

Space-time varying Poisson auto-regressive model

An application to Covid-19 cases at the England local authority level

P. Alaimo Di Loro S. Sahu D. Bohning

StanConnect through Space and Time

October 31, 2022



Outline

1 Introduction

2 Poisson Auto-Regression

- The model
- The data and the heterogeneity
- Extension to space-time varying epidemics
- The CAR-AR prior

3 Estimation and results: STAN

- Why STAN?
- Efficient implementation of the CAR-AR prior
- Final application

4 Final remarks

Data-driven models for the COVID-19 pandemic

Data-driven models for the COVID-19 pandemic

- **Semi-parametric** models and **ensembles** (Hastie et al., 2009)
 - Extremely flexible and accurate (short-term) if well-tuned
 - Non-interpretable
 - Farcomeni et al. (2021); Sherratt et al. (2022)

Data-driven models for the COVID-19 pandemic

- **Semi-parametric** models and **ensembles** (Hastie et al., 2009)
 - Extremely flexible and accurate (short-term) if well-tuned
 - Non-interpretable
 - Farcomeni et al. (2021); Sherratt et al. (2022)
- **Growth** models (Tjørve and Tjørve, 2010)
 - Fairly flexible, interpretable, accurate (medium-term)
 - Need a minimal observation window and model single waves
 - Alaimo Di Loro et al. (2021); Girardi et al. (2021); Mingione et al. (2022)

Data-driven models for the COVID-19 pandemic

- **Semi-parametric** models and **ensembles** (Hastie et al., 2009)
 - Extremely flexible and accurate (short-term) if well-tuned
 - Non-interpretable
 - Farcomeni et al. (2021); Sherratt et al. (2022)
- **Growth** models (Tjørve and Tjørve, 2010)
 - Fairly flexible, interpretable, accurate (medium-term)
 - Need a minimal observation window and model single waves
 - Alaimo Di Loro et al. (2021); Girardi et al. (2021); Mingione et al. (2022)
- **Poisson Auto-regression** (Fokianos et al., 2009; Held and Paul, 2012)
 - Open to many different specifications and interpretable
 - The most complex versions may be difficult to estimate
 - Celani and Giudici (2022); Fritz et al. (2022); Giudici et al. (2022)

Modeling the COVID-19 pandemic

- **Semi-parametric models and ensembles** (Hastie et al., 2009)
 - Extremely flexible and accurate (short-term) if well-tuned
 - Non-interpretable
 - Farcomeni et al. (2021); Sherratt et al. (2022)
- **Growth** models (Chowell, 2017; Tjørve and Tjørve, 2010)
 - Fairly flexible, interpretable, accurate (medium-term)
 - Need a minimal observation window and model single waves
 - Alaimo Di Loro et al. (2021); Girardi et al. (2021); Mingione et al. (2022)
- **Poisson Auto-regression** (Fokianos et al., 2009)
 - Open to many different specifications and interpretable
 - The most complex versions may be difficult to estimate
 - Celani and Giudici (2022); Fritz et al. (2022); Giudici et al. (2022)

Standard Poisson auto-regressive specification

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_I regions
Poisson auto-regressive model of order 1:

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st})$$

$$\lambda_{st} = y_{s(t-1)} \cdot \tilde{r}_s + b_s$$

Standard Poisson auto-regressive specification

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_I regions
Poisson auto-regressive model of order 1:

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st})$$

$$\lambda_{st} = y_{s(t-1)} \cdot \tilde{r}_s + b_s$$

- \tilde{r}_s are region-specific **auto-regressive coefficients**
It determines the **epidemic growth rate** r_s

Standard Poisson auto-regressive specification

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_I regions
Poisson auto-regressive model of order 1:

$$\begin{aligned}Y_{st} \mid \mathbf{y}_{s(1:t-1)} &\sim Poi(\lambda_{st}) \\ \lambda_{st} &= y_{s(t-1)} \cdot \tilde{r}_s + b_s\end{aligned}$$

- \tilde{r}_s are region-specific **auto-regressive coefficients**
It determines the **epidemic growth rate** r_s
- b_s are location specific **baseline rates**
It is a long-term **endemic rate** and accounts for **hidden infections**

Standard Poisson auto-regressive specification

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_I regions
Poisson auto-regressive model of order 1:

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st})$$
$$\lambda_{st} = y_{s(t-1)} \cdot \tilde{r}_s + b_s$$

- \tilde{r}_s are region-specific **auto-regressive coefficients**
It determines the **epidemic growth rate** r_s
- b_s are location specific **baseline rates**
It is a long-term **endemic rate** and accounts for **hidden infections**

$$\tilde{r}_s \neq r_s = \frac{\lambda_{st}}{\lambda_{s(t-1)}}$$

Standard Poisson auto-regressive specification

In the **standard framework**

- \tilde{r}_s is **constant** in time
- $\tilde{r}_s < 1$ to ensure **stationarity**
- \tilde{r}_s 's **independently estimated** across regions

Standard Poisson auto-regressive specification

In the **standard framework**

- \tilde{r}_s is **constant** in time
- $\tilde{r}_s < 1$ to ensure **stationarity**
- \tilde{r}_s 's **independently estimated** across regions

There are many extensions (INGARCH, log-INGARCH, INAR), but...

Standard Poisson auto-regressive specification

In the **standard framework**

- \tilde{r}_s is **constant** in time
- $\tilde{r}_s < 1$ to ensure **stationarity**
- \tilde{r}_s 's **independently estimated** across regions

There are many extensions (INGARCH, log-INGARCH, INAR), but...
... few go in the following directions:

Standard Poisson auto-regressive specification

In the **standard framework**

- \tilde{r}_s is **constant** in time
- $\tilde{r}_s < 1$ to ensure **stationarity**
- \tilde{r}_s 's **independently estimated** across regions

There are many extensions (INGARCH, log-INGARCH, INAR), but...

... few go in the following directions:

- $\tilde{r}_s \neq \tilde{r}_{st}$, may **vary over time** because of **observed or unobserved** factors

Standard Poisson auto-regressive specification

In the **standard framework**

- \tilde{r}_s is **constant** in time
- $\tilde{r}_s < 1$ to ensure **stationarity**
- \tilde{r}_s 's **independently estimated** across regions

There are many extensions (INGARCH, log-INGARCH, INAR), but...

... few go in the following directions:

- $\tilde{r}_s \neq \tilde{r}_{st}$, may **vary over time** because of **observed or unobserved** factors
- Epidemics **are not always stationary**: $\tilde{r}_{st} < 1$ and $\tilde{r}_{st*} > 1$

Standard Poisson auto-regressive specification

In the **standard framework**

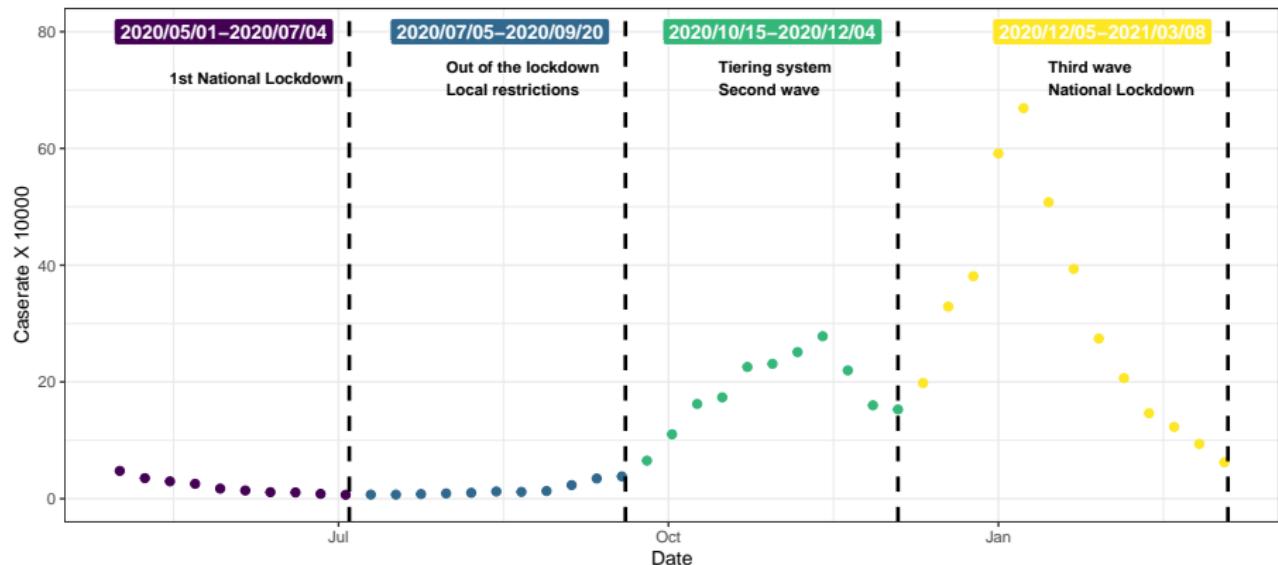
- \tilde{r}_s is **constant** in time
- $\tilde{r}_s < 1$ to ensure **stationarity**
- \tilde{r}_s 's **independently estimated** across regions

There are many extensions (INGARCH, log-INGARCH, INAR), but...

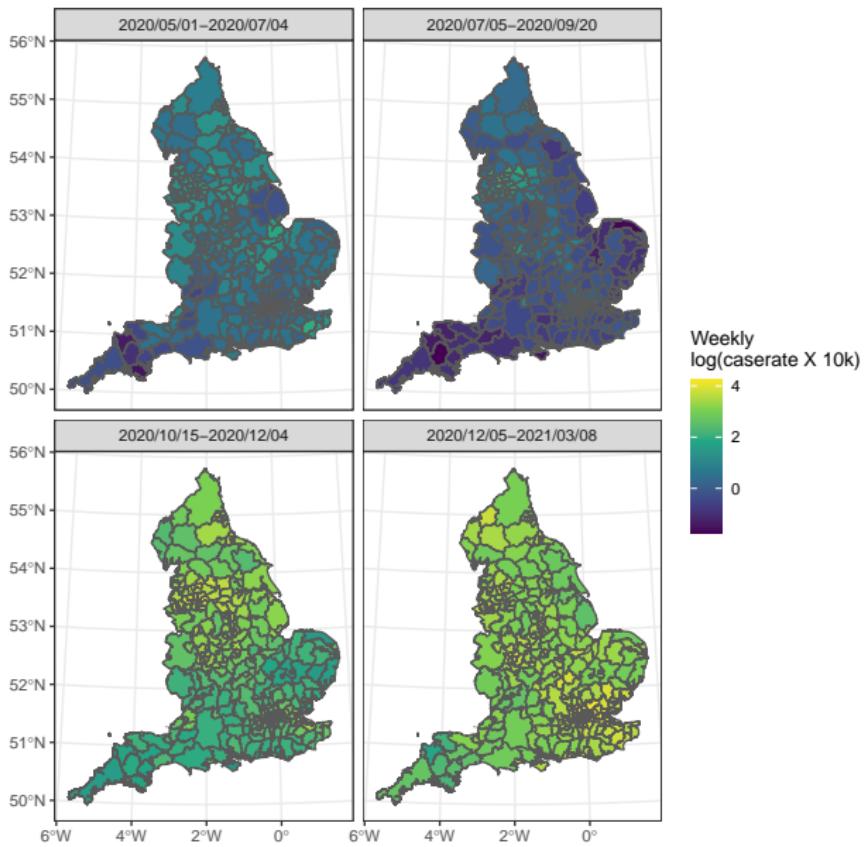
... few go in the following directions:

- $\tilde{r}_s \neq \tilde{r}_{st}$, may **vary over time** because of **observed or unobserved** factors
- Epidemics **are not always stationary**: $\tilde{r}_{st} < 1$ and $\tilde{r}_{st*} > 1$
- When **regions** are small we want to **borrow information**

England COVID-19 cases May 2021 - Mar 2022



England COVID-19 cases May 2021 - Mar 2022



Extending Poisson auto-regression for epidemics

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_I regions

$$\begin{aligned} Y_{st} | \mathbf{y}_{s(1:t-1)} &\sim Poi(\lambda_{st}) \\ \lambda_{st} &= y_{s(t-1)} \cdot \tilde{r}_s + b_s \end{aligned}$$

Extending Poisson auto-regression for epidemics

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_I regions

$$\begin{aligned} Y_{st} \mid \mathbf{y}_{s(1:t-1)} &\sim Poi(\lambda_{st}) \\ \lambda_{st} &= y_{s(t-1)} \cdot \tilde{\mathbf{r}}_{st} + \mathbf{b}_{st} \end{aligned}$$

- Time-varying location-specific rates $\tilde{\mathbf{r}}_{st}, \mathbf{b}_{st}$

Extending Poisson auto-regression for epidemics

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_l regions

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st})$$
$$\lambda_{st} = \left(\sum_{j=1}^{\tau} \mathbf{w}_j \cdot \mathbf{y}_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st}$$

- Time-varying location-specific rates \tilde{r}_{st}, b_{st}
- Include **infection latency** up to some lag τ

Extending Poisson auto-regression for epidemics

Let $\mathbf{y}_1, \dots, \mathbf{y}_T$ be series of **detected cases** over s_1, \dots, s_l regions

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st})$$

$$\lambda_{st} = \left(\left(\sum_{j=1}^{\tau} \mathbf{w}_j \cdot \mathbf{y}_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

- Time-varying location-specific rates \tilde{r}_{st}, b_{st}
- Include infection latency up to some lag τ
- Non-exploding condition through a **non-reinfection window**:

$$d_{st} = 1 - \frac{\sum_{i=1}^{\tilde{t}} y_{s(t-i)}}{\text{pop}_s}$$

Modeling the coefficients

Model **observed (covariates X, V)** and **unobserved (random effects)** heterogeneity

Modeling the coefficients

Model **observed (covariates X, V)** and **unobserved (random effects)** heterogeneity

- **generalized linear mixed model**
- **log-link**

Modeling the coefficients

Model **observed (covariates X, V)** and **unobserved (random effects)** heterogeneity

- generalized linear mixed model
- log-link

Auto-regressive coefficient

$$\log(\tilde{r}_{st}) = \ln_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}$$

Modeling the coefficients

Model **observed (covariates X, V)** and **unobserved (random effects)** heterogeneity

- generalized linear mixed model
- log-link

Auto-regressive coefficient

$$\log(\tilde{r}_{st}) = \text{Int}_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}$$

Baseline rate

$$\log(b_{st}) = \text{Offset}_s + \mathbf{V}\boldsymbol{\eta} + \psi_{st}$$

CAR-AR on Φ, Ψ (Rushworth et al., 2014)

Independence across random effects $\Phi \perp \Psi$

Independence across random effects $\Phi \perp \Psi$

Spatial dependence \leftrightarrow **Leroux model**

$$\phi_t^*, \psi_t^* \sim \mathcal{N} \left(\mathbf{0}, \sigma^2 \mathbf{Q}(\alpha, \mathbf{W})^{-1} \right)$$

with $\mathbf{Q}(\alpha, \mathbf{W}) = (\alpha(\mathbf{D} - \mathbf{W}) + (1 - \alpha)\mathbf{I}_I)$

CAR-AR on Φ, Ψ (Rushworth et al., 2014)

Independence across random effects $\Phi \perp \Psi$

Spatial dependence \leftrightarrow **Leroux model**

$$\phi_t^*, \psi_t^* \sim \mathcal{N} \left(\mathbf{0}, \sigma^2 \mathbf{Q}(\alpha, \mathbf{W})^{-1} \right)$$

with $\mathbf{Q}(\alpha, \mathbf{W}) = (\alpha(\mathbf{D} - \mathbf{W}) + (1 - \alpha)\mathbf{I}_I)$

Temporal dependence \leftrightarrow **AR(1) structure**

$$\phi_1, \psi_1 \sim \mathcal{N} \left(\mathbf{0}, \sigma^2 \mathbf{Q}(\alpha, \mathbf{W})^{-1} \right)$$

$$\phi_t, \psi_t \sim \mathcal{N} \left(\rho \cdot \phi_{t-1}, \sigma^2 \mathbf{Q}(\alpha, \mathbf{W})^{-1} \right), \quad t = 2, \dots, T$$

Model estimation

Full model specification

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st}), \quad \lambda_{st} = \left(\left(\sum_{j=1}^{\tau} w_j \cdot y_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

$$\log(\tilde{r}_{st}) = Ir_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}, \quad \log(b_{st}) = IOff_s + \mathbf{V}\boldsymbol{\eta} + \psi_{st}$$

$$\phi_t, \psi_t \sim \mathcal{N} \left(\rho \cdot \phi_{t-1}, \sigma^2 \mathbf{Q}(\alpha, \mathbf{W})^{-1} \right), \quad t = 2, \dots, T$$

$$\boldsymbol{\theta} \stackrel{\text{ind}}{\sim} \pi_{\boldsymbol{\theta}}(\cdot), \quad \boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\eta}, \sigma, \alpha, \rho)$$

Model estimation

Full model specification

$$Y_{st} \mid \mathbf{y}_{s(1:t-1)} \sim Poi(\lambda_{st}), \quad \lambda_{st} = \left(\left(\sum_{j=1}^{\tau} w_j \cdot y_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

$$\log(\tilde{r}_{st}) = Ir_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}, \quad \log(b_{st}) = IOff_s + \mathbf{V}\boldsymbol{\eta} + \psi_{st}$$

$$\phi_t, \psi_t \sim \mathcal{N} \left(\rho \cdot \phi_{t-1}, \sigma^2 \mathbf{Q}(\alpha, \mathbf{W})^{-1} \right), \quad t = 2, \dots, T$$

$$\boldsymbol{\theta} \stackrel{\text{ind}}{\sim} \pi_{\boldsymbol{\theta}}(\cdot), \quad \boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\eta}, \sigma, \alpha, \rho)$$

Estimation is carried out in a **Bayesian framework** through the STAN software (Carpenter et al., 2017)

Why STAN?

The real question is... **why not?**

Why STAN?

The real question is... **why not?** STAN is like your **best friend**

Why STAN?

The real question is... **why not?** STAN is like your **best friend**

- It is willing to fit **whatever you ask** him
 - ① Accepts **user-specified likelihoods**
 - ② Extremely **flexible** programming

Why STAN?

The real question is... **why not?** STAN is like your **best friend**

- It is willing to fit **whatever you ask** him
 - ① Accepts **user-specified likelihoods**
 - ② Extremely **flexible** programming
- Doesn't ask much **in exchange**
 - ① Implements the **NUTS** that adaptively sets sampling paths
 - ② **Prior** choice is driven only by **prior information or regularization**

Why STAN?

The real question is... **why not?** STAN is like your **best friend**

- It is willing to fit **whatever you ask** him
 - ① Accepts **user-specified likelihoods**
 - ② Extremely **flexible** programming
- Doesn't ask much **in exchange**
 - ① Implements the **NUTS** that adaptively sets sampling paths
 - ② **Prior** choice is driven only by **prior information or regularization**
- It is **honest** and **sincere**
 - ① It **efficiently explore** the full posterior (even when complex)
 - ② Provides **incredibly useful diagnostic tools**

Why STAN?

The real question is... **why not?** STAN is like your **best friend**

- It is willing to fit **whatever you ask** him
 - 1 Accepts **user-specified likelihoods**
 - 2 Extremely **flexible** programming
- Doesn't ask much **in exchange**
 - 1 Implements the **NUTS** that adaptively sets sampling paths
 - 2 **Prior** choice is driven only by **prior information or regularization**
- It is **honest** and **sincere**
 - 1 It **efficiently explore** the full posterior (even when complex)
 - 2 Provides **incredibly useful diagnostic tools**
- It is **up-to-date** and open to **improvement**
 - 1 It is **continuously updated** by an active team
 - 2 It has a very **responsive** community

Why STAN?

Estimation through STAN

- Requires way **less MCMC iterations** than the competitor algorithms
- Between **4000 — 12000** are usually **enough** even for complex models

Why STAN?

Estimation through STAN

- Requires way **less MCMC iterations** than the competitor algorithms
- Between **4000 — 12000** are usually **enough** even for complex models

Nevertheless, the **iteration time** scales with

- The computational complexity of the **posterior evaluation**
- The **geometry** of the posterior

Why STAN?

Estimation through STAN

- Requires way **less MCMC iterations** than the competitor algorithms
- Between **4000 — 12000** are usually **enough** even for complex models

Nevertheless, the **iteration time** scales with

- The computational complexity of the **posterior evaluation**
- The **geometry** of the posterior

How can we facilitate STAN job for the proposed model?

Sparse Leroux

The Leroux model requires the **determinant** and a **quadratic form on**:

$$\mathbf{Q}(\alpha, \mathbf{W}) = (\alpha(\mathbf{D} - \mathbf{W}) + (1 - \alpha)\mathbf{I}_l)$$

- These two operations scale with l^3 and l^2
- When l increases they become **rapidly costly**

Sparse Leroux

The Leroux model requires the **determinant** and a **quadratic form on**:

$$\mathbf{Q}(\alpha, \mathbf{W}) = (\alpha(\mathbf{D} - \mathbf{W}) + (1 - \alpha)\mathbf{I}_l)$$

- These two operations scale with l^3 and l^2
- When l increases they become **rapidly costly**

We can exploit the **sparsity pattern** in $\mathbf{Q}(\alpha, \mathbf{W})$

- Sparse-representation $\rightarrow \phi_t^\top \mathbf{Q}(\alpha, \mathbf{W}) \phi_t$ only non-zero entries

Sparse Leroux

The Leroux model requires the **determinant** and a **quadratic form on**:

$$\mathbf{Q}(\alpha, \mathbf{W}) = (\alpha(\mathbf{D} - \mathbf{W}) + (1 - \alpha)\mathbf{I}_l)$$

- These two operations scale with l^3 and l^2
- When l increases they become **rapidly costly**

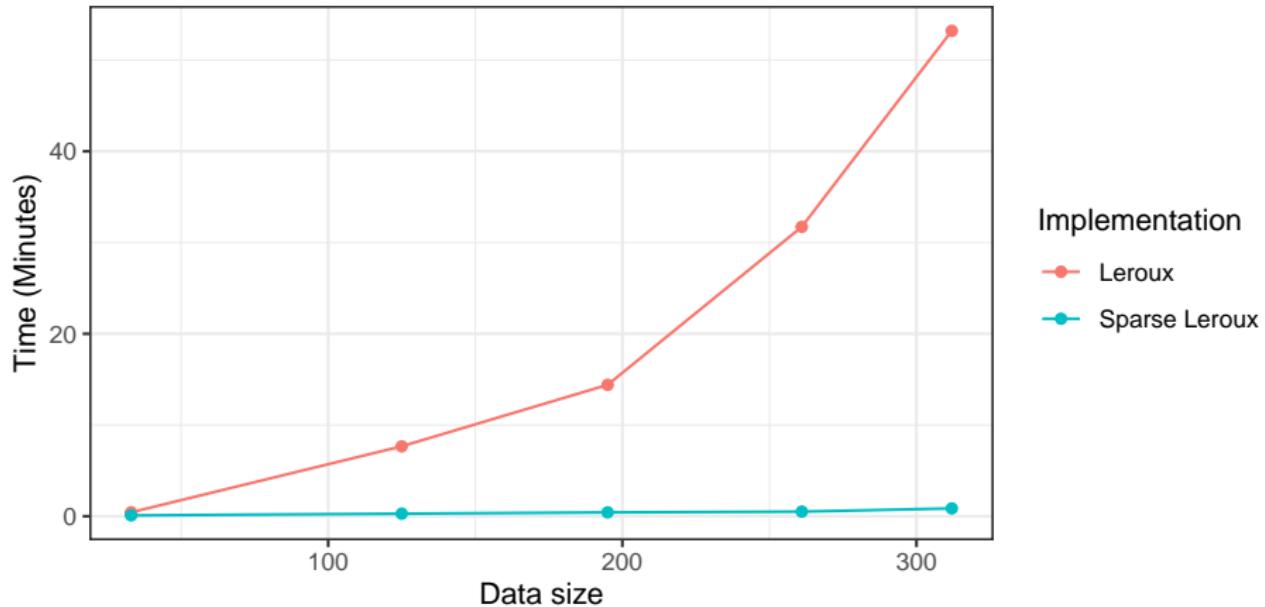
We can exploit the **sparsity pattern** in $\mathbf{Q}(\alpha, \mathbf{W})$

- Sparse-representation $\rightarrow \phi_t^\top \mathbf{Q}(\alpha, \mathbf{W}) \phi_t$ only non-zero entries
- Drawing from Jin et al. (2005) in the **proper CAR**:

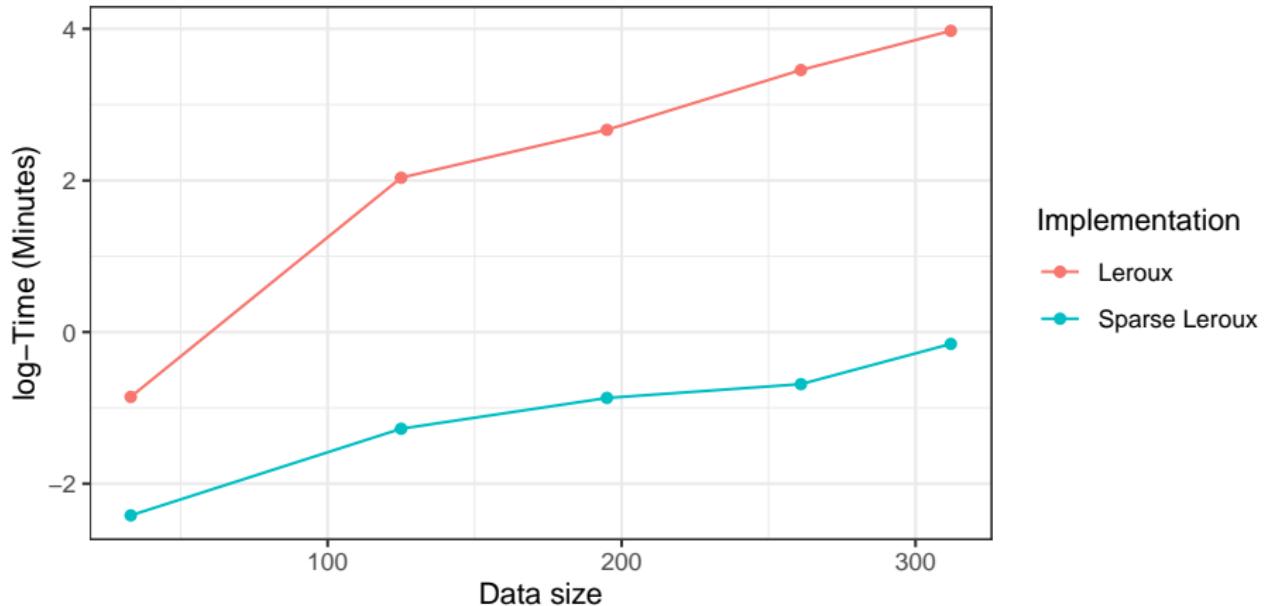
$$\begin{aligned}\mathbf{Q}(\alpha, \mathbf{W}) &= (\mathbf{I}_l - \alpha(\mathbf{I}_l - \mathbf{D} + \mathbf{W})) \rightarrow \\ &\rightarrow |\mathbf{Q}(\alpha, \mathbf{W})| \propto \prod_{i=1}^l (1 - \alpha\lambda_i)\end{aligned}$$

where λ_i are the eigen-values of $(\mathbf{I}_l - \mathbf{D} + \mathbf{W})$

Sparse Leroux - 1000 iterations



Sparse Leroux - 1000 iterations



Non-centered parametrization

- The **NUTS** samples and adapts its tuning parameters in the **unconstrained parameters space**
- Sampling is the more efficient the more the posterior in the unconstrained space is **uncorrelated/approximately Gaussian**

Non-centered parametrization

- The **NUTS** samples and adapts its tuning parameters in the **unconstrained parameters space**
- Sampling is the more efficient the more the posterior in the unconstrained space is **uncorrelated/approximately Gaussian**

The **non-centered** parametrization helps (often a lot)

Model block

$$\phi_t^*, \psi_t^* \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(\alpha, \mathbf{W})^{-1}), \quad t = 1, \dots, T$$

Transformed parameters block

$$\phi_1, \psi_1 = \phi_1^*$$

$$\phi_t, \psi_t = \rho \cdot \phi_{t-1} + \phi_t^*, \quad t = 2, \dots, T$$

Prior and hyper-prior settings

$$\lambda_{st} = \left(\left(\sum_{j=1}^{\tau} w_j \cdot y_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

$$\log(\tilde{r}_{st}) = l r_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}, \quad \log(b_{st}) = l Off_s + \mathbf{V}\boldsymbol{\eta} + \psi_{st}$$

Prior and hyper-prior settings

$$\lambda_{st} = \left(\left(\sum_{j=1}^{\tau} w_j \cdot y_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

$$\log(\tilde{r}_{st}) = Ir_0 + \mathbf{X}\beta + \phi_{st}, \quad \log(b_{st}) = IOff_s + \mathbf{V}\eta + \psi_{st}$$

Priors

$$Ir_0 \sim \mathcal{N}(-0.5, 1), \quad \beta \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_k), \quad \eta \sim \mathcal{N}_\nu(\mathbf{0}, \mathbf{I}_\nu),$$

$$\alpha_\phi, \alpha_\psi \sim Beta(2, 1), \quad \rho_\phi, \rho_\psi \sim Unif(0, 1),$$

$$\sigma_\phi \sim \mathcal{N}^+(0, 0.1), \quad \sigma_\psi \sim \mathcal{N}^+(0, 0.5),$$

$$\mathbf{w} \sim Dir(\mathbf{1}_\tau)$$

The model and data

$$\lambda_{st} = \left(\left(\sum_{j=1}^{\tau} w_j \cdot y_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

$$\log(\tilde{r}_{st}) = Ir_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}, \quad \log(b_{st}) = IOff_s + \mathbf{V}\boldsymbol{\eta} + \psi_{st}$$

Fixed components

- No reinfections
- Latency up to $\tau = 3$

The model and data

$$\lambda_{st} = \left(\left(\sum_{j=1}^{\tau} w_j \cdot y_{s(t-j)} \right) \cdot \tilde{r}_{st} + b_{st} \right) \cdot d_{st}$$

$$\log(\tilde{r}_{st}) = Ir_0 + \mathbf{X}\boldsymbol{\beta} + \phi_{st}, \quad \log(b_{st}) = IOff_s + \mathbf{V}\boldsymbol{\eta} + \psi_{st}$$

Fixed components

- No reinfections
- Latency up to $\tau = 3$

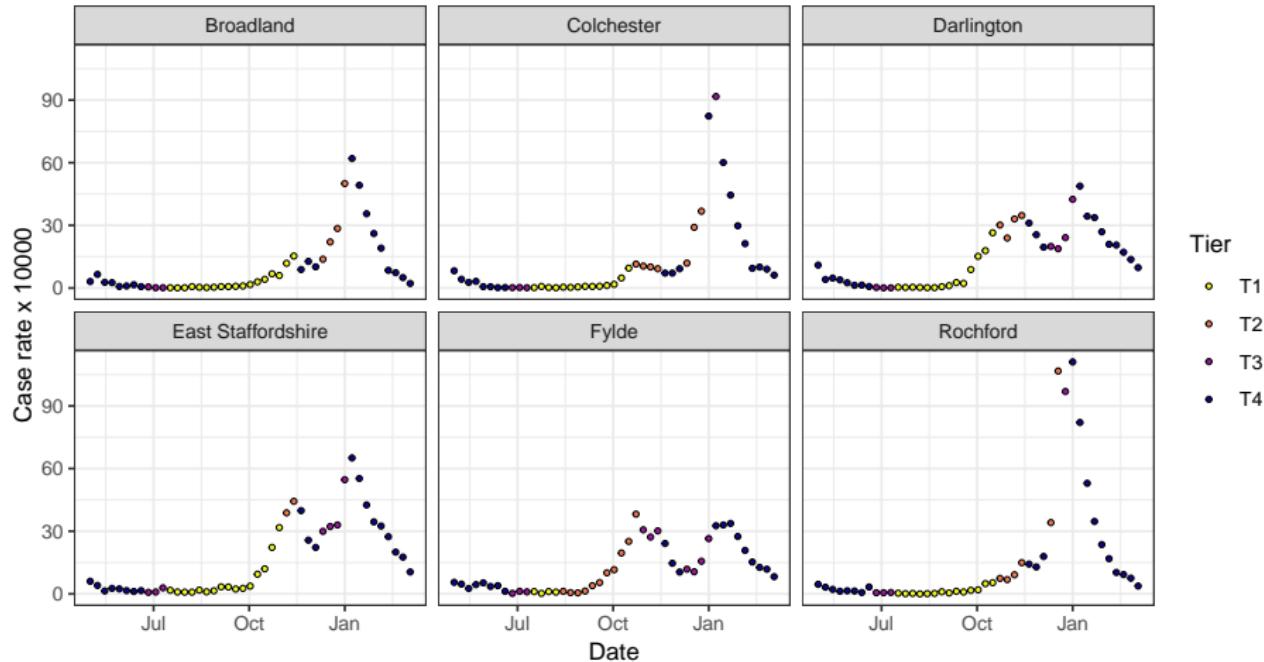
Covariates

\mathbf{X} Population density, jsa, median house price
Tier level (II, III, IV)

\mathbf{V} Only intercept

$IOff_s$ Log-population size divided by 10000

The model and data



Application on the real data

Full dataset

- $I = 313$ districts of England
- 8th of May 2020 to the 8th of March 2021, $T = 43$
- $N = 13'459$ records

Application on the real data

Full dataset

- $I = 313$ districts of England
- 8th of May 2020 to the 8th of March 2021, $T = 43$
- $N = 13'459$ records

Estimation

- 4 parallel chains
- 8000 iterations each
- Thinning of 2 to reduce memory burden
- 80% – 20% train-test split

Application on the real data

Full dataset

- $I = 313$ districts of England
- 8th of May 2020 to the 8th of March 2021, $T = 43$
- $N = 13'459$ records

Estimation

- 4 parallel chains
- 8000 iterations each
- Thinning of 2 to reduce memory burden
- 80% – 20% train-test split

Runtime of ≈ 40 hours

Application - Parameter estimates

Baselines	$\hat{\theta}$	Cl_{95}	$Rhat$
lr_0	-0.55	(-0.64, -0.45)	1.01
η_0	0.43	(0.35, 0.53)	1.00

Coefficients	$\hat{\beta}$	Cl_{95}	$Rhat$
Log-pop density	0.06	(0.05, 0.08)	1.00
Log-jsa	0.18	(0.01, 0.03)	1.00
Log-hprice	-0.06	(-0.08, -0.04)	1.00
Tier II	-0.08	(-0.14, -0.03)	1.00
Tier III	-0.23	(-0.30, -0.16)	1.00
Tier IV	-0.24	(-0.32, -0.16)	1.01

Application - Parameter estimates

CAR-AR	ϕ			ψ		
	$\hat{\theta}$	Cl_{95}	$Rhat$	$\hat{\theta}$	Cl_{95}	$Rhat$
α	0.998	(0.997, 0.999)	1.01	0.98	(0.97, 0.99)	1.00
ρ	0.64	(0.58, 0.70)	1.01	0.57	(0.54, 0.61)	1.01
σ	0.18	(0.17, 0.20)	1.01	0.98	(0.92, 1.05)	1.00

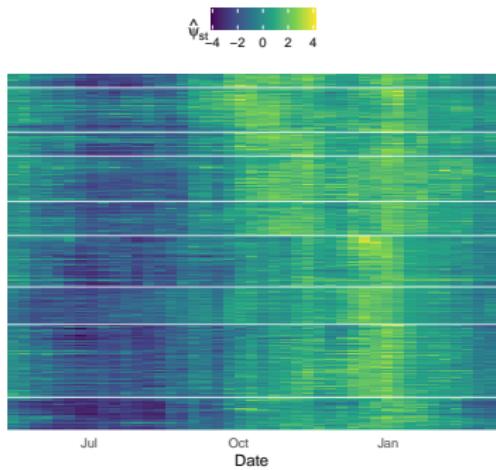
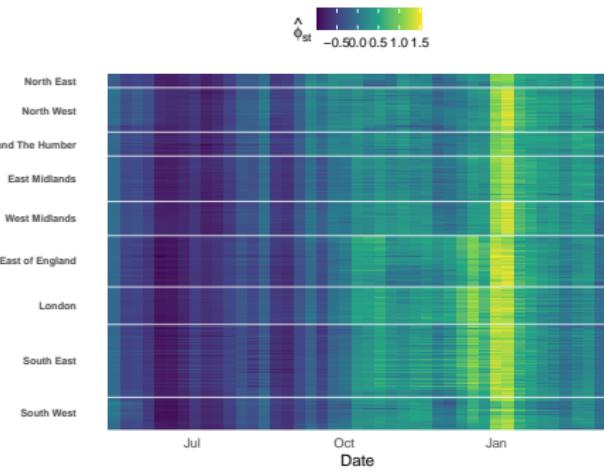
Application - Parameter estimates

CAR-AR	ϕ			ψ		
	$\hat{\theta}$	Cl_{95}	$Rhat$	$\hat{\theta}$	Cl_{95}	$Rhat$
α	0.998	(0.997, 0.999)	1.01	0.98	(0.97, 0.99)	1.00
ρ	0.64	(0.58, 0.70)	1.01	0.57	(0.54, 0.61)	1.01
σ	0.18	(0.17, 0.20)	1.01	0.98	(0.92, 1.05)	1.00

Weights	$\hat{\theta}$	Cl_{95}	$Rhat$
w_1	0.70	(0.67, 0.69)	1.00
w_2	0.19	(0.16, 0.18)	1.00
w_3	0.11	(0.08, 0.10)	1.00

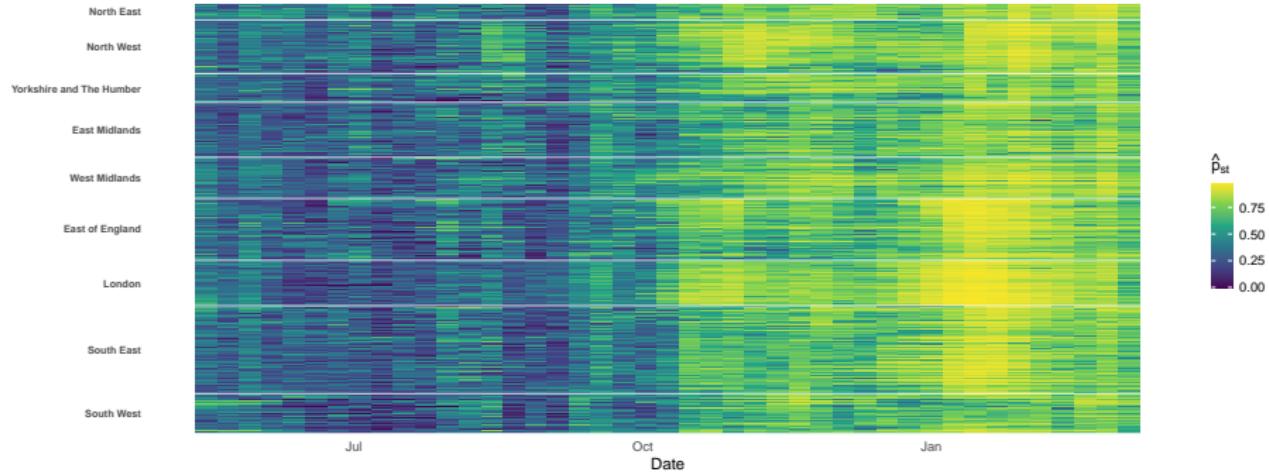
Application - Random effects

LADCUA



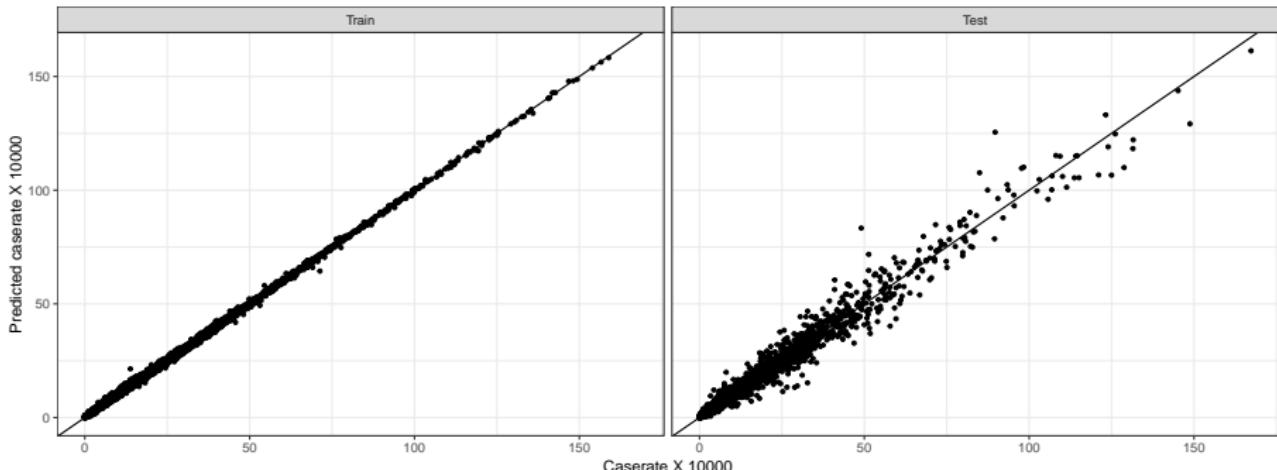
Application - Autoregressive and baseline components

$$\hat{p}_{st} = \frac{\sum_{i=1}^{\tau} (w_i \cdot y_{s,t-i}) \cdot \hat{r}_{s,t}}{\sum_{i=1}^{\tau} (w_i \cdot y_{s,t-i}) \cdot \hat{r}_{s,t} + \hat{b}_{st}}$$

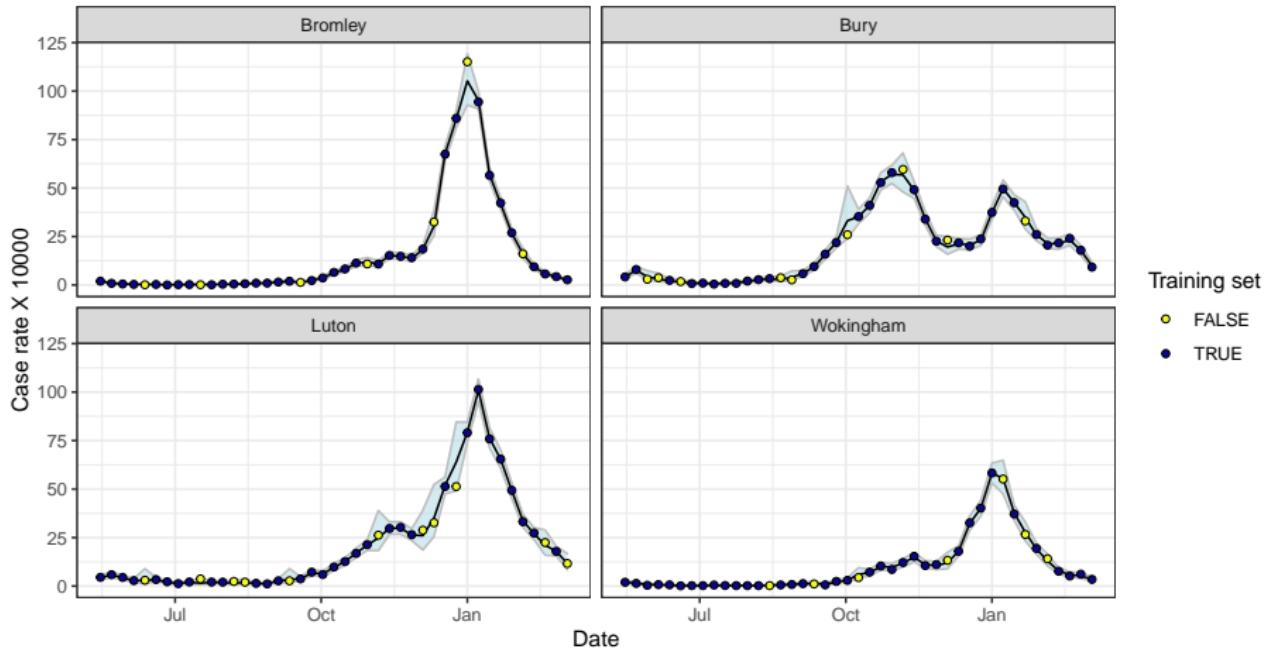


Application - Predictive performances

Metric	In-sample	Out-Of-sample
RMSE	≈ 0	0.02
\bar{Cov}_{95}	0.99	0.94
\bar{IW}_{95}	64.3	159



Application - Predictive performances



Concluding remarks

Conclusion

- ✓ Extension of the Poisson auto-regression to suite space-time varying epidemic processes
- ✓ Efficient implementation of the model in STAN
- ✓ Estimation of the efficacy of the tiering policies and description of the space-time heterogeneity

Concluding remarks

Conclusion

- ✓ Extension of the Poisson auto-regression to suite space-time varying epidemic processes
- ✓ Efficient implementation of the model in STAN
- ✓ Estimation of the efficacy of the tiering policies and description of the space-time heterogeneity

Work in progress

- ✗ Estimate (rather than fix) the latency τ
- ✗ Include dependence across the two sets of random effects Φ and Ψ
- ✗ Implement some variable-selection scheme

Main references

- P. Alaimo Di Loro, F. Divino, A. Farcomeni, G. Jona Lasinio, G. Lovison, A. Maruotti, and M. Mingione. Nowcasting covid-19 incidence indicators during the italian first outbreak. *Statistics in Medicine*, 40(16):3843–3864, 2021.
- B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell. Stan: a probabilistic programming language. *Grantee Submission*, 76(1):1–32, 2017.
- A. Celani and P. Giudici. Endemic–epidemic models to understand covid-19 spatio-temporal evolution. *Spatial Statistics*, 49:100528, 2022.
- G. Chowell. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infectious Disease Modelling*, 2(3):379–398, 2017.
- A. Farcomeni, A. Maruotti, F. Divino, G. Jona-Lasinio, and G. Lovison. An ensemble approach to short-term forecast of COVID-19 intensive care occupancy in italian regions. *Biometrical Journal*, 63:503–513, 2021.
- K. Fokianos, A. Rahbek, and D. Tjøstheim. Poisson autoregression. *Journal of the American Statistical Association*, 104(488):1430–1439, 2009.
- C. Fritz, G. D. Nicola, M. Rave, M. Weigert, Y. Khazaei, U. Berger, H. Küchenhoff, and G. Kauermann. Statistical modelling of covid-19 data: Putting generalized additive models to work. *Statistical Modelling*, 0(0):1471082X221124628, 2022. doi: 10.1177/1471082X221124628. URL <https://doi.org/10.1177/1471082X221124628>.
- P. Girardi, L. Greco, and L. Ventura. Misspecified modeling of subsequent waves during covid-19 outbreak: A change-point growth model. *Biometrical Journal*, 63:In press, 2021.
- P. Giudici, B. Tarantino, and A. Roy. Bayesian time-varying autoregressive models of covid-19 epidemics. *Biometrical Journal*, n/a(n/a), 2022. doi: <https://doi.org/10.1002/bimj.202200054>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/bimj.202200054>.
- T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.
- L. Held and M. Paul. Modeling seasonality in space-time infectious disease surveillance data. *Biometrical Journal*, 54(6):824–843, 2012.
- X. Jin, B. P. Carlin, and S. Banerjee. Generalized hierarchical multivariate car models for areal data. *Biometrics*, 61(4):950–961, 2005.
- M. Mingione, P. Alaimo Di Loro, A. Farcomeni, F. Divino, G. Lovison, A. Maruotti, and G. J. Lasinio. Spatio-temporal modelling of covid-19 incident cases using richards' curve: An application to the italian regions. *Spatial Statistics*, 49:100544, 2022.
- A. Rushworth, D. Lee, and R. Mitchell. A spatio-temporal model for estimating the long-term effects of air pollution on respiratory hospital admissions in Greater London. *Spatial and spatio-temporal epidemiology*, 10:29–38, 2014.
- K. Sherratt, H. Gruson, H. Johnson, R. Niehus, B. Prasse, F. Sandman, J. Deusel, D. Wolffram, S. Abbott, A. Ullrich, et al. Predictive performance of multi-model ensemble forecasts of covid-19 across european nations. *medRxiv*, 2022.
- E. Tjørve and K. M. Tjørve. A unified approach to the richards-model family for use in growth analyses: why we need only two model forms. *Journal of theoretical biology*, 267(3):417–425, 2010.

THANK YOU FOR YOUR ATTENTION!



THANK YOU FOR YOUR ATTENTION!



Questions?

APPENDIX

Simulations - Prior predictive check

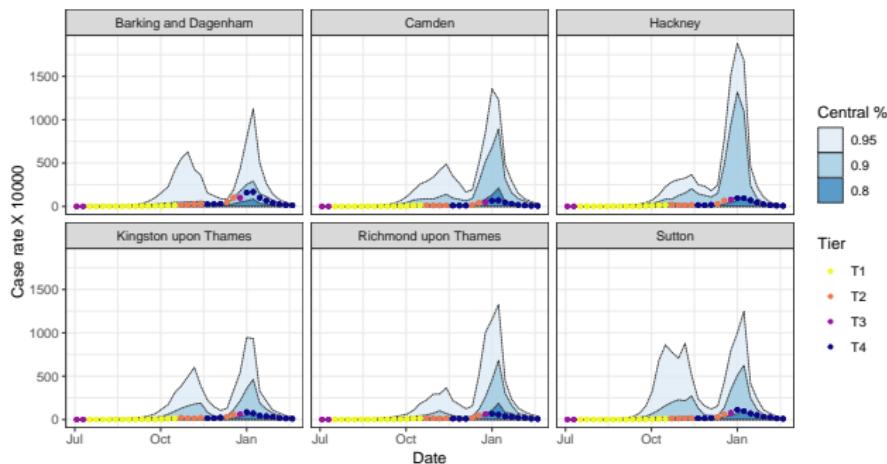
Restricted dataset

- 33 districts within the region of *London*
- 3rd of July 2020 to the 21st of February 2021
- $B = 300$ simulated sets

Simulations - Prior predictive check

Restricted dataset

- 33 districts within the region of *London*
- 3rd of July 2020 to the 21st of February 2021
- $B = 300$ simulated sets



Simulations - Parameter recovery

Estimation

- 2 parallel chains with 2000 iterations each
- 80% – 20% train-test split

Simulations - Parameter recovery

Estimation

- 2 parallel chains with 2000 iterations each
- 80% – 20% train-test split

Metric	lr_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
rMSE	0.09	0.02	0.02	0.07	0.08	0.08	0.08	0.16	0.12
Cov ₉₅	0.95	0.94	0.96	0.95	0.94	0.95	0.92	0.92	0.94

Metric	α_ϕ	α_ψ	ρ_ϕ	ρ_ψ	σ_ϕ	σ_ψ
rMSE	0.21	0.18	0.21	0.17	0.05	0.08
Cov ₉₅	0.93	0.94	0.93	0.93	0.91	0.94

Simulations - Predictive performances

Metric	In-Sample	Out-of-Sample
RMSE	0.03	0.09
Cov ₉₅	0.99	0.96