Initialization of a Disease Transmission Model [★]

Håkan Runvik [*] **Alexander Medvedev** [*] **Robin Eriksson** [*]
Stefan Engblom [*]

[*] *Information Technology, Uppsala University, Uppsala, SWEDEN,*
e-mail: {`hakan.runvik, alexander.medvedev, robin.eriksson,`
`stefan.engblom`}`@it.uu.se.`

Abstract: Approaches to the estimation of the full state vector of a larger epidemiological model for the spread of Covid-19 in Sweden at the initial time instant from available data and with a simplified dynamical model are proposed and evaluated. The larger epidemiological model is based on a time-continuous Markov chain and captures the demographic composition of and the transport flows between the counties of Sweden. Its intended use is to predict the outbreak development in temporal and spatial coordinates as well as across the demographic groups. It can also support evaluations and comparisions of prospective intervention strategies in terms of, e.g., lockdown in certain areas or isolation of specific age groups. The simplified model is a discrete time-invariant linear system that has cumulative infectious incidence, infected population, asymptomatic population, exposed population, and infectious pressure as the state variables. Since the system matrix of the model depends on a number of transition rates, structural properties of the model are investigated for suitable parameter ranges. It is concluded that the model becomes unobservable for some parameter values. Two contrasting approaches to the initial state estimation are considered. One is a version of Rauch–Tung–Striebel smoother and another is based on solving a batch nonlinear optimization problem. The benefits and shortcomings of the considered estimation techniques are analyzed and compared.

Keywords: Mathematical models, initial states, linear systems, smoothing filters, Markov models, model approximation.

1. INTRODUCTION

This paper makes use of publicly available epidemiological data for estimating suitable initial conditions for a large mechanistic general Susceptible-Exposed-Infectious-Recovered (SEIR) model of the Swedish Covid-19 outbreak. The model incorporates spatial communication between the Swedish municipalities and also includes the Swedish demographics. Following Widgren et al. (2018); Engblom et al. (2020), the viral contraction is driven by an infectious pressure. Fig. 1 provides an overview of the modeling approach and included compartments.

The dynamics of the disease transmission are modeled by a discrete-state continuous-time Markov chain (MC) implemented in SimInf, Widgren et al. (2019). A continuous state variable, the environmental compartment, is included to model the infectious pressure. A Bayesian approach is utilized to infer the model parameters since it allows for the use of empirical measures as prior knowledge.

The problem of estimating the state vector of a dynamical system backwards in time is known as smoothing. A minimal variance fixed-interval smoother for a linear time-invariant (LTI) model under additive Gaussian noise assumption was derived in Rauch et al. (1965), followed up by various generalizations including state-dependent Gaussian noise (Aravkin and Burke (2012)) and non-Gaussian noise sources (Wang et al. (2020)). In the present
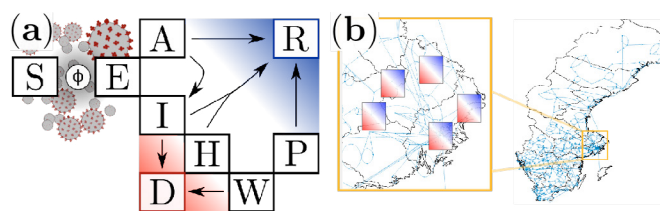

Fig. 1. **(a)** Illustration of the compartment model of the Swedish COVID-19 outbreak. Arrows denote the flows of individuals between the compartments: susceptible (S), exposed (E), asymptomatic (A), symptomatic (I), hospitalized (H), intensive care (W), post-intensive care (P), deceased (D), recovered (R). **(b)** The full model is a network of compartment models emulating the commuting network between the Swedish municipalities. The yellow rectangle shows a zoom-in view of the network within the greater Stockholm region.

work, the latter two circumstances occur combined, as the process noise is Poisson-distributed rather than Gaussian, and also dependent on the plant state. Therefore, none of the approaches found in in the literature is readily applicable here. Instead, empirical initialization algorithms are developed and compared performance-wise on synthetic data produced by models of increasing complexity.

The rest of the paper is organized as follows. First, the model initialization problem is formulated and the properties of the LTI model that is used to calculate the initial condition are explored. Then, three model-based approaches to solving the initialization problem are pre-

sented. Finally, performance of the considered approaches is evaluated on synthetic data and conclusions are drawn.

2. MODEL INITIALIZATION PROBLEM

The MC model inputs are the parameters inferred from data and an initial chain state. The latter consists of the epidemiological states in all compartments, including the hidden ones. The estimation of the infected, exposed, and asymptomatic populations at a given point in time (model initialization point) is therefore investigated, based on the data for cumulative incidence measured over a fixed time horizon. Thus, the problem at hand constitutes a fixed-interval smoothing problem. The remaining MC model compartments do not influence the infected, exposed or asymptomatic populations and are therefore not included at present in the considered estimation problem.

2.1 Initialization data

Epidemiological mathematical models are typically designed in terms of populations and face difficulties in capturing situations when only a few individuals are infected. This is logically the case in the beginning of an outbreak. Besides, an epidemic is not readily recognized until the number of patients in the healthcare system becomes significant, thus making initial data scarce and unreliable. In stochastic epidemiological modeling, simulations are conventionally started when the system has reached some (fairly large) threshold number (Allen, 2017; Giordano et al., 2020). A threshold of a 100 reported cases reached in Sweden on March 6 is applied here whereas the data up until March 12 have been used for smoothing.

As there were no deaths from the disease and very few individuals were in intensive care prior to the chosen point of initialization, the measurements that are used as input to the MC model cannot be used for its initialization. Instead, to obtain a county-wise initialization, the cumulative infected cases data reported by the Swedish public health agency were employed (Folkhälsomyndigheten, 2020a). A full disease testing strategy was in effect until March 12 in Sweden, after which testing was heavily restricted (Folkhälsomyndigheten, 2020b). The reported cases during full testing are thus assumed to hold the true number of cumulative infected cases, while incidence data from later times are significantly less reliable and therefore not utilized.

2.2 Initialization model

Since direct inversion of a continuous Markov chain is not easily apprehended, the following linear time-invariant approximation is utilized for the initialization of the model for each county, whereas the model states are lumped over the considered age groups. The latter simplification is introduced since the cases were few in the beginning of the outbreak and patient age was not specified in the data.

The model is derived as a normal approximation of the Poisson distributed forward steps and written as

$$x_{k+1} = Fx_k + w_k, \tag{1}$$

where

$$F = \begin{bmatrix} 1 & 0 & \gamma_A F_1 & \sigma F_0 & 0 \\ 0 & 1 - \gamma_I & \gamma_A F_1 & \sigma F_0 & 0 \\ 0 & 0 & 1 - \gamma_A & \sigma(1 - F_0) & 0 \\ 0 & 0 & 0 & 1 - \sigma & \beta \\ 0 & 1 - e^{-\rho} & \theta_A(1 - e^{-\rho}) & \theta_E(1 - e^{-\rho}) & e^{-\rho} \end{bmatrix},$$

$k = 0, 1, \ldots$ corresponds to daily sampling and w_k is the process noise sequence, whose properties are clarified in Section 2.3. The state vector elements

$$x_k = [I_c(k) \ I(k) \ A(k) \ E(k) \ \phi(k)]^\mathsf{T}$$

stand for the populations of the model compartments according to:

I_c cumulative infectious incidence,
I infected,
A asymptomatic,
E exposed,
ϕ infectious pressure.

The parameters of the model are specified below

σ expected rate of transition from the exposed state,
γ_A expected rate of transition from asymptomatic state,
γ_I expected rate of transition from infected state,
F_0 fraction of transition from exposed reaching the infected state; the remaining fraction reaches the asymptomatic state,
F_1 fraction of transition from asymptomatic state reaching the infected state, The remaining fraction corresponds to the recovery from the disease (not included in (1)),
β indirect transmission rate of the environmental infectious pressure,
ρ infections pressure decay rate,
θ_A asymptomatic viral shedding rate,
θ_E exposed viral shedding rate.

The parameters are positive and so are the elements of the state matrix F. Therefore, model (1) is also positive, i.e. the state vector belongs to the positive quadrant provided the initial condition x_0 and $w_k, k = 0, 1, \ldots$ do. The latter condition restricts the distribution of the process noise.

To obtain the parameter values for model (1), prior distributions for the Bayesian parameter estimation algorithm of the Markov chain model are utilized. The prior distributions are based on empirical data or published estimates.

For parameter values from these distributions, the matrix F tends to have one eigenvalue with magnitude larger than one and is therefore unstable. This is expected, since exponential growth is observed during the early phase of a disease outbreak.

Since the cumulative incidence is the only measured signal, the output of the model is

$$y_k = Hx_k + v_k, \tag{2}$$

where

$$H = [1 \ 0 \ 0 \ 0 \ 0],$$

and v_k is the measurement noise with zero mean and variance R_k. The introduction of measurement noise is a matter of complying with the standard assumptions of Kalman filtering and not an actual model property.

Model (1), (2) does not possess structural observability for the whole range of the parameter values. Some parameter values sampled from the prior distribution make the observability matrix

$$\mathcal{O}^\mathsf{T} = \begin{bmatrix} H^\mathsf{T} & F^\mathsf{T}H^\mathsf{T} & F^{2\mathsf{T}}H^\mathsf{T} & F^{3\mathsf{T}}H^\mathsf{T} & F^{4\mathsf{T}}H^\mathsf{T} \end{bmatrix}.$$

lose rank.

2.3 Process noise covariance

In order to evaluate the process noise covariance, the vector w_k is separated into two terms:

$$w_k = w_{1k} + w_{0k},$$

where w_{1k} describes the error of approximating the stochasticity of the full MC model by the linear dynamics of (1), and w_{0k} captures any other model uncertainty, including both differences between the models (e.g. the spread between counties) and differences between the complete model and the true outbreak dynamics. The process noise covariance matrix Q_k is split accordingly as

$$Q_k = Q_{1k} + Q_0.$$

The model uncertainty is assumed to be additive, independent of k, and uncorrelated between the components. Therefore, Q_0 is diagonal and constant.

The evaluation of the approximation error covariance Q_{1k} is more challenging. When the MC model is sampled, the distribution of an element of w_{1k} is given by a sum of Poisson processes that are shifted to have zero mean, and with variance that depends on the population size in the corresponding compartment. The matrix Q_{1k} is thus state-dependent and evaluated to (3). To avoid confusion with pure time-varying case, the explicit notation is utilized

$$Q_k = Q(x_k).$$

3. SMOOTHING PROBLEM

Let $I_D = [0, d]$ define a finite interval of discrete time instants corresponding to the measurements $y_k, k \in I_D$, and $m \in I_D$ be the point of the MC model initialization.

An estimate $\hat{x}_{m|d}$ of $x_k, k = m$ defined by (1) is then sought from the output data $y_k, k \in I_D$. The problem at hand was approached in three ways presented next.

3.1 Rauch-Tung-Striebel smoother

The Rauch-Tung-Striebel (RTS) smoother (Rauch et al., 1965) is a recursive method for solving fixed-interval smoothing problems. Being an optimal smoother when the noise sources are Gaussian and independent of the system states, the RTS smoother lacks theoretical justification in the present case and even its stability is not readily guaranteed. However, as the results of Section 4 demonstrate, it can nonetheless be used empirically. The stability concerns are not critical as the estimation is performed with a discrete LTI model and on a finite time interval.

The RTS smoother is a two-pass algorithm consisting of a Kalman filter (KF) that is run for the full interval in a forward pass, followed by a backwards pass, when the state estimates are smoothed. The KF equations are solved recursively from the initial conditions $\hat{x}_0$ and $P_{0|0}$

$$\hat{x}_{k|k-1} = F\hat{x}_{k-1|k-1},$$
$$P_{k|k-1} = FP_{k-1|k-1}F^\intercal + Q_k,$$
$$\tilde{y}_k = y_k - H\hat{x}_{k|k-1},$$
$$S_k = HP_{k|k-1}H^\intercal + R_k,$$
$$K_k = P_{k|k-1}H^\intercal S_k^{-1},$$
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\tilde{y}_k$$
$$P_{k|k} = (I - K_kH)P_{k|k-1},$$

where $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$ are the *a priori* and *a posteriori* state estimates, $P_{k|k-1}$ and $P_{k|k}$ are the *a priori* and *a posteriori* estimate covariances, $\tilde{y}_k$ is the innovation, S_k is the innovation covariance, and K_k is the Kalman gain.

Notice that the KF requires knowledge of the covariance matrix Q_k for $1 \le k \le d$. Here, the covariance matrix depends on the unknown states of the system and is not available. Therefore, the plant state is replaced by its estimate, and the covariance matrix is approximated as

$$\hat{Q}_k = Q(\hat{x}_{k|k-1}).$$

The *a priori* and *a posteriori* state and covariance estimates at each time are are saved and the algorithm proceeds backwards from the last time point d. The smoothed estimate $\hat{x}_{k|d}$ is calculated recursively according to

$$\hat{x}_{k|d} = \hat{x}_{k|k} + C_k(\hat{x}_{k+1|d} - \hat{x}_{k+1|k}),$$
$$P_{k|d} = P_{k|k} + C_k(P_{k+1|d} + P_{k+1|k})C_k^\intercal,$$

where $C_k = P_{k|k}F^\intercal P_{k+1|k}^{-1}$ and $P_{k|d}$ is the smoothed estimate covariance.

3.2 Ordinary least squares

The problem of estimating $\hat{x}_{m|d}$ can be approached as an optimization problem and solved once, rather than recursively. The simplest setup is based on the linear relation between the measurement and the state (i.e. backcasting) and leads to the algebraic system

$$y_d = HF^{k-d}x_k + \tilde{w}_k,$$

where the properties of the noise $\tilde{w}_k$ will be elaborated upon in Section 3.3. The state estimation problem is then formulated as

$$\hat{x}_{m|d} = \arg\min_{x_m} ||Y - \Phi x_m||^2, \qquad (4)$$

where

$$Y = [y_1 \ y_2 \ \cdots \ y_d]^\intercal,$$

$$\Phi^\intercal = \left[(F^{1-m})^\intercal H^\intercal \ (F^{2-m})^\intercal H^\intercal \ \cdots \ (F^{d-m})^\intercal H^\intercal \right].$$

Optimization problem (4) can be solved by standard least squares. Furthermore, positivity of the state estimation can be enforced by using constrained least squares.

3.3 Nonlinear least squares

The method of Section 3.2 can be refined by taking into account the correlation of $\tilde{w}_k$. Let $S_Q = \{Q_k\}_{k=0}^d$, and define the matrix $\Omega(S_Q)$ by specifying its elements

$$\omega_{k,l} = \sum_{n=0}^{r_1-1} HF^{r_2}Q_{r_3}F^{\intercal\tilde{n}}H^\intercal + \delta_{kl}R_k,$$

where

$$\tilde{n} = \begin{cases} n, & k < m \\ -n-1, & k \ge m \end{cases},$$

$$r_1 = \begin{cases} \min(|k-m|, |l-m|), & (k-m)(l-m) \ge 0 \\ 0, & (k-m)(l-m) < 0 \end{cases},$$

$$Q_{1k} = \begin{bmatrix} \gamma_A F_1 A(k) + \sigma F_0 E(k) & \gamma_A F_1 A(k) + \sigma F_0 E(k) & -\gamma_A F_1 A(k) & -\sigma F_0 E(k) & 0 \\ \gamma_A F_1 A(k) + \sigma F_0 E(k) & \gamma_A F_1 A(k) + \sigma F_0 E(k) + \gamma_I I(k) & -\gamma_A F_1 A(k) & -\sigma F_0 E(k) & 0 \\ -\gamma_A F_1 A(k) & -\gamma_A F_1 A(k) & \gamma_A A(k) + \sigma(1-F_0)E(k) & -\sigma(1-F_0)E(k) & 0 \\ -\sigma F_0 E(k) & -\sigma F_0 E(k) & -\sigma(1-F_0)E(k) & \sigma E + \beta\phi(k) & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (3)$$

$$r_2 = \begin{cases} \tilde{n} - k + l, & |k - m| \le |l - m| \\ \tilde{n} - l + k, & |k - m| > |l - m| \end{cases},$$

$$r_3 = \begin{cases} k - 1 - \tilde{n}, & |k - m| \le |l - m| \\ l - 1 - \tilde{n}, & |k - m| > |l - m| \end{cases}.$$

Then, $\Omega(S_Q)$ is the covariance matrix of

$$W = [\tilde{w}_0 \ \tilde{w}_1 \ \dots \ \tilde{w}_d]^\mathsf{T}.$$

The terms $\tilde{w}_k$ are thus neither uncorrelated nor homoscedastic and the Gauss-Markov theorem does not apply. If the process noise covariance matrices were independent of the system state, the best linear unbiased estimator would be obtained by solving the optimization problem

$$\hat{x}_{m|d} = \arg\min_{x_m} (Y - \Phi x_m)^\mathsf{T} \Omega(S_Q)^{-1} (Y - \Phi x_m).$$

Since the process noise is state-dependent in this case, the state estimation problem cannot be approached directly. Instead, an estimate

$$\hat{S}_Q(x_m) = \{Q(F^{k-m} x_m)\}_{k=0}^d.$$

is used in the weight matrix of a nonlinear least-squares setup. A solution to the estimation problem is then

$$\hat{x}_{m|d} = \arg\min_{x_m} (Y - \Phi x_m)^\mathsf{T} \Omega(\hat{S}_Q(x_m))^{-1} (Y - \Phi x_m). \quad (5)$$

Since $\Omega(\hat{S}_Q(x_m))$ depends on x_m, optimization problem (5) is nonlinear and solved iteratively by Algorithm 1.

Algorithm 1 Nonlinear least squares

Solve $x = \arg\min_{x_m} ||Y - \Phi x_m||^2$
let $s = \infty$, set s_{tol}
repeat
    let $s_0 = s$
    Calculate $\Omega(\hat{Q}(x))$
    Solve $\hat{x} = \arg\min_{x_m} (Y - \Phi x_m)^\mathsf{T} \Omega(\hat{Q}(x))^{-1} (Y - \Phi x_m)$
    Let $s = (Y - \Phi\hat{x})^\mathsf{T} \Omega(\hat{Q}(x))^{-1} (Y - \Phi\hat{x})$
    Let $x = \hat{x}$
until $|s_0 - s| < s_{tol}$
Let $\hat{x}_{m|d} = x$

3.4 Numerical consideration

For the parametrizations of F in this work, the observability matrix Φ that appears in the introduced least-squares state estimation problems becomes numerically infeasible for large m. The matrix F has eigenvalues that are significantly smaller than one in magnitude, so that repeated inversions of F result in very large elements in Φ. Therefore, the number of estimated states was limited. For parametrizations where one eigenvalue of F is close to zero, the corresponding state was removed through truncation, thus treating the state as identically zero.

The approximation $\hat{x}_k = F^{k-m} x_m$ in the nonlinear least squares formulation can also pose problems, when k is significantly smaller than m. For this reason, a simple regularization was implemented, where $\hat{x}_k$ is set to zero whenever any element of $F^{k-m} x_m$ becomes negative.

4. EXPERIMENTAL RESULTS

The three estimation algorithms introduced above were evaluated using two types of synthetic data. First, linear model (1) was used to generate the data, with the same Poisson-distributed state dependent noise sources as derived for the estimators. Then, the data were generated
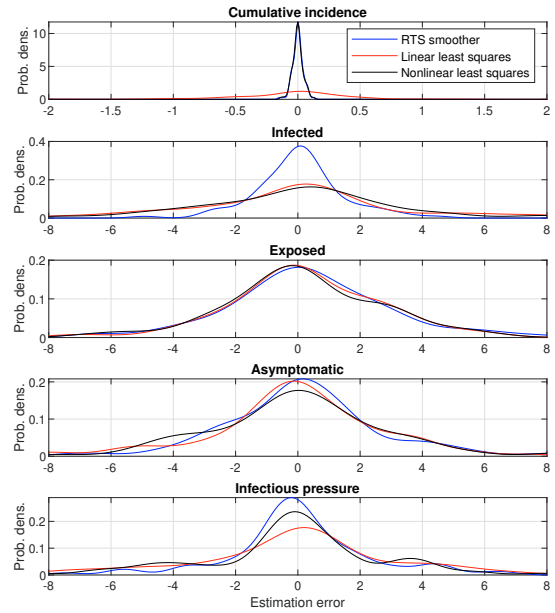

Fig. 2. Estimation error probability density function for synthetic data from simulations of linear model (1).

from stochastic simulations of the MC model. In both cases, the models were simulated repeatedly over a time horizon of 42 days ($d = 42$), from identical initial conditions (distinct between the two cases) and with identical parameter values (identical between the two cases), that were randomly selected from the prior parameter distributions. The probability distributions of the state estimation errors for $m = 30$ were estimated by fitting kernel distribution and compared to each other.

4.1 Synthetic data from linear time-invariant model

The state estimation was performed with the three algorithms for 100 realizations. The process noise covariance was calculated with the diagonal elements of Q_0 set to 0.1, and $R_k = 0.1$. Measurements for indices $k < 19$ were neglected in the batch optimization approaches.

The estimated distributions for all model states are shown in Fig. 2. The RTS smoother appears to perform the best, mostly through lower uncertainty in the infected population estimate. The main difference in performance between the linear and nonlinear least squares is the significantly higher uncertainty in the cumulative incidence estimation for the linear method. This is due to the linear method not exploiting the low uncertainty of the measurement of this state, that is encoded in the covariance model.

4.2 Synthetic data from the Markov chain model

In this case, 50 realizations were generated and data from the three counties with spread of the disease in the highest number of realizations (33, 33 and 31) were analyzed. To reflect the larger modelling error, the diagonal elements of Q_0 were set to 2 and $R_k = 0.5$. As above, indices $k < 19$ were neglected in the batch optimizations. The estimated estimation error distributions for the states I, E and A in the three counties are shown in Fig. 3 – Fig. 5.

Similarly to Section 4.1, the RTS smoother is generally better at estimating the infected population. It is hard to
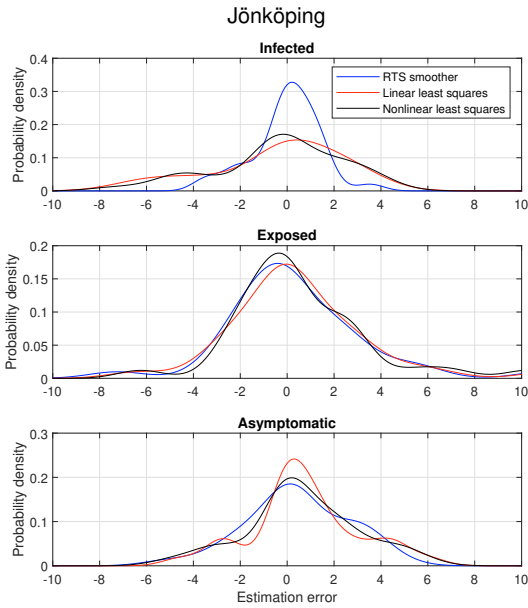
Fig. 3. Estimation error probability density function for Jönköping county based on synthetic data from simulations of the MC model.
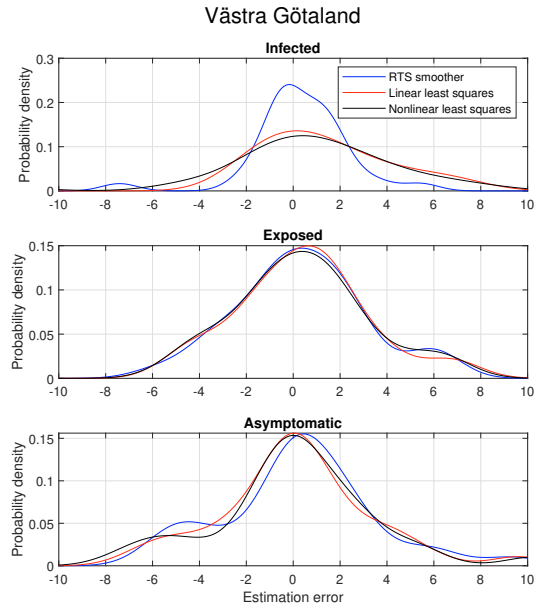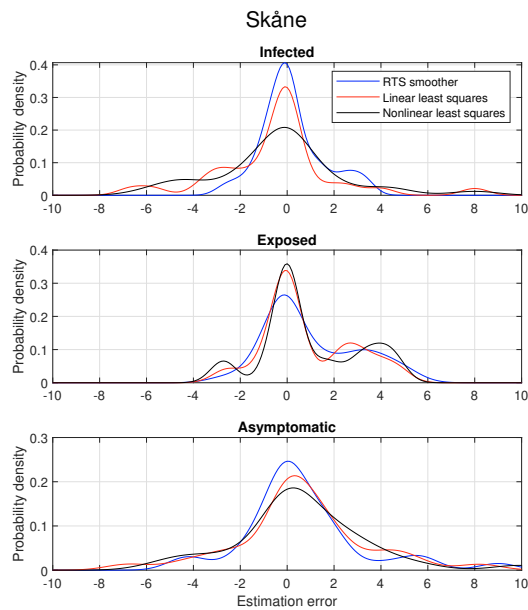

Fig. 4. Estimation error probability density function for Skåne county based on synthetic data from simulations of the MC model.

draw conclusions apart from this, as the characteristics of the distributions vary between the counties.

4.3 Markov chain initialization with different estimates

To investigate the effect of the initial estimation on the MC model, it was simulated using estimated states as initial conditions. The initial state estimates were produced by the three estimation methods from one realization of the simulation of the MC model, see Table 1. The MC model was run 50 times for 42 days from each of the three sets of initial conditions.

Fig. 5. Estimation error probability density function for Västra Götaland county based on synthetic data from simulations of the MC model.

Table 1. State estimates in initalization experiment. Counties not included in the list had no estimated spread of the disease. Västra G. denotes Västra Götaland county and RTS, OLS and NLS denote state estimations using the RTS smoother, ordinary least squares and nonlinear least squares respectively.

	RTS			OLS			NLS		
County	I	E	A	I	E	A	I	E	A
Stockholm	2	3	3	0	4	4	0	4	4
Skåne	24	28	24	7	30	31	12	30	29
Västra G.	24	39	30	36	40	28	29	40	31

The probability distributions of the logarithm of the infected, exposed and asymptomatic populations, in the three counties listed in Table 1, were then estimated using kernel distribution fitting, see Fig. 6 – Fig. 8.

The main conclusion that can be drawn from these results is that the variations between the considered estimation algorithms have limited effect on the states of the system in the end of the simulation, compared to the variations due the stochastic simulation. A greater variance in the states can be observed for the initial conditions generated by the RTS smoother compared to the other, but no general conclusion regarding the initialization methods can be drawn from this, as the results are based on a single estimation instance.

5. CONCLUSION

Three approaches to a fixed-interval linear smoothing problem, one based on the RTS smoother and two on batch optimization methods, have been compared with respect to the initialization of a larger MC epidemiological model. The non-Gaussian state-dependent noise in the linear model makes standard stochastic approaches inapplicable and covariance estimates have been employed in two of the methods. The smoother outperforms the other methods on synthetic data, despite the lack of theoretical justification.
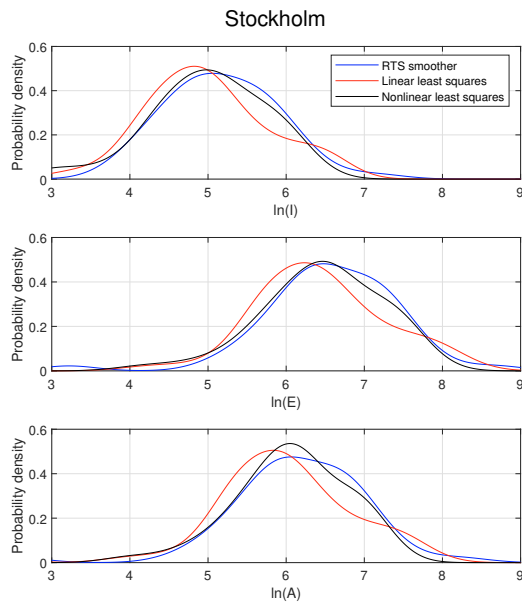
Fig. 6. Probability distributions of model states for Stockholm county according to simulation from estimated initial conditions.
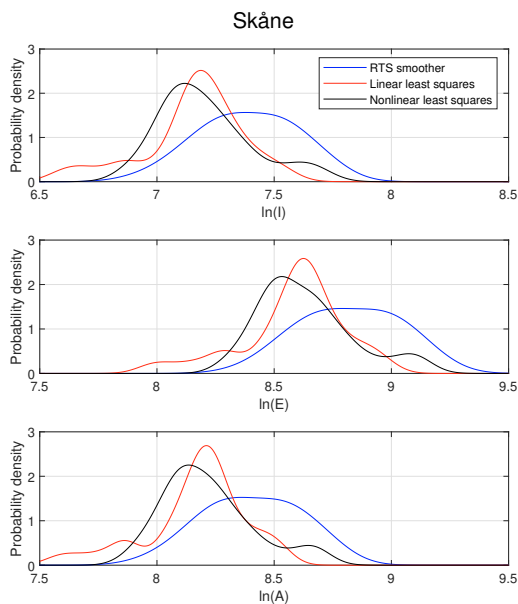

Fig. 7. Probability distributions of model states for Skåne county according to simulation from estimated initial conditions.

Simulations of the MC model from estimated initial conditions indicate that the effect of minor estimation errors is limited compared to the variability due to the model stochasticity. Therefore, computational complexity, robustness, and ease of implementation might be of greater importance than a high initialization accuracy. However, the incoherent impact of the initial conditions on the distributions produced by on the MC model for different considered counties warrants further investigation into the implications of the initialization method selection.
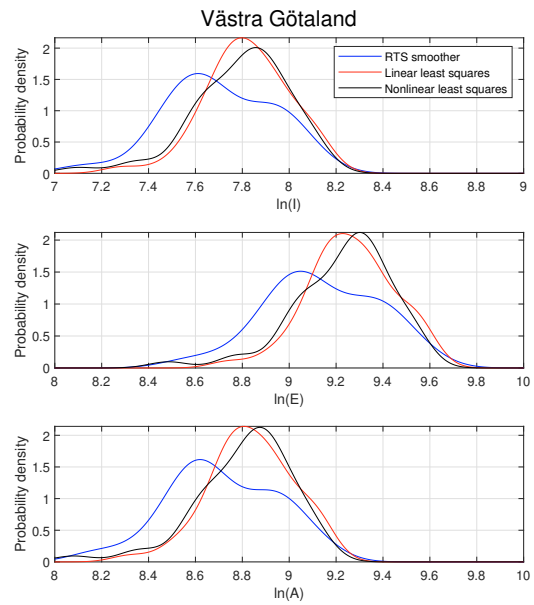

Fig. 8. Probability distributions of model states for Västra Götaland county according to simulation from estimated initial conditions.

REFERENCES

Allen, L.J. (2017). A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infectious Disease Modelling*, 2(2), 128–142.

Aravkin, A. and Burke, J. (2012). Smoothing Dynamic Systems with State-Dependent Covariance Matrices. *Proceedings of the IEEE Conference on Decision and Control*, 2015. doi:10.1109/CDC.2014.7039913.

Engblom, S., Eriksson, R., and Widgren, S. (2020). Bayesian epidemiological modeling over high-resolution network data. *Epidemics*, 32, 100399.

Folkhälsomyndigheten (2020a). Bekräftade fall i Sverige. https://www.arcgis.com/sharing/rest/content/items/b5e7488e117749c19881cce45db13f7e/data. Online; accessed: 2020-06-05.

Folkhälsomyndigheten (2020b). Ny fas kräver nya instatser mot covid 19. https://www.folkhalsomyndigheten.se/nyheter-och-press/nyhetsarkiv/2020/mars/ny-fas-kraver-nya-insatser-mot-covid-19/. Online; accessed: 2020-06-30.

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., and Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 1–6.

Rauch, H., Tung, F., and Striebel, C.T. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8), 1445–1450.

Wang, G., Zhang, Y., and Wang, X. (2020). Maximum correntropy Rauch-Tung-Striebel smoother for nonlinear and non-gaussian systems. *IEEE Transactions on Automatic Control*, PP, 1–1.

Widgren, S., Bauer, P., Eriksson, R., and Engblom, S. (2019). SimInf: An R package for data-driven stochastic disease spread simulations. *J. Stat. Softw.*, 91(12), 1–42.

Widgren, S. et al. (2018). Spatio-temporal modelling of verotoxigenic *E. coli* O157 in cattle in Sweden: Exploring options for control. *Veterinary Res.*, 49(78). doi: 10.1186/s13567-018-0574-2.