ORIGINAL ARTICLE

Modeling spatio-temporal relationships: retrospect and prospect

Daniel A. Griffith

Abstract   Interest in space–time modeling is experiencing a resurgence, in part because more and more sizeable space–time datasets are becoming readily available. Currently techniques to describe these data, many of which have existed for years, are being utilized and improved. This paper surveys general categories of these techniques (i.e., autoregressive-integrated-moving-average models, space–time autoregressive models, three-dimensional geostatistical models, and panel data models), in retrospect, demonstrates a future prospect (i.e., spatial filtering models), and suggests important topics for incorporation into a research agenda, including ones pertaining to non-normal random variables, panel data models, space–time heterogeneity, missing data, and distributional properties of space–time filters.

Keywords   Random effects · Space–time · Spatial filter · STAR

JEL Classification   C1 · C4 · C5 · C21

1 Introduction

Correlated samples data have been a topic treated in statistics and econometrics for many years, frequently in terms of repeated measures. Special cases of this class of data include time series (a topic of econometrics), spatial series (a topic of spatial econometrics), and space–time series. The purpose of this article is to comment on the state-of-the-art, future prospects, and remaining challenges concerning spatio-

D. A. Griffith is an Ashbel Smith Professor at the University of Texas at Dallas.

D. A. Griffith (✉)
University of Texas at Dallas, Richardson, TX, USA
e-mail: dagriffith@utdallas.edu

temporal modeling of space–time series. Some of these research themes constitute the topics of this special issue.

2 Retrospect

Methodological developments for spatio-temporal modeling were challenged for many years by the lack of available data. This situation has been changing, and today many space–time data series are available, although some may include rather short time series or considerable missing data. These kinds of data can be subjected to the following five basic analysis approaches: autoregressive-integrated-moving-average (ARIMA) models for multiple time series (an impulse-response type of specification), space–time autoregressive (STAR) models for a single space–time series, three-dimensional geostatistical models, panel data models with fixed and random effects, and spatial filter models.

ARIMA models require lengthy time series, and the impulse-response function specification aligns time lags of a response variable with those of a covariate. In this case, the response variable is an attribute variable at some location (i.e., a location-specific time series), and the covariates are the same attribute variable at other, frequently nearby, locations. This specification helps uncover lead and lag locations for some time series process and is especially informative for spatial diffusion processes. Much of its theory is based upon the general distributed lag model of econometrics. Hartfield and Gunst (2003) furnish an informative application of this approach. Meanwhile, Elhorst (2001) furnishes a treatment of space–time data transitioning between the multivariate ARIMA and the STAR models.

STAR models (see Cliff et al. 1975) are explicit specifications describing how some attribute variable jointly varies in space and time, and differ from the multivariate ARIMA specifications in that geographically varying leads and lags are not incorporated. This specification can take on two different forms: a value at a given location is cast as a function of the preceding in situ value at that location as well as the preceding neighboring values, a lagged specification, and a value at a given location is cast as a function of the preceding in situ value at that location as well as the contemporaneous neighboring values, a spatially contemporaneous specification. A white noise term, which is uncorrelated in space and time, is added to either of these trend specifications. Correlation in space and time is captured by the autoregressive structure of the model (i.e., the response variable appears in both sides of the equations). Feedback loops or cycles fundamentally differentiate these two forms. To avoid time and space boundary effects, these formulations assume that sufficient time has transpired and that the spatial extent is sufficiently large to allow negligible effects from boundary conditions. Griffith (1996) discusses numerical simplifications for the estimation of STAR models. Heuvelink and Griffith (2010) furnish an informative application of this approach.

Geostatistical model specifications commonly used to describe space–time variation frequently parallel STAR specifications. Resulting expressions for space–time correlograms derived from especially first- and second-order STAR models can be solved numerically and then linked to appropriate space–time semivariogram

models. Kriging in space and time can be done in much the same way that it is done in a purely spatial setting. The main difficulties are in defining a realistic stochastic model that is assumed to have generated observed data and in characterizing and estimating the space–time correlation of that model. An important issue in the space–time geostatistical literature concerns whether or not the space and the time components of a formulated function are: separable such that they factor (Gneiting 2006) or non-separable such that they form a linear combination (Ma 2008). In other words, the space–time variable of interest may be treated as a sum of independent stationary spatial, temporal, and spatio-temporal components, which leads to a summetric space–time variogram model, or it may be treated as some non-linear, multiplicative version of spatial, temporal, and spatiotemporal components, which leads to either a product- or a product-sum-metric space–time variogram. Stein (2005) furnishes an overview of space–time covariance and aspects of spatial–temporal interaction within the context of geostatistics.

Data forming a short-length-in-time space–time series—too short to utilize, say, an ARIMA model—can be described in a panel data context by including a random effects term in a spatiotemporal model specification. This term is spatially structured in order to account for both serial and spatial autocorrelation; the spatially structured component accounts for spatial autocorrelation, and the sum of the spatially structured and unstructured components accounts for serial correlation. The resulting specification is a mixed model (Baltagi 2005) containing both fixed and random effects. The random effects can be estimated by treating each locational time series as a set of repeated measures, allowing them to be separated from residuals. Practical situations often constitute quasi-panel datasets because attributes for the same set of objects (e.g., farms) within an areal unit are not being measured through time; rather, attributes for areal unit aggregates of changing collections of objects are being measured. Griffith (2008) furnishes an informative application of this approach.

Similarly, panel data consisting of a small spatial series with a lengthy time series can have the full spatial covariance matrix estimated by exploiting the time dimension and using it to establish asymptotics. This type of analysis can be achieve with, for example, a spatial seemingly unrelated regression (SUR) model. Panel data also can be described with a spatial vector autoregressive (VAR) model, which includes spatial as well as temporal lags among a vector of stationary state variables in order to capture joint spatial and temporal dynamics. An estimated spatial VAR model can be used to calculate impulse responses between variables over time and across space. This specification converts a situation of spatial dynamics with temporal statics (i.e., a repeated measures perspective) to one with temporal dynamics.

In this mixed-models context, space–time heterogeneity can be accounted for in various ways, including specifications involving spatial filtering methodology. Currently, the development of eigenvector spatial filtering theory and methodology focuses on spatial autocorrelation. It can be employed with either linear or generalized linear model descriptions. This methodology can be generalized, with extensions made to serial correlation in time series and to autocorrelation in space–time series. Griffith and Paelinck (2009) furnish a simple mixed-model extension when considering the explicit joint estimation of both space and time lags. Their specifications specifically focus on parametric parsimony together with specification

richness. They show that the bivariate Poisson frequency distribution specification is equivalent to a spatial filter random effects specification.

3 Eigenvector spatial filtering

Filtering of data has a long history. Impulse-response function filtering of time series data predates a parallel approach for spatial filtering and motivates the development of spatial autoregressive linear operators for prewhitening (Cochrane and Orcutt 1949; Tobler 1975). Kalman filtering is another approach to this problem (Bennett 1979). Data analytic spatial filtering views spatial autocorrelation as the outcome of missing variables in a model specification and, accordingly, constructs synthetic proxy variables as surrogates for these missing variables in order to obtain more robust findings in georeferenced data analyses (see Tiefelsdorf and Griffith 2007). Mathematical operators are utilized to separate geographically structured noise from both trend and random noise in georeferenced data, enhancing results by allowing clearer visualization and sounder statistical inference. Nearby/adjacent values are manipulated to adjust the value at a given location, smoothing, reducing variability, and retaining the local features of georeferenced data.

A spatial filter exploits the inter-observational correlation structure of georeferenced data. This structure is represented by matrix $\mathbf{C}$, an $n \times n$ geographic connectivity/weights matrix (e.g., $c_{ij} = 1$ if areal units i and j are nearby/adjacent and $c_{ij} = 0$ otherwise; $c_{ii} = 0$), where n is the number of locations. Often c_{ij} is defined in terms of chess moves: rook adjacency—if two areal units (i.e., locations) share a non-zero length common boundary; and, queen's adjacency—if two areal units share either a zero (i.e., a point) or a non-zero length common boundary. This structure also can be defined in terms of inter-point distances (e.g., geostatistics). Griffith (e.g., 2003) formulates a transformation procedure that depends on mathematical expressions, known as eigenfunctions, that characterize matrix

$$\left(\mathbf{I} - \mathbf{1}\mathbf{1}^{\mathrm{T}}/n\right)\mathbf{C}\left(\mathbf{I} - \mathbf{1}\mathbf{1}^{\mathrm{T}}/n\right) \tag{1}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{1}$ is an $n \times 1$ vector of ones, and T denotes the matrix transpose operator—a matrix appearing in the numerator of the Moran Coefficient (MC) spatial autocorrelation index. Matrix $(\mathbf{I} - \mathbf{1}\mathbf{1}^{\mathrm{T}}/n)$ is a projection matrix commonly encountered in multivariate and regression analysis theory. This decomposition also could be based upon the Geary Ratio, the other popular spatial autocorrelation index, and rests on the following decomposition theorem:

> the first eigenvector, say $\mathbf{E}_1$, is the set of numerical values that has the largest MC achievable by any set for the spatial arrangement defined by the geographic connectivity matrix $\mathbf{C}$; the second eigenvector is the set of values that has the largest achievable MC by any set that is uncorrelated with $\mathbf{E}_1$; the third eigenvector is the third such set of values; and so on through $\mathbf{E}_n$, the set of values that has the largest negative MC achievable by any set that is uncorrelated with the preceding($n-1$) eigenvectors.

As such, these eigenvectors furnish distinct map pattern descriptions of latent spatial autocorrelation in georeferenced variables, because they are both orthogonal and uncorrelated. Their corresponding eigenvalues index the nature and degree of spatial autocorrelation portrayed by each eigenvector (Tiefelsdorf and Boots 1995). This decomposition is similar to that for principal components analysis, except that the eigenvectors themselves are used to construct spatial filters, whereas eigenvectors in principal components analysis supply the linear combination coefficients and the linear combinations are of original attribute variables.

The resulting spatial filter is constructed from some linear combination of a subset of these eigenvectors. Griffith (2008) furnishes an informative application of this approach.

4 Prospects

One important prospect is to extend spatial filtering methodology to space–time datasets. The utility of this venture can be illustrated first with a comparison of the extension of filtering methodology to simple time series analysis.

4.1 A comparative time series analysis: sugarcane production in Puerto Rico

The annual tons of milled sugarcane for the entire island of Puerto Rico are available for the consecutive years beginning in 1828 and ending in 1996, forming a time series with $T = 169$ observations. The magnitude range of these weight figures is 9,391–1,359,841. Their time series plot appears in Fig. 1a. Its logarithmic transformation, which better aligns the values with a bell-shaped curve, appears in Fig. 1b. A first-order differencing (i.e., the difference between the values of two observations adjacent in time) of these transformed values (Fig. 1c) removes some of the non-stationarity in this series but also uncovers several parts of the series that are highly volatile. Two years appear as anomalies: 1879, which was an overproduction year because of a failure of plantation owners to understand capitalistic agricultural markets, and 1899, which was an underproduction year because of the war that year between Spain and the United States (US), which resulted in the US takeover of the island. The respective Shapiro–Wilk (SW) normality diagnostics are as follows: 0.7782 ($p < 0.0001$), 0.9353 ($p < 0.0001$), and 0.9606 ($p = 0.0001$). Griffith (2008) analyzes a slightly modified version of these data in considerable detail.

A rather sophisticated ARIMA model was estimated for this time series involving: a 1st differencing, two autoregressive components for lags 1 and 2, and a periodic moving average component for lag 5. The pseudo-R^2 for the back-transformed values produced by the estimated ARIMA model is 0.9614. Only trace levels of serial correlation can be detected in the residuals for this estimated ARIMA model (Table 1).

In this case, the eigenfunction temporal filter is conditional on including the time index as a covariate in order to maintain an equivalency with the ARIMA specification. In the empirical example analyzed here, the ARIMA specification
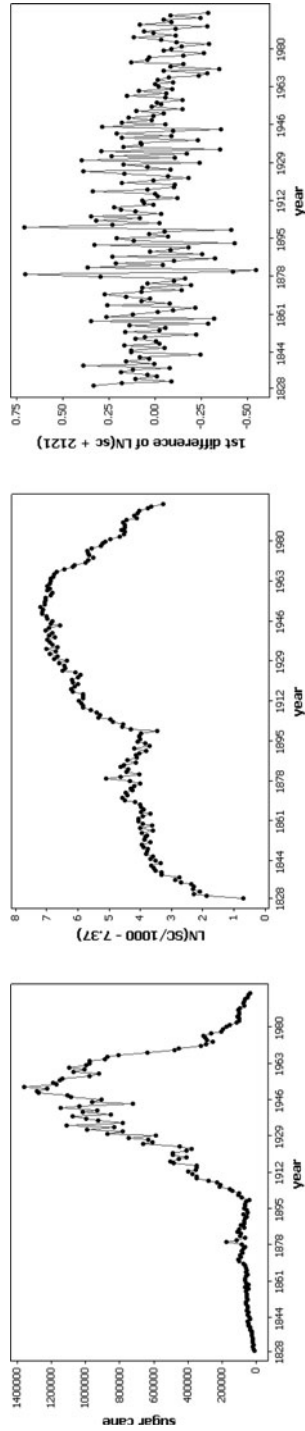
Fig. 1 Time series plots of annual sugarcane production in Puerto Rico. *Left* (**a**): original figures. *Middle* (**b**): logarithmic transformation of figures. *Right* (**c**): 1st—differences of figures. Observed values denoted by *black dots*, and time adjacencies denoted by *gray lines*

Table 1 ARIMA statistics for the tons of sugarcane time series portrayed in Fig. 1

Parameter	Estimate	Standard error	PR > |t|	Lag	Residual serial correlation			
					To lag	Chi-square	df	PR > chi-square
Intercept	0.01605	0.02025	0.4279	0	6	1.85	3	0.6044
AR-1	−0.25742	0.07802	0.0010	1	12	15.53	9	0.0773
AR-2	−0.18860	0.07416	0.0110	2	18	18.30	15	0.2473
SMA	0.18346	0.07708	0.0173	5	24	26.22	21	0.1982

includes a first-order differencing; the linear regression equivalent to this ARIMA specification is detrending a time series with a linear function of the time index, which assumes that this trend is fixed and deterministic. The constructed temporal filter, whose eigenvectors were selected with a stepwise regression procedure, is a linear combination of 42 positive and 5 negative temporal autocorrelation eigenvectors (47 of 141); the candidate subset (from 169) contains 70 positive and 71 negative autocorrelation eigenvectors. Analytical eigenvectors of matrix **C**, which is tridiagonal, for a first-order time series structure are known (Basilevsky 1983); in this unadjusted case, factor analysis can be used to rotate these vectors to the required modified ones. The deterministic trend accounts for roughly 43% of the variation in log-transformed production and 8% in the back-transformed production; the temporal filter accounts for roughly 55% of this variation in log-transformed production and 90% in the back-transformed production. The negative temporal autocorrelation component accounts for only a trace amount of these variations. These constructed filter components are portrayed in Fig. 2a. This filter mirrors the time series pattern: the time trend shifts values upward through time; the positive autocorrelation temporal filter replicates much of the more local variation through time, tracking the 3rd-order polynomial type of trend; and the negative autocorrelation temporal filter accounts for some of the variation in the most volatile part of the series, the period during which growth switches to decline.

A linear regression analysis of these data may contain specification error because the relationships uncovered are non-linear. Given that the transformation is logarithmic, that tons of sugarcane are recorded in integer units (i.e., they always can be rounded to integers), and that tons of sugarcane have a natural lower bound of 0, a rounded off version of this quantity can be approximated with a Poisson RV, avoiding a need to calculate back transformations —which are complicated because an additive error term converts to a multiplicative error term. As an aside, attempts were unsuccessful to use non-linear least squares to estimate an exponential model specification coupled with an additive normally distributed error term. The denominator of 1,000 can be included as an offset in a generalized linear model (GLM) estimation of parameters, and over-dispersion can be accommodated by introducing a gamma-distribute mean for the Poisson RV, converting it to a negative binomial RV. Temporal filtering exhibits an advantage over ARIMA techniques in this context because of its ability to be incorporated in a simple and straight-forward fashion into GLM specifications. Time series detrending still requires inclusion of the time index. The constructed temporal filter is a linear combination of 21 positive
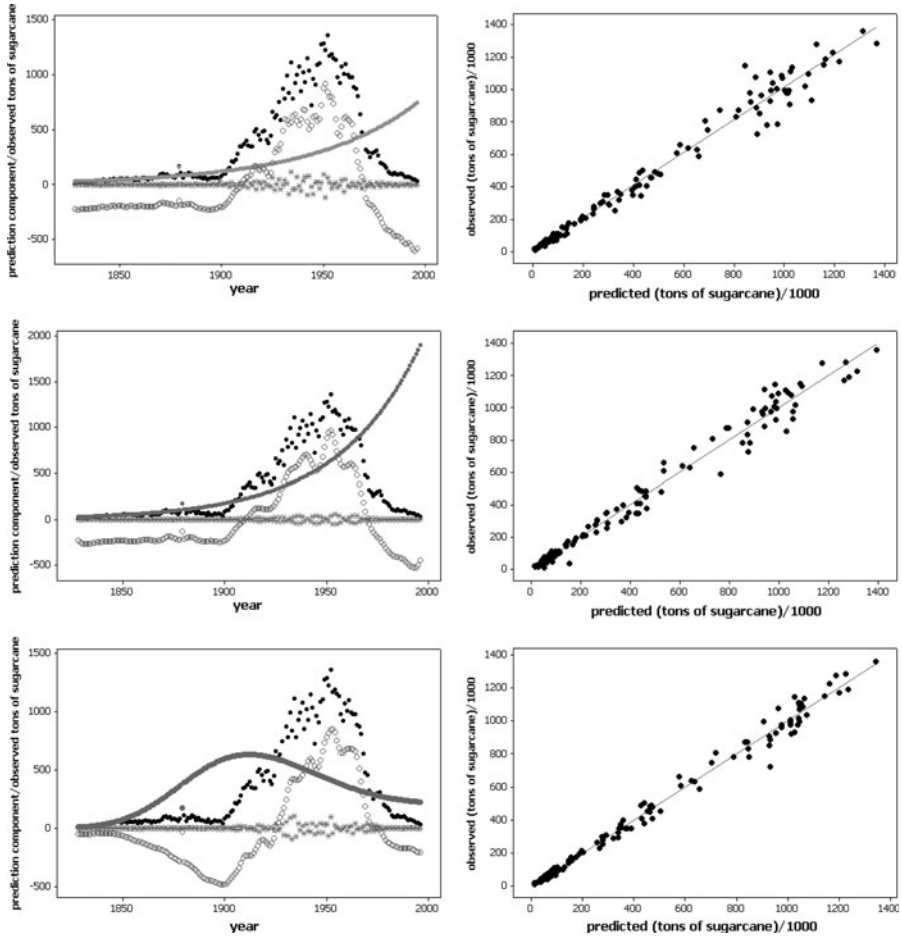
Fig. 2 Decomposition of time series description; *black dots* denote observed, and gray dots denote predicted, tons of sugarcane production, where the time trend is denoted by *solid gray circles*, the positive temporal filter component is denoted by *gray asterisks*, and the negative temporal filter component is denoted by *hollow gray circles*. Top *left* (**a**): back-transformed normal approximation description. *Top right* (**b**): observed versus back-transformed normal approximation predicted values. *Middle left* (**c**): negative binomial GLM description. *Middle right* (d): observed versus GLM predicted values. *Bottom left* (**e**): negative binomial GLM with cubic time trend description. *Bottom right* (**f**): observed versus GLM with cubic time trend predicted values

and 1 negative temporal autocorrelation eigenvectors (22 of 141). In other words, considerable simplicity is achieved with this Poisson approximation, which is accompanied by a deviance statistic of 1.21 (which is very close to 1), vis-à-vis the normal approximation. The deterministic trend accounts for roughly 4% of the variation in production; the temporal filter accounts for roughly 94% of this variation in production. The negative temporal autocorrelation component accounts for only a trace amount of these variations. These constructed filter components are portrayed in Fig. 2c. This filter mirrors the time series pattern: the time trend shifts

values upward through time, more so than for the log-linear normal approximation specification; the positive autocorrelation temporal filter replicates much of the more local variation through time, tracking the 3rd-order polynomial type of trend; and the negative autocorrelation temporal filter accounts for some of the variation in the most volatile part of the series, but less than does its log-linear normal approximation counterpart.

The resulting predictions are extremely good (Table 2; Figs. 2b, d), although the residuals fail to conform to a normal frequency distribution [the SW diagnostic statistic remains significant at the 0.0001 level, deteriorating from 0.937 for the log-transformed to 0.807 for the back-transformed values, and being 0.911 for residuals calculated with the negative binomial GLM expected values]. The global Durbin-Watson statistic (DW; the serial dependency diagnostic index that has an expected value of 2) is as follows for the various variates:

Sugarcane: 0.044
Log-normal approximation residuals: 1.854
Back-transformed residuals: 2.120
Negative binomial GLM residuals: 2.021

The original data display considerable positive temporal autocorrelation, with a DW value close to 0, and the negative binomial GLM residuals display only trace temporal autocorrelation, with a DW value very close to 2.

As a benchmark, a 3rd-order polynomial-in-time negative binomial specification was estimated (Table 2; Figs. 2e, f). It achieves a slightly better overall description ($R^2 = 0.9915$), with 17 positive and 6 negative temporal autocorrelation eigenvectors. The time trend comprises a square and a cubic term, but no linear term (i.e., the first-order differencing factor is gone), and accounts for roughly 18% of the variation in production. Therefore, little is gained with this specification, which requires 25 rather than 23 trend parameters to be estimated.

In conclusion, eigenvector filtering is effective in accounting for autocorrelation in time series. It posits a deterministic model description—functions of the time index need to be used to account for trends that are handled by differencing for ARIMA models, and the set of eigenvectors remains unchanged for a given time horizon and time intervals, regardless of the response variable measured in this

Table 2 Stepwise percentage of variance accounted for by descriptive components

Component	Log-normal approximation	Back trans-formation	Negative binomial GLM	Negative binomial GLM with $(time)^2$ and $(time)^3$
Time trend	0.4343 (3)	0.0813 (3)	0.0426 (2)	0.1792 (3)
Positive temporal autocorrelation	0.5468 (42)	0.8984 (42)	0.9395 (21)	0.8063 (17)
Negative temporal autocorrelation	0.0025 (5)	0.0016 (5)	0.0019 (1)	0.0060 (6)
Total	0.9836	0.9813	0.9836	0.9915

The number of covariates contained in a component is in parentheses

Year-specific indicator variables for 1879 and 1899 sometimes constitute part of a time trend

context—in contrast to a stochastic description furnished by ARIMA models. Its advantage is that it is much simpler to implement, especially for non-normal random variables (RVs), for which it may furnish a sounder implementation. For short time series (i.e., small T), the relevant analytical eigenvectors can be converted to the modified ones needed to construct a temporal filter; these analytical eigenvectors are asymptotically equivalent to the needed ones. Regardless, the advantage of ARIMA models is that a well-developed statistical theory exists for them. Nevertheless, this comparison contributes to a validation of the filtering methodology.

4.2 A space–time series extension of eigenvector filtering methodology: sugarcane production in Puerto Rico

The three-dimensional matrix counterpart to the geographic weights matrix $\mathbf{C}$ appearing in expression (1) for space–time additive components is given by

$$\mathbf{C} = \mathbf{I}_T \otimes \mathbf{C}_s + \mathbf{C}_T \otimes \mathbf{I}_s \tag{2}$$

whereas the three-dimensional matrix counterpart for space–time multiplicative components is given by

$$\mathbf{C} = \mathbf{C}_T \otimes \mathbf{C}_s + \mathbf{C}_T \otimes \mathbf{I}_s \tag{3}$$

where matrix $\mathbf{C}_T$ has 1s in its upper and lower off-diagonals and 0s elsewhere, s denotes space, $\mathbf{C}_s$ is the binary geographic weights matrix, $\otimes$ denotes Kronecker product, $\mathbf{I}_T$ denotes the T × T identify matrix, and $\mathbf{I}_s$ denotes the $n \times n$ identity matrix. Expression (2) is a Kronecker sum; expression (3) is a Kronecker product. Both expressions yield the same set of eigenvectors, namely

$$\mathbf{E}_T \otimes \mathbf{E}_s \tag{4}$$

where $\mathbf{E}_T = \begin{pmatrix} \mathrm{COS}[1t/(T+1)] \\ \vdots \\ \mathrm{COS}[Tt/(T+1)] \end{pmatrix}$, $t = 1, 2, \ldots, T$, and $\mathbf{E}_s$ is the set of n eigenvectors for the geographic surface partitioning under study. These are asymptotic results, when the 1st eigenvector is replaced with a vector of 1s, and assume a data structure in which the geographic data for each point in time are concatenated by time sequence.

A reanalysis of the space–time sugarcane production data analyzed by Griffith (2008) illustrates the utility of this extension. This dataset comprises 16 time periods and 73 areal units (16 × 73 = 1,168). The response variable is hectares of sugarcane harvested, which can be analyzed as a binomial RV when adjusted for the size of areal units (i.e., a percentage of the total area available). The 1,168 × 1,168 modified space–time connectivity matrix [see expression (1)] has 129 eigenvectors with a relative eigenvalue greater than 0.25, a minimum threshold value for positive space–time autocorrelation. Of these 129 vectors, 30 describe space–time autocorrelation in these data, accounting for roughly 53% of the variation in percentage of sugarcane area harvested. The preceding time series example suggests that negative

autocorrelation also should be considered. This phenomenon is rarely encountered in spatial series but is commonly encountered in time series. Including the 130 additional eigenvectors describing negative autocorrelation results in none being selected by the SAS stepwise logistic regression procedure. In other words, no negative space–time autocorrelation is detected in these data. The constructed positive autocorrelation space–time filter reduces the extra-binomial variation, as measured by the deviance statistic, from 3,384 to 1,540, clearly failing to account for all of this overdispersion while accounting for a substantial amount of it. The analysis overlooking autocorrelation yields the following estimate of the probability p: $\hat{p} = 0.89$; when autocorrelation is accounted for, this value becomes 0.92—as expected, the principal impact is on the variance, with little impact on the mean response. Overall, then, this approach exhibits considerable promise.

5 Challenges suggesting a research agenda

A number of challenges for space–time data analysis research are highlighted or alluded to in this article. Foremost is that panel data models need to be able to accommodate lag relationships in both space and time. A spatial filter specification offers one way to achieve this end, allowing spatial autocorrelation to be accounted for by a spatially structured random effects term and serial correlation to be accounted for by a time lag. Currently, a random effects term needs to be estimated first, followed by decomposing it into its spatially structured and spatially unstructured components. Explicit estimation of these components is preferable. In addition, non-normal RVs need to be treated directly, rather than subjected to some variable transformation. Again, filtering methodology should prove helpful here. Accordingly, distribution theory needs to be fully developed for space–time spatial filters.

One weakness of the STAR model is that it posits either a single, global space–time autocorrelation parameter or two parameters, one for spatial and the other for temporal autocorrelation. Most space–time data display far more heterogeneity than can be adequately captured by only one or two parameters. This is one reason why the multivariate ARIMA models can furnish better space–time data descriptions. Again, the space–time spatial filter represents a more heterogeneous trend descriptor, with temporal structure eigenvectors accounting for variation over time and spatial structure eigenvectors accounting for variation over space. The Kronecker product of these two sets of eigenvectors allows for interaction effects between the time and space dimensions to be accounted for, too. But a current limitation of space–time filtering methodology is a lack of visualization tools for it, another area demanding subsequent research.

Many space–time data series include missing data: instruments malfunction, government agencies suppress or temporarily discontinue collecting data for selected locations, events obscure or corrupt data collection. Regardless, spatial scientists wish to analyze data series whether or not they are incomplete. Imputation techniques to handle this situation also benefit cross-validation and jackknife resampling evaluations in the presence of space–time autocorrelation. For example, results for five replications of a single randomly suppressed value in the Puerto Rico

Table 3 Space-time filter imputations for selected suppressed hectares of harvested sugarcane data, based upon a binomial regression model

Time	Areal unit	Observed hectares	Cross-validation estimate	Imputation
1971	Hormigueros	4,284	1,756	1,743
1982	Utuado	0	1,433	1,469
1975	Patillas	4,039	5,024	5,083
1979	Ciales	51	2,420	2,543
1981	Yauco	2,437	1,543	1,526

sugarcane space–time data series appear in Table 3. The space–time filter imputations are consistent with the cross-validation predictions. Imputations were calculated by setting the probability to ½ and including a missing value indicator variable in the binomial regression equation. Of note is that the imputations can be bolstered by including meaningful covariates in the space–time model specification and can be easily extended to the case of multiple missing data values.

Finally, more comparative analyses between the various space–time model specifications need to be undertaken. And more empirical case studies need to be produced.

References

Baltagi B (2005) Econometric analysis of panel data. Wiley, Hoboken, NJ

Basilevsky A (1983) Applied matrix algebra in the statistical sciences. North-Holland, New York

Bennett R (1979) Spatial time series. Pion, London

Cliff A, Haggett P, Ord J, Bassett K, Davies R (1975) Elements of spatial structure. Cambridge University Press, Cambridge

Cochrane D, Orcutt G (1949) Application of least squares regression to relationships containing autocorrelated error terms. J Am Stat Assoc 44:32–61

Elhorst J (2001) Dynamic models in space and time. Geogr Anal 33:119–140

Gneiting, T, Genton M, Guttorp P (2006) Chapter 4—statistical methods of spatio-temporal systems. In: Finkenstaedt B, Held L, Isham V (eds). Boca Raton, Chapman-Hall, pp 151–175

Griffith D (1996) Spatial statistical analysis and GIS: exploring computational simplifications for estimating the neighborhood spatial forecasting model. In: Longley P, Batty M (eds) Spatial analysis: modelling in a GIS environment. Longman GeoInformation, UK, pp 255–268

Griffith D (2003) Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization. Springer, Berlin

Griffith D (2008) A comparison of four model specifications for describing small heterogeneous space-time datasets: sugar cane production in Puerto Rico, 1958/59–1973/74. Pap Reg Sci 87:341–356

Griffith D, Paelinck J (2009) Specifying a joint space- and time-lag using a bivariate Poisson distribution. J Geograph Syst 11:23–36

Hartfield M, Gunst R (2003) Identification of model components for a class of continuous spatiotemporal models. J Agric Biol Environ Stat 8:105–121

Heuvelink G, Griffith D (2010) Space-time geostatistics for geography: a case study of radiation monitoring across parts of Germany. Geogr Anal 42: in press

Ma C (2008) Recent developments on the construction of spatio-temporal covariance models. Stoch Env Res Risk Assess 22:S39–S47

Stein M (2005) Space-time covariance functions. J Am Stat Assoc 100:310–321

Tiefelsdorf M, Boots B (1995) The exact distribution of Moran's I. Environ Plan A 27:985–999

Tiefelsdorf M, Griffith D (2007) Semi-parametric filtering of spatial autocorrelation: the eigenvector approach. Environ Plan A 39:1193–1221

Tobler W (1975) Linear operators applied to areal data. In: Davis J, McCullagh M (eds) Display and analysis of spatial data. Wiley, London, pp 14–37