SIRS-EAKF Model In German Flu Forecasting

Presented by Xi Meng, Yukun Li, Nutcha Wattanachit

03/06/2019

What is the SIRS-EAKF model (again)?

The SIRS-EAKF model is a mechanistic that relies on state space estimation methods. The model has two frameworks to make forecast: SIRS and EAKF. Forecast accuracy of the model depends on factors in the prediction system:

- Model initial conditions
- Stochastic observation error
- Model misspecification

In particular, the first and the third factors can lead to error growth. In addition, the sensitivity to initial conditions can be crucial for forecasting.

First part of the model: Humidity-driven SIRS

This is a simple SIRS model modulated by local absolute humidity (AH) conditions is employed to simulate influenza outbreak dynamics. Assuming a perfectly mixed population, the model equations describing a local outbreak are:

$$\frac{dS}{dt} = \frac{N - S - I}{L} - \frac{\beta(t)IS}{L} - \alpha,$$

$$\frac{dI}{dt} = \frac{\beta(t)IS}{L} - \frac{I}{D} + \alpha,$$

where $S =$ susceptible population, $I =$ infected population are model variables. $N =$ total population, $\beta(t) =$ contact rate at time t, $L =$ average duration of immunity, $D =$ mean infectious period, and $\alpha =$ the rate of infection introduction from external sources are model parameters.

First part of the model: Humidity-driven SIRS (continued)

The contact rate, $\beta(t)$, is forced by local AH conditions through

$$R_0(t) = \beta(t)D = e^{a \times q(t) + b} + R_{0min},$$

where $q(t)$ is observed specific humidity (a measure of AH), $R_0(t)$ is the basic reproductive number and R_{0min} are the minimum daily basic reproductive number.

First part together with the second part of the model: The SIRS-EAKF framework

The SIRS-EAKF system iteratively optimizes the distribution of state variables and parameters of the SIRS model using a sequential ensemble filtering technique called the Ensemble Adjustment Kalman Filter (EAKF).

The state vector at time t is $\mathbf{x}_t = (S_t, I_t, R_{0max}, R_{0min}, L, D)$. Once the observation O_t at time t is observed, the posterior distribution of the system state is obtained by incorporating the information from the new observation through Bayes' rule:

$$p(\mathbf{x}_t|O_{1:t}) \propto p(\mathbf{x}_t|O_{1:t-1})p(O_t|\mathbf{x}_t),$$

where $p(\mathbf{x}_t|O_{1:t-1})$ is the prior distribution of the system state, $p(O_t|\mathbf{x}_t)$ is the likelihood of observing O_t given the state $\mathbf{x}_t$, and $O_{1:t}$ are the observations taken up to time t.

The SIRS-EAKF framework (continued)

- In the EAKF, unobserved variables, such as the susceptible population S, and model parameters, are adjusted depending on covariant relationships with the observed variables.
- In Kalman filtering the intervariable relationships are assumed to be linear. As a consequence, the adjustments of unobserved variables and parameters are linearly related to the adjustment of the observed variable through their covariance, which is computed directly from the ensemble.
- To initialize the SIRS-EAKF system, an ensemble of state vectors was randomly selected from a collection of possible variable and parameter combinations.

Error Correction in Perturbed SIRS Simulations

- 1. Input: $\mathbf{x}^{t-1} = (S^{t-1}, I^{t-1}, R_{0max}^{t-1}, R_{0min}^{t-1}, D_{t-1}, L_{t-1})^T$, and observation I_{obs}^t at t week.
- 2. Breeding method: impose Gaussian-distributed random errors on S_{t-1} to form multiple pertubed trajectories $\mathbf{x}_p^{t-1}$, integrate both $\mathbf{x}^{t-1}$ and $\mathbf{x}_p^{t-1}$ for one week to get $\mathbf{x}^t$ and $\mathbf{x}_p^t$.
- 3. The bred errors at t week are $\Delta S = S_p^t - S^t$ and $\Delta I = I_p^t - I^t$, then fit the error structure with a $3^{rd}-$order polynomial.
- 4. The discrepency in observed variable is $\Delta I = I^t - I_{obs}^t$. Calculate the structure error ΔS using the fitted error structure.
- 5. Output:

$$\mathbf{x}_{adj}^t = (S^t - \Delta S, I_{obs}^t, R_{0max}^{t-1}, R_{0min}^{t-1}, D_{t-1}, L_{t-1})^T$$

Note: Here we focus on variable S, the same analysis on R_{0max} can be applied.

Error Correction in Perturbed SIRS Simulations (continued)

However, the application in a realistic setting is complicated by several issues:

- 1. Effective error correction is only possible whe the system state is not far from the truth.
- 2. Structural errors can only be clearly diagnosed for sensitive state variables.

Therefore, the error correction process is implemented in conjunction with Ensemble Adjustment Kalman Filter(EAKF) which can provide an accurate initial estimate of the system state.

Trajectory Adjustment with EAKF

- 1. For each observation of weekly incidence, the EAKF calculates an adjustment of the observed state variable using Bayes' Rule.
- 2. The unobserved state variables and parameters are then adjusted based on their prior covariance with the observed state variable.

The update, i.e., the adjustment of the prior state, can be interpreted as the EAKF estimate of the error of the state variables and parameters.

Implementation of Error Correction with EAKF (EAKFC)

Subtract EAKF-estimated errors prior to application of the breeding method. For instance, when diagnosing the structure error in S

- 1. Remove the EAKF-estimated errors in R_{0max}, R_{0min}, D and L from the posterior trajectory at t-1.
- 2. Impose random errors on S that are evolved using the SIRS model to generate the error structure at time t.

Implementation of Error Correction with EAKF (EAKFC)

Once the error structure is obtained, the discrepancy Δobs of the observed variable from observation should be determined.

Instead of using the observation deirectly as the truth, we use EAKF posterior, obs_{post}, which is a weighted average of the prior and observation. Therefore, the discrepancy in observation is simply $\Delta obs = obs_{bred} - obs_{post}$, where obs_{bred} is the incidence of the unperturbed trajectory at time t.

Methods: Optimization of Adaptive Error Correction

- To avoid inappropriate error corrections that might undermine accurate EAKF estimation, the correction procedure is applied selectively rather than indiscriminately to all ensemble members.
- At each iteration, a new configuration θ' is generated by the original configuration θ. Whether the new configuration θ' is accepted depends on the acceptance probability function

$$P(E(\theta), E(\theta'), T_k) = \begin{cases} 1 & E(\theta') > E(\theta) \\ exp\{-\frac{100(E(\theta')-E(\theta))}{T_k}\} & E(\theta') \leq E(\theta) \end{cases}$$

Always accept new configurations when there is improvement; worse configurations also have a chance to be accepted.

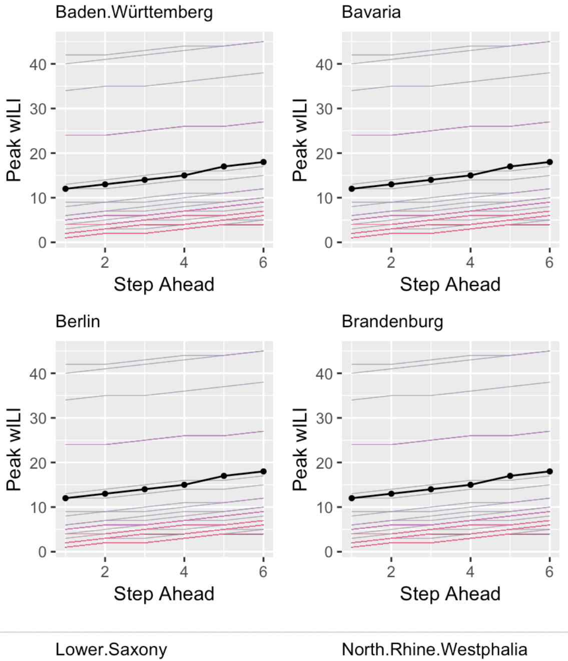
The Application of SIRS-EAKF in German Flu Forecasting Using Forecast Framework

We used the authors' initial parameters in the SIRS. Average daily absolute humidity was obtained from a representative station in Germany from 2012-2015.

Coding and Challenges

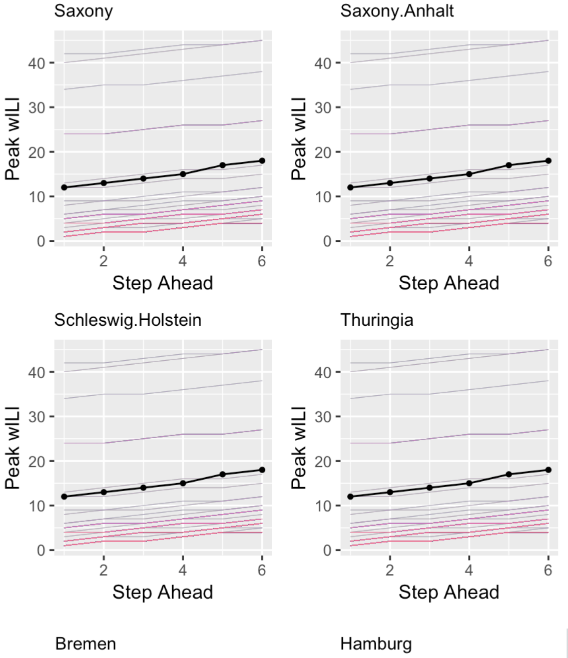
- There are a large number of fixed initial parameters in the model that were generated using Latin hypercube. This could lead to error growth in the prediction system. Absolute humidity is also not a commonly available measurement.
- The model has high complexity. Particularly, the error correction part in the paper was not implemented since it involves imaginary numbers that R could not handle.
- Due to the complexity of the model and the fact that it relies on a large number of simulations to generate or update parameters, this model is difficult to replicate and requires a lot of fine-tuning.

Results



Baden.Württemberg
Bavaria
Berlin
Brandenburg
Peak wILI
Step Ahead
0
10
20
30
40
2
4
6
Lower.Saxony
North.Rhine.Westphalia

Results (continued)



Discussion

Reference

- Pei S., Shaman J. Counteracting structural errors in ensemble forecast of influenza outbreaks. Nature Communicationsvolume 8, Article number: 925 (2017)