

stgam: An R package for GAM-based varying coefficient models

Lex Comber^{1*}, Paul Harris^{2*}, and Chris Brunsdon^{3*}

¹ School of Geography, University of Leeds, UK ² Rothamsted Research, North Wyke, UK ³ National Centre for Geocomputation, Maynooth University, Ireland * These authors contributed equally.

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Open Journals](#)

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Very often we are interested in quantifying how and where statistical relationships vary over space, and how they change over time. This can be done using *varying coefficient* models. Our ability to undertake such space-time analyses is enhanced by the increased production and availability of data describing a wide range of phenomenon that include both spatial and temporal attributes in the form of GPS coordinates and time-stamps. The stgam package provides a framework for creating regression models in which the relationships between the response (dependent) variable and individual predictor (independent) variables is allowed to vary over space, time or both, in order to create spatially, temporally or spatio-temporally varying coefficient models.

However a key question is what form of space-time interaction should be specified in our statistical models? Most current approaches require this to be specified in advance. when the The stgam package provides w

Statement of need

Generalized Additive Models (GAMs) ([Hastie & Tibshirani, 1990](#)) provide a framework for constructing regression models in which the response variable can have Gaussian and non-Gaussian distributions and linear or non-linear relationships with the predictor variables. GAMs handle varying responses and non-linear relationships using smooths. When smooths include location as well as individual predictor variables, then the result is a spatially varying coefficient model, with time a temporally varying one, and the inclusion of space *and* time results in a spatially and temporally varying coefficient model.

There are a number of approaches for quantifying spatial characteristics such as spatial autocorrelation ([Cliff & Ord, 1973, 2009](#); [Getis, 2008](#); [Moran, 1950](#)) and processes spatial heterogeneity ([Brunsdon et al., 1996](#); [Casetti, 1972](#); [Fotheringham et al., 2002](#); [Gelfand et al., 2003](#); [Griffith, 2008](#); [Jones, 1991](#); [McMillen, 1996](#)). Tools also exist that quantify the temporal dynamics of spatial processes ([Crespo et al., 2007](#); [Di Giacinto, 2006](#); [Elhorst, 2003](#); [Gelfand et al., 2004](#); [Huang et al., 2010](#); [Pace et al., 1998, 2000](#)). Griffith (2010) groups these into approaches that model autocorrelation effects such as space-time autoregressive integrated moving average models, spatial panel regressions and geo-statistical space-time models, and temporal extensions to approaches that model relationship spatial heterogeneity such as eigenvector filtering, geographically weighted regression and Bayesian spatially varying coefficient models.

All of the space-time approaches described above make a critical assumption about the presence and nature of the space-time relationships in the data, as documented in Comber et al. ([submitted](#)): space-time autoregressive models assume that measurements collected over

space and time are spatially and temporally correlated, spatial panel regression models assume the presence of serial and spatial autocorrelation and geo-statistical models are concerned with prediction uncertainty not process understanding. The other approaches, eigenvector spatio-temporal filtering (Chun, 2014; Patuelli et al., 2011), geographically and temporally weighted regression (GTWR) (Fotheringham et al., 2015; Huang et al., 2010) and Bayesian models extended to space-time (Amaral-Turkman et al., 2011; Gelfand et al., 2005; Paez et al., 2008), all assume the presence of latent spatial and temporal autocorrelation in variables.

This paper describes an R package and a workflow for undertaking space-time varying coefficient (STVC) models using GAMs. It parameterises GAMs with Gaussian Process (GP) smooths for spatially varying coefficient (SVC), temporally varying coefficient (TVC) and STVC models, and explicitly investigates spatio-temporal interactions in the data in order to determine the appropriate spatio-temporal model form. This workflow supports user investigations of the presence and the nature of data spatial interactions, temporal interactions, as well as space-time ones. This builds on the work describing the development for spatially varying coefficient models with GP-GAMs in Comber et al. (2024) and the extension to spatial and temporal coefficient modelling in Comber et al. (submitted). # Package overview

Worked example

Reference

- Amaral-Turkman, M., Turkman, K., Le Page, Y., & Pereira, J. (2011). Hierarchical space-time models for fire ignition and percentage of land burned by wildfires. *Environmental and Ecological Statistics*, 18, 601–617.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), 281–298.
- Casetti, E. (1972). Generating models by the expansion method: Applications to geographical research. *Geographical Analysis*, 4(1), 81–91.
- Chun, Y. (2014). Analyzing space-time crime incidents using eigenvector spatial filtering: An application to vehicle burglary. *Geographical Analysis*, 46(2), 165–184.
- Cliff, A. D., & Ord, J. K. (1973). *Spatial autocorrelation* (Vol. 41). Pion, London.
- Cliff, A. D., & Ord, J. K. (2009). What were we thinking? *Geographical Analysis*, 41(4), 351–363.
- Comber, A., Harris, P., & Brunsdon, C. (2024). Multiscale spatially varying coefficient modelling using a geographical gaussian process GAM. *International Journal of Geographical Information Science*, 38(1), 27–47.
- Comber, A., Harris, P., Tsutsumida, N., Gray, J., & Brunsdon, C. (submitted). Where, when and how? Specifying spatially and temporally varying coefficient models using generalized additive models with gaussian process splines. *International Journal of Geographical Information Science*.
- Crespo, R., Fotheringham, S., & Charlton, M. (2007). Application of geographically weighted regression to a 19-year set of house price data in london to calibrate local hedonic price models. *Proceedings of the 9th International Conference on Geocomputation*.
- Di Giacinto, V. (2006). A generalized space-time ARMA model with an application to regional unemployment analysis in italy. *International Regional Science Review*, 29(2), 159–198.
- Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26(3), 244–268.

- 86 Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. John Wiley & Sons.
- 88 Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Geographical and temporal weighted regression (GTWR). *Geographical Analysis*, 47(4), 431–452.
- 90 Gelfand, A. E., Banerjee, S., & Gamerman, D. (2005). Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics: The Official Journal of the International Environmetrics Society*, 16(5), 465–479.
- 93 Gelfand, A. E., Ecker, M. D., Knight, J. R., & Sirmans, C. (2004). The dynamics of location in home price. *The Journal of Real Estate Finance and Economics*, 29, 149–166.
- 95 Gelfand, A. E., Kim, H.-J., Sirmans, C., & Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98(462), 387–396.
- 98 Getis, A. (2008). A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical Analysis*, 40(3), 297–309.
- 100 Griffith, D. A. (2008). Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). *Environment and Planning A*, 40(11), 2751–2769.
- 102 Griffith, D. A. (2010). Modeling spatio-temporal relationships: Retrospect and prospect. *Journal of Geographical Systems*, 12, 111–123.
- 104 Hastie, T., & Tibshirani, R. (1990). Generalized additive models. Chapman hall & CRC. *Monographs on Statistics & Applied Probability*. Chapman and Hall/CRC, 1.
- 106 Huang, B., Wu, B., & 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.
- 109 Jones, K. (1991). Specifying and estimating multi-level models for geographical research. *Transactions of the Institute of British Geographers*, 148–159.
- 111 McMillen, D. P. (1996). One hundred fifty years of land values in chicago: A nonparametric approach. *Journal of Urban Economics*, 40(1), 100–124.
- 113 Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17–23.
- 115 Pace, R. K., Barry, R., Clapp, J. M., & Rodriguez, M. (1998). Spatiotemporal autoregressive models of neighborhood effects. *The Journal of Real Estate Finance and Economics*, 17, 15–33.
- 118 Pace, R. K., Barry, R., Gilley, O. W., & Sirmans, C. (2000). A method for spatial-temporal forecasting with an application to real estate prices. *International Journal of Forecasting*, 16(2), 229–246.
- 121 Paez, M. S., Gamerman, D., Landim, F. M., & Salazar, E. (2008). Spatially varying dynamic coefficient models. *Journal of Statistical Planning and Inference*, 138(4), 1038–1058.
- 123 Patuelli, R., Griffith, D. A., Tiefelsdorf, M., & Nijkamp, P. (2011). Spatial filtering and eigenvector stability: Space-time models for german unemployment data. *International Regional Science Review*, 34(2), 253–280.