

Accommodating Space-Time Scaling Issues in GAM-Based Varying Coefficient Models

Alexis Comber¹  

School of Geography, University of Leeds, UK

Paul Harris  

Sustainable Agriculture Sciences, Rothamsted Research, Harpenden, UK

Chris Brunsdon  

National Centre for Geocomputation, National University of Ireland, Maynooth, Ireland

Abstract

The paper describes modifications to spatial and temporal varying coefficient (STVC) modelling, using Generalized Additive Models (GAMs). Previous work developed tools using Gaussian Process (GP) thin plate splines parameterised with location and time variables, and has presented a space-time toolkit in the `stgam` R package, providing wrapper functions to the `mgcv` R package. However, whilst thin plate smooths with GP bases are acceptable for working with spatial problems they are not for working with space *and* time combined. A more robust approach is to use a tensor product smooth with GP basis. However, these in turn require correlation function length scale or range parameters (ρ) to be defined. These are distances (in space or time) at which the correlation function falls below some value, and can be used to indicate the scale of spatial and temporal dependencies between response and predictor variables (similar to geographically weighted bandwidths). The paper describes the problem in detail, illustrates an approach for optimising ρ and methods for determining model specification.

2012 ACM Subject Classification Mathematics of computing → Regression analysis; Information systems → Spatial-temporal systems; Mathematics of computing → Spline models

Keywords and phrases Spatial Analysis, Spatiotemporal Analysis

Digital Object Identifier 10.4230/LIPIcs.GIScience.2025.15

Funding This research was supported by UKRI (ESRC) funding ES/Y006259/1 under the Digital Footprints scheme.

1 Introduction

Previous research has described the use of Generalized Additive Models (GAMs) [13, 12] with Gaussian Process (GP) smooths as an approach for spatially varying coefficient (SVC) modelling [5, 6]. The proposed geographical GP-GAM has been shown to have all of the advantages of the SVC brand leader, geographically weighted regression (GWR) [2] and its multiscale variant (MGWR) [18, 15, 11] in modelling and capturing any spatial dependencies between the target variable and individual predictor variables (hence *multiscale*), and none of the disadvantages: GWR-based approaches are subject to local collinearity, they generate a collection of local models rather than a single one, until recently MGWR was only specified for Gaussian responses, MGWR is unable to support out of sample predictions and all GWR-based approaches re-use individual observations in multiple local models. Ideas extending the use of GAMs with GPs for SVC model construction into space-time analyses for *spatial and temporally* varying coefficient (STVC) models have been proposed [4] and, at the time of writing, are currently under review [7]. In parallel the `stgam` package [8] was developed to

¹ Corresponding author



© Alexis Comber, Paul Harris, and Chris Brunsdon;
licensed under Creative Commons License CC-BY 4.0

13th International Conference on Geographic Information Science (GIScience 2025).

Editors: Katarzyna Sila-Nowicka, Antoni Moore, David O'Sullivan, Benjamin Adams, and Mark Gahegan;
Article No. 15; pp. 15:1–15:9



Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

provide a framework for informed SVC and STVC model construction, through functions that wrapped around the `mgcv` packages `gam()` function [17] to fit a range of spatial, temporal, and spatiotemporal varying coefficient models, to investigate the nature of any space-time dependencies present in the data and to inform SVC and STVC model specification.

However, we have subsequently become aware of a number of problems with simply extending the `stgam` approach to SVC modelling to STVC models. This paper describes these, some potential solutions and their potential implementation in a revised `stgam` package.

2 Background

2.1 The original big idea

Increasingly the space-time data we analyse and use are *not* collected under some grand over-arching experimental design, nor for the purposes we intend to use them for. As such, the big idea behind `stgam` workflows is that it is naive to construct models that make assumptions about the presence and nature of data space-time dependencies, whether for the purposes of prediction or inference (process understanding). Instead these need to be explicitly examined and the most appropriate model form determined. This position is in contrast to a classic statistics perspective where data are considered to be a realisation of carefully considered data collection activity, constructed in such a way as to allow specific hypothesis to be tested.

In this context, most widely used approaches for capturing spatial and temporal dependencies in data and process heterogeneity are flawed because they assume the presence space-time interactions and dependencies. Examples include the alignment of lagged responses to nearby lagged variables in autoregressive moving average models, and existing GAM-based approaches that consider variable interaction over space but with only the aim of model selection and penalization and not process inference [10].

GAMs can be specified with smooths or splines to model non-linear relationships. These are constructed from basis functions that can include single or multiple predictor variables. If a predictor variable is included in a smooth with geographic location (X and Y) then non-linear relationships over space can be modelled - an SVC model. If the smooth is specified with geographic location and time of observation (X , Y and T) then an STVC model is specified. To illustrate this, consider each predictor variable in a SVC or STVC model. It can be specified in 3 different ways:

- i. It is excluded from the model.
- ii. It is included in the SVC / STVC model as a standard parametric response (as in an OLS regression).
- iii. It is included in the SVC / STVC model in a smooth with location (X and Y) but not time.

There are a further 3 ways that each covariate can be specified in an STVC model:

- iv. It is included in the STVC model in a smooth with time (T) but not location.
- v. It is included in the STVC model in a smooth with location and time (X , Y and T).
- vi. It is included in the STVC model in 2 separate smooths, one with location (X and Y) and the other with time (T).

The intercept is treated in a similar way, but without it being absent. Thus for any SVC with k predictor variables there are $2 \times 3^{k+1}$ potential models and for any STVC there are $5 \times 6^{k+1}$ potential models.

2.2 The `stgam` R package

The `stgam` package [8] was created to provide a wrapper around GAM functionality in the `mgcv` package [17], and to allow the user to investigate the different ways each predictor variable could be specified within the GP smooth (as described above), and thus model selection and specification. Its workflow determines the most plausible model given the data and it includes functions that i) create multiple SVC and STVC models defined in different ways as described above, ii) determine the probability of each being the correct model given the data, iii) combine multiple plausible models and, iv) generate spatially and temporally varying coefficient estimates.

`stgam` is underpinned by 2 core concepts. First, the need to test for the presence and nature of any spatial and temporal dependences and thus to determine whether to include each predictor variable in the model, including within GAM smooths, thereby informing SVC / STVC model specification. In `stgam` and related initial work [3], this is done by creating multiple models with each predictor variable specified in different ways, as described above. The probability of each model being the best model given the data is then determined and the most probable model is selected. The second core concept in `stgam` is to evaluate the probability of each potential model being the correct model given the data. This is determined from the model Bayesian Information Criterion (BIC) value [16]. It has been shown elsewhere that the marginal posterior probability of observing the data D given the model M_i – $\Pr(D|M_i)$ – can be approximated using BIC [3], and used to derive the probability of any individual model M_i being the correct model given the data. Of course this is under the assumption that the true model, although unknown, is one of the potential models being evaluated. If multiple models are highly probable, then model averaging can be performed with model weighting from the BIC-derived posterior beliefs in each model.

However, since the release of the first version of the `stgam` package and paper submission to conferences and peer reviewed journals, some issues with the approach to STVC model construction have been identified. These derive from the way that the approach for SVC modelling GAM with GP smooth was extended to STVC models and the use of GP basis from the `mgcv` package. Whilst `mgcv` GP bases may be appropriate for spatial problems and SVC model construction, they are not for STVC models. This is because the `mgcv` GP basis results in spatial models with a potentially suboptimal estimate of the length scale of the correlation basis functions used in the GP basis, and thus temporal models that are erroneous for all time series except those with the longest of length scales.

2.3 The problem in a bit more detail

The substantive problem with the `mgcv` GP basis is that it only fits within the penalized spline class of models that can be estimated by `mgcv` if the basis function parameters are specified. The critical parameter here is the correlation function length scale or range parameter, ρ which may be specified as spherical, power exponential, or as one of three forms of Matérn correlation with κ equal to 1.5, 2.5, or 3.5. The ρ parameter determines the distance at which the correlation function falls below some value. If ρ is not supplied, then `mgcv` fits a smooth using Matérn correlation functions with $\kappa = 1.5$ and $\rho = \max_{i,j} \|\mathbf{x}_i - \mathbf{x}_j\|$ as the basis functions, for any pair of points x_i and x_j , following [14]. If an order penalty is specified then one of the correlation functions is implemented but again using Kammann and Wand's recommendation for the length scale, the maximum distance attained by any pair of points x_i and x_j .

Evidently for SVCs (and similar spatial analyses) this is unlikely to be problematic except in the presence of long or short spatial dependencies. However, for temporal and spatiotemporal data, this specification of length scale in this way **is** problematic: it implies that the correlation between pairs of points will only fall below some small value when those points are separated by an amount of time equivalent to the time series itself. A space-time GP smooth specified in this way will be isotropic, with similar levels of non-linearity in all dimensions, such that the spatial heterogeneity is equal to temporal heterogeneity. This is clearly not correct but the current implementation of `stgam` has no way to control this, because the functions that create different models hardcode a GP basis in the smooths, and there is currently no option to specify an order penalty or ρ . The result is GP smooths specified with pure Kammann and Wand basis functions, despite their unsuitability for time series or space-time problems.

3 Methods

3.1 Proposed improvements

The proposed high level solutions to these problems are 1) to allow users to specify the length scales of the different components, and 2) to fit the spatio-temporal terms via a tensor product smooth of a marginal spatial smooth and a marginal time smooth, both specified with a GP basis, and effectively allowing space-time to be treated in a three-dimensional way. This can be done by specifying a 2D spatial GP smooth for the first margin of the tensor product, and a 1D temporal GP smooth for the second margin. These two bases will smoothly interact creating the desired spatio-temporal varying term due to the tensor product construction.

In `mgcv` syntax, the change in how the GP smooths for each predictor variable `var` are specified as follows:

```
# FROM: a mgcv smooth with GP basis
s(x, y, t, bs = "gp", by = var)
# TO: a mgcv tensor product smooth with GPs
te(x, y, t, d = c(2, 1), bs = rep("gp", 2), by = var)
```

The tensor product smooth requires the marginal basis dimensions to be specified (`d` above). Here the space-time smooth contains 3 covariates composed of a tensor product of a 2-D thin plate regression spline basis for location, and 1-D basis for time.

However, the correlation function length scale or range parameter, ρ also needs to be specified as, both in terms of its form (spherical, power exponential, etc) and distance. In the `mgcv` GAM implementation this is done for each of the margins through the `m` argument passed to the tensor product smooth. This expands the specification of the tensor product smooth with GPs to:

```
# TO: a mgcv tensor product smooth with GPs
# WITH: user specified scale form & length
te(x, y, t, d = c(2, 1), bs = rep("gp", 2), by = var,
  m = list(c(3, rho1), c(3, rho2)))
```

Here `rho1` is the length range for the spatial marginal smooth and `rho2` is that for the temporal marginal. Both are specified with a power of 3, indicating the distance decay of the range. For the spatial margin intervals of 100km for a case study in a large country (USA, China, Brazil, etc) could be explored, 10 km for a small country (UK, Germany, Vietnam, etc) or 1km for a local one. These may be the equivalent of predictor variable bandwidths

in MGWR. For the temporal dimension, intervals of 1, 2, 3 etc years could be explored, depending on how time was recorded in the data. The revised **stgam** package will support investigation of different forms and length scales for spatial and temporal margins.

The ability to vary the scales and lengths ranges in the tensor smooths as above allows each predictor variable to be specified in a way that treats implicitly space-time as three-dimensional. Therefore a second set of investigations will be supported in the revised **stgam** package to allow users to determine which of the possible six forms described above is the appropriate way to include each predictor variable in SVC and STVC models. For example an alternative to the **mgcv** tensor product smooth with GPs, and illustrated above is one that has separate 2-D spatial smooths and 1-D temporal ones:

```
te(x, y, d = 2, bs = "gp", by = var, m = c(3, rho1)) +
  te(t, d = 1, bs = "gp", by = var, m = c(3, rho2))
```

Essentially, this replaces the investigation of different model forms using **mgcv** smooths specified (**s()** in **mgcv**), with ones specified with **mgcv** tensor product smooths (**te()** in **mgcv**). The analysis undertaken in this short paper describes initial work investigating how both of these considerations (space-time lengths and predictor variable form) can be optimised and potentially included in a revised **stgam** package.

3.2 Data and model

The **stgam** package includes two datasets describing annual economic productivity for the 48 contiguous US states (with Washington DC merged into Maryland), 1970 to 1985 from the **plm** R package [9] and the spatial dataset of the 48 contiguous US states **spData** package [1]. The productivity data contains variables describing Private Capital stock (**privC**), Unemployment % (**unemp**) and Public capital investment (**pubC**). The Unemployment variable over time is shown in Figure 1, with 1986 not plotted for aesthetic reasons. The aim of the analysis was

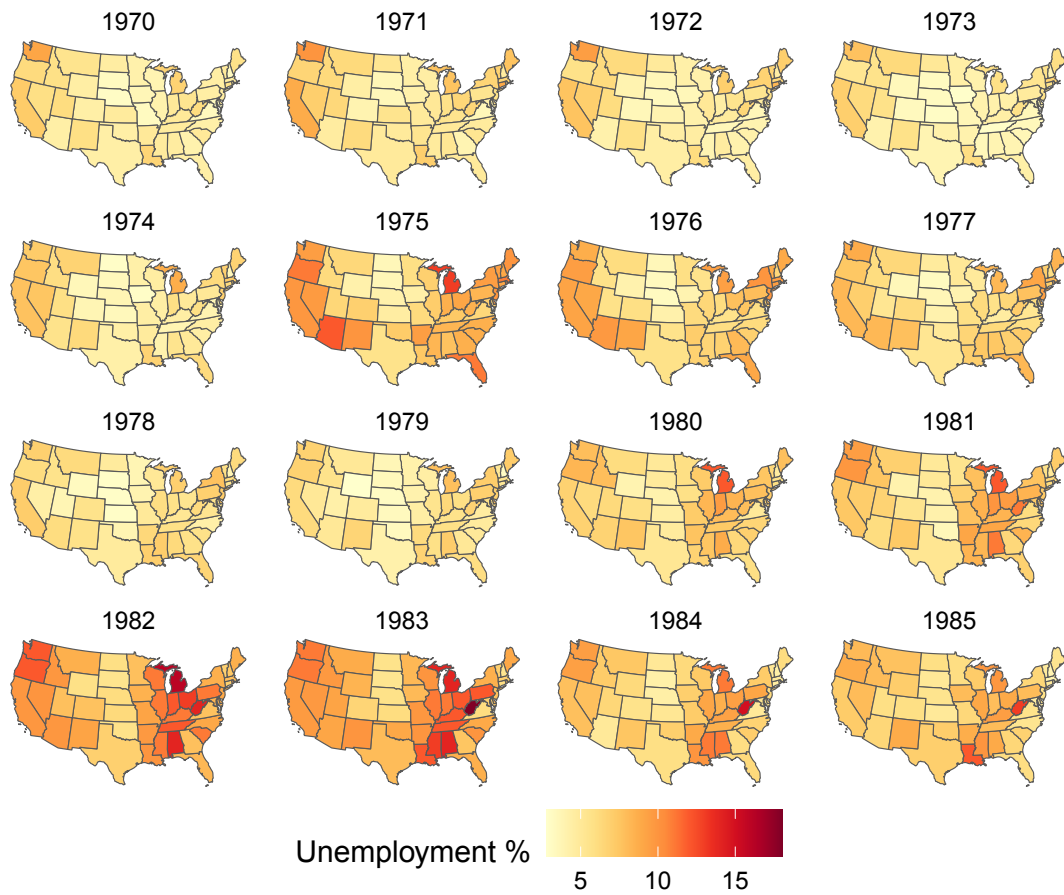
1. to determine optimal ρ values for the correlation function length scales, ρ , a 2D spatial margin and a 1D temporal margin (ρ_s and ρ_t , **rho1** and **rho2** respectively in the above code snippet).
2. to then determine the most appropriate STVC model of Private Capital stock (**privC**), with Unemployment and Public capital investment as predictor variables, specify the tensor product smooths in different ways as described in Section 2.1.

In both cases, ρ and model form were evaluated from the model BIC.

4 Results

The optimal ρ values were determined by creating multiple GAM models with tensor product smooths with GP bases, each with different ρ_s , for the 2D spatial margin and ρ_t for the 1D temporal margin. After investigation of the distances and time series lengths in the data, values for ρ_s from 0 km to 4,500 km in steps of 250 km were explored with values for ρ_t from 0 to 17 years in 1 year steps. A total of 42 models were created with different combinations of ρ_s and ρ_t , and the BIC value for each model extracted. The top ρ values are shown in Table 1 and indicate a ρ_s of 250 km and ρ_t of 9 years. What is interesting is that 250 km is the optimal spatial length range in all of the top 10 models and there is some similar consistency in the temporal length range.

These values for ρ_s and ρ_t specify the length ranges, could be plugged into to a model with each parameter specified in the manner indicated in Section 3.1. However this still assumes that the predictor variables exhibit simultaneous dependencies space-time dependencies with



■ **Figure 1** The % unemployed over US States, 1970-1985.

■ **Table 1** The values of rho1 and rho2 that resulted in the best 10 models.

rho1	rho2	BIC
250	9	5074.882
250	10	5074.893
250	7	5075.093
250	8	5075.252
250	11	5075.675
250	12	5076.220
250	13	5076.938
250	6	5077.692
250	14	5077.719
250	15	5078.704

the target variable (option *v.* in the list in Section 2.1 - *in a smooth with location and time* (X, Y and T)). The second stage in the analysis evaluated different model forms, with each predictor variable specified in one of 6 ways, and the intercept in one of five ways (i.e. 180 models). The models were evaluated using BIC from which the probabilities of the model being the best model were determined in the approach outlined in [3]. The results are

shown in Table 2 and indicate that there is a 4.5% chance that the second ranked model is better than the first, suggesting that the top ranked model can be used. This omits the unemployment predictor variable (`unemp`) and specifies space-time tensor product smooths. It is possible to extract the spatially and temporally varying coefficient estimates from these and to examine the nature of the spatial and temporal dependencies in the data. This is left for future research.

■ **Table 2** The 10 best models, and how the predictor variables were specified within the model, where “—” indicates the absence of a predictor, “Fixed” that a parametric form was specified, “te_S” a spatial tensor product smooth, “te_T” a temporal tensor product smooth and “s_ST” a spatio-temporal tensor product smooth.

Rank	Intercept	unemp	pubC	BIC	Pr(M)
1	te_ST	—	te_ST	4920.984	—
2	te_ST	Fixed	te_ST	4927.090	0.045
3	te_ST	te_T	te_ST	4942.752	0.000
4	te_S	—	te_ST	4963.511	0.000
5	te_S	Fixed	te_ST	4971.211	0.000
6	te_ST	te_S	te_ST	4972.547	0.000
7	te_T + te_S	—	te_ST	4975.040	0.000
8	te_ST	te_T + te_S	te_ST	4976.802	0.000
9	te_T + te_S	Fixed	te_ST	4981.094	0.000
10	te_S	te_T	te_ST	4982.354	0.000

5 Discussion

This short paper describes the next stage in the evolution of an approach for varying coefficient modelling based on GAMs with smooths. It unpicks work described in some published papers [5, 6, 7] for undertaking spatially and temporally varying coefficient modelling, wrapping functionality from the `mgcv` package [17]. Initial work developed spatially varying coefficient (SVC) models using Gaussian Process (GP) smooths that included observation locations, and this framework was extended to include observation location and time. The focus of this extension into space-time modelling was to determine the most appropriate way to specify space-time interactions (dependencies) in the smooths, for example in a single combined space-time smooth or in separate ones for space and for time. However, the type of the smooth is also important. Whilst GP smooths are appropriate for location, they are not for space-time interactions due to assumptions. This is because of the default way that the GP basis correlation function length scale or range parameters (ρ (the distance at which the correlation function falls below some value)) are determined in the `mgcv` GP basis if they are not specified.

This paper details a revised approach to STVC modelling using GAMs with tensor product smooths. Combining a 2D marginal spatial smooth and a 1D marginal time smooth each specified with a GP basis, within tensor product smooth is a more robust way to treat space-time interactions and dependencies. The correlation function length scale, ρ , still needs to be specified. The first part of the analysis optimally determined ρ for both the 2D spatial margin and the 1D temporal margin. In this illustrative case study, these were found to be 250km and 9 years. The second part of the analysis sought to determine the most appropriate model form as described in Section 2.1, with using tensor product smooths and

the optimally determined ρ values. Here the model with the greatest probability of being correct was one with a combined tensor product smooth for the Intercept and for the `pubC` (public capital) predictor variable, and that omitted unemployment from the model.

The determination of the optimal ρ values and the probability of each model being the correct model given the data, both used the model BIC value as described in [7, 3]. However this is not without controversy, as there is a potential lack of a theoretical foundation for the use of BIC for penalized spline models of the form fitted by `mgcv`. For future work, it will be important to establish the viability of a BIC approach to model selection and ρ determination. A second issue is that for many researchers undertaking model selection based on optimising some fit, parsimony or error measure is itself fallacious. The counter argument is that users should just fit the most complex model that they believe is valid for their task in hand. If their understanding of the process being examined is that it is a spatio-temporally varying process then a model specifying that interaction should be fitted. However a counter argument is that increasingly researchers and analysts are working with secondary data, collected for a different purpose and actually part of their job is to determine what kind of spatial, temporal and space-time dependencies are present in the data. Finally, the optimisation of ρ for both the 2D spatial margin and the 1D temporal margin in the tensor product smooths is exciting as this indicates the nature of scales of interactions between the predictor variables and the target variable in the same way that bandwidths do in geographically weighted approaches. Future work will explore how these can be optimised for each predictor variable in a similar way to multiscale geographically weighted approaches and of course their interaction with model selection.

References

- 1 Roger Bivand, Jakub Nowosad, and Robin Lovelace. *spData: Datasets for Spatial Analysis*, 2024. R package version 0.3. URL: <https://CRAN.R-project.org/package=spData>.
- 2 Chris Brunsdon, A Stewart Fotheringham, and Martin E Charlton. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4):281–298, 1996.
- 3 Chris Brunsdon, Paul Harris, and Alexis Comber. Smarter than your average model-bayesian model averaging as a spatial analysis tool (short paper). In *12th International Conference on Geographic Information Science (GIScience 2023)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.GIScience.2023.17.
- 4 Alexis Comber, Paul Harris, and Chris Brunsdon. Multiscale spatially and temporally varying coefficient modelling using a geographic and temporal gaussian process gam (gtgp-gam). In *Proceedings of 12th International Conference on Geographic Information Science (GIScience 2023)*, volume 277, page 22. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023.
- 5 Alexis Comber, Paul Harris, and Chris Brunsdon. Multiscale spatially varying coefficient modelling using a geographical gaussian process gam. *International Journal of Geographical Information Science*, 38(1):27–47, 2024. doi:10.1080/13658816.2023.2270285.
- 6 Alexis Comber, Paul Harris, Daisuke Murakami, Tomoki Nakaya, Narumasa Tsutsumida, Takahiro Yoshida, and Chris Brunsdon. Encapsulating spatially varying relationships with a generalized additive model. *ISPRS International Journal of Geo-Information*, 13(12):459, 2024. doi:10.3390/IJGI13120459.
- 7 Alexis Comber, Paul Harris, Naru Tsutsumida, Jennie Gray, and Chris Brunsdon. Specifying spatially and temporally varying coefficient models using gams with gaussian process splines. *Transactions in GIS*, submitted.
- 8 Lex Comber, Paul Harris, and Chris Brunsdon. *stgam: Spatially and Temporally Varying Coefficient Models Using Generalized Additive Models*, 2024. R package version 0.0.1.2. URL: <https://CRAN.R-project.org/package=stgam>.

- 9 Yves Croissant, Giovanni Millo, and Kevin Tappe. *plm: Linear Models for Panel Data*, 2025. R package version 2.6-5. URL: <https://CRAN.R-project.org/package=plm>.
- 10 Jakob A Dambon, Fabio Sigrist, and Reinhard Furrer. Joint variable selection of both fixed and random effects for gaussian process-based spatially varying coefficient models. *International Journal of Geographical Information Science*, 36(12):2525–2548, 2022. doi: 10.1080/13658816.2022.2097684.
- 11 A Stewart Fotheringham, Wenbai Yang, and Wei Kang. Multiscale geographically weighted regression (mgwr). *Annals of the American Association of Geographers*, 107(6):1247–1265, 2017.
- 12 T Hastie and R Tibshirani. Generalized additive models. chapman hall & crc. *Monographs on Statistics & Applied Probability. Chapman and Hall/CRC*, 1, 1990.
- 13 Trevor Hastie and Robert Tibshirani. Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398):371–386, 1987.
- 14 EE Kammann and Matthew P Wand. Geoadditive models. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 52(1):1–18, 2003.
- 15 Binbin Lu, Chris Brunsdon, Martin Charlton, and Paul Harris. Geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, 31(5):982–998, 2017. doi:10.1080/13658816.2016.1263731.
- 16 Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, pages 461–464, 1978.
- 17 Simon Wood. *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*, 2023. R package version 1.9-1. URL: <https://CRAN.R-project.org/package=mgcv>.
- 18 Wenbai Yang. *An extension of geographically weighted regression with flexible bandwidths*. PhD thesis, University of St Andrews, 2014.