Research Your Researcher

Bayesian Inference in Epidemiology

Zhen-Yen Chan Erwan Delorme Konstanty Kowalewski
Jonathan Lee Patricia Yin

Outline

Introduction
- Our researcher and research topic
- Overview of three papers

Dureau et al. (2013)
- Methodology
- Results
- Evolution

Conclusion
- Coronavirus

Our researcher and research topic

Researcher: Dr Kostas Kalogeropoulos

- Computational methods for Bayesian Inference

Our researcher and research topic

Researcher: Dr Kostas Kalogeropoulos

- Computational methods for Bayesian Inference

Research topic: Epidemic Modelling

- Transmission dynamics of disease
- Predict potential outbreaks
- Assess effectiveness of public health interventions

Overview: Malesios et al. (2017)

Bayesian epidemic models for spatially aggregated count data. C. Malesios, N. Demiris, K. Kalogeropoulos, and I. Ntzoufras. 2017.

- Problem tackled: To address the spatial nature of the disease spread
- Measures: Bayesian modelling, branching process-based methods

Overview: Dureau et al. (2016)

A Bayesian approach to estimate changes in condom use from limited human immunodeficiency virus prevalence data. J. Dureau, K. Kalogeropoulos, P. Vickerman, M. Pickles, M. Boily. 2016.

- Aim: to estimate condom use time trends from HIV prevalence data to assess the effectiveness of public health interventions
- Problem tackled: bias, limited self-reported data
- Measures: a Bayesian inference methodology that incorporates an HIV transmission dynamics model

Overview: Dureau et al. (2013)

Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. J. Durau, K. Kalogeropoulos, M. Baguelin. 2013.

- The methodology illustrated on data from A/H1N1 (2009) pandemic in England
- A flexible modelling framework encompassing time-varying aspects of the epidemic

Outline

Introduction
- Our researcher and research topic
- Overview of three papers

Dureau et al. (2013)
- Methodology
- Results
- Evolution

Conclusion
- Coronavirus

Methodology: SEIR

The **SEIR** model divides the population into four groups:

- **S**usceptible
- **E**xposed (infected but not infective)
- **I**nfective
- **R**emoved (immune or recovered)

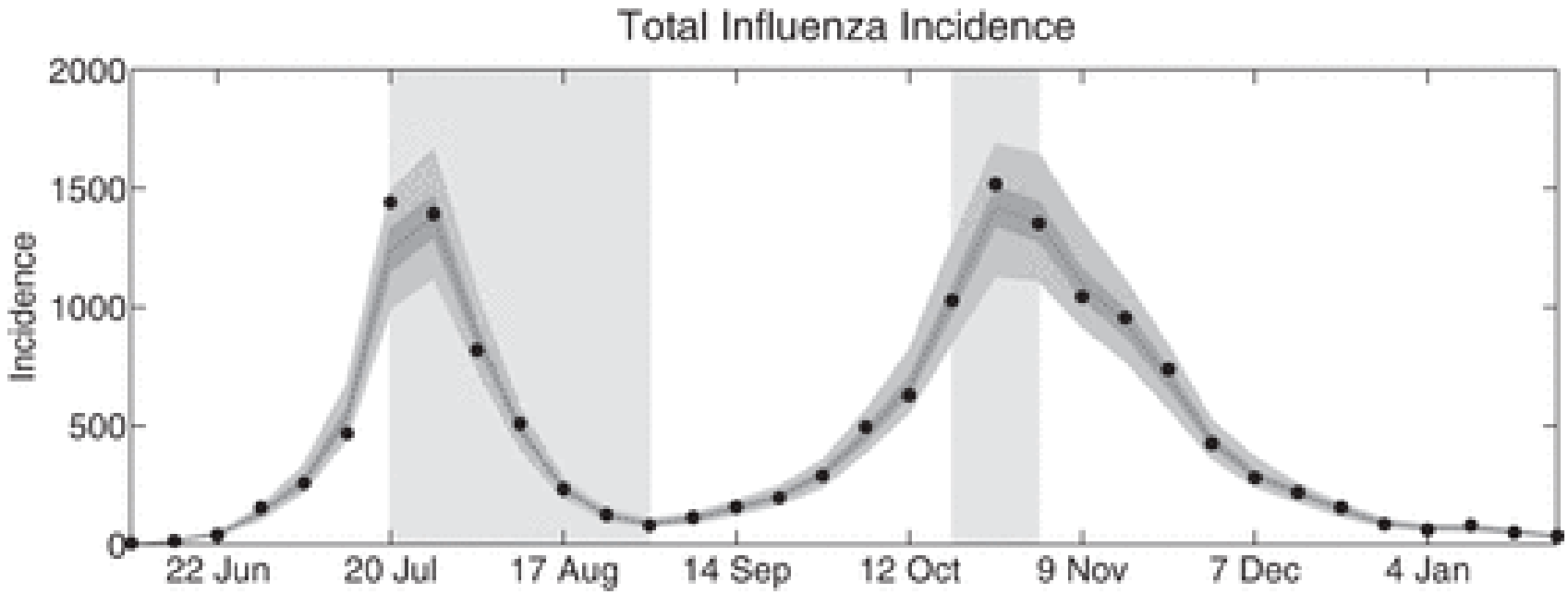
$$\frac{dS_t}{d_t} = -\beta S_t \frac{I_t}{N}, \ \frac{dE_t}{d_t} = -\beta S_t \frac{I_t}{N} - kE_t, \ \frac{dI_t}{d_t} = kE_t - \gamma I_t, \ \frac{dR_t}{dt} = \gamma I_t$$

Assumption: all individual characteristics are fixed over time

Reaction	Effect	Rate
infections	$(S_t, E_t, I_t, R_t) \rightarrow (\boldsymbol{S_t - 1}, \boldsymbol{E_t + 1}, I_t, R_t)$	$\beta_t I/N$
onset of symptoms	$(S_t, E_t, I_t, R_t) \rightarrow (S_t, \boldsymbol{E_t - 1}, \boldsymbol{I_t + 1}, R_t)$	k
recovery	$(S_t, E_t, I_t, R_t) \rightarrow (S_t, E_t, \boldsymbol{I_t - 1}, \boldsymbol{R_t + 1})$	γ

β is the effective contact rate which directly reflects transmissibility

Methodology: suitability of the model



Time-varying factors: behaviour changes, preventive measures, seasonal effects, holidays

Methodology: adding randomness

Introduce randomness in the parameters (Bayesian) and model β as a probability distribution. Assumption: population model, the random individual effects (biological factors, mode and frequency of travel) are aggregated in the data and hence captured in β_t.

$$\begin{cases} \frac{dS_t}{d_t} = -\beta_t S_t \frac{I_t}{N}, \ \frac{dE_t}{d_t} = -\beta_t S_t \frac{I_t}{N} - kE_t, \ \frac{dI_t}{d_t} = kE_t - \gamma I_t, \ \frac{dR_t}{dt} = \gamma I_t \\ dx_t = \mu_x(x_t, \theta_x)dt + \sigma_x(x_t, \theta_x)dB_t, \ x_t = h(\beta_t) \end{cases}$$

Add diffusion process x_t for β_t which will solve the stochastic differential equations in our model.

$$\pi(x_{0:n}|y_{1:n}) \propto f(y_{1:n}|V_{0:n}, \theta_y) \times d\mathbb{P}_x \times \pi(\theta)$$

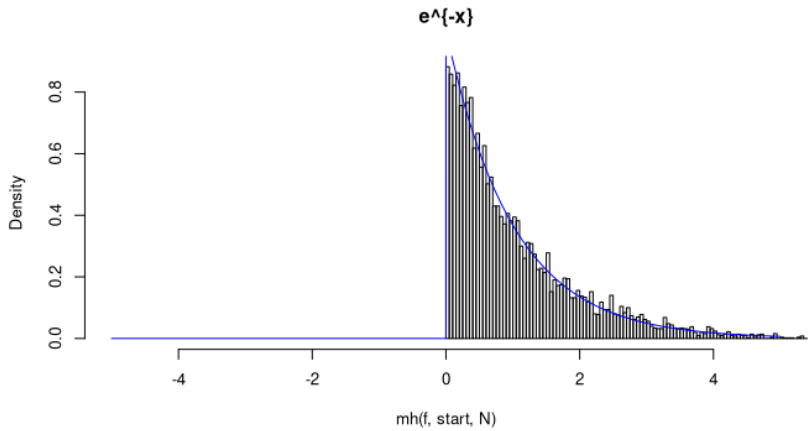
Sample the posterior of x given the priors using a specific Markov Chain Monte Carlo (MCMC) type algorithm.

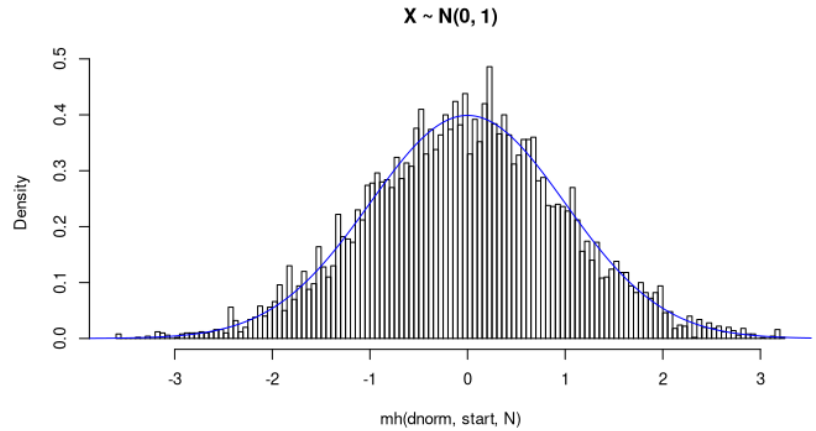
Motivation for Markov Chain Monte Carlo (MCMC)

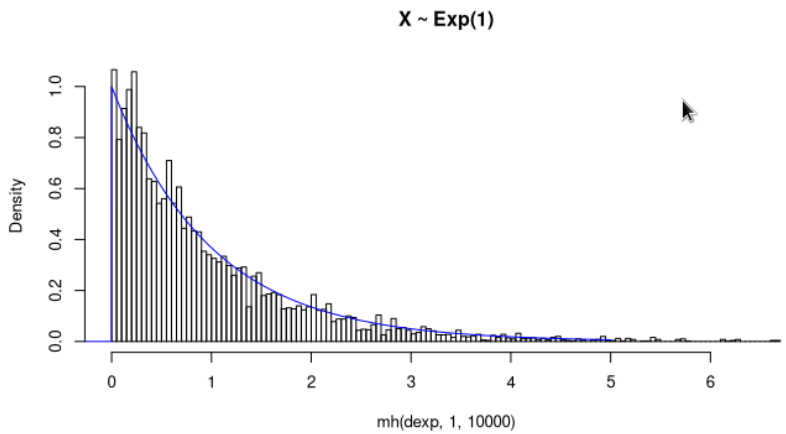
- Intractable integrals $\rightarrow$ approximate by sampling (Monte Carlo)
- No closed form of the density to sample from (by e.g. reverse integral transform) $\rightarrow$ use Markov chains
- A general name for algorithms based on this is Markov Chain Monte Carlo (MCMC)

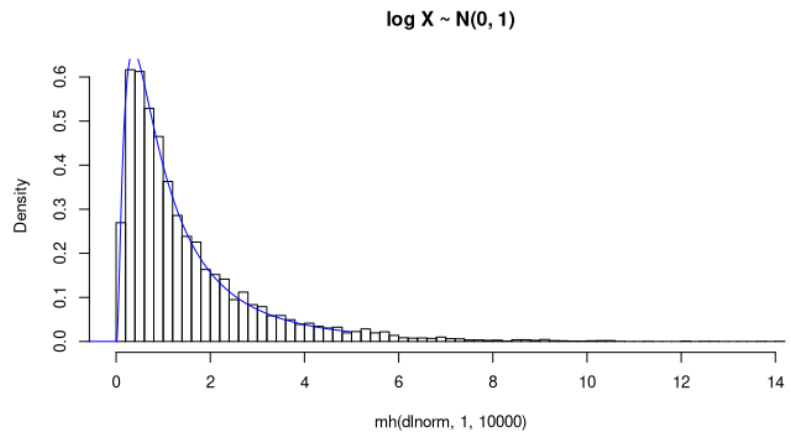
Metropolis-Hastings (random walk Metropolis) example

```
mh <- function(target, starting, N) {
    x <- rep(0, N)
    x[1] <- starting
    for(i in 2:N) {
        last <- x[i-1]
        # Proposal from N(last, 1)
        p <- rnorm(1, mean=last, sd=1)
        # Compute acceptance probabilty
        accept <- target(p) / target(last)
        # Either with acceptance probability
        if(runif(1) < accept) {
            x[i] <- p
        } else {
            x[i] <- last
        }
    }
    return(x)
}
```


e^{-x}
Density
0.0
0.2
0.4
0.6
0.8
-4
-2
0
2
4
mh(f, start, N)


X ~ N(0, 1)
Density
0.0
0.1
0.2
0.3
0.4
0.5
-3
-2
-1
0
1
2
3
mh(dnorm, start, N)


X ~ Exp(1)
Density
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
mh(dexp, 1, 10000)


log X ~ N(0, 1)
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14
mh(dlnorm, 1, 10000)

Challenges of MCMC

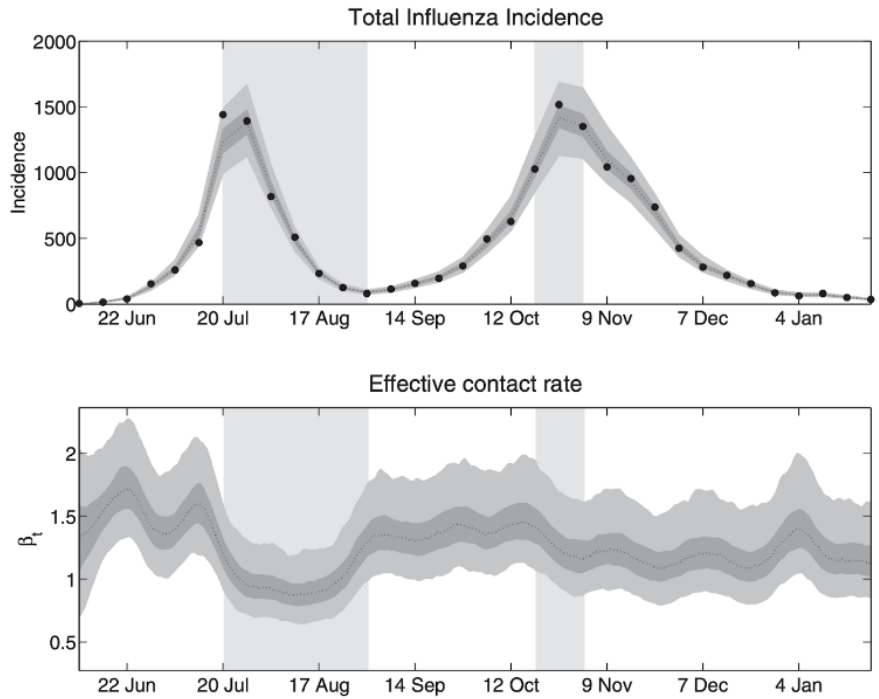
- Mixing – relates to dependency of samples obtained with MCMC algorithm and how fast we can get them, can be assessed by Estimated Sample Size (ESS)
- Let $x_t = h(\beta_t)$ where h is a positive-valued function, authors use natural log
- We want to obtain joint posterior density $p(x, \theta \mid y)$ – this is a non-standard problem since
  - The x is high-dimensional (it's the path of β)
  - There is some dependence between x and θ
  - In effect, classic methods are highly inefficient in mixing

Solution proposed by authors

- Particle filters (PF) and Particle MCMC (PMCMC)
  - PF gives us $p(x \mid y, \theta)$ and $p(y \mid \theta)$ (Doucet and Johansen, 2009)
  - PMCMC gives us $p(x, \theta \mid y)$ (Andrieu et al., 2010)
- There are still some issues with efficiency
- Adaptive Metropolis (Roberts and Rosenthal, 2009)
- Other methods such as Extended Kalman Filter (EKF) are considered but they require some assumptions which would contrain the generality of the proposed solution

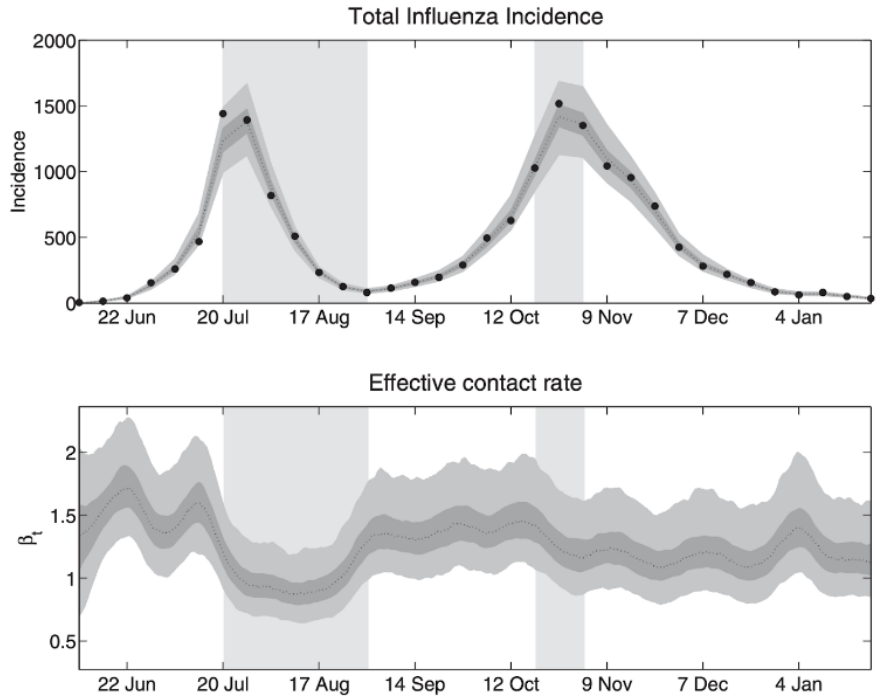
Results: effect of school closures

- Black dots indicate estimates of observed incidence
- Grey rectangular areas indicate school breaks
- Dark and light area indicates the credible intervals
- Offline estimates of the effective contact rate



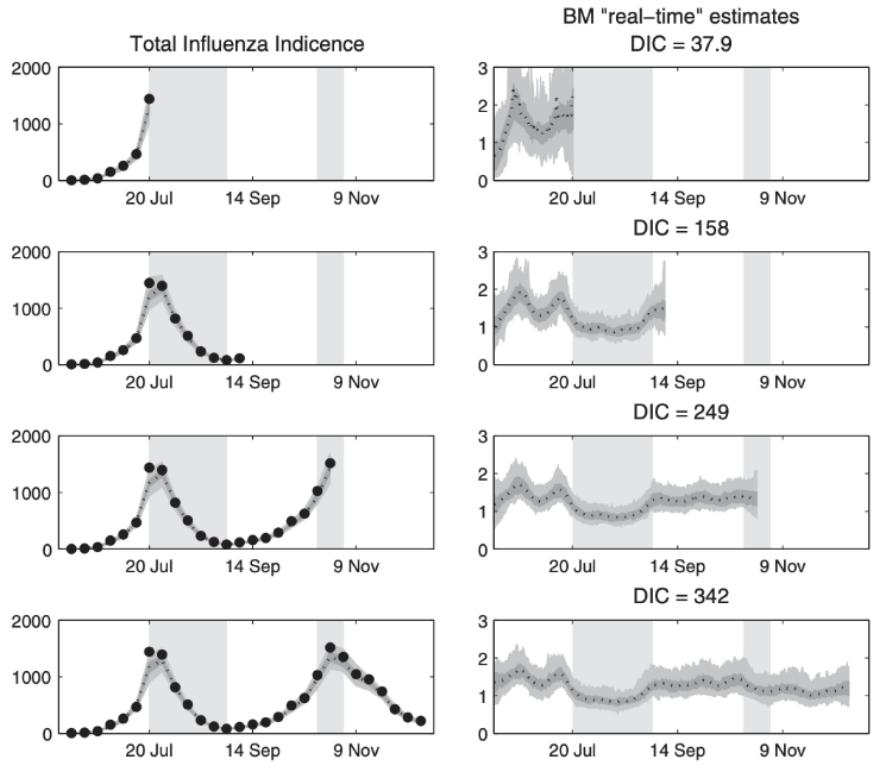
Results: effect of school closures

- Holidays periods have shown less total incidences.
- This could also mean that the epidemic had stopped because a critical population already had been infected, conferring that there is groups that have the immunity to stop the epidemic.
- The rise in influenza in the second grey area could contribute to the fact that schools were reopened and consequently effective contact rate increased them again.



Results: relationship between influenza cases and the real-time estimates

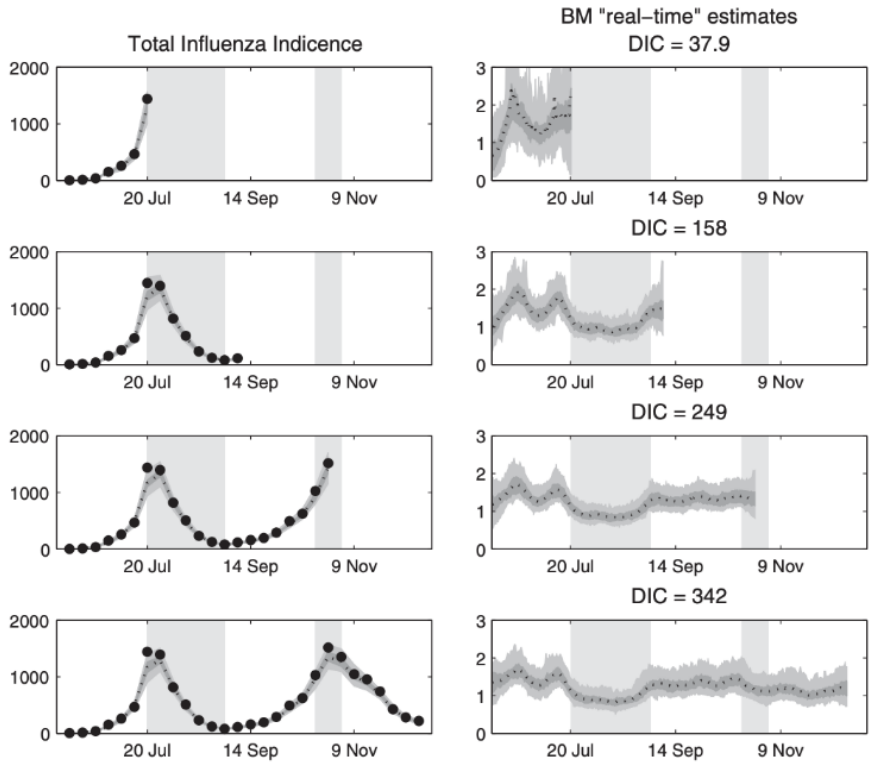
- Rise of H1N1 incidence fell throughout the school holidays and started rising again.
- An assumption could be made that decision-makers did not have access to the effective contact rate estimates data.
- From the data, it could be seen that the real-time estimates were still rising at the first grey area.
- Reopening schools while the effective contact rate was still rising contributed to another rise in total incidence.



Results: relationship between influenza cases and the real-time estimates

Some things to consider

- The effects of school holidays differ from children and adults
- We assumed here a homogeneous population.
- Will be more precise to consider a model with two age groups



Evolution of the subject

- SSM: Inference for time series analysis with State Space Models. Dureau et al. 2013.
- Optimal control and the value of information for a stochastic epidemiological SIS-model. Grandits et al. 2019.
- Efficient real-time monitoring of an emerging influenza pandemic: how feasible? Birrell et al. 2019.

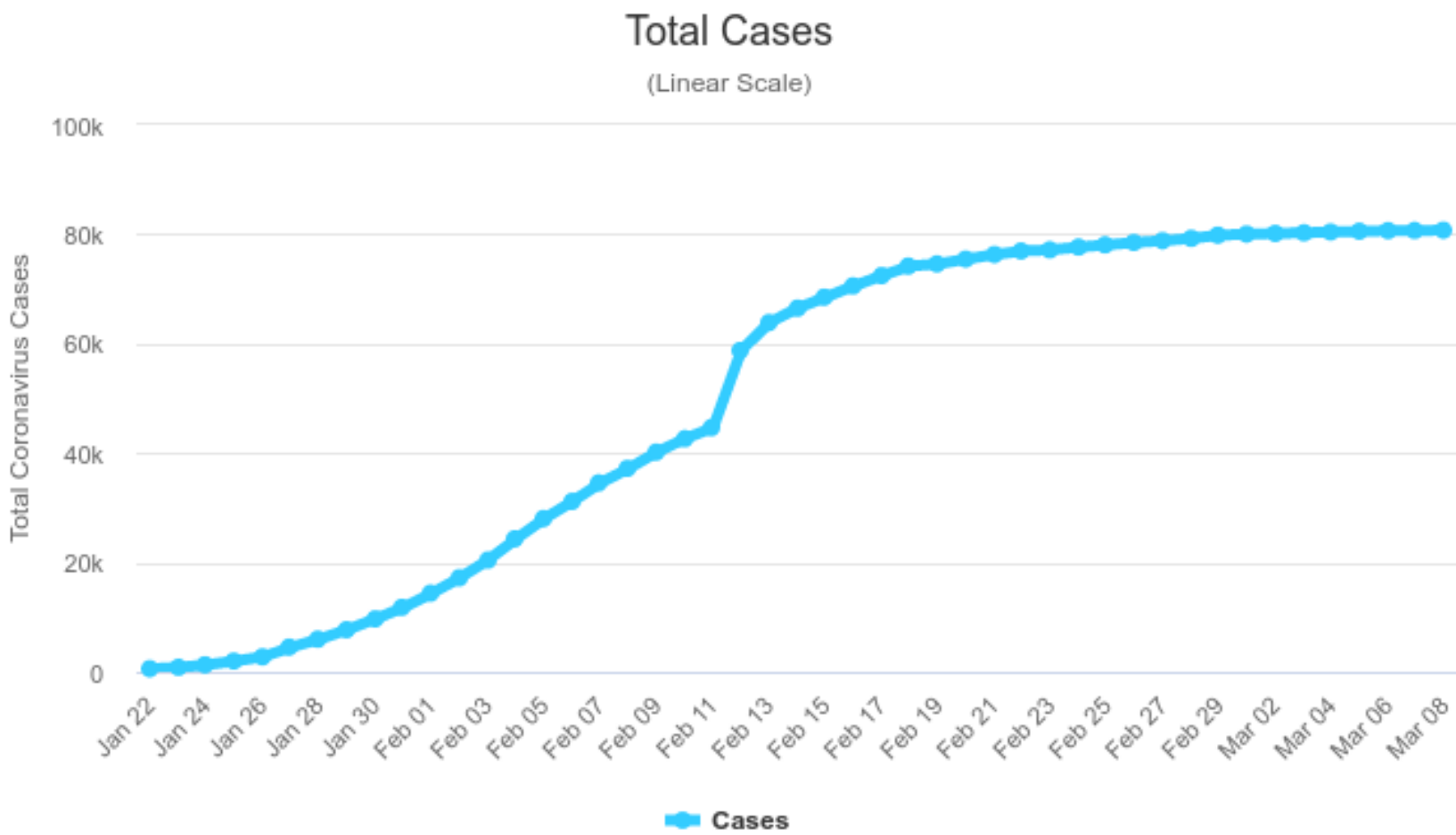
Outline

Introduction
- Our researcher and research topic
- Overview of three papers

Dureau et al. (2013)
- Methodology
- Results
- Evolution

Conclusion
- Coronavirus

Conclusion: current coronavirus cases in China



Conclusion: general remarks and questions

Bayesian Inference is a powerful tool, applicable to a vast range of problems. If you like the subject, consider taking ST308, a course on Bayesian Inference lectured by Dr Kalogeropoulos.

We're now happy to take any questions.

Feel free to message us at konstanty@kszk.eu

Thanks!