AGILE: GIScience Series, 6, 19, 2025. https://doi.org/10.5194/agile-giss-6-19-2025
Proceedings of the 28th AGILE Conference on Geographic Information Science, 10–13 June 2025.
Eds.: Auriol Degbelo, Serena Coetzee, Carsten Keßler, Monika Sester, Sabine Timpf, Lars Bernard
This contribution underwent peer review based on a full paper submission.
© Author(s) 2025. This work is distributed under the Creative Commons Attribution 4.0 License.

How much time to include in multiscale space-time regressions? Optimising predictor variable temporal lags

Alexis Comber [1], Paul Harris [2], and Chris Brunsdon [3]

[1] School of Geography, University of Leeds, Leeds, UK
[2] Sustainable Agriculture Sciences,Rothamsted Research, UK
[3] National Centre for Geomcomputation, Maynooth University, Maynooth, Ireland

Correspondence: Alexis Comber (a.comber@leeds.ac.uk)

Abstract. Generalised Additive Models (GAMs) with Gaussian process bases have been proposed as a framework for constructing spatially varying coefficient (SVC) and spatially and temporally varying coefficient (STVC) regression models, that overcome many of the theoretical problems and technical limitations associated with geographically weighted approaches. Recent work has considered the SVC case in detail and this is being extended to the temporal case. However, while spatial lags and dependencies are well handled by many existing methods, one of the critical issues in space-time modelling is how to determine appropriate temporal lags for individual predictor variables that may exhibit different temporal dependencies with the target variable. This paper demonstrates an outline approach for optimising these. Additionally, lags determined in this way may be used to inform on the temporal margins used to parameterise space-time tensor products smooths in GAM based STVC approaches.

Keywords. Space-time relationships, Coefficient non-stationarity, Process heterogeneity

1 Introduction

Spatial analysis increasingly involves working with time-series spatial data – repeated measurements made at or over the same locations. This extends analysis considerations into the temporal dimension and requires methods able to determine how statistical relationships vary temporally as well as spatially, and that capture process heterogeneity and any spatial and / or temporal dependencies (non-stationarities). Approaches based on Geographically Weighted Regression (GWR) (Brunsdon et al., 1996) are commonly used to construct spatially varying coefficient (SVC) models (Comber et al. (2022) document this), and have been extended into time via Geographically and Temporally Weighted Regression (GTWR) (Huang et al., 2010; Fotheringham et al., 2015). However, there are a number of long standing operational concerns and methodological limitations to geographically weighted (GW) approaches, around their re-use of data, limited form (often only Gaussian responses) and their inability to undertake out-of-sample prediction.

To address this situation, a number of researchers have started to develop varying coefficient approaches using Generalised Additive Models (GAMs) (Fan and Huang, 2022; Comber et al., 2024a, b) with some recent extensions into the temporal domain (Hong et al., 2025b, a; Comber et al., submitted) Ongoing work by the authors reported elsewhere describes the thinking around informed GAM-based regression approaches for constructing spatially and temporally varying coefficient (STVC) models. This includes the need for workflows that i) explicitly capture the nature of any space-time dependencies (interactions) in the data, and ii) then specifies them appropriately within the model (e.g. as parametric terms or within GAM smooths), ensuring appropriate model specification (see @stgam), rather than making assumptions about space-time dependencies, as is done in all other space-time modelling approaches in geography including GTWR.

In space-time GAMs with smooths, each predictor variable may take one of 6 forms: it may be absent, it may take a standard parametric form, be specified in a smooth with location, a smooth with time, in a single smooth with location and time, or in 2 separate space and time smooths. However, a further consideration is how to specify temporal dependencies - i.e. how much time and space to include local coefficient estimates. GWR based approaches, for example, determine spatial dependencies by optimising the bandwidth (size) of a moving window (kernel). The various GTWR extensions take different approaches. Huang et al. (2010) define a parameter ratio (ρ) to match and join the temporal and spatial distances, facilitating a linear combination of temporal and spatial distances. This difficulty in identifying a suitable space-time metric is similarly found in specifying space-time variograms for geostatistical models with respect to separability (Gneiting et al., 2006). Fotheringham et al. (2015) took a different approach. They proposed a kernel function to identify optimal spatio-temporal bandwidths to account for variations in relationships over space and time. However, because of the computational overheads of calibrating across this, 5 years of temporal data were selected from the 19 years available and introduced sequentially into the model as lagged data, from which optimal spatial bandwidths were determined for each year. However, both of the approaches to STVC modelling with GTWR are predicted on the assumption that the relationship between the target and predictor variables do indeed vary over time and space simultaneously, and any temporal dependencies are the same for each predictor variable. That is, the space-time bandwidths are not multi-scale. Some previous work has dealt with model form, and GAMs with smooths have been shown to deal with *how much space* (Comber et al., 2024a, b). Here the questions of *how much time* is addressed by optimising lag periods for each predictor variable and the exploratory research described in this paper investigates whether and how temporal lags should to be accommodated in space-time regression models, rather than treating each predictor variable has having a simultaneous association with the target variable.

2 Data and Methods

This study uses a dataset describing average annual house prices over 13 years (2008 to 2020) in 792 neighbourhoods (lower super output areas - LSOAs) in Leeds and Bradford. The data includes predictor variables of annual data describing population proportions of unemployment (JSA), neighbourhood population change (Churn), and new national insurance registrations (NINO) as a measure of in-migration. By way of illustration, the spatial distribution of mean house price for 2017 is shown in Figure 1.

A general GAM-based STVC was specified with Tensor Product (TP) smooths with Gaussian Process (GP) bases, parameterised with observation location and time for each predictor variable, of the following form:

$$y = \beta_0(u_i, v_i, t_i) + \sum_{m}^{j=1} \beta_j(u_i, v_i, t_i) x_{ij} + \epsilon(u_i, v_i, t_1)$$

where for observations $i = 1 \dots n$, y is the response variable, (u_i, v_i, t_i) describes the spatial location of the i^{th} observation at time lag t_i, β_0 is the intercept term, β_j is the regression coefficient for the j_{th} predictor variable, with $\beta_j(u_i, v_i, t_i)$ a realization of the continuous function $\beta_j(u, v, t)$ at point i, and ϵ is the independent random error. Note that no spatial lags were applied or investigated.

The spatio-temporal terms were fitted via a TP smooth that contained marginal spatial and time smooths, both specified with a GP basis. This treated space-time in a three-dimensional way, that is, an *a priori* assumption of space-time dependencies was made. A 2D spatial GP smooth was specified for the first margin and a 1D temporal GP smooth for the second. (Note that the use of tensor products here is a methodological shift to from our previous, and it turns out fallacious, investigations of STVC GAM models using GP smooths - for example see Comber et al. (2023) - and we are in the process of revising these approaches and the accompanying ‘stgam‘ R package (Comber et al., 2024c)).

For tensor product smooths, the length scales (penalty or range parameters) (ρ) need to be specified. These determine the distance at which the correlation function falls below some small value, and in space-time smooths they define the spatial and temporal scales of interaction (i.e. the space-time dependencies). Here, spatial and temporal ranges were specified with a power exponential of 3, indicating the a spherical range distance decay: ρ_{space} was set to 25km for the spatial smooths and rho_{time} to 7 years for the temporal smooths. These were based on the average distances and time frames (mid-points) in the data. A series of GAM models with the TP smooths specified in this way were implemented in R using the ‘mgcv‘ package [@mgcv], with each predictor variable lagged from 0 to 5 years. Thus for k predictor variables there are 6^k potential lagged STVC motels, resulting in the investigation of 216 models in this case. Each model was evaluated using a probability based approach using the Bayesian Information Criterion (BIC) (Schwarz, 1978). BIC can be used to derive the probability of each individual model being the correct model given the data, and the the *marginal* or *relative* or probability between each model and the best model. The derivation of this is described in Brunsdon et al. (2023). Here BIC was calculated for each lagged model from which probabilities were calculated allow different model specifications to be ranked and compared.

3 Results

Table 1 shows the relative probabilities the best 10 space-time models (i.e. with the lowest BIC scores). In this case, the probabilities indicate that the second ranked house price model has a probability of less than 1% chance (<0.00)) of providing a better lagged model than the first. This indicates the optimum lag periods for each individual predictor variable, and potential suggests the periodicity of the *cause and effect* response in house prices:

- a lag of 4 years for neighbourhood population change (Churn).
- a lag of 4 years for new national insurance registrations (NINO) indicating in-migration.
- a lag of 5 years for unemployment benefit recipients (JSA).

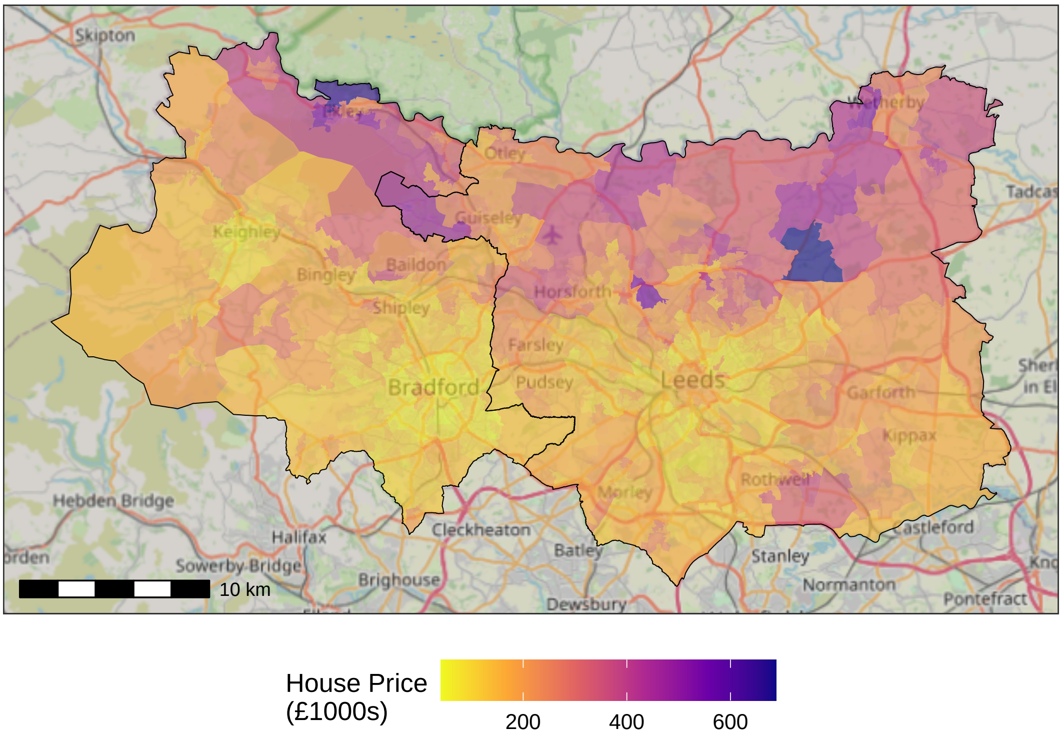

Figure 1. LSOAs in the study area and mean house price for 2017, with an OpenStreetMap backdrop.

These were used to construct a final model, and the spatially and temporally varying coefficient estimates were extracted using functions in the `stgam` R package (Comber et al., 2024c), which in turn can be summarised and mapped.

The optimised lagged model is reasonably accurate, with an R^2 of 0.78, a RMSE of 38.4 and a MAE of 26.3. Summaries of the STVC model estimates and their variation over space and time are shown in Table 2. Interestingly, each predictor variable flips in sign at some point in space-time. It is instructive to examine some of this temporal and spatial variation as in Figures 2 and 3. The plots in Figure 2 show the median coefficient estimates for each year over LSOAs. The Intercept increases over time, the relationship between house price and Churn (4 year lag) is generally less negative until 2014, and then plateaus from 2014 to 2019 (is this Brexit?), JSA (5 year lag) is moderately negative, declining after 2017 and NINO (4 year lag) is negative but increasing to 0 in 2020 . The changes of the JSA coefficient estimates over time are shown in Figure 3 and show the highly localised emergence of an increasingly negative relationship with house price in certain locations.

4 Discussion

This paper introduces an approach for determining how to specify the temporal lags in space-time regressions. It uses GAMs with Tensor Product smooths (Hastie and Tibshirani, 1990) specified with GP bases. These were parameterised with location and time and used to create multiple models with different temporal lags for each predictor variable. In doing so, it sought to answer the question ‘how much time’ to include in spatio-temporal regressions, by determining the most probable lagged model. This is one in which the time taken for the potentially causal effects of each predictor variable to manifest themselves is explicitly identified. Here, a lag of 4 years was found for the effects of `Churn` (neighbourhood in- and out-migration) on house price was found. For unemployment (`JSA`) it was 5 years and for national insurance registrations (`NINO`) the lag was 4 years.

The framework for determining how to most appropriately determine temporal lag time series spatial data described in this short paper is part of a wider set of activities that are seeking to move away from plug-and-play approaches, in which little consideration is given to the model form in space-time analyses. One potential advantage of this approach, is that it explicitly captures the scales of lagged temporal dependencies. Optimising temporal lags as done in this paper, could be used to determine the individual ρ parameters for each temporal predictor variable, in a space-time tensor product smooth, that is the distance at which the correlation function falls below some value.

A key objective the approach taken to space-time analysis described in this paper concerns process inference (understanding) rather than prediction. Thus, it is important to consider the nature and reasons for any observed variation n the coefficient estimates. High variation in these (large changes in magnitude, flipping signs) is commonly found in spatially varying coefficient models. These may be found for a number of reasons including a poor conceptual

Table 1. The 10 best models, and the time period used to lag each predictor variable, the model BIC value and the associated probabaility of the model being the correct given the data.

Churn	NINO	JSA	BIC	Pr(M)
4	4	5	66206.22	–
5	1	5	66226.59	0.00
5	1	4	66227.41	0.00
5	4	5	66244.77	0.00
5	3	4	66247.20	0.00
4	5	4	66248.29	0.00
5	2	4	66250.14	0.00
4	2	5	66253.39	0.00
5	5	5	66253.47	0.00
4	5	5	66253.83	0.00

Table 2. Summaries of the STVC coefficient estimates and their variation over space and time.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
middlehline ntercept	-527.5	119.6	173.6	187.5	227.3	789.6
Churn	-2171.9	-380.5	-120.4	-208.2	-11.2	2102.7
NINO	-6874.8	-368.8	-116.5	-168.6	53.6	8697.9
JSA	-28042.1	-2872.3	-1470.7	-2145.5	-589.1	14324.0

bottomhline

model, variations due to human dimensions (behaviours, preferences, attitudes etc - geography is not physics), bad measurements, unaccounted for local factors, or because of process spatial heterogeneity. Therefore in such situations, the analysts has to decide whether, for example, it is reasonable that factors such as unemployment are sometimes very negatively and sometimes very positively correlated, whether the model is missing some key driving variable - i.e. whether variations in coefficient estimate (and thus change in house price) may rather be due to unobserved effects - or potentially whether the model is overfitting.

It is also importation to note that the results lags were determined using arbitrarily set length ranges (ρ parameters) for space and time of 25km for ρ_{space} for the spatial smooths and rho_{time} to 7 years for the temporal ones. These specify the distances spatial and temporal dependencies ranges, in the same way as bandwidths in GWR-based approaches. Future work will develop approaches to optimise these in tensor product space-time smooths. Other work is also developing methods to determine appropriate model specification and how each predictor variable should be included in models, for example in parametric form, in separate space and time GP smooths, or in a combined space-time smooth. This is because most approaches to space-time analysis do not explicitly examine either model form or how variables should be lagged, and thereby implicitly assume the presence and nature of space-time dependencies i.e. that the measurements depend on previous values and those of nearby locations, without testing that assumption. A final limitations is that BIC values are potentially problematic for comparing `mgcv` GAM models - something that has emerged since submission and acceptance of this short paper - and current ongoing work is exploring the use of un-biased risk estimators and Generalized Cross-Validation (GCV) scores to do this. These developments will be included in future updates to the `stgam` (Comber et al., 2024c) package and future research papers. Finally, although temporal lags are commonly used, for example in econometric models, ongoing work continues to combine and refine these model specification investigations, considering how best to lag temporal variables and to determine the nature of the space-time interactions and dependencies present in the data. For example it is still unclear how optimal time lags should be used to inform modelling decisions beyond the specific case study, and the degree to which they are context-dependent.

5 Acknowledgements, DSA, Generative AI and Contributions

-

The authors contributed the ideas, theory and writing in this paper to differing degrees. AC did the bulk of the thinking, data collection and analysis. NM provided the interface into the LLM, with other authors contributing to discussions and brainstorming activities. No generative AI was used in the writing or ideas. This research is supported by UKRI (ESRC) funding ES/Y006259/1 under the Digital Footprints scheme. The data for this research

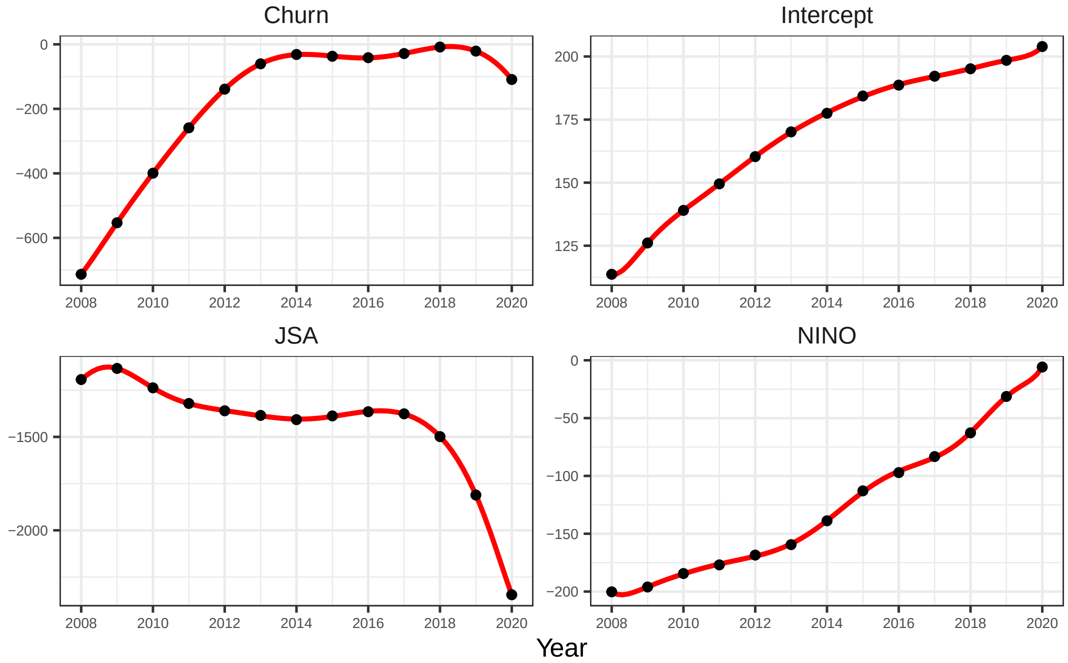

Figure 2. The temporal variation of the median coefficient estimates, 2008 to 2020.

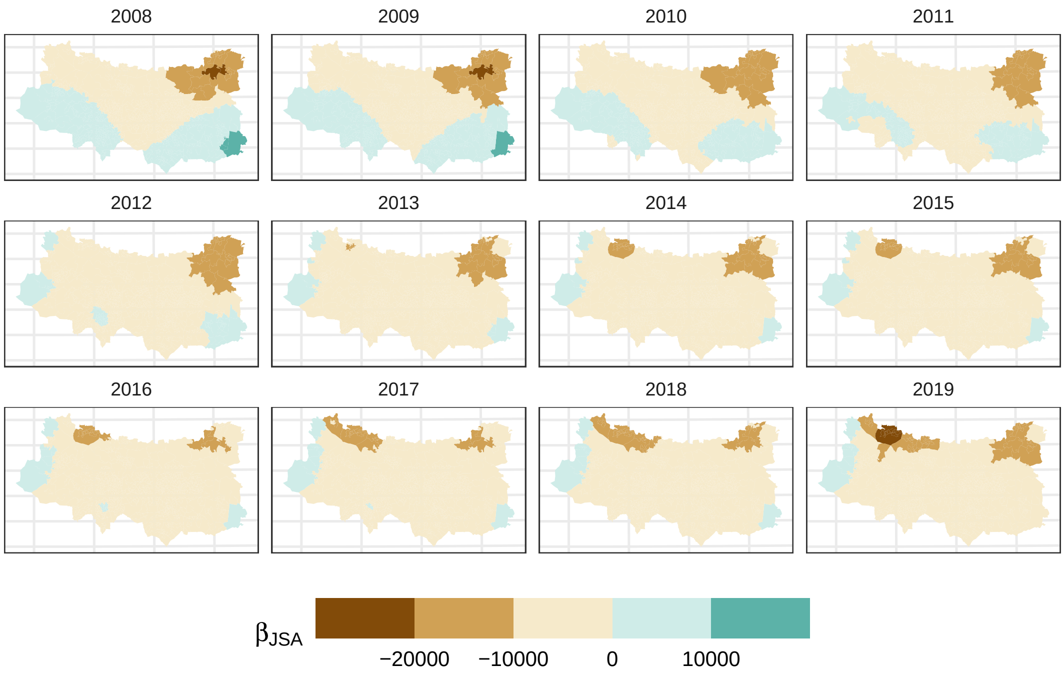

Figure 3. Detail of the spatial and temporal variation of the JSA (unemployment) coefficient estmates, 2008 to 2019.

have been provided by the Consumer Data Research Centre, an ESRC Data Investment, under project ID CDRC, ES/L011840/1; ES/L011891/1. The data used in this analysis cannot be shared publicly, but this may be possible on request. Please contact the authors.

References

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E.: Geographically weighted regression: a method for exploring spatial nonstationarity, Geographical Analysis, 28, 281–298, 1996.

Brunsdon, C., Harris, P., and Comber, A.: 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.

Comber, A., Callaghan, M., Harris, P., Lu, B., Malleson, N., and Brunsdon, C.: gwverse: A Template for a New Generic Geographically Weighted R Package, Geographical Analysis, 54, 685–709, 2022.

Comber, A., Harris, P., and Brunsdon, C.: 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), vol. 277, p. 22, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2023.

Comber, A., Harris, P., and Brunsdon, C.: Multiscale spatially varying coefficient modelling using a Geographical Gaussian Process GAM, International Journal of Geographical Information Science, 38, 27–47, 2024a.

Comber, A., Harris, P., Murakami, D., Nakaya, T., Tsutsumida, N., Yoshida, T., and Brunsdon, C.: Encapsulating Spatially Varying Relationships with a Generalized Additive Model, ISPRS International Journal of Geo-Information, 13, 459, 2024b.

Comber, A., Harris, P., Tsutsumida, N., Gray, J., and Brunsdon, C.: Specifying spatially and temporally varying coefficient models using GAMs with Gaussian Process splines, Transactions in GIS, submitted.

Comber, L., Harris, P., and Brunsdon, C.: stgam: Spatially and Temporally Varying Coefficient Models Using Generalized Additive Models, https://CRAN.R-project.org/package=stgam, r package version 0.0.1.0, 2024c.

Fan, Y.-T. and Huang, H.-C.: Spatially varying coefficient models using reduced-rank thin-plate splines, Spatial Statistics, 51, 100 654, 2022.

Fotheringham, A. S., Crespo, R., and Yao, J.: Geographical and temporal weighted regression (GTWR), Geographical Analysis, 47, 431–452, 2015.

Gneiting, T., Genton, M. G., and Guttorp, P.: Geostatistical space-time models, stationarity, separability, and full symmetry, Monographs On Statistics and Applied Probability, 107, 151, 2006.

Hastie, T. and Tibshirani, R.: Generalized Additive Models. Chapman Hall & CRC, Monographs on Statistics & Applied Probability. Chapman and Hall/CRC, 1, 1990.

Hong, Z., Wang, R., Wang, Z., and Du, W.: Inference issue in multiscale geographically and temporally weighted regression, Statistics and Computing, 35, 30, 2025a.

Hong, Z., Wang, Z., Wang, H., and Wang, R.: A scale-adaptive estimation for mixed geographically and temporally weighted regression models, Journal of Geographical Systems, pp. 1–27, 2025b.

Huang, B., Wu, B., and Barry, M.: Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices, International journal of geographical information science, 24, 383–401, 2010.

Schwarz, G.: Estimating the dimension of a model, The annals of statistics, pp. 461–464, 1978.