State Estimation for Non-linear State-space Transmission Models of Tuberculosis

Duayne Strydom[1], Johan Derik le Roux[1], and Ian Keith Craig [*1]

[1]Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, South Africa.

January 24, 2022

ABSTRACT

Given the high prevalence of Tuberculosis (TB) and the mortality rate associated with the disease, numerous models, such as the Gammaitoni and Nucci (GN) model, were developed to model the risk of transmission. These models typically rely on a quanta generation rate as a measurement of infectivity. Since the quanta generation rate cannot be measured directly, the unique contribution of this work is to develop state estimators to estimate the quanta generation rate from available measurements. To estimate the quanta generation rate, the GN model is adapted into an augmented single-room GN model and a simplified two-room GN model. Both models are shown to be observable, i.e. it is theoretically possible to estimate the quanta generation rate given available measurements. Kalman filters are used to estimate the quanta generation rate. First, a continuous-time extended Kalman filter (CEKF) is used for both adapted models using a simulation and measurement sampling rate of 60 s. Accurate quanta generate rate estimates are achieved in both cases. A more realistic scenario is also considered with a measurement sampling rate of 1 day. For these estimates, a hybrid extended Kalman filter (HEKF) is used. Accurate quanta generation rate estimates are achieved for the more realistic scenario. Future work could potentially use the HEKFs, the adapted models, and real-time measurements in a control system feedback loop to reduce the transmission of TB in confined spaces such as hospitals.

Index terms— extended Kalman filter, hybrid extended Kalman filter, modelling, non-linear observability, state and parameter estimation, Tuberculosis quanta estimation.

1 INTRODUCTION

Tuberculosis (TB) is a bacterial disease caused by *Mycobacterium tuberculosis* (*Mtb*) (World Health Organization, 1999). In most cases the disease infects the lungs (pulmonary TB). TB is most commonly spread when droplet nuclei, containing *Mtb* bacilli, are expelled from persons with active pulmonary TB and inhaled by persons that are uninfected. Latently infected persons are infected with the disease but are not symptomatic. Approximately 5-10% of these persons will become actively infected, with an estimated 10 million world-wide cases of TB in 2018 (World Health Organization, 1999, 2015, 2019; Küsel et al., 2019).

Given the high prevalence of TB and the mortality rate associated with the disease, numerous models were developed over the years to model risk of transmission, especially for confined spaces. The Wells-Riley (WR) model is commonly used to model such risk (Wells et al., 1955; Nardell, 2016; Yates et al., 2016). Another popular model is the Gammaitoni and Nucci (GN) model (Gammaitoni & Nucci, 1997). It was shown that the WR and GN models are fundamentally the same, but that unlike the WR model, the GN model allows for the use of nonsteady-state conditions of airborne infectious particles (Beggs et al., 2003). Additionally, the GN model is in state-space format, which makes it suitable for the design of a state estimator that can estimate unknown states (Simon, 2006).

An important parameter in TB transmission models is the quanta generation rate. A quantum is defined as the number of infectious droplet nuclei required to infect 63.2% of susceptible individuals (Wells et al., 1955). It quantifies the infectiousness of the airborne agent. In other words, the infectiousness

[*]Address correspondence to Ian Craig, University of Pretoria, South Africa; tel: +27 12 420 2172; ian.craig@up.ac.za.

of the disease and the number of the infectious agent in the room can be expressed by the number of quanta (Beggs et al., 2003). The difference of each individual's immunological response to the disease and the characteristics of the pathogen make it impossible to measure quanta directly. Although quanta is a theoretical unit of measure, it allows one to mathematically compare different scenarios of risk of transmission and how control mechanisms may affect the risk of transmission (Li et al., 2015; Taylor et al., 2016).

There is considerable variability in infectivity between patients with respiratory diseases (Nardell, 2016). The quanta generation rate is defined as the rate at which quanta is produced by infective people (Gammaitoni & Nucci, 1997). Quanta generation rate estimates range from 1 - 10 $quanta \cdot h^{-1}$ for rhinovirus and 15-128 $quanta \cdot h^{-1}$ for influenza (Rudnick & Milton, 2003). In the case of TB, one study indicate 1 - 50 $quanta \cdot h^{-1}$ as a suitable range for the quanta generation rate (Noakes et al., 2006), whereas others indicate a range of 1.25 - 60 $quanta \cdot h^{-1}$ (Noakes & Sleigh, 2009; Nardell et al., 1991). In the case of an intubation-related outbreak, the estimate can go as high as 30840 $quanta \cdot h^{-1}$ (Beggs et al., 2003; Nardell et al., 1991; Gammaitoni & Nucci, 1997). If the quanta generation rate in a TB transmission model can be estimated, it will provide a better indication of the risk of transmission.

The WR model can be manipulated to back-calculate a quanta generation rate. For example, the spread of influenza in a Boeing 737 was represented with both a steady-state and a dynamic model. A quanta generation rate ranging between 15 and 77 $quanta \cdot h^{-1}$ was back-calculated using the steady-state model, whereas a quanta generation rate ranging between 79 and 128 $quanta \cdot h^{-1}$ was back-calculated using the dynamic model from the same data (Rudnick & Milton, 2003). This shows that back-calculating the quanta generation rate can result in the calculated rate being much less than half of its actual value (Nardell et al., 1991; Noakes & Sleigh, 2009).

Because of the variability of quanta generation rate and the ranges of the estimates, the uncertainty of this parameter is large. The problem using backwards calculation is that much of the internal dynamics are lost and only the end result is taken into account. Backwards calculating the quanta generation rate can lead to an incorrect estimate thereof and most often does not account for non-steady-state cases for varying ventilation rates or varying number of infectors.

For experiments conducted at the AIR (Airborne Infections Research) facility in eMalahleni, South Africa, a quanta generation rate of 2.5 $quanta \cdot h^{-1}$ was calculated using a Nelder-Mead search algorithm (Küsel et al., 2019). Although this method is more accurate than back-calculation, it is not suited for a control system feedback loop or for real-time estimates.

The aim of this work is to estimate the quanta generation rate for risk of transmission of TB models as applied to a multi-room environment and to reduce the uncertainty that is present when estimating this parameter. The multi-room environment considered is an AIR facility in eMalahleni, South Africa (Mphaphlele et al., 2015; Strydom et al., 2017; Küsel et al., 2019). The standard GN model only works if the infectious individuals and susceptible individuals are in the same room (Küsel et al., 2019). Therefore, a modified two-room GN model is used in this work where the infectious individuals and susceptible individuals are separated and placed in adjacent rooms. This two-room GN model is used to simulate the spread of the disease in the AIR facility (Mphaphlele et al., 2015; Küsel et al., 2019). The two-room GN model also incorporates an incubation period which is not accounted for in the standard GN model (Küsel et al., 2019).

Simulation data are used to estimate the quanta parameters for risk of transmission of TB models as applied to a single and multi-room environment from published papers of experiments conducted at the AIR facility in eMalahleni, South Africa (Mphaphlele et al., 2015; Strydom et al., 2017; Küsel et al., 2019). Sentinel guinea pigs are used to measure the risk of transmission of TB. The AIR facility has three 2-bed inpatient wards that are connected by airtight ventilation systems to two identical rooms containing sentinel guinea pigs (Mphaphlele et al., 2015; Küsel et al., 2019). The air is vented from the wards into the animal rooms. The air in the wards is assumed to be well mixed with the help of paddle fans to circulate the air in the room. A basic layout of the facility is shown in Fig. 1.

Different risk of transmission models from literature are presented in Section 2 together with model modifications that allow for improved parameter estimation. Observability analyses are done in Section 3 to determine whether the quanta parameter can be theoretically estimated. Section 4.1 and Section 4.2 show the development of continuous-time and hybrid non-linear state estimators. Section 5 shows the simulation results obtained. Sections 6 and 7 discuss the results and conclude the findings respectively.

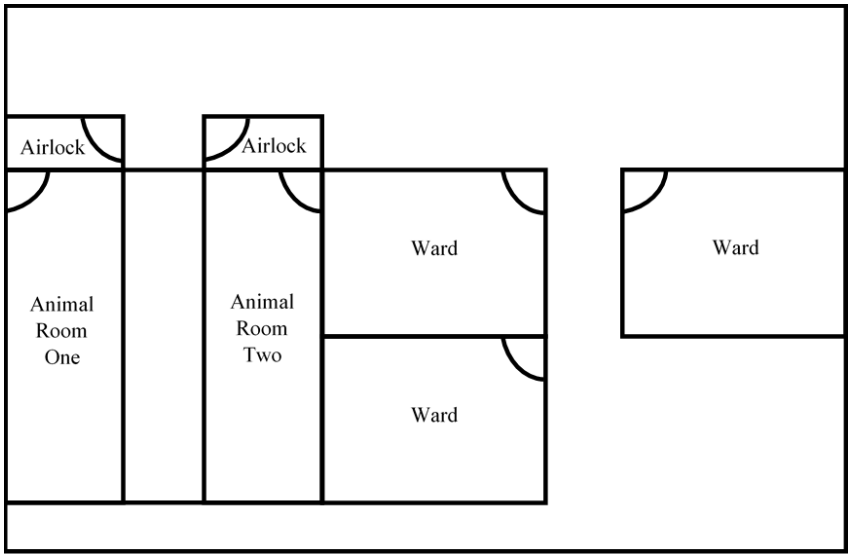

Figure 1: Basic layout of AIR facility in eMalahleni, South Africa (Küsel et al., 2019).

2 MODEL DESCRIPTIONS

Two distinct risk of transmission model scenarios are presented: a single-room and a two-room case. In each case a standard and a modified model is presented. Table 1 summarises the states and parameters of the single-room models in Section 2.1. Table 2 summarises the states and parameters of the two-room models in Section 2.2.

2.1 Single-room GN models

2.1.1 Standard single-room GN model

For the GN model (obtained from literature), it is assumed that the susceptible guinea pigs and the infected individuals (patients) are in the same room such that the room volumes are combined into a single space. It is assumed the guinea pigs do not become infectious as their contribution to new quanta in the room is negligible compared to the contribution of quanta by infected individuals.

The GN transmission model is shown in (1) (Gammaitoni & Nucci, 1997),

$$\begin{aligned} \dot{S} &= -\left(\tfrac{p}{V}\right) CS \\ \dot{C} &= \phi I_w - \tfrac{F}{V} C, \end{aligned} \tag{1}$$

where S is the number of susceptible animals and C is the number of quanta in the room. As per the model definition in Gammaitoni & Nucci (1997), variables S and C are considered as continuous and real. The pulmonary ventilation rate p, the room volume V, and the quanta generation rate per infectious individual ϕ, are constants. The number of infectious individuals I_w in a room, i.e., the number of sick patients, is known a-priori and is regarded as a time-varying input to the model. The flow rate of air F is measured by means of a SCADA system.

The number of infected animals is measured through the diagnosis of TB. Therefore, the number of susceptible animals S is equal to the difference between the initial number of susceptible animals in the room and the number of infected animals. Tuberculin skin tests (TST) are used at the AIR facility to measure whether an animal is infected (Mphaphlele et al., 2015).

Rewriting the model in (1) in state-space format with $S = x_1$ and $C = x_2$ gives,

$$\begin{aligned} \dot{x_1} &= -\beta x_1 x_2 + w_{x_1} \\ \dot{x_2} &= \phi I_w - \tfrac{F}{V} x_2 + w_{x_2}, \end{aligned} \tag{2}$$

where $\beta = p/V$ is the pulmonary ventilation rate over room volume, ϕ is the quanta generation rate, and w_{x_i} is additive zero-mean Gaussian process noise. The state x_1 is assumed measured as,

$$y = h(x) = x_1 + v_{x_1}, \tag{3}$$

where v_{x_1} is additive zero-mean Gaussian measurement noise.

Table 1: Single-room GN model parameters.

Parameter	Unit	Description
S or x_1	$animals$	number of susceptible animals
C or x_2	$quanta$	number of quanta in the room
p	$m^3 \cdot d^{-1}$	pulmonary ventilation rate
V	m^3	room volume
ϕ or x_3	$quanta \cdot d^{-1} \cdot ind.^{-1}$	quanta generation rate per infectious individual
F	$m^3 \cdot d^{-1}$	ventilation flow rate
I_w	$individuals$	number of infectious individuals

Table 2: Two-room GN model parameters.

Parameter	Unit	Description
x_1	$animals$	number of susceptible animals
x_2	$animals$	number of exposed animals
x_3	$animals$	number of infected animals
x_4	$quanta$	number of quanta in ward
x_5	$quanta$	number of quanta in animal room
p	$m^3 \cdot d^{-1}$	pulmonary ventilation rate
V_w	m^3	ward volume
V_1	m^3	animal room volume
ϕ_w	$quanta \cdot d^{-1} \cdot ind.^{-1}$	quanta generation rate per infectious individual
F_w	$m^3 \cdot d^{-1}$	ward flow rate
F_{in}	$m^3 \cdot d^{-1}$	animal room intake flow rate
F_{out}	$m^3 \cdot d^{-1}$	animal room outlet flow rate
I_w	$ind.$	number of infectious individuals
α	d^{-1}	incubation period delay rate

2.1.2 Augmented single-room GN model

The model states in (2) cannot be accurately estimated if there is a parameter mismatch between the estimation and simulation models. Therefore, an updated model shown in (4) is suggested that makes use of one additional quanta generation rate state,

$$\begin{aligned} \dot{x_1} &= -\beta x_1 x_2 + w_{x_1} \\ \dot{x_2} &= x_3 I_w - \tfrac{F}{V} x_2 + w_{x_2} \\ \dot{x_3} &= w_{x_3}, \end{aligned} \tag{4}$$

where x_1 and x_2 are as in (2), and x_3 represents the quanta generation rate. The state x_1 is assumed measured as,

$$y = h(x) = x_1 + v_{x_1}. \tag{5}$$

It is important to include the additional state x_3. If x_3 is disregarded and if x_1 is measured at a typical sampling rate of 24 hours (1 day), the number of quanta in the room will reach a steady-state value faster than an estimator can estimate. This can be seen from converting the equation for $\dot{x_2}$ in (2) to the Laplace domain,

$$X_2 = \frac{\frac{\phi I_w V}{F}}{\frac{V}{F} s + 1}. \tag{6}$$

The shortest time constant of this first-order system is $\frac{V}{F} = 513$ seconds (the parameter values are $V = 112.84\ m^3$ as given in Table 3, and the maximum value of $F = 0.22\ m^3 \cdot s^{-1}$ as shown in Fig. 10). The settling time for such a system in response to a step input is $\frac{4V}{F} = 2052$ seconds or just over 30 minutes (Nise, 2011). This is much faster than what an estimator can estimate if x_1 is measured once every 24 hours. The inclusion of the additional state x_3 prevents x_2 form reaching the incorrect steady-state which makes it possible for an estimator to estimate the value of x_2.

The number of quanta in the room x_2, reaches a steady-state value of $x_2 = \frac{\phi I_w V}{F}$, determined by setting $\dot{x_2}$ in (2) equal to zero. Thus, if the generation rate ϕ is increased, so does the quanta x_2 present in the room. Inversely, if the ventilation rate F is increased, the quanta in the room x_2 decreases.

Therefore, if the ventilation rate F and number of infected individuals I_w are measured, the number of quanta in the room x_2 can be determined by estimating the quanta generation rate parameter ϕ.

2.2 Two-room GN models

2.2.1 Standard two-room GN model

For the two-room GN model (which was also obtained from literature) (Küsel et al., 2019), the room in which the susceptible animals are in is called the animal room (animal room one from Fig. 1) and the room in which the infectious individuals are in is called the ward. The wards in Fig. 1 were combined into a single space since the air containing TB bacilli vented from each ward was combined before being vented to the animal rooms. In this model the individuals who were exposed to the disease and are not yet infected are labelled exposed individuals. The two-room GN model is described as,

$$\begin{aligned}
\dot{x_1} &= -\frac{p}{V_1}x_5x_1 + w_{x_1} \\
\dot{x_2} &= \frac{p}{V_1}x_5x_1 - \alpha x_2 + w_{x_2} \\
\dot{x_3} &= \alpha x_2 + w_{x_3} \\
\dot{x_4} &= \phi_w I_w - \frac{F_w}{V_w}x_4 + w_{x_4} \\
\dot{x_5} &= \frac{F_{in}}{V_w}x_4 - \frac{F_{out}}{V_1}x_5 + w_{x_5},
\end{aligned} \tag{7}$$

where x_1 is the number of susceptible animals in the animal room, p is the pulmonary ventilation rate of the susceptible animals, V_1 is the volume of the animal room, x_5 is the number of quanta in the animal room, x_2 is the number of exposed animals, α is the incubation period of the disease, x_3 is the number of infected animals, x_4 is the number of quanta in the ward, ϕ_w is the quanta generation rate per infectious individual, I_w is the number of infectious individuals in the ward, F_w is the air extraction ventilation rate of the ward, V_w is the volume of the ward, F_{in} is the ventilation rate of the air flowing into the animal room from the ward, and F_{out} is the ventilation rate of the air extracted from the animal room. The measurement is,

$$y = h(x) = \begin{bmatrix} x_1 + v_{x_1} \\ x_3 + v_{x_3} \end{bmatrix}, \tag{8}$$

where v_{x_1} and v_{x_3} are additive zero-mean Gaussian measurement noise.

2.2.2 Simplified GN two-room model

The two-room GN model has the same estimation problem as the single-room GN model. Because the quanta in the ward x_4 tends to a steady-state value, the state will simply return to that steady-state value when estimated. If the two-room GN model is modified by removing the incubation period α and adding the quanta generation rate per infectious individual ϕ_w as an additional state, and it is assumed the number of susceptible animals x_1 is measurable, one can estimate the generation rate ϕ_w and the number of quanta in the two rooms x_4 and x_5. The model is then given as,

$$\begin{aligned}
\dot{x_1} &= -\frac{p}{V_1}x_5x_1 + w_{x_1} \\
\dot{x_4} &= x_6 I_w - \frac{F_w}{V_w}x_4 + w_{x_4} \\
\dot{x_5} &= \frac{F_{in}}{V_w}x_4 - \frac{F_{out}}{V_1}x_5 + w_{x_5} \\
\dot{x_6} &= w_{x_6},
\end{aligned} \tag{9}$$

where $x_6 = \phi_w$. The measurement is now given as,

$$y = h(x) = x_1 + v_{x_1}. \tag{10}$$

It is important to note that the modified models are fundamentally the same. For the single-room GN model, if $\dot{x}_3 = 0$, then the modified model is the same as the unmodified model. The same can be said for the two-room models with the additions that the measurement of the two-room model is no longer the infected animals but the susceptible animals and that the incubation period is not taken into account. These alterations to the models allow for better state estimation as discussed in Section 4.

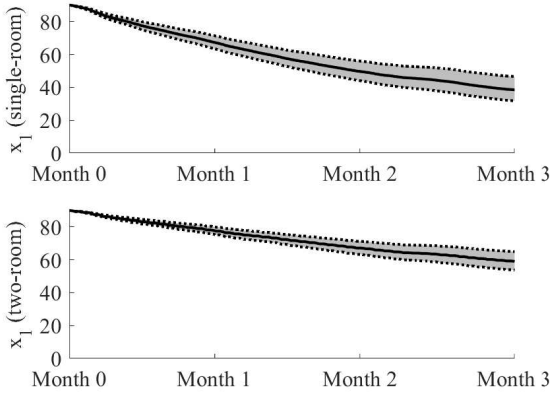

Figure 2: Sensitivity of number of susceptible guinea pigs to deviation of the number of infectious individuals for the augmented single-room and simplified two-room GN models. The deviations at the end of the simulations are 21% and 10% respectively.

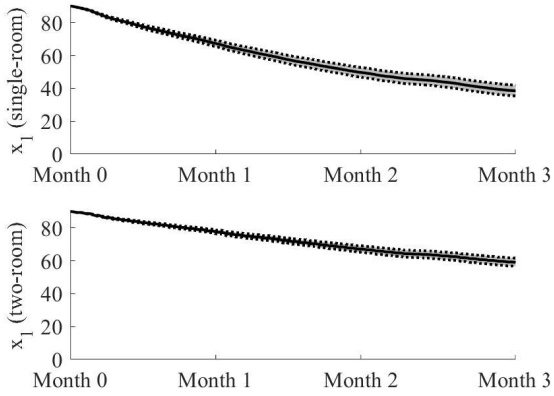

Figure 3: Sensitivity of the number of susceptible guinea pigs to deviation of the quanta generation rates for the augmented single-room and simplified two-room GN models. The deviations at the end of the simulations are 8.9% and 4.3% respectively.

2.3 Sensitivity

A sensitivity analysis was done in literature by deviating the standard single-room model parameters by 10% and the generation rates by 30% for GN and dose-response models. The large generation rates deviation resulted in a large deviation in the predicted number of infected animals (Strydom et al., 2017). The same procedure was followed for the standard two-room model (Küsel et al., 2019).

For the augmented single-room and the augmented two-room models, the simulation parameters were also deviated by 10% from the values in Table 3. The results are shown in Figs. 2 to 8, and the sensitivities for these models are plotted separately.

Fig. 2 shows a deviation of a single patient instead of a 10% deviation, seeing that the number of patients needs to be a positive integer. Fig. 3 shows the effect of a 10% deviation of the quanta generation rates. The sensitivity to a deviation of the pulmonary ventilation rate is shown in Fig. 4. The effect of a 10% deviation of the ward extraction ventilation rate is shown in Fig. 5. A 10% deviation of the ventilation rate into and out of the animal room is shown in Figs. 6 and 7, respectively.

Fig. 8 shows the effect of a larger deviation of 30% in the ranges of the quanta generation rates. The deviation for the augmented single-room GN model resulted in a deviation of up to 29% in the number of susceptible animals and the simplified two-room GN model showed a deviation of up to 13% when the generation rates are deviated by 30%.

3 OBSERVABILITY

In this section, observability analyses are done for each of the models in Section 2 to determine whether the states are observable and can therefore be estimated from a theoretical point of view. Measurement and process noise is disregarded in this section as it does not influence the final result.

3.1 Observability theory

A general non-linear state-space model can be written as,

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x), \end{aligned} \tag{11}$$

with $dim(x) = n$ and $dim(y) = m$. The system in (11) is locally (weakly) observable at x_0 if there exists a neighbourhood X_0 of x_0 such that for every x_1, which is an element of the neighbourhood $X_1 \subset X_0$ of x_0, the indistinguishability of the states x_0 and x_1 implies that $x_0 = x_1$. The two states, x_0 and x_1, are said to be indistinguishable if for every admissible input, u, the output y of (11) for the initial state x_0 and for the initial state x_1 is identical (Hermann & Krener, 1977; Isidori, 1995; Sastry, 1999).

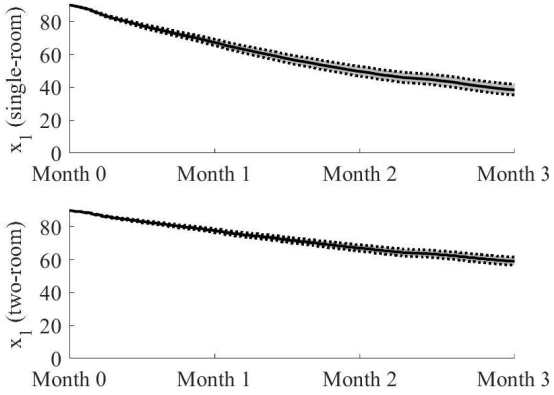

Figure 4: Sensitivity of the number of susceptible guinea pigs to deviation of the pulmonary ventilation rate for the augmented single-room and simplified two-room GN models. The deviations at the end of the simulations are 8.9% and 4.3% respectively.

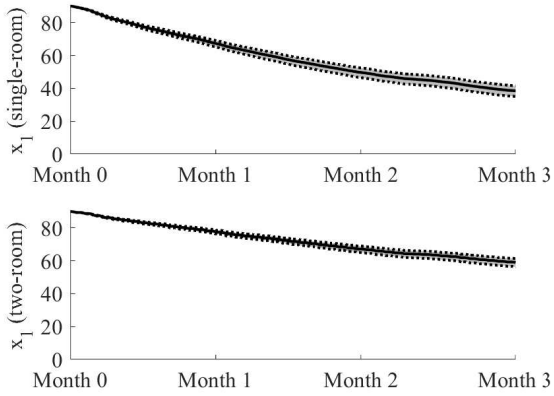

Figure 5: Sensitivity of the number of susceptible guinea pigs to deviation of the ward ventilation rate for the augmented single-room and simplified two-room GN models. The deviations at the end of the simulations are 8.0% and 3.9% respectively.

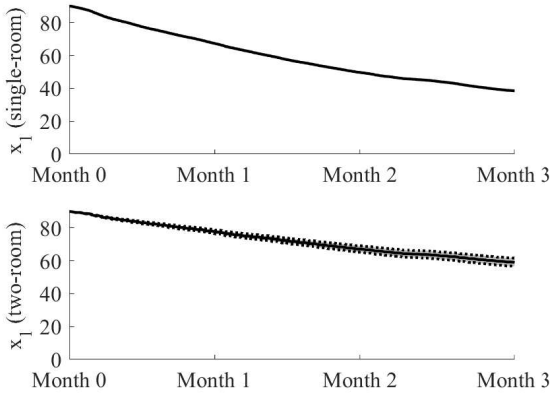

Figure 6: Sensitivity of the number of susceptible guinea pigs to deviation of the ventilation rate into the animal room for the augmented single-room and simplified two-room GN models. The deviation at the end of the simulation is 4.3%.

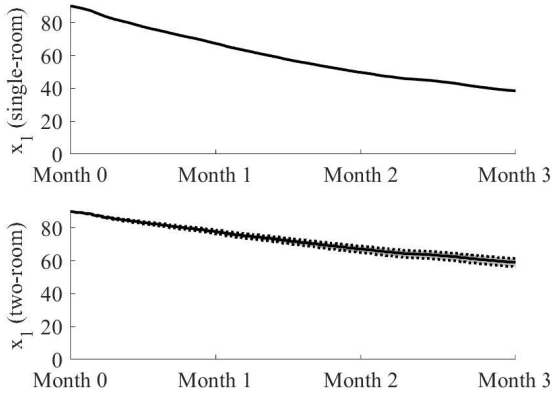

Figure 7: Sensitivity of the number of susceptible guinea pigs to deviation of the ventilation rate out of the animal room for the augmented single-room and simplified two-room GN models. The deviation at the end of the simulation is 3.9%.

To determine if a system is locally observable, the system has to satisfy the so-called observability rank condition, i.e. the observability codistribution $d\mathcal{O}$ must have dimension n at x_0. The observability codistribution $d\mathcal{O}$ of x_0 is given by (Hermann & Krener, 1977),

$$d\mathcal{O} = span\{dh_j, dL_f h_j, ..., dL_f^{n-1} h_j\}; j = 1...m. \tag{12}$$

where $L_f h$ denotes the Lie derivative of h with respect to f,

$$L_f h = \sum_{i=1}^{n} \frac{\partial h}{\partial x_i} f_i. \tag{13}$$

The Lie derivative of dh with respect to f is given as,

$$L_f dh = \left(\frac{\partial dh^T}{\partial x} f \right)^T + dh \frac{\partial f}{\partial x} = dL_f h. \tag{14}$$

Once the state observability is determined and the system is indeed observable, the states of the model can be theoretically estimated using state estimators (Simon, 2006).

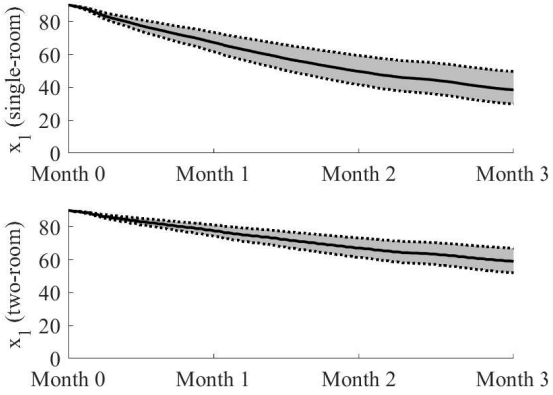

Figure 8: Sensitivity of the number of susceptible guinea pigs to deviation of the range of the quanta generation rate for the augmented single-room and simplified two-room GN models. The deviations at the end of the simulations are 29% and 13% respectively.

3.2 Single-room GN model observability

3.2.1 Observability of standard single-room GN model

The standard single-room GN model in (2) and (3) can be written in the state-space form of (11) as,

$$\begin{array}{lll} \dot{x_1} & = f_1 & = -\beta x_1 x_2 \\ \dot{x_2} & = f_2 & = \phi I_w - \frac{F}{V} x_2 \\ y & = h & = x_1. \end{array} \tag{15}$$

The observability codistribution for (15) is,

$$d\mathcal{O} = \begin{bmatrix} dh \\ dL_f h \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\beta x_2 & -\beta x_1 \end{bmatrix}, \tag{16}$$

and the determinant is $\det(d\mathcal{O}) = -\beta x_1$.

The rank of $d\mathcal{O}$ in (16) is 2 if $\beta x_1 \neq 0$ and is equal to the number of states in (15). Therefore, the standard single-room GN model is observable as long as $\beta x_1 \neq 0$.

3.2.2 Observability of augmented single-room GN model

The augmented single-room GN model given in (4) and (5) can be written in the state-space form of (11) as,

$$\begin{array}{lll} \dot{x_1} & = f_1 & = -\beta x_1 x_2 \\ \dot{x_2} & = f_2 & = x_3 I_w - \frac{F}{V} x_2 \\ \dot{x_3} & = f_3 & = 0 \\ y & = h & = x_1. \end{array} \tag{17}$$

The observability codistribution for (17) is,

$$d\mathcal{O} = \begin{bmatrix} 1 & 0 & 0 \\ -\beta x_2 & -\beta x_1 & 0 \\ \beta^2 x_2^2 - \beta\left(I_w x_3 - \frac{F}{V} x_2\right) & 2x_1 x_2 \beta^2 + \frac{F}{V} x_1 \beta & -\beta I_w x_1 \end{bmatrix}, \tag{18}$$

and the determinant is $\det(d\mathcal{O}) = \beta^2 I_w x_1^2$.

The rank of $d\mathcal{O}$ in (18) is 3 if $\beta^2 I_w x_1^2 \neq 0$ and is equal to the number of states in (17). Therefore, the augmented single-room GN model is observable as long as $\beta^2 I_w x_1^2 \neq 0$.

3.3 Two-room GN model observability

3.3.1 Observability of standard two-room GN model

The standard two-room GN model given in (7) and (8) can be written in the state-space form of (11) as,

$$\begin{array}{lll} \dot{x_1} & = f_1 & = -\frac{p}{V_1}x_5x_1 \\ \dot{x_2} & = f_2 & = \frac{p}{V_1}x_5x_1 - \alpha x_2 \\ \dot{x_3} & = f_3 & = \alpha x_2 \\ \dot{x_4} & = f_4 & = \phi_w I_w - \frac{F_w}{V_w}x_4 \\ \dot{x_5} & = f_5 & = \frac{F_{in}}{V_w}x_4 - \frac{F_{out}}{V_1}x_5, \\ y & = h & = [x_1, x_3]^T . \end{array} \tag{19}$$

Considering only until the second repeated Lie derivative in (12), the observability codistribution matrix for (19) is,

$$d\mathcal{O} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -\frac{p\,x_5}{V_1} & 0 & 0 & 0 & -\frac{p\,x_1}{V_1} \\ 0 & \alpha & 0 & 0 & 0 \\ \frac{p\left(V_w\,p\,{x_5}^2 + F_{out}\,V_w\,x_5 - F_{in}\,V_1\,x_4\right)}{{V_1}^2\,V_w} & 0 & 0 & -\frac{F_{in}\,p\,x_1}{V_1\,V_w} & \frac{p\,x_1\,(F_{out}+2\,p\,x_5)}{{V_1}^2} \\ \frac{\alpha\,p\,x_5}{V_1} & -\alpha^2 & 0 & 0 & \frac{\alpha\,p\,x_1}{V_1} \end{bmatrix}, \tag{20}$$

and the determinant of the full-rank portion (rows 1 to 5 and columns 1 to 5) is $\det(d\mathcal{O}_{1:5;1:5}) = -\frac{F_{in}\alpha p^2 x_1^2}{V_1^2 V_w}$.

The rank of $d\mathcal{O}$ in (20) is 5 if $\frac{F_{in}\alpha p^2 x_1^2}{V_1^2 V_w} \neq 0$ and is equal to the number of states in (19). Therefore, the standard two-room GN model is observable as long as $\frac{F_{in}\alpha p^2 x_1^2}{V_1^2 V_w} \neq 0$.

3.3.2 Observability of simplified two-room GN model

The simplified two-room GN model in (9) and (10) can be written in the state-space form of (11) as,

$$\begin{array}{lll} \dot{x_1} & = f_1 & = -\frac{p}{V_1}x_5x_1 \\ \dot{x_4} & = f_2 & = x_6 I_w - \frac{F_w}{V_w}x_4 \\ \dot{x_5} & = f_3 & = \frac{F_{in}}{V_w}x_4 - \frac{F_{out}}{V_1}x_5 \\ \dot{x_6} & = f_4 & = 0 \\ y & = h & = x_1. \end{array} \tag{21}$$

The observability codistribution for (21) is,

$$d\mathcal{O} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{px_5}{V_1} & 0 & -\frac{px_1}{V_1} & 0 \\ \frac{pA}{V_1} + \frac{p^2{x_5}^2}{{V_1}^2} & -\frac{F_{in}px_1}{V_1V_w} & O_{33} & 0 \\ O_{41} & O_{42} & O_{43} & -\frac{F_{in}I_wpx_1}{V_1V_w} \end{bmatrix}, \tag{22}$$

where,

$$\begin{array}{lll} A & = & \frac{F_{out}x_5}{V_1} - \frac{F_{in}x_4}{V_w} \\ O_{33} & = & \frac{F_{out}px_1}{{V_1}^2} + \frac{2p^2x_1x_5}{{V_1}^2} \\ O_{41} & = & -A\left(\frac{2p^2x_5}{{V_1}^2} + \frac{F_{out}p}{{V_1}^2}\right) - \frac{px_5\left(\frac{pA}{V_1} + \frac{p^2{x_5}^2}{{V_1}^2}\right)}{V_1} - \frac{F_{in}p\left(I_wx_6 - \frac{F_wx_4}{V_w}\right)}{V_1V_w} \\ O_{42} & = & \frac{F_{in}(O_{33})}{V_w} + \frac{F_{in}F_wpx_1}{V_1{V_w}^2} + \frac{F_{in}p^2x_1x_5}{{V_1}^2V_w} \\ O_{43} & = & -\frac{F_{out}(O_{33})}{V_1} - \frac{2p^2x_1A}{{V_1}^2} - \frac{px_1\left(\frac{pA}{V_1} + \frac{p^2{x_5}^2}{{V_1}^2}\right)}{V_1} - \frac{px_1x_5\left(\frac{2p^2x_5}{{V_1}^2} + \frac{F_{out}p}{{V_1}^2}\right)}{V_1} \end{array} . \tag{23}$$

The determinant is $\det(d\mathcal{O}) = \frac{F_{in}^2 I_w p^3 x_1^3}{V_1^3 V_w^2}$.

The rank of $d\mathcal{O}$ in (22) is 4 if $\frac{F_{in}^2 I_w p^3 x_1^3}{V_1^3 V_w^2} \neq 0$ and is equal to the number of states in (21). Therefore, the simplified two-room GN model is observable as long as $\frac{F_{in}^2 I_w p^3 x_1^3}{V_1^3 V_w^2} \neq 0$.

4 STATE ESTIMATORS

A Kalman filter (KF) can be used to accurately estimate states of a process where the number of measurements are limited. The KF can also account for uncertainties in the models and the effects of unmeasured disturbances. The KF generates the maximum likelihood estimates for a linear dynamic system subjected to additive process and measurement noises with a multivariate Gaussian distribution (Kalman, 1960; Simon, 2006; Valappil & Georgakis, 2000).

The most commonly used non-linear state estimator is the extended Kalman filter (EKF) (Bavdekar et al., 2011; Schneider & Georgakis, 2013; Simon, 2006). Implementing the filter requires knowledge about the process states and the process model. These models often contain model-plant-mismatch in both the parameters and the structure of the model (Valappil & Georgakis, 2000; Olivier & Craig, 2015).

An EnKF has been used for parameter estimation for the infection rate and fraction of smear positive cases in India. The study uses a deterministic model of TB that models the risk of infection of TB across India and makes use of least squares estimation and an EnKF to determine the infection rate and the number of new cases (Narula et al., 2016). As opposed to the study in India, this study aims to estimate the generation rate of quanta for models in confined spaces such as hospitals and workplaces.

Continuous-time EKFs and hybrid EKFs are used to estimate the states in the models given in Section 2 from the available measurements.

4.1 Continuous-time EKF

Continuous-time EKFs (CEKF) were designed for the systems in (2), (4), (7) and (9). The system equations for the CEKFs are defined as,

$$\begin{aligned} \dot{x} &= f(x,u,w,t) \\ y &= h(x,v,t) \\ w(t) &\sim \mathcal{N}(0,Q) \quad ; \quad v(t) \sim \mathcal{N}(0,R), \end{aligned} \tag{24}$$

where $f(x,u,w,t)$ is the continuous-time system, $h(x,v,t)$ is the measurement model, $w(t)$ and $v(t)$ is the zero-mean Gaussian process noise and measurement noise respectively, and Q and R is the process noise covariance matrix and the measurement noise covariance matrix respectively. Matrices Q and R are positive definite and constant.

The continuous-time EKF algorithm consists of two steps, the linearization step and the measurement update step (Simon, 2006). The linearization step, which uses Taylor series approximations, is given as,

$$\begin{aligned} C &= \frac{\partial f}{\partial x}|_{\hat{x}} \\ L &= \frac{\partial f}{\partial w}|_{\hat{x}} \\ H &= \frac{\partial h}{\partial x}|_{\hat{x}} \\ M &= \frac{\partial h}{\partial v}|_{\hat{x}}, \end{aligned} \tag{25}$$

and the measurement update step is,

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x},u,w_0,t) + K[y - h(\hat{x},v_0,t)] \\ K &= PH^TR^{-1} \\ \dot{P} &= CP + PC^T + LQL^T - PH^TR^{-1}HP, \end{aligned} \tag{26}$$

where $\hat{x}$ is the estimate of x, K is the KF gain, and P is the estimation-error covariance of $\hat{x}$ (Simon, 2006; Bavdekar et al., 2011). The estimation algorithm can be changed for cases where measurement updates occur at irregular sampling intervals, which itself occurs at multiples of the sampling time to form a hybrid EKF.

The EKF is initialised using,

$$\begin{aligned} \hat{x}_0 &= E[x_0] \\ P_0 &= E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T], \end{aligned} \tag{27}$$

where x_0 is a vector containing the initial states and $\hat{x}_0$ is the initial estimates of x_0.

4.2 Hybrid EKF

Increasing the simulation step time and the measurement time for the models can result in numerical integration inaccuracies when doing simulations (Labuschagne et al., 2013). To circumvent this, a hybrid EKF (HEKF) can be used. The HEKF will allow one to simulate the system and update the filter at different time intervals or when measurements are available.

HEKFs are used for continuous-time systems where discrete-time measurements are available. The HEKF uses continuous-time update equations to evaluate the discrete-time values of the model states, x, and the covariance of the estimation error, P, and updates the state estimate using a discrete-time KF.

The HEKF process is similar to the EKF. However, the update step is split into the continuous-time update equations (where the models are simulated in continuous-time) and the discrete measurement update equations (where the filter estimates are updated).

The continuous-time update equations are,

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, u, \omega_0, t) \\ \dot{P} &= CP + PC^T + LQL^T, \end{aligned} \tag{28}$$

and the discrete measurement update equations are,

$$\begin{aligned} K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + M_k R_k M_k^T)^{-1} \\ x_k^+ &= x_k^- + K_k[y_k - h_k(\hat{x}_k^-, v_0, t_k)] \\ P_k^+ &= (I_D - K_k H_k) P_k^- (I_D - K_k H_k)^T + K_k M_k R_k M_k^T K_k^T, \end{aligned} \tag{29}$$

where k denotes the k-th discrete-time measurement, P_k^- is the covariance of x_k before the measurement update, and P_k^+ is the covariance of x_k after the measurement update. I_D is an identity matrix. x_k^- is the estimate of x before the measurement update is made and x_k^+ is the estimate of x after the measurement update is made.

5 SIMULATIONS AND RESULTS

The simulation setup and results for the continuous-time and hybrid EKFs (CEKF and HEKF) for the modified models are given in this section.

5.1 CEKF for the augmented single-room GN model

5.1.1 Simulation scenario and setup

Considering the models given in (2) and (4), the variable x_1 and parameters p, V, I_w and F can all be measured. The variable x_2 and parameter ϕ (or x_3) can not be measured but can be estimated. Since x_2 is a function of the measured parameters, I_w, V, and F, and the estimated parameter ϕ, one only needs to estimate ϕ (or x_3) to obtain x_2.

Table 3 shows the model parameters taken from literature. The data were obtained from an experiment that was conducted at the AIR (Airborne Infections Research) facility in eMalahleni, South Africa (Küsel et al., 2019) in which sentinel guinea pigs were used to measure the risk of transmission. The model in (4) was simulated with a sampling time of 60 s. Measurement and process noise with distributions as shown in Table 4 was added to (4) and (5) in the simulation at an interval of 60 s.

The infectious individual excitation data (number of infectious individuals in the ward I_w) is shown in Fig. 9. Fig. 10 shows the ventilation rate out of the ward F_w. The ventilation rate into and out of the animal rooms (F_{in} and F_{out}) are shown in Fig. 11 (Küsel et al., 2019).

It was assumed that each infectious individual generated the same number of quanta x_2 and that each of the animals have the same susceptibility. It was also assumed that the infectious individuals I_w remained in the room for the duration of the experiment. The model parameters were time-scaled to the sampling time used in the simulations and the simulation duration was 3 months. The simulation time step size for the CEKF is the same as the measurement time of 60 s.

The measurement and process noise covariance matrices R and Q as in (24) for the CEKF, as well as the initialization of the state and estimation-error covariance for the CEKF are shown in Table 5. The initial number of susceptible animals is taken as $\hat{x}_1 = 81$ *animals* and the initial quanta in the room as $\hat{x}_2 = 17.14$ *quanta*. The quanta generation rate is taken as $\hat{x}_3 = 150$ *quanta*$\cdot d^{-1}$. The initial EKF

Table 3: Initial parameters for the augmented single-room GN model.

Parameter	Values
S_0	90 *animals* (Mphaphlele et al., 2015)
ϕ	60 $quanta \cdot d^{-1}$ (Strydom et al., 2017; Küsel et al., 2019)
p	0.23 $m^3 \cdot d^{-1}$
V	112.84 m^3

Table 4: Augmented single-room GN simulation model process and measurement noise distributions.

Parameter	Distribution
w_{x_1}	$\mathcal{N}\left(0, 0.3312^2\right)$
w_{x_2}	$\mathcal{N}\left(0, 22.603^2\right)$
w_{x_3}	$\mathcal{N}\left(0, 48^2\right)$
v_{x_1}	$\mathcal{N}\left(0, 1.8^2\right)$

Table 5: CEKF and HEKF initialization parameters for augmented single-room GN model.

Parameter		Value
EKF process noise	Q	$\begin{bmatrix} 0.1838 & 0 & 0 \\ 0 & 766.34 & 0 \\ 0 & 0 & 3456 \end{bmatrix}$
EKF measurement noise	R	$\begin{bmatrix} 4.86 \end{bmatrix}$
State condition	$\hat{x}_0$	$[81,\ 17.14,\ 150]^T$
Estimation error-covariance	P_0	$\begin{bmatrix} 81 & 0 & 0 \\ 0 & 311.22 & 0 \\ 0 & 0 & 22500 \end{bmatrix}$

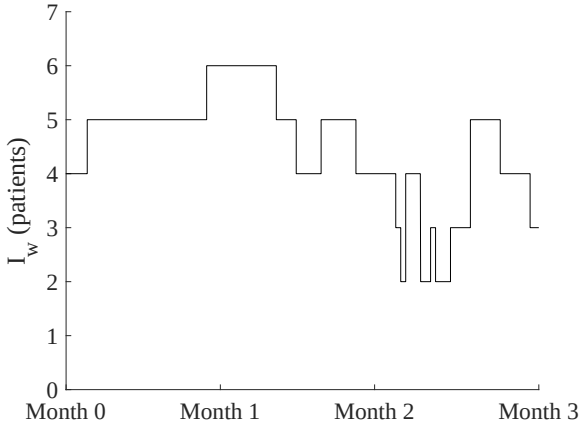

Figure 9: Measured number of infectious individuals in the ward (Küsel et al., 2019).

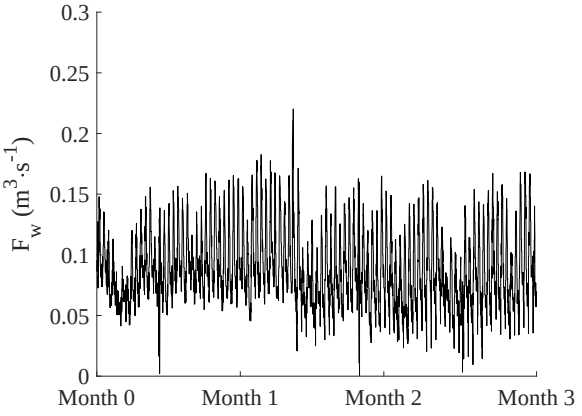

Figure 10: Measured air flow out of the ward, measured in $m^3 \cdot s^{-1}$ (Küsel et al., 2019). The flow rate was measured in $m^3 \cdot s^{-1}$ but time scaled to $m^3 \cdot d^{-1}$ in the simulations.

state and noise estimates in Table 5 are larger than the actual state values in Table 3 and the noise distributions in Table 4 to ensure there is a degree of uncertainty between the simulation model and the EKF. The CEKF measurement update matrix K and the estimation-error covariance matrix P are updated as in (26).

5.1.2 Simulation results

The number of susceptible animals x_1 decays from 90 to 38 animals in 3 months for a quanta generation rate of $x_3 = 60$ $quanta \cdot d^{-1}$ per infectious individual as shown in Fig. 12. The CEKF estimate of the number of susceptibles x_1 starts at 81 and converges to the measured number of susceptibles within 0.1542 days. The number of quanta x_2 in the room is simulated in Fig. 13 and a zoomed version is

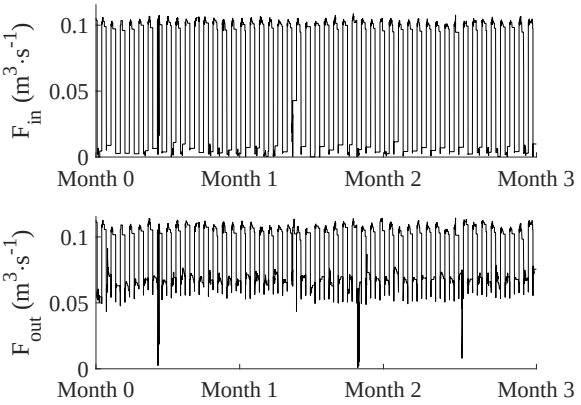

Figure 11: Measured air flow out of and into the animal room, measured in $m^3 \cdot s^{-1}$ (Küsel et al., 2019). The flow rate was measured in $m^3 \cdot s^{-1}$ but time scaled to $m^3 \cdot d^{-1}$ in the simulations.

shown in Fig. 14. The estimated number of quanta converges to the simulated number of quanta after 3.9 days.

As seen in Fig. 15, the estimated quanta generation rate x_3 converges to within 2% of the measured values after just 3.9 days of simulation. This shows that the CEKF can converge to an accurate estimate of the quanta generation rate of the model.

Although it is not shown, the CEKF was also simulated for measurement sampling times of both twice per day and once per day in order to view the effects of different sampling times on the estimates. It is important to note that if the sampling time becomes fairly large, some integration error occurs in the simulation. The number of susceptible animals x_1 reduces to 38 animals in the simulation with a measurement sampling time of 60 s and in the worst case only reduces to 44 animals in the simulation with a measurement sampling time of 1 day. The simulation with a measurement sampling time of 1 day also only converged after 2 months. To improve on this situation, an HEKF is used for simulations where the time between measurements of the number of susceptibles is greater than the time between measurements of the other parameters, as is the case in Section 5.2.

In order to obtain the best sampling rate (measurement time), one can use a fast Fourier transform of the simulated quanta in the room (or ward) to determine the Nyquist frequency. The Nyquist frequency results in a measurement rate of twice per day and the second harmonic gives a measurement rate of once per day. Sampling at the Nyquist frequency is a minimum requirement for preserving the information contained in the quanta generation rate (Lathi, 2010).

The generation rate parameter $x_3 = \phi$ only influences the gain of the state x_2 as seen in (6), and is a time-independent constant. Because quanta state x_2 is dependent on the generation rate x_3 (which is constant) and the ventilation rate F (which is measured at a sampling rate of 60 s, much faster than the Nyquist frequency), the estimated number of quanta in the room x_2 can be updated at a rate slower than the Nyquist frequency, seeing as the only uncertain parameter is a constant and that the simulation time is the same as the ventilation sampling rate. Therefore, sampling at the second harmonic would be sufficient to estimate the quanta generation rate x_3 and thereby the number of quanta x_2. Thus, a measurement sampling rate of once per day, and not twice per day, is used.

The CEKF performs well seeing that the estimated states rapidly converge to the measurements when using a measurement sampling time of 60 s. However, taking measurements every 60 s is unrealistic, but is used here to see what estimates can be obtained under ideal circumstances. A more realistic scenario, with a measurement rate of once per day, is discussed in the next section.

It is important to note that the artificial zero-mean Gaussian noise parameter, w_{x_3}, allows the EKF to better adjust the estimate of the quanta generation rate (Simon, 2006).

5.2 HEKF for augmented single-room GN model

5.2.1 Simulation scenario and setup

Similar to Section 5.1, the model as given by (4) and (5) along with the process and measurement noise distributions as in Table 4 are used here. The HEKF is initialized with the values shown in Table 5, and the simulation input data, Figs. 9 to 11, are the same as in Section 5.1.

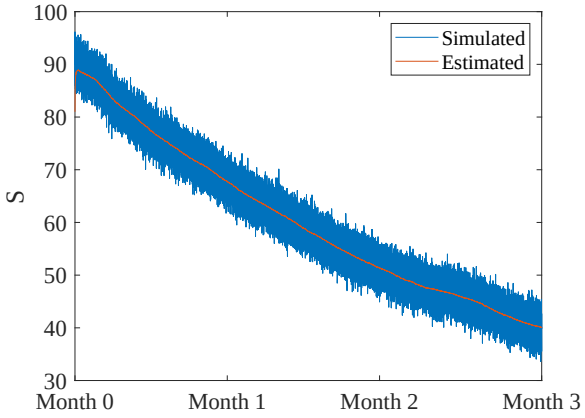

Figure 12: Number of susceptible animals of the augmented single-room GN model CEKF simulation with a sampling time of 60 s. The measured number of susceptible animals is shown in blue and the estimated number of susceptible animals is shown in orange.

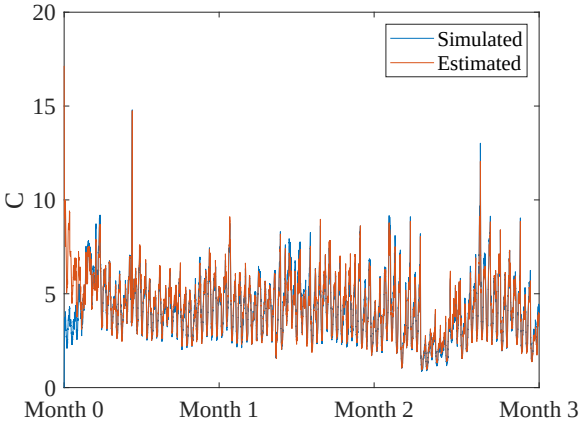

Figure 13: Number of quanta in the room of the augmented single-room GN model CEKF simulation with a sampling time of 60 s. The simulated number of quanta in the room is shown in blue and the estimated number of quanta in the room is shown in orange.

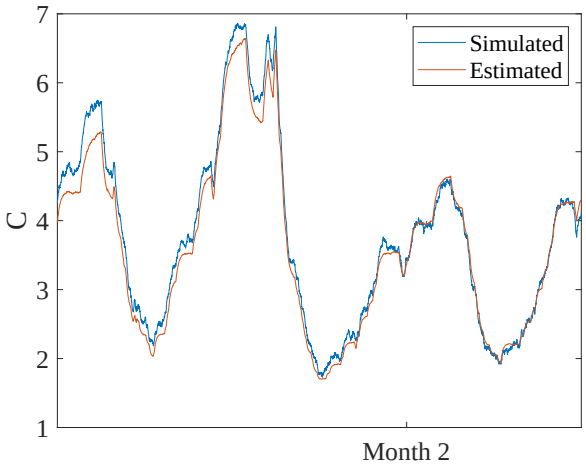

Figure 14: Zoomed number of quanta in the room of the augmented single-room GN model CEKF simulation with a sampling time of 60 s. The blue line indicates the simulated number of quanta in the room and the orange line represents the estimated number of quanta in the room. The period shown is 3 days.

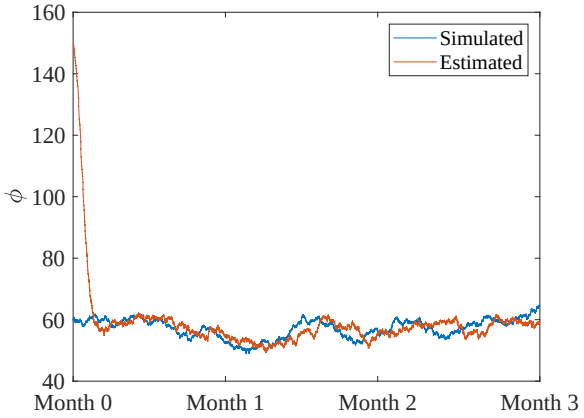

Figure 15: Quanta generation rate of the augmented single-room GN model CEKF simulation with a sampling time of 60 s. The blue line indicates the actual quanta generation rate per infectious individual and the orange line represents the estimated quanta generation rate per infectious individual.

5.2.2 Simulation results

The model in (4) is simulated with a simulation step size of 60 s and the measurement sampling rate for (5) is 1 day because the measurement data are not available at 60 s sampling intervals as assumed in Section 5.1. The resulting simulations are shown in Figs. 16 to 19.

Figs. 16 to 19 show that the HEKF quanta generation rate x_3 estimate converges using the same noise covariance matrices as used in Section 5.1. However, the estimate is not very accurate, but remains within $0.5w_{x_1}$ after 40.031 days. The number of susceptible animals x_1 also reduces to 38 as it should. A large amount of deviation can be seen in the number of quanta in the room due to the discontinuities present in the measured state x_1.

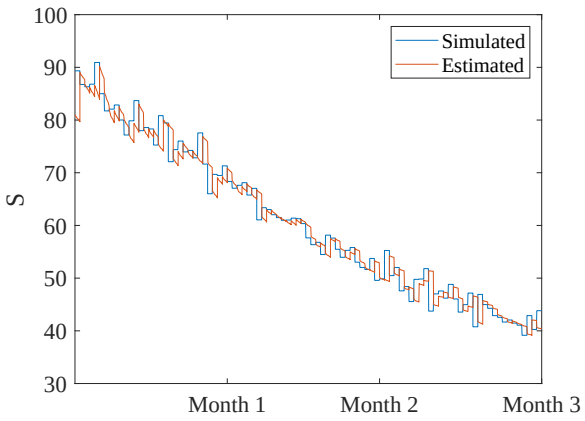

Figure 16: Number of susceptible animals of the augmented single-room GN model HEKF parameter estimation simulation with a sampling time of 1 day and simulation step size of 60 s. The measured number of susceptible animals is shown in blue and the estimated number of susceptible animals is shown in orange.

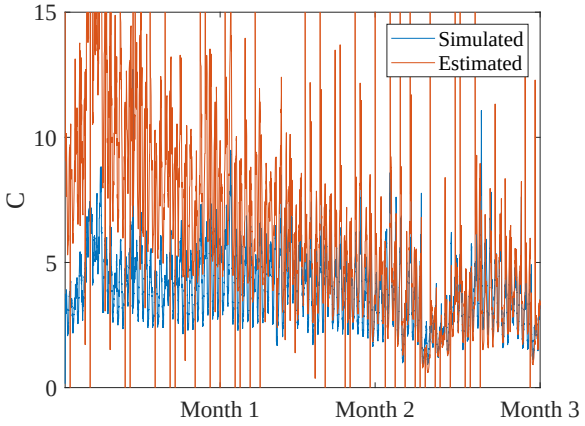

Figure 17: Number of quanta in the room of the augmented single-room GN model HEKF parameter estimation simulation with a measurement time of 1 day and simulation step size of 60 s. The simulated number of quanta in the room is shown in blue and the estimated number of quanta in the room is shown in orange.

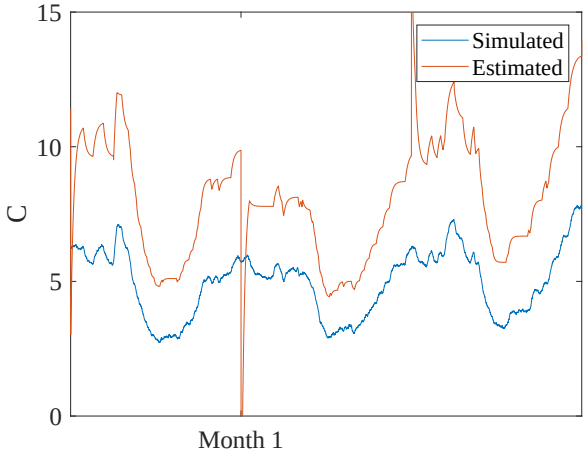

Figure 18: Time-zoomed number of quanta in the room of the augmented single-room GN model HEKF parameter estimation simulation with a measurement time of 1 day and simulation step size of 60 s. The simulated number of quanta in the room is shown in blue and the estimated number of quanta in the room is shown in orange. The period shown is 3 days.

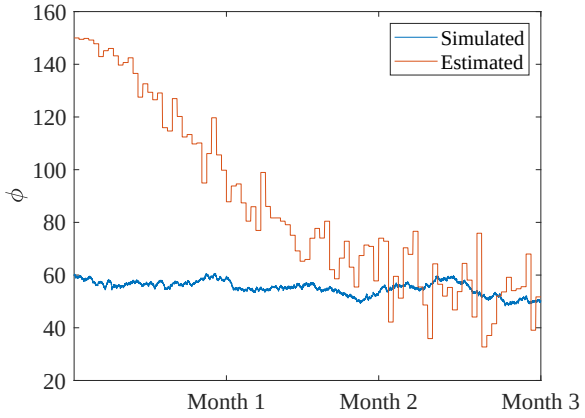

Figure 19: Quanta generation rate per infectious individual of the augmented single-room GN model HEKF parameter estimation with a measurement time of 1 day and simulation step size of 60 s. The simulated quanta generation rate per infectious individual is shown in blue and the estimated quanta generation rate per infectious individual is shown in orange.

5.3 CEKF for the simplified two-room GN model

5.3.1 Simulation scenario and setup

The simplified two-room GN model as given by (9) and (10) is used to simulate the process. The initial parameters are given in Table 6 and the process and measurement noise distributions are shown in Table 7.

The CEKF simulation sampling time and the measurement sampling time is 60 s. The CEKF measurement updates are made using (26). The same simulation input data (Figs. 9 to 11) is used as in Section 5.1. The CEKF is initialized with the data as shown in Table 8.

The initial number of susceptible animals is taken as $\hat{x}_1 = 81$ *animals*, the initial quanta in the ward as $\hat{x}_4 = 17.14$ *quanta*, the initial quanta in the animal room as $\hat{x}_5 = 0.3333$ *quanta* and the initial quanta generation rate is taken as $\hat{x}_6 = 150$ *quanta*$\cdot d^{-1}$.

Table 6: Initial parameters for simplified two-room GN model (Küsel et al., 2019).

Parameter	Values
S_0	90 *animals*
p	0.23 $m^3 \cdot d^{-1}$
V_w	112.84 m^3
V_1	3.4965 m^3

Table 7: Simplified two-room GN simulation model process and measurement noise distributions.

Parameter	Distribution
w_{x_1}	$\mathcal{N}\left(0, 0.6912^2\right)$
w_{x_4}	$\mathcal{N}\left(0, 23.603^2\right)$
w_{x_5}	$\mathcal{N}\left(0, 0.6918^2\right)$
w_{x_6}	$\mathcal{N}\left(0, 48^2\right)$
v_{x_1}	$\mathcal{N}\left(0, 1.8^2\right)$

Table 8: CEKF and HEKF initialization parameters for the simplified two-room GN model.

Parameter		Value
EKF process noise	Q	$\begin{bmatrix} 0.35 & 0 & 0 & 0 \\ 0 & 22.603 & 0 & 0 \\ 0 & 0 & 0.6918 & 0 \\ 0 & 0 & 0 & 48 \end{bmatrix}$
EKF measurement noise	R	$\begin{bmatrix} 4.86 & 0 \\ 0 & 0.003375 \end{bmatrix}$
State condition	$\hat{x}_0$	$[81,\ 17.14,\ 0.3333,\ 150]^T$
Estimation error-covariance	P_0	$\begin{bmatrix} 81 & 0 & 0 & 0 \\ 0 & 302.76 & 0 & 0 \\ 0 & 0 & 0.1111 & 0 \\ 0 & 0 & 0 & 22500 \end{bmatrix}$

5.3.2 Simulation results

Although not explicitly shown in Fig. 20, the estimate of the number of susceptible animals x_1 converges to the correct number of susceptible animals after 16 days. Figs. 21 and 23 show that the estimates of the number of quanta in the ward x_4 and the animal room x_5 converge within 16 days to the simulated number of quanta in the ward and the animal room respectively. Figs. 22 and 24 are time-zoomed figures of x_4 and x_5 respectively. The quanta generation rate x_6 estimate also converges after 16 days as seen in Fig. 25.

5.4 HEKF for simplified two-room GN model

5.4.1 Simulation scenario and setup

Similar to Section 5.3, the model as given by (9) and (10) along with the process and measurement noise distributions as in Table 7 are used here. The HEKF is initialized with the values shown in Table 8, and the simulation input data, Figs. 9 to 11, are the same as in Section 5.1.

5.4.2 Simulation results

The model in (9) is simulated with a simulation step size of 60 s and a measurement time of 1 day, as in Section 5.2. Figs. 26 to 31 show the results.

Fig. 26 shows that the number of susceptible animals x_1 is estimated fairly well. Figs. 27 and 29 show that the filter does not estimate the correct quanta in the ward x_4 and animal room x_5. Time-zoomed figures are shown in Figs. 28 and 30. The quanta generation rate x_6 estimate is shown in Fig. 31, with the estimate converging to within $1w_{x_6}$ after 1 month. A large amount of deviation can be seen in the number of quanta in the rooms due to the discontinuities present in the measured state x_1.

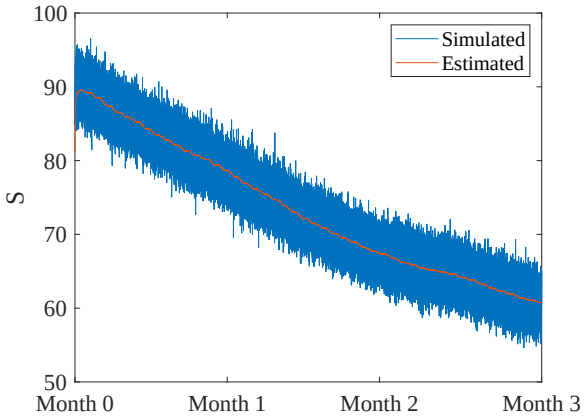

Figure 20: Number of susceptible animals of the simplified two-room GN model estimation with a sampling time of 60 s. The measured number of susceptible animals is shown in blue and the estimated number of susceptible animals is shown in orange.

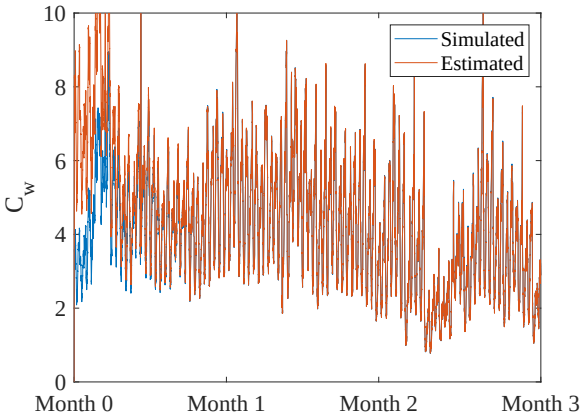

Figure 21: Number of quanta in the ward of the simplified two-room GN model estimation with a sampling time of 60 s. The simulated number of quanta in the ward is shown in blue and the estimated number of quanta in the ward is shown in orange.

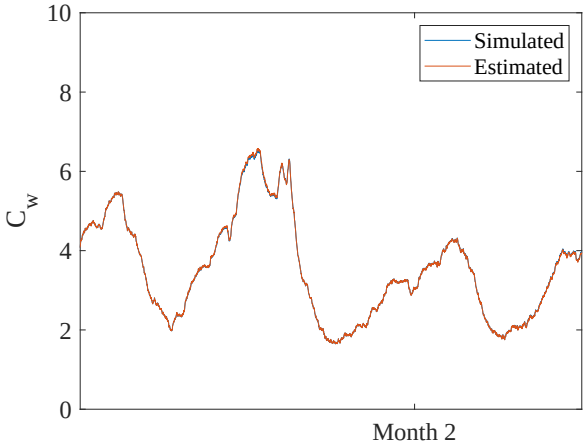

Figure 22: Time-zoomed number of quanta in the ward of the simplified two-room GN model estimation with a sampling time of 60 s. The simulated number of quanta in the ward is shown in blue and the estimated number of quanta in the ward is shown in orange. The period shown is 3 days.

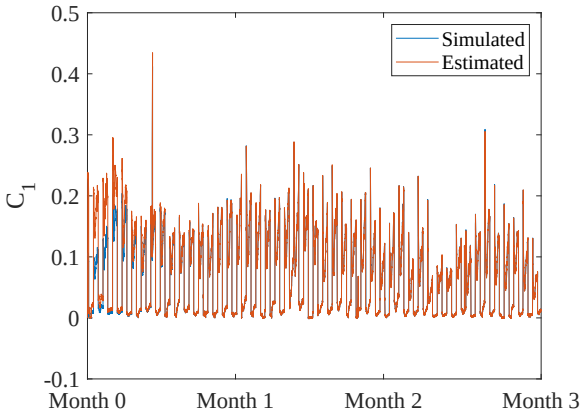

Figure 23: Number of quanta in the animal room of the simplified two-room GN model estimation with a sampling time of 60 s. The simulated number of quanta in the animal room is shown in blue and the estimated number of quanta in the animal room is shown in orange.

5.5 Summary of results

Table 9 shows a summary of the results and a normalised root mean square error (NRMSE) of each of the individual simulations shown above for the number of quanta in the ward and animal room. These are normalised since the two models will have different levels of quanta and are compared directly. The NRMSE is calculated as:

$$NRMSE = \frac{\sqrt{\sum \frac{(y-\hat{y})^T \cdot (y-\hat{y})}{N}}}{max(y)}. \tag{30}$$

The analysis is performed using the data after the estimate has reached steady-state because of the large deviation of the estimate at the start of the simulation. The NRMSE is calculated for the error between the estimated and simulated quanta after 1 month because this is close to the longest settling time of the CEKF models with 10% mismatch, which is 29 days. The NRMSEs of the unmodified single-room and two-room GN models, for which the simulations are not given in this article, are also given in the table. The unmodified single-room and two-room GN model estimates do not converge.

Table 9 indicates that the CEKF outperforms the HEKF for both the unmodified single-room and

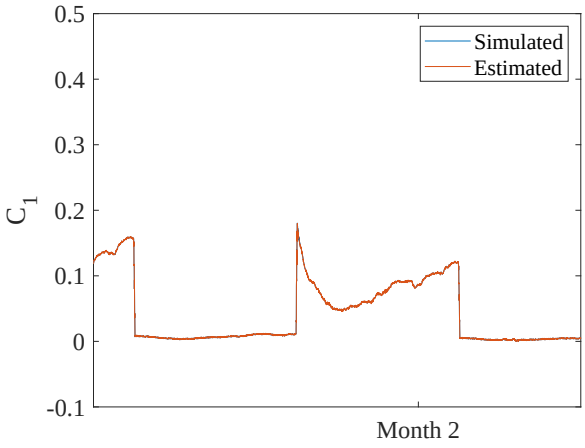

Figure 24: Time-zoomed number of quanta in the animal room of the simplified two-room GN model estimation with a sampling time of 60 s. The simulated number of quanta in the animal room is shown in blue and the estimated number of quanta in the animal room is shown in orange. The period shown is 3 days.

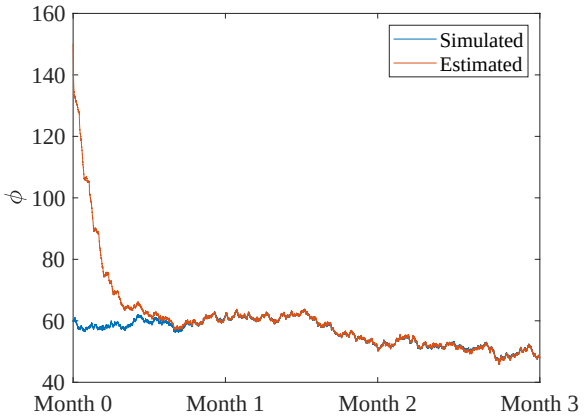

Figure 25: Quanta generation rate of the simplified two-room GN model estimation with a sampling time of 60 s. The simulated quanta generation rate is shown in blue and the estimated quanta generation rate is shown in orange.

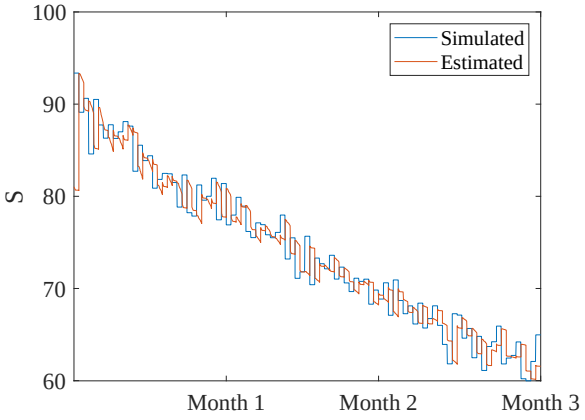

Figure 26: Number of susceptible animals of the simplified two-room GN model HEKF estimation simulation with a measurement time of 1 day and simulation step size of 60 s. The measured number of susceptible animals is shown in blue and the estimated number of susceptible animals is shown in orange.

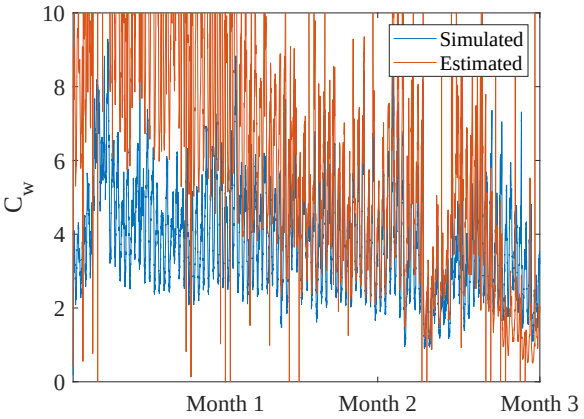

Figure 27: Number of quanta in the ward of the simplified two-room GN model HEKF estimation with a measurement time of 1 day and simulation step size of 60 s. The simulated quanta in the ward is shown in blue and the estimated quanta in the ward is shown in orange.

two-room GN models. However, the HEKF represents a more realistic scenario where measurements are made available once a day. Therefore, the accuracy of the HEKF can be improved if measurements become available at a higher frequency.

6 DISCUSSION

6.1 Single-room GN models

The GN model is in state-space format which allows for simulation with non-zero initial conditions, the use of non-steady-state parameters (such as the ventilation rate), and the design of state estimators. The state-space format also allows one to easily add additional states to the model, such as the quanta generation rate per infectious individual.

The single-room GN model, as opposed to the two-room GN model, has fewer states but the same

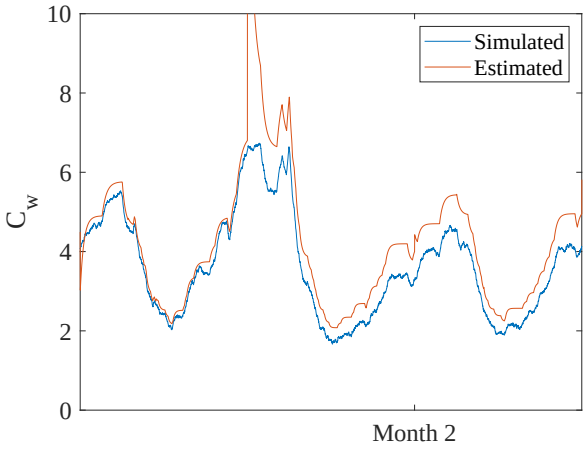

Figure 28: Time-zoomed number of quanta in the ward of the simplified two-room GN model HEKF estimation with a measurement time of 1 day and simulation step size of 60 s. The simulated quanta in the ward is shown in blue and the estimated quanta in the ward is shown in orange.

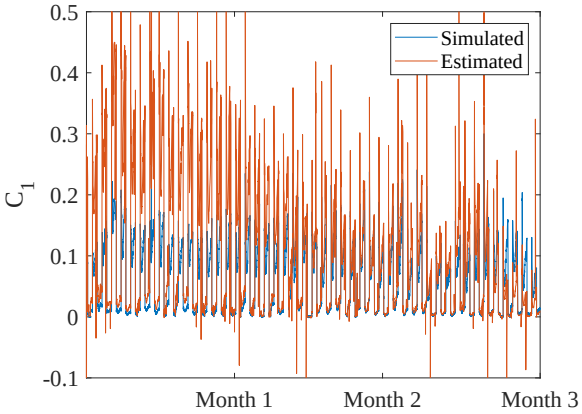

Figure 29: Number of quanta in the animal room of the simplified two-room GN model HEKF estimation simulation with a measurement time of 1 day and simulation step size of 60 s. The simulated number of quanta in the room is shown in blue and the estimated number of quanta in the room is shown in orange.

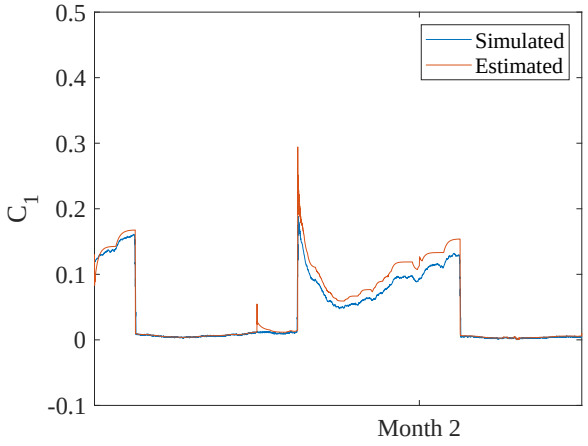

Figure 30: Time-zoomed number of quanta in the animal room of the simplified two-room GN model HEKF estimation simulation with a measurement time of 1 day and simulation step size of 60 s. The simulated number of quanta in the room is shown in blue and the estimated number of quanta in the room is shown in orange.

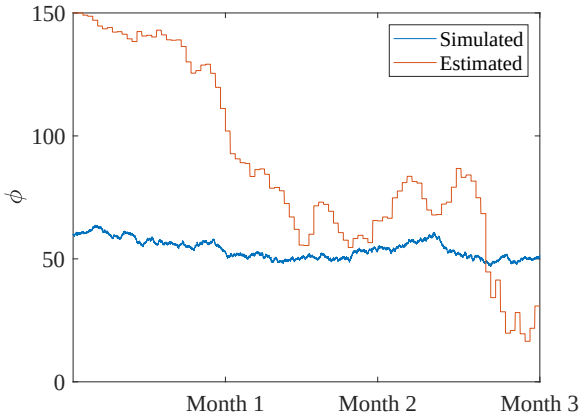

Figure 31: Number of quanta in the ward of the simplified two-room GN model HEKF estimation with a measurement time of 1 day and simulation step size of 60 s. The simulated quanta in the ward is shown in blue and the estimated quanta in the ward is shown in orange.

number of measurements. Fewer states make state estimation easier as it implies fewer parameters that are susceptible to noise and unknown disturbances. The single-room GN model describes the infection of susceptible entities by infectious entities who occupy the same room. This is often the case in hospital settings, but it does not account for the ventilation of infectious droplets to other rooms.

The single-room GN model assumes that the susceptible animals are immediately infected upon exposure to the droplet nuclei and therefore incorporates the incubation period of the disease into the quanta parameter. The other assumptions made are that the air in the room is well mixed, the infectious entities are equally infectious, the susceptible animals are equally susceptible and that they have the same pulmonary ventilation rate. The quanta parameter essentially contains the uncertainty of each of the assumptions made, the measurement noise, and unknown disturbances.

The single-room GN model was augmented with the quanta generation rate as an additional state and EKFs were used to estimate the states. The observability analysis showed that the single-room GN model is observable. Results from the CEKF, for which a measurement time of 60 s is used, are quite

Table 9: Normalised root mean square error of the simulated and estimated number of quanta in the ward. (SR - Single-room; TR - Two-room)

Model	Estimator	Covariance	Initial generation rate estimate	Simulation step size	Measurement time	Time before estimate converges	NRMSE of Quanta in ward after 30 days	NRMSE of Quanta in animal room after 30 days
Standard SR model	CEKF	22500	150	60 s	60 s	-	0.527	-
Augmented SR model	CEKF	22500	150	60 s	60 s	4 d	0.0179	-
Augmented SR model	HEKF	22500	150	60 s	86400 s	40 d	0.194	-
Standard TR model	CEKF	22500	150	60 s	60 s	-	0.524	0.265
Simplified TR model	CEKF	22500	150	60 s	60 s	16 d	0.00125	0.00161
Simplified TR model	HEKF	22500	150	60 s	86400 s	40 d	0.197	0.128

favourable. An NRMSE of the quanta in the ward of 0.01793 was obtained. This is 29.4 times less than when using the standard single-room GN model (for which the quanta in the ward estimate could not converge), also using a measurement time of 60 s. The HEKF, for which a measurement time of 1 day is used, resulted in an NRMSE of 0.1939 which is 2.72 times greater than the standard single-room GN model. The convergence time for the CEKF and HEKF were 3.9 and 40 days, respectively.

Therefore, the quanta generation rate and thereby the number of quanta in the ward can be estimated using the augmented GN model CEKF, but performance is degraded for the HEKF. The CEKF results can be used to better determine the number of quanta present in a ward and are less susceptible to measurement noise and unknown disturbances than when determining the number of quanta and generation rates when using for example methods such as a Nelder-Mead search algorithm.

The estimators for the augmented single-room GN model require measurements of the ventilation rate out of the ward, the pulmonary ventilation rate of the susceptible individuals (measured or estimated), the total room volume, measurements of the number of susceptible individual and the time between each measurement. Most of these measurements only have to be made once or can be obtained from sensors.

6.2 Two-room GN models

The two-room GN model adds the incubation period and the transmission of the disease between two-rooms through ventilation. The model is able to simulate the transmission of the disease as it spreads from one room to another, assuming that the ventilated air is not sterilised and the ventilation rates between the rooms are measured. The model can also be expanded to more than two-rooms, making it possible to model disease transmission throughout a building.

The number of quanta ϕ in the two-rooms and the incubation period α of the disease cannot be measured and the uncertainty of the model lies in these three parameters. The same assumptions are made as for the single-room GN models, with the added assumption that the incubation period of the disease is the same for each individual. The uncertainty that is contained in these assumptions are spread across the number of quanta in the ward and number of quanta in the adjacent rooms.

The simplified two-room GN model removed the incubation period, added the quanta generation rate as a state, and assumed that once an animal has been exposed to the disease it shows signs of infection. However, this assumption causes the estimate of the number of quanta to be less than it actually is because there might be more animals who have not tested positive but have been exposed, meaning the quanta would be greater if those animals are included in the measurements.

For the two-room GN augmented model, a measurement of the change in number of susceptible

animals in addition to the number of susceptible animals is required to make the two-room GN model observable. The CEKF (measurement time 60 s) and HEKF (measurement time 1 day) estimates of the quanta in the ward and animal rooms converge after 16 and 40 days respectively. The NRMSEs of the quanta in the ward and the quanta in the animal rooms for the CEKF are 420.3 and 164 times smaller, respectively, than for the standard two-room GN model. The HEKF NRMSEs for the quanta in the ward and animal rooms are 2.66 and 2.07 times smaller respectively than for the standard two-room GN model.

The number of quanta in the two rooms and the quanta generation rate can therefore be estimated reasonably well. This model requires the same measurements as the single-room GN model in addition to the added ventilation rate and volume of the animal room.

7 CONCLUSION

A reduction in the transmission of TB is essential to reduce new cases and thereby reduce the number of deaths from TB in the future. Towards this end, two TB risk of transmission models in literature were investigated (Gammaitoni & Nucci, 1997; Strydom et al., 2017). Both models rely on a quanta generation rate that cannot be directly measured, as a measurement of infectiousness. If this quanta generation rate can be estimated using nonlinear state estimators, a feedback control system can potentially be used to reduce the transmission of TB by controlling room ventilation rates and, where applicable, also use ultraviolet germicidal irradiation (UVGI) to prevent transmission (Li et al., 2015; Beggs et al., 2000; Beggs & Sleigh, 2002; Li et al., 2007).

In this paper, the standard Gammaitoni and Nucci (GN) model was adapted into an augmented single-room GN model, and a two-room GN model from literature was adapted into a simplified two-room GN model. Both modified models were shown to be observable, which means that it is theoretically possible to estimate the quanta generation rate in these models given the available measurements.

The models were simulated based on parameter data from an AIR facility (Mphaphlele et al., 2015; Küsel et al., 2019) to generate data. Kalman filters were used to estimate the quanta generation for the simulated data. First, a continuous-time extended Kalman filter (CEKF) was used for both adapted models using a simulation sample time and a measurement update time of 60 s. Accurate quanta generation rate estimates were achieved in both cases. Taking measurements every 60 s is however unrealistic, but was used to see what estimates can be obtained under ideal circumstances.

A more realistic scenario, with a measurement rate of once per day, was used next. For these estimates, a hybrid extended Kalman filter (HEKF) was required as it is able to deal with simulations where the time between model output measurements is greater than the time between measurements of other model parameters. The quanta generation rate for both the augmented single-room and simplified two-room GN models was estimated reasonably well.

Real-time measurement of the number of quanta in a room is not possible since the parameter contains uncertainty related to an individual's immune response (Nardell, 2016) and TST measurements can take up to 72 hours to indicate infection (Küsel et al., 2019). The measurement of the concentration of infectious particles in the air is possible using bioaerosol particle sensors (Ghosh et al., 2015; Bhangar et al., 2014; Tobias et al., 2005). Pairing the infectious particle concentration measurement with a modified dose-response model, more accurate estimates of the risk of transmission can be made (Sze To & Chao, 2010; Nardell, 2004).

The measurements would only be able to identify the concentration of bio-aerosols of a specific size, 1-5 μ m (Nardell, 2016; Issarow et al., 2015; Liu et al., 2017; Yang et al., 2007), and not the exact composition thereof (Ghosh et al., 2015). The estimates could therefore be further improved using filters. The effects of control measures could then also be investigated in real-time.

The efficacy of UV control measures and the placement thereof relative to the ventilation in- and outlets should also be taken into account (Sung & Kato, 2011; Noakes et al., 2015). Zonal models can be used to model the interaction between these control measures.

REFERENCES

Bavdekar, V., Deshpande, A., & Patwardhan, S. (2011). Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *Journal of Process Control*, *21*(4), 585–601.

Beggs, C., Kerr, K., Donnelly, J., Sleigh, P., Mara, D., & Cairns, G. (2000). The resurgence of tuberculosis in the tropics. An engineering approach to the control of Mycobacterium tuberculosis and other airborne pathogens: a UK hospital based pilot study. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, *94*(2), 141–146.

Beggs, C., & Sleigh, P. (2002). A quantitative method for evaluating the germicidal effect of upper room UV fields. *Journal of Aerosol Science*, *33*(12), 1681 – 1699.

Beggs, C. B., Noakes, C. J., Sleigh, P. A., Fletcher, L. A., & Siddiqi, K. (2003). The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models. *International Journal of Tuberculosis and Lung Disease*, *7*, 1015–1026.

Bhangar, S., Huffman, J., & Nazaroff, W. (2014). Size-resolved fluorescent biological aerosol particle concentrations and occupant emissions in a university classroom. *Indoor Air*, *24*(6), 604–617.

Gammaitoni, L., & Nucci, M. C. (1997). Using a mathematical model to evaluate the efficacy of TB control measures. *Emerging Infectious Diseases*, *3*, 335–342.

Ghosh, B., Lal, H., & Srivastava, A. (2015). Review of bioaerosols in indoor environment with special reference to sampling, analysis and control mechanisms. *Environment International*, *85*, 254–272.

Hermann, R., & Krener, A. (1977). Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, *22*(5), 728–740.

Isidori, A. (1995). *Nonlinear Control Systems*. Springer, London, 3 ed.

Issarow, C. M., Mulder, N., & Wood, R. (2015). Modelling the risk of airborne infectious disease using exhaled air. *Journal of Theoretical Biology*, *372*, 100 – 106.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45.

Küsel, R. R., Craig, I. K., & Stoltz, A. C. (2019). Modeling the airborne infection risk of tuberculosis for a research facility in eMalahleni, South Africa. *Risk Analysis*, *39*(3), 630–646.

Labuschagne, A., van der Merwe, A., van Rensburg, N., & Zietsman, L. (2013). *An introduction to numerical analysis*. Pretoria: Hatfield Press.

Lathi, B. P. (2010). *Linear systems and signals*. New York: Oxford University Press.

Li, Y., Leung, G. M., Tang, J., Yang, X., Chao, C., Lin, J. Z., Lu, J., Nielsen, P. V., Niu, J., Qian, H., et al. (2007). Role of ventilation in airborne transmission of infectious agents in the built environment-a multidisciplinary systematic review. *Indoor Air*, *17*(1), 2–18.

Li, Y., Tang, J., Noakes, C., & Hodgson, M. J. (2015). Engineering control of respiratory infection and low-energy design of healthcare facilities. *Science and Technology for the Built Environment*, *21*(1), 25–34.

Liu, L., Wei, J., Li, Y., & Ooi, A. (2017). Evaporation and dispersion of respiratory droplets from coughing. *Indoor Air*, *27*(1), 179–190.

Mphaphlele, M., Dharmadhikari, A. S., Jensen, P. A., Rudnick, S. N., Reenen, T. H. V., Pagano, M. A., Leuschner, W., Sears, T. A., Milonova, S. P., Walt, M. V. D., Stoltz, A. C., Weyer, K., & Nardell, E. A. (2015). Institutional tuberculosis transmission: controlled trial of upper room ultraviolet air disinfection: a basis for new dosing guidelines. *American Journal of Respiratory and Critical Care Medicine*, *192*, 477–484.

Nardell, E. (2016). Indoor environmental control of tuberculosis and other airborne infections. *Indoor air*, *26*(1), 79–87.

Nardell, E. A. (2004). Catching droplet nuclei: toward a better understanding of tuberculosis transmission. *American Journal of Respiratory and Critical Care Medicine*, *169*(5), 553–554.

Nardell, E. A., Keegan, J., Cheney, S. A., & Etkind, S. (1991). Airborne infection. Theoretical limits of protection achievable by building ventilation. *American Review of Respiratory Disease*, *144*, 302–306.

Narula, P., Piratla, V., Bansal, A., Azad, S., & Lio, P. (2016). Parameter estimation of tuberculosis transmission model using ensemble Kalman filter across Indian states and union territories. *Infection, Disease & Health*, *21*(4), 184–191.

Nise, N. S. (2011). *Control systems engineering*. Hoboken: John Wiley & Sons, 6th ed.

Noakes, C., Beggs, C., Sleigh, P., & Kerr, K. (2006). Modelling the transmission of airborne infections in enclosed spaces. *Epidemiology & Infection*, *134*(5), 1082–1091.

Noakes, C. J., Khan, M. A. I., & Gilkeson, C. A. (2015). Modeling infection risk and energy use of upper-room ultraviolet germicidal irradiation systems in multi-room environments. *Science and Technology for the Built Environment*, *21*(1), 99–111.

Noakes, C. J., & Sleigh, P. A. (2009). Mathematical models for assessing the role of airflow on the risk of airborne infection in hospital wards. *Journal of the Royal Society Interface*, *6*(suppl_6), S791–S800.

Olivier, L. E., & Craig, I. K. (2015). Development and application of a model-plant mismatch expression for linear time-invariant systems. *Journal of Process Control*, *32*, 77–86.

Rudnick, S., & Milton, D. (2003). Risk of indoor airborne infection transmission estimated from carbon dioxide concentration. *Indoor Air*, *13*(3), 237–245.

Sastry, S. (1999). *Nonlinear Systems: Analysis, Stability, and Control*. Springer, New York, NY.

Schneider, R., & Georgakis, C. (2013). How to not make the extended Kalman filter fail. *Industrial & Engineering Chemistry Research*, *52*(9), 3354–3362.

Simon, D. J. (2006). *Optimal state estimation*. Hoboken: John Wiley & Sons.

Strydom, D., Küsel, R. R., & Craig, I. K. (2017). When is it appropriate to model transmission of tuberculosis using a dose response model? *IFAC-PapersOnLine*, *50*(2), 31–36.

Sung, M., & Kato, S. (2011). Estimating the germicidal effect of upper-room UVGI system on exhaled air of patients based on ventilation efficiency. *Building and Environment*, *46*(11), 2326–2332.

Sze To, G. N., & Chao, C. Y. (2010). Review and comparison between the Wells-Riley and dose-response approaches to risk assessment of infectious respiratory diseases. *Indoor Air*, *20*(1), 2–16.

Taylor, J., Yates, T., Mthethwa, M., Tanser, F., Abubakar, I., & Altamirano, H. (2016). Measuring ventilation and modelling M. tuberculosis transmission in indoor congregate settings, rural KwaZulu-Natal. *The International Journal of Tuberculosis and Lung Disease*, *20*(9), 1155–1161.

Tobias, H. J., Schafer, M. P., Pitesky, M., Fergenson, D. P., Horn, J., Frank, M., & Gard, E. E. (2005). Bioaerosol mass spectrometry for rapid detection of individual airborne Mycobacterium tuberculosis H37Ra particles. *Applied and Environmental Microbiology*, *71*(10), 6086–6095.

Valappil, J., & Georgakis, C. (2000). Systematic estimation of state noise statistics for extended Kalman filters. *AIChE Journal*, *46*(2), 585–601.

Wells, W. F., et al. (1955). Airborne contagion and air hygiene. An ecological study of droplet infections. *Airborne Contagion and Air Hygiene. An Ecological Study of Droplet Infections.*.

World Health Organization (1999). Tuberculosis, www.who.int/topics/tuberculosis/en/, Accessed 28 March 2016.

World Health Organization (2015). Global tuberculosis report 2015, 20th ed. World Health Organization. https://apps.who.int/iris/handle/10665/191102.

World Health Organization (2019). Global tuberculosis report 2019, World Health Organization. https://www.who.int/tb/publications/global_report/en/ .

Yang, S., Lee, G. W., Chen, C., Wu, C., & Yu, K. (2007). The size and concentration of droplets generated by coughing in human subjects. *Journal of Aerosol Medicine*, *20*(4), 484–494.

Yates, T. A., Khan, P. Y., Knight, G. M., Taylor, J. G., McHugh, T. D., Lipman, M., White, R. G., Cohen, T., Cobelens, F. G., Wood, R., et al. (2016). The transmission of Mycobacterium tuberculosis in high burden settings. *The Lancet Infectious Diseases*, *16*(2), 227–238.