectious case is reported C_r, and a state I_k – calculating the accumulative pertussis infectious cases per unit time (per Month in this project) – are implemented. The final part is the Poisson process associated with the incidence of infection. This process reflects the small number of cases that occur over each small unit of time – Δt (0.01 in this model). I also treat the natural logarithm of the transmissible contact rate (denoted by β) and the logit of C_r as undergoing a random walk according to a Wiener Process (Brownian Motion) [56, 28]. It is notable that in this project, it is assumed that the individuals under the medium infectious (I_m) and weak infectious (I_w) also have the probability to be confirmed and reported. The rates of the medium infectious (I_m) and weak infectious (I_w) that have symptoms are also considered as ρ_m and ρ_w. Finally, the state space model of the aggregate pertussis particle filtering model is listed as follows:

$$\begin{aligned}
\frac{dS}{dt} &= Nv - A_I - \mu S + \iota R_1 \\
\frac{dI}{dt} &= A_I - (\gamma + \mu) I \\
\frac{dI_m}{dt} &= A_{I_m} - (\gamma + \mu) I_m
\end{aligned}$$

$$
\begin{aligned}
&\frac{dI_w}{dt} = A_{I_w} - (\gamma + \mu)\, I_w \\
&\frac{dR_1}{dt} = \alpha R_2 - A_{I_m} - (\mu + \iota) R_1 \\
&\frac{dR_2}{dt} = \alpha R_3 - A_{I_w} - (\mu + \alpha) R_2 \\
&\frac{dR_3}{dt} = \alpha R_4 - (\lambda + \mu + \alpha) R_3 \\
&\frac{dR_4}{dt} = \gamma(I + I_m + I_w) + \lambda R_3 - (\alpha + \mu) R_4 \\
&\lambda = \frac{\beta(I + \rho_m I_m + \rho_w I_w)}{N} \\
&N = S + I + I_m + I_w + R_1 + R_2 + R_3 + R_4 \\
&dln(\beta) = s_\beta dW_t \\
&d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t \\
&I_k = \int_{k-1}^{k} (A_I + \rho_m A_{I_m} + \rho_w A_{I_w})\, dt \\
&I_{rk} = I_k C_r \\
&A_I = \frac{Poisson(\lambda S \Delta t)}{\Delta t} \\
&A_{I_m} = \frac{Poisson(\lambda R_1 \Delta t)}{\Delta t} \\
&A_{I_w} = \frac{Poisson(\lambda R_2 \Delta t)}{\Delta t}
\end{aligned}
\tag{5.11}
$$

The parameters related to the transmission of pertussis in this model are referred from the research of Hethcote (1997) [50]. Similar to the aggregate model of measles introduced in chapter 3, the demographic parameters of this model are got from the Annual Report of the Saskatchewan Department of Public Health [21] and the age pyramid of Saskatchewan [16]. Then, the parameters of the pertussis aggregate state space model – Equations (5.11) are specified in Table 5.1, while the initial values of the stocks are listed in Table 5.2.

The age-structured model of 2 age groups ($n = 2$)

Equations (5.2) is employed as the base model of the state space model of the age-structured model with 2 age groups. Then, the pure ODEs model – Equations (5.2) is extended by several stochastic processes. Except for the similar three stochastic processes considered in the aggregate state space model – the infectious contact rate of the child age group (denoted as β_c), the report rate of pertussis cases (denoted as C_r), and the Poisson process related to the incidence of the infectious – two other stochastic processed are also considered. These two stochastic processes are related to the parameter of the multiplier of the adult age group model (M_a) of the infectious contact rate and the fraction of children's infectious contacts that occur with other children (f_{cc}). Specifically, the natural logarithm of the multiplier of the infectious contact rate of the adult age group (M_a) and the logit of f_{cc} are treated as undergoing a random walk according to a Wiener Process (Brownian Motion) [56, 28], similar to the age-structured model of 2 age groups of measles introduced in

Table 5.1: Table showing the value of parameters in the pertussis aggregate particle filtering model.

Parameter	Description	Value	Units
γ^{-1}	mean time for infectives to recover from pertussis	21	Day
v	birth rate of the total population	0.03	1/Year
μ	death rate of the total population	0.03	1/Year
N	total population	863,545	Person
ι^{-1}	mean time to lose immunity from the stock of R_1 to S	10	Year
α^{-1}	mean time to lose immunity from R_i down to R_{i-1}	5	Year
ρ_m	the relative infectivities of the individuals in the stock I_m	0.5	Dimensionless
ρ_w	the relative infectivities of the individuals in the stock I_w	0.25	Dimensionless
s_β	the diffusion parameter of $ln(\beta)$	0.5	Dimensionless
s_r	the diffusion parameter of $ln(\frac{C_r}{1-C_r})$	0.05	Dimensionless

Table 5.2: Table showing initial values of the stocks in the pertussis aggregate particle filtering model.

Parameter	Value	Unit
S_0	Uniform[5000, 30000)	Person
I_0	Uniform[500, 5000)	Person
I_{m0}	1000	Person
I_{w0}	2500	Person
R_{10}	Uniform[10, 10000)	Person
R_{20}	10000	Person
R_{30}	20000	Person
R_{40}	$N - S_0 - I_0 - I_{m0} - I_{w0} - R_{10} - R_{20} - R_{30}$	Person
β	Uniform[5,100)	Person/Month
C_r	Uniform[0,0.2)	Dimensionless

chapter 4. Finally, the state space model of the pertussis age-structured model of 2 age groups is listed as follows:

$$
\begin{aligned}
&\begin{bmatrix} \frac{dS_c}{dt} \\ \frac{dS_a}{dt} \end{bmatrix} = \begin{bmatrix} N_a v_a \\ 0 \end{bmatrix} + \begin{bmatrix} -\omega S_c \\ \omega S_c \end{bmatrix} - \begin{bmatrix} A_{I_c} \\ A_{I_a} \end{bmatrix} - \begin{bmatrix} \mu_c S_c \\ \mu_a S_a \end{bmatrix} + \begin{bmatrix} \iota_c R_{1c} \\ \iota_a R_{1a} \end{bmatrix} \\
&\begin{bmatrix} \frac{dI_c}{dt} \\ \frac{dI_a}{dt} \end{bmatrix} = \begin{bmatrix} -\omega I_c \\ \omega I_c \end{bmatrix} + \begin{bmatrix} A_{I_c} \\ A_{I_a} \end{bmatrix} - \gamma \begin{bmatrix} I_c \\ I_a \end{bmatrix} - \begin{bmatrix} \mu_c I_c \\ \mu_a I_a \end{bmatrix} \\
&\begin{bmatrix} \frac{dI_{mc}}{dt} \\ \frac{dI_{ma}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega I_{mc} \\ \omega I_{mc} \end{bmatrix} + \begin{bmatrix} A_{I_{mc}} \\ A_{I_{ma}} \end{bmatrix} - \gamma \begin{bmatrix} I_{mc} \\ I_{ma} \end{bmatrix} - \begin{bmatrix} \mu_c I_{mc} \\ \mu_a I_{ma} \end{bmatrix} \\
&\begin{bmatrix} \frac{dI_{wc}}{dt} \\ \frac{dI_{wa}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega I_{wc} \\ \omega I_{wc} \end{bmatrix} + \begin{bmatrix} A_{I_{wc}} \\ A_{I_{wa}} \end{bmatrix} - \gamma \begin{bmatrix} I_{wc} \\ I_{wa} \end{bmatrix} - \begin{bmatrix} \mu_c I_{wc} \\ \mu_a I_{wa} \end{bmatrix} \\
&\begin{bmatrix} \frac{dR_{1c}}{dt} \\ \frac{dR_{1a}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega R_{1c} \\ \omega R_{1c} \end{bmatrix} + \alpha \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} - \begin{bmatrix} A_{I_{mc}} \\ A_{I_{ma}} \end{bmatrix} - \iota \begin{bmatrix} R_{1c} \\ R_{1a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{1c} \\ \mu_a R_{1a} \end{bmatrix} \\
&\begin{bmatrix} \frac{dR_{2c}}{dt} \\ \frac{dR_{2a}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega R_{2c} \\ \omega R_{2c} \end{bmatrix} + \alpha \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \begin{bmatrix} A_{I_{wc}} \\ A_{I_{wa}} \end{bmatrix} - \iota \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{2c} \\ \mu_a R_{2c} \end{bmatrix} \\
&\begin{bmatrix} \frac{dR_{3c}}{dt} \\ \frac{dR_{3a}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega R_{3c} \\ \omega R_{3c} \end{bmatrix} + \alpha \begin{bmatrix} R_{4c} \\ R_{4a} \end{bmatrix} - \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \iota \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{3c} \\ \mu_a R_{3a} \end{bmatrix} \\
&\begin{bmatrix} \frac{dR_{4c}}{dt} \\ \frac{dR_{4a}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega R_{4c} \\ \omega R_{4c} \end{bmatrix} + \gamma \begin{bmatrix} I_c + I_{mc} + I_{wc} \\ I_a + I_{ma} + I_{wa} \end{bmatrix} + \begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} \circ \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} - \alpha \begin{bmatrix} R_{4c} \\ R_{4a} \end{bmatrix} - \begin{bmatrix} \mu_c R_{4c} \\ \mu_a R_{4a} \end{bmatrix} \\
&\begin{bmatrix} \lambda_c \\ \lambda_a \end{bmatrix} = \begin{bmatrix} \beta_c f_{cc} & \beta_c f_{ca} \\ \beta_a f_{ac} & \beta_a f_{aa} \end{bmatrix} \times \begin{bmatrix} \frac{I_c + \rho_m I_{mc} + \rho_w I_{wc}}{N_c} \\ \frac{I_a + \rho_m I_{ma} + \rho_w I_{wa}}{N_a} \end{bmatrix} \\
&\begin{bmatrix} N_c \\ N_a \end{bmatrix} = \begin{bmatrix} S_c \\ S_a \end{bmatrix} + \begin{bmatrix} I_c \\ I_a \end{bmatrix} + \begin{bmatrix} I_{mc} \\ I_{ma} \end{bmatrix} + \begin{bmatrix} I_{wc} \\ I_{wa} \end{bmatrix} + \begin{bmatrix} R_{1c} \\ R_{1a} \end{bmatrix} + \begin{bmatrix} R_{2c} \\ R_{2a} \end{bmatrix} + \begin{bmatrix} R_{3c} \\ R_{3a} \end{bmatrix} + \begin{bmatrix} R_{4c} \\ R_{4a} \end{bmatrix} \qquad (5.12) \\
&d(ln\beta_c) = s_{\beta_c} dW_t \\
&d(ln(\frac{f_{cc}}{1-f_{cc}})) = s_{cc} dW_t \\
&d(lnM_a) = s_{M_a} dW_t \\
&\beta_a = M_a \beta_c \\
&d(ln(\frac{C_r}{1-C_r})) = s_r dW_t \\
&f_{ca} = 1 - f_{cc} \\
&f_{ac} = \begin{cases} \frac{N_c \beta_c}{N_a \beta_a}(1 - f_{cc}), & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a}(1 - f_{cc})\right] < 1.0 \\ 1.0, & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a}(1 - f_{cc})\right] \geq 1.0 \end{cases} \\
&f_{aa} = 1 - f_{ac} \\
&\mu_c = \frac{N_a}{N_c} v_a - \omega \\
&\mu_a = \frac{N_c}{N_a} \omega
\end{aligned}
$$

$$I_{ck} = \int_{k-1}^{k} (A_{I_c} + \rho_m A_{I_{mc}} + \rho_w A_{I_{wc}})\, dt$$
$$I_{ak} = \int_{k-1}^{k} (A_{I_a} + \rho_m A_{I_{ma}} + \rho_w A_{I_{wa}})\, dt$$
$$\begin{bmatrix} I_{rck} \\ I_{rak} \end{bmatrix} = C_r \begin{bmatrix} I_{ck} \\ I_{ak} \end{bmatrix}$$
$$A_{I_c} = \frac{Poisson(\lambda_c S_c \Delta t)}{\Delta t}$$
$$A_{I_a} = \frac{Poisson(\lambda_a S_a \Delta t)}{\Delta t}$$
$$A_{I_{mc}} = \frac{Poisson(\lambda_c R_{1c} \Delta t)}{\Delta t}$$
$$A_{I_{ma}} = \frac{Poisson(\lambda_a R_{1a} \Delta t)}{\Delta t}$$
$$A_{I_{wc}} = \frac{Poisson(\lambda_c R_{2c} \Delta t)}{\Delta t}$$
$$A_{I_{wa}} = \frac{Poisson(\lambda_a R_{2a} \Delta t)}{\Delta t}$$

In this project, I have built one two-age-group particle filtering model, where the individuals in the age group of "child" are from newborn up to the end of 4 years. The parameters with constant values related to the pure compartmental model (γ, ι, α, ρ_m and ρ_2) in the two-age-group pertussis model are the same as the aggregate model. All these parameters and the parameters related to the demographic model and the stochastic processes of the two-age-group particle filtering model are listed in Table 5.3. The initial values of each stocks in this two-age-group particle filtering model are listed in Table B.1 of Appendix B.

Table 5.3: Table showing the value of parameters (only related to the demographic model and stochastic processes) in pertussis two-age-groups particle filtering model.

Parameter	**Description**	**Value**	**Units**
v_a	birth rate of the adult age group	0.034	1/Year
N_c	the population of the child age group	98743	Person
N_a	the population of the adult age group	764802	Person
ω	the aging rate from child to adult age group	0.2	Dimensionless
s_β	the diffusion parameter of $ln(\beta)$	0.5	Dimensionless
s_r	the diffusion parameter of $ln(\frac{C_r}{1-C_r})$	0.05	Dimensionless
s_{M_a}	the diffusion parameter of $ln(M_a)$	0.2	Dimensionless
s_{cc}	the diffusion parameter of $ln(\frac{f_{cc}}{1-f_{cc}})$	0.15	Dimensionless

The age-structured model of 32 age groups ($n = 32$) with the Hethcote contact matrix

I employ the pure ODEs model – the age-structured model of 32 age groups introduced in the paper of Hethcote (1997) [50] as the base model. This mathematical model is also introduced in the previous part

in this chapter. Similarly, three stochastic processes are added to the base model as the state space model. These three stochastic processes are related to the Poisson process related to the incidence of infectious, the contact rate of the first age group and the reporting process of the pertussis cases. Similarly, the natural logarithm of the contact rate of the first age group (denoted as l_1) and the logit of the report rate (denoted as C_r) are treated as undergoing a random walk according to a Wiener Process (Brownian Motion) [56, 28]. Then, a vector represents the fraction of the contact rate of each age group compared with the first age group is imported to calculate the contact rate of each age group (denoted as f_{l_i}). f_{l_i} is calculated from the value of contact rate of all age groups in the paper of Hethcote (1997) [50]. The value of f_{l_i} is (1, 6.03, 8.03, 10.03, 12.04, 15.06, 20.08, 28.10, 47.18, 47.18, 47.18, 47.18, 47.18, 25.09, 25.09, 25.09, 25.09, 25.09, 15.06, 15.06, 15.06, 15.06, 15.06, 15.06, 15.06, 15.06, 10.03, 10.03, 5.02, 5.02, 5.02, 5.02). Moreover, another vector p_i is imported to represent the change rate of the transmission probability of pertussis of each age group in the state space model. The parameter of transmission probability is not used in the original mathematical model in [50]. However, normally the transmission probability is different among different age groups. For example, the transmission probability of the young children is usually higher than the adults, such as due to hygienic disparities. Thus, the parameters of p_i are imported to represent the difference of the transmission probability of each age group. The value of p_i is (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0.5, 0.5, 0.5, 0.5, 0.5, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.05) in the 32-age-group models of pertussis in this project. And the population in each age group, referred from the age pyramid of Saskatchewan [16] is (3349, 3330, 3320, 9950, 19843, 19733, 19647, 19571, 19486, 19394, 19289, 19161, 19002, 18809, 18577, 18318, 18033, 17724, 17386, 17021, 16629, 16218, 15802, 73256, 65935, 117771, 97621, 70964, 44313, 19332, 4377, 387). To let the numbers of newborns of each pertussis particle filtering models per unit time (Month in this project) similar to the other models – the pertussis particle filtering models with aggregate population structure and with two-age-groups population structure, the yearly birth rate of the 32-age-group models are estimated as (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.03, 0.03, 0.03, 0.03, 0.03, 0.1, 0.1, 0.1, 0, 0, 0, 0, 0, 0). Finally, the state space model of the the age-structured pertussis particle filtering model of 32 age groups with contact matrix in the paper of Hethcote (1997) [50] is listed as follows:

$$
\begin{aligned}
\frac{dS_1}{dt} &= \sum_{j=1}^{n} v_j N_j + \iota R_{11} - A_{I_1} - (c_1 + \mu_1) S_1 \\
\frac{dS_i}{dt} &= c_{i-1} S_{i-1} + \iota R_{1i} - A_{I_i} - (c_i + \mu_i) S_i \quad 2 \le i \le n \\
\frac{dI_1}{dt} &= A_{I_1} - (c_1 + \gamma + \mu_1) I_1 \\
\frac{dI_i}{dt} &= c_{i-1} I_{i-1} + A_{I_i} - (c_i + \gamma + \mu_i) I_i \quad 2 \le i \le n \\
\frac{dI_{m1}}{dt} &= A_{I_{m1}} - (c_1 + \gamma + \mu_1) I_{m1} \\
\frac{dI_{mi}}{dt} &= c_{i-1} I_{m,i-1} + A_{I_{mi}} - (c_i + \gamma + \mu_i) I_{mi} \quad 2 \le i \le n
\end{aligned}
$$

$$
\begin{aligned}
&\frac{dI_{w1}}{dt} = A_{I_{w1}} - (c_1 + \gamma + \mu_1) I_{w1} \\
&\frac{dI_{wi}}{dt} = c_{i-1} I_{w,i-1} + A_{I_{wi}} - (c_i + \gamma + \mu_i) I_{wi} \quad 2 \le i \le n \\
&\frac{dR_{11}}{dt} = \alpha R_{21} - A_{I_{m1}} - (\iota + c_1 + \mu_1) R_{11} \\
&\frac{dR_{1i}}{dt} = c_{i-1} R_{1,i-1} + \alpha R_{2i} - A_{I_{mi}} - (\iota + c_i + \mu_i) R_{1i} \quad 2 \le i \le n \\
&\frac{dR_{21}}{dt} = \alpha R_{31} - A_{I_{w1}} - (\iota + c_1 + \mu_1) R_{21} \\
&\frac{dR_{2i}}{dt} = c_{i-1} R_{2,i-1} + \alpha R_{3i} - A_{I_{wi}} - (\iota + c_i + \mu_i) R_{2i} \quad 2 \le i \le n \\
&\frac{dR_{31}}{dt} = \alpha R_{41} - (\lambda_1 + \iota + c_1 + \mu_1) R_{31} \\
&\frac{dR_{3i}}{dt} = c_{i-1} R_{3,i-1} + \alpha R_{4i} - (\lambda_i + \iota + c_i + \mu_i) R_{3i} \quad 2 \le i \le n \\
&\frac{dR_{41}}{dt} = \gamma (I_1 + I_{m1} + I_{w1}) + \lambda_1 R_{31} - (\alpha + c_1 + \mu_1) R_{41} \\
&\frac{dR_{4i}}{dt} = c_{i-1} R_{4,i-1} + \gamma (I_i + I_{mi} + I_{wi}) + \lambda_i R_{3i} - (\alpha + c_i + \mu_i) R_{4i} \quad 2 \le i \le n \qquad (5.13) \\
&N_i = S_i + I_i + I_{mi} + I_{wi} + R_{1i} + R_{2i} + R_{3i} + R_{4i} \quad 1 \le i \le n \\
&\mu_1 = \frac{\sum_{j=1}^{n} v_j N_j - c_1 N_1}{N_1} \\
&\mu_i = \frac{c_{i-1} N_{i-1} - c_i N_i}{N_i} \quad 2 \le i \le n \\
&\lambda_j = p_j \sum_{k=1}^{n} \frac{l_k l_j}{D} \frac{I_k + \rho_m I_{mk} + \rho_w I_{wk}}{\sum_{k=1}^{n} N_k} \quad 1 \le j \le n \\
&D = \sum_{k=1}^{n} l_k N_k / \sum_{k=1}^{n} N_k \\
&d(lnl_1) = s_{l_1} dW_t \\
&l_i = l_1 * f_{l_i} \quad 2 \le i \le n \\
&d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t \\
&I_{ki} = \int_{k-1}^{k} (A_{I_i} + \rho_m A_{I_{mi}} + \rho_w A_{I_{wi}})\, dt \quad 1 \le i \le n \\
&I_{rki} = I_{ki} C_r \quad 1 \le i \le n \\
&A_{I_i} = \frac{Poisson(\lambda_i S_i \Delta t)}{\Delta t} \quad 1 \le i \le n \\
&A_{I_{mi}} = \frac{Poisson(\lambda_i R_{1i} \Delta t)}{\Delta t} \quad 1 \le i \le n \\
&A_{I_{wi}} = \frac{Poisson(\lambda_i R_{2i} \Delta t)}{\Delta t} \quad 1 \le i \le n
\end{aligned}
$$

The values of the parameters are the same as the ones listed in the aggregate particle filtering models and two-age-group particle filtering model, and the initial values of the stocks in this particle filtering model are listed in the Table B.2 of Appendix B.

The age-structured model of 32 age groups ($n = 32$) with re-balanced contact matrix

The age-structured model (introduced in Equations (5.3)) of 32 age groups with re-balanced contact matrix (introduced in Equations (5.10)) are employed as the base model of the state space model of the age-structured particle filtering model of 32 age groups with re-balanced contact matrix. Except for the three stochastic processes – the same as the state space model of age-structured model of 32 age groups with Hethcote contact matrix (introduced in Equations (5.13)), the logit of the six mixing parameters ($\epsilon_i, 1 \leq i \leq 6$) are treated as undergoing a random walk according to a Wiener Process (Brownian Motion) [56, 28]. The reason the total number of mixing parameters is 6, instead of 32 (a mixing parameter related to an age group each), is because the yearly empirical datasets could only be split to 6 age groups – less than 1 year, 1 to 4 year, 5 to 9 year, 10 to 14 year, 15 to 19 year and 20 year and plus, which will be introduced in next part in detail. Moreover, two parameters of f_{l_i} and p_i are also employed, same as the model in Equation (5.13). Finally, the state space model of the the age structured pertussis particle filtering model of 32 age groups with re-balanced contact matrix is listed as follows:

$$
\begin{aligned}
\frac{dS_1}{dt} &= \sum_{j=1}^{n} v_j N_j + \iota R_{11} - A_{I_1} - (c_1 + \mu_1) S_1 \\
\frac{dS_i}{dt} &= c_{i-1} S_{i-1} + \iota R_{1i} - A_{I_i} - (c_i + \mu_i) S_i \quad 2 \leq i \leq n \\
\frac{dI_1}{dt} &= A_{I_1} - (c_1 + \gamma + \mu_1) I_1 \\
\frac{dI_i}{dt} &= c_{i-1} I_{i-1} + A_{I_i} - (c_i + \gamma + \mu_i) I_i \quad 2 \leq i \leq n \\
\frac{dI_{m1}}{dt} &= A_{I_{m1}} - (c_1 + \gamma + \mu_1) I_{m1} \\
\frac{dI_{mi}}{dt} &= c_{i-1} I_{m,i-1} + A_{I_{mi}} - (c_i + \gamma + \mu_i) I_{mi} \quad 2 \leq i \leq n \\
\frac{dI_{w1}}{dt} &= A_{I_{w1}} - (c_1 + \gamma + \mu_1) I_{w1} \\
\frac{dI_{wi}}{dt} &= c_{i-1} I_{w,i-1} + A_{I_{wi}} - (c_i + \gamma + \mu_i) I_{wi} \quad 2 \leq i \leq n \\
\frac{dR_{11}}{dt} &= \alpha R_{21} - A_{I_{m1}} - (\iota + c_1 + \mu_1) R_{11} \\
\frac{dR_{1i}}{dt} &= c_{i-1} R_{1,i-1} + \alpha R_{2i} - A_{I_{mi}} - (\iota + c_i + \mu_i) R_{1i} \quad 2 \leq i \leq n \\
\frac{dR_{21}}{dt} &= \alpha R_{31} - A_{I_{w1}} - (\iota + c_1 + \mu_1) R_{21} \\
\frac{dR_{2i}}{dt} &= c_{i-1} R_{2,i-1} + \alpha R_{3i} - A_{I_{wi}} - (\iota + c_i + \mu_i) R_{2i} \quad 2 \leq i \leq n \\
\frac{dR_{31}}{dt} &= \alpha R_{41} - (\lambda + \iota + c_1 + \mu_1) R_{31} \\
\frac{dR_{3i}}{dt} &= c_{i-1} R_{3,i-1} + \alpha R_{4i} - (\lambda + \iota + c_i + \mu_i) R_{3i} \quad 2 \leq i \leq n \\
\frac{dR_{41}}{dt} &= \gamma (I_1 + I_{m1} + I_{w1}) + \lambda R_{31} - (\alpha + c_1 + \mu_1) R_{41}
\end{aligned}
$$

$$\frac{dR_{4i}}{dt} = c_{i-1}R_{4,i-1} + \gamma(I_i + I_{mi} + I_{wi}) + \lambda R_{3i} - (\alpha + c_i + \mu_i)R_{4i} \quad 2 \leq i \leq n \tag{5.14}$$

$$N_i = S_i + I_i + I_{mi} + I_{wi} + R_{1i} + R_{2i} + R_{3i} + R_{4i} \quad 1 \leq i \leq n$$

$$\mu_1 = \frac{\sum_{j=1}^{n} v_j N_j - c_1 N_1}{N_1}$$

$$\mu_i = \frac{c_{i-1}N_{i-1} - c_i N_i}{N_i} \quad 2 \leq i \leq n$$

$$l_{ij}f_{ij} = l_i \left(\frac{\epsilon_j}{\epsilon_i}\right)^{0.5} \left[(1.0 - \epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j}\right)\right] \quad 1 \leq i \leq n, \quad 1 \leq j \leq n$$

$$\lambda_i = p_i \sum_{j=1}^{n} \frac{l_{ij}f_{ij}(I_j + \rho_m I_{mj} + \rho_w I_{wj})}{N_j} \quad 1 \leq i \leq n$$

$$d(lnl_1) = s_{l_1} dW_t$$

$$l_i = l_1 * f_{l_i} \quad 2 \leq i \leq n$$

$$d(logit(\epsilon_i)) = d(ln(\frac{\epsilon_i}{1-\epsilon_i})) = s_{\epsilon_i} dW_t \quad 1 \leq i \leq 6$$

$$d(logit(C_r)) = d(ln(\frac{C_r}{1-C_r})) = s_r dW_t$$

$$I_{ki} = \int_{k-1}^{k} (A_{I_i} + \rho_m A_{I_{m\,i}} + \rho_w A_{I_{w\,i}})dt \quad 1 \leq i \leq n$$

$$I_{rki} = I_{ki} C_r \quad 1 \leq i \leq n$$

$$A_{I_i} = \frac{Poisson(\lambda_i S_i \Delta t)}{\Delta t} \quad 1 \leq i \leq n$$

$$A_{I_{m\,i}} = \frac{Poisson(\lambda_i R_{1i} \Delta t)}{\Delta t} \quad 1 \leq i \leq n$$

$$A_{I_{w\,i}} = \frac{Poisson(\lambda_i R_{2i} \Delta t)}{\Delta t} \quad 1 \leq i \leq n$$

The values of the parameters are the same as the ones listed in the aggregate particle filtering models and two-age-group particle filtering model, and the initial values of the stocks in this particle filtering model are also listed in the Table B.2 of Appendix B.

5.3.2 The measurement model, likelihood function, empirical datasets and evaluating particle filtering performance

The implementation of the measurement model, likelihood function, empirical datasets and evaluating particle filtering performance of the pertussis particle filtering models are similar as the particle filtering models of measles (introduced in chapter 3 and chapter 4). The Condensation Algorithm [14, 57, 74] is also used as the method of choosing the proposal distribution of the particle filtering algorithm.

Specifically, in the aggregate particle filtering model of pertussis, only one empirical dataset – monthly pertussis reported cases of the province of Saskatchewan from the year 1921 to 1956 is employed [21]. The RMSE between the monthly empirical data and the related data in the particle filtering model is calculated as the discrepancy among the sampled particles by importance sampling. The negative binomial distribution

is employed as the likelihood function, with the value of the dispersion parameter of r is 10. In the age-structured particle filtering model of pertussis of 2 age groups, three empirical datasets are used – monthly pertussis reported cases among the total population, the yearly pertussis reported cases of the child age group and the yearly pertussis reported cases of the adult age group. The weight update rule (likelihood function) is the same as the aggregate model, except for the close of each year – same as the Equations (4.17) in the measles two-age-groups models. The discrepancy is the sum of the monthly RMSE and the yearly RMSE between each age group over 12 (to convert the unit to Month).

In the pertussis age-structured particle filtering models of 32 age groups, 7 empirical datasets are empoyed – monthly pertussis reported cases among all population, and yearly pertussis reported cases of six age groups, including less than 1 year, from 1 to 4 year, from 5 to 9 year, from 10 to 14 year, from 15 to 19 year and 20 year and plus. The update weight rule is similar to the age structured pertussis model of two age groups, except for the close of each year that it is the multiplication among the 7 empirical datasets. The discrepancy is also the sum of the monthly RMSE and the yearly RMSE between each age group over 12 (to convert the unit to Month).

5.4 Results

5.4.1 Results of models incorporating empirical datasets across all timeframe

To explore the performance of different compartmental pertussis models incorporating with the particle filtering algorithm, four particle filtering models have been built in this research – the aggregate particle filtering pertussis model (denoted as $PF_{aggregate}$), the age-structured particle filtering model of 2 age groups (denoted as PF_{age_2}), the age-structured particle filtering model of 32 age groups with the Hethcote contact matrix (denoted as $PF_{age\_32\_Hethcote}$), and the age-structured particle filtering model of 32 age groups with the re-balanced contact matrix (denoted as $PF_{age\_32\_rebalance}$). Specifically, in the age-structured model of 2 age groups (child and adult), the child age group including the individuals from newborn to the end of 4 years old, while the adult age group including the individuals from the beginning of 5 years old. In all the four particle filtering models, the total number of particles in the particle filtering algorithm, in the plots of the 2D histogram and in the sampling process of calculating the discrepancy are all 3000. To explore the accuracy improvement of the particle filtering models of pertussis, I have also built a calibrated model with aggregate population. In this calibrated model, the values of the parameters got from calibration with the empirical dataset are listed as follows. The initial value in class S, I and R1 are relatively 19420, 500, 9960. The value of the infectious contact rate (β) is 56.692, and the report rate of pertussis is 0.01. Finally, this calibration model is denoted by *Calibrated*.

By comparing the discrepancy of these models, I sought to identify the model offering the greatest predictive validity. I then used the most favorable model to perform prediction analysis. To assess model results, each of the four particle filtering models was run 5 times with random seeds generated from the

Table 5.4: Comparison of the average discrepancy of all four pertussis models by incorporating empirical data across all observation points.

Model	Monthly	Yearly in Month	Total
$Calibrated$	34.2	NONE	NONE
$PF_{aggregate}$	20.9 (20.0, 21.9)	NONE	NONE
PF_{age_2}	19.9 (18.8, 21.0)	21.0 (19.2, 22.7)	40.9 (38.1, 43.6)
$PF_{age\_32\_Hethcote}$	20.6 (20.1, 21.2)	25.8 (23.0, 28.6)	46.4 (43.1, 49.7)
$PF_{age\_32\_rebalance}$	19.8 (19.5, 20.1)	28.1 (24.2, 31.9)	47.9 (43.9, 51.9)

Each of the five particle filtering models was run 5 times (the random seed generated from the same set). Shown here are the average and 95% confidence intervals (in parentheses) of the mean discrepancy for each model variant.

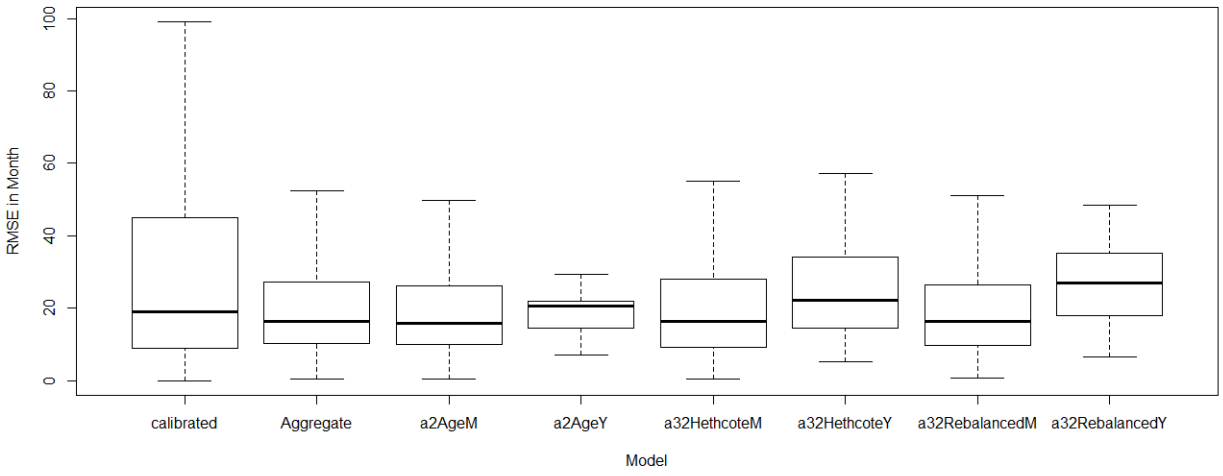


Figure 5.2: Box plots of monthly and yearly discrepancy of all models by incorporating empirical data across all observation points.

same set. I then calculated the average and 95% confidence intervals of the mean discrepancy. Table 5.4 displays the average discrepancies of the four pertussis particle filtering models and the deterministic pertussis compartmental model with calibrated parameters across all timeframe. The results of table 5.4 suggests that the particle filtering models have improved the accuracy compared with the traditional deterministic compartmental model with calibrated parameters. It is notable that both the deterministic model with calibrated parameters and the particle filtering model with aggregate population only offer monthly average discrepancy, because the age structures are not implemented in these two models. Table 5.4 indicates that the accuracy of the particle filtering models are increased significantly when compared with the calibration model – the average discrepancies of the particle filtering models are decreased markedly compared to the average discrepancy of the deterministic model with parameters derived via calibration. Moreover, although the monthly average discrepancies among the four particle filtering model with different population structure and contact matrix structure are quite close, the particle filtering models of PF_{age_2} and $PF_{age\_32\_rebalance}$ exhibit smaller average discrepancies. With respect to the yearly average discrepancies, table 5.4 shows that the age-structured model with two age groups offers better predictive performance than the models with 32 age groups. (In this regard, it bears emphasis that the particle filtering model with aggregate population does not support calculation of the yearly average discrepancy, due to the inability to compare against age-specific empirical data). It is notable that the total number of the yearly empirical datasets is different compared with the age-structured model with 2 age groups (with 2 yearly empirical datasets) and the age-structured models with 32 age groups (with 6 empirical yearly datasets). The yearly average discrepancies listed in table 5.4 are the sum of the average discrepancy with each empirical dataset. Thus, this difference may contribute to the result that the yearly average discrepancies of the model with 32 age groups are greater than the model with 2 age groups. However, I also employ the particle filtering model with 2 age groups as the minimum average discrepancy model to explore the performance of the outbreak prediction of pertussis below.

Figure 5.2 shows the box plot of the distribution of the datasets of discrepancies among the calibrated model and the four particle filtering models, where the boxplot summarizes monthly discrepancy estimates for a given model at different times. Each of the five particle filtering models was run 5 times (the random seed generated from same set). Then the average monthly and yearly discrepancy among these five runs at each time between the particle filtering models and the empirical data were plotted. It is notable that the datasets of the discrepancies of the box plot in the particle filtering model are calculated by the average value among five realization at each time point. Both the monthly and yearly distribution of the discrepancies of each age structured models are plotted in figure 5.2. This box plot also indicates that the discrepancies of all the particle faltering models are smaller than for the calibrated model. This indicates that the particle filtering models have improved the estimation accuracy compared to the traditional deterministic model with calibrated parameters. Moreover, the datasets of the discrepancy of the model PF_{age_2} has a narrow distribution, especially for the dataset of the yearly discrepancy.

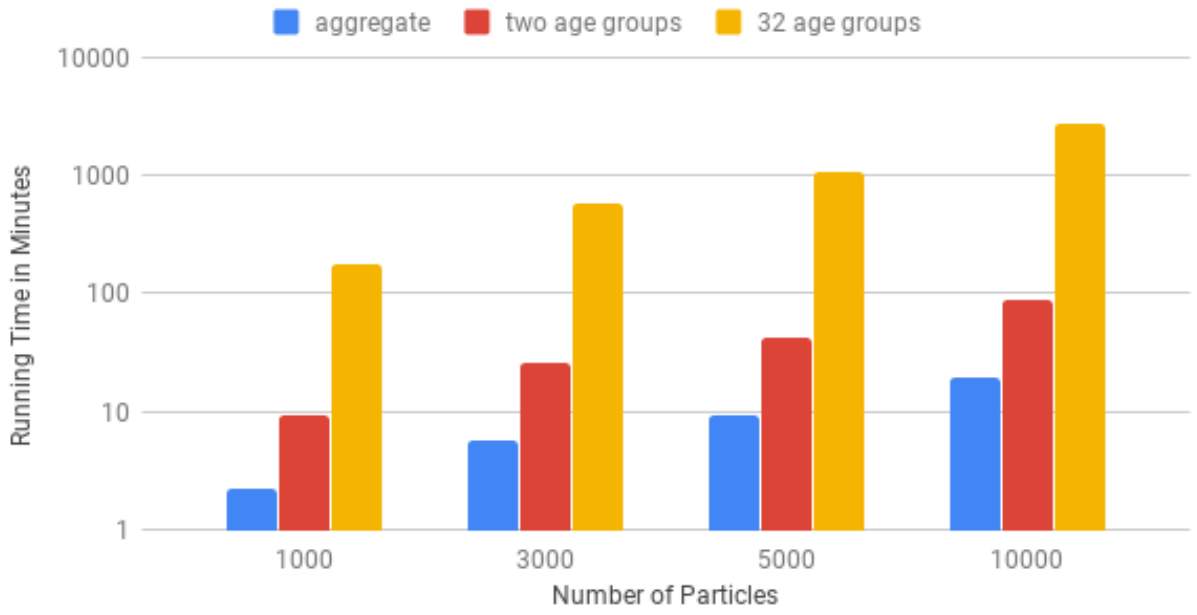


Figure 5.3: Running time of the pertussis particle filtering models.

Figure 5.3 shows the running time of the pertussis particle filtering models with three population structures – aggregate population, with two age groups and with 32 age groups (taking the one using an unbalanced contact matrix as an example). All of the models are built using Anylogic 8.1.0 software, a Java-based simulation platform. Similar to figure 4.3, figure 5.3 indicates that the running time grows linearly with an increase in the number of particles. The running time of the deterministic model with aggregate population structure – and without particle filtering mechanisms – is only 0.028 minutes, which is far fast than the particle filtering models. It is notable that the running time of the pertussis models and the measles models are not comparable, because the running time calculated in this thesis also includes the time of plotting all the results of the models for analysis, and the figures of the plots of the measles and pertussis models are different. Finally, all the pertussis particle filtering models are run on the author's another personal computer, with a processor type is intel i7 3.5 GHZ, and a memory size of 16GB DRAM.

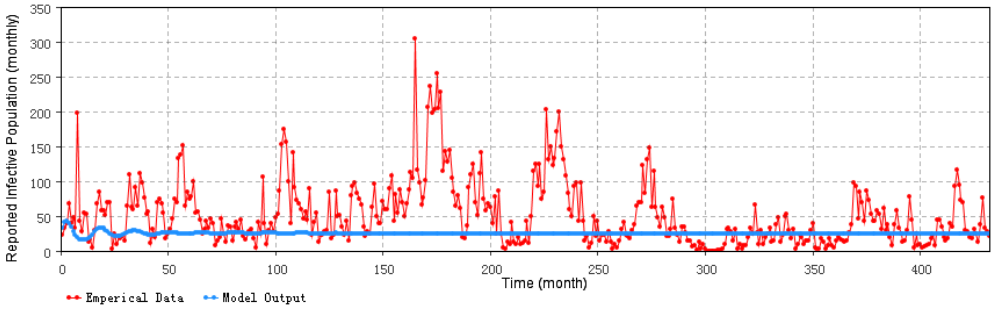


Figure 5.4: Reported pertussis cases of the calibration model (monthly).

Figure 5.4 compares the output of the calibration model and the empirical data. Similar to the situation

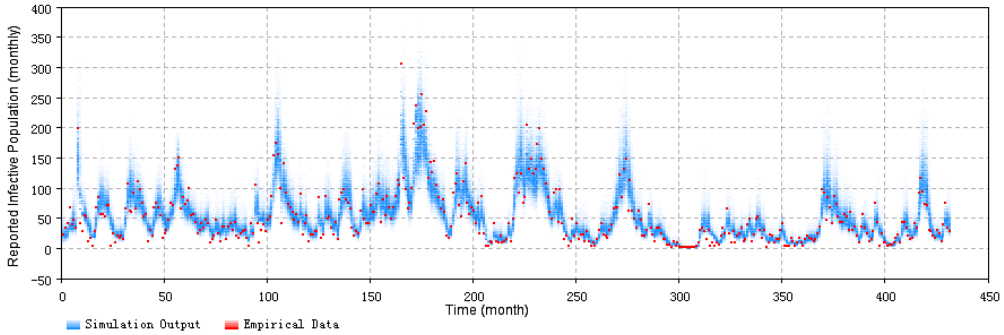


Figure 5.5: 2D histogram posterior result over the total timeframe of the aggregate particle filtering model of pertussis.

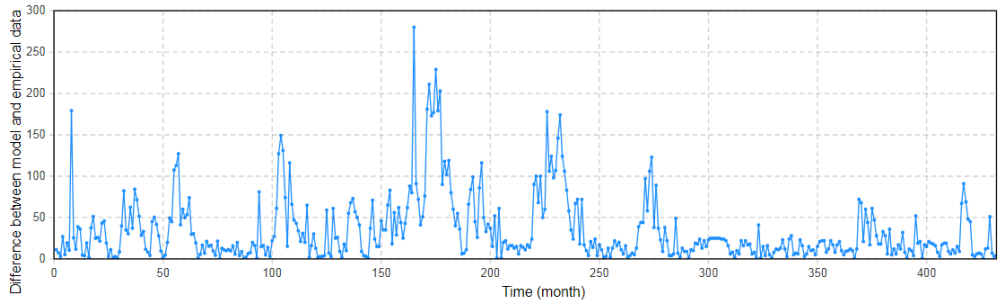


Figure 5.6: Difference between the results of the calibration model and the empirical data of pertussis (monthly).

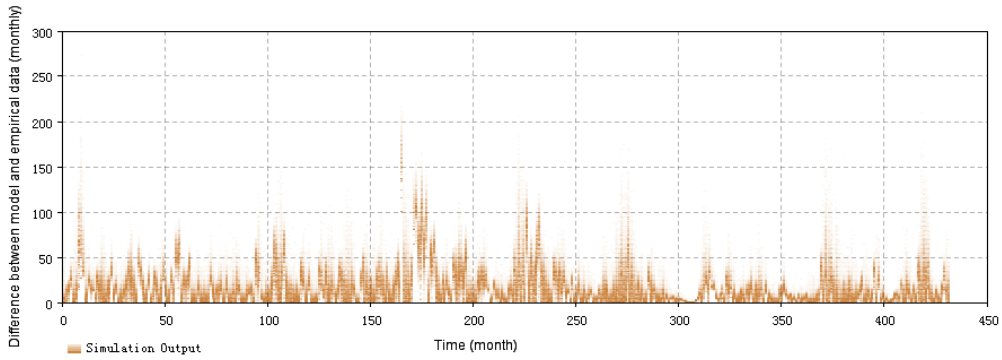


Figure 5.7: 2D histogram plot of the difference between the posterior result of the particle filtering model and the empirical data of pertussis (monthly).

with the measles model, figure 5.4 indicates that the deterministic model with calibrated parameters encounters difficulties in tracking the oscillation of the outbreak of pertussis almost across the entire model time horizon, because the deterministic model of pertussis would approaches a stable equilibrium.

Figure 5.5 presents the posterior results of the pertussis particle filtering model with aggregate population structure for the entire timeframe. For this diagram, the results of the particle filtering model are sampled according to the weight of all particles after the weights are updated by incorporating the empirical data of the current time point. The values of empirical data points are shown in red, while the sampled posterior distribution of particle filtering model is shown in blue at each time. Figure 5.5 demonstrates that most of the empirical data points are located in or near the high density location of the results of the particle filtering model. It further indicates that the particle filtering model has the capability to track the outbreak of pertussis, especially compared with the calibrated model whose results are shown in figure 5.4.

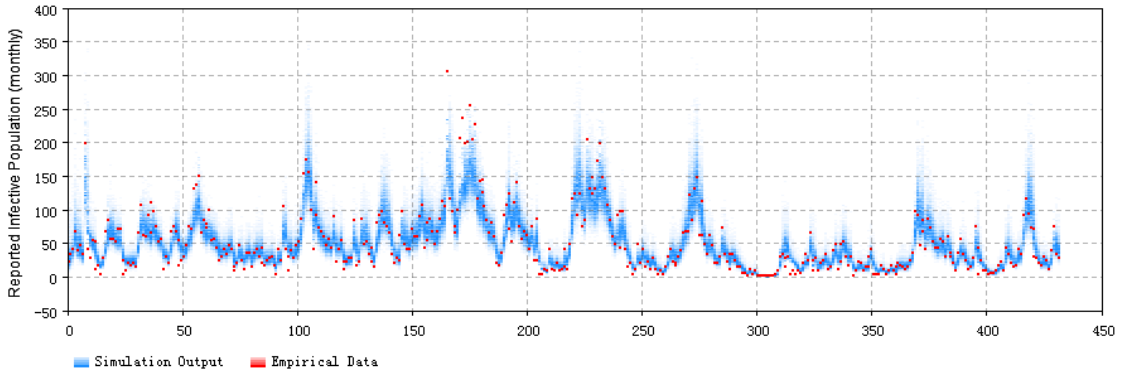
Figure 5.6 shows the difference between the results of the deterministic model with calibrated parameters and the empirical data of pertussis in each time point.

Figure 5.7 represents the 2D histogram plot between the empirical data and the sampled posterior distributions of the pertussis particle filtering model with aggregate population structure by incorporating the empirical data over the whole model time horizon. It is notable that the particles sampled in figure 5.5 are the same as the particles samples in figure 5.7.

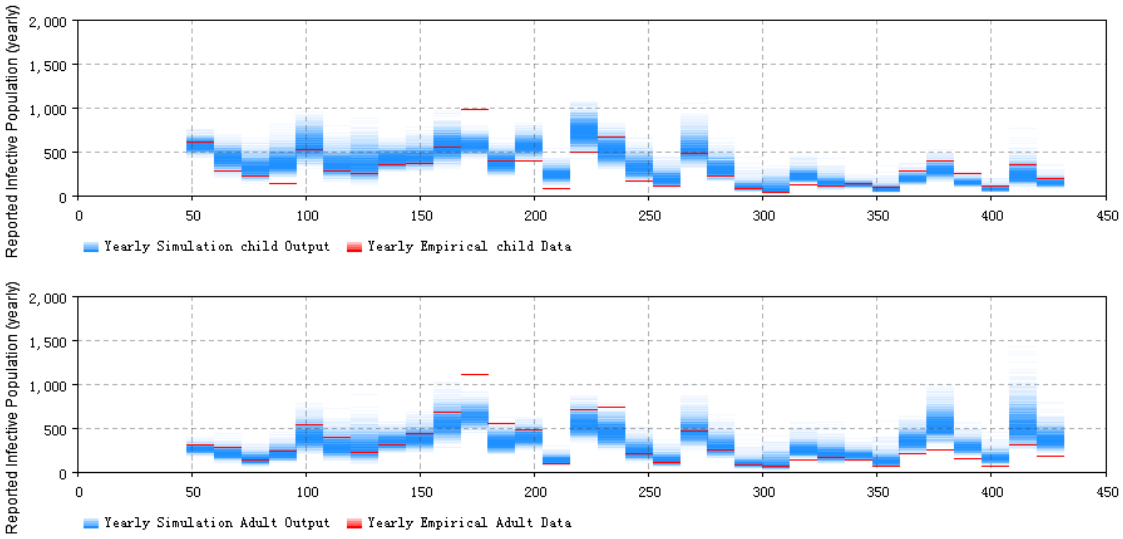
The previous results indicate that the particle filtering models considered here can not only decrease the discrepancy between the model results and the empirical data (see table 5.4, figure 5.2 and the comparison between figure 5.6 and figure 5.7), but can further track the oscillation of the outbreaks of pertussis – by comparing the results of particle filtering model (figure 5.5) and deterministic model with calibrated parameters (figure 5.4). All of these results suggest that incorporating particle filtering in the compartmental model of pertussis could enhance the simulation accuracy and support more accurate outbreak tracking.

Figure 5.8 displays the 2D histogram plots comparing both the monthly and yearly empirical datasets (on the one hand) with the distributions of samples from the posterior distribution of incident cases from the pertussis age structured pertussis particle filtering model with 2 age groups and where the individuals in the children age group are those in the first 5 years of life (denoted as PF_{age_2}) (on the other). This figure demonstrates that the model PF_{age_2} is capable to tracking and simulating the outbreaks of pertussis, considered here because most of the monthly and yearly empirical data (shown in the red dashes) in each month are located in or near the high density area of the sampled distribution of the particle filtering model (shown in blue in the 2D histogram plots).

Similarly, figure 5.9 displays the 2D histogram plots comparing both the monthly and yearly empirical datasets (on the one hand) with the sampled posterior distribution of incident cases from the pertussis age structured pertussis particle filtering model with 32 age groups and the Hethcote contact matrix (denoted as $PF_{age\_32\_Hethcote}$) (on the other). It is notable that the total number of the yearly empirical datasets employed is 6. This figure also demonstrates that the model $PF_{age\_32\_Hethcote}$ is capable of tracking and

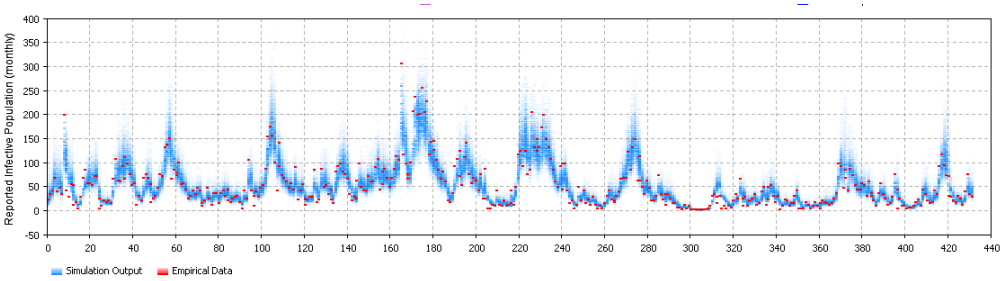


(a)

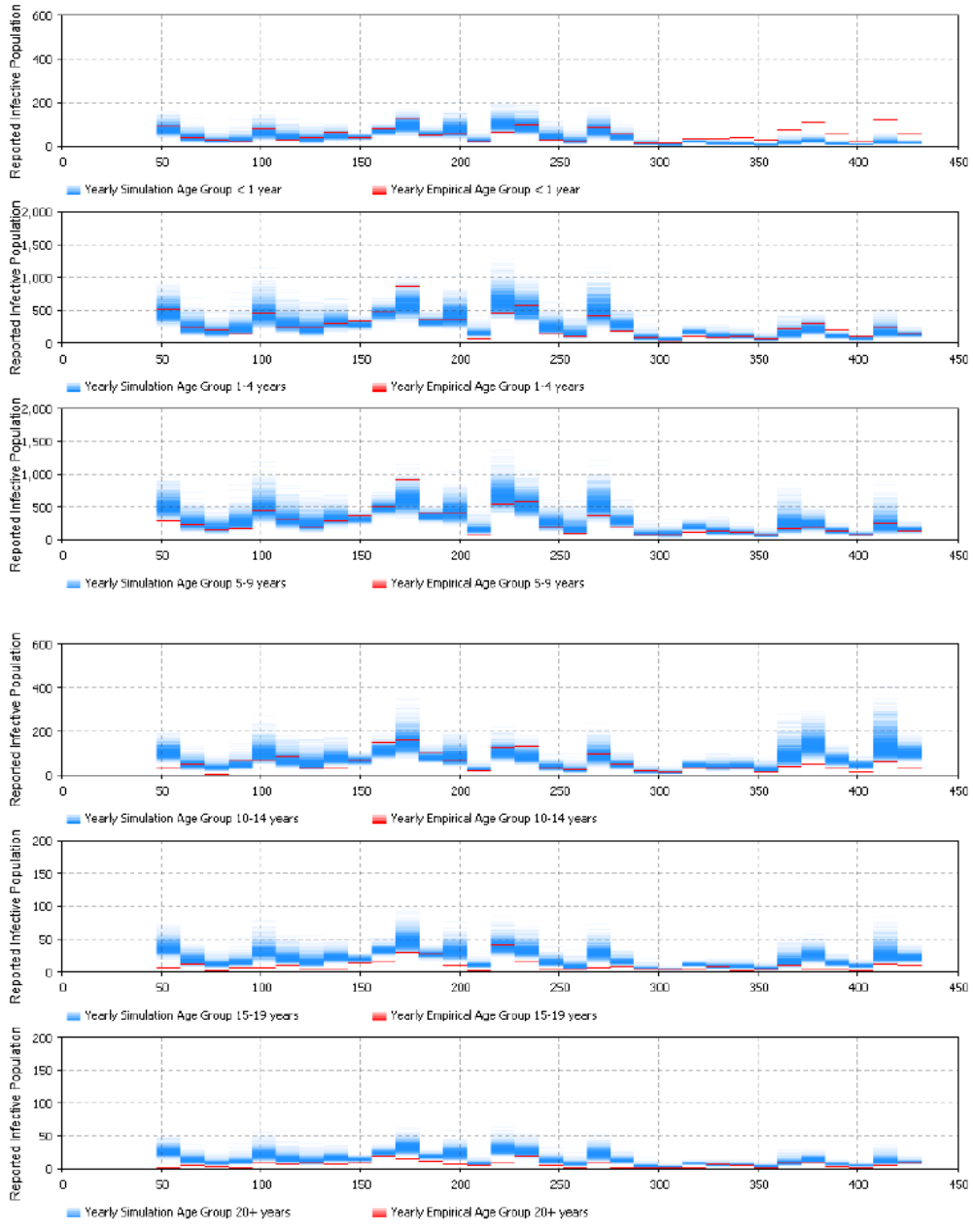


(b)

Figure 5.8: 2D histogram posterior result over the total timeframe of the age structured model of 2 age groups. (a) the monthly particle filtering result across all population. (b) the yearly particle filtering result of the children and adult age groups.



(a)



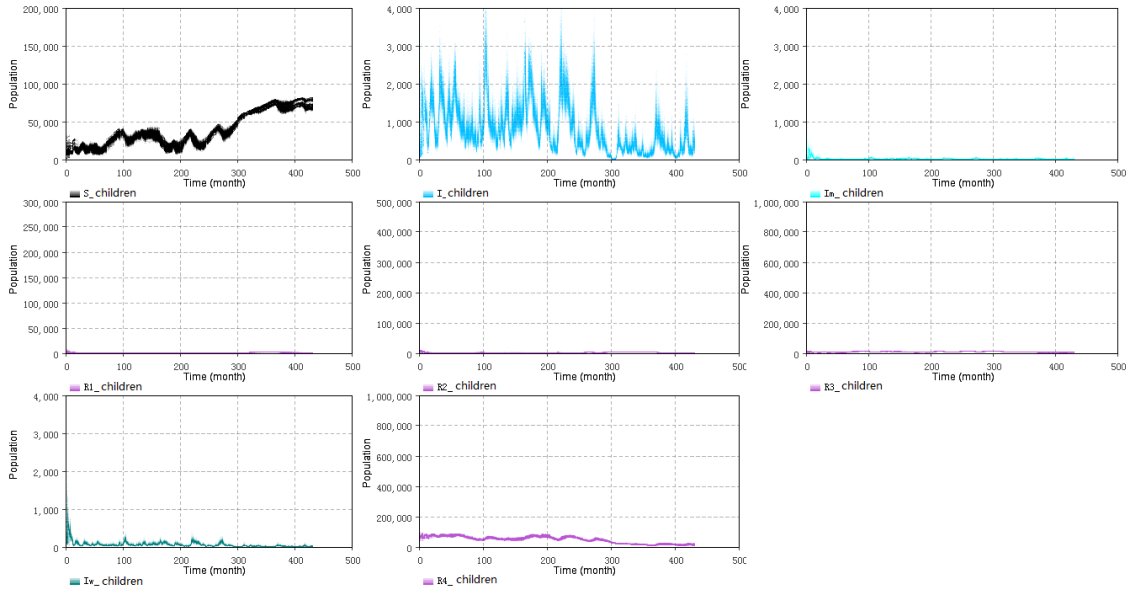
(b)

Figure 5.9: 2D histogram posterior result over the total timeframe of the age structured model of 32 age groups with the Hethcote contact matrix. (a) the monthly particle filtering result across all population. (b) the yearly particle filtering results of each age group of empirical datasets.

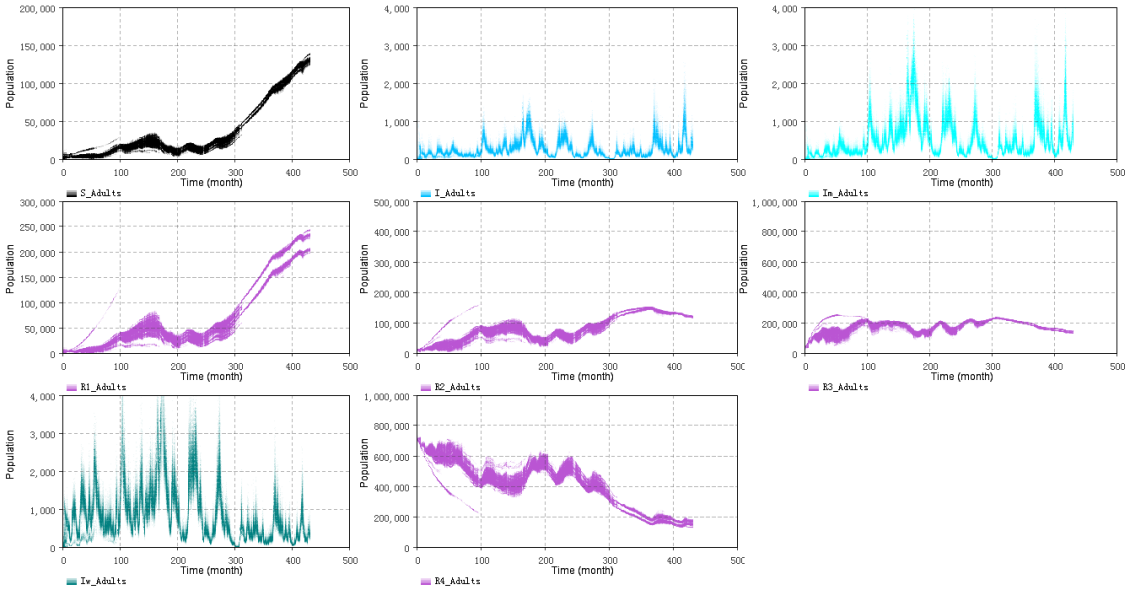
simulating the outbreaks of pertussis, because most of the monthly and yearly empirical data in each time are located in or near the high density area of the results of the particle filtering model.

Figure 5.5, figure 5.8 and figure 5.9 represent the 2D histogram posterior result of all the particle filtering models, except for the age-structured model of 32 age groups with re-balanced contact matrix. The reason I do not plot the result of this model is that it is quite similar to the 32-age-group model with Hethcote contact matrix, which is shown in figure 5.9. These 2D histogram plots indicate that both the age-structured particle filtering models and the aggregate population particle filtering model have the capability to simulate the outbreak pattern of pertussis. The results of the models could match the empirical datasets quite well, including both monthly empirical dataset and yearly empirical datasets. Compared with the result of the calibration model which is shown in figure 5.4, the particle filtering models are capable of localizing the model's prediction of empirical data lies near those data. Incorporating the discrepancy of each model shown in the table 5.4 and figure 5.2, although all these four pertussis particle filtering models are capable of tracking and estimating the pertussis outbreaks, I still employ this minimum discrepancy model – the age-structured particle filtering pertussis model of 2 age groups, divided at the age of 5 – to perform the prediction and intervention analysis.

Similar to what was discussed in chapter 3 and chapter 4, the pertussis particle filtering models can contribute to the estimation of entire state – including the latent and observable state of the dynamic models during the period of incorporating the empirical datasets. In the investigation of pertussis considered here, the empirical data – monthly pertussis reported cases across all population and yearly reported pertussis cases of different age groups, are only related to an aggregation of the latent states of weakly infectious (I_w), medium infectious (I_m), and fully infectious (I). However, by applying particle filtering to the system dynamics models of pertussis and the empirical data, the pertussis particle filtering models are capable of estimating the entire state across the whole model time horizon during which empirical datasets are available by incorporation of multiple lines of empirical data. Figure 5.10 shows the 2D histogram plots of samples from the distributions of latent stocks with the minimum discrepancy model (the age structured model with 2 age groups where the child age group consists of children in their first 5 years and incorporating both the monthly and yearly empirical datasets) as an example. Figure 5.10 indicates that most of the fully infectious individuals are estimated to be located amongst the children (less than 5 years of age) age group, while most of the weak infectious and medium infectious individuals are located in the "adult" (equal and greater than 5 years) age group. Most of the recovered individuals are also located in the adult age group. This lies in accordance with the expectations for pertussis transmission in the real world and builds confidence in the capacity of the model to meaningfully estimate the latent state. As noted below, estimation of latent state can also be an important enabler for understanding of the effects of interventions, for understanding the current epidemiological context, as well as for providing insights into historical circumstances.



(a)



(b)

Figure 5.10: 2D histogram results for the latent states in the dynamic models with the two age groups of the pertussis age structured particle filtering model with 2 age groups. (a) the child age group (those within their first 5 years of life). (b) the adult age group (years 5 and up).

5.4.2 Prediction with the minimum discrepancy model

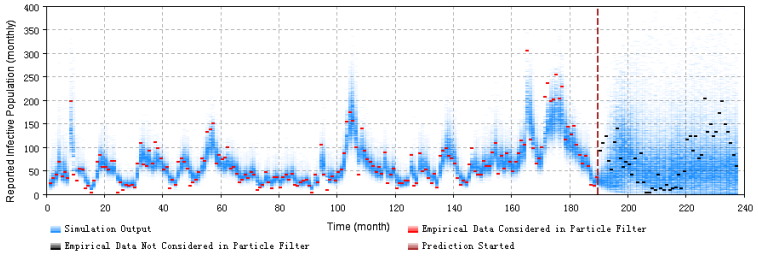
Similar to the analysis of the measles particles filtering model introduced in chapter 4, to assess the outbreak predictive capacity of the pertussis particle filtering models, I have performed the prediction experiments, by changing different Prediction Start Time of T^*, with the same different archetypal situations list as follows. It is notable that the minimum discrepancy model – the age structured model with 2 age groups where the child age group is up to 5 year and incorporating both the monthly and yearly empirical datasets, which is identified in the previous section is employed to perform these experiments.

(1) Prediction started from the first or second time points of an outbreak.

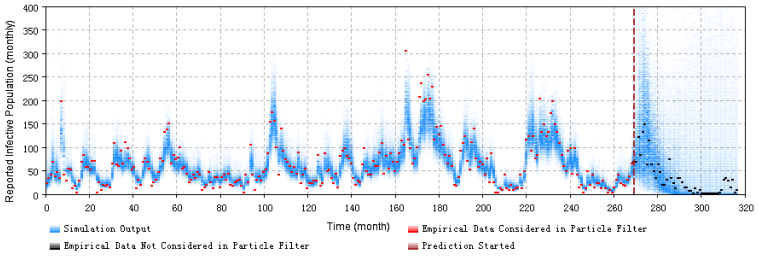
(2) Prediction started before the next outbreak.

(3) Prediction started from the peak of an outbreak.

(4) Prediction started from the end of an outbreak.



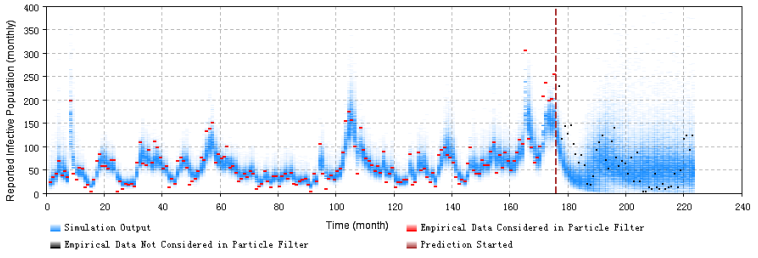
(a)



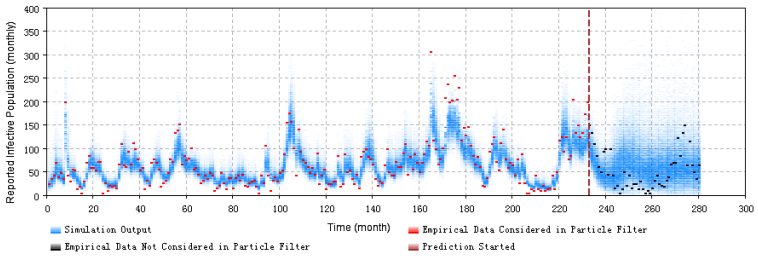
(b)

Figure 5.11: 2D histogram depicting prediction from the first or second time points of an outbreak of the minimum discrepancy model. (a) predicted from the month 190. (b) predicted from the month 269.

Figure 5.11 displays the prediction results of the minimum discrepancy model in situations in which the prediction started from the first or second time points of an outbreak. It does so with two experiments, one in which prediction started from month 190 ($T^* = 190$) – with s_{β_c} is 0.6, and monthly prediction discrepancy 72.6, the sum of yearly prediction discrepancy of all age groups per month is 53.3 – and one in which such prediction initiated at month 269 ($T^* = 269$) – with s_{β_c} is 0.6, and monthly prediction discrepancy 87.4, the

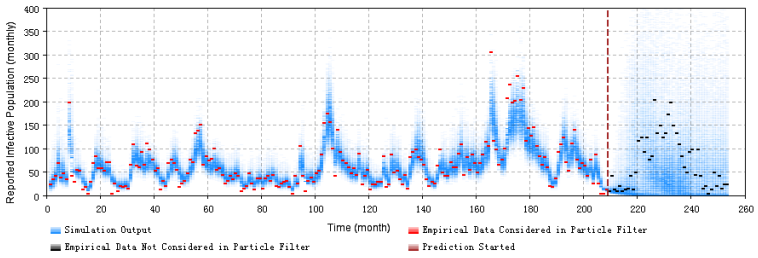


(a)

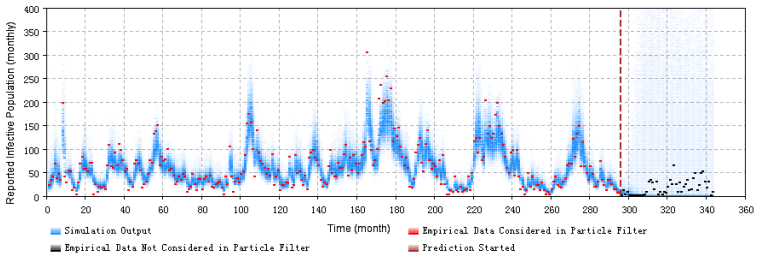


(b)

Figure 5.12: 2D histogram depicting prediction from the peak of an outbreak of the minimum discrepancy model. (a) predicted from the month 176. (b) predicted from the month 233.

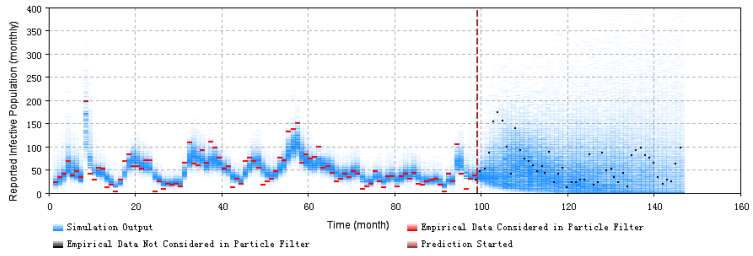


(a)

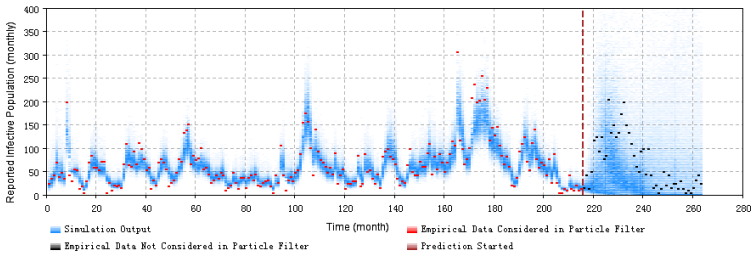


(b)

Figure 5.13: 2D histogram depicting prediction from the end of an outbreak of the minimum discrepancy model. (a) predicted from the month 209. (b) predicted from the month 296.



(a)



(b)

Figure 5.14: 2D histogram depicting prediction before the next outbreak of the minimum discrepancy model. (a) predicted from the month 99. (b) predicted from the month 216.

sum of yearly prediction discrepancy of all age groups per month is 63.6 – respectively. In the prediction process of the particle filtering model, the weights of particles will stop updating at the "Prediction Started Time" (T^*) by incorporating the empirical data. Following that point, all of the particles run without new empirical data coming. In this chapter, all of the prediction experiments are run 4 years after the "Prediction Start Time" T^*. In the 2D histogram plot of figure 5.11, the empirical data having been considered in the particle filtering process are shown in red (incorporated in training the models), while the empirical data not having been considered in the particle filtering process (and only displayed to compare with model results) are shown in black. The vertical straight line labels the "Prediction Start Time" of T^* of each experiment.

Figure 5.12 displays the prediction results of the particle filtering model in situations in which the prediction started from the peak of an outbreak. As above, it examines two experiments, one in which prediction started from month 176 ($T^* = 176$) – with s_{β_c} is 0.5, and monthly prediction discrepancy 54.9, the sum of yearly prediction discrepancy of all age groups per month is 55.4 – and another in which prediction initiates at month 233 ($T^* = 233$) – with s_{β_c} is 0.5, and monthly prediction discrepancy 56.8, the sum of yearly prediction discrepancy of all age groups per month is also 56.8 – respectively. The layout of the 2D histogram plot of figure 5.12 is the same as for figure 5.11.

Figure 5.13 displays the prediction results of the particle filtering model in situations in which prediction started from the end of an outbreak. As above, it investigates this with two experiments, one in which prediction starts from month 209 ($T^* = 209$) – with s_{β_c} is 0.6, and monthly prediction discrepancy 79.6, the sum of yearly prediction discrepancy of all age groups per month is 56.7 – and month 296 ($T^* = 296$) – with s_{β_c} is 0.6, and monthly prediction discrepancy 99.5, the sum of yearly prediction discrepancy of all age

groups per month is 70.3 – respectively. The layout of the 2D histogram plot of figure 5.13 is the same as in figure 5.11.

Figure 5.14 displays the prediction results of the particle filtering model in situations in which it initiates predictions before the next outbreak. As above, this is investigated with two experiments. In the first, prediction starts from month 99 ($T^* = 99$) – with s_{β_c} is 0.4, and monthly prediction discrepancy is 75.3, the sum of yearly prediction discrepancy of all age groups per month is 56.9 – and month 216 ($T^* = 216$) – with s_{β_c} is 0.6, and monthly prediction discrepancy 79.2, the sum of yearly prediction discrepancy of all age groups per month is 55.3 – respectively. The layout of the 2D histogram plot of figure 5.14 is the same as for figure 5.11.

Figures 5.11–5.14 display the prediction results of these situations with the monthly 2D histogram of reported cases of the total population. These prediction results suggest that the pertussis particle filter model offers the capacity to probabilistically anticipate pertussis dynamics with a fair degree of accuracy. From the 2D histogram plots, empirical data lying after Prediction Start Time – and thus not considered by the particle filtering machinery – mostly lie within the high-density range of the particles. Notably, in such examples, the particle filter model appears to be able to accurately anticipate a high likelihood of a coming outbreak and non-outbreak. Such an ability could offer substantial value for informing the public health agencies with accurate predictions of the anticipated evolution of pertussis over the coming months.

5.4.3 Intervention with the minimum discrepancy model

Based on the previous discussion, particle filtering models examined here are capable of estimating the entire latent state and projecting outbreak progression. These features also support particle filtering models in simulating intervention strategies.

In this section, I have implemented several intervention experiments to simulate public health intervention policies, based on the minimum discrepancy particle filtering pertussis model discussed above. The intervention strategies are normally performed before or at the very beginning of an outbreak. Moreover, to support easy comparison with the baseline prediction result of the minimum discrepancy model absent any interventions, all of the intervention strategies are simulated starting at the start month of an outbreak (month 269) in this project. The baseline prediction result of the minimum discrepancy model without any interventions are shown in figure 5.11 (b). We examine below the impact of two stylized intervention policies discussed in the context of the chapters on measles models – quarantine and vaccinating.

Figure 5.15 and figure 5.16 display simulation results of quarantine intervention strategies by decreasing the contact rate parameter to be 20% and 50% less than the value before intervention, respectively. Similarly to the 2D histogram plot of the prediction result without any intervention shown in figure 5.11 (b), the red dots represent the empirical data incorporated in the particle filtering model, while the black dots represent the empirical data not incorporated in the model, but presented for comparison purposes. It is notable that the empirical data shown in black are under the situation of absent any intervention (that is, in a baseline

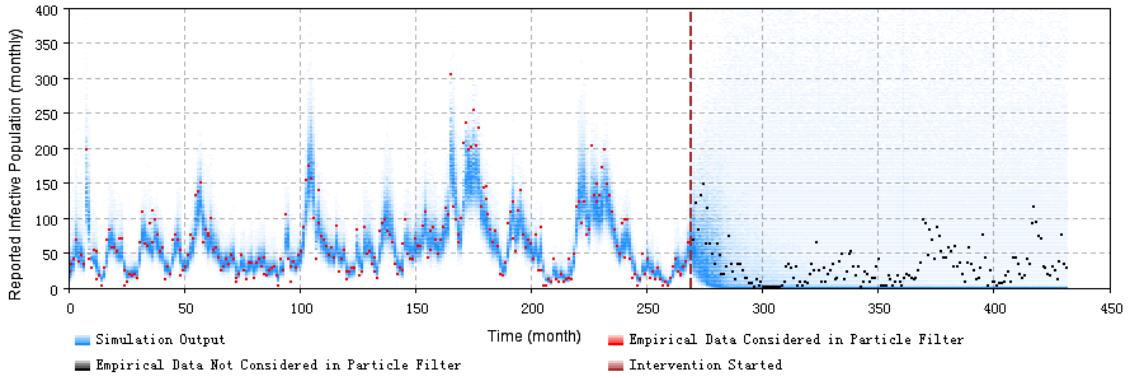


Figure 5.15: 2D histogram of simulating quarantine during a pertussis outbreak.. This is realized by decreasing the contact rate by 20%.

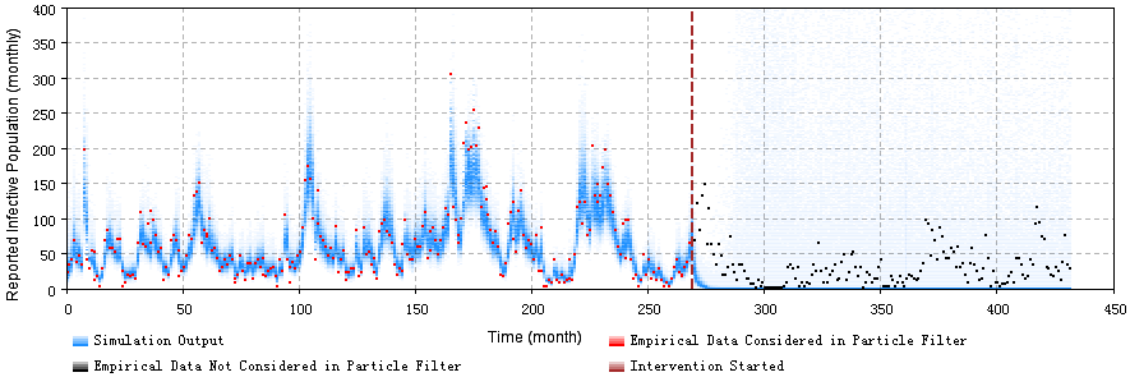


Figure 5.16: 2D histogram of simulating quarantine during a pertussis outbreak. This is realized by decreasing the contact rate by 50%.

context that lacked an intervention of the sort simulated here). By comparing the quarantine intervention results (see figure 5.15 and figure 5.16) with the model result without intervention shown in figure 5.11 (b) and the empirical data during the intervention period (not incorporated in the particle filtering model), we can see that, although the interventions are implemented in a stylized fashion, by virtue of its ability to estimate the underlying epidemiological state, the pertussis particle filtering model is capable of evaluating public health intervention policies to reduce or even avoid the outbreak of pertussis.

To simulate a immunization intervention during a pertussis outbreak, a vaccination parameter is incorporated to represent the fraction of the population moved with respect to their immunity status. Specifically, recall that the pertussis model characterizes a chain of successively higher levels of vaccine-induced protection; this parameter moves a certain fraction of the population from the previous stock (before vaccination) to the stock representing the next higher level of vaccination (following vaccination). Normally, there are 4 doses of pertussis vaccination in total. Thus, in the simulation of vaccination, I just simply move the fraction of individuals specified by the vaccination parameter from the stock of S to R_1 to simulate the first dose; I similarly move an identical fraction of the population from the stock of R_1 to R_2 to simulate the second dose; I further move an identical fraction of individuals from the stock of R_2 to R_3 to simulate the third

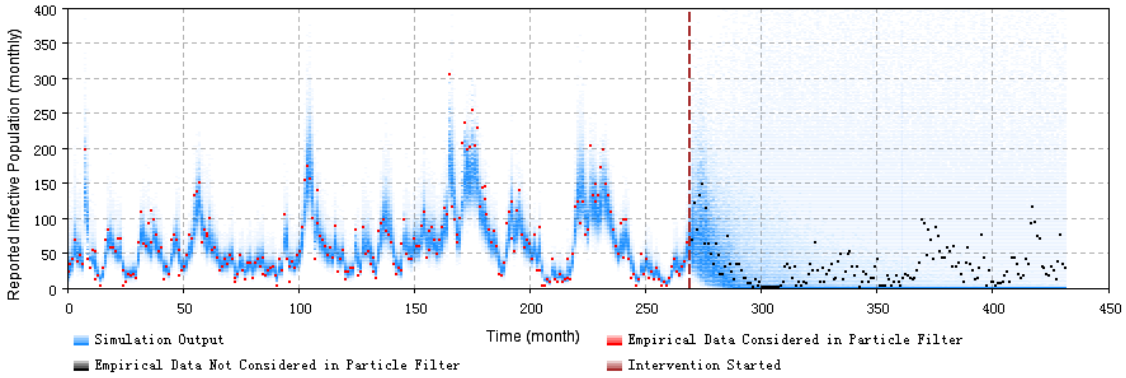


Figure 5.17: 2D histogram of simulating an immunization campaign during an outbreak.. This is realized by increasing the vaccine-induced protection level among 20% of the population.

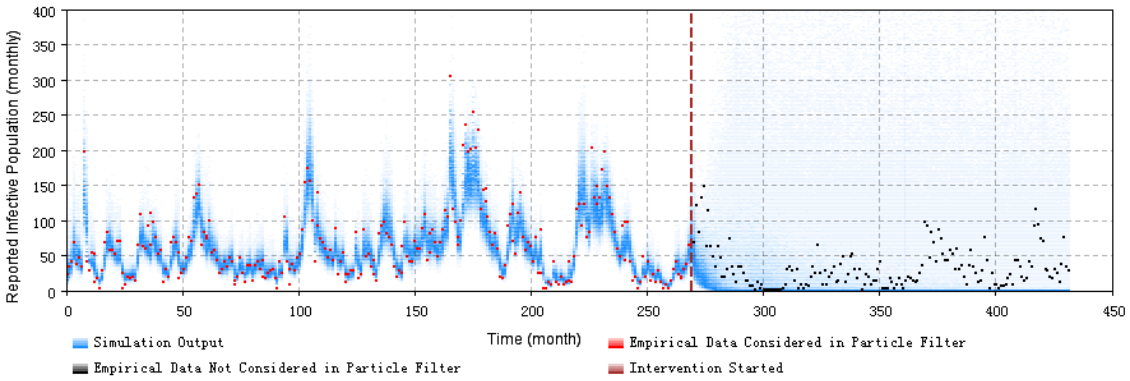


Figure 5.18: 2D histogram of simulating an immunization campaign during an outbreak. This is realized by increasing the vaccine-induced protection level among 50% of the population.

dose; finally, I move an identical fraction of individuals from the stock of R_3 to R_4 to simulate the fourth dose. Figure 5.17 and figure 5.18 show the results of the vaccination intervention. The layout of the 2D histogram plots of the vaccination interventions with figure 5.17 and figure 5.18 is same as the quarantine plots of figure 5.15 and figure 5.16. Similar to the intervention results of the measles particle filtering model, these results of pertussis interventions demonstrate that by virtue of its ability to estimate the underlying epidemiological state, the use of particle filtering with pertussis models supports evaluating public health intervention policies to reduce or even avoid the outbreak of pertussis, etc.

5.5 Discussion and conclusion

This chapter presents a new method of tracking pertussis outbreak patterns by integrating a particle filtering algorithm with a pertussis compartmental model and empirical incidence data. Specifically, four particle filtering models have been researched, including an aggregate population model, a two-age-groups population model, and a 32-age-group population models with both a contact matrix derived from Hethcote (1997) [50] and a re-balanced contact matrix. Although the results of all these four particle filtering models could match

the empirical data quite well, the minimum discrepancy model emerged from the 2-age-group age-structured particle filtering model, where the individuals in the child age group represent children in the first 4 years of life, and incorporating with both monthly and yearly empirical datasets.

It is notable that according to Equation (2.36), with more age groups considered in the age-structured model, the model could simulate the aging rate (c_i) more accurately. However, in this project, the discrepancies associated with both of the 32-age-groups particle filtering models do not show evident decrease compared with the two-age-group particle filtered models. I think that the possible reasons are listed as follows. Firstly, the stochastic processes considered in both the 32-age-groups age-structured models and the two-age-groups age-structured model are different, especially in characterizing the stochastic evolution of contact rate. Secondly, the likelihood functions employed in this project – the multiplication across all empirical datasets based on the negative binomial distribution with dispersion parameters set to a uniform value – may be too naive to capture the difference between the age groups within the empirical datasets. Thirdly, as the number of age groups increases, the dimensions of the particle filtering models increase dramatically. As introduced in the section of literature review in chapter 2, the particle filtering algorithm has several limitations. An important one is that the particle filtering method with naive condensation algorithm may encounter problems in high-dimensional systems. In high-dimensional systems, the probability density functions would be more complex. Then, the variance of the density functions may be high, because it is difficult to be represented by simple probability density functions. Finally, the relationship between the number of particles and the dimensions of the models may merit additional research.

Then, the prediction experiments have been calculated based on the minimum discrepancy model – the age-structured model using two age groups. The results suggest that particle filtering approaches offer notable strengths in predicting occurrence of the pertussis outbreak in the subsequent months.

Moreover, it is notable that the pertussis particle filtering models could support effective estimation of the entire state of the dynamics models during the time incorporating the empirical datasets. Combined with the capability to perform outbreak projections, the particle filtering models can support effective evaluation of intervention strategies.

Similar to the measles particle filtering models introduced in the chapter 3 and chapter 4, the noise in the stochastic processes in the state space models could impact the particle filtering model differently during the estimation and prediction period. Thus, the noise in the particle filtering models in this chapter should also be controlled in a proper range, by tuning the parameters of diffusion coefficients in the stochastic processes related to the Brownian motion.

Also similar to the measles particle filtering models introduced in the chapter 3 and chapter 4, the initial values of the age-structured population models are estimated both manually and by the particle filtering algorithm. Specifically, the population distribution among the different age groups are tuned manually, while the population distribution among different stocks within a given age group is estimated by the particle filtering algorithm by setting the initial values of stocks in a proper range following a uniform distribution,

but maintaining a total number of individuals for that age group across the stocks. Especially in building the 32-age-groups particle filtering models, much time and efforts was been used in estimating the population distribution among the different age groups.

Finally, the hypotheses proposed in the thesis statements are demonstrated for pertussis models. Firstly, the 2D histogram plots comparing the empirical data and samples from the posterior distributions of the particle filtering models' output data (the monthly and yearly reported cases of pertussis) indicates that the high probability density region of the model's prediction of empirical data encompasses or lies near those data. Secondly, the discrepancy of the pertussis particle filtering model's predictions vs. observed data is reduced with a factor of 1.6 when compared with the traditional calibration model. This reduction indicates that the particle filtering algorithm is capable of enhancing model predictions when compared with predictions resulting from the traditional calibration process. Additionally, it is notable that the deterministic model with calibrated parameters encounters difficulties in tracking the fluctuation of the outbreak pattern of measles, while the particle filtering model is capable to tracking the oscillation of the outbreak of pertussis, and moreover, could predict the outbreak projections and simulate intervention strategies.

Chapter 6

Classifying outbreak occurrence with the prediction results of the particle filtering models

6.1 Introduction

The previous three chapters have explored how particle filtering models of measles and pertussis offer the capability to probabilistically anticipate measles and pertussis dynamics with a fair degree of accuracy. Specifically, the particle filtering models appear to be able to accurately anticipate a high likelihood of a coming outbreak and non-outbreak over the following months. In this chapter, I investigate the effectiveness of using the particle filtering models in predicting the outbreak of measles in the next time unit (Month in the models in this thesis).

6.2 Methodology

The goal of the classification task investigated here is to map from the cases counts of measles or pertussis predicted by particle filtering models in the next month to the class of outbreak or non-outbreak. This mapping can be represented by the following equation [74]:

$$z_k = f(I_{\text{rpk}}^{(i)}) \tag{6.1}$$

where $\left\{\left\{I_{\text{rpk}}^{(i)}\right\}_{i=1}^{N_s}\right\}_{k=1}^{T_f}$ indicates the matrix of the reported cases of measles or pertussis predicted by the particle filtering models of particle i ($1 \leq i \leq N_s$) (sampled by importance sampling) at time k ($1 \leq k \leq T_f$). T_f is the total running time of the model, and N_s is the total number of particles in the particle filtering models. $\{z_k\}_{k=1}^{T_f}$ is the vector of the predicted classes. Within this project, $z_k \in \{0,\ 1\}$, where 0 indicates non-outbreak, and 1 indicates the outbreak. The value I_{rpk} used here is generated by the particle filtering, which is the total reported cases of measles or pertussis across all subpopulations in the particle filtering models.

The method employed to address this binary classification problem is as follows: At time step *k-1*, we only calculate the state space model from time *k-1* to *k*, without the weight of particles being updated by incorporating y_k. As a result, the weight of each particle is the one calculated in time step *k-1*. Hence, the

state vector $x_k^{N(i)}$ with the weight ${w_{k-1}}^{(i)}$ represents the prediction results for time step k, on the basis only of data available until time *k-1*. As a result, I could get the predicted reported cases of particles generated by importance sampling of the particles – that is, sampling those particles according to weight. Following this, all the $I_{\text{rpk}}^{(i)}$ will make up a matrix. One dimension is the time steps (k). The other dimension is the particles (i).

To perform the classification analysis, two processes need to be conducted. In the first process, I define the threshold of a particle i at time step k above which to consider that particle as positing an outbreak (θ_{ik}). In this project, I impose a reported count threshold above which I consider there as being an outbreak is "mean plus 1.5 times the standard deviation of the empirical monthly reported cases". As given by Equation (6.2) and (6.3), if for a particle i where $I_{\text{rpk}}^{(i)} \geq \theta_{\text{ik}}$, I label this data $I_{\text{rpk}}^{(i)}$ as indicative of an "outbreak". Otherwise, it is considered to not be indicative of an outbreak. In the second process, I define a threshold of fraction (θ_k) of particles required as positing an outbreak at time k for us to consider there as being an outbreak at time k. The equations for classification are then listed as follows:

$$z_{\text{ik}} = \begin{cases} 1, I_{\text{rpk}}^{(i)} \geq \mu^E + 1.5\sigma^E \\ 0, I_{\text{rpk}}^{(i)} < \mu^E + 1.5\sigma^E \end{cases} \tag{6.2}$$

$$z_k = \begin{cases} 1, \sum_{i=1}^{N_s} z_{\text{ik}} \geq \ \theta_k * N_s \\ 0, \sum_{i=1}^{N_s} z_{\text{ik}} < \ \theta_k * N_s \end{cases} \quad (0 \leq \theta_k \leq 1) \tag{6.3}$$

where μ^E is the mean of empirical data, σ^E is the standard deviation of the empirical data, and N_s is the total number of particles.

I could get the predicted classification vector $\{z_k\}_{k=1}^{T_f}$ by conducting the particle filtering algorithm and applying equation (6.2) and (6.3). I further denote $\{y_k\}_{k=1}^{T_f}$ as the empirical vector of whether a measles (or pertussis) outbreak indeed obtained at time k, $y_k \in \{0, 1\}$. The calculation method of y_k is similar to that of z_{ik}:

$$y_k = \begin{cases} 1, y_k^M \geq \mu^E + 1.5\sigma^E \\ 0, y_k^M < \mu^E + 1.5\sigma^E \end{cases} \tag{6.4}$$

where y_k^M is the empirical data of measles or pertussis reported cases. Finally, to summarize the performance of the classifier, I employ metrics such as the area under the Receiver Operating Characteristic (ROC) curve, etc.

6.2.1 The algorithm of particle filtering with the next month prediction output

With the previous introduction of particle filtering, the generic particle filter algorithm considering the output of the reported cases for the next month employed in this paper is given as follows [26, 6, 84]:

1. At time k=0:

(1) Sample $X_0^{N(i)}$ from $q_0(x_0^N)$;

(2) Compute a *weight* for each particle $w_0^{(i)} = \frac{1}{N_s}$. This indicates that the weight at initial time step follows uniform distribution.

2. At time $k \geq 1$, perform a recursive update as follows:

(1) Advance the sampled state by sampling $X_k^{N(i)} \sim q_k(x_k^N | y_k, X_{0:k-1}^{N(i)})$ and set $X_{0:k}^{N(i)} = (X_{0:k-1}^{N(i)}, X_k^{N(i)})$;

(2) To support classification, output $I_{\text{rpk}}^{(i)}$ by importance sampling, where $I_{\text{rpk}}^{(i)}$ is the sum of all the age groups in the age structured model;

(3) Update the weights to reflect the probabilistic and state update models as follows:

$$w_k^{(i)} = W_{k-1}^{(i)} \frac{p(y_k^M | X_k^{N(i)}) p(X_k^{N(i)} | X_{k-1}^{N(i)})}{q(X_k^{N(i)} | X_{k-1}^{N(i)}, y_k^M)}.$$

Normalize the weights $W_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_s} w_k^{(i)}}$

(3) Calculate the S_{eff} : $\frac{1}{\sum_{i=1}^{N_s} (w_k^{(i)})^2}$

(4) If $S_{\text{eff}} < S_T$ (S_T is the minimum effective sample size – the threshold of resampling), perform resampling to get a new set of $X_k^{N(i)}$. And set the weight of the new particles as $\frac{1}{N_s}$.

6.3 Prediction results of next month breakout classification of the minimal discrepancy model

6.3.1 Results of measles

By incorporating the prediction results of the lowest discrepancy particle filter measles model $PF_{age\_15\_both}$, I could perform a classification-based prediction of whether measles will break out or not in the next month. Figure 6.1 displays the ROC curve showing the prediction results. The Area Under the Curve (AUC) of the ROC curve is 0.89, indicating a favourable classification ability.

Figure 6.2 displays the scatter plot between the monthly empirical data and the mean and median of the model predicted next month results over all sampled particles. The figure further displays the results of a linear regression result (the regression result is: y_mean=0.80x+84.80, y_median=0.78x+72.31, where x indicates the monthly the empirical data, and the y_{mean} and y_{median} specify values from the model). The slopes of these two regression lines are 0.80 and 0.78. Theoretically, the best slope is 1.0 – that is, one would hope for the model predictions of case count in the next month to very closely match the empirically reported case count for that month.

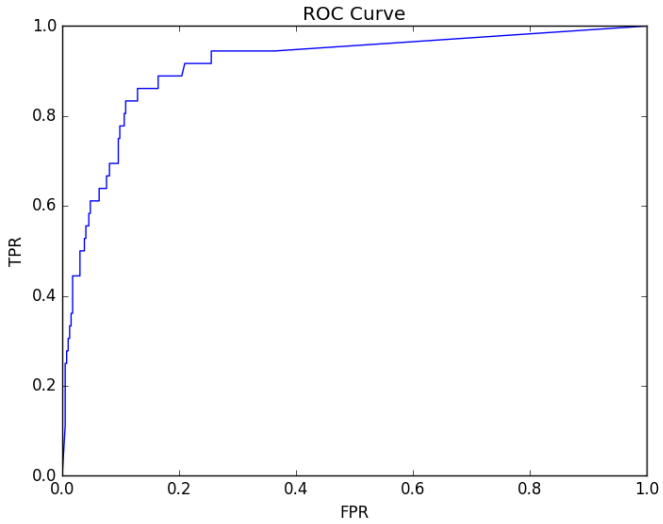


Figure 6.1: ROC curve of the prediction classification result of the minimum discrepancy model of measles. AUC is 0.893.

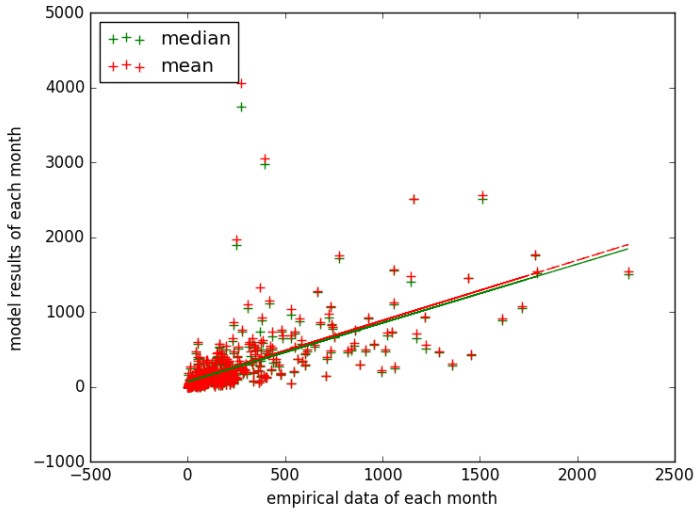


Figure 6.2: Scatter plot and regression result of the empirical data vs. mean and median data of the model calculated results of the minimum discrepancy model of measles.

6.3.2 Results of pertussis

Similar to the analysis of outbreak prediction for the particle filtering model of measles above, I perform a classification-based prediction of whether a pertussis outbreak will break out or not in the next month, by incorporating the prediction results of the lowest discrepancy particle filter pertussis model, previously noted to be PF_{age_2}. Figure 6.3 displays the ROC curve showing the prediction results. The Area Under the Curve (AUC) of the ROC curve is 0.913.

Figure 6.4 displays the scatter plot between the monthly empirical data and the mean and median of the model predicted next month results over all sampled particles (the regression result is: y_mean=0.52x+27.2, y_median=0.51x+25.0). The slopes of these two regression line are 0.52 (mean) and 0.51 (median), which are far smaller than the best value – 1.0. I think there are two reasons. Firstly, there are two outliers in the monthly empirical data of pertussis – at time-point (month) 8 and 165. Thus, the predicted results at these two time are far smaller than the empirical data. Secondly, the particle filtering model is under-estimated at several points especially where the empirical data are greater than 200. However, Figure 6.4 shows that the mean and median values across all the sampled particles at the time where the empirical data are less than 200 are located at a better range – close to the the value of the empirical data at the same time-point.

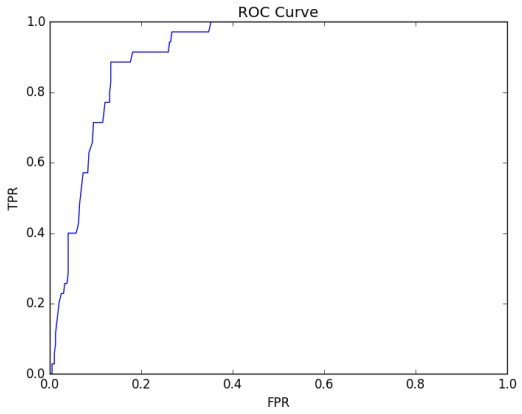


Figure 6.3: ROC curve of the prediction classification Result of the minimum discrepancy model of pertussis. AUC is 0.913

6.4 Discussion and conclusion

In this chapter, I have performed prediction and prospective outbreak classification analysis, based on the most competitive models as evaluated by predictive accuracy of measles and pertussis, which have been introduced in the previous chapters (from chapter 3 to chapter 5). The prediction results demonstrated that the particle filtering models could predict the transmission patterns and classify whether there will be an outbreak or not in the next month to considerable degree of accuracy (area under the ROC Curve is 0.89 for measles, and 0.91 for pertussis). I conclude that anticipating the outbreak dynamics of measles and pertussis

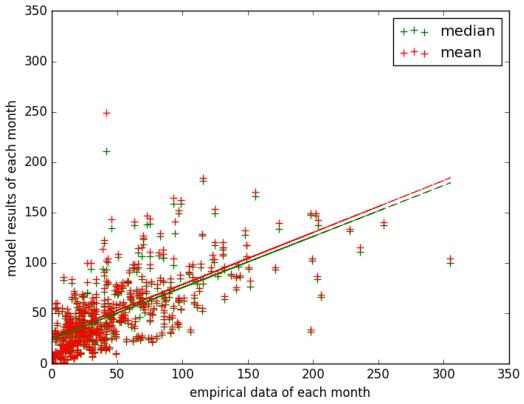


Figure 6.4: Scatter plot and regression result of the empirical data vs. mean and median data of the model calculated results of the minimum discrepancy model of pertussis.

in low vaccination regions by applying particle filtering with simple transmission models, and incorporating time series of reported case counts, is a valuable technique to assist public health authorities in estimating risk and magnitude of outbreaks.

It is notable that while the classification method I employed here is quite naive – only employing a single thresholding approach to classify the category of outbreak and non-outbreak, the results of the classification are acceptable both for the measles and pertussis outbreak classification prediction. However, the strong results demonstrated for even such a simple model suggest that it is worth investigating more sophisticated machine learning classification methods to perform the outbreak classification prediction – a priority constituting an important element of my future work related to this chapter.

Moreover, while the classification results of the next month prediction have demonstrated high accuracy, it indicates that it is possible to perform the research of the classification of the outbreak prediction related to a longer prediction period, such as next three months, half year, even one year. Such longer-range predictions can further contribute to priorities for future work in classification research with regards to outbreak prediction results of both measles and pertussis particle filtering models.

Chapter 7

Conclusion

7.1 Discussion and conclusion

In this thesis, a new method for tracking patterns of measles and pertussis outbreaks in low vaccination regions is presented by applying particle filtering with measles and pertussis compartmental transmission models, and incorporating empirical data on case reports. Particle filtering offers many attractive features for epidemiological models. Firstly, it relaxes the stiff assumptions of normality with respect to the process and observation noise required by the older statistical filtering techniques of Kalman Filtering and its variants; such assumptions are often particularly problematic in epidemiological contexts with small sample counts. Secondly, particle filtering is especially well suited to non-linear models such as that used here, because it foregoes a focus on a single Maximum Likelihood Estimate seen in the Kalman Filter – which can be particularly problematic in the context of state uncertainty that can span multiple basins of attraction – and instead samples from a distribution of possible states for a given time-point.

In this study, the particle filtering algorithm has mitigated significant weaknesses and simplifications associated with aggregate compartmental models and noisy empirical data. By incorporating ongoing arriving empirical data, the particle filter model has the capability to correct for distortions that accompany compartmental model aggregation, such as assumptions of random mixing and homogeneity. It is notable that while age-structured compartmental models capture heterogeneous mixing among the different age groups, the individuals within the same age group are engaged in homogeneous mixing. In the datasets examined here, the particle filter offered strong performance in estimating the outbreak pattern of measles and pertussis and predicting future trends.

A key benefit of particle filtering lies in its capacity to estimate the latent state of the system – the state that cannot be directly measured, but which is jointly implied by the combination of empirical time series and the hypothesized structure of the system, as captured in the mathematical model. It is important to stress that a key motivation for conducting particle filtering to infer latent state in this way lies in the fact that reliable understanding of such latent state is important for estimating the impact of interventions enacted at that point. By estimating the latent state of the system using particle filtering, we can then conduct "what if" scenarios forward from that point, each of which characterize the effects of distinct interventions. Accurate estimation of the state of the system prior to initiation of different intervention strategies will frequently be

an important enabler for accurately assessing the differential effects of those interventions.

The particle filtering algorithm in general also has important limitations, such as information loss and particle collapse [27] during the evolutionary process of the particles. These limitations are also inherited in the algorithm applying in this paper that combines particle filtering with a compartmental transmission model. Moreover, the treatment here further suffers from some additional challenges. For simplicity, the condensation algorithm is employed in calculating the proposal distribution. However, this algorithm may contribute to a loss of the diversity of particles. These limitations could be relieved by employing more particles during the calculation, or restricting particle dispersion by selecting the values of parameters in more tightly informed ranges. Moreover, in high-dimensional systems with complex patterns of probability density functions, the variance of the density functions may be high, because it is difficult to be represent by simple probability density functions. Thus, the particle filtering method with the condensation algorithm in this thesis maybe difficult in simulating a high-dimensional systems. Finally, a key limitation in terms of practical implications for the findings in the context of the developed world reflects the fact that we have focused on prediction in a non-vaccine context; there remains an important uncertainty as to the degree to which the approach proposed here will offer high predictive capacity in the context of sporadic, low attack rate outbreaks characteristic of measles and pertussis epidemiology in developed countries within the vaccination era.

It is notable that there are distinctive different consequences of the stochastic processes added in the state space model during the estimation period and prediction period. In the estimation period in which successive elements of the empirical datasets are incorporated by the particle filtering algorithm, if these stochastic processes are more noisy (i.e., exhibit larger stochastic fluctuations), particle filtering models find it easier to trace the empirical datasets, given a sufficient particle count. This reflects the fact that when a new datapoint arrives, the support of the distribution of particle values is sufficiently broad that there are likely to be particles closer in value to the observed particle, allowing the distribution of particles to shift towards consistency with the new datapoint. However, during the prediction period, if the stochastic processes are too noisy, the support of the distribution of particle values within the particle filtering model will grow very rapidly with increasing time since the last observation was reserved. This can allow the model to rapidly become highly uncertain about the values of coming empirical data points. For a properly performing particle filter to be produced, it is thus particularly important that the stochastic processes in the state space model should be controlled in a proper range. Achieving that balance is likely to require some tuning.

When estimation of the initial values of each stock in the system dynamic models, the initial values in the age-structured population models are estimated both manually and by the particle filtering algorithm. Specifically, the population distribution among the different age groups are tuned manually, while the population distribution among different stocks within an age group is estimated by the particle filtering algorithm by setting the initial values of stocks in a proper range following a uniform distribution. As is the case for

the rates of stochastic variations seen in the model, the design of the initial state of the types of particle filtering models examined in this thesis can also require considerable manual tuning.

The work of the thesis establishes that the hypothesis advanced in the thesis statement has been successfully tested, and that it does hold. Firstly, the 2D histogram plots with the particle filtering models' output data (the monthly or yearly reported cases of measles or pertussis) and the empirical data indicates that the high probability density region of the model's prediction of empirical data lies near those data, especially compared with the results of the output of the traditional calibration models. Secondly, the discrepancy of the measles particle filtering model's predictions vs. observed data is reduced with a factor of 2.0, while the discrepancy of the pertussis particle filtering model's predictions vs. observed data is reduced with a factor of 1.6. This proved that the incorporating particle filtering algorithm could build better models compared with the traditional calibration methods.

I conclude that anticipating the outbreak pattern of measles and pertussis in low vaccination regions by applying particle filtering with simple measles and pertussis transmission models so as to recurrently incorporate successive elements of time series of reported case counts is a valuable technique to assist public health authorities in estimating risk and magnitude of measles and pertussis outbreaks. Such approaches offer a particularly strong value proposition for other pathogens with little-known dynamics, latent drivers, and in the context of the growing emergence of high-velocity electronic data sources. While it remains to be evaluated, we anticipate that additional strong benefits will be realized by extending the application of this technique to highly vaccinated populations.

Finally, it is notable that all the models in this thesis are built using Anylogic 8.1.0 software, a Java-based simulation platform. When compared to their non-particle filtered analogue, particle filtering models are much more resource and time consuming, especially for the models with high-dimensional states. Normally, the running time grows linearly with the increasing of the number of particles. By contrast, the corresponding run time of deterministic models with both measles and pertussis would complete less than a minute.

7.2 Future work

Much work remains to be undertaken. While particle filtering techniques investigated in this thesis have immediate application in populations with low vaccine coverage (including isolated pockets of population or individuals who refuse vaccination in jurisdictions with otherwise high vaccination coverage), in the future, I will consider vaccination state in the measles and pertussis particle filter model, so as to support simulation of the outbreak patterns in high vaccination regions in the vaccination era1. Such a model could be helpful for predicting outbreaks of measles and pertussis in regions suffering from borderline or waning vaccination rates due to vaccine hesitancy, health disparities or other causes. I further plan to apply more powerful techniques, such as Particle Markov Chain Monte Carlo methods that can allow for jointly estimating (via sampling) the latent state of the model and static parameter values whose values are poorly known.

Moreover, I also plan to investigate more sophisticated means of predicting outbreak occurrence based on particle filtering results. It appears likely that such refinements will further enhance the already strong advantages conferred by particle filtering methods and variants in measles and pertussis transmission modeling.

Finally, I think it would be helpful in the estimation of the contact matrix to build a single particle filtering model for both measles and pertussis that uses a common mixing matrix and demographic structure, while matching empirical data for both pertussis and measles. Compared with the particle filtering models of measles and pertussis that have been built already, this model would support using twice the amount of data to be used in support of estimation of the contact matrix.

7.3 Contributions

The first contribution of this thesis lies in conducting the first known experiments with particle filtering on both measles and pertussis transmission models to enhance the accuracy of the prediction of measles or pertussis in the context of incoming epidemiological data. Moreover, both aggregate and age-structured population particle filtering models of measles and pertussis have been researched.

Secondly, this thesis provides the first known application of particle filtering algorithm to state space models corresponding to age-structured population models with a relatively large number of age groups. Although Dureau *et al.* (2013) [28] implemented the particle filtering algorithm with a state space model with 2 age groups, this research applies particle filtering in a complex age-structured model with 32 age groups.

Thirdly, several different mathematical methods to calculate the contact matrix of the age-structured compartmental measles and pertussis models have been presented in this thesis. It is particularly notable that a new method which could decrease the number of free parameters in calculating the contact matrix with a relatively large number of age groups is explored in this thesis. For a number of age groups n, this method reduces the count of such parameters from $O(n^2)$ to $O(n)$.

Fourthly, particle filtering supports estimation via sampling from the entire state of the corresponding dynamic model – including both latent and observable states – for each time point, incorporating the empirical datasets as they occur. This advantage can help the researchers to perform researches related to those latent states, including so as to better understand the current epidemiological context, and to better understand historical epidemiological situations.

Fifthly, the thesis constitutes the first time that the particle filtering has been used to simulate the intervention strategies for health models. The intervention experiments investigated here indicate that particle filtered measles and pertussis models are capable of evaluating interventions by examining the outcome of potential intervention strategies using model results. In the current investigations, such interventions were characterized naively by alternation of model parameters – such as the contact rate and the total number of

individuals in each stocks – but the same kinds of investigations could be undertaken with respect to more sophisticated intervention investigations.

The last but not the least contribution is that this thesis proposes, applies and evaluates the performance of a binary classification method of predicting whether there will be an outbreak in the next time (Month in this research), with the data generated from the particle filtering models. It is notable that although the evaluation suggests positive performance, the generality of the approach further suggests ready extension.

7.4 Publications of this thesis

Till the date of the submission of this thesis, an article has been published in the Journal of PLoS ONE, based on the content of chapter 3, chapter 4 and the measles part of chapter 6 in 2nd November, 2018. The information of this article is listed as follows:

Li X, Doroshenko A, Osgood ND (2018) Applying particle filtering in both aggregated and age-structured population compartmental models of pre-vaccination measles. PLoS ONE 13(11): e0206529. https://doi.org/10.1371/journal.pone.0206529

References

[1] R M Anderson and R M May. Population biology of infectious diseases: Part I. *Nature*, 280(5721):361–367, August 1979.

[2] Roy M Anderson, Robert M May, and B Anderson. *Infectious diseases of humans: dynamics and control*, volume 28. Wiley Online Library, 1992.

[3] Christophe Andrieu, Nando de Freitas, Arnaud Doucet, and Michael I Jordan. An introduction to MCMC for machine learning. *Mach. Learn.*, 50(1):5–43, January 2003.

[4] Christophe Andrieu, Arnaud Doucet, and Roman Holenstein. Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 72(3):269–342, June 2010.

[5] J Arino, C McCluskey, and P van den Driessche. Global results for an epidemic model with vaccination that exhibits backward bifurcation. *SIAM journal on applied mathematics*, 64(1):260–276, January 2003.

[6] M S Arulampalam, S Maskell, N Gordon, and T Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE transactions on signal processing: a publication of the IEEE Signal Processing Society*, 50(2):174–188, February 2002.

[7] Amy H Auchincloss and Ana V Diez Roux. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *American journal of epidemiology*, 168(1):1–8, July 2008.

[8] Norman T J Bailey. The mathematical theory of epidemics. Technical report, 1957.

[9] Sylvain Baize, Delphine Pannetier, Lisa Oestereich, Toni Rieger, Lamine Koivogui, N'faly Magassouba, Barrè Soropogui, Mamadou Saliou Sow, Sakoba Keïta, Hilde De Clerck, Amanda Tiffany, Gemma Dominguez, Mathieu Loua, Alexis Traoré, Moussa Kolié, Emmanuel Roland Malano, Emmanuel Heleze, Anne Bocquin, Stephane Mély, Hervé Raoul, Valérie Caro, Dániel Cadar, Martin Gabriel, Meike Pahlmann, Dennis Tappe, Jonas Schmidt-Chanasit, Benido Impouma, Abdoul Karim Diallo, Pierre Formenty, Michel Van Herp, and Stephan Günther. Emergence of zaire ebola virus disease in guinea. *N. Engl. J. Med.*, 371(15):1418–1425, October 2014.

[10] David Barber, A Taylan Cemgil, and Silvia Chiappa. Inference and estimation in probabilistic time series models. *Bayesian Time Series Models*, 1, 2011.

[11] M S Bartlett. Deterministic and stochastic models for recurrent epidemics. In *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, volume 4, page 109, 1956.

[12] Daniel Bernoulli. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. *Histoire de l'Acad. , Roy. Sci. (Paris) avec Mem*, pages 1–45, 1760.

[13] Ottar N Bjørnstad, Bärbel F Finkenstädt, and Bryan T Grenfell. Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecological monographs*, 72(2):169–184, 2002.

[14] Andrew Blake and Michael Isard. The condensation algorithm-conditional density propagation and applications to visual tracking. In *Advances in Neural Information Processing Systems*, pages 361–367, 1997.

[15] B Bolker and B Grenfell. Space, persistence and dynamics of measles epidemics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 348(1325):309–320, May 1995.

[16] Statistics Canada. Historical age pyramid. `http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/pyramid/pyramid.cfm?geo1=47&type=1`.

[17] O Cappe, S J Godsill, and E Moulines. An overview of existing methods and recent advances in sequential monte carlo. *Proceedings of the IEEE*, 95(5):899–924, May 2007.

[18] Shi Chen, John Fricks, and Matthew J Ferrari. Tracking measles infection through non-linear state space models: Tracking measles infection. *J. R. Stat. Soc. Ser. C Appl. Stat.*, 61(1):117–134, January 2012.

[19] Maria Yui Kwan Chow, Gulam Khandaker, and Peter McIntyre. Global childhood deaths from pertussis: A historical review. *Clin. Infect. Dis.*, 63(suppl 4):S134–S141, December 2016.

[20] Suruchi Deodhar, Keith R Bisset, Jiangzhuo Chen, Yifei Ma, and Madhav V Marathe. An interactive, web-based high performance modeling environment for computational epidemiology. *ACM Transactions on Management Information Systems (TMIS)*, 5(2):7, 2014.

[21] Department of Public Health, Saskatchewan, CA. *Annual Report*, 1921-1956.

[22] O Diekmann and J A P Heesterbeek. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. John Wiley & Sons, April 2000.

[23] Ilaria Dorigatti, Simon Cauchemez, Andrea Pugliese, and Neil Morris Ferguson. A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: Application to the italian 2009–2010 A/H1N1 influenza pandemic. *Epidemics*, 4(1):9–21, March 2012.

[24] Alexander Doroshenko, Weicheng Qian, and Nathaniel D Osgood. Evaluation of outbreak response immunization in the control of pertussis using agent-based modeling. *PeerJ*, 4:e2337, August 2016.

[25] Arnaud Doucet, Nando de Freitas, and Neil Gordon. An introduction to sequential monte carlo methods. In Arnaud Doucet, Nando de Freitas, and Neil Gordon, editors, *Sequential Monte Carlo Methods in Practice*, pages 3–14. Springer New York, New York, NY, 2001.

[26] Arnaud Doucet and Adam M Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering*, 12(656-704):3, 2009.

[27] Vanja Dukic, Hedibert F Lopes, and Nicholas G Polson. Tracking epidemics with google flu trends data and a State-Space SEIR model. *Journal of the American Statistical Association*, 107(500):1410–1426, December 2012.

[28] Joseph Dureau, Konstantinos Kalogeropoulos, and Marc Baguelin. Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. *Biostatistics*, 14(3):541–555, July 2013.

[29] Jonathan Dushoff, Joshua B Plotkin, Simon A Levin, and David J D Earn. Dynamical resonance can account for seasonality of influenza epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, 101(48):16915–16916, November 2004.

[30] D J Earn, P Rohani, B M Bolker, and B T Grenfell. A simple model for complex dynamical transitions in epidemics. *Science*, 287(5453):667–670, January 2000.

[31] Centers for Disease Control and Prevention. Causes and transmission. `https://www.cdc.gov/pertussis/about/causes-transmission.html`.

[32] Centers for Disease Control and Prevention. Epidemiology and prevention of vaccine-preventable diseases. https://www.cdc.gov/vaccines/pubs/pinkbook/pert.html.

[33] Centers for Disease Control and Prevention. Measles data and statistics. https://www.cdc.gov/measles/downloads/MeaslesDataAndStatsSlideSet.pdf.

[34] Centers for Disease Control and Prevention. Measles history in centers for disease control and prevention. https://www.cdc.gov/measles/about/history.html.

[35] Centers for Disease Control and Prevention. Signs and symptoms. https://www.cdc.gov/measles/about/signs-symptoms.html.

[36] Centers for Disease Control and Prevention. Transmission of measles. https://www.cdc.gov/measles/about/transmission.html.

[37] Centers for Disease Control and Prevention. Whooping cough and the vaccine (shot) to prevent it. https://www.cdc.gov/vaccines/parents/diseases/child/pertussis.html.

[38] E Frias-Martinez, G Williamson, and V Frias-Martinez. An Agent-Based model of epidemic spread using human mobility and social network information. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pages 57–64, October 2011.

[39] G P Garnett and F J Bowden. Epidemiology and control and curable sexually transmitted diseases: opportunities and problems. *Sex. Transm. Dis.*, 27(10):588–599, November 2000.

[40] Anne Gershon. Measles virus (rubeola) in infectious disease advisor. https://www.infectiousdiseaseadvisor.com/infectious-diseases/measles-virus-rubeola/article/610297/.

[41] Walter R Gilks and Carlo Berzuini. Following a moving target—monte carlo inference for dynamic bayesian models. *J. R. Stat. Soc. Series B Stat. Methodol.*, 63(1):127–146, 2001.

[42] N J Gordon, D J Salmond, and A F M Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F (Radar and Signal Processing)*, 140(2):107–113, April 1993.

[43] B T Grenfell. Chance and chaos in measles dynamics. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 54(2):383–398, 1992.

[44] Bryan T Grenfell, Ottar N Bjørnstad, and Bärbel F Finkenstädt. Dynamics of measles epidemics: scaling noise, determinism, and predictability with the TSIR model. *Ecological monographs*, 72(2):185–202, May 2002.

[45] W N Hammer. Epidemic disease in England-The evidence of variability and the persistence of type. *The Lancet, II*, pages 733–739, 1906.

[46] healthychildren.org. Overview of infectious diseases. https://www.healthychildren.org/English/health-issues/conditions/infections/Pages/Overview-of-Infectious-Diseases.aspx.

[47] Soren Henriksen, Adrian Wills, Thomas B Schön, and Brett Ninness. Parallel implementation of particle MCMC methods on a GPU. *IFAC Proceedings Volumes*, 45(16):1143–1148, 2012.

[48] H Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, January 2000.

[49] H W Hethcote. Measles and rubella in the united states. *American journal of epidemiology*, 117(1):2–13, January 1983.

[50] H W Hethcote. An age-structured model for pertussis transmission. *Mathematical biosciences*, 145(2):89–136, October 1997.

[51] Herbert W Hethcote and James A Yorke. *Gonorrhea Transmission Dynamics and Control*. Lecture Notes in Biomathematics. 1984.

[52] Y Ho and R Lee. A bayesian approach to problems in stochastic estimation and control. *IEEE Trans. Automat. Contr.*, 9(4):333–339, October 1964.

[53] Jack B Homer and Gary B Hirsch. System dynamics modeling for public health: background and opportunities. *Am. J. Public Health*, 96(3):452–458, March 2006.

[54] F Hoppensteadt. Mathematical theories of populations: Demographics, genetics and epidemics, SIAM reg. In *Conf. Series in Appl. Math*, 1975.

[55] Weipeng Huang and Gregory Provan. An improved state filter algorithm for SIR epidemic forecasting. In *ECAI 2016: 22nd European Conference on Artificial Intelligence*. cora.ucc.ie, 2016.

[56] Timothy P Hubbard and Yigit Saglam. Stochastic processes, itô calculus, and applications in economics. *Lecture notes, Department of Mathematics, University of Iowa*, 2006.

[57] M Isard and A Blake. Condensation–Conditional density propagation for visual tracking,(1998). *International Journal of Computer Vision Publ*, pages 5–28.

[58] A H Jazwinski. Stochastic processes and filtering theory (academic, new york, 1970). *Google Scholar*, 1981.

[59] Yu Jin, Wendi Wang, and Shiwu Xiao. An SIRS model with a nonlinear incidence rate. *Chaos Solitons Fractals*, 34(5):1482–1497, December 2007.

[60] Mcd Jong, O Diekmann, and H Heesterbeek. How does transmission of infection depend on population size? epidemic models. *Publication of the Newton Institute*, pages 84–94, 1995.

[61] W O Kermack and A G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Mathematical and physical sciences*, 115(772):700–721, August 1927.

[62] Genshiro Kitagawa and Seisho Sato. Monte carlo smoothing and Self-Organising State-Space model. In Arnaud Doucet, Nando de Freitas, and Neil Gordon, editors, *Sequential Monte Carlo Methods in Practice*, pages 177–195. Springer New York, New York, NY, 2001.

[63] James S Koopman. Infection transmission science and models. *Japanese journal of infectious diseases*, 58(6):S3–8, December 2005.

[64] K Kreuger and N Osgood. Particle filtering using agent-based transmission models. In *2015 Winter Simulation Conference (WSC)*, pages 737–747, December 2015.

[65] Anthony Lee, Christopher Yau, Michael B Giles, Arnaud Doucet, and Christopher C Holmes. On the utility of graphics cards to perform massively parallel simulation of advanced monte carlo methods. *Journal of computational and graphical statistics: a joint publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America*, 19(4):769–789, December 2010.

[66] Fengchen Liu, Wayne T A Enanoria, Jennifer Zipprich, Seth Blumberg, Kathleen Harriman, Sarah F Ackley, William D Wheaton, Justine L Allpress, and Travis C Porco. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for california. *BMC public health*, 15(1):447, May 2015.

[67] L Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Trans. Automat. Contr.*, 24(1):36–50, February 1979.

[68] A J Lotka. The stability of the normal age distribution. *Proc. Natl. Acad. Sci. U. S. A.*, 8(11):339–345, November 1922.

[69] C M Macal and M J North. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 4(3):151–162, September 2010.

[70] Charles Macal and Michael North. Introductory tutorial: Agent-based modeling and simulation. In *Proceedings of the 2014 Winter Simulation Conference*, WSC '14, pages 6–20, Piscataway, NJ, USA, 2014. IEEE Press.

[71] Ishikawa Masaaki. Optimal strategies for vaccination using the stochastic SIRV model. , 25(12):343–348, 2012.

[72] Hamish McCallum, Nigel Barlow, and Jim Hone. How should pathogen transmission be modelled? *Trends in ecology & evolution*, 16(6):295–300, June 2001.

[73] Robert Moss, Alexander Zarebski, Peter Dawson, and James M McCaw. Forecasting influenza outbreak dynamics in melbourne from internet search query surveillance data. *Influenza and other respiratory viruses*, 10(4):314–323, July 2016.

[74] Kevin P Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, August 2012.

[75] Christian Musso, Nadia Oudjane, and Francois Le Gland. Improving regularised particle filters. In Arnaud Doucet, Nando de Freitas, and Neil Gordon, editors, *Sequential Monte Carlo Methods in Practice*, pages 247–271. Springer New York, New York, NY, 2001.

[76] L F Olsen, G L Truty, and W M Schaffer. Oscillations and chaos in epidemics: a nonlinear dynamic study of six childhood diseases in copenhagen, denmark. *Theor. Popul. Biol.*, 33(3):344–370, June 1988.

[77] Jimmy Boon Som Ong, Mark I-Cheng Chen, Alex R Cook, Huey Chyi Lee, Vernon J Lee, Raymond Tzer Pin Lin, Paul Ananth Tambyah, and Lee Gan Goh. Real-time epidemic monitoring and forecasting of H1N1-2009 using influenza-like illness from general practice and family doctor clinics in singapore. *PloS one*, 5(4):e10036, April 2010.

[78] R Oraji, V H Hoeppner, A Safarishahrbijari, and N D Osgood. Combining particle filtering and transmission modeling for TB control. In *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 392–398. ieeexplore.ieee.org, October 2016.

[79] World Health Organization. Pertussis, vaccines and diseases, immunization, vaccines and biologicals. `http://www.who.int/immunization/diseases/pertussis/en/`.

[80] World Health Organization. World health organization measles - fact sheet. `http://www.who.int/mediacentre/factsheets/fs286/en/`, 2017.

[81] Nathaniel Osgood. Representing heterogeneity in complex feedback system modeling: computational resource and error scaling. In *22nd International Conference of the System Dynamics Society*, 2004.

[82] Nathaniel Osgood. Using traditional and agent based toolsets for system dynamics: present tradeoffs and future evolution. *System Dynamics*, 2007.

[83] Nathaniel Osgood. Representing progression and interactions of comorbidities in aggregate and individual-based systems models. In *Proceedings of the 27th International Conference of the System Dynamics Society. Albuquerque, New Mexico*, 2009.

[84] Nathaniel Osgood and Juxin Liu. Towards closed loop modeling: Evaluatng the prospects for creating recurrently regrounded aggregate simulation models using particle filtering. In *Proceedings of the 2014 Winter Simulation Conference*, WSC '14, pages 829–841, Piscataway, NJ, USA, 2014. IEEE Press.

[85] Jsm Peiris, C M Chu, Vcc Cheng, K S Chan, Ifn Hung, Llm Poon, K I Law, Bsf Tang, Tyw Hon, C S Chan, K H Chan, Jsc Ng, B J Zheng, W L Ng, Rwm Lai, Y Guan, and K Y Yuen. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet*, 361(9371):1767–1772, May 2003.

[86] Liliana Perez and Suzana Dragicevic. An agent-based approach for modeling dynamics of contagious disease spread. *International journal of health geographics*, 8:50, August 2009.

[87] Michael K Pitt and Neil Shephard. Filtering via simulation: Auxiliary particle filters. *J. Am. Stat. Assoc.*, 94(446):590–599, June 1999.

[88] Weicheng Qian, Nathaniel D Osgood, and Kevin G Stanley. Integrating epidemiological modeling and surveillance data feeds: A kalman filter based approach. In *Social Computing, Behavioral-Cultural Modeling and Prediction*, pages 145–152. Springer International Publishing, 2014.

[89] L R Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE*, 77(2):257–286, February 1989.

[90] Hazhir Rahmandad and John Sterman. Heterogeneity and network structure in the dynamics of diffusion: Comparing Agent-Based and differential equation models. *Management science*, 54(5):998–1014, May 2008.

[91] Carmen L Vidal Rodeiro and Andrew B Lawson. Online updating of space-time disease surveillance models via particle filters. *Statistical methods in medical research*, 15(5):423–444, 2006.

[92] P Rohani, D J Earn, B Finkenstädt, and B T Grenfell. Population dynamic interference among childhood diseases. *Proc. Biol. Sci.*, 265(1410):2033–2041, November 1998.

[93] Ronald Ross. *The prevention of malaria*. John Murray; London, 1911.

[94] D B Rubin. Using the SIR algorithm to simulate posterior distributions. *Bayesian statistics*, 1988.

[95] A Safarishahrbijari, T Lawrence, R Lomotey, J Liu, C Waldner, and N Osgood. Particle filtering in a SEIRV simulation model of H1N1 influenza. In *2015 Winter Simulation Conference (WSC)*, pages 1240–1251. ieeexplore.ieee.org, December 2015.

[96] Erinn Sanstead, Cynthia Kenyon, Seth Rowley, Eva Enns, Claudia Miller, Kristen Ehresmann, and Shalini Kulasingam. Understanding trends in pertussis incidence: An Agent-Based model approach. *American journal of public health*, 105(9):e42–7, September 2015.

[97] D Schenzle. An age-structured model of pre- and post-vaccination measles transmission. *IMA journal of mathematics applied in medicine and biology*, 1(2):169–191, 1984.

[98] Bernard W Silverman. *Density estimation for statistics and data analysis*. Routledge, 2018.

[99] H E Soper. The interpretation of periodicity in disease prevalence. *Journal of the Royal Statistical Society*, 92(1):34–73, 1929.

[100] Marc A Suchard, Quanli Wang, Cliburn Chan, Jacob Frelinger, Andrew Cron, and Mike West. Understanding GPU programming for statistical computation: Studies in massively parallel massive mixtures. *Journal of computational and graphical statistics: a joint publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America*, 19(2):419–438, January 2010.

[101] F S Tabataba, B Lewis, M Hosseinipour, F S Tabataba, S Venkatramanan, J Chen, D Higdon, and M Marathe. Epidemic forecasting framework combining Agent-Based models and smart beam particle filtering. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 1099–1104, November 2017.

[102] W Tang and W P Tay. A particle filter for sequential infection source estimation. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4094–4098. ieeexplore.ieee.org, March 2017.

[103] the Pan American Health Organization and the World Health Organization. Fact sheet of measles elimination in the americas. `https://www.paho.org/hq/index.php?option=com_content&view=article&id=12526&Itemid=40721&lang=en`.

[104] Milton C Weinstein. Recent developments in decision-analytic modelling for economic evaluation. *PharmacoEconomics*, 24(11):1043–1053, 2006.

[105] H J Whitaker and C P Farrington. Infections with varying contact rates: application to varicella. *Biometrics*, 60(3):615–623, September 2004.

[106] S Wolfe. Epidemiology and prevention of vaccine-preventable diseases. `https://www.cdc.gov/vaccines/pubs/pinkbook/downloads/table-of-contents.pdf`, April 2015. Editors: Hamborsky, J and Kroger, A . US Department of Health & Human Services, Centers for Disease Control and Prevention.

[107] Wan Yang, Alicia Karspeck, and Jeffrey Shaman. Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. *PLoS computational biology*, 10(4):e1003583, April 2014.

Appendix A

Proof of the n-square grows of the unknown parameters

In this part, I prove that the unknown parameters grows with n-squared with the total number of age groups in the model.

The contact matrix has been introduced in chapter 2 , which is

$$\begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix} \circ \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{bmatrix} \tag{A.1}$$

where $\circ$ indicates the Hadamard (element-wise) product; the parameter of l_i $(1 \leq i \leq n)$ is the contact rate of age group i. In this research, the l_i is known variables; the parameter of f_{ij} $(1 \leq i \leq n, 1 \leq j \leq n)$ indicates the fraction of the age group j of the contact rate of the age group i.

The f_{ij} are normally unknown. And the total number of f_{ij} is n^2. However, there are two relationships under this method. One relationship is that the sum of the fraction to all the age groups of the age group (e.g. i) is 1.0. The other relationship, related to the characteristics of balance of the contact matrix, is that the total contacts of the age group i to the age group j should be equal to the total contacts of the age group j to the age group i. Based on these two relationships, two equations could be generated as follows:

$$\begin{aligned} &\sum_{j=1}^{n} f_{ij} = 1 \\ &N_i l_i f_{ij} = N_j l_j f_{ji} \end{aligned} \tag{A.2}$$

the total number of equations in Equation (A.2) is $n + \binom{2}{n} = n + \frac{n(n-1)}{2} = \frac{n^2+n}{2}$. Finally, in this method of calculating the contact matrix, the number of unknown parameters is $\frac{n^2-n}{2}$. It indicates that the number of the unknown parameters grows in n-squared with the total number of age groups (n) in the model.

Appendix B

Initial values of stocks in the particle filtering models of pertussis

Table B.1: Table showing initial values of the stocks in the pertussis two-age-groups particle filtering model.

Parameter	**Value**	**Unit**
S_{c0}	Uniform[500, 35000)	Person
S_{a0}	Uniform[10, 10000)	Person
I_{c0}	Uniform[30, 2500)	Person
I_{a0}	Uniform[0, 500)	Person
I_{mc0}	50	Person
I_{ma0}	50	Person
I_{wc0}	100	Person
I_{wa0}	100	Person
R_{1c0}	Uniform[5, 10000)	Person
R_{1a0}	Uniform[0, 10000)	Person
R_{2c0}	10000	Person
R_{2a0}	10000	Person
R_{3c0}	10000	Person
R_{3a0}	10000	Person
R_{4c0}	$N_c - S_{c0} - I_{c0} - I_{mc0} - I_{wc0} - R_{1c0} - R_{2c0} - R_{3c0}$	Person
R_{4a0}	$N_a - S_{a0} - I_{a0} - I_{ma0} - I_{wa0} - R_{1a0} - R_{2a0} - R_{3a0}$	Person
β_c	Uniform[5,100)	Person/Month
M_a	Uniform[5,100)	Dimensionless
C_r	Uniform[0,0.2)	Dimensionless
f_{cc}	Uniform[0,0.2)	Dimensionless

Table B.2: Table showing initial values of the stocks in the pertussis 32-age-groups particle filtering models.

Parameter	Value	Unit
$S_{i0} \quad 1 \leq i \leq 32$	Uniform[1000, 3000), Uniform[1000, 3000), Uniform[1000, 3000), Uniform[1000, 9000), Uniform[1000, 10000), Uniform[1000, 10000), Uniform[1000, 10000), Uniform[100, 5000), Uniform[100, 2000), Uniform[100, 2000), Uniform[100, 2000), Uniform[100, 2000), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[10, 500), Uniform[100, 2000), Uniform[100, 2000), Uniform[100, 2000), Uniform[100, 5000), Uniform[100, 5000), Uniform[0, 2000), Uniform[0, 1000), Uniform[0, 500), Uniform[0, 100)	Person
$I_{i0} \quad 1 \leq i \leq 32$	Uniform[0, 10), Uniform[0, 10), Uniform[0, 10), Uniform[0, 10), Uniform[0, 20), Uniform[0, 20), Uniform[0, 20), Uniform[0, 20), Uniform[0, 20), Uniform[0, 20), Uniform[0, 20), Uniform[0, 20), Uniform[0, 10), Uniform[0, 10), Uniform[0, 10), Uniform[0, 10), Uniform[0, 10), 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0	Person
$I_{mi0} \quad 1 \leq i \leq 32$	5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0	Person
$I_{wi0} \quad 1 \leq i \leq 32$	10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0	Person
$R_{1i0} \quad 1 \leq i \leq 32$	0, 0, 0, 0, Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 20000), Uniform[0, 20000), Uniform[0, 20000), Uniform[0, 20000), Uniform[0, 20000), Uniform[0, 20000), Uniform[0, 2000), Uniform[0, 2000), Uniform[0, 100)	Person
$R_{2i0} \quad 1 \leq i \leq 32$	0, 0, 0, 0, 100, 100, 100, 100, 100, 100, 100, 100, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 10000, 10000, 10000, 10000, 10000, 10000, 5000, 500, 50	Person
$R_{3i0} \quad 1 \leq i \leq 32$	0, 0, 0, 0, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 6000, 10000, 10000, 10000, 10000, 6000, 5000, 0, 0, 0	Person
$R_{4i0} \quad 1 \leq i \leq 32$	$N_i - S_{i0} - I_{i0} - I_{mi0} - I_{wi0} - R_{1i0} - R_{2i0} - R_{3i0}$	Person
l_1	Uniform[0.001, 0.5)	Person/Day
C_r	Uniform[0, 0.15)	Dimensionless
$\epsilon_i \quad 1 \leq i \leq 6$	Uniform[0, 1)	Dimensionless