A Mechanism-based Outbreak Projection Study of Pertussis (Whooping Cough): Combining Particle Filtering and Compartmental Models with Pre-vaccination Surveillance data

Xiaoyan Li, Nathaniel D. Osgood

Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Abstract

Particle filtering is a contemporary Sequential Monte Carlo state inference and identification methodology that allows filtering of general non-Gaussian and non-linear models in light of time series of empirical observations. Several previous lines of research have demonstrated the capacity to effectively apply particle filtering to low-dimensional compartmental transmission models. We demonstrate here implementation and evaluation of particle filtering to more complex compartmental transmission models for pertussis – including application with models involving 1, 2, and 32 age groups and with two distinct functional forms for contact matrices – using over 35 years of monthly and annual pre-vaccination provincial data from the mid-western Canadian province. Following evaluation of the predictive accuracy of these four particle filtering models, we then performed prediction, intervention experiments and outbreak classification analysis based on the most accurate model. Using that model, we contribute the first full-paper description of particle filter-informed intervention evaluation in health. We conclude that applying particle filtering with relatively high-dimensional pertussis transmission models, and incorporating time series of reported counts, can serve as a valuable technique to assist public health authorities in predicting pertussis outbreak evolution and classify whether there will be an outbreak or not in the next month (Area under the ROC Curve of 0.9) in the context of even aggregate monthly incoming empirical data. Within this use, the particle filtering models can moreover perform counterfactual analysis of interventions to assist the public health authorities in intervention planning. With its grounding in an understanding of disease mechanisms and a representation of the latent state of the system, when compared with other emerging applications of artificial intelligence techniques in outbreak projection, this technique further offers the advantages of high explanatory value and support for investigation of counterfactual scenarios.

Keywords: Particle Filter, Mathematical Modelling, Pertussis, Age-structured Model, Contact Matrix, Outbreak Prediction

1. Introduction

Pertussis is a common childhood disease, which is a highly contagious disease of the respiratory tract that caused by the bacterium *Bordetella pertussis* [1]. It is most dangerous for infants, due to risks of severe complications, post-paroxysm apnia [1]. The most frequent complication is pneumonia, while seizures and encephalopathy occur more rarely [1]. Pertussis is a highly contagious disease only found in humans, and spreads from person to person by coughing, sneezing, and prolonged proximity [2]. Evidence indicates a secondary attack rate of 80% among susceptible household contacts [3]. In contrast to some other prevalent childhood diseases, immunity conferred by natural exposure or vaccination 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 [2]. Pertussis incidence shows no distinct seasonal pattern. However, it may increase in the summer and fall [3].

In the pre-vaccination era, pertussis was one of the most common childhood infectious diseases and a major cause of childhood mortality. In 1860, the mortality rate of all-age pertussis in Demark was 0.015% [4],

but that burden fell heavily on infants and children. Research into historical mortality rates from pertussis indicate that the death rate in infancy is higher than in other groups [4]. In recent years globally, there are an estimated 24.1 million cases of pertussis, and about 160,700 deaths per year [5]. Since the 1980s, there has been a rising trend in the reported cases of pertussis in the United States [5]. 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 are believed to go undiagnosed and unreported [5]. Research aimed at estimating the level of population susceptibility and predicting the transmission dynamics of pertussis could aid outbreak prevention and control efforts by health agencies, such as performing intervention before the predicted next outbreak, and in targeted outbreak response immunization campaigns [6].

Dynamic modelling has long served as an important tool for understanding the spread of the infectious diseases in population [7], including pertussis, and for evaluating the impacts of interventions such as immunization and hygeine-enhancing. In recent years, particle filtering as a machine learning method has been employed for incorporating empirical time series data (such as surveillance [8] and online communicational behavior data [9]) to ground the hypothesis as to the underlying model state models in some previous researches [10, 11, 12, 13], especially for the infectious diseases of influenza [14, 15] and measles [8]. In this paper, we apply the particle filtering algorithm in a more complex and widely used compartmental model [16] of pertussis by incorporating the reported pertussis cases in Saskatchewan during the pre-vaccination era. Particle filtering for pertussis is different than for other pathogens on account of the need for state estimation to estimate the population segments at varying levels of immunity. Another need concerns extends from the heterogenous nature of the mixing and incidence burden between different age groups. For this reason, age-structured models are examined here. Specifically, we have examined two categories of age-structured particle filtering models – with 2 age groups and with 32 age groups. Moreover, we have proposed and explored three methods for calculating the contact matrix, so as to reduce the degrees of freedom associated with characterization of the contact matrix. This contribution compares the results obtained from all the particle filtering models by incorporating the empirical data across the whole timeframe evaluating the predictive accuracy of the models. Finally, using the minimum discrepancy particle filtering model, we demonstrate how we can evaluate intervention effects in a fashion that leverages the capacity of particle filtering to perform state estimation.

2. Methods and materials

2.1. Mathematical epidemiological models

As noted above, the dynamics of pertussis in the population is more complex than for infectious diseases that confer lifelong immunity – including other prominent childhood infectious diseases such as measles – due to the temporary character of the immunity acquired by *Bordetella pertussis* infection. As the time since the most recent pertussis infection increases, the immunity of a person wanes [7]. People with lower immunity generally tend to be more easily infected, and exhibit a higher risk of transmitting the infection once infected.

In this paper, we have employed the structure of the popular pertussis mathematical model of Hethcote [16]. To capture the characteristics of pertussis in waning of immunity and the different level of infectiousness and susceptibility involved with infection in light of pre-existing immunity, the compartmental model in [16] further divides the infectious population into three groups: infective with weak-disease (I_w), mild-disease (I_m), and full-disease (I). In a similar fashion, the recovered population is divided into four groups of successively increasing immune system strength: R_1, R_2, R_3 and R_4.

Figure 1 shows the mathematical structure of our compartmental pertussis model adapted from [16]; readers interested in further introduction of this structure are referred to Appendix A. It is notable that the model of Hethcote (1997) [16] employs 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 easing comparison against empirical data – the pertussis reported cases in the province of Saskatchewan, Canada during pre-vaccination era (from 1921 to 1956) – two parts are modified compared to the original model in [16]. Firstly, the model in this paper is represented in a re-dimensionalized fashion, with the state variables representing counts of persons based on the structure in Figure 1. Secondly, because of the focus

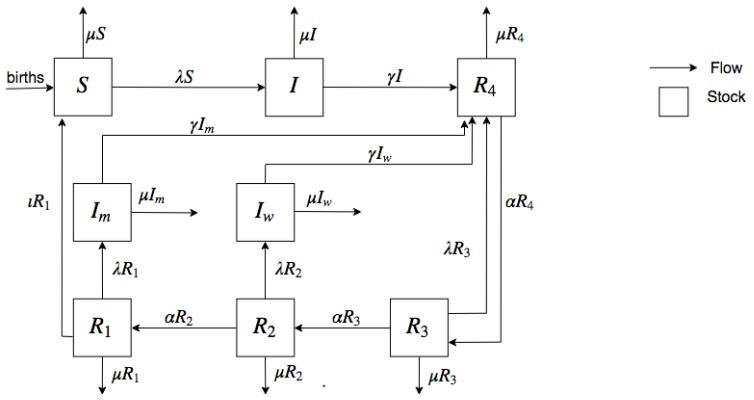


Figure 1: **The transfer diagram for the pertussis model without vaccination.** adapted from [16]

of this paper on the pre-vaccination error, all vaccinated-related elements of the original model of [16] are removed.

Finally, four models are considered in this research. Using n to denote the count of age groups incorporated in the models, we consider models of the aggregate population ($n = 1$), of two age groups ($n = 2$), 32 age groups model ($n = 32$) with the contact matrix introduced in the paper of Hethcote (1997) [16], and a final model with 32 age groups ($n = 32$) model with a re-balanced contact matrix. The mathematical models are introduced in separate sections below.

2.1.1. Aggregate population epidemiological model ($n = 1$)

In the aggregate model, discordant contacts – contacts of infectious individuals (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 (Figure 1) adapted from Hethcote (1997) [16], the equations of the aggregate compartmental model of pertussis are as follows:

$$
\begin{aligned}
\frac{dS}{dt} &= Nv - (\lambda + \mu)S + \iota R_1 \\
\frac{dI}{dt} &= \lambda S - (\gamma + \mu)\, I \\
\frac{dI_m}{dt} &= \lambda R_1 - (\gamma + \mu)\, I_m \\
\frac{dI_w}{dt} &= \lambda R_2 - (\gamma + \mu)\, I_w \\
\frac{dR_1}{dt} &= \alpha R_2 - (\lambda + \mu + \iota) R_1 \\
\frac{dR_2}{dt} &= \alpha R_3 - (\lambda + \mu + \alpha) R_2 \\
\frac{dR_3}{dt} &= \alpha R_4 - (\lambda + \mu + \alpha) R_3 \\
\frac{dR_4}{dt} &= \gamma(I + I_m + I_w) + \lambda R_3 - (\alpha + \mu) R_4 \\
\lambda &= \frac{lp(I + \rho_m I_m + \rho_w I_w)}{N} \\
N &= S + I + I_m + I_w + R_1 + R_2 + R_3 + R_4
\end{aligned}
\tag{1}
$$

The meaning of the states and parameters are as follows: Compartment S is the count of susceptible individuals. Compartments I, I_m and I_w are the count of individuals having full-disease infectious cases

with all of the usual symptoms, with mild disease and weak disease infectious cases, respectively, and with correspondingly decreasing infectivity. It is notable that individuals in both class I_m and I_w lack usual symptoms of pertussis, and thus exhibit atypical pertussis [16]. Compartments R_1, R_2, R_3 and R_4 are the count of recovered people in the population, with correspondingly increasing levels of immunity. N is the population size. v is the overall population birth rate, while μ is the death rate. It is notable that this paper follows [16] in assuming that all model compartments share identical values for the mortality rate (μ), although the death rates in the stocks of the infectives (I_m, I_w, and especially I) are theoretically higher than for the other stocks, due to risk of pertussis-induced mortality. The mean time for waning of immunity from the stock of R_1 to S, and that for successive waning of immunity from successive pairs of R_3, R_2, and R_1, are ι^{-1} and α^{-1}, respectively. The three infectious compartments – I, I_m, I_w – share an identical mean infectious periods of γ^{-1}. However, the infectiousness of an individual varies across the three infectious compartments (I, I_m, I_w), with individuals in compartment I, I_m and I_w having highest, middle and lowest infectiousness, respectively. Parameters ρ_m and ρ_w represent the ratios of the infectiousness of those in the mild-disease (I_m) and weak-disease (I_w) infectious classes to those in the full-disease infectious classes of I [16]. The force of infection parameter λ characterizes the hazard rate – the probability density with which a susceptible (a person in the stocks of S, R_1, R_2 and R_3) is subject to infection from an infective, and is governed by the mass action principles [17, 16]. Parameter λ is related to the total effectively infectious cases ($I + \rho_m I_m + \rho_w I_w$), contact rate (denoted as l) and per-discordant-contact transmission probability (denoted as p).

2.1.2. General age-structured epidemiological model

To capture the difference of the contact pattern among different age groups – for example, the fact that children in school age primarily contact with peers, while babies contact more closely with their parents or caregivers – and adapt the simulation models with the empirical datasets (both monthly pertussis reported cases across the whole population and age-group-specific yearly pertussis reported cases), we extended the pertussis model in Equation (1) to an age-structured model.

The age-structured demographic model. Before introducing the epidemiological age-structured pertussis mathematical model, we first introduce the age-structured demographic model. The demographic model mainly captures the age structure and the birth and death in the population related to the empirical data (pertussis reported cases in province of Saskatchewan in Canada during the pre-vaccination era – from 1921 to 1956) employed in this paper. Suppose we have n age groups in the whole population, by divided by a sequence of ages $a_i, 1 \leq i \leq n-1$. The age groups can be characterized as a series of n intervals – $[0, a_1), [a_1, a_2), \cdots, [a_{n-1}, \infty)$. The demographic model can then be written as follows [8, 16, 18].

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

where N_i is the number of people in age group i; v_i and μ_i are the birth and death rate of age group i, respectively; c_i is the aging rate of age group i, given by $c_i = (a_i - a_{i-1})^{-1}$, and $c_n = 0$.

In this paper, we assume that the population is in equilibrium; this reflects the fact that the empirical Saskatchewan population size from 1921 to 1956 does not change dramatically [19], as will be discussed below in greater detail. This approximation assumes that the total population $N_i(t)$ of age group i will remain invariant over the model time horizon, that is, $dN_i(t)/dt = 0$. Thus, according to Equation (2), for this simplified context, the death rate μ_i can be calculated as follows:

$$\begin{aligned} \mu_1 &= \frac{\sum_{j=1}^{n} v_j N_j(t) - c_1 N_1(t)}{N_1(t)} \\ \mu_i &= \frac{c_{i-1} N_{i-1} - c_i N_i}{N_i(t)}, \quad i \geq 2 \end{aligned} \tag{3}$$

The values of parameters in the demographic model are estimated from the empirical data. Specifically, the population in each age group N_i is estimated from the age pyramid of Saskatchewan [19], and the birth rates v_i are estimated from the Public Health Annual Report of Saskatchewan [20] published yearly by the Government of Saskatchewan.

The age-structured pertussis epidemiological model. By incorporating the age-structured demographic model shown in Equation (2), and the structure of the compartmental epidemiological model shown in Figure 1, we obtain the age-structured pertussis epidemiological model given below. Readers interested in the detailed mathematical derivation are referred to our previous contribution [8].

$$
\begin{aligned}
\frac{dS_1}{dt} &= \sum_{j=1}^{n} v_j N_j + \iota R_{11} - (c_1 + \lambda_1 + \mu_1) S_1 \\
\frac{dS_i}{dt} &= c_{i-1} S_{i-1} + \iota R_{1i} - (c_i + \lambda_i + \mu_i) S_i \quad 2 \le i \le n \\
\frac{dI_1}{dt} &= \lambda_1 S_1 - (c_1 + \gamma + \mu_1) I_1 \\
\frac{dI_i}{dt} &= c_{i-1} I_{i-1} + \lambda_i S_i - (c_i + \gamma + \mu_i) I_i \quad 2 \le i \le n \\
\frac{dI_{m1}}{dt} &= \lambda_1 R_{11} - (c_1 + \gamma + \mu_1) I_{m1} \\
\frac{dI_{mi}}{dt} &= c_{i-1} I_{m,i-1} + \lambda_i R_{1i} - (c_i + \gamma + \mu_i) I_{mi} \quad 2 \le i \le n \\
\frac{dI_{w1}}{dt} &= \lambda_1 R_{21} - (c_1 + \gamma + \mu_1) I_{w1} \\
\frac{dI_{wi}}{dt} &= c_{i-1} I_{w,i-1} + \lambda_i R_{2i} - (c_i + \gamma + \mu_i) I_{wi} \quad 2 \le i \le n \\
\frac{dR_{11}}{dt} &= \alpha R_{21} - (\lambda_1 + \iota + c_1 + \mu_1) R_{11} \\
\frac{dR_{1i}}{dt} &= c_{i-1} R_{1,i-1} + \alpha R_{2i} - (\lambda_i + \iota + c_i + \mu_i) R_{1i} \quad 2 \le i \le n \\
\frac{dR_{21}}{dt} &= \alpha R_{31} - (\lambda_1 + \iota + c_1 + \mu_1) R_{21} \\
\frac{dR_{2i}}{dt} &= c_{i-1} R_{2,i-1} + \alpha R_{3i} - (\lambda_i + \iota + c_i + \mu_i) R_{2i} \quad 2 \le i \le n \\
\frac{dR_{31}}{dt} &= \alpha R_{41} - (\lambda_1 + \iota + c_1 + \mu_1) R_{31} \\
\frac{dR_{3i}}{dt} &= c_{i-1} R_{3,i-1} + \alpha R_{4i} - (\lambda_i + \iota + c_i + \mu_i) R_{3i} \quad 2 \le i \le n \\
\frac{dR_{41}}{dt} &= \gamma (I_1 + I_{m1} + I_{w1}) + \lambda_1 R_{31} - (\alpha + c_1 + \mu_1) R_{41} \\
\frac{dR_{4i}}{dt} &= c_{i-1} R_{4,i-1} + \gamma (I_i + I_{mi} + I_{wi}) + \lambda_i R_{3i} - (\alpha + c_i + \mu_i) R_{4i} \quad 2 \le i \le n \\
N_i &= S_i + I_i + I_{mi} + I_{wi} + R_{1i} + R_{2i} + R_{3i} + R_{4i} \quad 1 \le i \le n \\
\mu_1 &= \frac{\sum_{j=1}^{n} v_j N_j - c_1 N_1}{N_1} \\
\mu_i &= \frac{c_{i-1} N_{i-1} - c_i N_i}{N_i} \quad 2 \le i \le n
\end{aligned}
\tag{4}
$$

In this age-structured epidemiological model, the definition of most quantities are consistent with (*mutatis mutandis*) the aggregate population epidemiological model (Equation (1)) and the age-structured demographic model (Equation (3)), with the notable exception of the force of infection λ_i for age group i.

As noted above, this work followed [16] in characterizing transmission of pertussis infection between infectives and susceptibles according to mass action principles. The force of infection defined as the hazard rate with which susceptibles are infected by infectives, and is related to contact rate, transmission probability, and the fraction of infectives in the whole population. In the model with aggregate population, the individuals are assumed contact homogeneously, and the force of infection can be simply calculated as in Equation (1). However, in the age-structured model, contacts between individuals are assumed to occur homogeneously within age groups and heterogeneously across different age groups. Thus, the calculation of force of infection in the age-structured models are considerably more complex than for aggregate population model, being mediated by a contact matrix. Readers interested in the mathematical representation of the contact matrix could are referred to our previous contribution [8]. In the current paper, we have employed three different methods of calculating the contact matrix and – by extension – the force of infection in the models. These three methods are introduced as follows.

2.1.3. Force of infection models

General mass action-based contact matrix. Under the assumption of mass action, the force of infection – the hazard rate (probability density) with which a susceptible is transmitted the pathogen by infectives – can be calculated by the sum of the hazard rates associated with transmission from infectives in each age group in turn. The force of infection of each age group is correspondingly represented as follows.

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} l_1 f_{11} & l_1 f_{12} & \cdots & l_1 f_{1n} \\ l_2 f_{21} & l_2 f_{22} & \cdots & l_2 f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_n f_{n1} & l_n f_{n2} & \cdots & l_n f_{nn} \end{bmatrix} \times \begin{bmatrix} p_1 \frac{I_1 + \rho_m I_{m1} + \rho_w I_{w1}}{N_1} \\ p_2 \frac{I_2 + \rho_m I_{m2} + \rho_w I_{w2}}{N_2} \\ \vdots \\ p_n \frac{I_n + \rho_m I_{mn} + \rho_w I_{wn}}{N_n} \end{bmatrix} \tag{5}$$

The above can be rewritten as the following equation:

$$\lambda_i = l_i p_i \sum_{j=1}^{n} f_{ij} \frac{I_j + \rho_m I_{mj} + \rho_w I_{wj}}{N_j} \tag{6}$$

where λ_i, l_i, p_i, I_i, I_{mi} and I_{wi} are the force of infection, contact rate, transmission probability, number of persons in full-disease infectious, number of persons in mild-disease and weak-disease infectious classes in age group i, respectively. For an individual in age group i, f_{ij} is the fraction of that individual's contacts that occur with others in the age group of j. Thus, for a given age group i: $\sum_{i=1}^{n} f_{ij} = 1$. $l_i f_{ij}$ are then the elements in the contact matrix.

An advantage of this method in calculating the contact matrix in the age-structured model is that the contacts between any two age groups (e.g., i and j) is balanced (symmetric) – 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; that is, $N_i l_i f_{ij} = N_j l_j f_{ji}$. However, this method has a notable disadvantage that the count of unknown parameters in calculating the contact matrix grows quadratically with the count of age groups (denoted as n) in the model; a demonstration of the super-linear growth of the total number of unknown parameters in the contact matrix with the total number of age groups is shown in Appendix B. This disadvantage makes challenging parameter estimation for models incorporating a large number of age groups. To address this challenge, we have explored two other methods for characterizing the contact matrix and force of infection in which the count of parameters grows sub-linearly or linearly with the total number of age groups. The first is a method of obtaining an un-balanced contact matrix contributed by Hethcote (1997) with a constant number of unknown parameters [16]. The second approach calculates a re-balanced contact matrix in which the number of unknown parameters grows linearly with the total number of age groups. Each of these approaches are characterized below.

The Unbalanced Contact Matrix. This unbalanced contact matrix is introduced in the research of Hethcote (1997) [16], which assumes that only adequate contacts are sufficient to transmit the disease. This method based on a simple proportional mixing assumption that the number of total persons contacted by one person

in the age group j is distributed among the population in the age group i in proportion to the fractions l_i/D^*, where D^* is the total number of contacts per unit time received by all people, l_i is the contact rate – average number of persons contacted by a person per unit time – of age group i, and $D^* = \sum_{j=1}^{n} l_j N_j$ [16]. The elements in the contact matrix are $l_i l_j / D^*$ [16]. Finally, the re-dimensionalized force of infection (λ) used in Equation (4) and in [16] is given as follows:

$$\lambda_i = \sum_{j=1}^{n} \frac{l_i l_j}{D^*} (I_j + \rho_m I_{mj} + \rho_w I_{wj}) \tag{7}$$

However, in this paper, we employ the dimensionless representation of the "force of infection" in Equation (8), which is consistent with [16], instead of the re-dimensionalized one in Equation (7). The motivation for this lies in our use of the values of parameters related to the mixing matrix from [16], which will be detailed below in the section "particle filtering implementation". The dimensionless equation of force of infection in [16] is as follows:

$$\lambda_i = \sum_{j=1}^{n} \frac{l_i l_j}{D} \frac{I_j + \rho_m I_{mj} + \rho_w I_{wj}}{\sum_{j=1}^{n} N_j} \tag{8}$$

where D is the dimensionless total contacts across all population, and $D = \sum_{j=1}^{n} l_j N_j / \sum_{j=1}^{n} N_j = D^* / \sum_{j=1}^{n} N_j$.

The advantage of this method is that – if one adopts the values of the contact rate in each age group given in [16] – there are no unknown parameters required for calculating the contact matrix. And it is straightforward to calculate the contact matrix as long as those age-specific contact rate parameters are known. However, this method of calculating the contact matrix suffers from a notable disadvantage – a lack of guaranteed symmetry between the contacts exerted between pairs of age groups. Specifically, it can be readily shown that the value of the total contacts occurring from age group i to age group j is not in general equal to the value of the total contacts occurring from age group j to age group i. This reflects the fact that 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 $N_k l_j l_k / D^*$. In general, these two quantities need not be equal.

To address this shortcoming, we explored a previously contributed method to calculate a balanced contact matrix. While the above method does not require additional parameters, for the balanced method, the total number of the unknown parameters grows linearly with the number of age groups.

The Re-balanced Contact Matrix. To calculate the balanced contact matrix, we have employed the method introduced in research by Garnett and Bowden (2000) [21]. The elements of the contact matrix $l_{ij} f_{ij}$ and force of infection λ_i are as follows; readers interested in the detailed mathematical deduction of the re-balanced contact matrix can refer to Appendix C:

$$\begin{aligned} l_{ij} f_{ij} &= l_i \left(\frac{\epsilon_j}{\epsilon_i} \right)^{0.5} \left[(1.0 - \epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j} \right) \right] \\ \lambda_i &= p_i \sum_{j=1}^{n} \frac{l_{ij} f_{ij} (I_j + \rho_m I_{mj} + \rho_w I_{wj})}{N_j} \end{aligned} \tag{9}$$

where p_i is the transmission probability of age group i, f_{ij} is the fraction of the contacts of an individual in age group i that are made with others in age group j, δ_{ij} is the identity matrix, mixing parameter ϵ_i determines where mixing occurs on a scale from fully homophilic – persons only contact with the individuals in the same age group (representing $\epsilon_i = 0$) – to random mixing in which the contact among the total population is non-preferential (representing $\epsilon_i = 1.0$).

Finally, based on the above discussion, we have employed four pertussis epidemiological models as the state-space models to be used in corresponding applications of particle filtering – the aggregate population model (shown in Equation (1)), the age-structured model with two age groups (Equation (4) with $n = 2$)

with the general contact matrix based the mass action assumption (Equation (6)), the age-structured model with 32 age groups (Equation (4) with $n = 32$) with an un-balanced contact matrix (Equation (8)) and the age-structured model with 32 age groups (Equation (4) with $n = 32$) with a re-balanced contact matrix (Equation (9)). It is notable that the 32 age group division applied is directly adopted from Hethcote (1997) [16], with age groups from 0–1 month, 2–3 months, 4–5 months, 6–11 months, integer ages for 1 through and including 19, 20–24 years, 25–29 years, 30–39 years, 40–49 years, 50–59 years, 60–69 years, 70–79 years, 80–89 years, 90 years and older. The age structured model with two age groups dichotomizes the population into 0–4 years and 5 years and older age categories. It bears noting that while more detailed age structure can better capture both the effects of population aging and inter-group heterogeneity, in terms of particle filtering, it entails estimation of a larger underlying model state space – potentially adversely affecting the accuracy of that estimation; in many models, it also requires specification of additional parameter values.

2.2. Particle Filter Implementation

Particle filtering is a contemporary state inference and identification methodology that allows filtering of general non-Gaussian and non-linear state space models in light of time series of empirical observations [10, 11, 22, 15, 8, 23]. This 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. The state space model characterizes the processes governing evolution over time of internal states with stochastics consisting of random perturbations. The states in the state space model are assumed in general to be latent and unobservable. Information concerning the latent states is obtained from a noisy observation vector. The means by which the particle filter method operates includes the Recursive Bayes Filter [22], Sequential Importance Sampling [22, 23, 10], and Resampling [22, 23, 10].

Sequential importance sampling (SIS) is the most basic Monte Carlo method used to sample when the predict-and-update equations of the recursive Bayes filter are not analytically tractable [22]. The key idea of SIS is to estimate the posterior distribution at a given time with a weighted set of samples. SIS then recursively updates these prior samples to obtain samples approximating the posterior distribution at the next time step. These importance-weighted samples are also named particles [22]. 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. This occurs because we are sampling in a high dimensional space, using a myopic proposal distribution [23].

The key idea underlying resampling is a variant of the principle of "survival of the fittest". To achieve this, the resampling step will monitor the effective sample size following each observational update. Whenever the effective sample size drops below a threshold, the algorithm will draw a new set of particles from the existing set, where the probability of drawing a given particle is – in accordance with the principle of importance sampling – proportional to its weight. Within such resampling, particles with higher weight will tend to be reproduced, and particles with lower weight will tend to die out. The new particles inherit their parent's values but carry a uniform normalized weight. At a given time, each particle contributing to the distribution (represented collectively by the particles according to the principles of sequential importance sampling [23]) can be seen as representing a competing hypothesis concerning the underlying state of the system at that time. The particle filtering method can be viewed as undertaking a "survival of the fittest" of these hypotheses, with fitness of a given particle being determined by the consistency between the expectations of the hypothesis associated with that particle and the empirical observations.

Interested readers are referred to more detailed treatment in [22, 24, 23, 10].

2.2.1. State Space Model

The state space model depicts the processes governing the state – both latent and observable – of a noisy system evolving with time. In this paper, we employ the deterministic pertussis epidemiological models as base models. Reflecting the fact that particle filtering offers value in the context of underlying state equation models exhibiting stochastic variability, we then extend these deterministic models by adding random perturbations in some processes or parameters, so as to represent the stochastic processes in the real world; the extended, stochastic model then serves as the basis for a corresponding particle filter. Thus, we have built four particle filtering models based on the respective pertussis compartmental epidemiological models introduced previously – the aggregate population model, two-age group model with the general

contact matrix, and the 32-age group models with both un-balanced contact matrix and re-balanced contact matrix.

Stochastic Adaptation of the Aggregate Population State Space Model. In the aggregate population state space model, we employ the aggregate population pertussis compartmental epidemiological model (equation (1)) as the base model. Stochastics are added to this base model in three areas – in the rate of new infections, for the contact process between susceptibles and infectives, and in the reporting process for infected cases. The mathematical structure of the pertussis aggregate population state space model is shown in Figure 2. The stochastics associated with these factors represents a composite of two factors. Firstly, there is expected to both be stochastic variability in the pertussis infection processes and some evolution in the underlying transmission dynamics in terms of an evolving reporting rate, as well as 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 to provide the model with the requisite openness to correction when observing a new empirical datum [8].

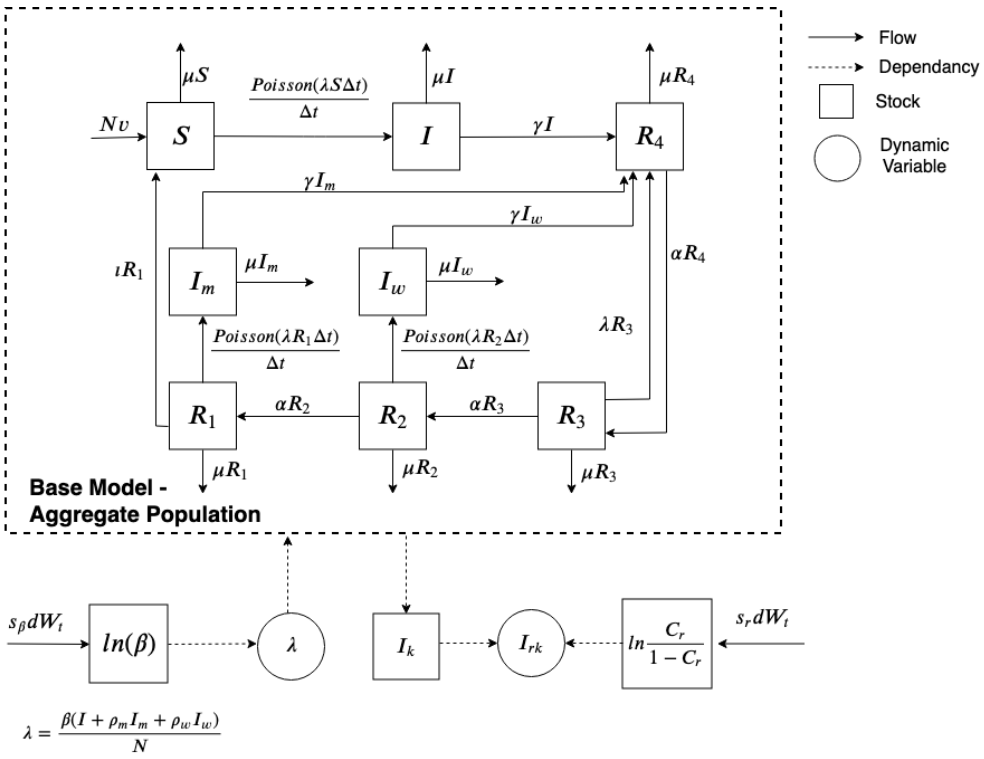


Figure 2: **The mathematical structure of the aggregate particle filtered model.**

In characterizing transmission process, we consider a stochastic process – specifically, a Poisson process – associated with incidence of infection, including cases of full-disease infectives (I), mild-disease infectives (I_m) and weak-disease infectives (I_w). This process reflects the small number of cases that occur over each small unit of time – denoted as Δt (carrying the value of 0.01 months in all models in this paper, or roughly 7.3 days) [10, 8]. The new infection flows incorporating stochastic process (Poisson process) are correspondingly listed as follows:

$$A_I = \frac{Poisson(\lambda S \Delta t)}{\Delta t}$$

$$A_{I_m} = \frac{Poisson(\lambda R_1 \Delta t)}{\Delta t}$$
$$A_{I_w} = \frac{Poisson(\lambda R_2 \Delta t)}{\Delta t} \tag{10}$$

The second stochastic process considered in the aggregate population pertussis state space model is the mixing process between susceptibles and the infectious. We know that the transmission probability of the disease of pertussis is normally a constant. Thus, to simplify the model, we incorporated an effective contact rate parameter, denoted as β, where the effective contact rate is the multiplication of a per-month contact rate and transmission probability (of unit dimension), denoted as l and p ($\beta = lp$) in the deterministic aggregate population compartmental pertussis epidemiological model characterized in Equations (1). We posit that parameter will undergo some evolution in value in accordance with contact rates – such as due to social distancing, as the school year starts or stops, and enhanced hygenic awareness during outbreaks. We thus characterized effective contact rate β as evolving stochastically within the model.

To estimate changing values of the stochastic effective contact rate parameter β, and to investigate the capacity of the particle filter to adapt to parameters whose effective values evolve over simulation, we incorporated the parameter β into the state of the particle filter model, as seen in Figure 2. Moreover, reflecting the fact that the effective contact rate β is conceptually bounded to the non-negative real numbers, we treat the natural logarithm of the effective contact rate β as undergoing a random walk according to Brownian Motion, as characterized by a Wiener Process [25, 26, 8]. The stochastic differential equation of the effective contact rate β can thus be described according to Stratonovich notation as:

$$dln(\beta) = s_\beta dW_t \tag{11}$$

where dW_t is a standard Wiener process whose perturbations follow a normal distribution with 0 of mean and unit rate of variance; s_β is the diffusion coefficient. Thus, the perturbations in the value of $ln(\beta)$ are normally distributed with 0 of mean and variance ${s_\beta}^2$.

The third stochastic process considered in the noisy state space model relates to the reporting process for infected pertussis cases. Over the multi-decadal model time horizon (as circumscribed by the span of the empirical data from 1921 to 1956), 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, we incorporated another stochastically evolving parameter – the fraction of underlying pertussis cases that are reported (denoted as C_r); as for the above parameters, this parameter is also treated as an element of evolving model state. Reflective of the fact that the reporting rate C_r is a probability limited to the range $[0, 1]$, we characterize the logit of C_r as also undergoing Brownian Motion according to Stratonovich notation [8] as follows:

$$d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t \tag{12}$$

where dW_t is as above; s_r is the diffusion coefficient. Perturbations in the value of $ln(\frac{C_r}{1-C_r})$ with time follow a normal distribution with mean 0 and variance ${s_r}^2$.

Moreover, to calculate the reported number of pertussis cases in the particle filtering model, which is used in the measurement model discussed below, we incorporated an extra state, denoted as I_k, which accumulates the count of pertussis infectious cases from time $k-1$ to time k. It is notable that the state of cumulative infectious cases from time $k-1$ to k – I_k – is different from the original infectious states I, I_m or I_w in the deterministic compartmental model in Equations (1). Specifically, the state of the cumulative count of infectious cases I_k purely integrates all the inflows to the infectious states as a whole (and without all the outflows), so as to simulate a similar process of successively tallying up the pertussis cases over the course of some period of time as is undertaken in the real world. Moreover, we further assume that the individuals with mild-disease infectious cases (I_m) and weak-disease infectious cases (I_w) are also subject to reporting. The reporting rates of the mild-disease infectious cases (I_m) and weak-disease infectious cases (I_w) that have symptoms are considered to be ρ_m and ρ_w, in this paper. It is notable that the sequence of the values of k correspond to the sequence of historical reporting times (per Month in this paper). Then, the cumulative infectious cases from time $k-1$ to k in state I_k is represented as follows:

$$I_k = \int_{k-1}^{k} \left(\frac{Poisson(\lambda S \Delta t)}{\Delta t} + \rho_m \frac{Poisson(\lambda R_1 \Delta t)}{\Delta t} + \rho_w \frac{Poisson(\lambda R_2 \Delta t)}{\Delta t} \right) dt \tag{13}$$

It bears emphasis that the model implementation of Equation (13) made use of identical values drawn for the stochastic components used in the flows that it serves to accumulate. Thus, the reported pertussis cases calculated in the state space model at the measure time k, denoted as I_{rk}, can be represented as follows:

$$I_{rk} = I_k C_r \tag{14}$$

Finally, we obtained the noisy state space model of the pertussis particle-filtered aggregate model by incorporating into the base model – as given by the deterministic compartmental epidemiological model in equations of (1) – the adjusted stochastic parts in equations of (10), (11), (12), (13) and (14) (Figure 2). Readers interested in the complete mathematical equations, parameter values, and initial state assumptions for the state space model can refer to Appendix D; those seeking better understanding of basis for the parameters related to the transmission of pertussis in this model are referred to the research of Hethcote (1997) [16]. The demographic parameters of this model are sourced from the Annual Report of the Saskatchewan Department of Public Health [20] and the age pyramid of Saskatchewan [19]. The initial values of states in this model are estimated by tuning the particle-filtered model (the assumptions regarding the distribution of the initial states, as given by constants) and sampled by the particle filtering algorithm. Both the values of parameters and initial values of states in this model are listed in Appendix D.

The two-age group population structure state space model. In the two-age-group population structure state space model, we employ the age- and population-structured pertussis compartmental epidemiological model (Equation (4)) with $n = 2$ as the base model, where the variable of "force of infection" is calculated according to the mass-action based formulation of the general contact matrix (Equation (5)). In this model variant, we use subscripts "c" and "a" to denote the child- and adult-specific values, respectively, where the child age group includes all individuals from newborns to the end of the fourth year, and the remaining individuals are in the adult age group. Similar to the state space model with an aggregate population, noise is imparted to this base model in three elements – the new infectious occurrence process, the contact process between susceptibles and infectives, and the reporting process for infected cases. The mathematical structure of the pertussis aggregate population state space model is shown in Figure 3.

As discussed in the aggregate population state space model, we consider occurrence of infections within a given small interval to be characterized by a Poisson process. Then, the flows of new infections incorporated into the model are given by the following equations:

$$\begin{aligned}
A_{I_c} &= \frac{Poisson(\lambda_c S_c \Delta t)}{\Delta t} \\
A_{I_a} &= \frac{Poisson(\lambda_a S_a \Delta t)}{\Delta t} \\
A_{I_{mc}} &= \frac{Poisson(\lambda_c R_{1c} \Delta t)}{\Delta t} \\
A_{I_{ma}} &= \frac{Poisson(\lambda_a R_{1a} \Delta t)}{\Delta t} \\
A_{I_{wc}} &= \frac{Poisson(\lambda_c R_{2c} \Delta t)}{\Delta t} \\
A_{I_{wa}} &= \frac{Poisson(\lambda_a R_{2a} \Delta t)}{\Delta t}
\end{aligned} \tag{15}$$

Characterization of the stochastic mixing process between susceptibles and infectives within the stratified model is more involved than the same process in the aggregate population model, due to the need to include both homogeneous mixing within the same age group and heterogeneous mixing amongst different age groups.

In the two-age structured model, we assume that all the differences in transmission from an infected adult vs. an infected child is due to differences in contact rates, and thus that the transmission probability

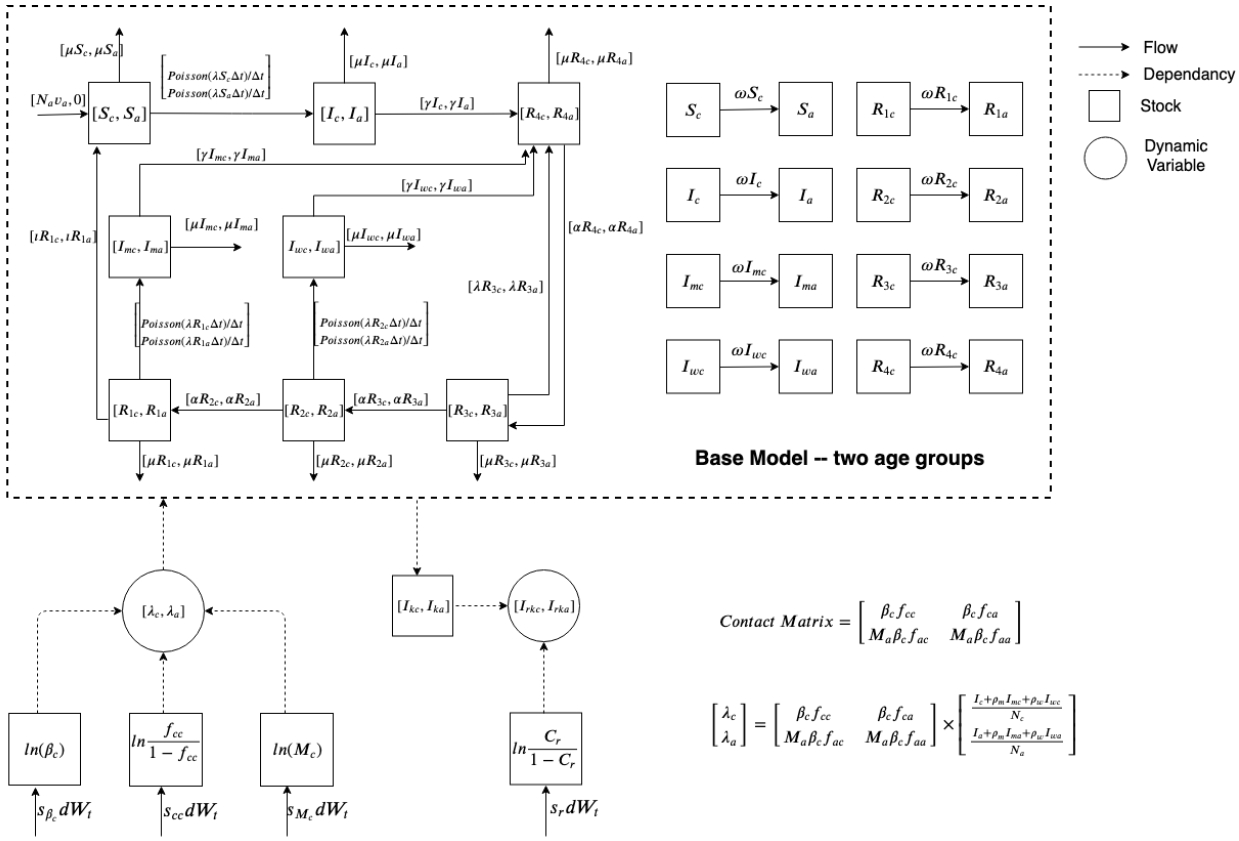


Figure 3: **The mathematical structure of the particle filtering age-structured model with two age groups.**

of pertussis (denoted as p_i in the force of infection model of Equation (5) and Equation (6) for age group i) are the same between child and adult age groups (i.e., that $p_c = p_a$). Then, according to the general contact matrix model based on mass action introduced previously, we obtain the following equations:

$$\begin{aligned} f_{cc} + f_{ca} &= 1 \\ f_{ac} + f_{aa} &= 1 \\ N_c l_c f_{ca} &= N_a l_a f_{ac} \end{aligned} \tag{16}$$

where l_c and l_a are the contact rates of child and adult age groups; N_c and N_a are the total populations of the child and adult age groups; $f_{ij}, i, j \in [c, a]$ indicates the fraction of the contacts of age group i occur with age group j.

Then, similarly to the aggregate population state space model, we import the parameters – effective contact rates of the child and adult age groups – denoted as β_c and β_a, respectively. We know the effective contact rate is the multiplication of the parameter of contact rate and transmission probability. Then, we get $\beta_c = l_c p_c$ and $\beta_a = l_a p_a$. Substituting the equation with β_c and β_a to the Equation (15), we can get [8]:

$$\begin{aligned} f_{ca} &= 1 - f_{cc} \\ f_{ac} &= \begin{cases} \frac{N_c \beta_c}{N_a \beta_a}\left(1 - f_{cc}\right), & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a}\left(1 - f_{cc}\right)\right] < 1.0 \\ 1.0, & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a}\left(1 - f_{cc}\right)\right] \geq 1.0 \end{cases} \\ f_{aa} &= 1 - f_{ac} \end{aligned} \tag{17}$$

To represent the stochastic characteristics of the mixing process of the two-age group state space model, we allowed three parameters to change with time according to a random walk (with the values of these parameters being estimated as part of model state upon each observation during particle filtering) – the effective contact rate of the child age group β_c, the fraction of the contacts of the child age group that occur with the child age group f_{cc}, and the ratio of the adult age group's effective contact rate (β_a) to that of the child age group (β_c), denoted as M_a. Reflecting the fact that both β_c and M_a vary over the entire range of positive real numbers and f_{cc} varies in the range of $[0, 1]$, we treat the natural logarithm of each of β_c and M_a, as well as the logit of f_{cc}, as undergoing a random walk according to a Wiener Process, and thus undergoing Brownian Motion) [25, 26, 8]. Drawing on notation from the Stratonovich calculus for the random walks involved, we obtain the equations as follows:

$$\begin{aligned} d(ln\beta_c) &= s_{\beta_c} dW_t \\ d(ln(\frac{f_{cc}}{1-f_{cc}})) &= s_{cc} dW_t \\ d(lnM_a) &= s_{M_a} dW_t \\ \beta_a &= M_a \beta_c \end{aligned} \tag{18}$$

The third stochastic component in the two-age group model relates to calculation of the reported cases of pertussis in the model. As in the aggregate population model, for comparison with reported case counts, we also make use of two convenience states – denoted as I_{kc} and I_{ka} – to accumulate pertussis infectious cases from time $k-1$ to k for the child and adult age groups. Moreover, we assume that the pertussis reporting rates of child and adult age groups are the same. Thus, the equation of reporting rate – denoted as C_r – is identical to that in the aggregate model in Equation (12). The mathematical equations characterizing the reporting process are listed as follows:

$$\begin{aligned} d(logit(C_r)) &= d(ln(\frac{C_r}{1-C_r})) = s_r dW_t \\ I_{kc} &= \int_{k-1}^{k} (A_{I_c} + \rho_m A_{I_{m_c}} + \rho_w A_{I_{w_c}})\, dt \\ I_{ka} &= \int_{k-1}^{k} (A_{I_a} + \rho_m A_{I_{m_a}} + \rho_w A_{I_{w_a}})\, dt \\ I_{rck} &= C_r I_{kc} \\ I_{rak} &= C_r I_{ka} \end{aligned} \tag{19}$$

where dynamic variables I_{rck} and I_{rak} indicate the reported pertussis cases calculated from the two-age group model.

Finally, the noisy state space model of the two-age group pertussis particle-filtered transmission model is the combination of the base model of the deterministic compartmental epidemiological model in Equations (4) and the adjusted stochastic parts in Equations (15), (18) and (19) (Figure 3). Readers interested in the full mathematical equations of the state space model, values of parameters, and initial states can refer to Appendix D.

32-age group population structure state space models. In this paper, we have explored two pertussis particle filtering models with 32-age group population structure – with the unbalanced contact matrix introduced by [16] (Equation (8)) and re-balanced contact matrix (Equation (9)) – taking the deterministic epidemiological model of Equation (4) with $n = 32$ as the base model. As in the state space models above, we also incorporated three stochastic elements within the 32-age group state space models – the new infectious occurrence process, the contact process between susceptibles and infectives, and the infected case reporting process (Figure 4).

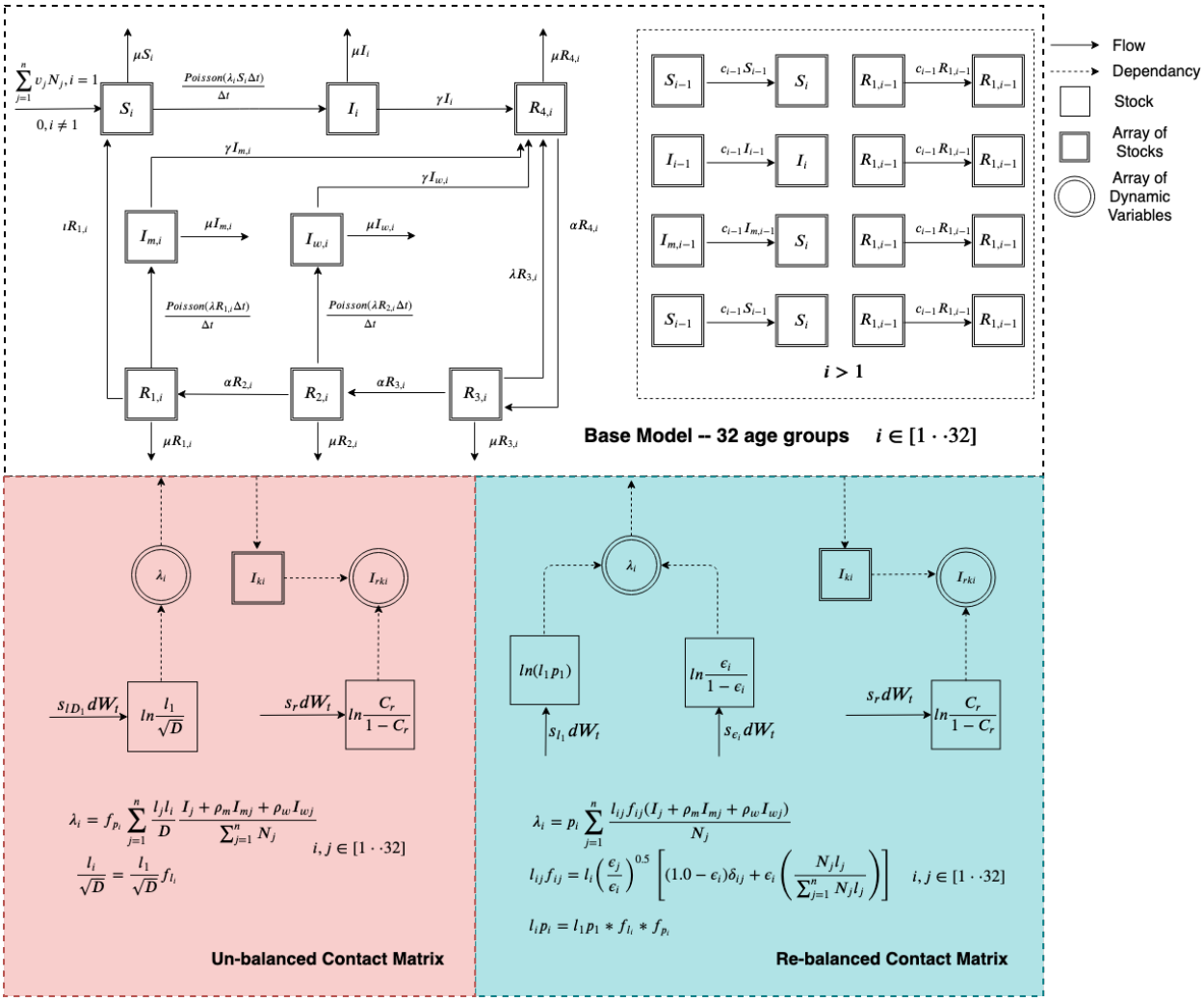


Figure 4: **The mathematical structure of the particle filtering age-structured model with 32 age groups.**

Similar to the aggregate and two-age group population state space models introduced above, we consider the new infectious individuals occurrence processes follows the Poisson process, and the mathematical equations are listed as follows:

$$
\begin{aligned}
A_{I_i} &= \frac{Poisson(\lambda_i S_i \Delta t)}{\Delta t} \quad 1 \le i \le 32 \\
A_{I_{m\,i}} &= \frac{Poisson(\lambda_i R_{1i} \Delta t)}{\Delta t} \quad 1 \le i \le 32 \\
A_{I_{w\,i}} &= \frac{Poisson(\lambda_i R_{2i} \Delta t)}{\Delta t} \quad 1 \le i \le 32
\end{aligned}
\tag{20}
$$

Similarly to those previous models, in the stochastic process of reporting the pertussis cases in the 32-age group state space models, we consider the reporting rate of each age group to be the same, denoted as C_r, the logit of C_r undergoing Brownian Motion. The resulting mathematical equations related to the reporting process are listed as follows:

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

$$I_{ki} = \int_{k-1}^{k} (A_{I_i} + \rho_m A_{I_{m\,i}} + \rho_w A_{I_{w\,i}})\, dt \quad 1 \le i \le 32 \tag{21}$$
$$I_{rki} = I_{ki} C_r \quad 1 \le i \le 32$$

where I_{ki} represent the new states incorporated into the state space model to capture the accumulative number of pertussis cases from time $k-1$ to time k for age group i, and dynamic variable I_{rki} is the estimated occurrence of reported cases calculated by the state space model for age group i.

Then, in characterizing the mixing process between susceptibles and infectives, we separately implement the 32-age group state space model with the un-balanced contact matrix introduced in [16] and re-balanced contact matrix.

In the un-balanced contact matrix method introduced in [16], we consider the parameter of $l_1/\sqrt{D}$ to evolve stochastically in the state space model (i.e., the natural logarithm of $l_1/\sqrt{D}$ undergoes Brownian Motion). Then, a vector represents how the contact rate of each successive age group compares with that of the first age group; specifically, f_{l_i} represents the ratio between the contact rate for the first age group and the contact rate of age group i. This vector is then used to calculate the parameter of $l_i/\sqrt{D}$ for each age group i. f_{l_i} is calculated from the value assumed for contact rate of all age groups, which are taken from [16]. 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 f_{p_i} is incorporated to represent the ratio of the transmission probability of pertussis of each age group compared to the first age group in the state space model. The original mathematical model of [16] lacks a dedicated transmission probability parameter. However, one would expect transmission probabilities to different among different age groups. For example, transmission probability from a young child is usually higher than that of the adults due to hygienic disparities. The mathematical equations of the stochastic mixing process are listed as follows:

$$\begin{aligned}
&d(ln\frac{l_1}{\sqrt{D}}) = s_{lD_1} dW_t \\
&\frac{l_i}{\sqrt{D}} = l_1 * f_{l_i} \quad 2 \le i \le 32 \\
&\lambda_i = f_{p_i} \sum_{j=1}^{n} \frac{l_j l_i}{D} \frac{I_j + \rho_m I_{mj} + \rho_w I_{wj}}{\sum_{j=1}^{n} N_j} \quad 1 \le i \le 32 \\
&D = \sum_{k=1}^{n} l_k N_k / \sum_{k=1}^{n} N_k
\end{aligned} \tag{22}$$

The value of f_{p_i} in the 32-age-group models of pertussis in this paper is (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). We assume that the transmission probability of individuals under 15 years old are the same and is the highest, while the transmission probabilities of individuals in the age groups from 15 to 19 years and over 20 years are half and 1/20 compared to the individuals under 15 years, respectively. The population of each age group – which is collected from the age pyramid of Saskatchewan [19] and is assumed to be invariant – 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 arrival rate of newborns in each pertussis particle filtering model per unit time (here, month) be the same across all models, the yearly birth rate of the 32-age-group models are assumed as (0, 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). This is done to ensure the new born population each year are the same in all the models. The values of birth rates are informed from the [20]. Readers interested in a complete characterization of the mathematical equations of 32-age group state space model with an un-balanced contact matrix [16] and the initial values of all states can refer to Appendix D.

In the re-balanced contact matrix method, to represent the stochastic mixing process, we assume that the changes of the logarithm of $l_1 p_1$ (the effective contact rate of the first age group) undergoes a random

walk according to a Wiener Process (Brownian Motion) [25, 26, 8]. The logit of the six mixing parameters $(\epsilon_i, 1 \leq i \leq 6)$ are similarly treated as evolving according to a Wiener Process. The reason that the total number of mixing parameters is 6, instead of 32 – as might be expected if there are a mixing parameter related to each age group each – lies in the fact that the yearly empirical datasets could only be split into 6 age groups – less than 1 year, 1 to 4 years, 5 to 9 years, 10 to 14 years, 15 to 19 years and 20 years and older, as is characterized in detail below. Finally, the force of infection for the the 32-age group structured pertussis particle filtering model with a re-balanced contact matrix is given as follows:

$$
\begin{aligned}
& l_{ij} f_{ij} = l_i \left(\frac{\epsilon_j}{\epsilon_i}\right)^{0.5} \left[(1.0-\epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j}\right)\right] \quad 1 \leq i \leq 32, \quad 1 \leq j \leq 32 \\
& \lambda_i = p_i \sum_{j=1}^{n} \frac{l_{ij} f_{ij} (I_j + \rho_m I_{mj} + \rho_w I_{wj})}{N_j} \quad 1 \leq i \leq n \\
& d[ln(l_1 p_1)] = s_{l_1} dW_t \\
& l_i p_i = l_1 p_1 * f_{l_i} * f_{p_i} \quad 2 \leq i \leq n \\
& d(logit(\epsilon_i)) = d(ln(\frac{\epsilon_i}{1-\epsilon_i})) = s_{\epsilon_i} dW_t \quad 1 \leq i \leq 6
\end{aligned}
\tag{23}
$$

where f_{l_i} and f_{p_i} are the ratios of the contact rate and transmission probability between age group i and the first age group, respectively. Both the values of f_{l_i} and f_{p_i} are the same as in the un-blanced contact matrix model. It is notable that we treat the effective contact rate $(l_i p_i)$ – the multiplication of the contact rate and the transmission probability of age group i as a single parameter in this re-balanced model, to be simplify and consistent with the previous models. Thus, we use $N_j l_j p_j / \sum_{j=1}^{n} N_j l_j p_j$ to approximately represent the value of $N_j l_j / \sum_{j=1}^{n} N_j l_j$ during implementing the model. Readers interested the complete mathematical equations of the 32-age group state space model with a re-balanced contact matrix [16] and the initial values of all states can refer to Appendix D.

2.2.2. Likelihood function

In the condensation method version [23] of the particle filtering method [22], the weight update rule for a particle given a new observation y_k involves multiplying the previous weight by the value of the likelihood function $p(y_k|x_k)$, where the latter represents the probability of observing the empirical data (denoted as y_k) given the particle state x_k at time k. In this paper, following several past contributions [10, 14, 8, 27], we select the negative binomial distribution as the basis for the likelihood function. We treat the likelihood of observing y_k individuals at time k given an estimated count of incident individuals from the model i_k as follows:

$$
p(y_k|i_k) = \binom{y_k + r - 1}{y_k} p^{y_k} (1-p)^r \tag{24}
$$

where y_k is the empirical data (reported pertussis cases) at time k; $p = i_k/(i_k + r)$ represents the probability that a given reported case is in fact a true incident case, and r is a dispersion parameter. In all scenarios reported in this paper, the value of r is chosen to be 10.

Aggregate model. Because the aggregate particle filtering model lacks the capacity to distinguish between individuals with different age groups as necessary to compare to the yearly age-stratified reported values, the measured data for that model consists of a one-dimensional vector giving the reported cases for successive months. The likelihood function in the aggregate model can then correspondingly be calculated by the value of $p(y_{mk}|I_{rk})$, where y_{mk} is the empirical data as given by the monthly reported measles cases at time k, and I_{rk} is the expected reported cases as calculated by the particle filtering model for each particle.

Age structured model. The weight update rule in the age structured model is similar to that in the aggregate model, with the exception of the updates associated with the close of each year. Specifically, we take the likelihood function at the close of the last month (December) of each year as the product of the likelihood

functions as formulated for each empirical dataset – including both monthly pertussis reported cases across the whole population and the yearly reported cases related to each of the reported age groups considered. The likelihood formulation of age structured models is as follows:

$$
\begin{aligned}
L_{AgeStructuredModel} &= L_{month} * \prod_{i=1}^{n} L_{yearly_i} \\
L_{month} &= p(y_{mk} | \sum_{i=1}^{n} I_{rik}) \\
L_{yearly_i} &= \begin{cases} 1, & \text{if } (k \mod 12) \neq 0 \\ p(y_{yik} | \sum_{i=k-12}^{k} I_{rik}), & \text{if } (k \mod 12) = 0 \end{cases}
\end{aligned}
\tag{25}
$$

where L_{month} is the likelihood function based on the monthly empirical data across the total population, L_{yearly_i} is the likelihood function based on the yearly empirical data for group i, y_{yik} is the yearly empirical data for age group i, and I_{rik} is the reported pertussis cases of age group i at time k.

In the two-age group particle filtering model, we have three empirical datasets – the monthly reported pertussis cases across the whole population and yearly reported cases for each of the two age groups ($n = 2$ in Equations (25)). In the 32-age group particle filtering models, we have employed seven empirical datasets – the monthly reported pertussis cases across the whole population and six datasets of yearly reported cases ($n = 6$ in Equations (25)). As noted previously, the yearly empirical datasets could only be split into 6 age groups.

2.2.3. Proposal distribution

The Condensation Algorithm [28, 23] is applied in this project to implement the particle filter model. It is the simplest and most widely used proposal algorithm, making use of the prior as the proposal distribution [23, 22].

2.3. Empirical data resources

2.3.1. The surveillance data

This paper benefits from the fact that pertussis is formally classified as a notifiable illness for the mid-western Canadian province of Saskatchewan. Pertussis reporting data for Saskatchewan are used as empirical data for the particle filtering models. These data are public aggregate data obtained from the Government of Saskatchewan's "Annual Report of Department of Public Health in the Province of Saskatchewan" [20]. This paper employs two categories of datasets drawn from that report – monthly reported cases aggregated across the entire population, and yearly reported cases in each age group. The latter reflects the fact that in the yearly empirical datasets, the annual reported cases are split into different age groups. Within this dataset, age stratification is inconsistent; as a result, the splitting in some years fails to precisely match stratification of the age groups in the models. For these cases, we proportionally split the yearly empirical reported cases into overlapping age groups within the model. Readers interested the detailed introduction of age deviation of the empirical data can refer to Appendix E.

This study employs pertussis reported cases in Saskatchewan specifically during the pre-vaccination era. The monthly empirical data extends from Jan. 1921 to Dec. 1956, with the dataset offering a total of 432 records. Reporting of age-specific data initiated in 1925, and continued through 1956. Every record contains three features – date, reported cases and population size [19]. 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, as shown in Figure 5, yielding estimated incidence rates rather than incident case counts. 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.

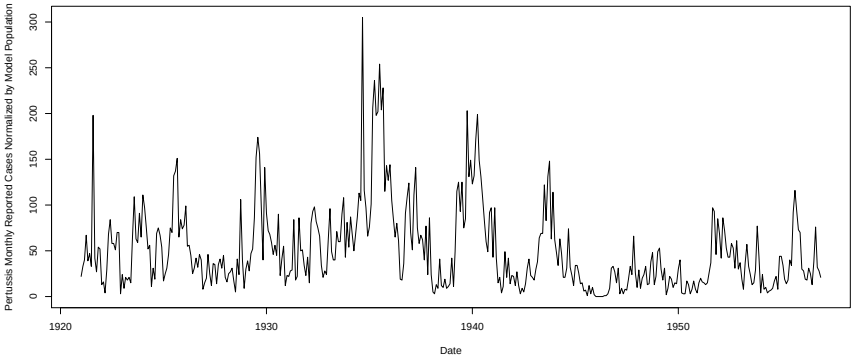


Figure 5: **The monthly reported pertussis cases in Saskatchewan from 1921 to 1956 normalized by the population employed in the model (863,545)**

2.3.2. The demographic data

The demographic parameters play a significant role in the models, particularly the age structured variants. The parameters related to the population are abstracted from the empirical population of Saskatchewan from 1921 to 1956 [19]. The empirical demographic data indicate that the total population of Saskatchewan does not show drastic fluctuation [19] over the year range from 1921 to 1956. During these years, the empirical population lie in the interval from 757,000 to 932,000. The population of Saskatchewan from 1921 to 1956 of each age is depicted in Figure 6. Thus, we let the model population constantly stay in 863,545, which is the average reported population over the years 1921 to 1956 within the Saskatchewan age pyramid [19]. It bears emphasis that for simplicity, we assumed an equilibrium in the population structure – the total population and population among each age group (in the age-structured models) – remain invariant. Similarly, the model assumes fixed values of the population in each age group, according to the previously noted average population.

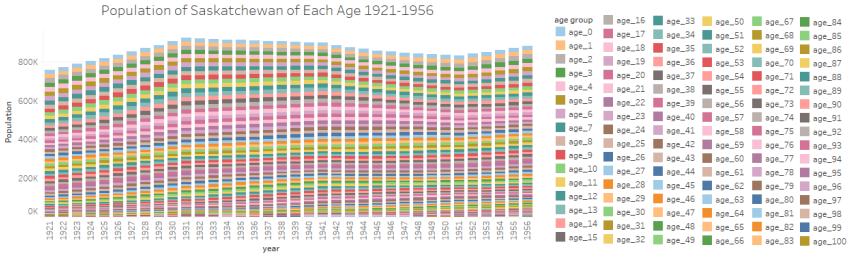


Figure 6: **The age-specific and overall population of Saskatchewan from 1921 to 1956.**

2.4. Introduction of the aggregate population model with calibrated parameters

To evaluate the performance of the particle filtering model when compared to the traditional calibration method, combining with the empirical data, we further constructed a calibration model with the aggregate population using the deterministic epidemiological compartmental model of Equation (1). To be consistent with the particle filtering aggregate model, the parameters and initial values sampled in the particle filtering model are estimated in the calibration model, which are the effective contact rate β, reporting rate C_r and the initial value of the stocks of S, I and R_1. In this calibrated model, the values of the parameters obtained from calibration against the empirical dataset are listed below. The initial value estimated from the calibration process in class S, I and R_1 are 19420, 500, 9960. The value of the effective contact rate (β) is 56.692; it bears emphasis that this value incorporates both a rate of contact and the probability of

transmission. The calibrated value of the reporting rate of pertussis is 0.01. The other parameters are the same as the particle filtering models.

2.5. Classifying outbreak occurrence

The pertussis particle filtering models – combining the particle filtering algorithm and the compartmental models with empirical data – are capable to estimate and predicting the full (continuous) model state over time. Moreover, in this paper, on the basis of having particle filtered up to a certain month, we further perform classification outbreak (outbreak vs. non-outbreak) analysis based on the predicted results (in the next time unit – month) of the particle filtered models. Referred from our previous contribution [8], the function mapping from the continuous predicted results of particle filtering models – predicted reported pertussis cases in the next month – to dichotomous categories of outbreak and non-outbreak can be represented as follows [23, 8]:

$$z_k = f(I_{\mathrm{rk}}^{(i)}) \tag{26}$$

where $\left\{\left\{I_{rk}^{(i)}\right\}_{i=1}^{N_s}\right\}_{k=1}^{T_f}$ indicates the matrix of reported cases of pertussis predicted by the particle filtering model of particle i ($1 \leq i \leq N_s$) at time k ($1 \leq k \leq T_f$). T_f is the total running time of the model.$\{z_k\}_{k=1}^{T_f}$ is the vector of dichotomous predicted classes – $z_k \in \{0,\ 1\}$, where 0 indicates non-outbreak, and 1 indicates outbreak. The value I_{rk} is generated by the particle filtering models. Specifically, I_{rk} equals $\sum_{i=1}^{n} I_{rik}$ (where I_{rik} is the reported pertussis cases of age group i at the time k) in the particle filtering models introduced above.

Two processes are then used to perform classification analysis of the results from the particle filtering models [8]. In the first process, we define a threshold (θ) – mean plus 1.5 times the standard deviation of the empirical monthly reported cases, above which that particle is considered as positing an outbreak. In the second process, we define a threshold of the fraction (θ_k) of particles required to posit an outbreak at time k for us to consider there as being an outbreak. Then, the vector determining whether there is an outbreak of measles in each month – z_k – is calculated. We further denote $\{y_{lk}\}_{k=1}^{T_f}$ as the binary empirical vector of whether a pertussis outbreak indeed obtained at time k, $y_{lk} \in \{0, 1\}$. The calculation method of y_{lk} is similar to that of each particle. If the count of measles reported cases is greater or equal to the threshold θ, the related element in vector y_{lk} is labeled to be outbreak (the value is 1). Otherwise, a non-outbreak is assumed (the value is 0).

Finally, to summarize the performance of the classifier, we employ as a metric the area under the Receiver Operating Characteristic (ROC) curve. Readers interested in additional detail are referred to our previous contribution utilizing a comparable methodology for measles [8].

3. Results

3.1. Results of models incorporating empirical datasets across all timeframe

Recall that to explore the predictive performance of particle filtering in different compartmental pertussis models, four distinct 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 with 2 age groups (denoted as PF_{age_2}), the age-structured particle filtering model with 32 age groups with the original Hethcote contact matrix (denoted as $PF_{age\_32\_Hethcote}$), and the age-structured particle filtering model with 32 age groups with the re-balanced contact matrix (denoted as $PF_{age\_32\_rebalanced}$). In each of the four particle filtering models, 3000 particles are used in the particle filtering algorithm; for clarity in exposition, we sampled the same number when generating the plots of the 2D histogram and for calculating the discrepancy. To compare the accuracy of a particle filtered model against that of a traditional model of pertussis calibrated against comparable data, we have further built a calibrated model of the aggregate population, henceforth denoted *Calibrated*.

By comparing the discrepancy – the root mean square error (RMSE) between the model results and the empirical data – associated with each model, we sought to identify the model offering the greatest predictive validity. We then use the most favorable model to perform prediction and intervention analysis.

Table 1: **Comparison of the average discrepancy (RMSE) for the calibrated model and all four particle filtered pertussis models, considering empirical data across all observation points; parentheses give the 95% confidence intervals.**

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)
$PF_{age\_32\_rebalanced}$	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.

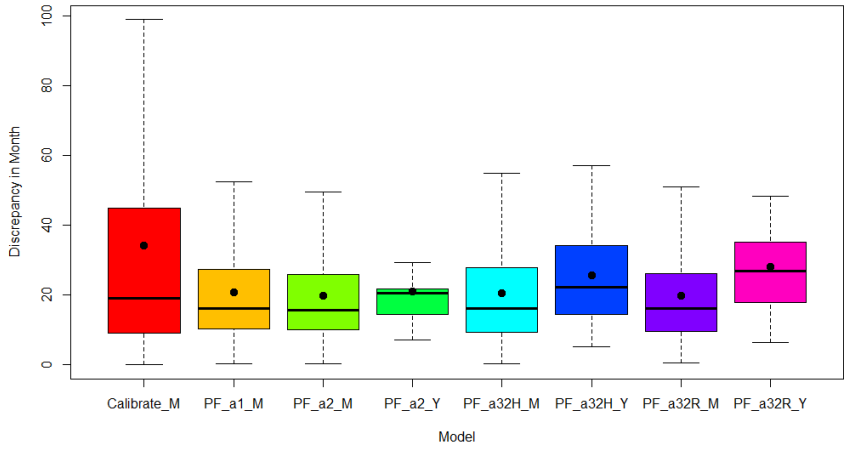


Figure 7: **Boxplot of monthly and yearly discrepancy of all models at monthly observation points, considering empirical data across all observation points.** "Calibrate" indicates the calibration model with aggregate population structure; "PF_a1" indicates the particle filtering model with aggregate population structure; "PF_a2" indicates the particle filtering model with 2 age groups; "PF_a32H" indicates the particle filtering model with 32 age groups and the contact matrix introduced in [16]; "PF_a32R" indicates the particle filtering model with 32 age groups and the re-balanced contact matrix. "_M" indicates the discrepancy of the model comparing model-based monthly results with the monthly empirical data – the pertussis reported cases among all population; "_Y" indicates the sum of discrepancy (of each age group) of the models comparing model-based yearly results with the yearly empirical data – the pertussis reported cases classified into age groups and having adjusted the unit to Month by dividing by 12. It is also notable that the dot in the boxplot indicates the mean value, while the horizontal line indicates the median value.

To assess model results, each of the four particle filtering models was run 5 times with random seeds generated from the same set. We calculated the average and 95% confidence intervals of the mean discrepancy. Table 1 displays the average discrepancies of the four pertussis particle filtering models and the calibrated deterministic pertussis compartmental model, where the discrepancy considers the entire timeframe. These results suggest that particle filtering models significantly improve the predictive accuracy beyond what is achieved via calibration. It is notable that both the calibrated deterministic model and the aggregate particle filtering model only offer monthly average discrepancy, because the yearly observations are stratified by age, but age stratification absent in both such models. Table 1 indicates that the particle filtering models are significantly more accurate than the calibrated model – the average discrepancies of the particle filtering models are significantly lower than those for the calibrated deterministic model. Moreover, although the monthly average discrepancies among the four particle filtering models with different population structure and contact matrix structure are quite close, the particle filtering models PF_{age_2} and $PF_{age\_32\_rebalanced}$ exhibit smaller average discrepancies. With respect to the yearly average discrepancies, Table 1 shows that the age-structured model with two age groups offers better predictive performance than the model with 32 age groups; as noted, the aggregate model lacks the age stratification required to calculate yearly discrepancies. It is notable that the total number of the yearly empirical datasets against which the calibration is assessed is different between the age-structured models with 2 age groups (which is compared with 2 yearly empirical datasets) and that with with 32 age groups (which is compared with 6 empirical yearly datasets). The yearly average discrepancies listed in Table 1 are the sum of the average discrepancy across 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; at the same time, this effect will tend to be limited by the fact that both the model and the empirical values will tend to have smaller counts when applied to a greater number of age groups, yielding a smaller per-age-group discrepancy. On balance, we chose to employ the particle filtering model with two age groups as the minimum average discrepancy model to explore the performance of pertussis outbreak prediction.

Figure 7 shows a boxplot of the distribution of discrepancies among the calibrated model and the four particle filtering models, where a given box in the boxplot summarizes monthly discrepancy estimates for a given model, where those discrepancies are considered over different points in time. Each of the particle filtering models was run 5 times (with the random seed being generated from same set). Then the average monthly and yearly discrepancy among these five runs at each observation time between the particle filtering models and the empirical data are recorded for the boxplot. Both the monthly and yearly (adjusted to units of one Month by dividing by 12) distributions of the discrepancies of each of the age structured models are plotted in Figure 7. This boxplot also indicates that when considered over time, the the discrepancies of all the particle faltering models tend to be smaller than for the calibrated model, although there are similar median discrepancy values. More notable yet is the fact that the discrepancies associated with the calibrated model are significantly more variable than those for the particle filtered models. This suggests that particle filtering improves the consistency of the model's match against empirical data, when compared to a traditional deterministic model with calibrated parameters. Finally, it bears note that the datasets of the discrepancy of the model PF_{age_2} have a particularly narrow distribution, especially when judged in terms of yearly discrepancy.

Figure 8 compares the output of the calibration model and the empirical data. It indicates that the deterministic model even with parameters calibrated against the entire scope of data encounters difficulties in tracking oscillations associated with waning and waxing of pertussis almost across the entire model time horizon, reflecting the approach of the deterministic model towards a stable equilibrium. These results indicate that the particle filtering models considered here can not only decrease the discrepancy between model results and the empirical data, but can further track the oscillation of outbreaks of pertussis.

Taken together, the results shown in Figure 7 and Figure 8 suggest that incorporating particle filtering in the compartmental model of pertussis could enhance simulation accuracy and support more accurate outbreak tracking.

Figure 9 presents the posterior results of the pertussis particle filtering model with aggregate population structure over the entire timeframe. For this diagram at time t, the results of the particle filtering model at time t are sampled according to the weight of all particles following the update to those weights resulting from incorporating the empirical data from time t. Those time-specific values are then plotted; the values of

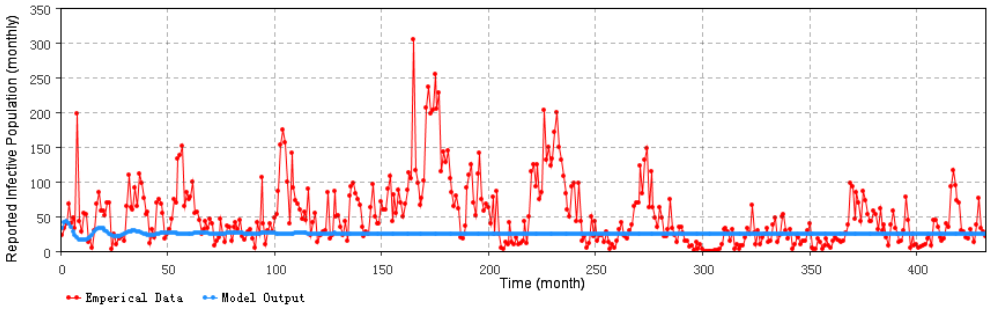


Figure 8: **Reported pertussis cases predicted by the calibration model (monthly).**

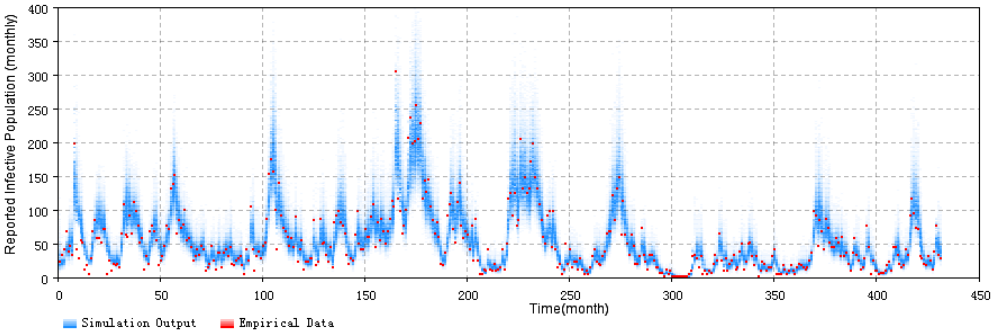


Figure 9: **2D histogram posterior result over the total timeframe for the aggregate particle filtering pertussis model.** The posterior result is sampled following weight updates in light of observations of empirical data arriving at each unit time.

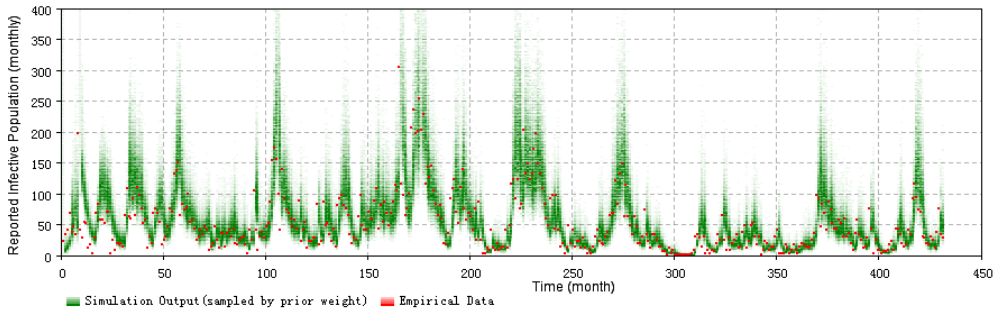
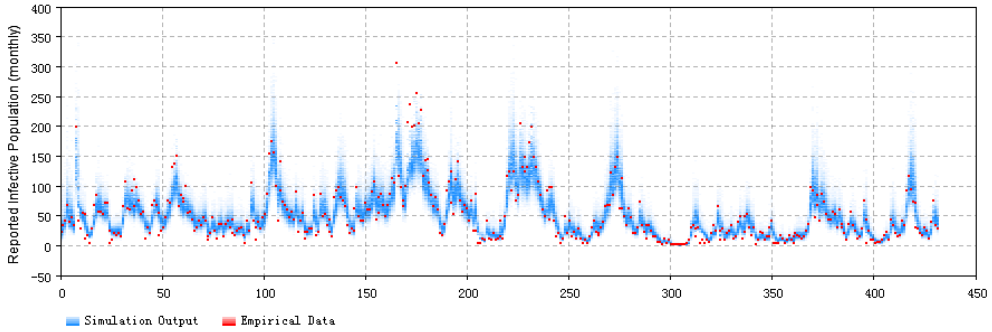
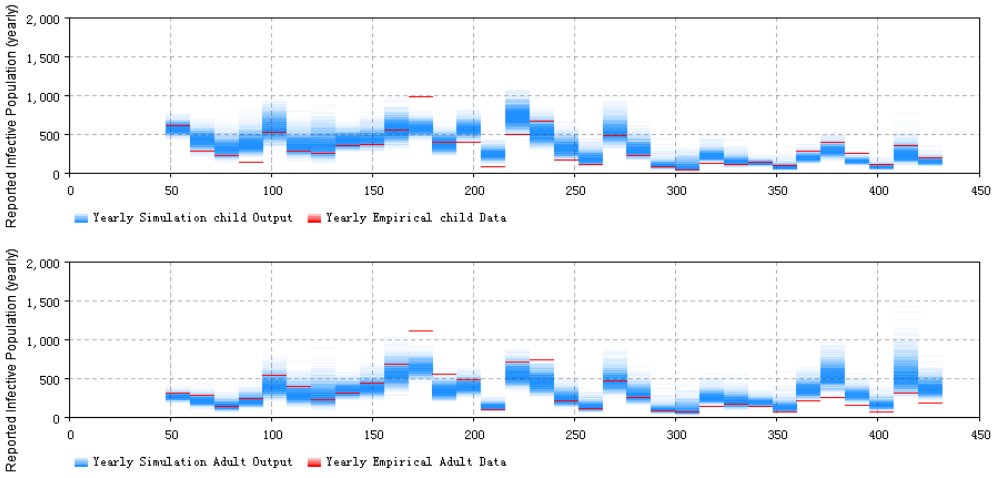


Figure 10: **2D histogram prior result over the total timeframe for the aggregate particle filtering pertussis model.** The prior result is sampled before the weight updates in light of observations of empirical data arriving at each unit time.

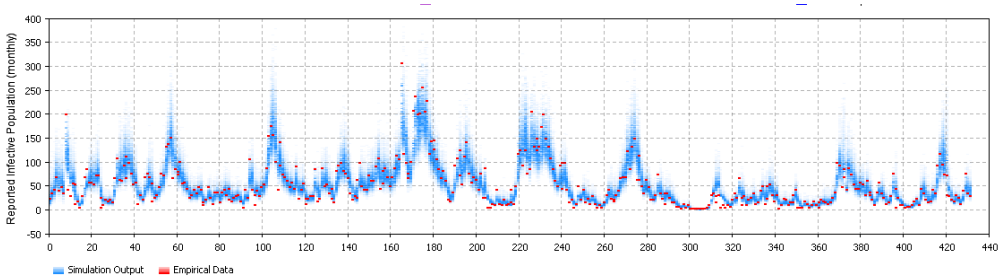


(a)

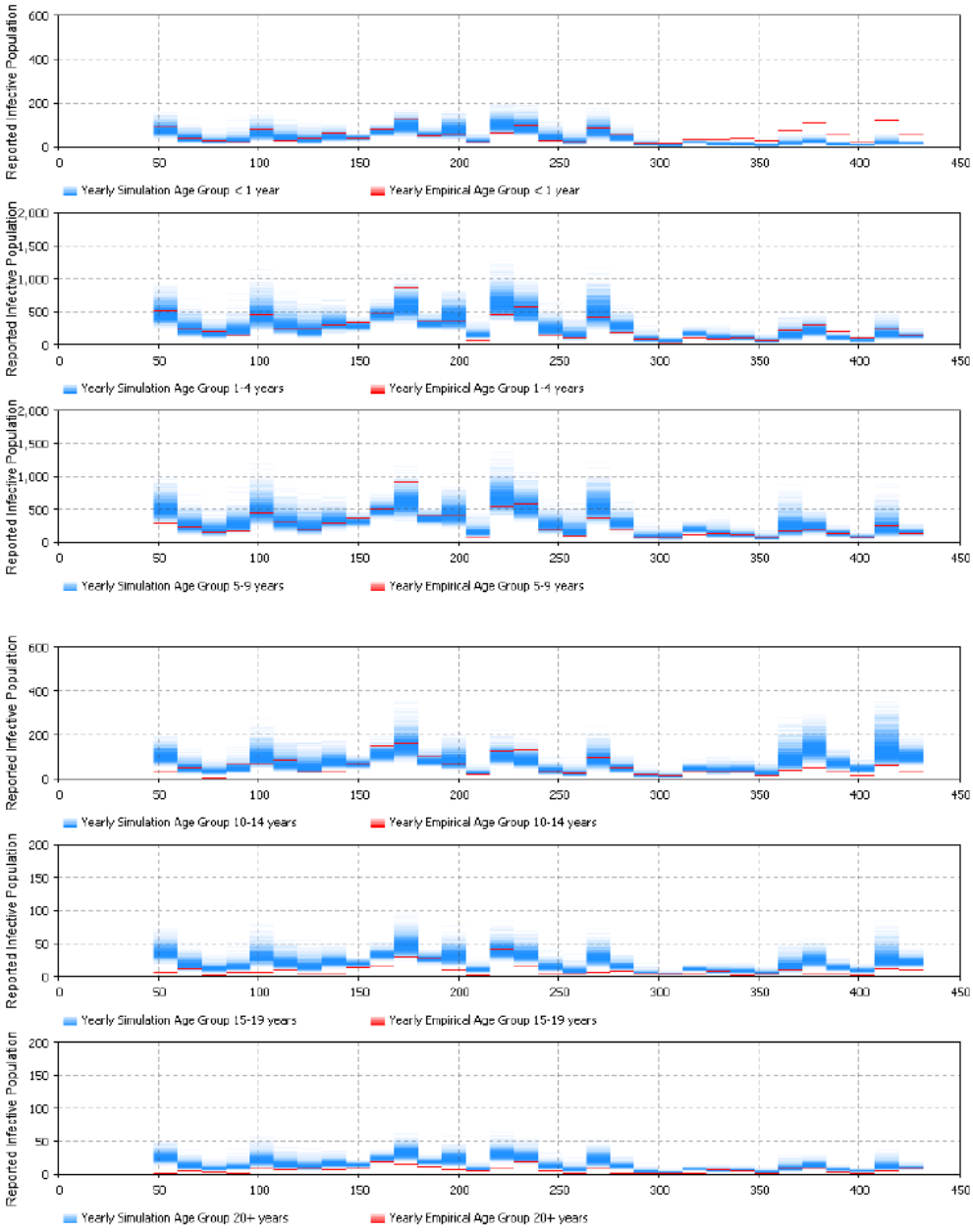


(b)

Figure 11: **2D histogram posterior result over the total timeframe of the two-age stratified pertussis model.** (a) the monthly particle filtering result summed over the entire population. (b) the yearly particle filtering result for the child (top) and adult (bottom) age groups.



(a)



(b)

Figure 12: **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 summed over the entire population. (b) the yearly particle filtering results of each age group of empirical datasets; age groups are successively older from top to bottom.

empirical data points are shown in red, while the sampled posterior distribution of particle filtering model are shown in blue. The blue color saturation indicates the relative density of sampled points within a given 2D bin. Figure 9 demonstrates that most of the empirical data points are located in or near the high density region of the posterior distribution of the particle filtering model. The results shown in that figure further indicate that the particle filtering model has the capability to track the outbreak of pertussis over time, especially compared with the calibrated model whose results are shown in Figure 8. It bears emphasis that the particle filtered results can follow the patterns of empirical data as they arrive; this capacity to update its estimate of model state – both latent and observed – in with new arriving data is central to the function of particle filtering. By contrast, calibration lacks a means of updating the estimate of the model state over time, and is instead relegated to estimating parameter values, rather than the values of the state at varying points in time.

Figure 10 shows the prior results of the pertussis particle filtering model with aggregate population structure for the entire timeframe. For the prior diagram, the results are sampled before the weight update step triggered by arrival of an empirical data point. Compared with the posterior results shown in Figure 9, the prior values of sampled particles of Figure 10 are distributed over a wider range. This difference in dispersion indicates that the weight update process of particle filtering algorithm in this paper has the capability to use an empirical datum to concentrate the distribution of particles in the state space of the particle filtering model into a tighter range offering greater consistency with the empirical datum.

Figure 11 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 age structured pertussis particle filtering model containing 2 age groups (denoted as PF_{age_2}) (on the other). This figure demonstrates that the model PF_{age_2} is capable of tracking and simulating outbreaks of pertussis, as evidenced by the fact that most of the monthly and yearly empirical data (shown in the red dashes) in each month are located in or near the high density region of the sampled distribution of the particle filtering model (shown in blue in the 2D histogram plots).

Figure 12 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 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 simulating the outbreaks of pertussis, as reflected in the fact that most of the monthly and yearly empirical data for each observation point are located in or near the high density area of the results of the particle filtering model.

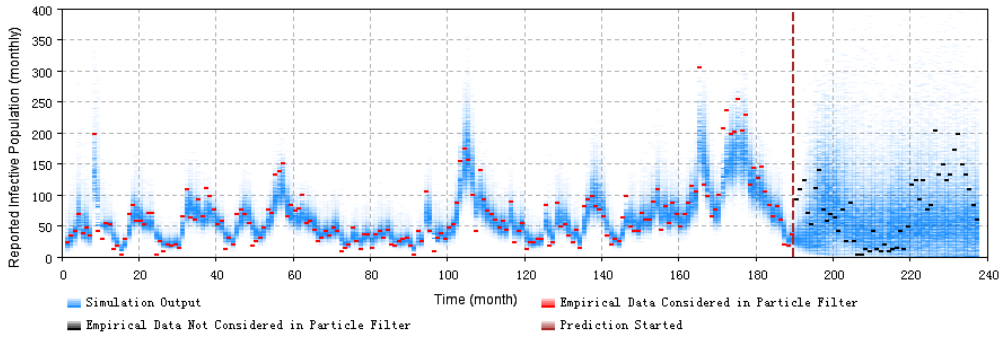
Figure 9, Figure 11 and Figure 12 represent the 2D histogram posterior result of all the particle filtering models, except for the age-structured model of 32 age groups with a re-balanced contact matrix. Results are omitted for this final model as they are highly similar to those for the 32-age-group model using the Hethcote contact matrix, which is itself shown in Figure 12. The 2D histogram plots shown indicate that both the age-structured particle filtering models and the aggregate population particle filtering model have the capability to closely track 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. In contrast to the calibrated model whose results are shown in Figure 8, the particle filtering models are capable of localizing the model's prediction of empirical data near the empirical data, as achieved by concentrating the distribution of particles across the underlying state space. Although the results in Table 1, Figure 7 (for discrepancy), Figure 9, Figure 11, and Figure 12 (for posterior distribution) suggest that all four pertussis particle filtering models are capable of tracking and estimating the pertussis outbreaks, in the interest of brevity of exposition, we selected the minimum discrepancy model – the age-structured particle filtering pertussis model with 2 age groups – to perform the prediction and intervention analysis below.

3.2. Prediction of outbreaks with the minimum discrepancy model

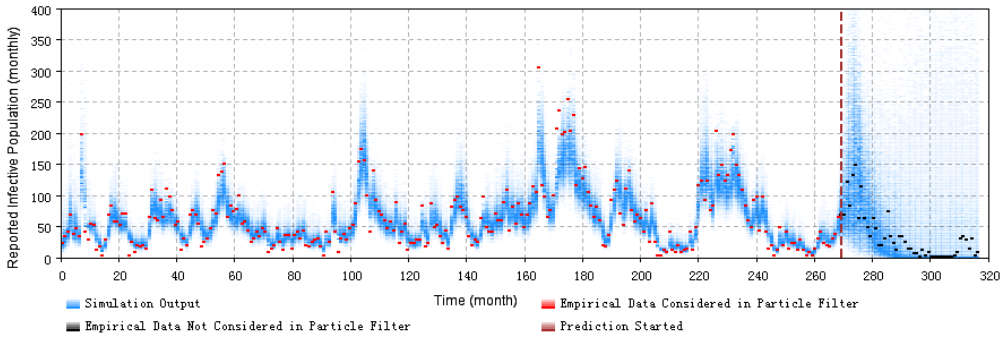
To assess the predictive capacity of the pertussis particle filtering models in anticipating outbreaks, we performed out-of-sample prediction experiments. Informally, each such experiment examines the capacity of the model to project results into the future, having considered data only to some "current" time. That is, the model is particle filtered so as to incorporate data only to up to – but not including – a "Prediction Start Time" (T^*), and then begins projecting (predicting) forward, starting at T^*. More specifically, in

this process, the weights of particles will cease updating in response to observations at time T^*; following that point, all of the particles run without new empirical data being considered. In this paper, all of the prediction experiments are run for 4 years following the "Prediction Start Time" T^*. To evaluate the predictive capacity of the model, we examined the effects of changing the prediction start time T^* so as to pose different archetypal types of prediction challenges. It is notable that the minimum discrepancy model – the age structured model with 2 age groups where the child age group represents children in the first 5 years of their life, and incorporating both the monthly and yearly empirical datasets, as identified in the previous section – is employed to perform all of 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)

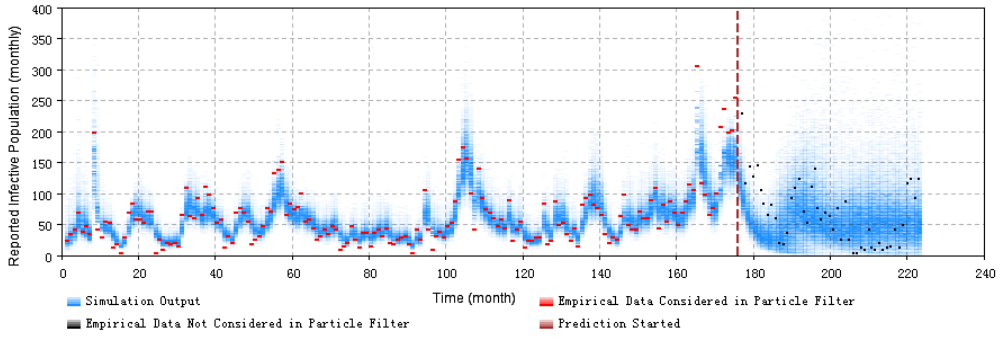


(b)

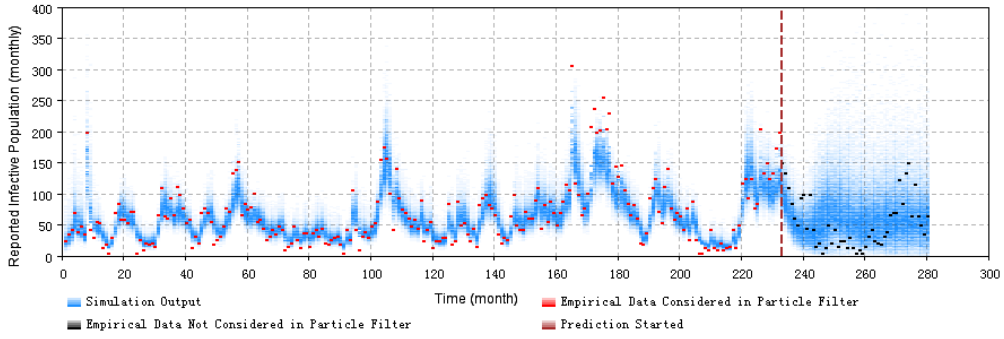
Figure 13: **2D histogram depicting prediction using the minimum discrepancy model from the first or second time points of an outbreak.** (a) prediction from month 190. (b) prediction from month 269.

Figures 13–16 display the prediction results of these situations with respect to the monthly 2D histogram of population-wide reported case counts. In the 2D histogram plots of Figures 13–16, the empirical data having been considered in the particle filtering process (i.e., incorporated in training the models) are shown in red, while the empirical data 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" (T^*) of each experiment.

These prediction results suggest that the pertussis particle filter model offers the capacity to probabilistically anticipate pertussis dynamics with a fair degree of accuracy over a year or so. From the 2D histogram plots, empirical data lying in the projection interval after the prediction start time – and thus not considered by the particle filtering machinery – mostly lie within the high-density range of the particles. Reflecting the fact an ability to accurately anticipate a high likelihood of a coming outbreak could offer substantial value

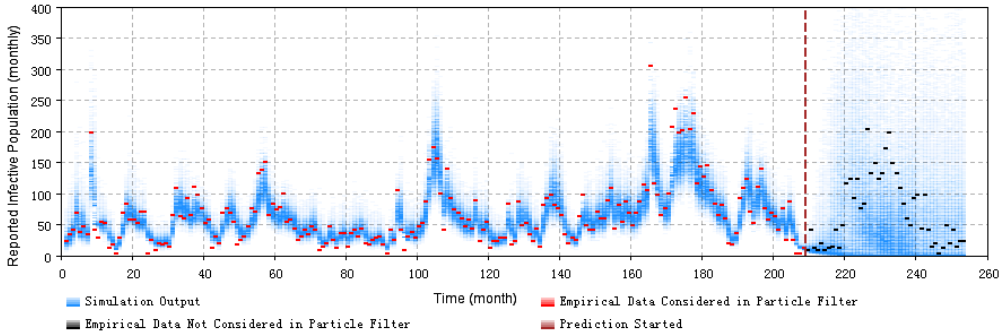


(a)

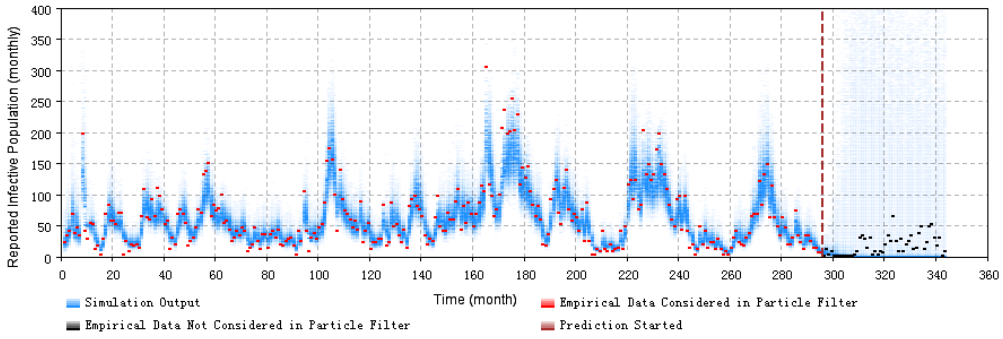


(b)

Figure 14: **2D histogram depicting prediction using the minimum discrepancy model from the peak of an outbreak.** (a) prediction from month 176. (b) prediction from month 233.

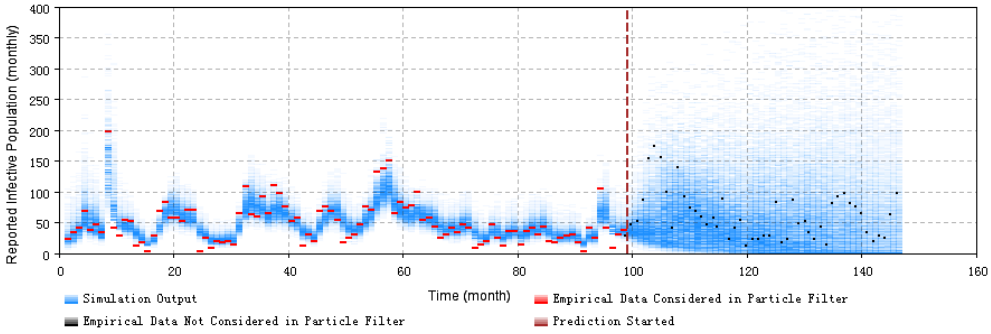


(a)

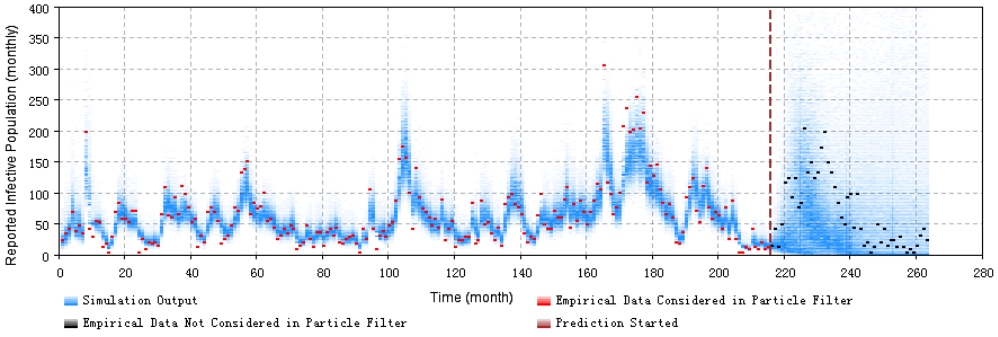


(b)

Figure 15: **2D histogram depicting prediction using the minimum discrepancy model from the end of an outbreak.** (a) prediction from month 209. (b) prediction from month 296.



(a)



(b)

Figure 16: **2D histogram depicting prediction using the minimum discrepancy model prior to the next outbreak.** (a) prediction from month 99. (b) prediction from month 216.

for informing public health agencies with accurate predictions of the anticipated evolution of pertussis over coming months, the next section formally evaluates the performance a simple classifier as to whether the next month will be subject to an outbreak or not, where that classifier uses a very simple prediction scheme constructed atop the particle filter model.

3.3. Prediction of classifying outbreak occurrence of the minimal discrepancy model

Beyond assessing the use of particle filtering models for predicting forward pertussis transmission more generally, we also used the lowest discrepancy particle filter pertussis model (PF_{age_2}) to dichotomously predict occurrence of a pertussis outbreak within the next month.

Figure 17 displays an evaluation of the predictive performance in the form of an ROC curve. The Area Under the Curve (AUC) of the ROC curve is 0.913, suggesting that it is possible to achieve both high specificity and high sensitivity. Figure 18 shows the boxplot of residuals (difference between predicted model result and empirical data) of sampled particles (by weight) at each time point where empirical data comes in (each month). Two points bear emphasis. Firstly, these results depict prior model predictions – that is, those predicted by the model before the new data is observed. Secondly, Figure 18 excludes the first 10 months (empirical data points) of the time horizon, during which the particle filtering model is not stable enough due to insufficient incorporation of empirical data. Figure 18 indicates that for results of the next time point (month in this paper), the prior prediction of the particle filtering model are quite close to those of the empirical data – although the empirical data at each predicted time point are not yet incorporated to ground the model.

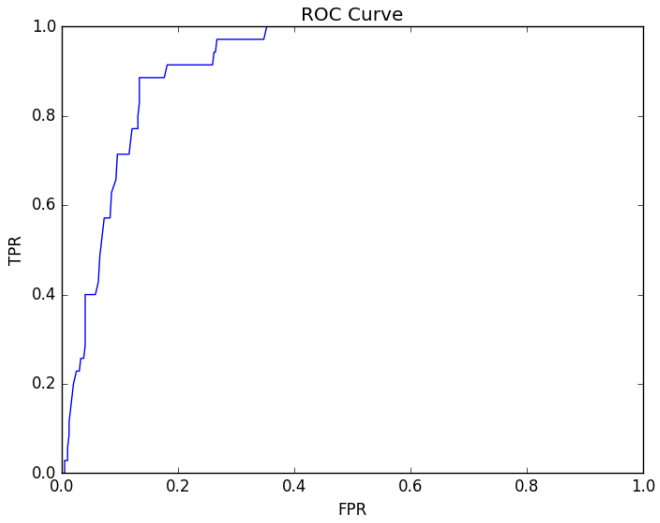


Figure 17: **ROC curve of the binary outbreak classifier of the minimum discrepancy model.**

3.4. Intervention with the minimum discrepancy model

The capacity of particle filtering to accurately estimate (sample from) the latent state of a pertussis model makes this technique capable of both estimating the entire latent state and using that estimation to project patterns of pertussis spread and waxing and waning of incidence in the near term, and to anticipate outbreak occurrence. The capacity to perform such state estimation within a mechanistic model also supports particle filtering models in more accurate simulation of the tradeoffs between intervention strategies, despite their counterfactual character.

In this section, we have implemented several experiments to simulate stylized public health intervention policies, based on the minimum discrepancy particle filtering pertussis model identified above. The stylized intervention strategies are characterized in an abstract way for demonstration purposes, and are typically performed before or at the very beginning of an outbreak. For simplicity, we examine them as a historical

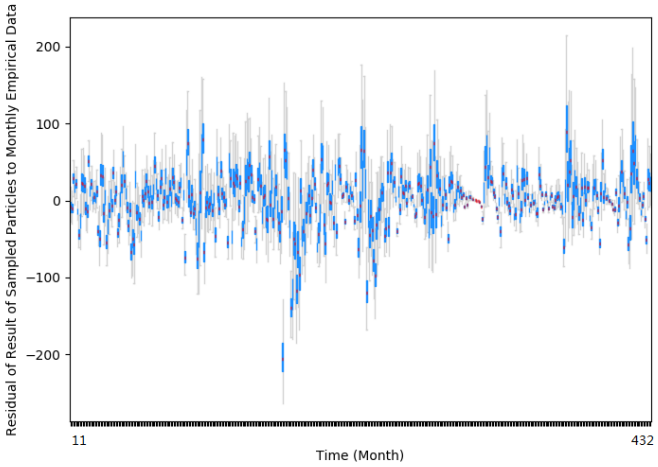


Figure 18: **Boxplot of the residuals of results of prior prediction by sampled particles of the minimum discrepancy model.**

counterfactual that takes place at a certain historic context. Moreover, to support easy comparison with the baseline prediction results 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. Moreover, in order to appropriately characterize how such techniques could be employed in public health scenario planning, we assume here that the start month of the intervention (month 269) is the "current time" in the scenario – that we wish to asses the effects of that intervention considering only the data available up to but not including month 269, and simulate the results of the intervention forward from that point. The baseline prediction result of the minimum discrepancy model absent any interventions is shown in Figure 13 (b). We examine below the impact of two stylized intervention policies – hygeine-enhancing and vaccination.

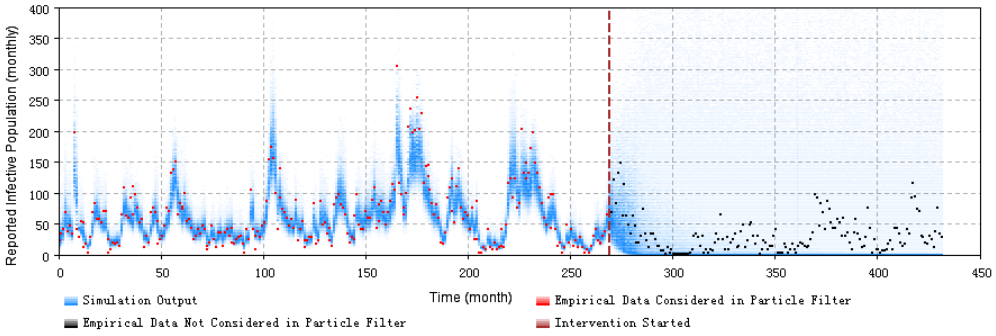


Figure 19: **2D histogram of model-based projections of pertussis incident case counts when simulating a hygeine-enhancing intervention during a pertussis outbreak.** This is realized by decreasing the contact rate by 20%.

Figure 19 and Figure 20 display results from simulation of hygeine-enhancing intervention strategies [29] whose effects are characterized as decreasing the contact rate parameter by 20% and 50% when compared to its pre-intervention value, respectively. Similarly to the 2D histogram plot of the baseline prediction result shown in Figure 13 (b), the red dots represent the empirical data incorporated into the particle filtering model (here, up to just prior to the point of intervention); by contrast, the black dots represent empirical data not incorporated in the model, but presented for comparison purposes. It bears emphasis that because the interventions being characterized are counterfactual in character – i.e., did not in fact take place historically – the empirical data shown in black reflect the baseline context, which lacked an intervention of the sort simulated here. By comparing the hygeine-enhancing intervention results (see Figure 19 and

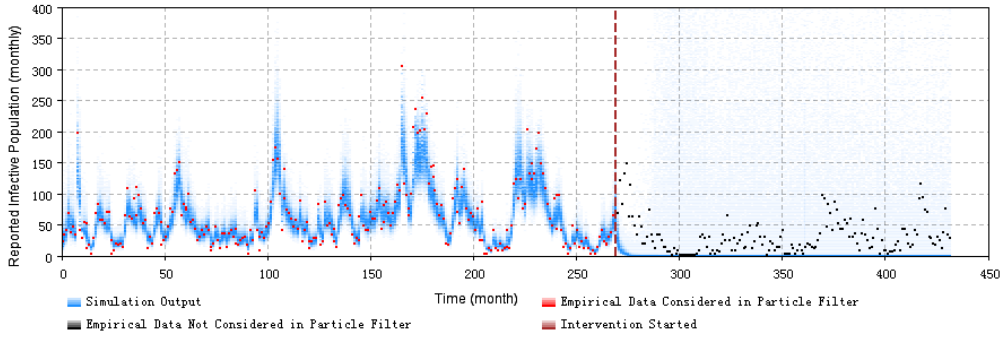


Figure 20: **2D histogram of model-based projections of pertussis incident case counts when simulating a hygeine-enhancing intervention during a pertussis outbreak.** This is realized by decreasing the contact rate by 50%.

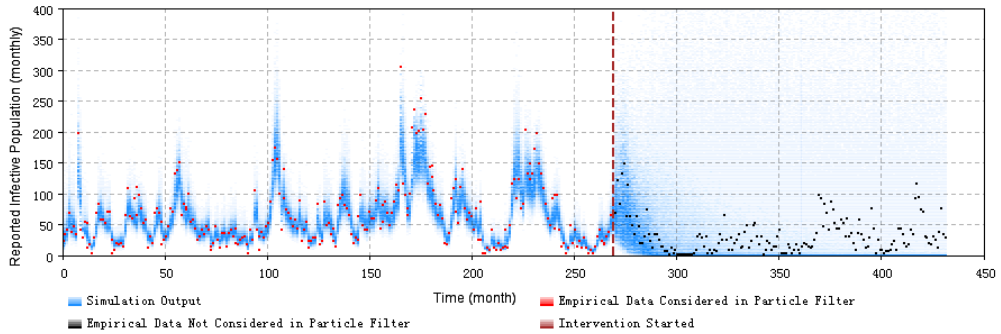


Figure 21: **2D histogram of model-based projections of pertussis incident case counts when simulating an outbreak-response immunization campaign.** This is realized by characterizing a stylized elevated vaccine-induced protection level among 20% of the population.

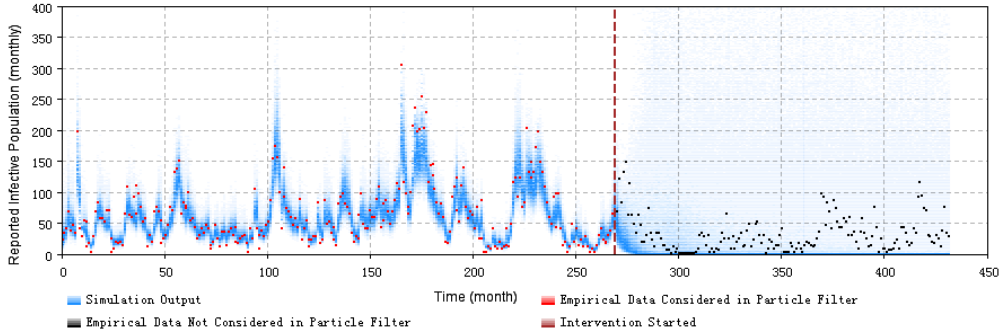


Figure 22: **2D histogram of model-based projections of pertussis incident case counts when simulating an outbreak-response immunization campaign.** This is realized by characterizing a stylized elevated vaccine-induced protection level among 50% of the population.

Figure 20) with the baseline model result without intervention shown in Figure 13 (b) and the empirical data during the intervention period (the black-markers indicating historic data points lying after the triggering of the intervention, and not incorporated into the particle filtering model), we can see that, although the interventions are implemented in a stylized fashion, by virtue of the particle filter's ability to estimate the underlying epidemiological state at the point of intervention through the transmission model, the particle filtered pertussis model is capable of using the estimated latent state to serve as the basis for probabilistically evaluating pertussis related intervention policies.

To simulate an immunization intervention during a pertussis outbreak, a vaccination parameter is incorporated into the simulation model, so as to represent the fraction of the population whose immunity

status is elevated as a result of the intervention. Specifically, recall that the pertussis model characterizes a chain of successively higher levels of vaccine-induced protection. This parameter specifies the fraction of the population that should be moved from their pre-vaccination classification – as characterized by the model compartment in which they reside – to the compartment representing the next higher level of vaccination (following vaccination). Figure 21 and Figure 22 show the results of the vaccination intervention. The layout and organization of the 2D histogram plots of the vaccination interventions with Figure 21 and Figure 22 mirrors that of the hygeine-enhancing plots of Figure 19 and Figure 20.

The results of pertussis interventions demonstrate that by virtue of its ability to estimate the underlying epidemiological state of the model (and thus the system characterized by that model), the use of particle filtering with pertussis models supports evaluation of public health intervention policies to prevent or control pertussis outbreaks.

4. Discussion and conclusion

This paper contributes a new method for anticipating, tracking, controlling and preventing pertussis outbreak patterns by integrating a particle filtering algorithm with a mechanistic pertussis compartmental model and empirical incidence data. This contribution represents the first time that particle filtering has applied to pertussis transmission dynamics, and demonstrates the great promise of this technique. The models examined here demonstrated a notable degree of accuracy in predicting pertussis dynamics over multi-month timeframes – the 2D histogram plots comparing the empirical data and samples from the posterior distributions of the particle filtering models' projected 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 the historic data. The results of prediction analysis based on the minimum discrepancy model suggest that particle filtering approaches offer notable strengths in predicting of occurrence of pertussis outbreak in the subsequent month. Moreover, the discrepancy of the pertussis particle filtering model's predictions vs. observed data is reduced by approximately 60% when compared with a traditional calibration model, demonstrating a significant enhancement in model prediction ability. Additionally, it bears emphasis that the calibrated deterministic model encounters marked difficulties in tracking the fluctuation of the outbreak pattern of the calibration model; by contrast, the particle filtered model is capable to tracking stochastic fluctuations associated with pertussis, while still mechanistically capturing the impacts of such stochastics on the latent underlying dynamics of susceptibles, exposed individuals, etc. Further to this point, it is of great significance to the success and promise of these methods that pertussis particle filtering models support effective estimation of the entire state of the pertussis transmission models – and thus the systems that they represent – during those periods when the empirical datasets are available, including latent states of strong interest, such as those associated with waning of natural immunity and differing levels of infection severity. Combined with the capability to perform outbreak projections, such particle filtering models can serve as powerful tools for understanding the current epidemiology of pertussis in the population, for projecting forward evolution of pertussis spread – including occurrence of outbreaks.

Beyond that, in a further contribution that also benefits strongly from the capacity to estimate latent state, this research further marks the first instance of research demonstrating the capacity to perform public health intervention experiments using particle filtered models.

Despite the strengths of these contributions, there remain a number of important limitations of this work, and priorities for future research. We briefly comment on several below.

This work investigated the performance of four particle filtering models, including an aggregate population model, a two age group-stratified population model, and 32 age group population models using – alternatively – a contact matrix derived from Hethcote (1997) [16] and (separately) a re-balanced contact matrix. Although the results of all four of these particle filtering models matched the empirical data quite well, the minimum discrepancy model proved to be the 2 age group age-stratified particle filtering model in which individuals in the child age group represent children in the first 5 years of life, and which incorporates both monthly and yearly empirical datasets. In this regard, it is notable that according to the mathematical deduction of the age structured population model introduced in [8] – and adapted to pertussis in this research – the model can simulate the aging rate (c_i) more accurately with more age groups considered in the age-structured model. However, in this paper, the 32 age group particle filtering models fail to demonstrate

improved performance – as measured by the discrepancy of model predictions from the empirical data – when compared with the two age group particle filtered models. We provide below some comments on possible reasons. Firstly, the stochastic processes considered in both the 32 age group age-structured model and the two age-group stratified model are different, especially in their characterization of the stochastic evolution of the contact rate. Secondly, the likelihood functions employed in this project – which are captured as the product of negative binomial density functions across all empirical datasets and sharing a common dispersion parameter – may be too naïve to capture the difference between the age groups within the empirical datasets. Thirdly – and perhaps most significantly – as the number of age groups increase, the the state space dimensionality of the particle filtering models increases dramatically. This latter issue must be considered in light of the limitations of the particle filtering algorithm, particularly the fact that the particle filtering method employing the condensation algorithm may encounter problems in high-dimensional systems. In such systems, the probability density functions would be more involved; addressing this using the functional form of the likelihood functions employed may require high dispersion, due to the difficulty of representing the details of the multivariate likelihood function using the product of simple probability density functions. Research is needed into more effective multivariate likelihood function design. The relationship between the the nominal state space dimensionality and the number of particles required for effective particle filtering also merits additional research, particularly in light of observed limitations in the benefit of particle filtering for high dimensional models [13]. Finally, when comparing the discrepancy for distinct models, our lack of normalization for the count of datasets used may lead to artificially stacking the comparison against the 32 age group model; while the 32 age group does not exhibit markedly better discrepancy against the monthly aggregate observations than does the 2 age group model, this consideration suggests that it may be stronger than the yearly discrepancy numbers would suggest.

It is also worth emphasizing the critical role that stochastic process noise within the state space models plays within successful particle filtering, and the practical challenges associated with managing such noise. The stochastics associated incorporated into the model represent a composite of two factors. Firstly, there is expected to be both stochastic variability in the measles infection processes (e.g., those that are prominent for small incident case counts) and some evolution in the underlying transmission dynamics in terms of changes in mixing and the reporting rate. Secondly, incorporation of such stochastic variability into the particle filtered model allows for 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 fallibility, and allowing a requisite variety in the distribution over particle states, such that the particle filtered model is more open to correction by new observations. While results in both the estimation and prediction periods are sensitive to the degree of stochastics involved, such model stochastics impact the particle filtered model in distinct distinct ways during these periods. Taking into account these influences, the investigations demonstrated the importance of keeping the noise in the particle filtering models controlled within a proper range, by tuning the parameters of diffusion coefficients in the stochastic processes related to the Brownian motion. The need to characterize and tune stochastic noise effectively can impose limits on the speed with which particle filtering models can be prepared for a new sphere of application.

The initial values of the age-structured population models in this paper 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 compartments within a given age group is estimated by the particle filtering algorithm by setting the initial values of compartments in a proper range following a uniform distribution, but maintaining a total number of individuals for that age group across the compartments. Especially in building the 32-age-groups particle filtering models, much time and efforts is dedicated to estimation of the population distribution among the latent states.

While application of particle filtering to pertussis dynamics is not without its challenges, the approach examined here demonstrates great promise for creating models that are automatically kept abreast of the latest evidence, for understanding the underlying epidemiology of pertussis in the population – including the balance of the population at varying levels of immunity – for projecting forward pertussis dynamics and outbreak prediction over a year's time, and for evaluation of counter-factual interventions. The results of this paper – which represents both the first application of particle filtering to pertussis, the first to demonstrate the capacity to accurately predict pertussis outbreaks in the pre-vaccination era, and the first to use particle filtering to assess the tradeoff between public health applications – suggest that particle

filtering may represent an important element in the arsenal of public health tools to address the increasingly difficult challenge of controlling pertussis in the context of vaccine hesitancy and waning of both natural- and vaccine induced- immunity.

5. Future work

The growing risk of pertussis outbreaks triggered by combinations of vaccine hesitancy and waning dynamics from earlier generations of pertussis vaccination, has elevated the urgency and prominence of questions about the rate at which immunity to pertussis wanes, especially about vaccine effectiveness over time [30, 31, 32, 33]. We identify here two notable needs for future work responsive to such dynamics Firstly, there is a keen need for application of the models presented here to data and dynamics from the vaccination era. While vaccination elements of the models discussed here are only glancingly tapped by this research (in the context of demonstrating capacity to reason about the effects of a stylized immunization intervention), because of their incorporation into the existing model structure, extension of this work to vaccine-era dynamics should require only limited changes to the models involved.

A second need relates to the fact that we choose to employ a constant value for the waning of immunity that is drawn from [16]. In the future, to arrive at more informed parameter estimates for models and to contribute to discussion concerning the empirical rate at which vaccine-induced as well as (separately) naturally induced pertussis immunity wanes, and drawing on the success of our past work in this area [34], we propose to formally estimate the value of waning immunity for the the particle filtered pertussis models from a posterior distribution within Particle Markov Chain Monte Carlo (PMCMC) techniques, by incorporating empirical data on reported pertussis cases from the vaccination era.

These and other lines of future work offer substantial promise in extending the already strong potential demonstrated here for using mechanistic transmission models informed by the machine learning approach of particle filtering to contribute to enhancements in pertussis prevention and control by providing a tool to improve understanding of underlying complex epidemiology of pertussis, to anticipate pertussis dynamics in the population, and to rigorously assess the tradeoffs between counterfactual intervention tradeoffs in light of uncertainties in both model and empirical data.

Appendix A. The introduction of mathematical models

In this compartmental model of pertussis in the pre-vaccination era, the total population is divided into 8 distinct epidemiological classes. Newborns enter directly into class S of fully susceptible individuals. If a fully susceptible individual contacts an infective individual and is successfully transmitted pertussis, this previous susceptible person becomes infectious and enters the class (state) I of full infectives. Infective individuals in state I of have full cases of pertussis, with all of the usual symptoms. When individuals recover from the state I of infectives, they achieve full immunity and enter state R_4. In this state, they are fully protected and can not be infected by pertussis. However, as time goes by, their immunity wanes and they enter into a less strong immunity class of R_3. When individuals in class R_3 are exposed to an infective, they are assumed to return to the highest immunity class of R_4 without becoming infectious. Otherwise, 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 is sufficiently (re-)exposed to an infective for transmission to occur, the infected individual enters the I_w state with weak infectivity. Individuals in the I_w class have the weakest infective capability to infect a susceptible. After they recover, the individuals in class of I_w then secure the highest immunity, (re-) entering the class of R_4 from which they originally waned. By contrast, if people in the class of R_2 are not re-exposed to the infectives, their immunity continues waning, and they enter the minimally immune class of R_1. Similarly, if a person in class R_1 is 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), but exhibit a lower infectious capability compared to the fully infective individuals in I. When recovered, the individuals in class I_m enter the class 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 whence they originated at birth. Given the presence of multiple infection states (I, I_w, I_m) as well as multiple levels of immunity (R_1, R_2, R_3, R_4), three invariants bears noting. Firstly, regardless of the

pre-existing level of immunity, following recovery from an infection, that individual always returns to the full level of natural immunity (R_4). Secondly, in any of the recovered states (R_1, R_2, R_3, R_4), immunity continues to wane absent re-exposure. Thirdly, as the level of immunity is reduced, the severity of resulting infectiousness rises, with no infectiousness being possible at all from exposure in states R_3 and R_4.

Appendix B. Proof of the n-square grows of the unknown parameters

In this part, we 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, which is

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

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 parameters 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:

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

the total number of equations in Equation (B.2) is $n + \binom{2}{n} = n + n(n-1)/2 = (n^2+n)/2$. Finally, in this method of calculating the contact matrix, the number of unknown parameters is $(n^2-n)/2$. 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 C. The mathematical deduction of the force of infection with re-balanced contact matrix

In the beginning, we introduce the method of calculating the basic contact matrix which is balanced already and with one unknown parameters. Before introduced, we 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 represented as follows:

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

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, we 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:

$$f_{ij} = (1.0 - \epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j} \right) \tag{C.2}$$

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, we have employed the method introduced in [21] 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:

$$\Delta_{ij} = \frac{N_i l_i f_{ij}}{N_j l_j f_{ji}} = \frac{\epsilon_i N_i l_i N_j l_j}{\epsilon_j N_i l_i N_j l_j} = \frac{\epsilon_i}{\epsilon_j}, \quad i \neq j \tag{C.3}$$

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 [21], the equations of l_{ij} and l_{ji} could be defined separately:

$$\begin{aligned} l_{ij} &= l_i \Delta_{ij}^{\theta} = l_i \left(\frac{\epsilon_j}{\epsilon_i} \right)^{\theta} \\ l_{ji} &= l_j \Delta_{ij}^{-(1-\theta)} = l_j \left(\frac{\epsilon_j}{\epsilon_i} \right)^{-(1-\theta)} \end{aligned} \tag{C.4}$$

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, we could get the value of the parameter of θ ($\theta = 0.5$). Substitute the value of θ ($\theta = 0.5$) to Equation (C.4), the matrix of contact rate – l_{ij} could be generated as follows:

$$l_{ij} = l_i \left(\frac{\epsilon_j}{\epsilon_i} \right)^{0.5} \tag{C.5}$$

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

$$\begin{aligned} l_{ij} f_{ij} &= l_i \left(\frac{\epsilon_j}{\epsilon_i} \right)^{0.5} \left[(1.0 - \epsilon_i)\delta_{ij} + \epsilon_i \left(\frac{N_j l_j}{\sum_{j=1}^{n} N_j l_j} \right) \right] \\ \lambda_i &= p_i \sum_{j=1}^{n} \frac{l_{ij} f_{ij} (I_j + \rho_m I_{mj} + \rho_w I_{wj})}{N_j} \end{aligned} \tag{C.6}$$

where p_i is the transmission probability of age group i.

Appendix D. The state space models in Particle filter implementation

The mathematical system dynamics models are employed as the governing equations underlying the state space model. 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 mathematical model – pure Ordinary Differential Equations (ODEs), models of infection transmission are often related to more complex dynamics – such as parameters evolving according to stochastic processes.

In this paper, we employ the identity method introduced in the previous contribution of [8] to let some constant parameters in the pure mathematical models change dynamically. Specifically, if the parameter varies over the entire range of positive real numbers, we treat the natural logarithm of this parameter as undergoing a random walk according to a Wiener Process (Brownian Motion) [25, 26, 8]. Otherwise, if the parameter varies over the range [0,1], we characterize the logit of this parameter as also undergoing Brownian Motion.

Appendix D.1. 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. Then, three stochastic processes are considered in the implementation of the aggregate particle filtering model. The first is 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). We 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) [25, 26, 8]. It is notable that we assume 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:

$$
\begin{aligned}
&\frac{dS}{dt} = Nv - A_I - \mu S + \iota R_1 \\
&\frac{dI}{dt} = A_I - (\gamma + \mu)\, I \\
&\frac{dI_m}{dt} = A_{I_m} - (\gamma + \mu)\, I_m \\
&\frac{dI_w}{dt} = A_{I_w} - (\gamma + \mu)\, I_w \\
&\frac{dR_1}{dt} = \alpha R_2 - A_{I_m} - (\mu + \iota) R_1 \\
&\frac{dR_2}{dt} = \alpha R_3 - A_{I_w} - (\mu + \alpha) R_2 \\
&\frac{dR_3}{dt} = \alpha R_4 - (\lambda + \mu + \alpha) R_3 \\
&\frac{dR_4}{dt} = \gamma(I + I_m + I_w) + \lambda R_3 - (\alpha + \mu) R_4 \\
&\lambda = \frac{\beta(I + \rho_m I_m + \rho_w I_w)}{N} \\
&N = S + I + I_m + I_w + R_1 + R_2 + R_3 + R_4 \\
&dln(\beta) = s_\beta dW_t \\
&d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t
\end{aligned}
\tag{D.1}
$$

$$
\begin{aligned}
I_k &= \int_{k-1}^{k} (A_I + \rho_m A_{I_m} + \rho_w A_{I_w})\, dt \\
I_{rk} &= I_k C_r \\
A_I &= \frac{Poisson(\lambda S \Delta t)}{\Delta t} \\
A_{I_m} &= \frac{Poisson(\lambda R_1 \Delta t)}{\Delta t} \\
A_{I_w} &= \frac{Poisson(\lambda R_2 \Delta t)}{\Delta t}
\end{aligned}
$$

The parameters related to the transmission of pertussis in this model are referred from the research of Hethcote (1997) [16]. The demographic parameters of this model are got from the Annual Report of the Saskatchewan Department of Public Health [20] and the age pyramid of Saskatchewan [19]. Then, the parameters of the pertussis aggregate state space model – Equations (D.1) are specified in Table D.2, while the initial values of the stocks are listed in Table D.3.

Table D.2: **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 D.3: **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

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

The mathematical model with two age groups 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 – mathematical model – 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

Table D.4: **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

Table D.5: **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

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) [25, 26, 8]. 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_{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}
\end{aligned}
\tag{D.2}
$$

$$
\begin{aligned}
d(ln\beta_c) &= s_{\beta_c} dW_t \\
d(ln(\frac{f_{cc}}{1-f_{cc}})) &= s_{cc} dW_t \\
d(lnM_a) &= s_{M_a} dW_t \\
\beta_a &= M_a \beta_c \\
d(ln(\frac{C_r}{1-C_r})) &= s_r dW_t \\
f_{ca} &= 1 - f_{cc} \\
f_{ac} &= \begin{cases} \frac{N_c \beta_c}{N_a \beta_a}(1-f_{cc}), & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a}(1-f_{cc})\right] < 1.0 \\ 1.0, & \text{if } \left[\frac{N_c \beta_c}{N_a \beta_a}(1-f_{cc})\right] \geq 1.0 \end{cases} \\
f_{aa} &= 1 - f_{ac} \\
\mu_c &= \frac{N_a}{N_c} v_a - \omega \\
\mu_a &= \frac{N_c}{N_a} \omega
\end{aligned}
$$

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

In this paper, we have built a 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_w) 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 D.4. The initial values of each stocks in this two-age-group particle filtering model are listed in Table D.5.

Appendix D.3. The age-structured model of 32 age groups ($n = 32$) with the Hethcote contact matrix

We employ the pure ODEs model – the age-structured model of 32 age groups introduced in the paper of Hethcote (1997) [16] as the base model. 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 parameter related to contact rate of the first age group (denoted as $l_1/\sqrt{D}$) 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) [25, 26, 8]. :

$$\frac{dS_1}{dt} = \sum_{j=1}^{n} v_j N_j + \iota R_{11} - A_{I_1} - (c_1 + \mu_1) S_1$$
$$\frac{dS_i}{dt} = c_{i-1} S_{i-1} + \iota R_{1i} - A_{I_i} - (c_i + \mu_i) S_i \quad 2 \le i \le n$$
$$\frac{dI_1}{dt} = A_{I_1} - (c_1 + \gamma + \mu_1) I_1$$
$$\frac{dI_i}{dt} = c_{i-1} I_{i-1} + A_{I_i} - (c_i + \gamma + \mu_i) I_i \quad 2 \le i \le n$$
$$\frac{dI_{m1}}{dt} = A_{I_{m1}} - (c_1 + \gamma + \mu_1) I_{m1}$$
$$\frac{dI_{mi}}{dt} = c_{i-1} I_{m,i-1} + A_{I_{mi}} - (c_i + \gamma + \mu_i) I_{mi} \quad 2 \le i \le n$$
$$\frac{dI_{w1}}{dt} = A_{I_{w1}} - (c_1 + \gamma + \mu_1) I_{w1}$$
$$\frac{dI_{wi}}{dt} = c_{i-1} I_{w,i-1} + A_{I_{wi}} - (c_i + \gamma + \mu_i) I_{wi} \quad 2 \le i \le n$$

$$
\begin{aligned}
&\frac{dR_{11}}{dt} = \alpha R_{21} - A_{I_{m1}} - (\iota + c_1 + \mu_1)R_{11} \\
&\frac{dR_{1i}}{dt} = c_{i-1}R_{1,i-1} + \alpha R_{2i} - A_{I_{mi}} - (\iota + c_i + \mu_i)R_{1i} \quad 2 \le i \le n \\
&\frac{dR_{21}}{dt} = \alpha R_{31} - A_{I_{w1}} - (\iota + c_1 + \mu_1)R_{21} \\
&\frac{dR_{2i}}{dt} = c_{i-1}R_{2,i-1} + \alpha R_{3i} - A_{I_{wi}} - (\iota + c_i + \mu_i)R_{2i} \quad 2 \le i \le n \\
&\frac{dR_{31}}{dt} = \alpha R_{41} - (\lambda_1 + \iota + c_1 + \mu_1)R_{31} \\
&\frac{dR_{3i}}{dt} = c_{i-1}R_{3,i-1} + \alpha R_{4i} - (\lambda_i + \iota + c_i + \mu_i)R_{3i} \quad 2 \le i \le n \\
&\frac{dR_{41}}{dt} = \gamma(I_1 + I_{m1} + I_{w1}) + \lambda_1 R_{31} - (\alpha + c_1 + \mu_1)R_{41} \\
&\frac{dR_{4i}}{dt} = c_{i-1}R_{4,i-1} + \gamma(I_i + I_{mi} + I_{wi}) + \lambda_i R_{3i} - (\alpha + c_i + \mu_i)R_{4i} \quad 2 \le i \le n \\
&N_i = S_i + I_i + I_{mi} + I_{wi} + R_{1i} + R_{2i} + R_{3i} + R_{4i} \quad 1 \le i \le n \\
&\mu_1 = \frac{\sum_{j=1}^{n} v_j N_j - c_1 N_1}{N_1} \\
&\mu_i = \frac{c_{i-1}N_{i-1} - c_i N_i}{N_i} \quad 2 \le i \le n \\
&\lambda_i = f_{p_i} \sum_{j=1}^{n} \frac{l_j l_i}{D} \frac{I_j + \rho_m I_{mj} + \rho_w I_{wj}}{\sum_{j=1}^{n} N_j} \quad 1 \le j \le n \\
&D = \sum_{k=1}^{n} l_k N_k / \sum_{k=1}^{n} N_k \\
&d(ln\frac{l_1}{\sqrt{D}}) = s_{lD_1} dW_t \\
&\frac{l_i}{\sqrt{D}} = \frac{l_1}{\sqrt{D}} * f_{l_i} \quad 2 \le i \le n \\
&d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t \\
&I_{ki} = \int_{k-1}^{k} (A_{I_i} + \rho_m A_{I_{mi}} + \rho_w A_{I_{wi}})\, dt \quad 1 \le i \le n \\
&I_{rki} = I_{ki} C_r \quad 1 \le i \le n \\
&A_{I_i} = \frac{Poisson(\lambda_i S_i \Delta t)}{\Delta t} \quad 1 \le i \le n \\
&A_{I_{mi}} = \frac{Poisson(\lambda_i R_{1i} \Delta t)}{\Delta t} \quad 1 \le i \le n \\
&A_{I_{wi}} = \frac{Poisson(\lambda_i R_{2i} \Delta t)}{\Delta t} \quad 1 \le i \le n
\end{aligned}
\tag{D.3}
$$

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 D.6.

Appendix D.4. The age-structured model of 32 age groups ($n = 32$) with re-balanced contact matrix

The age-structured model of 32 age groups with re-balanced contact matrix 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. The mathematical equations of state space model are listed as follows:

Table D.6: **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, 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, 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, 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
$l_1 p_1$	Uniform[0.001, 0.5)	Person/Day
$l_1/\sqrt{D}$	Uniform[0.005, 0.2)	$\sqrt{Person/Day}$
C_r	Uniform[0, 0.15)	Dimensionless
$\epsilon_i \quad 1 \leq i \leq 6$	Uniform[0, 1)	Dimensionless

$$
\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 \\
&\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 + \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 \le i \le n \\
&\frac{dR_{41}}{dt} = \gamma (I_1 + I_{m1} + I_{w1}) + \lambda 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 R_{3i} - (\alpha + c_i + \mu_i) R_{4i} \quad 2 \le i \le n \qquad \text{(D.4)} \\
&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 \\
&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 \le i \le n, \quad 1 \le j \le 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 \le i \le n \\
&d[ln(l_1 p_1)] = s_{l_1} dW_t \\
&l_i p_i = l_1 p_1 * f_{l_i} * f_{p_i} \quad 2 \le i \le n \\
&d(logit(\epsilon_i)) = d(ln(\frac{\epsilon_i}{1 - \epsilon_i})) = s_{\epsilon_i} dW_t \quad 1 \le i \le 6 \\
&d(logit(C_r)) = d(ln(\frac{C_r}{1 - C_r})) = s_r dW_t
\end{aligned}
$$

$$I_{ki} = \int_{k-1}^{k} (A_{I_i} + \rho_m A_{I_{m\,i}} + \rho_w A_{I_{w\,i}})dt \quad 1 \le i \le n$$
$$I_{rki} = I_{ki} C_r \quad 1 \le i \le n$$
$$A_{I_i} = \frac{Poisson(\lambda_i S_i \Delta t)}{\Delta t} \quad 1 \le i \le n$$
$$A_{I_{m\,i}} = \frac{Poisson(\lambda_i R_{1i} \Delta t)}{\Delta t} \quad 1 \le i \le n$$
$$A_{I_{w\,i}} = \frac{Poisson(\lambda_i R_{2i} \Delta t)}{\Delta t} \quad 1 \le i \le n$$

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 listed in Table D.6.

Appendix E. Further introduction of split the pertussis yearly reported cases to each age group

The yearly empirical data related to multiple age categories are available from year 1925 to 1956 [20]. During the process in preparing the yearly empirical data for the two-age-group and 32-age-group particle filtering models (the yearly empirical data divided into 6 groups), we need to split the data in some age categories in the original datasets [20] due to two reasons. The first reason is because the division of the age group in empirical dataset does not match the division of the age groups in the pertussis particle filtering models. Specifically, from year 1926 to 1941, we need to split the reported pertussis cases in age category "1-6 years" in age 5 proportionally (four fifths goes to the " 1-4 years" age group, and one fifth goes to "5-9 years" age group); from year 1942 to 1955, we need to split the reported pertussis cases in age category "5-14 years" in age 10 proportionally (half goes to the "5-9 years" age group, and half goes to "10-14 years" age group). The second reason is because there is a category in the empirical yearly dataset of "age not stated". Thus, we need to split the counts in this category to corresponding age groups in the particle filtering models proportionally (based on the proportion calculated by the age categories has labeled age clearly).

References

[1] W. H. Organization, Pertussis, vaccines and diseases, immunization, vaccines and biologicals., `http://www.who.int/immunization/diseases/pertussis/en/`, 2018.

[2] C. for Disease Control, Prevention, Causes and transmission., `https://www.cdc.gov/pertussis/about/causes-transmission.html`, 2017.

[3] C. for Disease Control, Prevention, Epidemiology and prevention of vaccine-preventable diseases., `https://www.cdc.gov/vaccines/pubs/pinkbook/pert.html`, 2019.

[4] M. Y. K. Chow, G. Khandaker, P. McIntyre, Global childhood deaths from pertussis: A historical review, Clin. Infect. Dis. 63 (2016) S134–S141.

[5] C. for Disease Control, Prevention, Whooping cough and the vaccine (shot) to prevent it., `https://www.cdc.gov/vaccines/parents/diseases/child/pertussis.html`, 2019.

[6] A. Doroshenko, W. Qian, N. D. Osgood, Evaluation of outbreak response immunization in the control of pertussis using agent-based modeling, PeerJ 4 (2016) e2337.

[7] H. Hethcote, The mathematics of infectious diseases, SIAM Review 42 (2000) 599–653.

[8] X. Li, A. Doroshenko, N. D. Osgood, Applying particle filtering in both aggregated and age-structured population compartmental models of pre-vaccination measles, PLoS ONE 13 (2018).

[9] O. N. Safarishahrbijari A, Social media surveillance improves outbreak projection via transmission models, JMIR Public Health and Surveillance (2019).

[10] N. Osgood, J. 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, IEEE Press, Piscataway, NJ, USA, 2014, pp. 829–841.

[11] V. Dukic, H. F. Lopes, N. G. Polson, Tracking epidemics with google flu trends data and a State-Space SEIR model, Journal of the American Statistical Association 107 (2012) 1410–1426.

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

[13] K. Kreuger, N. Osgood, Particle filtering using agent-based transmission models, in: 2015 Winter Simulation Conference (WSC), 2015, pp. 737–747.

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

[15] J. B. S. Ong, M. I.-C. Chen, A. R. Cook, H. C. Lee, V. J. Lee, R. T. P. Lin, P. A. Tambyah, L. G. 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 (2010) e10036.

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

[17] M. Jong, O. Diekmann, H. Heesterbeek, How does transmission of infection depend on population size? epidemic models, Publication of the Newton Institute (1995) 84–94.

[18] X. Li, Master's thesis (2019).

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

[20] Annual Report of Department of Public Health in the Province of Saskatchewan, Department of Public Health, Saskatchewan, CA, 1921-1956.

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

[22] M. S. Arulampalam, S. Maskell, N. Gordon, 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 (2002) 174–188.

[23] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.

[24] A. Doucet, N. de Freitas, N. Gordon, An introduction to sequential monte carlo methods, in: A. Doucet, N. de Freitas, N. Gordon (Eds.), Sequential Monte Carlo Methods in Practice, Springer New York, New York, NY, 2001, pp. 3–14.

[25] T. P. Hubbard, Y. Saglam, Stochastic processes, itô calculus, and applications in economics, Lecture notes, Department of Mathematics, University of Iowa (2006).

[26] J. Dureau, K. Kalogeropoulos, M. Baguelin, Capturing the time-varying drivers of an epidemic using stochastic dynamical systems, Biostatistics 14 (2013) 541–555.

[27] I. Dorigatti, S. Cauchemez, A. Pugliese, N. M. 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 (2012) 9–21.

[28] A. Blake, M. Isard, The condensation algorithm-conditional density propagation and applications to visual tracking, in: Advances in Neural Information Processing Systems, 1997, pp. 361–367.

[29] T. W. Tulu, B. Tian, Z. Wu, Modeling the effect of quarantine and vaccination on ebola disease, Adv. Difference Equ. 2017 (2017) 178.

[30] H. J. Wearing, P. Rohani, Estimating the duration of pertussis immunity using epidemiological signatures, PLoS pathogens 5 (2009) e1000647.

[31] F. Magpantay, M. D. De CellÈs, P. Rohani, A. King, Pertussis immunity and epidemiology: mode and duration of vaccine-induced immunity, Parasitology 143 (2016) 835–849.

[32] M. D. de Cellès, F. M. Magpantay, A. A. King, P. Rohani, The impact of past vaccination coverage and immunity on pertussis resurgence, Science translational medicine 10 (2018) eaaj1748.

[33] A. A. of Pediatrics, et al., Dtap effectiveness: Not perfect, but still good, AAP Grand Rounds 42 (2019) 18–18.

[34] X. Li, B. Keeler, R. Zahan, L. Duan, A. Safarishahrbijari, J. Goertzen, Y. Tian, J. Liu, N. Osgood, Illuminating the hidden elements and future evolution of opioid abuse using dynamic modeling, big data and particle markov chain monte carlo (2018).