

# Rothamsted Repository Download

## A - Papers appearing in refereed journals

Comber, A, Harris, P. and Brunsdon, C 2023. Multiscale spatially varying coefficient modelling using a Geographical Gaussian Process GAM.  
*International Journal Of Geographical Information Science*.  
<https://doi.org/10.1080/13658816.2023.2270285>

The publisher's version can be accessed at:

- <https://doi.org/10.1080/13658816.2023.2270285>

The output can be accessed at:

<https://repository.rothamsted.ac.uk/item/98y7z/multiscale-spatially-varying-coefficient-modelling-using-a-geographical-gaussian-process-gam>.

© 27 October 2023, Please contact [library@rothamsted.ac.uk](mailto:library@rothamsted.ac.uk) for copyright queries.

## Multiscale spatially varying coefficient modelling using a Geographical Gaussian Process GAM

Alexis Comber, Paul Harris & Chris Brunsdon

**To cite this article:** Alexis Comber, Paul Harris & Chris Brunsdon (27 Oct 2023): Multiscale spatially varying coefficient modelling using a Geographical Gaussian Process GAM, International Journal of Geographical Information Science, DOI: [10.1080/13658816.2023.2270285](https://doi.org/10.1080/13658816.2023.2270285)

**To link to this article:** <https://doi.org/10.1080/13658816.2023.2270285>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 27 Oct 2023.



Submit your article to this journal [↗](#)



Article views: 271



View related articles [↗](#)



View Crossmark data [↗](#)

# Multiscale spatially varying coefficient modelling using a Geographical Gaussian Process GAM

Alexis Comber<sup>a</sup> , Paul Harris<sup>b</sup>  and Chris Brunsdon<sup>c</sup> 

<sup>a</sup>School of Geography, University of Leeds, Leeds, UK; <sup>b</sup>Sustainable Agriculture Sciences, Rothamsted Research, North Wyke, UK; <sup>c</sup>National Centre for Geocomputation, Maynooth University, Maynooth, Ireland

## ABSTRACT

This paper proposes a novel spatially varying coefficient (SVC) regression through a Geographical Gaussian Process GAM (GGP-GAM): a Generalized Additive Model (GAM) with Gaussian Process (GP) splines parameterised at observation locations. A GGP-GAM was applied to multiple simulated coefficient datasets exhibiting varying degrees of spatial heterogeneity and out-performed the SVC brand-leader, Multiscale Geographically Weighted Regression (MGWR), under a range of fit metrics. Both were then applied to a Brexit case study and compared, with MGWR marginally out-performing GGP-GAM. The theoretical frameworks and implementation of both approaches are discussed: GWR models calibrate multiple models whereas GAMs provide a full single model; GAMs can automatically penalise local collinearity; GWR-based approaches are computationally more demanding; MGWR is still only for Gaussian responses; MGWR bandwidths are intuitive indicators of spatial heterogeneity. GGP-GAM calibration and tuning are also discussed and areas of future work are identified, including the creation of a user-friendly package to support model creation and coefficient mapping, and to facilitate ease of comparison with alternate SVC models. A final observation that GGP-GAMs have the potential to overcome some of the long-standing reservations about GWR-based regression methods and to elevate the perception of SVCs amongst the broader community.

## ARTICLE HISTORY

Received 20 January 2023  
Accepted 4 October 2023

## KEYWORDS

Spatial regression; GWR

## 1. Introduction

A standard linear regression seeks to model the relationship between a response variable and a series of predictor variables. Via ordinary least squares (OLS), it estimates a single set of unknown regression coefficients for each predictor (and intercept) together with errors that are assumed independent and identically distributed with mean zero and a uniform variance (usually denoted  $\sigma^2$ ). A standard OLS regression

**CONTACT** Alexis Comber  [a.comber@leeds.ac.uk](mailto:a.comber@leeds.ac.uk)

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

assumes that the errors have a Gaussian distribution. For non-Gaussian response variables (such as integer count data), the OLS regression can be replaced with a Generalized Linear Model (GLM) (Nelder and Wedderburn 1972; McCullagh and Nelder 2019) whose coefficients are typically estimated via maximum likelihood (ML). For greater flexibility in capturing non-linear relationships between the response and predictors, a GLM can be extended by a Generalized Additive Model (GAM) (Hastie and Pregibon 2017) where each regression relationship can be modelled non-linearly, as an arbitrary function, say through splines or some other smoothing tool.

OLS regression and classic forms of GLM assume fixed (constant) coefficient processes. However, in reality, the relationship between the response and predictor variables in a geographical context may change with observation location, potentially due to unaccounted-for local factors or due to spatial heterogeneity in the process being examined. In this context, spatially varying coefficient (SVC) models provide an alternative to 'whole map' or global regressions that may incorrectly assume spatial stationarity in regression coefficients (Openshaw 1996). GAM smoothing functions allow for varying (non-constant) coefficient processes in attribute-space. In geographic space, spatial settings for the GAM smoothing functions are possible leading to the characterisation of response-to-predictor relationships changing across space, i.e. a SVC model.

The most popular SVC model is Geographically Weighted Regression (GWR) (Brunsdon *et al.* 2010) which has been subject to a number of developments including Multiscale GWR (MGWR) (Yang 2014; Fotheringham *et al.* 2017) which is now recommended as the default GWR (Comber *et al.* 2023). Interestingly these can be viewed in terms of a GAM (Brunsdon *et al.* 1999; Yu *et al.* 2020), where coefficients are estimated from data subsets falling under distance-weighted kernels around any spatial location via a weighted least squares algorithm for GWR and an iterative back-fitting algorithm for MGWR.

The focus of this study is only for multiscale SVC models, that allow each predictor-to-response relationship to operate at its own spatial scale. MGWR is the most popular SVC model (Comber *et al.* 2021, 2023). Other widely applied multiscale SVC models include Bayesian Gaussian Process (GP) models that use co-kriging (via a linear Model of Co-regionalisation (LMC)) (Bayes-GP) to generate unknown coefficients that are assumed spatially co-dependent stochastic processes (Gelfand *et al.* 2003; Finley 2011; Kim and Lee 2017; Finley and Banerjee 2020) and Eigenvector Spatial Filtering (ESF) with Moran coefficients (ESF-MC) (Murakami and Griffith 2015; Murakami *et al.* 2017) that extends the deterministic ESF model (Griffith 2008) with random effects to model stochastic spatial processes. These SVC constructs can be shown to have a direct relationship to each other (Murakami *et al.* 2017), where comparisons with MGWR can be found in the simulated studies of Wolf *et al.* (2018) and Murakami *et al.* (2019) for Bayes-GP and ESF-MC, respectively. Such comparative studies concluded that no-one multiscale SVC model could be considered significantly better than another, when jointly considering model accuracy and precision, computational overheads, and ease of application for the novice user. More recently, multiscale SVC models include a model that uses triangulation with bivariate spline estimators (Mu *et al.* 2018), a GP-based model that uses a frequentist approach with ML estimation (Dambon *et al.* 2021), and a GLM-based model that uses a reduced-

rank spline methodology (Fan and Huang 2022), but at the time of writing community take up for these models has been limited.

For this study, a novel SVC model is proposed as an alternative, not only to MGWR, but also to Bayes-GP, ESF-MC and the more recent SVC models listed above: a Geographical Gaussian Process GAM (GGP-GAM), in which GP splines are parameterised at observation locations. The combination of a GAM with GP splines constitutes the key advance. The main theoretical limitation with MGWR (and GWR) is that its rudimentary kernel-based parameterisation means that only a collection of local models is used to capture and quantify the non-stationary processes (re-using data each time), whereas Bayes-GP, ESF-MC and GGP-GAM each provide a single, global model formulation that is still non-stationary in nature (e.g. Finley (2011); Wolf *et al.* (2018)). This has led some to argue that MGWR (and GWR) is best suited to exploratory rather than inferential analyses (Wheeler and Calder 2007; Finley 2011; Dambon *et al.* 2021). However, the practical consequences of this theoretical limitation can in part, be quantified and are often marginal (Li *et al.* 2020; Yu *et al.* 2020). The GGP-GAM is evaluated through a simulation experiment with MGWR as comparison, and also through an empirical case study with UK Brexit vote data.

Common challenges across all multiscale SVC models, include: (i) sensitivity to the parameterisation of the spatial smoothing function (e.g. sensitivity to the location and number of knots in GGP-GAM, sensitivity to kernel and bandwidth form in MGWR, ensuring a positive-definite fit for the LMC in Bayes-GP), (ii) collinearity in the predictors (e.g. for instances of local collinearity when not present globally), (iii) use as a spatial predictor, (iv) computational burden, and (v) ease of application for the novice user together with intuitive understandings of process spatial heterogeneity from model outputs. Each of these challenges are discussed for the proposed GGP-GAM model in relation to MGWR only.

Comparison with MGWR is important as currently such studies exist only for basic GWR, where all non-stationary relationships are naively assumed to operate at the same spatial scale. For example, Páez *et al.* (2011) and Fotheringham and Oshan (2016) examined GWR's sensitivity to local collinearity; Harris *et al.* (2010) and Harris *et al.* (2011) for GWR's use as a spatial predictor, including hybrids with kriging; Li *et al.* (2019) and Lu *et al.* (2022) describe computational solutions to implementations of GWR. Whilst MGWR has some benefits to dealing with local collinearity (Harris 2019), its value as a spatial predictor is unclear (Lu *et al.* 2019) and its underlying back-fitting algorithm is computationally expensive (Fan and Huang 2022). Both GWR and MGWR also suffer from an absence of a generic framework for any form of Geographically Weighted model (Comber *et al.* 2022), whereas frameworks for GAMs are relatively mature meaning that issues such as those for collinearity are more easily resolved through various penalised model options in GAM R packages.

## 2. GAMs with Gaussian Process splines

### 2.1. GAMs

Generalized Additive Models (GAMs) are general in that they can handle outputs with many types of distributions and not just linear relationships, polynomial or not (Wood

2006; Fahrmeir *et al.* 2022). They are additive and because they generate multiple model terms which are added together to generate predictions. Methodologically, GAMs offer a middle ground between simple but interpretable linear models and complex but opaque machine learning approaches, where it is difficult to understand how one variable relates to an outcome. Thus GAMs provide an intuitive approach to fit relatively complex relationships in data with complex interactions and non-linearities, and the outputs provide easily understood measures of the relationship between predictor and response variables and how the outcome is modelled. They are able to predict well from complex systems, to quantify relationships and to make inferences about these relationships, for example about what variables are important and at what range of values are they most influential. GAMs can perform as well or better than most machine learning models and they are relatively fast.

GAMs typically model non-linear relationships using splines. Other approaches to modelling can also fit into a GAM framework, but in many situations splines are the building blocks of GAMs. A spline can be represented as a linear combination of functions called *\*basis functions\**. Each of these is assigned a coefficient and these are linearly combined to generate the predictions of outcome ( $\hat{y}$ ). Basis functions can be single or multi-dimensional in terms of predictor variables. Thus, the form of a GAM is composed of linear sums of multi-dimensional basis functions and this is what allows complex relationships to be modelled.

Splines can be of different forms depending on the problem and data (see below) and need to include an appropriate degree of 'wiggleness' to model non-linear relationships. This is determined by the spline smoothing parameter. If this is too small then the model will overfit the data, resulting in too many bases and too much noise in the function. If it is too large then the functions will be overly simple and fail to represent the complexity of the relationship. Thus a core problem to resolve in GAMs is to determine the smoothing parameters and balance overfitting versus capturing the complex in the relationship. This is done automatically in the GAM functions of the 'mgcv' R package (Wood 2015) used in this work.

Some earlier work has combined GAMs with location. Kelsall and Diggle (1998) used GAMs to estimate covariates with a discrete smooth (spline) for location. However they did not look at the interactions of individual covariates over space as was done by Kammann and Wand (2003), who combined geostatistical approaches with a GAM by merging kriging, cast as a linear mixed model, with an additive model in order to account for non-linearity. However, this was only done for the intercept and for prediction rather than process understanding. In recent work, Fan and Huang (2022) proposed a SVC model via a GLM with reduced-rank thin-plate splines. In these, the rotations of the covariates are not changed and the splines are fitted with fewer coefficients than there are data to smooth using an Eigen-decomposition for coefficient (rank) reduction. An alternative approach to SVC modelling is to consider the predictor-to-response relationship over space as a GP. Here a GP is a random process over functions whose terms can be modelled using GP splines within a GAM. In this work, a low rank GP spline, parameterised with location was used rather than the more standard thin-plate spline. This is because GPs are good for specifying autocorrelation in spatially-varying random functions and GP-based

smoothing using observations at specific locations should detect any spatial trends in the data.

## 2.2. A Geographical Gaussian Process GAM for SVC modelling

GAMs provide a method for calibrating regression models with unspecified functions of the predictor variables, of the form:

$$y = \alpha + f_1(z_1) + f_2(z_2) + \dots + f_m(z_m) + \epsilon \quad (1)$$

where  $z_j$  may be a scalar or a vector.

These can be extended such that each  $f_j(z_j)$  is a linear regression coefficient on another scalar predictor  $x_j$ :

$$y = \alpha(z_0) + x_1 f_1(z_1) + x_2 f_2(z_2) + \dots + x_m f_m(z_m) + \epsilon \quad (2)$$

Finally, if  $z_0 = z_1 = \dots = z_m = z$  say, and  $z$  is a vector of locations then this specifies a SVC model:

$$y = \alpha(z) + x_1 f_1(z) + x_2 f_2(z) + \dots + x_m f_m(z) + \epsilon \quad (3)$$

One way of specifying  $\alpha(z) \dots f_m(z)$  is that each function is generated from a GP and each function estimate is an *a posteriori* estimate of a GP with a zero mean.

Importantly in the context of this work, GPs also have a covariance function:

$$\kappa_m(\delta) = \text{Cov}(f_m(\delta), f_m(z + \delta)) \quad (4)$$

These control the smoothness of  $f_m(z)$ , such that the more  $\kappa_m(\delta)$  decreases as  $\delta$  increases, the smoother  $f_m(z)$  tends to be. In this sense, these are similar to models based on kriging (that similarly use covariance functions via semivariograms) and similar to MGWR, as the covariance function for each  $f_m(z)$  is individually calibrated to optimise model fit. Thus a key task in the GAM, in the context of SVCs, is to estimate parameters in each  $\kappa_m(\delta)$  and so estimate  $f_m(z)$ .

A GAM uses smooth functions of the predictor variables in which the values of  $y$  are assumed to have an exponential family distribution:

$$f(y|\theta) = h(y) \exp(\nu(\theta)T(x) - A(\theta)) \quad (5)$$

where  $h(\cdot)$ ,  $\nu(\cdot)$ ,  $T(\cdot)$  and  $A(\cdot)$  are known functions, and  $\theta$  is a vector of parameters. This very flexible form can represent a wide range of commonly used distributions such as Gaussian, Poisson, Gamma, or Binomial with a known number of trials. If

$$y = f(x) + \epsilon \quad (6)$$

where  $f$  is the function being modelled, then in GAMs a space of functions, or *basis*, is chosen of which  $f$  is some element, rather than more specifically assuming  $y$  to be some linear function of  $x$ . This allows the basic formula above to be expanded:

$$y = f(x) + \epsilon = \sum_{j=1}^d \kappa_j(x) \gamma_j + \epsilon \quad (7)$$

where each  $\kappa_j$  is a basis function of the transformed  $x$  and the  $\gamma$  are the corresponding regression coefficient estimates. One example of a basis is a GP basis. If there are  $n$  distinct geographical locations in the data set, knowing the locations

and the covariance function  $\kappa$  allows the variance-covariance function of the values of  $\beta_j$  in each location to be found, giving a variance-covariance matrix  $R$ . This can be translated into a set of  $n$  basis vectors  $\kappa_j(x)$  (Hefley *et al.* 2017), and the GAM can be calibrated in this way. Thus, in contrast to standard linear models, the predictors in a GAM include smooth non-linear functions of some or all of the covariates, which allows for non-linear relationships between the predictors and the target variable.

It is this local fitting that starts to hint at how GAMs can be used with geographic data, if the non-linear relationships are constructed over geographical space. This can be done by constructing splines that represent the GP parameterised by location as well as attribute space. Specifically a GP spline with location can be specified for each covariate  $j$ , although because the GP has a mean zero, a further constant  $\alpha_j$  is added as an offset, creating a spatially varying coefficient  $\beta_j(z)$ . In the standard 2D case,  $z_i = (u_i, v_i)$ . This is the Geographical Gaussian Process GAM (GGP-GAM) described in this paper.

### 3. A simulation case study

Simulated spatial data sets with varying degrees of regression relationship heterogeneity were used to examine the performance of the GGP-GAM and to compare that with the performance of a standard MGWR. The simulated data were created in a similar way to Fotheringham *et al.* (2017) and used subsequently by others (e.g. Fan and Huang (2022)), with the aim of simulating the coefficient estimates ( $\beta_j(z)$ 's) for Equation (8):

$$y_i = \alpha_0 + \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \epsilon_i \quad (8)$$

Three coefficient surfaces were created over a  $25 \times 25$  regular square grid. Each of the surfaces was assigned to the coefficients of the three predictor variables as follows:

$$\beta_{zero} = \beta_0(u, v) = 3 \quad (9)$$

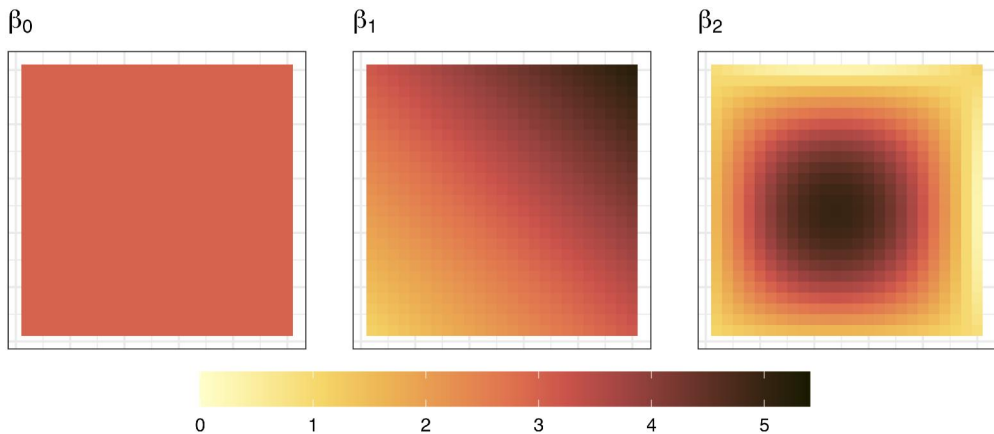
$$\beta_{low} = \beta_1(u, v) = 1 + \frac{1}{12}(u + v) \quad (10)$$

$$\beta_{high} = \beta_2(u, v) = 1 + \frac{1}{324} \left[ 36 - \left(6 - \frac{u}{2}\right)^2 \right] \left[ 36 - \left(6 - \frac{v}{2}\right)^2 \right] \quad (11)$$

The values for predictors  $x_1$  and  $x_2$  were generated from a normal distribution in the range  $[0, 1]$  (i.e.  $N(0.50, 0.17)$ ) and  $\epsilon$  from normal distribution in the range  $[0, 0.25]$  (i.e.  $\epsilon \sim N(0.04, 0.11)$ ), and 100 surfaces were generated. The values of the response  $y$  can be found directly from the simulated coefficients, the simulated predictor variables and the simulated error term. The simulated true regression coefficient surfaces are shown in Figure 1, which the analysis will seek to estimate using GGP-GAM and MGWR.

As discussed above, GP splines can be used within a GAM model. The `mgcv` R package (Wood 2015) was used to construct the GAM with GP splines with a GP smooth and for the simulation data, the spatial locations were extracted from  $(u, v)$  respectively. In this package, a parameter which controls the degree of smoothing of





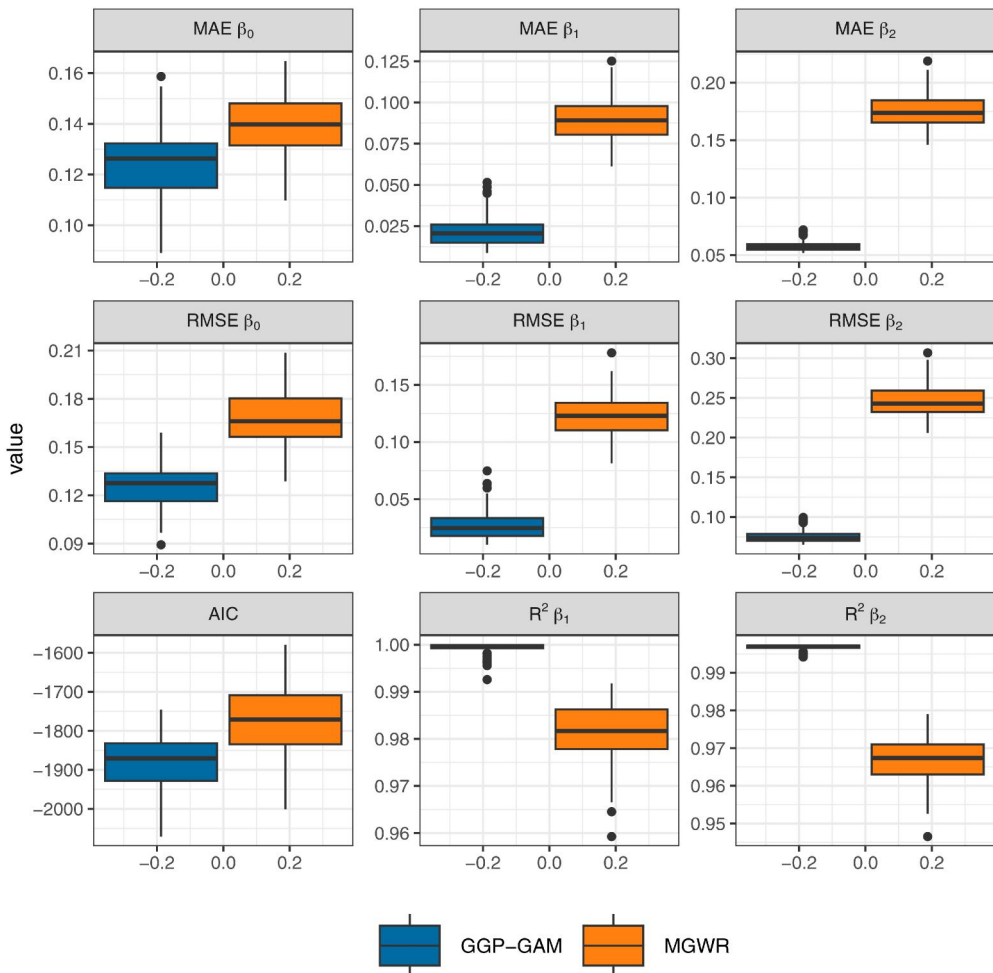
**Figure 1.** The simulated regression coefficient surfaces with zero, low and high spatial heterogeneity.

the data (via the GP's correlation function) is optimised and as such potentially indicates the locally varying nature of the coefficient estimate in a similar way to MGWR bandwidths. The GPs modelled in the GAM function all have a mean of zero, so for each covariate an extra fixed offset term is added along with the spatially smoothed terms. All of the MGWR models were fitted using the `GWmodel` R package (Lu *et al.* 2014; Gollini *et al.* 2015).

The simulated coefficient surfaces for  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  in Figure 1 and the 100 simulated values of  $x_1$ ,  $x_2$  and  $\epsilon$  were used to create a 100 sets of predictor variables over the 625 surface points,  $y$ . These values of true  $y$ ,  $x_0$ ,  $x_1$  and  $x_2$  were then used as inputs to the GGP-GAM model (i.e at 625 located observations with 4 fields) to generate coefficient estimates. For comparison, a MGWR was also undertaken with the same data. The spatially located coefficient estimates for both models were retained and used to generate measures of coefficient accuracy by comparing the model estimated coefficients with true ones in Figure 1, generating  $R^2$  for  $\beta_1$  and  $\beta_2$  (as  $\beta_0$  is stationary  $R^2$  cannot be computed) and RMSE and MAE for  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ . An assessment of model fit for predicted  $y$  ( $\hat{y}$ ) and true  $y$  for GGP-GAM and MGWR is found using AIC. This was done for each of the 100 sets of simulated  $y$ ,  $x_0$ ,  $x_1$  and  $x_2$ .

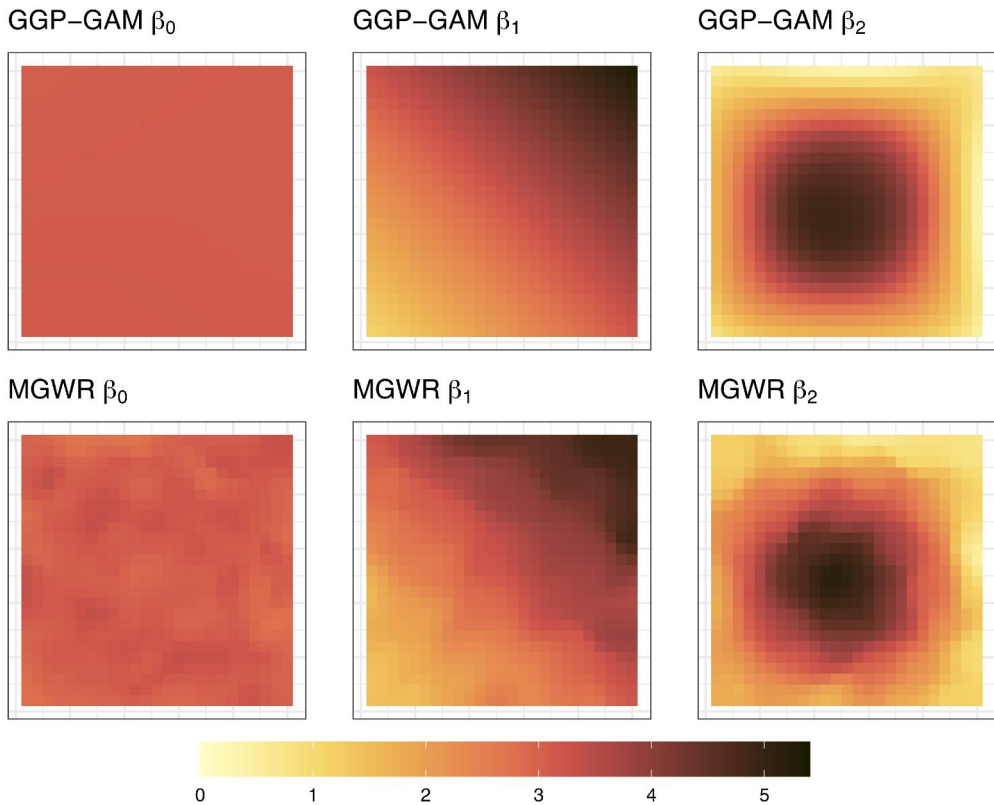
The results are shown in Figure 2. Under each fit measure, (AIC, RMSE, MAE,  $R^2$ ) the GPP-GAM generates more accurate estimates of the true coefficients than MGWR, with the difference in fit measures increasing with increasing degrees of spatial heterogeneity. The distributions of AIC values indicate the GGP-GAM to have consistently lowered AIC to that found with MGWR, indicating improved model parsimony and fit with GGP-GAM. This is also shown visually using an example set of  $\beta$  values from the 100 sets of coefficients estimated by the GGP-GAM and by the MGWR to recreate the surfaces for  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ . Figure 3 shows the coefficient surfaces estimated from the 4th set of simulations, with the same shading breaks as Figure 1. The better performance of the GGP-GAM in estimating the true  $\beta$ s is clearly demonstrated.

A final comparison concerns the estimated scale of process spatial heterogeneity. In a MGWR model this is indicated by the optimised parameter-specific bandwidths that



**Figure 2.** Evaluation of the accuracy of the GGP-GAM and MGWR regression coefficient estimates, when compared to the true coefficients, and with the distribution of AIC values for model fit.

explicitly describe the scale of the predictor-to-response relationship. In a GGP-GAM, this is indicated through the smoothing parameters (SPs) for each spline. Table 1 summarises the MGWR adaptive bandwidths and the GGP-GAM spline SPs for the 100 models using simulated data. What is interesting is the homogeneity of the bandwidths identified under the MGWR models and heterogeneity of the spline SPs values identified under the GGP-GAMs. However, this is related to the nature of the GAM SPs which are parameters in a covariance function, and thus not equivalent to a bandwidth. Here the GGP-GAM models were constructed using the `mgcv` package defaults. The result was that the number of knots,  $k$ , which defines the basis dimension in a spline, was automatically set at 33 in each GGP-GAM. Normally,  $k$  is iteratively determined (S. Wood 2015), as done in the empirical analysis below. Thus, the heterogeneity of the spline SPs in Table 1 is probably a result of under-fitting. This point is returned in the Discussion section.



**Figure 3.** The estimated regression coefficients from a single fit of the GGP-GAM and MGWR models, shaded using the same range as Figure 1.

**Table 1.** Summaries of MGWR bandwidths (BW) and GGP-GAM spline smoothing parameters (SP).

	Min.	Q1	Median	Q3	Max.
GGP-GAM SP $x_0$	9.2e-05	2.0e-02	2.3e-01	1.9e + 04	4.3e + 05
GGP-GAM SP $x_1$	3.8e-05	1.3e-02	1.9e + 02	1.3e + 03	1.1e + 05
GGP-GAM SP $x_2$	2.3e-07	2.5e-06	9.1e-06	1.7e-05	2.5e-05
MGWR BW $x_0$	10	10	10	10	13
MGWR BW $x_1$	22	25	30	46	73
MGWR BW $x_2$	10	13	13	25	26

## 4. An empirical case study: the UK Brexit

### 4.1. Data

A spatial analysis of the factors associated with the 2016 referendum on EU membership (Brexit) was used to empirically illustrate the proposed GGP-GAM approach and to compare it with MGWR. Census and voting data were obtained from the `parlitoools` R package (Odell 2020). This includes data on voting for the 632 Westminster parliamentary constituencies in Great Britain (i.e. excluding those in Northern Ireland, for which some data is missing). Information on the data and the package can be found at <https://docs.evanodell.com/parlitoools/>. The voting data were linked to socio-

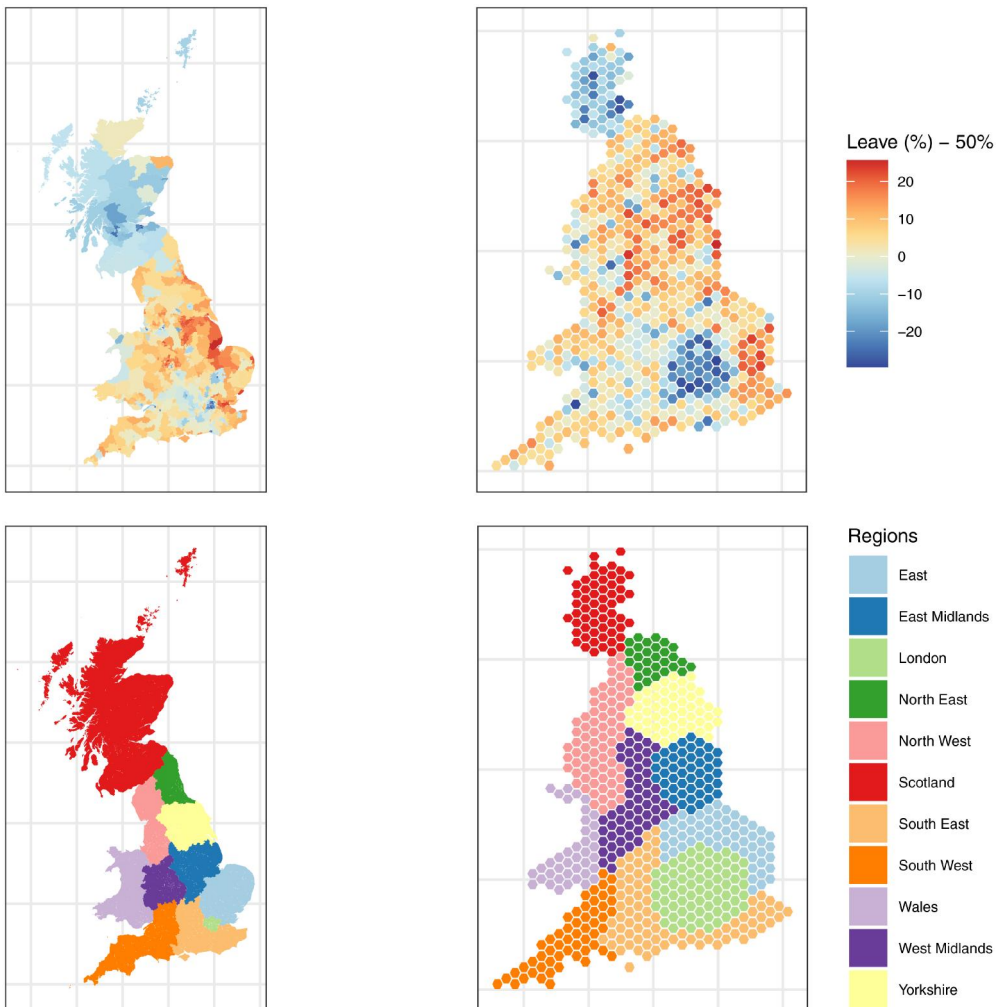
**Table 2.** The variables used to construct the GGP-GAM of leave voting rates, for each parliamentary constituency.

Variable	Description
Leave	Leave vote (%)
to15	Turn out in the 2015 general election (%)
over65	Over 65s (%)
ttu	Transport trade and utilities employment (%)
lhosp	Leisure & hospitality employment (%)
manu	Manufacturing employment (%)
Degree	Degree (level 4) qualification (%)
badhealth	Bad health (%)
bornuk	UK born (%)

economic factors that were associated with the Leave vote as described in Beecham *et al.* (2020) and contained in the `parlitoools` R package. A number of variables were chosen to illustrate the SVC model using a GGP-GAM relating to structural economic variables, age, and employment sector. These are summarised in Table 2. The constituency areas with an OSGM projection (<https://epsg.io/27700>) are available from the UK Office of National Statistics<sup>1</sup> and the hexagonal cartograms are contained in the `parlitoools` R package (Odell 2020). The response variable and the UK regions for reference are shown over both spatial structures in Figure 4. The predictor variables over the hexagonal cartograms (hex-bins) are shown in Figures 4 and 5. The GGP-GAM and MGWR spatial analyses used the original untransformed data and the coordinates from the OSGB projected data, with the results displayed using the hex-bins.

#### 4.2. A GGP-GAM analysis

Spatially varying coefficient models with the GGP-GAM were undertaken using the OSGB projected parliamentary constituency in Figure 4. The geometric centroids of each parliamentary constituency were extracted to generate  $X$  and  $Y$  (Easting and Northing) variables in kilometres. A GGP-GAM was specified. Recall that the splines optimise a parameter which controls the degree of smoothing and that the GPs modelled in the GAM function all have a mean of zero, so for each covariate an extra fixed offset term was added along with the spatially smoothed terms. In this case the GGP-GAM was fully tuned by specifying a sufficiently high number of knots through the  $k$  parameter, in contrast to the GGP-GAMs applied to the simulated data above for which the defaults were used ( $k = 33$ ). The value of  $k$  sets an upper limit on effective degrees of freedom associated with GP spline smooth and thus the splines bases. Higher values of  $k$  also increase computation time. However the choice of  $k$  is difficult to determine algorithmically or to optimise and a typical approach is to investigate different values of  $k$  and adjust up or down as need. As Simon Wood notes in the `mgcv` R package manual under the `choose.k` help page, '[the] exact choice of  $k$  is not generally critical: it should be chosen to be large enough that you are reasonably sure of having enough degrees of freedom to represent the underlying 'truth' reasonably well, but small enough to maintain reasonable computational efficiency' (Wood 2015). Here, the GGP-GAM was specified with a  $k = 60$ . This was large enough to accommodate the GGP-GAM parametric and spline parameters, and to ensure sufficient degrees of freedom. A summary of the fixed parametric coefficient

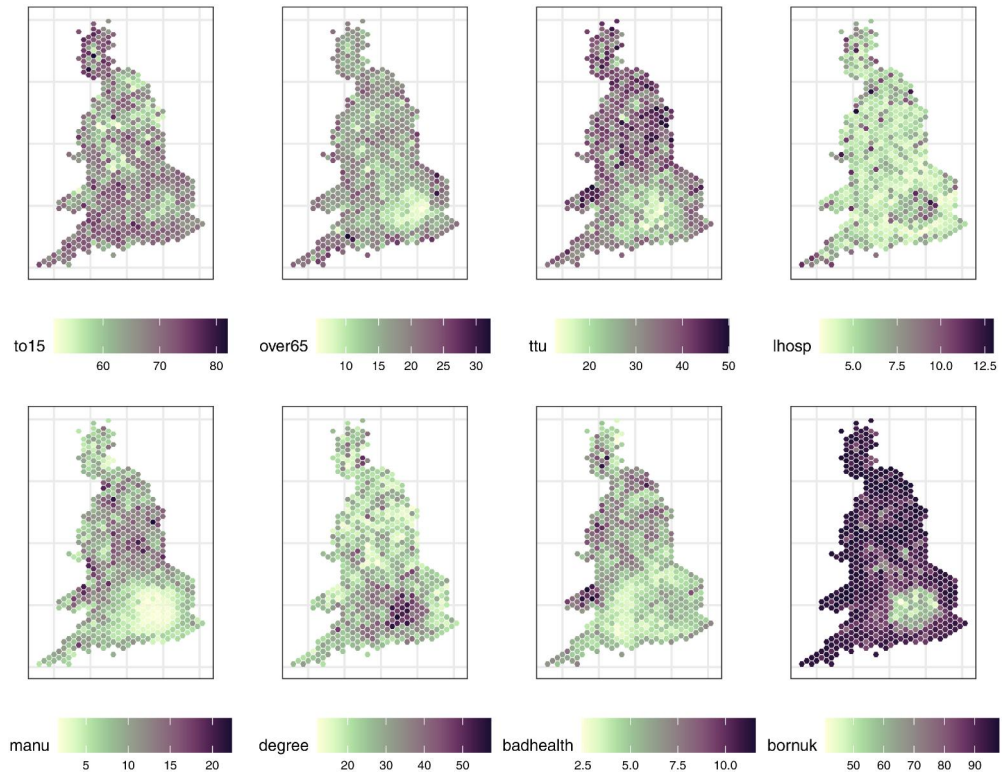


**Figure 4.** The Brexit leave vote majority over parliamentary constituencies with the UK regions, under an OSGB projection (left) and recast into hexagons (right).

estimates, the smooth terms of the GGP splines for the GGP-GAM model and the distribution of the spatial varying coefficient estimates are shown in Tables 3–5, respectively.

The GGP-GAM fixed (parametric) coefficient estimates in Table 3 can be interpreted as linear terms in the model, in a similar way as the results of a standard linear regression model. Here, only the Intercept and the covariate for the election 2015 turn out (*Turn out (2015)*) are globally significant.

Table 4 summarises the GGP-GAM smooth terms (GP splines) and the spline SPs. The SPs determine the ‘wiggleness’ of the SVCs and are controlled by the number of knots ( $k$ ). Together these control the characterisation of the response-to-predictor relationships over space. The spline SPs are parameters in a covariance function that penalise the complexity of the model and  $k$  defines the basis dimension in a spline (also referred to as a smooth), and are generally indicative of the degree of variation in the predictor to response spatial relationship. Table 4 indicates that the SP is



**Figure 5.** The predictor variables over parliamentary constituencies, recast into hexagons.

**Table 3.** The fixed (parametric) coefficients for the GGP-GAM.

	Estimate	Std. error	t Value	Pr(> t )
Intercept	63.566	6.087	10.443	0.000
Turn out (2015)	-0.096	0.026	-3.732	0.000
Over 65s (%)	0.192	0.869	0.221	0.825
Transport & utilities employment (%)	0.085	0.920	0.092	0.927
Leisure & hospitality employment (%)	3.067	8.552	0.359	0.720
Manufacturing employment (%)	0.861	3.974	0.217	0.829
Degree qualification (%)	-1.583	2.591	-0.611	0.542
Bad health (%)	-0.858	2.276	-0.377	0.706
UK born (%)	-0.029	1.233	-0.024	0.981

**Table 4.** The smooth terms of the GGP-GAM model, with the Spline smoothing parameters.

	Spline SPs	edf	Ref.df	F	p-Value
s(X,Y):Intercept	3.24e-13	4.108	4.121	4.367	0.002
s(X,Y):Turn out (2015)	8.71e+01	2.500	2.500	6.351	0.051
s(X,Y):Over 65s (%)	1.59e-03	4.386	5.327	0.354	0.851
s(X,Y):Transport & utilities employment (%)	1.41e-03	6.572	7.982	0.746	0.653
s(X,Y):Leisure & hospitality employment (%)	1.48e-05	11.607	13.499	1.584	0.073
s(X,Y):Manufacturing employment (%)	7.13e-05	9.966	12.155	1.109	0.351
s(X,Y):Degree qualification (%)	1.62e-04	16.926	20.090	1.679	0.033
s(X,Y):Bad health (%)	2.35e-04	3.810	4.436	1.176	0.271
s(X,Y):UK born (%)	7.21e-04	31.584	37.638	1.444	0.050

**Table 5.** The distributions of the GGP-GAM spatially varying coefficient estimates (MAE = 1.794 Adjusted  $R^2 = 0.945$ ; AIC = 3131.3).

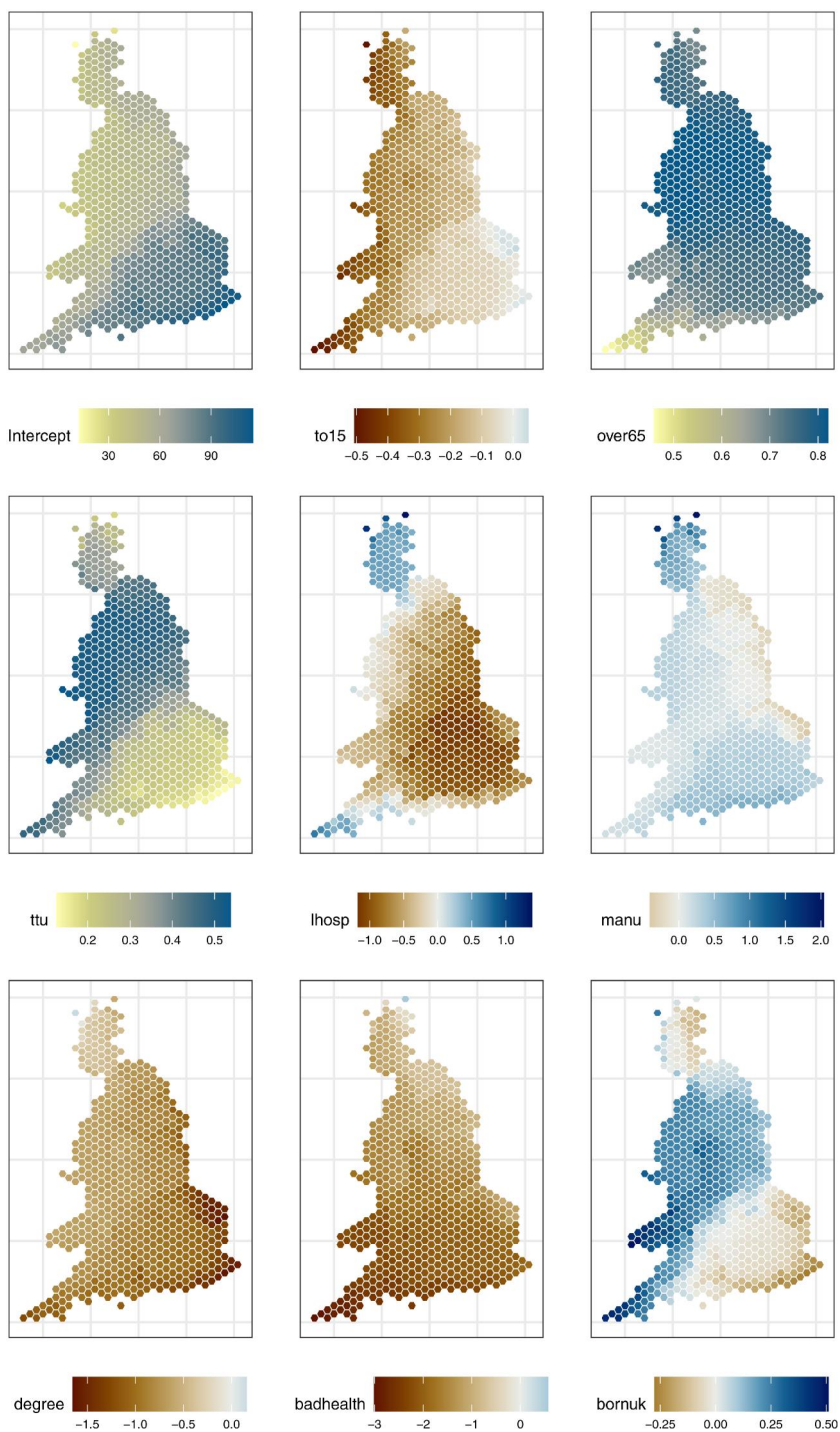
	Min.	1st Qu.	Median	3rd Qu.	Max.
Intercept	12.80	48.65	56.99	82.64	114.52
Turn out (2015)	-0.51	-0.27	-0.19	-0.10	0.05
Over 65s (%)	0.46	0.73	0.76	0.81	0.82
Transport & utilities employment (%)	0.13	0.24	0.38	0.47	0.54
Leisure & hospitality employment (%)	-1.17	-0.91	-0.61	-0.19	1.39
Manufacturing employment (%)	-0.41	0.07	0.20	0.35	2.04
Degree qualification (%)	-1.65	-0.99	-0.82	-0.68	0.16
Bad health (%)	-3.01	-2.04	-1.72	-1.25	0.58
UK born (%)	-0.28	-0.03	0.11	0.23	0.51

relatively small for the Intercept ( $10^{-13}$ ), relatively large for Turn out (2015) ( $10^1$ ) and between  $10^{-3}$  and  $10^{-5}$  for the other covariates. The effective degrees of freedom (edf in the table) can be interpreted as an indicator of the non-linearity of the relationship of the response (leave vote) with the predictor variables - i.e. the effective degrees of freedom would have a value of 1 if the model penalized the smooth term to a near linear relationship. Higher edf values indicate increasing non-linearity in the relationship over a 2-dimensional space defined by  $(X, Y)$ . All edf values are less than  $k$  indicating that it is adequately specified. The p-values relate to the splines defined over geographic space and their significance can be interpreted as indicating whether they vary locally. The p-values indicate that the Intercept, Degree qualification (%) and UK born (%) are significant locally. The insignificant p-values can be interpreted as indicating that although there are effects, these do not vary locally.

Table 5 summarises the distribution of the GGP-GAM SVCs. Most flip between being positive in some constituencies and negative in others, except for the Intercept, Over 65s (%) and Transport & utilities employment (%), which are consistently positively associated with the leave vote. However the interpretation of these should be alongside the interpretation of the fixed parameters and the spline smooth terms: the Intercept is significant globally and locally; Turn out (2015) is globally significant with Degree qualification (%) and UK born (%) significant locally.

It also possible to map the spatial variation in the coefficient estimates generated by the GGP-GAM as in Figure 6. The spatial trends in the coefficients show a number of spatial patterns associated with different population characteristics and the leave vote share. The 2015 Turn out is positive associated in the East; Over 65s is positively associated except in the South West; Transport and utilities employment is positively associated except in the South East and London; Leisure and hospitality is negatively associated except in Scotland and the South West; Manufacturing is positively associated except on the East, East Midlands and North East; Degree and Bad Health are negatively associated nearly everywhere; and Born UK is positively associated except in the South East, London and Scotland. So overall, for some covariates, there are some distinct differences between Scotland, the East, South East (including London) and the parts of the South West with the rest of the country, but in different directions for different covariates, suggesting how individual factors were interacting in different ways in these areas in relation to the leave vote.





**Figure 6.** The spatially varying coefficient (SVC) estimates from GGP-GAM.



**Table 6.** Comparison of the model fit diagnostics for the GGP-GAM and MGWR models.

	<i>R</i> -squared	MAE	AIC
GGP-GAM	0.953	1.794	3131.3
MGWR	0.966	1.582	2927.1

**Table 7.** The distributions of the MGWR spatially varying coefficient estimates and the adaptive bandwidths.

	Min.	1st Qu.	Median	3rd Qu.	Max.	Bandwidths
Intercept	27.202	48.550	63.901	106.228	130.530	18
Turn out (2015)	-0.078	-0.077	-0.075	-0.073	-0.070	630
Over 65s (%)	0.398	0.401	0.406	0.415	0.445	630
Transport & utilities employment (%)	-1.448	-1.167	-0.342	-0.224	0.088	88
Leisure & hospitality employment (%)	-2.433	-0.681	0.022	0.609	2.704	22
Manufacturing employment (%)	-0.291	0.278	0.425	1.415	2.012	65
Degree qualification (%)	-1.681	-1.669	-1.345	-1.209	-1.090	313
Bad health (%)	-0.710	-0.462	-0.178	0.365	0.398	295
UK born (%)	0.107	0.116	0.374	0.416	0.468	192

### 4.3. A MGWR analysis

Finally it also is possible to compare the GGP-GAM results with those from a MGWR. Summaries of model fit and accuracy are shown in Table 6. Comparing the diagnostics of AIC, adjusted  $R^2$ , and MAE for the GGP-GAM and MGWR models indicates that the MGWR model exhibits marginally better metrics: broadly they seem to be performing as well as each other. Table 7 summarises the MGWR SVCs and the parameter-specific adaptive bandwidths. It is instructive to examine the variation in these and to compare with the GGP-GAM SVCs in Table 5. Comparing the SVC inter-quartile ranges indicates that both models exhibit similar levels of variation for many of the covariates, in terms of their sign and range, except for the following:

- Transport & utilities employment (%): negative in MGWR, positive in GGP-GAM.
- Leisure & hospitality employment (%): flips sign in MGWER, negative in GGP-GAM.
- Bad Health (%): flips sign in MGWER, negative in GGP-GAM.
- UK born (%): positive in MGWR, flips sign in GGP-GAM.

This suggest that in general the inference (understanding) supported by each model is largely similar when the associations between the target variable and individual covariates are considered, but with some potentially important differences as described above.

## 5. Discussion and conclusion

Spatially varying coefficient (SVC) models explicitly accommodate process spatial non-stationarity, where statistical relationships expressed using regression coefficient estimates are allowed to vary with location. SVCs provide an explicit representation of process spatial heterogeneity and can be mapped to describe how and where statistical relationships vary. These in turn provide insights into variations in the underlying socio-economic, environmental or other kind of geographical process, and support the spatial 'detective work' that is core activity in much spatial and geographical analysis.

This paper proposes an alternative SVC based on GAMs with Gaussian Process splines parameterised with location - a Geographical GP GAM (GGP-GAM). It was applied to 100 simulated data generated from coefficients with known spatial heterogeneities and was shown to out-perform Multiscale Geographically Weighted Regression (MGWR) under a range of fit metrics. Then both models were applied to an empirical case study of the UK Brexit referendum. Here the MGWR model was shown to (marginally) outperform the GGP-GAM. Comparisons of the SVCs arising from both models indicated similar levels of variation for many of the covariates, with few differences in the predictor-to-response relationships. The correlations between the GGP-GAM and MGWR coefficients are summarised in [Table 8](#) with two covariates indicating different relationships. In general, however, this reinforces an emerging theme from the SVC literature that alternative SVC approaches generate results that are much more similar than they are different, indicating a general coherence across this family of models and supporting SVC modelling overall. The GGP-GAM covariates were mapped and these indicated some strong regional differences in their relationship with the leave vote, for specific variables in different parts of the study area, when compared to the rest of the country ([Figure 6](#)). These are reflected the spatial variation in the associations between socio-economic factors the process and the Brexit leave vote.

In terms of historic frequency of use, the SVC model ‘brand leader’ is Geographically Weighted Regression. One could say that SVC modelling is to ‘vacuum cleaner’ as GWR is to ‘Hoover’. More specifically Multiscale GWR is becoming the approach of choice. The increase in GWR and MGWR applications over recent years (as documented in [Comber \*et al.\* \(2022\)](#)) has in part been driven by the conceptual simplicity of GWR and MGWR: they generate a collection of local regression models, which are readily understood by both expert and non-expert users. However there are number of conceptual limitations to GWR-based approaches. Coefficients are estimated as though calibrating multiple models rather than a single one, whereas, for example, GP-based and ESF-based approaches within GLM or GAM frameworks provide a full single data model. Additionally, GW-based regressions may be sensitive to local collinearity even when this is not a problem across the whole dataset ([Wheeler and Páez 2010](#); [Páez \*et al.\* 2011](#)), although tools for GWR have been developed to compensate for this (see [Gollini \*et al.\* \(2015\)](#)). Finally, kernel based approaches like GWR and particularly MGWR are computationally demanding, as they require a series of

**Table 8.** The correlations between the GGP-GAM and MGWR spatially varying coefficient estimates.

Covariate	Correlation
Intercept	0.877
Turn out (2015)	-0.661
Over 65s (%)	0.230
Transport & utilities employment (%)	0.871
Leisure & hospitality employment (%)	0.314
Manufacturing employment (%)	0.437
Degree qualification (%)	0.754
Bad health (%)	-0.675
UK born (%)	0.798

optimisation problems to be solved to determine the individual scales of predictor-to-response relationships (Fan and Huang 2022).

In contrast, the proposed GGP-GAM has a rich and more comprehensive theoretical underpinning through its use of GP splines set within a GAM framework that has much greater flexibility than MGWR (especially as many GWR extensions have not been developed for MGWR). For example, GAM frameworks are relatively mature, and GAM functions commonly provide options for penalized terms for handling collinearity, robust options for treating outliers, and options for heteroskedastic or autocorrelated error terms. Further, MGWR is still only for Gaussian responses, despite being developed in 2014, whereas GGP-GAM can be used for a variety of response types and has much stronger theoretical background to support such 'within framework' extensions than GWR/MGWR. The one area where MGWR has an inferential advantage is in the way that the scales of predictor-to-response variable relationships are reported. In MGWR, the parameter specific bandwidths provide an explicit and intuitively understood indication of process spatial heterogeneity, which GGP-GAM outputs currently lack. The nearest similar measure is the GGP-GAM spline SPs as in Table 1 but these describe parameters in a covariance function whose stability is linked to the specification of the knots parameter,  $k$ , and are *not* equivalent to a bandwidth. They do however provide a measure of the spatial complexity of the predictor to response relationship. Wolf *et al.* (2018) compared MGWR and multiscale Bayesian SVC (BSVC) models and sought to unpick the conceptual differences between the bandwidths reported by the two approaches. Specifically, they were interested in expounding how the different bandwidth distances indicated process 'scale' and identifying some important ontological differences between them. Their summary was that the MGWR bandwidths describe the local heterogeneity of predictor to response relationships and BSVC bandwidths (ranges) provided a composite measure of predictor to response relationships heterogeneity combined with local variation. Understanding the scale of spatial variation in this way is a key geographic consideration. However, as with BSVC, weights in GGP-GAM will not decay uniformly from a calibration point as in MGWR, since accounting for spatial correlation in a BSVC or a GGP-GAM leads to desirable weighting effects related to the data's spatial configuration (Chiles and Delfiner 1999). In this respect, the spatial configuration of a dataset will directly affect any comparison between a GGP-GAM and MGWR (see also Harris *et al.* (2011) in the context of GWR versus kriging). Future work will explore these issues in depth and how to characterise the link between the spline SPs through the GP covariance function in order to increase the interpretability of process scale through the GGP-GAM. Specifically, it will seek to connect the spline SPs to an inference of geographic scale. A final observation is the potential for SVCs via a GGP-GAM to overcome some of the long-standing reservations about GWR-based regression methods and to elevate the perception of SVCs amongst the broader community. Typically, GWR studies are presented as informal, exploratory analyses, and thus of implicitly lower status, supporting informal inferences of observed spatial processes. The mature GAM framework offers a potential opportunity to elevate this class of work and maybe incentivise greater mainstream use of SVCs. Future work will provide a more thorough demonstration of how GGP-GAMs and their

outputs can be explored, tuned and reported, potentially using the diagnostic tools proposed in Sachdeva *et al.* (2022).

There are a number of other areas of further work that will be undertaken. The immediate next steps are to explore how the GGP-GAM handles highly complex spatial heterogeneities and apply it to larger datasets. Then a wrapper package for undertaking GGP-GAMs and extracting the coefficients at each locations will be constructed to sit above the `mcgv` package. Currently, a number of operations are needed to set up the GGP-GAMs and to extract the outputs in a form useful to a person interpreting a SVC model. Although R code can be written to do this, providing a simpler interface for frequent use would greatly increase accessibility. It is anticipated that there will be a number of further methodological developments. First, investigations will be undertaken to determine optimal spline smoothing parameters, which control the smoothing of the data and as such indicates the locally varying nature of the coefficient, and the number of knots used to ensure sufficient degrees of freedom across the data and the splines. In this paper, defaults were used, and the results are encouraging. However this suggests that potentially even better results could be achieved if the rubrics for tuning the GGP-GAP could be established. Second, further work will investigate how to provide explicit descriptions of the scale of predictor-to-response variable relationships. An approximation of these can be extracted from the semivariograms of the GP spline smoothing parameters which can be constructed for each GP spline. However, there is some uncertainty in the interpretation of the semivariograms in the `mcgv` package, in which the GGP-GAMs were constructed. Third, GAMs have options for penalised terms in order to handle predictor variable collinearity and selection. Investigations are needed to determine how and whether this needs to be tested for and then specified. Fourth, the wrapper package will include options to deal with outliers and dependent errors (i.e. heteroskedascity, autocorrelation) will also be investigated for GGP-GAMs. There are also a number of areas of SVC model development, particularly in the temporal domain. GP splines can be parameterised with time as well as as location. This suggests opportunities for methodological entrancement to current Geographically and Temporally Weighted Regression (GTWR) approaches (Huang *et al.* 2010; Fotheringham *et al.* 2015) and multiscale GTWR (Wu *et al.* 2019), which effectively determine parameter ( $\rho$ ) to link space and time. These, and other methodological opportunities, are possible because of the considerable developments that have taken place around GAMs in the temporal, spatial and spatio-temporal domains.

In conclusion, this paper demonstrates for the first time, the formulation and application of a GAM with GP splines calibrated via observation location in a multiscale spatially varying coefficient model. The Geographical Gaussian Process GAM (GGP-GAM) was applied to 100 simulated datasets with known variations in spatial heterogeneity, and were shown consistently to perform better than MGWR. GGP-GAM was applied to a Brexit case study to demonstrate how process GPP-GAMs could be used to understand how and where relationships (and therefore processes) vary spatially. The results were compared with MGWR which was shown to perform marginally better. A number of areas of further work have been identified, including the development of a user-friendly wrapper to the GGP-GAM functions that will incorporate spline

tuning, explicit descriptions of the scale of predictor-to-response variable relationships, handling variable collinearity, model selection (penalization), robust regression and handling error term dependencies. These are supported by the richer and more comprehensive theoretical framework of GAMs.

## Note

1. <https://geoportal.statistics.gov.uk/datasets/d3650c649bff40768badafdf0518a4af>.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was by Natural Environment Research Council [NE/S009124/1] and Biotechnology and Biological Sciences Research Council [BB/X010961/1 and BBS/E/RH/23NB0008].

## Notes on contributors

**Alexis Comber** is Professor of Spatial Data Analytics, with research interests in all areas of spatial analysis and geocomputation. His recent work explores methods for handling spatial scale and for supporting different scales of decision making.

**Paul Harris** is Professor of Agriculture Sciences with interests in environmental monitoring and modelling, and methodologies in spatial statistics. His recent research focuses on developing analytical methods for resilient farming systems.

**Chris Brunsdon** is Professor of Geocomputation at the National University of Ireland, Maynooth. He has research interests in developing spatial data analysis methods, and reproducibility and inference for these methods. His recent research has been applying these ideas to data related to Covid-19.

## ORCID

Alexis Comber  <http://orcid.org/0000-0002-3652-7846>

Paul Harris  <http://orcid.org/0000-0003-0259-4079>

Chris Brunsdon  <http://orcid.org/0000-0003-4254-1780>

## Data and codes availability statement

The data and code used to undertake the full GGP-GAM analysis and generate the results presented in this paper can be found at <https://github.com/lexcomber/GGP-GAM>.

## References

- Beecham, R., Williams, N., and Comber, A., 2020. Regionally-structured explanations behind area-level populism: an update to recent ecological analyses. *PLOS One*, 15 (3), e0229974.
- Brunsdon, C., Fotheringham, A.S., and Charlton, M., 1999. Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, 39 (3), 497–524.

- Brunsdon, C., Fotheringham, A.S., and Charlton, M.E., 2010. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, 28 (4), 281–298.
- Chiles, J.P., and Delfiner, P., 1999. *Geostatistics: modeling spatial uncertainty*. London: John Wiley & Sons.
- Comber, A., et al., 2021. The GWR route map: a guide to the informed application of GWR. Arxiv <https://arxiv.org/abs/2004.06070>.
- Comber, A., et al., 2023. A route map for successful applications of geographically weighted regression. *Geographical Analysis*, 55 (1), 155–178.
- Comber, A., et al., 2022. Gwverse: a template for a new generic geographically weighted R package. *Geographical Analysis*, 54 (3), 685–709.
- Dambon, J.A., Sigrist, F., and Furrer, R., 2021. Maximum likelihood estimation of spatially varying coefficient models for large data with an application to real estate price prediction. *Spatial Statistics*, 41, 100470.
- Fahrmeir, L., et al., 2022. Regression models. In: *Regression: models, methods and applications*. Berlin, Heidelberg: Springer, 23–84.
- Fan, Y.T., and Huang, H.C., 2022. Spatially varying coefficient models using reduced-rank thin-plate splines. *Spatial Statistics*, 51, 100654.
- Finley, A.O., 2011. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. *Methods in Ecology and Evolution*, 2 (2), 143–154.
- Finley, A.O., and Banerjee, S., 2020. Bayesian spatially varying coefficient models in the spBayes R package. *Environmental Modelling & Software*, 125, 104608.
- Fotheringham, A.S., Crespo, R., and Yao, J., 2015. Geographical and temporal weighted regression (GTWR). *Geographical Analysis*, 47 (4), 431–452.
- Fotheringham, A.S., and Oshan, T.M., 2016. Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*, 18 (4), 303–329.
- Fotheringham, A.S., Yang, W., and Kang, W., 2017. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*, 107 (6), 1247–1265.
- Gelfand, A.E., et al., 2003. Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98 (462), 387–396.
- Gollini, I., et al., 2015. GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, 63 (17), 1–50.
- Griffith, D.A., 2008. Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environment and Planning A: Economy and Space*, 40 (11), 2751–2769.
- Harris, P., 2019. A simulation study on specifying a regression model for spatial data: choosing between autocorrelation and heterogeneity effects. *Geographical Analysis*, 51 (2), 151–181.
- Harris, P., Brunsdon, C., and Fotheringham, A.S., 2011. Links, comparisons and extensions of the geographically weighted regression model when used as a spatial predictor. *Stochastic Environmental Research and Risk Assessment*, 25 (2), 123–138.
- Harris, P., et al., 2010. The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. *Mathematical Geosciences*, 42 (6), 657–680.
- Hastie, T.J., and Pregibon, D., 2017. Generalized additive models. In: J.M. Chambers and T.J. Hastie eds. *Statistical models in S*. Boca Raton: Chapman & Hall, 195–247.
- Hefley, T.J., et al., 2017. The basis function approach for modeling autocorrelation in ecological data. *Ecology*, 98 (3), 632–646.
- Huang, B., Wu, B., and Barry, M., 2010. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, 24 (3), 383–401.
- Kammann, E., and Wand, M.P., 2003. Geoadditive models. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 52 (1), 1–18.
- Kelsall, J.E., and Diggle, P.J., 1998. Spatial variation in risk of disease: a nonparametric binary regression approach. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 47 (4), 559–573.

- Kim, H., and Lee, J., 2017. Hierarchical spatially varying coefficient process model. *Technometrics*, 59 (4), 521–527.
- Li, Z., et al., 2019. Fast geographically weighted regression (FastGWR): a scalable algorithm to investigate spatial process heterogeneity in millions of observations. *International Journal of Geographical Information Science*, 33 (1), 155–175.
- Li, Z., et al., 2020. Measuring bandwidth uncertainty in multiscale geographically weighted regression using Akaike weights. *Annals of the American Association of Geographers*, 110 (5), 1500–1520.
- Lu, B., et al., 2019. A response to 'A comment on geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, 33 (7), 1300–1312.
- Lu, B., et al., 2014. The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-Spatial Information Science*, 17 (2), 85–101.
- Lu, B., et al., 2022. High-performance solutions of geographically weighted regression in R. *Geo-Spatial Information Science*, 25 (4), 536–549.
- McCullagh, P., and Nelder, J.A., 2019. *Generalized linear models*. London: Routledge.
- Mu, J., Wang, G., and Wang, L., 2018. Estimation and inference in spatially varying coefficient models. *Environmetrics*, 29 (1), e2485.
- Murakami, D., and Griffith, D.A., 2015. Random effects specifications in eigenvector spatial filtering: a simulation study. *Journal of Geographical Systems*, 17 (4), 311–331.
- Murakami, D., et al., 2019. The importance of scale in spatially varying coefficient modeling. *Annals of the American Association of Geographers*, 109 (1), 50–70.
- Murakami, D., et al., 2017. A Moran coefficient-based mixed effects approach to investigate spatially varying relationships. *Spatial Statistics*, 19, 68–89.
- Nelder, J.A., and Wedderburn, R.W., 1972. Generalized linear models. *Journal of the Royal Statistical Society: Series A*, 135 (3), 370–384.
- Odell, E., 2020. Package 'parlitoools. R Package Version, 12. <https://cran.r-project.org/web/packages/parlitoools/index.html>
- Openshaw, S., 1996. Developing GIS-relevant zone-based spatial analysis methods. In: P. Longley and M. Batty, eds. *Spatial Analysis: Modelling in a GIS Environment*. Cambridge: Geoinformation International, 55–73.
- Páez, A., Farber, S., and Wheeler, D., 2011. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environment and Planning A: Economy and Space*, 43 (12), 2992–3010.
- Sachdeva, M., et al., 2022. Are we modelling spatially varying processes or non-linear relationships? *Geographical Analysis*, 54 (4), 715–738.
- Wheeler, D.C., and Calder, C.A., 2007. An assessment of coefficient accuracy in linear regression models with spatially varying coefficients. *Journal of Geographical Systems*, 9 (2), 145–166.
- Wheeler, D.C., and Páez, A., 2010. Geographically weighted regression. In: M. Fischer and A. Getis, eds. *Handbook of applied spatial analysis*. Berlin, Heidelberg: Springer, 461–486.
- Wolf, L.J., Oshan, T.M., and Fotheringham, A.S., 2018. Single and multiscale models of process spatial heterogeneity. *Geographical Analysis*, 50 (3), 223–246.
- Wood, S., 2015. Package 'MGCV. R Package Version, 1 (29), 729. <https://cran.r-project.org/web/packages/mgcv/index.html>
- Wood, S.N., 2006. *Generalized additive models: an introduction with R*. Boca Raton: Chapman Hall/CRC.
- Wu, C., et al., 2019. Multiscale geographically and temporally weighted regression: Exploring the spatiotemporal determinants of housing prices. *International Journal of Geographical Information Science*, 33 (3), 489–511.
- Yang, W., 2014. *An extension of geographically weighted regression with flexible bandwidths*. Thesis (PhD). University of St Andrews.
- Yu, H., et al., 2020. On the measurement of bias in geographically weighted regression models. *Spatial Statistics*, 38, 100453.