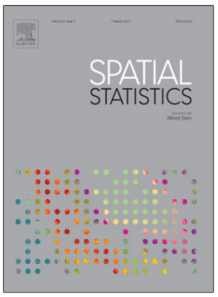

Quantifying the small-area spatio-temporal dynamics of the Covid-19 pandemic in Scotland during a period with limited testing capacity

Duncan Lee, Chris Robertson, Diogo Marques

PII: S2211-6753(21)00018-X
DOI: https://doi.org/10.1016/j.spasta.2021.100508
Reference: SPASTA 100508

To appear in: *Spatial Statistics*

Received date : 2 February 2021
Revised date : 31 March 2021
Accepted date : 31 March 2021

Please cite this article as: D. Lee, C. Robertson and D. Marques, Quantifying the small-area spatio-temporal dynamics of the Covid-19 pandemic in Scotland during a period with limited testing capacity. *Spatial Statistics* (2021), doi: https://doi.org/10.1016/j.spasta.2021.100508.

Quantifying the small-area spatio-temporal dynamics of the Covid-19 pandemic in Scotland during a period with limited testing capacity

Duncan Lee[a,*], Chris Robertson[b], Diogo Marques[c]

[a]*School of Mathematics and Statistics, University of Glasgow, Glasgow, G12 8SQ, Scotland.*
[b]*Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XH, Scotland, and Public Health Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, Scotland.*
[c]*Public Health Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, Scotland.*

Abstract

Modelling the small-area spatio-temporal dynamics of the Covid-19 pandemic is of major public health importance, because it allows health agencies to better understand how and why the virus spreads. However, in Scotland during the first wave of the pandemic testing capacity was severely limited, meaning that large numbers of infected people were not formally diagnosed as having the virus. As a result, data on confirmed cases are unlikely to represent the true infection rates, and due to the small numbers of positive tests these data are not available at the small-area level for confidentiality reasons. Therefore to estimate the small-area dynamics in Covid-19 incidence this paper analyses the spatio-temporal trends in telehealth data relating to Covid-19, because during the first wave of the pandemic the public were advised to call the national telehealth provider NHS 24 if they experienced symptoms of the virus. Specifically, we propose a multivariate spatio-temporal correlation model for modelling the proportions of calls classified as either relating to Covid-19 directly or having related symptoms, and provide software for fitting the model in a Bayesian setting using Markov chain Monte Carlo sim-

*Corresponding author
Email addresses: Duncan.Lee@glasgow.ac.uk (Duncan Lee), chris.robertson@strath.ac.uk (Chris Robertson), diogo.marques@nhs.net (Diogo Marques)

ulation. The model was developed in partnership with the national health agency Public Health Scotland, and here we use it to analyse the spatio-temporal dynamics of the first wave of the Covid-19 pandemic in Scotland between March and July 2020, specifically focusing on the spatial variation in the peak and the end of the first wave.

Keywords: Covid-19 pandemic, Gaussian Markov random field models, Scotland, Telehealth data.

1. Introduction

Covid-19 represents the biggest public health challenge in decades, and was declared a global pandemic by the World Health Organisation on 11th March 2020. The disease originated in the city of Wuhan in the People's Republic of China in December 2019, and reached the USA and Europe towards the end of January 2020. The first European epicentre for Covid-19 was in northern Italy in February 2020, and in Scotland, the focus of this paper, the first confirmed case occurred on the 2nd March 2020 (Public Health Scotland, `https://www.opendata.nhs.scot/dataset/covid-19-in-scotland`). Since then Covid-19 has spread across the world causing global health and economic devastation, and as of $30th$ March 2021 there have been over 127 million cases worldwide with over 2.7 milion people sadly dying from the disease (Johns Hopkins Coronavirus Resource Centre, `https://coronavirus.jhu.edu/map.html`).

Unsurprisingly, modelling the spread and dynamics of the Covid-19 pandemic has become a research priority, and there is a quickly growing research literature in this area. This literature has focused on a range of important epidemiological topics, including: (i) predicting the spread of the pandemic and its impacts on healthcare systems (Remuzzi and Remuzzi, 2020); (ii) identifying the factors that make people more at risk of displaying severe symptoms (Conticini et al., 2020, Wu et al., 2020 and Konstantinoudis et al., 2021); (iii) identifying the wider health impacts of the pandemic (Douglas et al., 2020); and (iv) developing surveillance systems for identifying the spatio-temporal dynamics in disease incidence (Dong et al., 2020). Developing a small-area surveillance system for monitoring the spatio-temporal trend in Covid-19 incidence is a vital tool in the fight against the virus, because it allows public health agencies to monitor its spread and identify hot-spots with high incidence, as well as providing vital clues as to how and why the

virus spreads more easily in certain areas.

The focus of this study is Covid-19 surveillance in Scotland, which is currently in its second wave of infection since September 2020. During this second wave the spatio-temporal spread of the pandemic can be measured using data on positive tests at the small-area scale, which is due to Scotland having a wide-spread testing programme during this period. This programme allows any member of the public to book a test at `https://www.gov.uk/get-coronavirus-test`, and well over 15,000 tests are conducted each day. However, during the first wave of the pandemic between March and July 2020 Covid-19 testing capacity was strictly limited to priority groups, because there was a lack of infrastructure to allow large-scale testing. For example, in March 2020 only 350 tests could be conducted each day (`https://www.gov.scot/publications/foi-202000084813/`), which rose to 1,900 in April 2020. Therefore in this first wave the public were not able to access a diagnostic test to determine if they had the virus unless a test was recommended by a doctor. Instead, anyone experiencing symptoms was advised to phone the national telehealth service NHS 24 for medical advice, and was then asked to self-isolate at home. As a result data on confirmed Covid-19 cases will not provide a detailed picture of the spatio-temporal spread of the virus during this first wave, because only a very small fraction of the actual cases were confirmed by a positive test.

Due to this massive under-reporting the aim of this paper is to use proxy indicators of disease incidence to quantify the small-area spatio-temporal dynamics of the Covid-19 pandemic in Scotland during its first wave of infections. Specifically, we aim to estimate both Scotland-wide and small-area temporal trends in disease incidence, focusing on both the peak and the end of this first wave. As people with symptoms during this first wave were advised to phone NHS 24 for medical advice, we model data on the numbers of NHS 24 calls categorised as Covid-19 or having related symptoms at the small-area scale on a weekly basis. The model we developed was run by analysts in Public Health Scotland (PHS) on this proxy measure of disease incidence on a weekly basis during the first wave of the pandemic, allowing them to better understand the spread of the virus and target public health interventions appropriately at areas likely to exhibit the greatest risks.

Our model is a multivariate binomial spatio-temporal random effects model, with inference in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation. It jointly models the spatio-temporal variation in the numbers of calls to NHS 24 directly categorised as Covid-19, as well as those

calls categorised with related symptoms such as fever and difficulty breathing, the latter ensuring that potential local outbreaks are not missed due to calls being misclassified. In developing this model the key methodological challenge we address is the complex multivariate spatio-temporal structure of the data, which means we need to capture spatial, temporal and between call type correlations.

The development of multivariate space-time (MVST) models for disease risk modelling is a relatively new advance, with Carroll et al. (2017) and Lawson et al. (2017) proposing innovative mixture models, Quick et al. (2017) proposing a fully MVST Gaussian Markov Random Field (GMRF, Rue and Held, 2005) model, while Jack et al. (2019) combine separate simpler multivariate spatial and multivariate temporal processes. The model we propose here is most similar to that proposed by Quick et al. (2017), because it uses a Gaussian Markov Random Field prior distribution applied to a set of random effects to model the multivariate spatio-temporal correlations inherent in the data. Our model extends that of Quick et al. (2017) by considering first and second order temporal autoregressive dependence structures, as well as allowing for varying strengths of spatial correlation via the Leroux spatial correlation model (Leroux et al., 2000). The NHS 24 telehealth data for the first wave of the pandemic that we analyse are described in Section 2, while our multivariate spatio-temporal model is presented in Section 3. Our surveillance model is applied to the Scottish telehealth data in Section 4, while Section 5 concludes the paper.

2. Covid-19 telehealth data in Scotland

2.1. NHS 24 and the study region

NHS 24 (`https://www.NHS24.scot/`) is Scotland's national telehealth service, and gives the public phone access to non-emergency medical advice 24 hours a day and 7 days a week when their regular primary health care providers are closed. NHS 24 deals with around 1.5 million calls per year and serves a population of around 5.4 million people, and at peak demand answers around 14,500 calls over the course of a weekend. Data were obtained from Public Health Scotland (PHS, `https://publichealthscotland.scot/`) on the weekly numbers of calls to NHS 24 for Covid-19 and other similar conditions during the first wave of the pandemic, which spanned $N = 22$ weeks from the week beginning $2nd$ March 2020 to the week beginning $27th$ July 2020 inclusive. A weekly temporal scale was used because it smooths out the

large amount of noise in the daily data caused by small numbers of calls and known day of the week effects, the latter including the fact that there are more calls during the weekends when doctors surgeries are closed.

The data have been aggregated to the 444 postcode districts (PD) within Scotland, and a shapefile containing the spatial boundary information for these PDs was obtained from the National Records for Scotland (`https://www.nrscotland.gov.uk`). This spatial boundary information did not include 8 of the PDs in the data set, but as these PDs only accounted for 44 NHS 24 calls out of a total of 524,036 calls they were removed from the study region. After removing these PDs there were 1005 instances (PD and week combinations) with no NHS 24 calls at all, which were spread relatively evenly across the 22 weeks with between 34 and 56 instances each week. Therefore, to ensure a rectangular data set for analysis, only the $K = 328$ PDs having at least 1 NHS 24 call (about any illness) per week were retained in the study region. The PDs removed from the data only accounted for 0.7% of the total calls to NHS 24, and were mostly sparsely populated rural or industrial / commercial areas.

2.2. Data available

For the kth PD and tth week the data comprise the following counts of the numbers of calls to NHS 24: (i) N_{kt} - the total number of calls to NHS 24; (ii) Y_{kt1} - the number of calls classified as Covid-19; and (iii) Y_{kt2} - the number of calls classified as `Simple Estimate 1` (hereafter `SE1`), which is a set of symptoms potentially related to Covid-19 including cold, flu, coughs, fever and difficulty breathing. The latter is modelled here to ensure that potential local outbreaks are not missed due to a misclassification of calls. The classification for Covid-19 was only initially available from $14th$ April onwards, but was back-predicted to $2nd$ March using a prediction model developed by PHS to allow trends to be modelled over the peak of the first wave of the pandemic. The prediction model was developed using NHS 24 call data from mid April to the end of May relating to respiratory and gastrointestinal syndromes plus the patients age. The prediction performance of this model had a specificity of 96% and a sensitivity of 75%, with an area under the curve (AUC) of 0.88. Therefore to ensure the Covid-19 series covers the peak of the first wave of the pandemic, we treat these predictions as observed data.

2.3. Limitations with the data

As discussed in the introduction wide-scale testing of Covid-19 was not available during the first wave of the pandemic, and the public were instead advised to phone NHS 24 if they developed Covid-like symptoms. These considerations motivate our use of the NHS 24 data as a proxy measure of disease incidence, but one must be cognisant of the issues that arise with these data not relating to laboratory confirmed cases. The main issue is misclassification of calls, because a person phoning NHS 24 with Covid-like symptoms does not mean they actually have the virus. Furthermore, the NHS 24 call handler may misdiagnose the patients symptoms, and hence wrongly classify them as having or not having Covid-19. This potential for misclassification is why we jointly model calls classified as Covid-19 and SE1, and examine the similarities and differences in the spatio-temporal dynamics of both classifications. Furthermore, each NHS 24 call can actually have multiple classifications, and as expected there is substantial overlap in the calls classified as Covid-19 and SE1. In fact, the total number of calls classified as Covid-19 or SE1 is sometimes greater than the total number of calls, i.e $Y_{kt1}+Y_{kt2} > N_{kt}$, particularly where N_{kt} is small. Thus in the next section we model these two classifications as a correlated multivariate binomial process rather than with a multinomial distribution.

A further potential issue with using the NHS 24 data as a proxy measure of disease incidence is that an individual may call NHS 24 more than once during a week, either for different or for the same reason. Hence the data we model relate to the numbers of calls to NHS 24 rather than the number of individuals who call NHS 24. However, the number of individuals who call NHS 24 multiple times for Covid-like symptoms within a week should be low, because the NHS 24 call handlers are trained to provide expert medical advice, precluding the need for multiple calls by the same individual. Thus despite these limitations the NHS 24 data provide the most comprehensive, if imperfect, data source for quantifying the spatio-temporal dynamics of the first wave of the Covid-19 pandemic across Scotland, which is why we model them here.

2.4. Exploratory analysis

The correlations between the proportions of calls, $\hat{\theta}_{ktj} = Y_{ktj}/N_{kt}$, classified as Covid-19 ($j = 1$) and SE1 ($j = 2$) across all PDs for each week range between 0.60 and 0.94, suggesting there is a strong relationship between them. This is further evidenced by the top panel (A) of Figure 1, which displays the

temporal trends in these raw proportions. In the figure jittering has been added to the week beginning (horizontal) dimension to improve the visibility of the points, and the proportions for Covid-19 are in red while those for SE1 are in blue. The trend line in each case has been estimated using generalised additive model (GAM) smoothing. The figure shows a number of key points, the first of which is large amounts of noise in the data arising from small numbers of calls in some PDs, with sample proportions equal to 0 or 1 in 6.4% (Covid-19) and 7.4% (SE1) of week and PD combinations respectively. Secondly, the temporal trends are broadly similar for Covid-19 and SE1, showing a rise in the proportions from the $2nd$ March, a peak around $23rd$ March, a decrease until $1st$ June, and a generally steady state since then. Thirdly, the figure shows that the dominant classification seems to change around the week beginning $6th$ April, with more calls classified as SE1 before that date and more Covid-19 calls after that date. This may be an artifact of the prediction model used to back-predict the Covid-19 classification before $14th$ April, or alternatively it may be that as the pandemic became more prevalent from late March onwards people might be more likely to mention Covid-19 directly when they called NHS 24.

The median lag-1 temporal autocorrelation coefficients across the $K = 328$ PDs are respectively 0.54 (Covid-19) and 0.70 (SE1), which suggests these data are likely to exhibit temporal autocorrelation as expected. The raw proportions also exhibit spatial autocorrelation, which was quantified for each week and call classification using Moran's I (Moran, 1950) statistics and a corresponding Monte-Carlo p-value to test the null hypothesis of no spatial autocorrelation. The computation of Moran's I statistic requires an adjacency or neighbourhood structure between the K PDs to be specified, and details of its construction that accounts for the fact that PDs with no NHS 24 calls have been removed is given in the model specification in Section 3.2. From these Moran's I tests 41% (Covid-19) and 23% (SE1) of these weekly p-values were significant at the 5% level, suggesting that despite the noise in these raw proportions, spatial autocorrelation is likely to be present in the data.

2.5. Aims of the analysis

Thus as the data exhibit spatio-temporal and between call type correlations contaminated by noise due to small numbers, a multivariate spatio-temporal smoothing model is proposed in the next section to estimate the underlying

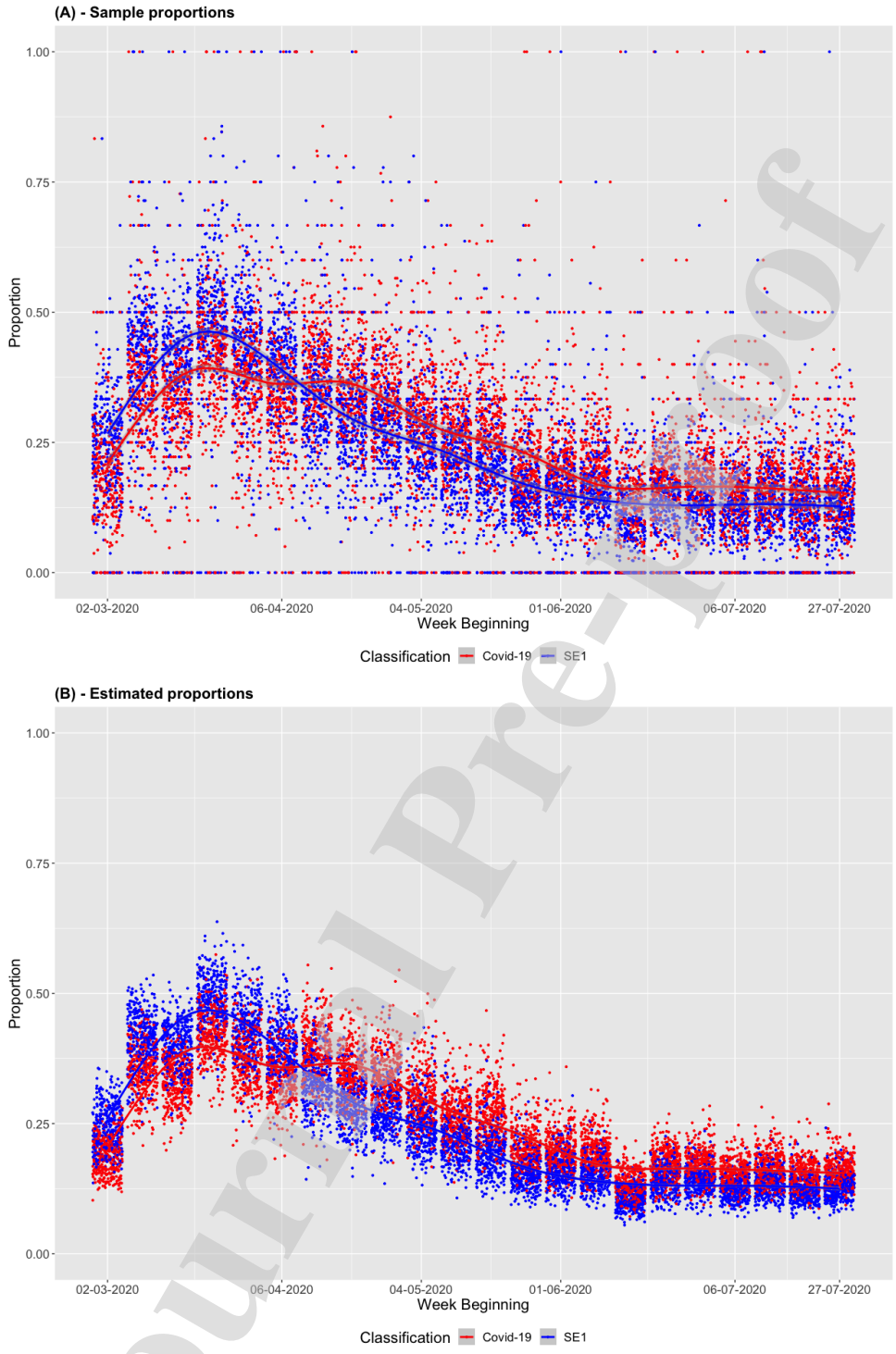

Figure 1: Scatterplots showing the temporal trends in the proportions of calls to NHS 24 that were related to Covid-19 (red) and SE1 (blue) for all PDs as points, with generalised additive model smoothed trend lines superimposed. The points have been jittered in the Week Beginning (horizontal) direction to improve their visibility. Panel (A) relates to the sample proportions and panel (B) to the estimated proportions from the final model (AR(2) Intrinsic CAR model with $D = 7$).

trends in these data. Specifically, our 2 underlying goals when modelling these data are to:

(a) Estimate the Scotland-wide spatio-temporal trend in disease incidence across the first wave of the pandemic.
(b) Estimate the spatial variation in this overall trend, particularly the extent of the spatial variation in when each PD in Scotland reached its peak and the end of its first pandemic wave.

3. Methodology

This section proposes a new multivariate spatio-temporal (MVST) model for estimating the spatio-temporal trends in the proportions of NHS 24 calls classified as either Covid-19 or having related symptoms (SE1). The model is fitted in a Bayesian setting using MCMC simulation, using a combination of Gibbs sampling and Metropolis-Hastings steps. Software to implement the model in `R` is available in the `CARBayesST` package (Lee et al., 2018), which allows others to apply the MVST models considered here to their own data.

3.1. Level 1 - Data likelihood model

Let Y_{ktj} denote the number of calls to NHS 24 in the kth PD ($k = 1, \ldots, K$) during the tth week ($t = 1, \ldots, N$) for the jth outcome ($j = 1, \ldots, J$), where for our data $j = 1$ is Covid-19 and $j = 2$ is SE1. Additionally, let N_{kt} denote the total number of NHS 24 calls in the kth PD and tth week. Then as the two outcomes (call classifications) are not disjoint as described in Section 2, a multinomial model is not appropriate for these data. Instead, we model these data as conditionally independent binomial distributions, where the spatio-temporal and between outcome (auto) correlations are modelled by random effects at the second level of the model hierarchy. The first level of the hierarchical model is given by:

$$\begin{aligned} Y_{ktj} &\sim \text{Binomial}(N_{kt}, \theta_{ktj}) \\ \ln\left(\frac{\theta_{ktj}}{1-\theta_{ktj}}\right) &= \beta_j + \phi_{ktj}. \end{aligned} \tag{1}$$

Here, θ_{ktj} is the true unknown proportion of calls (or probability that a single call) to NHS 24 in PD k during week t that is due to outcome j, and the

spatio-temporal variation in the estimated $\{\hat{\theta}_{ktj}\}$ provides a proxy measure of the incidence of the virus in the absence of comprehensive testing data. We do not include any covariates in our model for two reasons, the first of which is that our aim is to estimate the spatio-temporal trends in $\{\theta_{ktj}\}$ via the random effects $\{\phi_{ktj}\}$, rather than explaining what factors are associated with these trends. Secondly, up-to-date temporally varying covariate information is not available on a weekly basis, meaning that it would not be available to include in the model. The intercept terms β_j are outcome specific, which allows the two call types to have different average proportions over all PD and time period combinations. We assign weakly informative independent Gaussian prior distributions given by $\beta_j \sim \text{N}(0, 100,000)$ to these outcome specific intercept terms, which allow the data to play the dominant role in estimating their values.

3.2. Level 2 - Multivariate spatio-temporal random effects model

The remaining term in (1) $\{\phi_{ktj}\}$ are random effects, which are the mechanism for estimating the smooth multivariate spatio-temporal trends in $\{\theta_{ktj}\}$ for all outcomes. As such, the prior distribution for these random effects must induce (auto)correlations in time, space and between outcomes. The entire set of random effects are denoted by $\boldsymbol{\phi} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N)$, where $\boldsymbol{\phi}_t = (\boldsymbol{\phi}_{1t}, \ldots, \boldsymbol{\phi}_{Kt})$ denotes the set of $K \times J$ random effects at time t, while $\boldsymbol{\phi}_{kt} = (\phi_{kt1}, \ldots, \phi_{tkJ})$ denotes the subset of these effects at the kth PD for all J outcomes. As mentioned earlier MVST models are in their infancy for areal unit data, and we follow the general approach of Quick et al. (2017) and propose a zero-mean multivariate Gaussian Markov random field (Rue and Held, 2005) model for $\boldsymbol{\phi}$. The general form of the model is given by

$$\boldsymbol{\phi} \sim \text{N}\left(\mathbf{0}, \left[\mathbf{D}(\boldsymbol{\alpha}) \otimes \mathbf{Q}(\mathbf{W}, \rho) \otimes \boldsymbol{\Sigma}^{-1}\right]^{-1}\right), \tag{2}$$

where $\otimes$ denotes a Kronecker product. The precision matrix is given by $\mathbf{P}(\boldsymbol{\alpha}, \rho, \boldsymbol{\Sigma}) = \mathbf{D}(\boldsymbol{\alpha}) \otimes \mathbf{Q}(\mathbf{W}, \rho) \otimes \boldsymbol{\Sigma}^{-1}$, where $\mathbf{D}(\boldsymbol{\alpha})_{N\times N}$ controls the temporal autocorrelations, $\mathbf{Q}(\mathbf{W}, \rho)_{K\times K}$ controls the spatial autocorrelations and $\boldsymbol{\Sigma}_{J\times J}$ captures the between outcome correlations. The precision matrix $\mathbf{P}(\boldsymbol{\alpha}, \rho, \Sigma)$ is sparse because both $[\mathbf{D}(\boldsymbol{\alpha}), \mathbf{Q}(\mathbf{W}, \rho)]$ are sparse as they are built from specific cases of GMRFs (described below), which enables computationally efficient Bayesian inference by making use of their triplet form representation. As the model is defined in terms of its precision matrix

$\mathbf{P}(\boldsymbol{\alpha}, \rho, \boldsymbol{\Sigma})$ rather than its covariance matrix, multivariate Gaussian theory gives the following partial (auto)correlations for (ϕ_{ktj}, ϕ_{rsi}) conditional on the remaining random effects $\boldsymbol{\phi}_{-ktj,rsi}$:

$$\text{Corr}(\phi_{ktj}, \phi_{rsi}|\boldsymbol{\phi}_{-ktj,rsi}) = \frac{-\mathbf{D}(\boldsymbol{\alpha})_{ts}\mathbf{Q}(\mathbf{W}, \rho)_{kr}\left(\boldsymbol{\Sigma}^{-1}\right)_{ji}}{\sqrt{\left(\mathbf{D}(\boldsymbol{\alpha})_{tt}\mathbf{Q}(\mathbf{W}, \rho)_{kk}\right)\left(\boldsymbol{\Sigma}^{-1}\right)_{jj}\left(\mathbf{D}(\boldsymbol{\alpha})_{ss}\mathbf{Q}(\mathbf{W}, \rho)_{rr}\left(\boldsymbol{\Sigma}^{-1}\right)_{ii}\right)}}. \quad (3)$$

In what follows we now discuss the three components of the precision matrix in turn.

3.2.1. Between outcome correlation

The between outcome covariance matrix $\boldsymbol{\Sigma}$ is not assigned a specific structure, and is instead assigned the following conjugate Inverse-Wishart prior distribution

$$\boldsymbol{\Sigma} \sim \text{Inverse-Wishart}(d, \boldsymbol{\Omega}). \quad (4)$$

The hyperparameters are set at $(d = J + 1, \boldsymbol{\Omega} = 0.01\mathbf{I})$ where $\mathbf{I}$ is the identity matrix, and are chosen to ensure it is only weakly informative.

3.2.2. Spatial autocorrelation

Spatial autocorrelation is modelled by a conditional autoregressive (CAR) prior, which is a special case of a GMRF. The prior requires the specification of a $K \times K$ neighbourhood or adjacency matrix $\mathbf{W} = (w_{kr})$ that quantifies the spatial closeness between each pair of PDs. Here we adopt a binary specification where $w_{kr} = 1$ if PDs (k, r) are spatially close together, and $w_{kr} = 0$ otherwise, with $w_{kk} = 0 \; \forall \; k$. The most common approach in the literature is to specify $\mathbf{W}$ via the border sharing rule, that is $w_{kr} = 1$ if areas (k, r) share a common border and $w_{kr} = 0$ otherwise. However our study region has numerous islands, as well as additionally a number of mainland PDs with no NHS 24 calls that have therefore been removed. As a result this border sharing specification leads to a corresponding graph with 15 separate unconnected components, one main one containing most of the areas, 7 small components containing between 2 and 8 areas, and 7 additional isolates with no neighbours at all.

Therefore to obtain a neighbourhood structure with all the PDs in a single connected component we use the D-nearest neighbours rule (after removing

the PDs with no NHS 24 calls), which first represents the location of each PD by its centroid (central point). Then based on these centroids it specifies $w_{kr} = 1$ if the rth PD is one of the D nearest PDs to the kth PD, and $w_{kr} = 0$ otherwise. This leads to an asymmetric $\mathbf{W}$ matrix, which is made symmetric for the purposes of fitting the model by if $w_{kr} = 1$ and $w_{rk} = 0$ then setting $w_{rk} = 1$. In the analysis in the next section we consider $D = 3, 5, 7$ to assess the sensitivity of our results to this choice. Further details on specifying neighbourhood matrices can be found in Bivand et al. (2013). Based on $\mathbf{W}$ we model the spatial autocorrelation via the CAR prior proposed by Leroux et al. (2000), which corresponds to the following spatial precision matrix

$$\mathbf{Q}(\mathbf{W}, \rho) \quad = \quad \rho(\text{diag}[\mathbf{W1}] - \mathbf{W}) + (1 - \rho)\mathbf{I}. \tag{5}$$

Here $(\mathbf{1}, \mathbf{I})$ are a $K \times 1$ vector of ones and the $K \times K$ identity matrix respectively, while diag$[\mathbf{W1}]$ denotes a diagonal matrix with diagonal elements $\mathbf{W} \times \mathbf{1}$, so that the kth diagonal element is given by $\sum_{i=1}^{K} w_{ki}$. This specification models (ϕ_{ktj}, ϕ_{rtj}) as partially spatially autocorrelated if $w_{kr} = 1$ and conditionally independent if $w_{kr} = 0$, which can be seen from (3) and the fact that for $k \neq r$ $\mathbf{Q}(\mathbf{W}, \rho)_{kr} = -\rho w_{kr}$. This also illustrates that ρ is a global spatial dependence parameter, with a value of 0 corresponding to spatial independence. In contrast, if $\rho = 1$ the model captures strong spatial autocorrelation and simplifies to the intrinsic CAR model proposed by Besag et al. (1991), and this simplification was used to capture spatial correlation by Quick et al. (2017) within an MVST setting. We specify a non-informative uniform prior on the unit interval for ρ, i.e. $\rho \sim \text{Uniform}(0, 1)$, which provides equal prior weight for all allowable values of ρ and allows the data to play the dominant role in estimating its value.

3.2.3. Temporal autocorrelation

Temporal autocorrelation is modelled using either first order or second order autoregressive processes, which are both special cases of a GMRF. This extends the work of Quick et al. (2017) who only consider the first order case. The joint distribution for $\boldsymbol{\phi}$ from (2) in each case can be decomposed as described below.

A - First-order autoregressive process

For a first-order autoregressive process the joint prior distribution $f(\boldsymbol{\phi})$ can be decomposed as

$$
\begin{aligned}
f(\boldsymbol{\phi}) &= f(\boldsymbol{\phi}_1)\prod_{t=2}^{N} f(\boldsymbol{\phi}_t|\boldsymbol{\phi}_{t-1}) \qquad (6)\\
&= \mathrm{N}\left(\boldsymbol{\phi}_1\middle|\mathbf{0}, \left[\mathbf{Q}(\mathbf{W},\rho)\otimes\boldsymbol{\Sigma}^{-1}\right]^{-1}\right)\prod_{t=2}^{N}\mathrm{N}\left(\boldsymbol{\phi}_t\middle|\alpha\boldsymbol{\phi}_{t-1}, \left[\mathbf{Q}(\mathbf{W},\rho)\otimes\boldsymbol{\Sigma}^{-1}\right]^{-1}\right),
\end{aligned}
$$

which is combined with the improper non-informative prior $f(\alpha) \propto 1$. This specification corresponds to a tridiagonal matrix for $\mathbf{D}(\alpha)$ with entries

$$
\begin{aligned}
\mathbf{D}(\boldsymbol{\alpha})_{t,t} &= \begin{cases} 1+\alpha^2 & \text{for } t = 1,\ldots,N-1 \\ 1 & \text{for } t = N \end{cases},\\
\mathbf{D}(\boldsymbol{\alpha})_{t,t-1} &= -\alpha \quad \text{for } t = 2,\ldots,N.
\end{aligned}
$$

Thus from (3) it is clear that (ϕ_{ktj}, ϕ_{ksj}) are conditionally independent if $s \notin \{t-1, t, t+1\}$.

B - Second-order autoregressive process

For a second-order autoregressive process the joint prior distribution $f(\boldsymbol{\phi})$ can be decomposed as

$$
\begin{aligned}
f(\boldsymbol{\phi}) &= f(\boldsymbol{\phi}_1)f(\boldsymbol{\phi}_2)\prod_{t=3}^{N} f(\boldsymbol{\phi}_t|\boldsymbol{\phi}_{t-1},\boldsymbol{\phi}_{t-2}) \qquad (7)\\
&= \mathrm{N}\left(\boldsymbol{\phi}_1\middle|\mathbf{0}, \left[\mathbf{Q}(\mathbf{W},\rho)\otimes\boldsymbol{\Sigma}^{-1}\right]^{-1}\right)\mathrm{N}\left(\boldsymbol{\phi}_2\middle|\mathbf{0}, \left[\mathbf{Q}(\mathbf{W},\rho)\otimes\boldsymbol{\Sigma}^{-1}\right]^{-1}\right)\\
&\times \prod_{t=3}^{N}\mathrm{N}\left(\boldsymbol{\phi}_t\middle|\alpha_1\boldsymbol{\phi}_{t-1}+\alpha_2\boldsymbol{\phi}_{t-2}, \left[\mathbf{Q}(\mathbf{W},\rho)\otimes\boldsymbol{\Sigma}^{-1}\right]^{-1}\right),
\end{aligned}
$$

which is combined with the improper non-informative prior $f(\alpha_1, \alpha_2) \propto 1$. This specification corresponds to the following sparse matrix for $\mathbf{D}(\boldsymbol{\alpha})$ with

non-zero entries

$$\begin{aligned}
\mathbf{D}(\boldsymbol{\alpha})_{t,t} &= \begin{cases} 1+\alpha_2^2 & \text{for } t=1 \\ 1+\alpha_1^2+\alpha_2^2 & \text{for } t=2,\ldots,N-2 \\ 1+\alpha_1^2 & \text{for } t=N-1 \\ 1 & \text{for } t=N \end{cases}, \\
\mathbf{D}(\boldsymbol{\alpha})_{t,t-1} &= \begin{cases} \alpha_1\alpha_2 & \text{for } t=2 \\ \alpha_1\alpha_2-\alpha_1 & \text{for } t=3,\ldots,N-1 \\ -\alpha_1 & \text{for } t=N \end{cases}, \\
\mathbf{D}(\boldsymbol{\alpha})_{t,t-2} &= -\alpha_2 \quad \text{for } t=3,\ldots,N.
\end{aligned}$$

Thus from (3) it is clear that (ϕ_{ktj}, ϕ_{ksj}) are conditionally independent if $s \notin \{t-2, t-1, t, t+1, t+2\}$.

4. Spatio-temporal dynamics of Covid-19 in Scotland

This section presents the results of fitting the MVST models to the Covid-19 telehealth data in Scotland during the first wave of the pandemic. In modelling these data our aims are to: (a) estimate the Scotland-wide spatio-temporal trend in disease incidence; and (b) estimate when each PD in Scotland reached the peak and end of its first pandemic wave.

4.1. Model fitting

We fit 12 different models to the data that have varying spatio-temporal correlation structures, because it allows us to examine the sensitivity of the results to model choice. Specifically, we fit models with all possible combinations of: (i) first and second order temporal autoregressive structures; (ii) spatial autocorrelation structures defined by the Leroux (given by (5)) and intrinsic (where $\rho = 1$ in (5)) CAR models; and (iii) the neighbourhood matrix $\mathbf{W}$ defined by the $D = 3$, 5 and 7 nearest neighbours rule. The model with a temporal first order autoregressive process and the Intrinsic CAR structure is the closest to that proposed by Quick et al. (2017), while the models based on a second order autoregressive process and a Leroux CAR structure are the extensions considered here. In what follows AR(1) / AR(2) respectively denote models with first and second order temporal autoregressive structures, while (I, L) respectively denote models with intrinsic and Leroux CAR spatial structures.

Inference for each of these 12 models is based on 3,000 MCMC samples generated from 3 independent Markov chains. Each chain was burnt in for

50,000 samples by which time convergence was assessed to have been reached, and then run for a further 300,000 samples which were thinned by 300 to greatly reduce their autocorrelation. Convergence was visually assessed using traceplots and numerically assessed using the Gelman-Rubin diagnostic, and for the latter none of the values of $\hat{R}$ were above 1.1, which is suggested as a convergence criteria by Gelman et al. (2013).

4.2. Model assessment

A summary of the fit of each model to the data is presented in Table 1, which displays the deviance information criterion (DIC, Spiegelhalter et al., 2002), the effective number of independent parameters (p.d), and the log marginal predictive likelihood (LMPL, Geisser and Eddy, 1979). The DIC measures the overall fit of each model to the data, and the model with an intrinsic CAR spatial structure and a second order autoregressive temporal structure fits the data best as it minimises the DIC. However, the overall fits of all the models are relatively similar, as there is only a 0.8% difference between the largest and smallest DIC values. The LMPL measures the predictive ability of each model and is calculated as LMPL= $\sum_{ktj} \ln[f(Y_{ktj}|\mathbf{Y}_{-ktj})]$, where $\mathbf{Y}_{-ktj}$ denotes all observations except for Y_{ktj}. The best fitting model is the one that maximises the LMPL, which is also achieved by the model with an intrinsic CAR spatial structure and a second order autoregressive temporal structure. However, in common with the DIC the differences between the models by this measure are also small, being at most 1.3%

The residuals from all models were assessed for the presence of any remaining spatial autocorrelation using a Moran's I permutation test separately for each year, and in all cases no significant autocorrelation remained. The presence of residual temporal autocorrelation was also checked for each model and PD, by determining whether the lag 1 autocorrelation coefficient was significantly different from zero at the 5% level. We based on our assessment on the lag one coefficient only because the data only contain $N = 22$ time periods making estimation of higher lags less reliable, and also because the Moran's I test is also only based on first order neighbours. The models with a second order autoregressive process adequately capture the temporal autocorrelation in the data, as in all cases only 5% of the sets of temporal residuals contain significant (at the 5% level) autocorrelation at lag 1. In contrast, the corresponding percentages for the models with a first order autoregressive process are between 12% - 14%, suggesting that an AR(1)

Table 1: Summary of all models fitted to the data, including overall fit to the observed data via the DIC, model complexity via the effective number of independent parameters (p.d), and predicitve ability via the log marginal predictive likelihood (LMPL).

Quantity	**W matrix**	**Spatio-temporal correlation model**			
		AR(1) - I	**AR(1) - L**	**AR(2) - I**	**AR(2) - L**
DIC	D=3	68,424	68,461	62,276	68,313
	D=5	68,139	68,171	68,014	68,057
	D=7	67,982	68,028	67,888	67,915
p.d	D=3	2,330	2,372	2,487	2,524
	D=5	2,579	2,612	2,689	2,720
	D=7	2,735	2,757	2,802	2,834
LMPL	D=3	-34,050	-34,065	-33,928	-33,941
	D=5	-33,828	-33,842	-33,726	-33,739
	D=7	-33,694	-33,722	-33,619	-33,631

temporal autocorrelation structure is not entirely sufficient for capturing the temporal autocorrelation in the data.

Finally, the fitted values from each model were plotted against the observed values, and in all cases good agreement was seen with no large outliers suggesting a lack of fit for individual data points. The estimated proportions $\{\hat{\theta}_{ktj}\}$ were also relatively similar for all models, with for example the differences between the AR(1) Leroux CAR model with $D = 3$ and the AR(2) Intrinsic CAR model with $D = 7$ (the two most dissimilar models) ranging between -0.06 and 0.06 on the proportion scale for both Covid-19 and SE1 call classifications.

4.3. Multivariate spatio-temporal correlation structures

The spatio-temporal and between outcome correlations estimated by each model are summarised in Table 2, which presents point estimates (posterior medians) and 95% credible intervals for key model parameters. The table shows that the estimated proportions of calls classified as Covid-19 and SE1 have similar levels of spatio-temporal variation, as the posterior medians of $(\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22})$ are similar for both models, albeit slightly larger for SE1 calls in all cases. The values of both $(\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22})$ increase with increasing numbers of spatial neighbours D, which occurs because the conditional distribution of $\boldsymbol{\phi}_{kt}|\boldsymbol{\phi}_{-kt}$ has a covariance matrix including the elements of $\boldsymbol{\Sigma}$ divided by a function of $\sum_{r=1}^{K} w_{kr}$. Thus as the average number of neighbours (con-

trolled by D) increases the conditional variance is divided by a bigger number, leading to the inflation of $(\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22})$. The table also shows substantial between outcome (call classification) correlations, which are computed by $(\boldsymbol{\Sigma}_{12}/\sqrt{\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{22}})$ and are very close to one for all models.

The levels of spatial dependence estimated by the Leroux CAR models are high because the posterior medians for ρ are close to or equal to 1 for all models, which corresponds to the intrinsic CAR model (where ρ is fixed at 1) for strong spatial dependence. Thus for these data there is little difference between the Intrinsic and Leroux CAR models, with the former having a better DIC due to it having a lower p.d as it does not need to estimate ρ. Substantial temporal dependence is also present in these data, because in the AR(1) and AR(2) models the respective 95% credible intervals for α and (α_1, α_2) are not close to zero which would represent temporal independence.

4.4. (a) Scotland-wide spatio-temporal trend in the pandemic

The remainder of this section presents the estimated spatio-temporal trend in the Covid-19 pandemic during its first wave in Scotland. All results relate to the AR(2) Intrinsic CAR model with $D = 7$, because this was shown to be the best model via both the DIC and LMPL metrics, as well as adequately capturing both the temporal and spatial correlations in the data.

The estimated (posterior median) proportions of calls $\{\hat{\theta}_{kt1}, \hat{\theta}_{kt2}\}$ to NHS 24 classified as Covid-19 and SE1 are displayed in the bottom panel of Figure 1, which has the same format as the top panel of the same figure, with Covid-19 in red and SE1 in blue. The estimated proportions exhibit much less noise than the raw proportions due to the spatio-temporal smoothing applied by the model, and the peak in the average proportions is 0.42 for Covid-19 and 0.49 for SE1 in the week beginning $23rd$ March. The trends in the estimated proportions are shown by generalised additive model curves, and the curve for SE1 is unimodal and has a steeper ascent and descent compared to the Covid-19 curve.

In contrast, the Covid-19 curve exhibits a second local maximum on the week beginning $13th$ April, and the very limited available data on confirmed cases at a national level also suggests the existence of a double peak (for details see `https://public.tableau.com/profile/phs.covid.19#!/vizhome/COVID-19DailyDashboard_15960160643010/Overview`). This double peak in the confirmed cases occurs slightly later with around a 2 week lag compared to the NHS 24 calls, which is likely to be partially caused by testing and reporting delays as the testing infrastructure was less advanced

Table 2: Summary of the posterior medians and 95% credible intervals for the covariance parameters from each of the models.

Quantity	W matrix	Spatio-temporal correlation model			
		AR(1) - I	**AR(1) - L**	**AR(2) - I**	**AR(2) - L**
$\boldsymbol{\Sigma}_{11}$	D=3	0.059 (0.051, 0.068)	0.060 (0.052, 0.070)	0.074 (0.065, 0.084)	0.074 (0.065, 0.083)
	D=5	0.151 (0.132, 0.172)	0.152 (0.134, 0.173)	0.175 (0.155, 0.195)	0.172 (0.153, 0.192)
	D=7	0.262 (0.231, 0.295)	0.260 (0.230, 0.292)	0.287 (0.257, 0.319)	0.282 (0.253, 0.315)
$\boldsymbol{\Sigma}_{22}$	D=3	0.062 (0.054, 0.072)	0.063 (0.055, 0.074)	0.077 (0.068, 0.087)	0.079 (0.069, 0.089)
	D=5	0.157 (0.136, 0.178)	0.159 (0.140, 0.180)	0.178 (0.158, 0.198)	0.183 (0.163, 0.205)
	D=7	0.271 (0.238, 0.304)	0.272 (0.239, 0.307)	0.293 (0.262, 0.326)	0.302 (0.271, 0.337)
$\boldsymbol{\Sigma}_{12}/\sqrt{\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{22}}$	D=3	0.997 (0.996, 0.998)	0.994 (0.991, 0.996)	0.997 (0.996, 0.998)	0.994 (0.992, 0.996)
	D=5	0.998 (0.997, 0.999)	0.995 (0.993, 0.997)	0.998 (0.998, 0.999)	0.995 (0.993, 0.997)
	D=7	0.999 (0.998, 0.999)	0.996 (0.993, 0.997)	0.999 (0.998, 0.999)	0.996 (0.993, 0.998)
ρ	D=3	-	1.000 (1.000, 1.000)	-	1.000 (1.000, 1.000)
	D=5	-	1.000 (0.999, 1.000)	-	1.000 (0.999, 1.000)
	D=7	-	0.999 (0.999, 1.000)	-	0.999 (0.999, 1.000)
$\boldsymbol{\alpha}$	D=3	α 0.770 (0.724, 0.810)	α 0.762 (0.713, 0.802)	α_1 0.459 (0.394, 0.529) α_2 0.346 (0.272, 0.419)	α_1 0.461 (0.394, 0.527) α_2 0.337 (0.269, 0.407)
	D=5	α 0.689 (0.637, 0.739)	α 0.687 (0.639, 0.730)	α_1 0.419 (0.360, 0.480) α_2 0.333 (0.267, 0.396)	α_1 0.419 (0.362, 0.479) α_2 0.324 (0.259, 0.387)
	D=7	α 0.638 (0.583, x0.687)	α 0.640 (0.591, 0.687)	α_1 0.401 (0.342, 0.459) α_2 0.319 (0.255, 0.379)	α_1 0.403 (0.346, 0.457) α_2 0.311 (0.251, 0.369)

than it is now. The average (over Scotland) estimated proportions of calls classified as Covid-19 for the weeks beginning $15th$ June onwards are lower than the average for $2nd$ March (the first week of the data), suggesting that the majority of the first wave of the pandemic had come to an end by this point.

The spatio-temporal trend in the Covid-19 classifications is summarised in Figure 2, which displays maps for the first and last week of the study as well as for the two peaks in the estimated proportions ($23rd$ March and $13th$ April) highlighted above. The figure shows that most PDs have relatively low proportions of calls in the first and last weeks below 0.2, while most PDs have increased proportions between 0.3 and 0.6 during the two weeks of peak Covid-19 activity. The figure also shows that the proportions of NHS 24 calls classified as Covid-19 do not show a pronounced spatial trend for any of the weeks, and instead show pockets of higher proportions in different parts of the country.

4.5. (b) PD specific temporal trends

The previous section suggested that on average the first wave of the pandemic peaked in Scotland in the week beginning $23rd$ March, and had reduced back to baseline levels seen at the beginning of March by $15th$ June. However, our second motivating question is to assess whether the pandemic hit some parts of Scotland earlier than other parts. Our hypothesis is that the pandemic would be likely to affect more connected urban areas before it affected more remote rural ones, due to the former's greater levels of population density (and hence mixing) and easier access to travel via proximity to airports.

To assess this Figure 3 displays maps for each PD displaying: (A) the week that $\hat{\theta}_{kt1}$ was at its highest, which represents the peak of its first wave; and (B) the first week after this peak that $\hat{\theta}_{kt1}$ was smaller than its value in the first week (i.e. smaller than $\hat{\theta}_{k11}$), which approximately represents the end of its first wave of infection. The maps relate to Covid-19 rather than the SE1 classification, because the previous section highlighted that the double peak observed in the Covid-19 trend (see Figure 1) resembles the limited testing data at a national level more closely than the single peak from the SE1 trend.

The figure shows that 62% of the PDs exhibited their peak in Covid-related calls during the week beginning $23rd$ March, with the 7% of the PDs that exhibited their peak two weeks earlier mainly being located around the largest city of Glasgow. In contrast, those PDs exhibiting later peaks

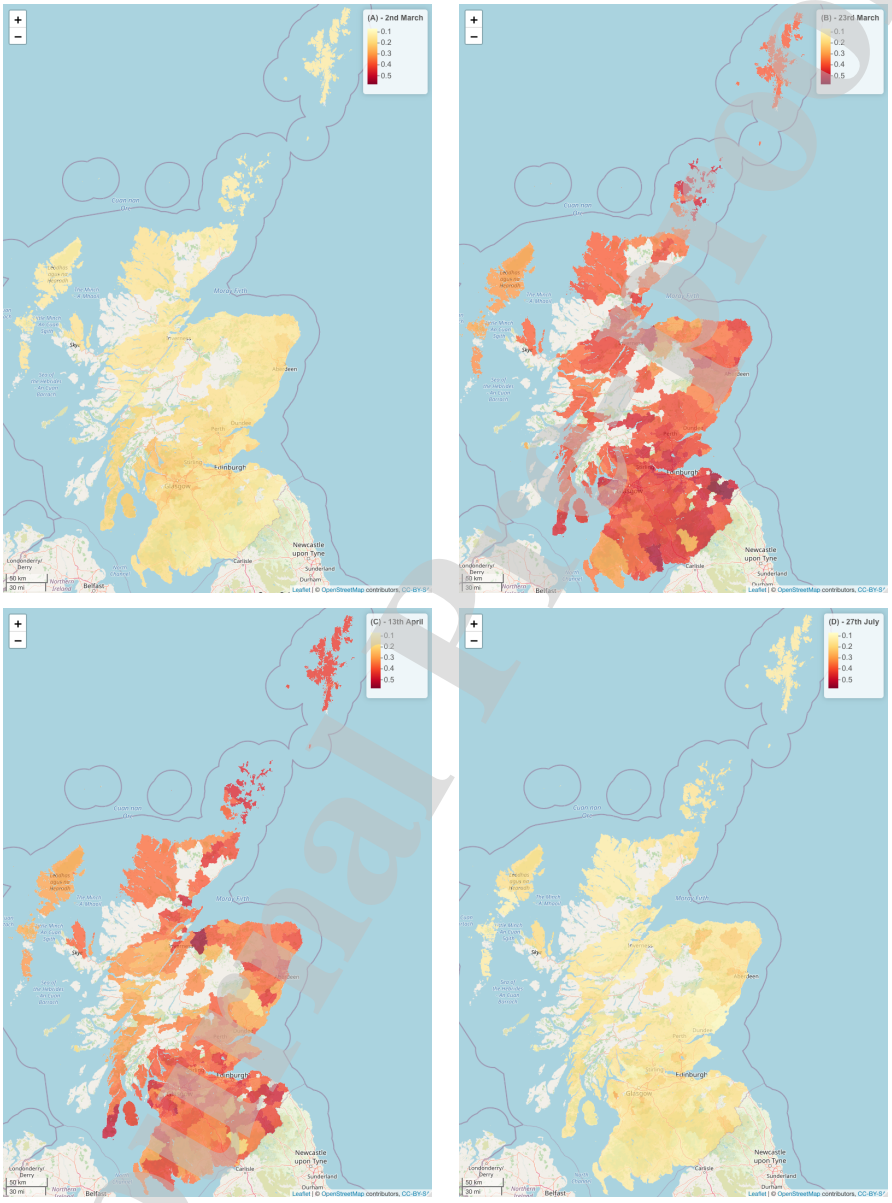

Figure 2: Maps displaying the proportions of NHS 24 calls classified as Covid-19 in four weeks of the pandemic.

(coloured red on the map) are mainly rural areas, with 20% of the PDs peaking in the week beginning 13th April. These PDs with later peaks are mostly in the more remote northern parts of Scotland that are away from the main cities. The right panel of Figure 3 displays the first week that the Covid-19 related calls were below their March 2nd levels, and a bimodal pattern is evident with 31% of PDs achieving this by 25th May while 40% met this by 15th June. In addition, 4% of the PDs had not seen their Covid-19 related calls drop below the 2nd March levels by the end of July, suggesting that in some areas the first wave of the pandemic had not yet finished by the end of our study. Finally, there is no clear urban-rural divide in these approximate end times of the first wave of the pandemic, which suggest that whilst urban areas were mainly affected first, they did not necessarily see the end of the wave first.

5. Discussion

This paper has developed a multivariate spatio-temporal model for quantifying the spread of Covid-19 in Scotland during the first wave of the pandemic, which was a period with limited testing capacity resulting in large numbers of infected people whose disease status was not confirmed by a diagnostic test. As a result we quantified the spatio-temporal dynamics of Covid-19 spread using proxy data from the national telehealth service NHS 24, who members of the public were advised to call if they experienced symptoms. The model estimates the joint spatio-temporal trends in the proportions of calls to NHS 24 classified as either Covid-19 directly or as having related symptoms (called SE1), and a simplification of the model using only the Covid-19 classification was run on a weekly basis by Public Health Scotland during the first wave of the pandemic as new data became available to monitor the likely locations of new outbreaks.

Modelling the spatio-temporal dynamics in the NHS 24 data allows us to study the spread of the pandemic at a small-area scale, albeit with a proxy measure of infection rates. However, as previously discussed testing capacity was severely limited in this initial stage of the pandemic, and hence data on confirmed cases would also only be a proxy measure of the true infection rates. Additionally, due to the small numbers of positive tests in this phase of the pandemic, small-area testing data are not available for confidentiality reasons, making it impossible to study the spread of the virus at the small-area scale using confirmed case data. Thus while telehealth data

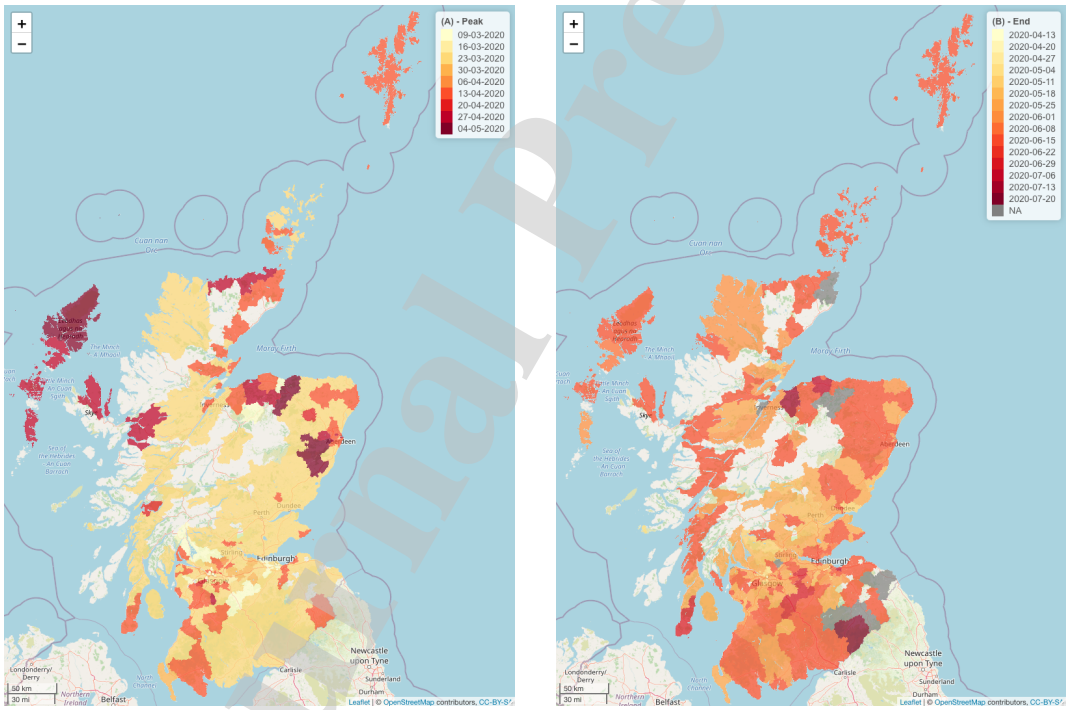

Figure 3: Maps displaying for each PD the weeks when the estimated proportions for the Covid-19 classification: (A) peaked; and (B) were below their 2*nd* March levels signifying the end of the first wave.

are imperfect as discussed above, we have illustrated the value of modelling them in early stage pandemic situations where reliable confirmed testing data are not available.

The paper has presented a number of findings from our data analysis, the first being that the first wave of the pandemic peaked in Scotland in the week beginning $23rd$ March, with a smaller peak 3 weeks later on $13th$ April. The $23rd$ March was the peak of the pandemic for 65% of the PDs, while the 19% of the PDs that peaked later than $13th$ April were largely rural areas in the north and west of Scotland. By the end of July all but 4% of the PDs had NHS 24 call levels for Covid-19 below the levels observed at the beginning of March when the first confirmed case was recorded ($2nd$ March) in Scotland, suggesting that the first wave of the pandemic was coming to an end by this point.

Our other main finding is the differential temporal trends in the Covid-19 and SE1 classifications, with the latter exhibiting a single peak and having a steeper decline in proportions after the pandemic peak. This steeper descent in its proportions may be because as the pandemic became more prevalent from late March onwards people might be more likely to mention Covid-19 directly when they called NHS 24, hence the proportions of calls attributed to SE1 declined to lower levels than those attributed to Covid-19.

The overarching aim of this paper was to estimate the key dynamics of the Covid-19 pandemic at a high spatio-temporal resolution in a retrospective manner, which is why no predictive modelling of the proportions of calls classified as Covid-19 or SE1 into the future was undertaken. However, the temporally autoregressive nature of the models would make such prediction straightforward via (6) or (7), and both the proportions $\{\theta_{k,T+1,j}\}$ and counts $\{Y_{k,T+1,j}\}$ could be predicted in this way, although for the latter $\{N_{k,T+1,j}\}$ would also need to be predicted. Thus an area of future work will be to utilise these MVST models to predict disease burden into the future, to allow NHS managers to predict the amount of health care resources (e.g. hospital beds) needed in the future.

Another area of future work would be to continue the development of spatio-temporal modelling tools for telehealth data, because it has clear future applications that extend beyond the early stage pandemic setting considered here. Other examples include the routine monitoring of ordinary seasonal flu and outbreaks of Norovirus, which would give the NHS better information on the likely prevalence of these diseases and where and when outbreaks are likely to occur, thus allowing targeted action to be taken in a

timely manner.

Acknowledgements

The authors gratefully acknowledge the helpful comments of an anonymous reviewer that improved the motivation for and content of this work. The NHS 24 data were provided by Public Health Scotland. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing Interests

The authors declare they have no competing interests.

References

Besag, J., J. York, and A. Mollié (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistics and Mathematics 43*, 1–59.

Bivand, R., E. Pebesma, and V. Gomez-Rubio (2013). *Applied Spatial Data Analysis with R* (2nd ed.). Springer.

Carroll, R., A. Lawson, R. Kirby, C. Faes, M. Aregay, and K. Watjou (2017). Space-time variation of respiratory cancers in South Carolina: a flexible multivariate mixture modeling approach to risk estimation. *Annals of Epidemiology 27*, 42 – 51.

Conticini, E., B. Frediani, and D. Caro (2020). Can atmospheric pollution be considered a co-factor in extremely high level of sars-cov-2 lethality in northern Italy? *Environmental Pollution 261*, 114465.

Dong, E., H. Du, and L. Gardner (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases 20*, 533–534.

Douglas, M., S. Katikireddi, M. Taulbut, M. McKee, and G. McCartney (2020). Mitigating the wider health effects of Covid-19 pandemic response. *BMJ 369*, m1557.

Geisser, S. and W. Eddy (1979). A predictive approach to model selection. *Journal of the American Statistical Association 74*, 153–160.

Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin (2013). *Bayesian Data Analysis* (3rd ed.). Chapman and Hall / CRC.

Jack, E., D. Lee, and N. Dean (2019). Estimating the changing nature of Scotland's health inequalities by using a multivariate spatiotemporal model. *Journal of the Royal Statistical Society Series A 182*, 1061–1080.

Konstantinoudis, G., T. Padellini, J. Bennett, B. Davies, M. Ezzati, and M. Blangiardo (2021). Long-term exposure to air-pollution and COVID-19 mortality in England: a hierarchical spatial analysis. *Environment International 146*, 106316.

Lawson, A., R. Carroll, C. Faes, R. Kirby, M. Aregay, and K. Watjou (2017). Spatiotemporal multivariate mixture models for Bayesian model selection in disease mapping. *Environmetrics 28*, e2465.

Lee, D., A. Rushworth, and G. Napier (2018). Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. *Journal of Statistical Software 84*(9), 1–39.

Leroux, B., X. Lei, and N. Breslow (2000). *Statistical Models in Epidemiology, the Environment and Clinical Trials, Halloran, M and Berry, D (eds)*, Chapter Estimation of Disease Rates in Small Areas: A New Mixed Model for Spatial Dependence, pp. 135–178. Springer-Verlag, New York.

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika 37*, 17–23.

Quick, H., L. Waller, and M. Casper (2017). Multivariate spatiotemporal modeling of age-specific stroke mortality. *Annals of Applied Statistics 11*, 2165–2177.

Remuzzi, A. and G. Remuzzi (2020). Covid-19 and Italy: what next? *The Lancet 395*, 1225–1228.

Rue, H. and L. Held (2005). *Gaussian Markov Random Fields: Theory and Applications*. Chapman and Hall / CRC.

Spiegelhalter, D., N. Best, B. Carlin, and A. Van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B 64*, 583–639.

Wu, X., R. Nethery, M. Sabath, D. Braun, and F. Dominici (2020). Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Science Advances 6*, eabd4049.