Bayesian Nonparametric Inference for Heterogeneously Mixing Epidemic Models

Rowland Seymour, Philip O'Neill, Theodore Kypraios, Thomas Hagenaars

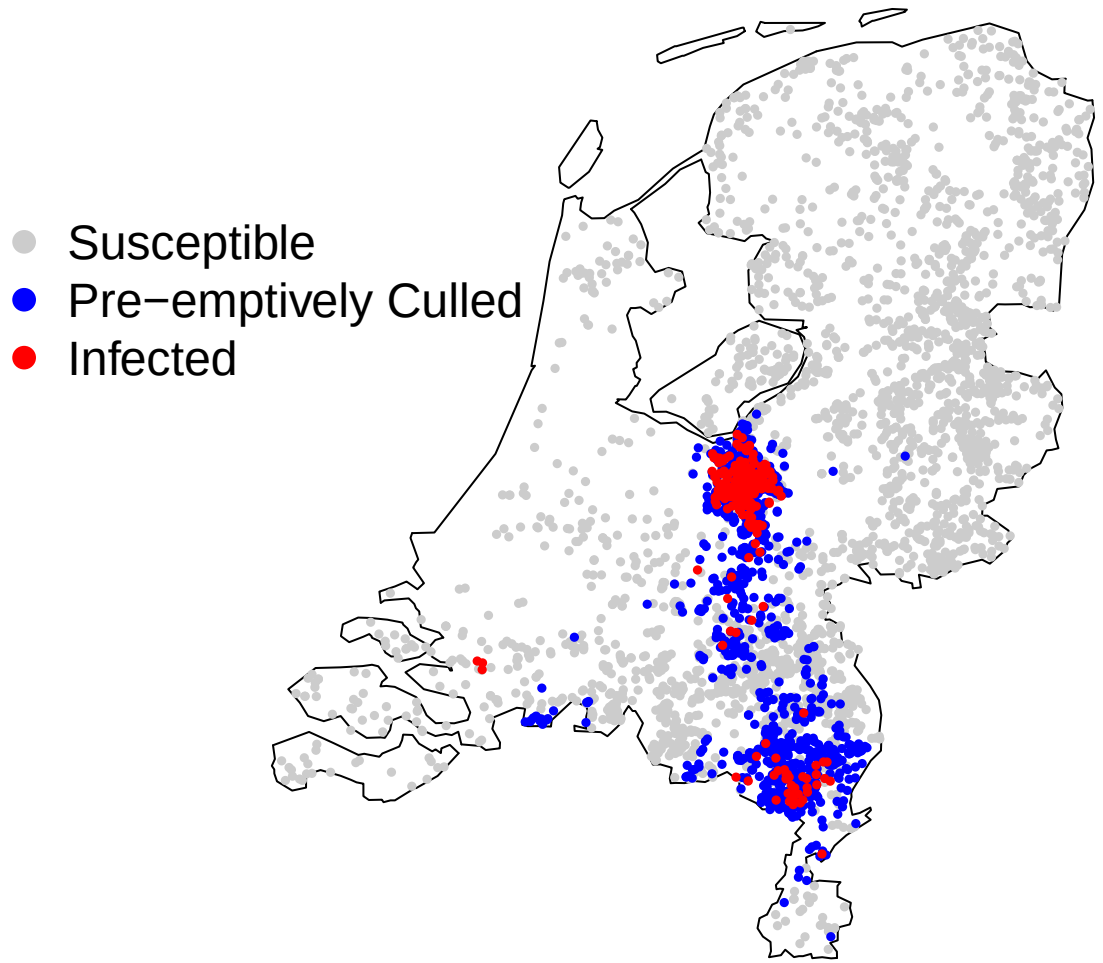
ModAH^2, 16^{th} September 2021

University of Nottingham
UK | CHINA | MALAYSIA

Engineering and Physical Sciences Research Council

Table of Contents

Avian Influenza



Approach and Aims

From the outbreak of Avian Influenza, we observe the following data:

ID	Coordinates	Status	Culling Date
1	(5.32, 18.82)	Susceptible	*NA*
2	(2.90, 15.67)	Susceptible	*NA*
3	(2.86, 17.99)	Pre-Emptively Culled	3^{rd} May
4	(4.56, 18.01)	Culled	30^{th} April
⋮	⋮	⋮	⋮

Using this data, we want to work out:

- If there is a spatial element to the spread of the disease?
- If so, how can we quantify this element?

Table of Contents

Homogeneous Mixing

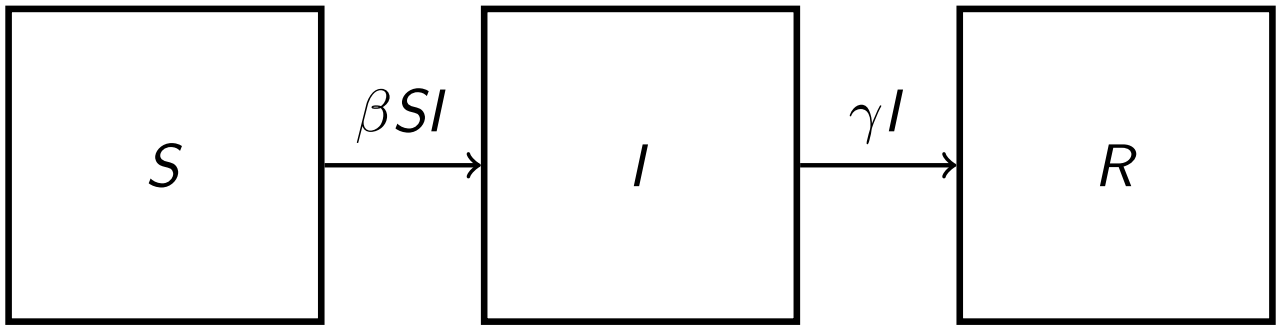


Figure: A homogeneously mixing SIR model.

Farms are either **S**usceptible, **I**nfected or **R**emoved.

- **Infections:** Infections occur according to a Poisson process with rate $\beta S_t I_t$.
- **Removals:** Individual j remains infectious for a time period drawn from an Exponential distribution, such that $r_j - i_j \sim Exp(\gamma)$.

Heterogeneous Mixing

We need to include the distance between farms in the model. We compute the euclidean distance between each pair of farms farm i and j, $d_{i,j}$ and we make β and function of the distance.

$$\beta_{i,j} = f(d_{i,j})$$

Our heterogeneously mixing model builds on the homogeneously model, as we specify the infection and culling rates as follows:

- **Infections:** Infections occur according to an inhomogeneous Poisson process with a unique rate for each pair of farms, specified by $\beta_{i,j} = f(d_{i,j})$.
- **Removals:** Individual j remains infectious for a time period drawn from an Exponential distribution, such that $r_j - i_j \sim Exp(\gamma)$.

Table of Contents

Parametric Infection Rates

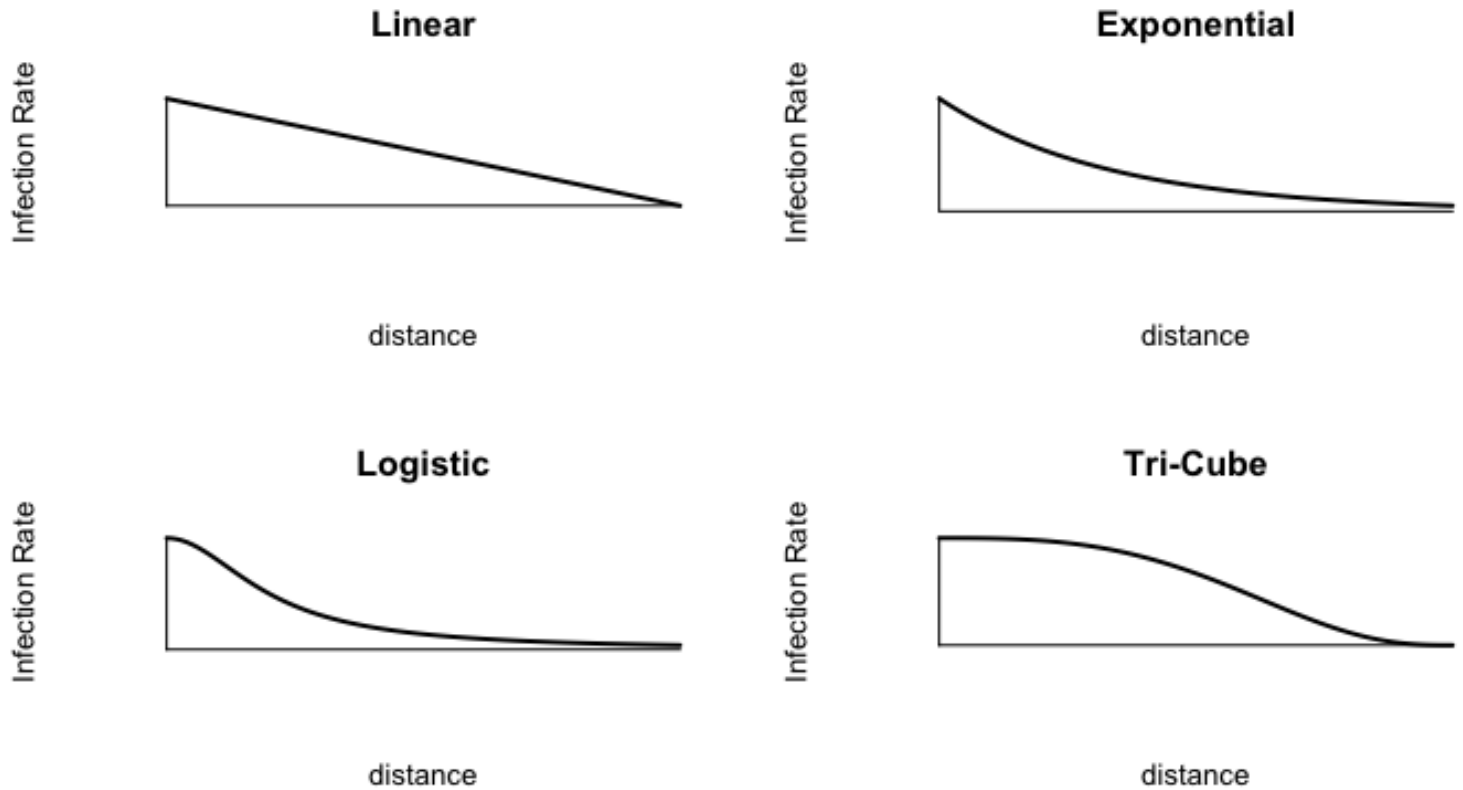


Figure: Four Possible Infection Rates.

Parametric Models for Avian Influenza

Parametric inference for this data set has already been carried out by Boender *et al.* (2007). They proposed five models and used AIC to choose the best of the proposed models.

Model	Infection Rate
1	$\beta_{i,j} = \beta_0$
2	$\beta_{i,j} = \frac{\beta_0}{1+d_{i,j}}$
3	$\beta_{i,j} = \frac{\beta_0}{1+d_{i,j}^2}$
4	$\beta_{i,j} = \frac{\beta_0}{1+d_{i,j}^\alpha}$
5	$\beta_{i,j} = \frac{\beta_0}{1+(d_{i,j}/\beta_1)^\alpha}$

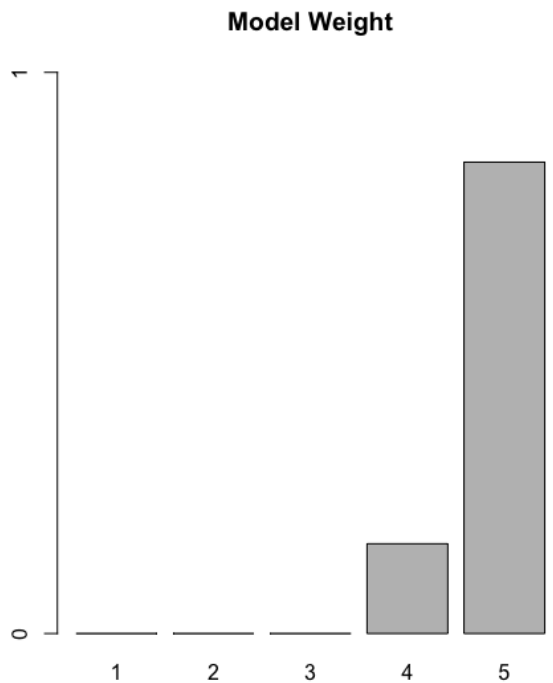


Figure: Model Weights for the different Infection Rates.

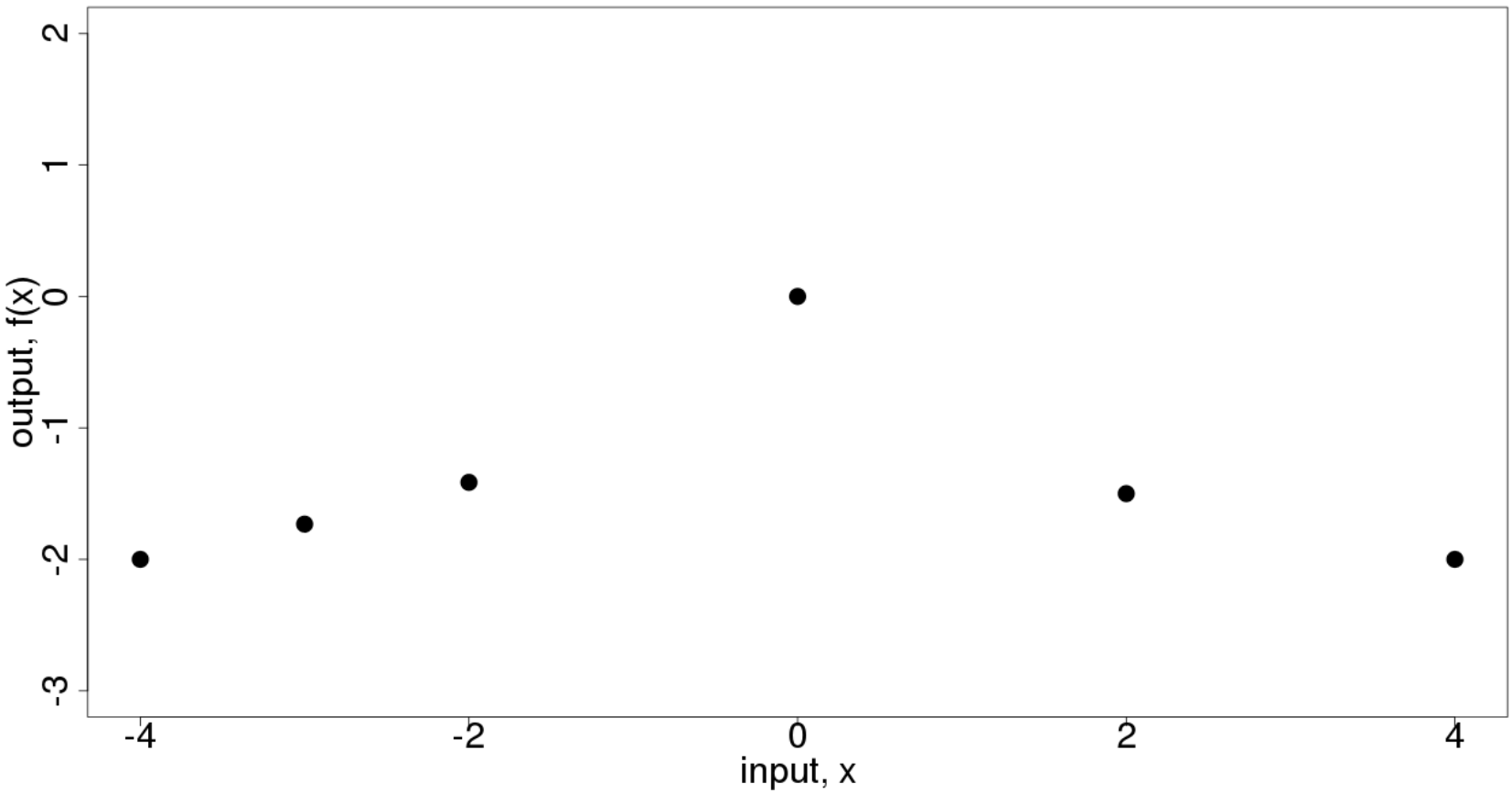
Inference for Nonparametric Models

Our nonparametric methodology use Gaussian Processes (GPs) to estimate $\beta_{i,j}$.

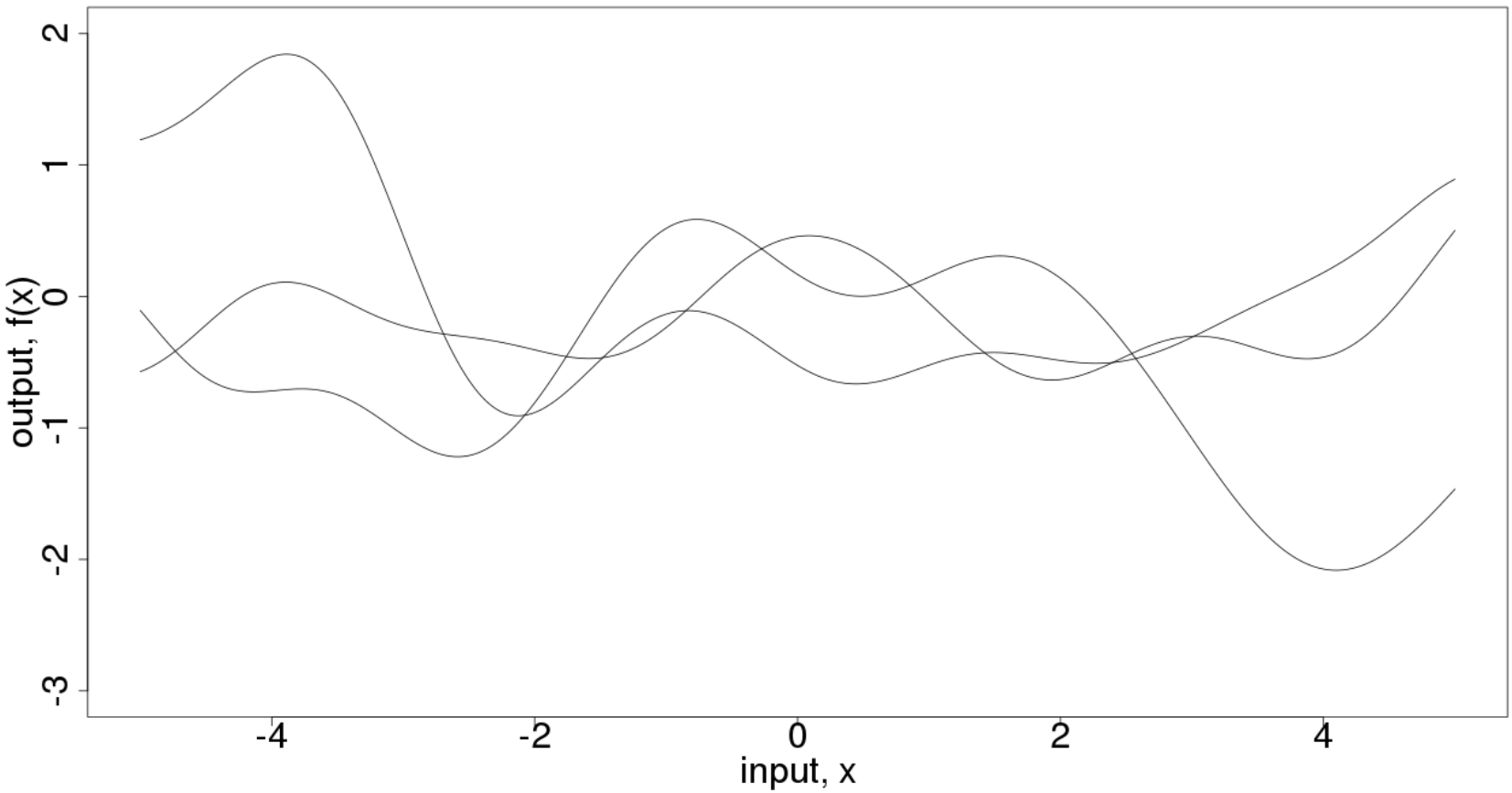
- GPs are a generalisation of the multivariate Gaussian distribution to a function space.
- We use GPs to assign a prior distribution over a function space.
- We need to control the signal variance and volatility.

$$f(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, \Sigma)$$

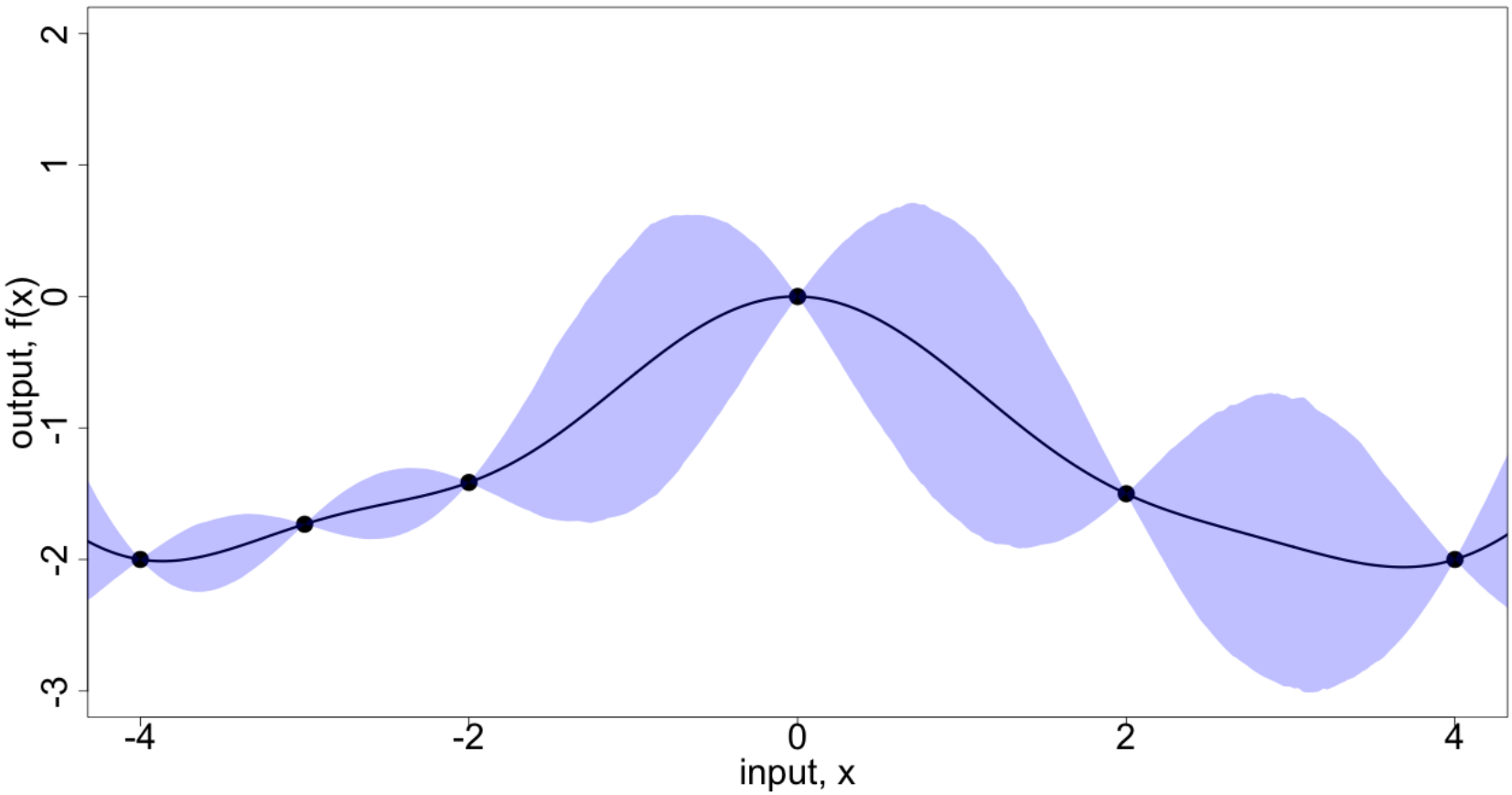
Data



GP Prior Distribution



Posterior Distribution



Avian Influenza Likelihood

The augmented likelihood function for this model is given by

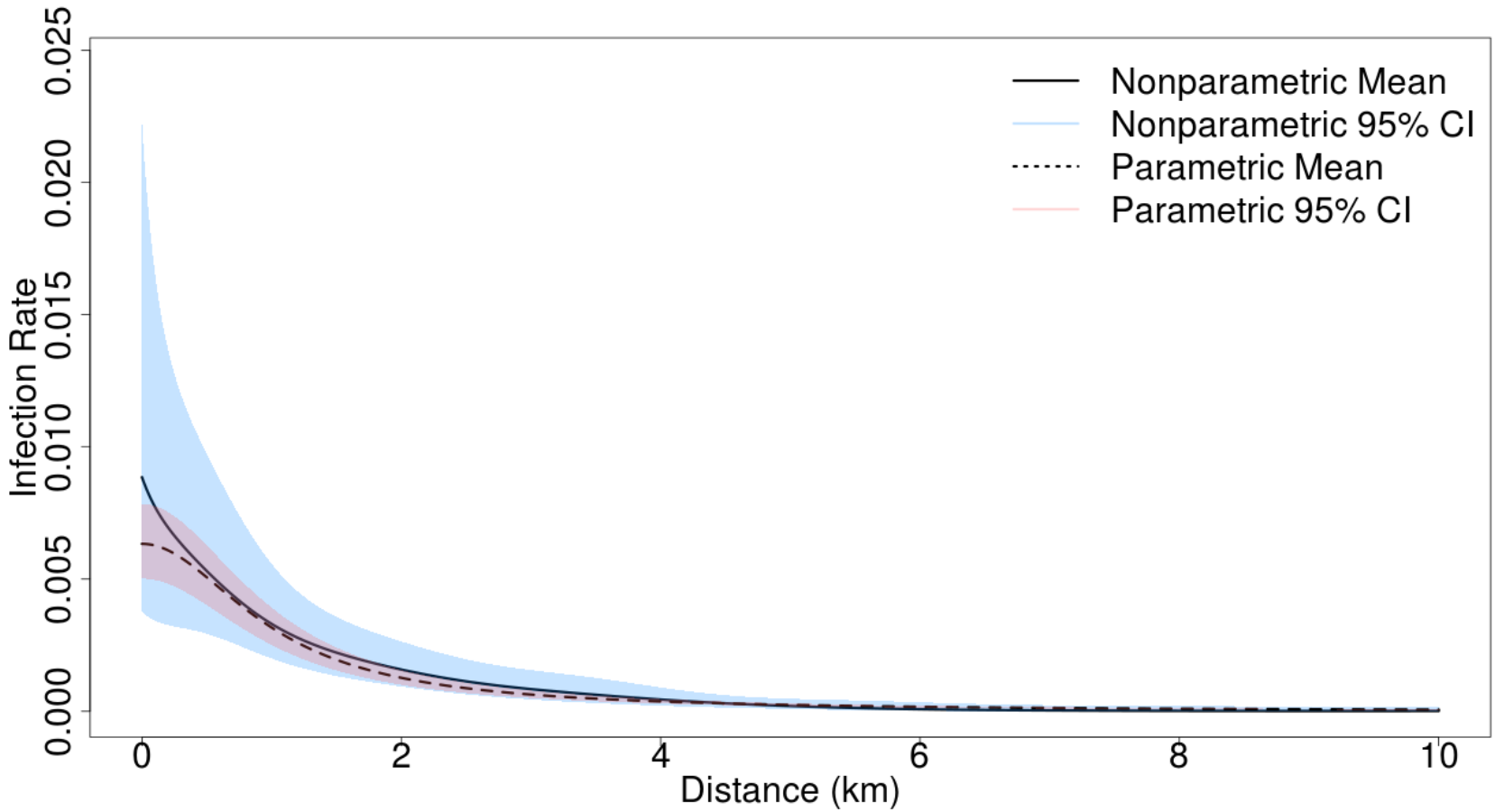
$$\pi(\mathbf{i}, \mathbf{r}|\boldsymbol{\beta}, \gamma) \propto \underbrace{\exp\Big(-\sum_{j=1}^{n}\sum_{k=1}^{N} \beta_{j,k}\big((r_j \wedge i_k) - (i_j \wedge i_k)\big)\Big)}_{\text{Total Infectious Pressure}}$$

$$\times \underbrace{\prod_{\substack{j=1 \\ j\neq\kappa}}^{n}\Big(\sum_{k\in\mathcal{Y}_j}\beta_{k,j}\Big)}_{\text{Pressure infectives put on each susceptible}}$$

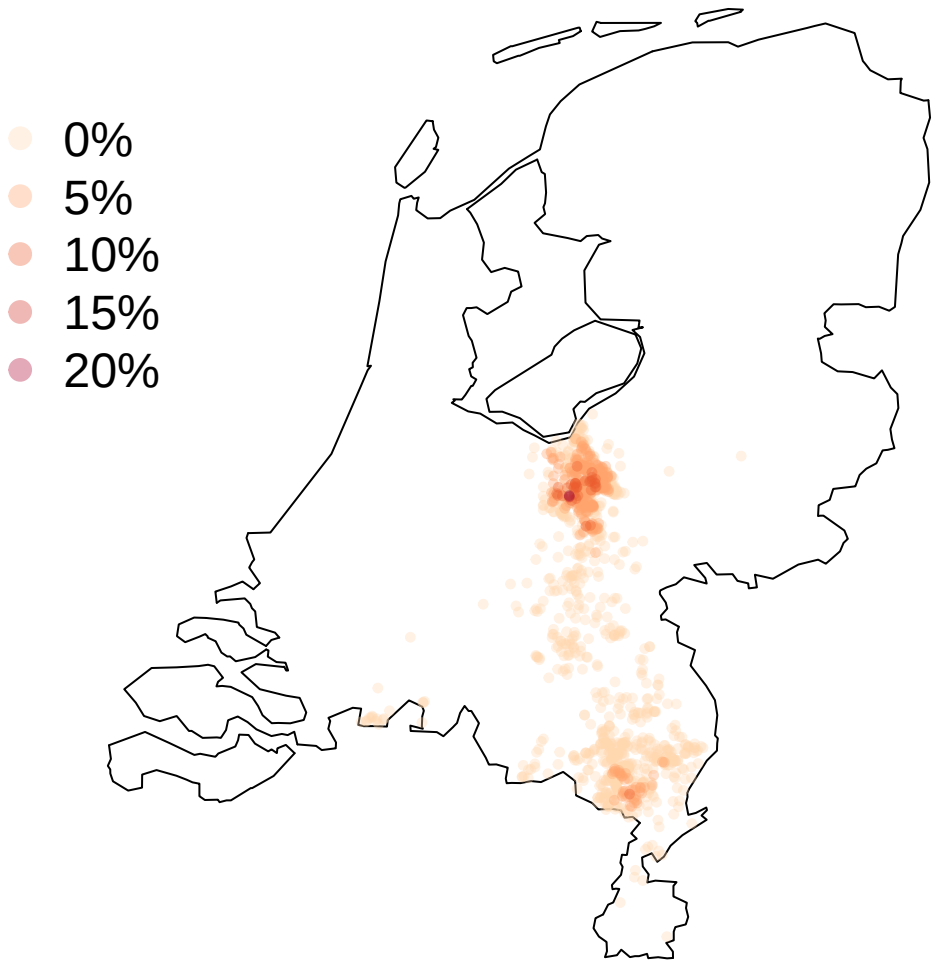
$$\times \underbrace{\prod_{j=1}^{n} f(r_j - i_j|\gamma)}_{\text{Infectious period distribution}} \quad .$$

Avian Influenza Results

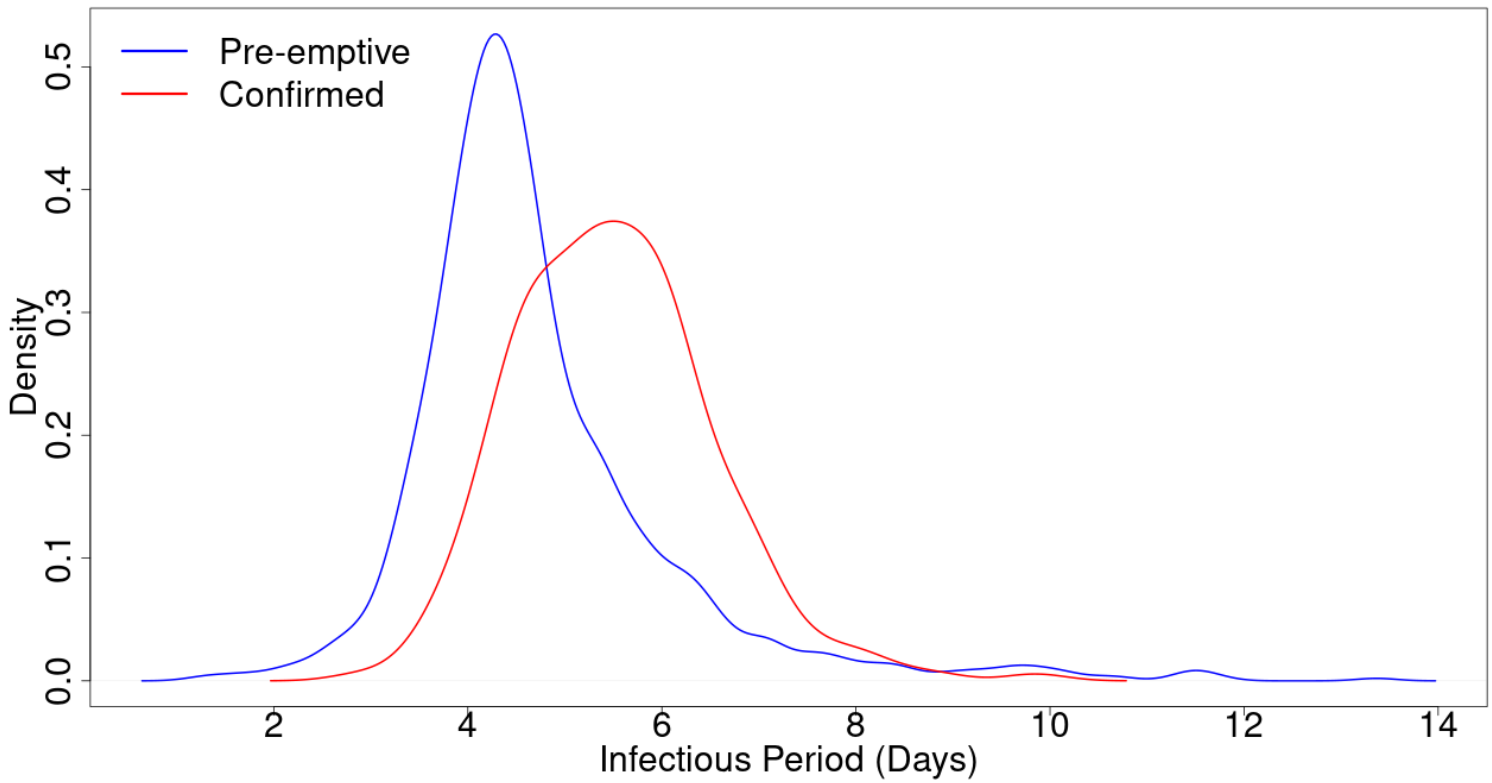
We then applied the GP inference method to the Avian Influenza data set and compared our model to the parametric model.



Avian Influenza Results



Avian Influenza Results



Culling Analysis

Table: Posterior predictive medians (95% probability intervals) for the number of infected and culled farms and the amount of compensation paid.

Radius (km)	Infected Farms	Culled Farms	Compensation (mil)
0	443 (151, 644)	443 (151, 644)	24.8 (8.62, 35.9)
1	297 (110, 535)	489 (215, 709)	27.2 (12.2, 38.9)
2	283 (108, 608)	488 (217, 740)	27.5 (12.2, 41.7)
3	283 (112, 582)	517 (242, 775)	29.0 (13.2, 43.1)
4	274 (105, 564)	512 (228, 793)	28.5 (12.3, 43.9)
5	280 (109, 549)	527 (226, 797)	39.2 (12.4, 41.9)

Table of Contents

Conclusion

Throughout this talk, we have shown that:

- We can model infection rates non-parametrically with GPs,
- We do not need to make restrictive assumptions about the parametric form of the infection rate,
- We can use this to analyse control strategies.

ORIGINAL ARTICLE | Open Access

A Bayesian nonparametric analysis of the 2003 outbreak of highly pathogenic avian influenza in the Netherlands

Rowland G. Seymour, Theodore Kypraios, Philip D. O'Neill, Thomas J. Hagenaars

First published: 10 August 2021 | https://doi.org/10.1111/rssc.12515

Gaussian Processes Overview

- A Gaussian Process is a collection of random variables, any finite number of which have a joint Gaussian Distribution. (Rasmussen and Williams, 2007).
- We describe the log of the infection rate fully through the GP, this ensure the infection rate is always positive.

$$\log \beta_{i,j} \sim \mathcal{GP}\big(\mathbf{0}, \Sigma\big).$$

- We use the squared exponential covariance function to define the covariance matrix.

$$\Sigma = k(\mathbf{d}, \mathbf{d}'; \alpha, l) = \alpha^2 \exp\Big(- \frac{(\mathbf{d} - \mathbf{d}')^2}{l^2}\Big).$$

GPs and Epidemic Models

We put a GP prior on $\beta_{i,j}$ and use under-relaxed MCMC to give estimate this rate.

1. Propose new infection rate based on current one.

$$\log \beta'_{i,j} = \sqrt{1+\delta^2}\log \beta_{i,j} + \delta\nu,$$

$$\nu \sim \mathcal{GP}\left(\mathbf{0}, \Sigma\right).$$

2. Evaluate likelihood of proposed infection rate.
3. Accept new rate based on probability of proposed against current rate being the true rate.

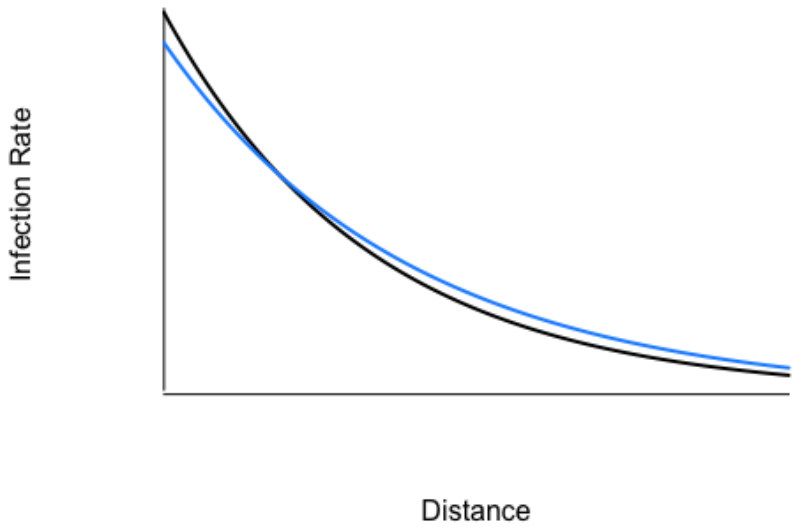


Figure: Current infection rate (black) and proposed rate (blue).