Spatial-Temporal Models and Computational Statistics Methods: A Survey

Erniel B. Barrios
Kevin Carl P. Santos
School of Statistics
University of the Philippines Diliman

We introduce panel models and identify its link to spatial-temporal models. Both models are characterized and differentiated through the variance-covariance matrix of the disturbance term. The resulting estimates or tests are as complicated as the nature of the said variance-covariance matrix. Some iterative methods typically used in computational statistics are also presented. These methods are used in conducting statistical inference for spatial-temporal models.

Keywords: *panel data, spatial-temporal model, forward search algorithm, additive models, backfitting algorithm, isotonic regression*

1. Panel Models

In finite population sampling, the sample selection procedure is carefully chosen to ensure independence of the resulting samples. Independence of the sample observations facilitates statistical inference and simplifies modeling procedure. Specific population patterns are also inherited and clearly manifested by the sample. Oftentimes, samples are visited more than once to monitor progress of indicators of interest and, as time progress, a short time series is generated among the sample units. While independence is maintained among sample units, temporal dependence is induced into the so-called panel data. When these data are used in statistical inference, a more complex form of the variance-covariance matrix (non-diagonal), often leading to estimates that has no closed form, will be required.

The roots of spatial-temporal models can be traced from the panel data models. Panel data models have been vital in analyzing the dynamics of the cross-sectional units at different time points that cannot be captured by simple cross-sectional data analysis. Baltagi (1995) defined panel data as a set of observations

such as households, companies, provinces, or countries, gathered across different time periods. The panel data analysis (in contrast to cross-section analysis, or time series analysis per unit) yields more information, more variability, less collinearity among variables, more degrees of freedom, and more efficiency. Panel data are usually gathered from microlevel units. This makes it more appealing because measurements can be more accurate at this level. Accuracy of measurements are achieved from the elimination of bias resulting from aggregation over individuals.

A panel data regression model can be written as follows:

$$Y_{it} = \alpha + \mathbf{X}_{it}'\beta + \varepsilon_{it},\ i = 1,2,...,N\ t = 1,2,...,T \tag{1}$$

The subscript i denotes the cross-sectional unit, whereas, t denotes the temporal dimension. The regression coefficient α is a scalar, $\boldsymbol{\beta}$ is a k x 1 vector, and $\mathbf{X}_{it}$ is the vector of the covariates for the i^{th} observation at time t. While the units i are indexed over space, they are independent, assured by the choice of a sample selection method. On the other hand, there is inherent temporal dependence because the same units are visited at several periods in time. This is captured into the structure of the variance-covariance matrix of ε_{it}.

There are two important models in panel data analysis, namely, the fixed effects model and the random effects model. The two models differ on the way dependencies of the error terms are specified.

Oftentimes, panel data analysis employs one-way error component model for the disturbance. It is given by

$$\varepsilon_{it} = \mu_i + \upsilon_{it},\ i = 1,2,...,N\ t = 1,2,...,T, \tag{2}$$

where μ_i represents the unobservable unit-specific effect and υ_{it} refers to the remainder disturbances. μ_i is time-invariant and accounts for any unit-specific effect not accounted for by the regression model. Meanwhile, υ_{it} varies with individuals and time. It is parallel to the usual disturbance term in regression models.

When time-specific effect characterizes the disturbance in addition to individual-specific effect, the two-way error component model can be specified to accompany Equation (1) and defined as follows:

$$\varepsilon_{it} = \mu_i + \lambda_t + \upsilon_{it},\ i = 1,2,...,N\ t = 1,2,...,T, \tag{3}$$

where μ_i is the same unobservable individual effect in Equation (2), λ_t is the unobservable time-specific effect, and υ_{it} is the remainder stochastic disturbance term. Note that λ_t is individual-invariant and it accounts for any time-specific effect that is not included in the regression model but varies over time. On the other hand, μ_i is constant over time among individuals i.

Panel data models assume independence of the cross-sectional units. However, several cross-sectional units are also spatially dependent with one another. In this case, the estimates become biased and inconsistent. Thus, Lee and Yu (2008) addressed this by allowing adjacent observations to be correlated, hence, generalizing the panel data model to include spatial autoregressive (SAR) structure. They proposed an alternative method based on transformation for the estimation of spatial autoregression panel models with fixed effects. Provided that either n or T are large enough, the proposed method produces consistent estimators for all the parameters of models with individual effects only and with both individual and time effects.

Furthermore, Baltagi et al. (2003) considered a spatial panel regression model with serial correlation over time for each cross-sectional unit and spatial dependence across these units at a particular time point. The model allowed for heterogeneity across the spatial units postulated as random effects.

2. Spatial-Temporal Models

As more complex dependency between units and among units over time becomes a reality than an exceptional case, panel data model has been generalized into a spatial-temporal model. Models that take into account space and time dependency are of great demand due to more efficient data collection. For instance, studying the prevalence of air pollutants in different locations across time would require spatial-temporal models, which can be used to unveil meaningful patterns. In epidemiology, such models can be used for disease mapping so as to prevent spread of certain disease. Monitoring food prices or inflation rates across provinces in different time points may also necessitate models accounting for space-time relationship. Patterns on crime incidents across time and space can also be recognized with the use of such models. Water levels in flood prone areas in different time periods can be mapped using spatial-temporal models. Hering et al. (2009) and Waller et al. (2007) illustrated the applications of spatial-temporal modeling in studying foot-and-mouth disease, wildfire ignition point patterns, and country-level incidence and reporting of Lyme disease, respectively. Indeed, spatial-temporal models are used to make smoother predictions of some response variable over certain spatial domain and to discern patterns and relationships across time and space.

Spatial-temporal models take into account the time and space correlation of observations. But initially, literature on statistical modeling treated space and time dependency separately. However, oftentimes, observations that are contiguous with one another in space are similar and so are those which are adjacent in time. Thus, with such data points, the error terms would be autocorrelated over time and spatially dependent over an area for a fixed time point. As a result, the spatial-temporal models are proposed. Landagan and Barrios (2007) explains agricultural production in terms of a spatio-temoral model, given in

$$Y_{it} = \mathbf{X}_{it}' \boldsymbol{\beta} + \mathbf{W}_{it} \boldsymbol{\gamma} + \varepsilon_{it},\ i = 1,2,...,N\ t = 1,2,...,T, \quad (4)$$

where $\mathbf{X}_{it}$ is the vector of covariates from location i at time t, $\mathbf{W}_{it,}$ the vector of spatial indicators for location i at time t, and an error term, $\varepsilon_{it} = \mu_i + \upsilon_{it}$. The individual effect, μ_i, is assumed to be independently and identically distributed with zero mean and constant variance, whereas, the remainder of the error, $\upsilon_{it,}$ is an autoregressive process of order p. The predictors, $\mathbf{X}_{it,}$ vary across individual units i, whereas, the spatial indicators, $\mathbf{W}_{it,}$ have common values for a group of units within a predefined spatial neighborhood. W_{it} emphasizes the spatial externalities common among units in the neighborhood. For example, $\mathbf{W}_{it}$ is the amount of rainfall, constant among farmers in a small community of contiguous farms, whereas, $\mathbf{X}_{it}$ are individual farmer's production capabilities, both sets affecting yield Y_{it} for a particular crop. It was shown that the proposed method is efficient and the iterative estimation procedure of parameters usually converges after second or third iterations. Dumanjug et al. (2010) modified the procedure by Landagan and Barrios (2007) by incorporating the block bootstrap into the hybrid of the backfitting algorithm and the Cochranne-Orcutt procedure. They introduced two bootstrap procedures, one using independent block bootstraps, and the other using overlapping blocks of consecutive vectors.

In analysis of the spatial dimension, the goal is to separate out the relationship existing between spatial units that is attributed to common factors, suggesting strong dependence, from that which is purely spatial. Recently, Bailey, Holly and Pesaran (2016) presented a two-stage approach to spatial-temporal analysis with strong and weak cross-sectional dependence. To do this, they extracted common factors first, which are source of strong cross-sectional dependence, using cross-unit averages, and compared these with the principal components approach. Once the observations are de-factored, they conducted multiple testing procedures to determine significant bilateral correlations between spatial units. They compared this to simply using distance to determine units that are neighbors.

Due to the dynamic nature of many real-world phenomena, the approach for spatial-temporal models that use the covariances to characterize the spatial-temporal dependence may be inferior compared to the use of dynamical models. However, the major problem associated with those models is the curse of dimensionality. Wilke and Hotten (2010) presented a general science-based framework for dynamical spatial-temporal model. In this framework, they demonstrated their proposed method for the linear case, and then developed a general nonlinear spatio-temporal framework, called general quadratic nonlinearity. They showed that this framework, which is presented in a hierarchical Bayesian framework, subsumes a large class of scientific-based parameterizations.

A more recent paper by Bradley, Holan, and Wikle (2015) proposed multivariate spatial-temporal models for high-dimensional areal data as a possible remedy for

practical issues when formulating statistical models. In particular, they postulated the so-called multivariate spatial-temporal mixed effects model (MSTM) that can efficiently model high-dimensional multivariate spatial-temporal areal datasets as illustrated by its good predictive ability. Using a fully Bayesian methodology, this model generalizes the notion of Moran's I basis functions to the multivariate spatial-temporal setting.

3. Spatial-Temporal Process

Enormous amount of data referenced through space and time can now be gathered due to recent advances in science and technology. Spatial-temporal models treat these observations as a partial realization of a spatial-temporal random process denoted by

$$\{Z(s;t): s \in D \subset \Re^d; t \in (0,\infty)\} \tag{5}$$

observed at N space-time coordinates $(s_1;t_1)$, $(s_2;t_2)$,..., $(s_N;t_N)$. As noted by Cressie and Huang (1999), the spatial-temporal process $Z(\cdot;\cdot)$ satisfies the regularity condition, $Var\left(Z(s;t)\right) < \infty$, i.e., the second moment of $Z(s;t)$ exists and is finite, for all $s \in D, t \geq 0$. The mean and covariance functions are defined as follow:

$$\mu(s;t) \equiv E\left(Z(s;t)\right) \text{ and } K(s,r;t,q) \equiv \text{cov}\left(Z(s;t), (r;q)\right) \tag{6}$$

$s,r \in D$, $t,q>0$ respectively. Gneiting et al. (2007) noted that optimal least-squares prediction, or kriging, relies on the suitable specification of the covariance structure of the spatial-temporal process. Furthermore, the covariance between $Z(s;t)$ and $Z(r;q)$, in general, depends on the space-time coordinates $(s;t)$ and $(r;q)$ and no further structure may exist. In practice, assumptions such as stationarity, separability, and full symmetry are required in estimation and modeling procedures.

If $\text{cov}\left(Z(s;t), (r;q)\right)$ depends on the spatial coordinates s and r only through the spatial separation vector, $s-r$, then the spatial-temporal random process has *spatially stationary* covariance. On the other hand, the spatial-temporal random process has *temporally stationary* covariance if $\text{cov}\left(Z(s;t), (r;q)\right)$ depends on the time points t and q only through the temporal lag, $t-q$. If a spatial-temporal process has both spatially and temporally stationary covariance, then the process is said to have a *stationary* covariance, that is, $K(s,r;t,q) \equiv C(s-r;t-q)$ for certain functions C. Cressie and Huang (1999) observed that this stationarity condition is often assumed so that the covariance function can be estimated from data. In order to be a valid covariance function, the function C should satisfy a positive-definiteness condition, i.e., for any $(r_1;q_1)$, $(r_2;q_2)$,...$(r_m;q_m)$, any real a_1, a_2 ,...a_m, and any positive integer m, C must satisfy

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j C(r_i - r_j; q_i - q_j) \geq 0. \tag{7}$$

Gneiting et al. (2007) further noted that the second-order structure of a spatial-temporal random process is sometimes modeled based on variances rather than covariances. Viewed as a function of space-time coordinates $(s;t)$ and $(r;q)$ in $\Re^d$ x $\Re$, the non-stationary variogram is given by

$$\frac{1}{2}\text{var}\big(Z(s;t) - Z(r;q)\big). \tag{8}$$

If the non-stationary variogram depends on the locations s and r only through the spatial separation vector, $s-r$, then the spatial-temporal random process has *spatially intrinsically stationary* variogram. Likewise, if the non-stationary variogram depends on time points t and q only through the temporal lag, $t-q$, then the spatial-temporal random process is said to have a *temporally intrinsically* variogram. Thus, if the process has both spatially intrinsically stationary and temporally intrinsically stationary variogram, then it has intrinsically *stationary* variogram. Under this condition, there exists a function γ defined on $\Re^d$ x $\Re$ such that

$$\frac{1}{2}\text{var}\big(Z(s;t) - Z(r;q)\big) = \gamma(s-r;\ t-q) \tag{9}$$

for all $(s;t)$ and $(r;q)$ in $\Re^d$ x $\Re$. The function γ is called the *stationary variogram* of the process $Z(\cdot;\cdot)$. Gneiting et al. (2007) also stated that variograms exist under less stringent assumptions than covariances.

Usually, the covariance function C is chosen to belong to a parametric family whose members are known to be positive-definite. That is, one assumes

$$\text{cov}\big(\text{Z}(s;t),\, \text{Z}(s+h;t+u) = C^0(h;u|\theta) \tag{10}$$

where C^0 satisfies positive-definiteness for all $\theta \in \Theta \subset \Re^p$.

Particularly, the spatial-temporal random process $Z(\cdot;\cdot)$ is said to have a *separable* covariance if the covariance between two distinct space-time coordinates can be expressed as a product of a purely spatial and a purely temporal covariance functions C^1 and C^2, respectively, that is,

$$C^0(h;u|\theta) = C^1(h|\theta_1)\, C^2(u|\theta_2) \tag{11}$$

where C^1 is a positive-definite function in $\Re^d$, C^2 is a positive-definite function in $\Re^1$ and $\theta' = (\theta_1'\theta_2')$. Equation (11) is called the *product model*. Separable models are usually chosen for convenience in estimation and inference rather than for their ability to fit the data well (Cressie and Huang, 1999).

However, spatial-temporal random process with a separable covariance is extremely limited because it does not take into account the interaction between space and time. For any two fixed spatial lags h_1 and h_2, notice that

$$C^0(h_1;u) \propto C^0(h_2;u),\ u \in \Re. \tag{12}$$

This means that, for two spatial locations, the cross-covariance function between the time series at each site will always have the same shape despite the relative displacement of the locations. The same holds true for any pair of time points and the cross-covariance of the two spatial processes.

Cressie and Huang (1999) described another type of separability by adding spatial and temporal covariances, i.e.,

$$C^0(h;u|\theta) = C^1(h|\theta_1) + C^2(u|\theta_2) \tag{13}$$

Model (13) is a linear model and the covariance matrices of certain configurations of spatial-temporal data are singular, which is unsatisfactory when using the optimal (minimum mean squared prediction error) linear predictor of $Z(s_0;t_0)$ given by

$$Z^*(s_0;t_0) = \mu(s_0;t_0) + \mathbf{c(s_0;t_0)'\Sigma^{-1}(Z-\mu)} \tag{14}$$

where $\mathbf{\Sigma} \equiv \mathbf{cov(Z)}$; $\mathbf{c(s_0;t_0)} \equiv \text{cov}(\mathbf{Z(s_0;t_0), Z})$, and $\boldsymbol{\mu} \equiv E(\mathbf{Z})$.

Furthermore, Cressi and Huang (1999) derived a new approach to obtain classes of nonseparable, spatial-temporal stationary covariance functions which is called the *nonseparable model*. The postulated model is given by

$$H(\omega,h_t) = \rho(\omega,h_t)\, K(\omega) \tag{15}$$

where $H(\omega,h_t) = \dfrac{1}{(2\pi)^d}\int e^{-ih_s^T\omega} C_{st}(h_s,h_t)dh_s$, and should satisfy the following conditions:

(1) for each $\omega \in \Re^d$, $\rho(\omega,\cdot)$ is a continuous autocorrelation function and $K(\omega) > 0$;
(2) the positive function $K(\omega)$ satisfies: $\int K(\omega)d\omega < \infty$.

De Cesare et al. (2001) introduced a class of product-sum covariance models for estimating and modeling space-time correlation structures. The very general product-sum model can be obtained in the following way:

$$c_{st}(h_s,h_t) = k_1 C_s(h_s)C_t(h_t) + k_2 C_s(h_s) + k_3 C_t(h_t). \tag{16}$$

or, equivalently,

$\gamma_{st}(h_s,h_t) = [k_{2+} k_1 C_t(0)]\,\gamma_s(h_s) + [k_{3+} k_1 C_s(0)]\,\gamma_t(h_t) - k_1\,\gamma_s(h_s)\,\gamma_t(h_t)$, where C_s and C_t are covariance functions and γ_s and γ_t are the corresponding variogram functions. Note that $C_{st}(0)$ is the "sill" of γ_{st} and $C_t(0)$ is the "sill" of γ_t. Moreover, by definition, $\gamma_{st}(0,0) = \gamma_t(0) = \gamma_s(0) = 0$. Note that the product-sum model is more flexible than the nonseparable covariance model for estimating and modeling spatial-temporal correlation structures. It is because the spatial locations need not be on a regular grid while the temporal points are regularly spaced; hence, it is not necessary to use temporal distance classes.

De Iaco et al. (2001) generalized the product-sum covariance model (De Cesare et al., 2001) into a non-separable model. Moreover, it cannot be derived from the Cressie-Huang representation because, in general, it is non-integrable in space-time. The generalized product-sum model (De Iaco et al., 2001) is given by

$$\gamma_{st}(h_s, h_t) = [k_2 + k_1 C_t(0)]\,\gamma_s(h_s) + [k_3 + k_1 C_s(0)]\,\gamma_t(h_t) - k_1\,\gamma_s(h_s)\,\gamma_t(h_t) + k_4\,\gamma_{1s}(h_s) + k_5\,\gamma_{2t}(h_t) \quad (17)$$

where γ_s and γ_t are (bounded) valid spatial and temporal variograms, γ_{1s} and γ_{2t} are valid spatial and temporal variograms (not necessarily bounded), k_4 and k_5 are non-negative constants. Because model (17) depends solely on the parameter k, De Iaco et al. (2001) addressed the problem of computing the parameter k by proposing simple methods for estimating and modeling the variogram components using data from realizations of spatial-temporal random process.

There were several proposals in modeling procedures for spatial-temporal random process. Gneiting (2002) considered general classes of nonseparable, stationary covariance functions for spatial-temporal random process. These classes do not depend on closed-form Fourier inversions unlike that of Cressie and Huang (1999). Ma (2003a) addressed the problem of incorporating the interaction of a purely spatial process and a purely temporal process by proposing a mixture method. In this method, only simple modeling procedure is needed to fit real data. Moreover, Ma (2003b) illustrated how to obtain nonstationary covariance functions that take into account the space-time interactions. Meanwhile, De Iaco et al. (2003) extended the use of marginal variograms in modeling space-time process to the multivariate cases. They also suggested the use of Linear Coregionalization Model (LCM) for co-krigging in spatial-temporal random process. Porcu et al. (2007) postulated to adapt the Dagum survival function to be a function of space and time. The resulting class has interesting mathematical properties.

4. Computational Statistics Methods

Computational Statistics is the development of statistical theory and methods using computer-intensive algorithms. Many approaches may exhibit interesting

ad hoc characteristics but closed-form solutions are not readily available. The solutions are often iterative and requires substantial empirical approximations. Thus, with the aid of high-speed computing, such methods become implementable and new theories are developed in the process. Some methods and techniques that are used in the study of spatial-temporal models in the perspective of computational statistics are summarized in this section.

4.1 Bootstrap methods

Introduced by Efron (1979), bootstrapping is described as a computationally intensive and nonparametric method for making probability-based inferences about a population parameter based on an estimator by using only a sample from the population of interest (Mooney and Duval, 1993). The nonparametric framework does away with stringent distributional assumptions. This method makes use of large numbers of repetitive computations to approximate the form of the empirical distribution function (EDF). It involves 'resampling' from the sample with replacement a large number of times to obtain an empirical estimate of the distribution of a parameter of interest. This method is widely used because of the significant reduction in numerical computational costs as noted by MacKinnon (2006). Mooney and Duval (1993) outlined the basic steps of the bootstrap procedure.

Let F be an unspecified probability distribution, $X_i \overset{ind}{\sim} F$, and x is the sample realization of X.

(1) Construct an empirical probability distribution, $\hat{F}(x)$, from the sample and draw a simple random sample (SRS) of size n with replacement B number of times, where B is a large number.

(2) Calculate the statistic of interest, $\hat{\theta}_b^*, b = 1, 2, ..., B,$ for each of the resample obtained in (1).

(3) Construct a probability distribution, $\hat{F}^*(\hat{\theta}^*)$, from B $\hat{\theta}_b^{*}$'s, with each estimate having a probability of $\frac{1}{B}\cdot$ $\hat{F}^*(\hat{\theta}^*)$, referred to as the bootstrap estimate of the sampling distribution of $\hat{\theta},$ is then used to make inferences about the parameter of interest $\theta.$

The bootstrap estimate and its variance are given by the following:

$$\hat{\theta}_{bootstrap} = \frac{\sum_{b=1}^{B} \hat{\theta}_b^*}{B} \text{ and } Var\left(\hat{\theta}_{bootstrap}\right) = \frac{\sum_{b=1}^{B} \left(\hat{\theta}_b^* - \hat{\theta}_{bootstrap}\right)^2}{B} \tag{18}$$

The bootstrap smoothens the estimate by reducing the bias that the nuisance of the sample selection might induce. The method is also applicable for variance estimation of statistics with complex sampling distributions and for statistical modeling such as regression analysis.

Analytical properties of this method were already established, e.g., consistency provided by the Glivenko-Cantelli Theorem, which states that: Given $x_1, x_2, ..., x_n \overset{\text{iid}}{\sim} F$, let $F_n(x) = \frac{1}{n}\sum_{m=1}^{n} I_{(x_m \le x)}$, as n$\to\infty$, $\sup_x \left|F_n(x) - F(x)\right| \to 0$ a.s. Davison and Hinkley (1997) observed that error in the bootstrap are classified into statistical error and simulation error. The statistical error is the small difference between the true distribution F and the estimate $F_n(x)$. The simulation error is a result of the approximation in the simulation of the properties of the statistics and of factors like resample size and replication size.

Using the EDF $\hat{F}(x)$, bootstrap procedures can be used to construct confidence intervals and tests of hypotheses. Paparoditis and Politis (2003), Park (2003), and Ioannidis (2005) used bootstrap methods in testing unit roots in time series models, whereas, Westerlund and Edgeworth (2007) proposed a bootstrap procedure for testing cointegration in panel data.

When Efron (1979) introduced this procedure, it was intended for models that assume independent observations. Eventually, the use of the procedure was extended to models assuming dependent observations such as those in time series data. Modifications were made because resampling of time series data is different from that of independent data. The generation of the empirical distribution in time series data is more complicated than in independent data (Bühlmann, 2002). The bootstrap procedure has become a powerful nonparametric method for estimation; however, it fails when the dependency in the observations is not regarded according to Bühlmann (1997).

Different bootstrap procedures proposed for time series data were compared by Bühlmann (2002). The common approaches for stationary time series are model-based approach and blockwise bootstrap. The former approach resamples from independently and identically distributed (*i.i.d.*) residuals. However, because of the assumed model, this approach is prone to model misspecification and the nonparametric nature of bootstrap is lost. On the other hand, blockwise or moving blocks bootstrap is a nonparametric and model-free approach. In this method, the time series is divided into blocks and then the blocks of observations are resampled instead of the individual observations. One advantage of this approach is its robustness to model misspecification. Nevertheless, the resampled series exhibits artifacts when joining randomly selected blocks. Bühlmann (2002) cautioned that the bootstrap samples may lose conditional stationarity once the block dependence is disregarded. Politis and Romano (1994) proposed a

modification of the blockwise bootstrap that produces conditionally stationary bootstrap samples.

Bühlmann (2002) presented sieve bootstrap which is another bootstrap procedure intended for time series data. This approach fits parametric models first, then, the residuals of the estimated model are resampled. The steps in constructing the sieve bootstrap are summarized by Bühlmann (2002) below:

(1) Choose the starting values for each replicate, e.g. equal to zero.

(2) Generate an AR(p(n)) process according to $\sum_{j=0}^{p(n)} \hat{\varphi}_{j,n}(X^*_{t-j} - \bar{X}) = \varepsilon^*_t$ until stationarity is reached. Then, the first set of generated values is discarded.

The bootstrap statistics T^*_n are defined by $T^*_n = T_n(X^*_1, X^*_2, ..., X^*_n)$, where $T_n = T_n(X_1, X_2, ..., X_n)$ is any statistic and T_n is a measurable function of n observations.

4.2 Additive models and the backfitting algorithm

Augmented to the traditional linear model, additive models are a nonparametric generalization of the linear regression model. Because of its nonparametric nature, nothing is imposed on the form of the dependency of the response variable Y on the predictors X_1, X_2, ..., X_p. Furthermore, they retain the important feature of a model, its interpretability, because they are additive in the predictor effects. These models can also help approximate the shape of each of the predictor effect. Hastie and Tibshirani (1990) defined an additive model as follows:

$$Y = \alpha + \sum_{j=1}^{p} f_j(X_j) + \varepsilon, \tag{19}$$

where the error terms ε are independent of the X_j's, $E(\varepsilon) = 0$ and Var $(\varepsilon) = \sigma^2$. The f_j's are arbitrary univariate functions, one for each explanatory variable. The f_j's can be viewed as smooth functions that can be individually estimated by a scatterplot smoother in an iterative manner. Landagan and Barrios (2007) described a smoother as a tool for summarizing the trend of a response measurement Y as a function of one or more explanatory variables X_1, X_2, ..., X_p. In addition, it yields an estimate of the trend that is less variable than Y itself.

Friedman and Stuelzle (1982) noted that one of the most widely used procedures for estimating additive models is the backfitting algorithm. It is a simple and iterative method where a response variable is expressed as a sum of functions of the explanatory variables. The algorithm cycles through the predictors and replaces each current function estimate by a curve based from smoothing a partial residual on each predictor (Hastie and Tibshirani, 1990).

Consider modeling the dependence of Y on two variables X_1 and X_2. The multiple linear regression model given by

$$E(Y \mid X_1, X_2) = \alpha + \beta_1 X_1 + \beta_2(X_2) \tag{20}$$

may be inappropriate if the regression surface $E(Y \mid X_1, X_2)$ cannot be represented by a plane, but an additive model can be considered as a possible solution for this problem. To illustrate the backfitting algorithm applied to modeling $E(Y|X_1, X_2)$. Hastie and Tibshirani (1990) consider the following additive model

$$E(Y|X_1, X_2) = \alpha + f_1(X_1) + f_2(X_2). \tag{21}$$

The algorithm assumes that given an estimate of $f_1(X_1)$ denoted by $\hat{f}_1(X_1)$, a logical approach in estimating $f_2(X_2)$ is to smooth the residuals $Y - \hat{f}_1(X_1)$ on X_2. With this estimate $\hat{f}_2(X_2)$, another estimate of $f_1(X_1)$ is cycled by smoothing the residuals $Y - \hat{f}_2(X_2)$ on X_1. This is continued until the estimates of $f_1(X_1)$ and $f_2(X_2)$ are such that the smooth of $Y - \hat{f}_2(X_2)$ on X_1 is $\hat{f}_1(X_1)$, and the smooth of $Y - \hat{f}_1(X_1)$ on X_2 is $\hat{f}_2(X_2)$. In other words, the cycle continues until the individual components $\hat{f}_1(X_1)$ and $\hat{f}_2(X_2)$ stabilize. Hastie and Tibshirani (1990) outlined the algorithm as follows:

(1) Initialize: $\alpha = ave(Y_i), f_i = f^0_j, j = 1,2,...,p$.

(2) Cycle: $j = 1, 2,..., p$

$$f_j = S_j[(Y - \alpha - \sum_{k \neq j} f_k \mid X_j],$$

where $S_j(Y|X_j)$ denotes a smooth of the response variable Y against the predictor X_j, and produces a function.

(3) Continue (2) until the individual functions do not change.

Given that the dependence structure of the data is not too strong, the backfitting algorithm can also be used to time series data according to Chen and Tsay (1993). Because of this, the algorithm can also be applied to spatial-temporal models; however, it poses caveats in the temporal dependence of the data.

4.3 Forward search algorithm

The forward search algorithm is a powerful technique for detecting multiple masked outliers. It can also determine the effects of the atypical observations on fitted models and minimize their effects in model estimation and yield robust estimates of the parameters. Moreover, it also tests the presence of systematic model inadequacy (Riani, 2004).

Atkinson and Riani (2002) summarized the steps in the implementation of the forward search algorithm. It begins with the choice of the initial subset,

which is a robustly chosen subset of the sample. The subset should be ideal or outlier-free. Afterwards, the search continuously fits the model to a larger subset by successively adding an observation in the estimation procedure. During the search, diagnostic checking is done using the quantities that indicate the adequacy of the model. If there are no outliers, the plots of the parameter estimates and the residuals should remain the same as the size of the subset becomes large. Atypical observations are included at the end of the search.

The forward search algorithm is more popularly used in model-building like those in regression analysis. As illustrated by Atkinson and Riani (2007) in regression analysis, the algorithm involves fitting the least squares regression model each time the subset of observations is increased until all the units are exhausted. To avoid refitting the least squares regression model to all units in each step, an algorithm was proposed by Konis and Laurini (2007). They addressed the problem by considering only the change in observations between the m^{th} and $(m+1)^{\text{th}}$ subsets when the model fit is updated.

Originally, the method is intended for models that assume independent observations (Atkinson and Riani, 2001; 2002), generalized linear models (Atkinson and Riani, 2001), and multivariate analysis (Riani and Atkinson, 2001). Riani (2004) extended the use of the forward search algorithm to the analysis of structural time series data. As opposed to the regression models, the initial subset using time series data is robustly selected among k blocks of adjacent observations of fixed dimension. Block sampling is done to preserve the dependence of the series.

In addition, the use of the algorithm was extended to latent variable modeling to detect outliers and extreme response patterns. Mavridis and Moustaki (2008, 2009) implemented in factor analysis both for continuous and binary data. Furthermore, Toman (2014) studied the viability of the algorithm in confirmatory factory analysis, whereas Zijlstra, van der Ark, and Sijtsma (2011) implemented FS algorithm in detecting outliers in the Mokken scale analysis. Santos, de la Torre and Barrios (2015, 2016) also implemented the forward search algorithm in cognitive diagnosis model framework to identify examinees with aberrant response patterns.

4.4 Nonparametric regression

Nonparametric regression relates the response variable with the explanatory variables without specifying in advance the functional form of the model. The model is given by:

$$y_t = f(t) + \varepsilon_t. \tag{22}$$

This method estimates the function, $f(t)$, directly rather than estimating the parameters of the model. The good thing about this method is it makes fewer

assumptions; thus, it has flexibility and more variety of applications. Nevertheless, if the parametric assumptions are indeed true, then the estimates using parametric models are more precise than that of a nonparametric model (Wright, 2010).

Nonparametric regression aims to minimize the following objective function:

$$\sum_{t}\left(y-f(t)\right)^2+\lambda\int_{x_{\min}}^{x_{\max}}\left(f''(t)\right)^2\,dt, \tag{23}$$

where λ is the smoothing parameter penalizing the curvature of the function. However, McMurry and Politis (2008) noted that this method yields biased estimates. Thus, confidence intervals constructed using nonparametric bootstrap are not likely to attain its expected level of coverage.

There are several nonparametric regression procedures and among them is the scatterplot smoother. It is defined by Hastie and Tibshirani (1990) as a function of X and Y whose result is a function S with the same domain as the values in $x: s=S(X|Y)$and is smoother or less variable than Y itself. It serves as a visual aid in studying the relationship between X and Y and it can also be used as a building block in estimating an additive model. Kernel estimation, local polynomial regression, and smoothing splines are other common nonparametric regression techniques.

4.5 Isotonic regression and PAV algorithm

A smoothing method that belongs to a class of regression models that satisfy order or inequality restriction is called *monotonic regression*. It restricts the fitted model to be a monotone function. This procedure is advantageous when prior information on the ordering of the dependent variable with respect to the independent variable is available. If the response variable increases (or decreases) as the independent variable increases (or decreases), it is called isotonic regression.

Let $f(\cdot)$ be a function such that $f(x_1) \leq f(x_2) \leq \ldots \leq f(x_k)$ on a finite set X with simple order $x_1 \leq x_2 \leq \ldots \leq x_k$, then $f(\cdot)$ is called isotonic as defined by Barlow et al. (1972) and Robertson et al. (1988). Furthermore, they also defined isotonic function h^* on X as an isotonic regression h with weights w with respect to the simple ordering $x_1 \leq x_2 \leq \ldots \leq x_k$, if it minimizes the following objective function

$$\sum_{x\in X}[h(x)-f(x)]^2 w(x) \tag{24}$$

in the class of all isotonic functions $f(\cdot)$ on X. The resulting function is a step function. Hanson et al. (1973) illustrated the consistency property of the isotonic regression estimator for the cases of one or two predictors.

Meanwhile, Cheng (2009) proposed a generalized additive model with both parametric and nonparametric components being estimated simultaneously. The

nonparametric component of the model is assumed to be a monotone function and is modeled using isotonic regression. Cheng (2009) demonstrated that the asymptotic distribution of the isotonic function is not influenced by the simultaneous estimation of the parameters of the other component of the model. Simulation study showed that the isotonic estimator has an oracle property. Even if the other estimators in the model are unknown, the component of the monotone function in the model can be estimated with asymptotic accuracy.

Pooled Adjacent Violators (PAV) algorithm is the most widely used algorithm for estimating isotonic regression model with one independent variable (Robertson, 1988). The resulting function is a step function and, in a way, automatically assigns the independent variable into categories.

Burkadov et al. (2004) described the recursive procedure in the following manner. Define the set containing the independent variable X and the dependent variable Y be $M_n = \{(x_i, y_i), i = 1,2,...,n\}$, with the independent variable arranged in increasing order. The solution for M_n,$h(x)$, starts with $h(x_1) = y_1$ then modification is performed to obtain a solution for M_2 until all the n observations are included.

(1) Extend M_r with data point (x_{r+1}, y_{r+1}), such that $x_{r+1} \leq x_i$ for all $i > r + 1$ and let $h(x_{r+1}) = y_{r+1}$.

(2) Pool adjacent $h(x)$ values that violate the monotonicity constraint by assigning a common value

$$\frac{h(x_{r+1}) + ... + h(x_{r+1-k})}{k+1} \tag{25}$$

where k is the smallest nonnegative integer that will form a nondecreasing $h(x_1),..., h(x_{r+1})$ sequence. Repeat (1) until all the n observations are included. An algorithm to compute for the isotonic regression in case of ties among the observations in the independent variables was investigated by De Leeuw et al. (2009). They illustrated different approaches for situations where the predictors has a partial ordering and extended the existing PAV algorithm for models with more than one explanatory variable through active set algorithm and demonstrated how the procedure is implemented in R software using the package Isotone. Burkdakov et al. (2004) extended the algorithm for isotonic regression models with more than one explanatory variable.

5. Recent Developments in Spatial-Temporal Models using Computational Statistics Methods

There are various methods available in computational statistics as presented in the previous section. With the help of high-speed computers, computationally intensive estimation and hypothesis testing procedures can now be done much faster. This section enumerates several researches that utilize methods on computational statistics in performing statistical inference specifically for spatial-

temporal models. These studies are indeed an artifact that the aforementioned computationally intensive methods work well and can be an alternative to the classical or even Bayesian approach in inferential statistics.

Santos and Barrios (2011) investigated the dynamic panel data (DPD) model, which treats the lagged values of the dependent variable as explanatory variables. It is expressed as

$$Y_{it} = \alpha Y_{i,t-1} + \mathbf{X}_{it}' \boldsymbol{\beta} + \varepsilon_{it},\ i = 1,2,...,N\ t = 1,2,...,T,$$

where α is a constant, $\mathbf{X}_{it}$ is the vector of explanatory variables, $\boldsymbol{\beta}$ is the vector of regression coefficients, and ε_{it} follows a two-way error component model. Using simulations, they examined both the small and large sample properties of the within-groups (WG) estimator and the first different generalized method of moments (FD-GMM) estimator for a DPD model. They found that WG estimator performs best for long panels, whereas, the FD-GMM is superior for panels that are long and wide. However, they employed parametric bootstrap in the estimation procedure both for WG and FD-GMM when the asymptotic optimality of WG and FD-GMM fail, i.e. for the boundary scenario. They demonstrated that the bootstrapped WG and FD-GMM estimators are optimal for small samples.

Meanwhile, Landagan and Barrios (2007) proposed an iterative procedure for estimating spatial-temporal models that incorporates the Cochranne-Orcutt procedure into the backfitting algorithm. They postulated an additive model for yield of cereals, given in Equation 4.

Moreover, Campano and Barrios (2011) used the forward search algorithm with nonparametric bootstrap in the estimation of ARIMA models with structural change. The method identifies and models the structural change. Afterwards, the fitted model is adjusted to incorporate the presence of structural change. The forward search algorithm employed to overlapping blocks and bootstrap employed to independent blocks produced comparable estimates. Furthermore, Campano and Barrios (2011) investigated the use of the forward search algorithm and nonparametric bootstrap in detecting presence of outliers and modeling ARIMA models.

The forward search algorithm together with maximum likelihood estimation was incorporated by Bastero and Barrios (2010) into the backfitting algorithm in the estimation of spatial-temporal models in epidemics. This modified and iterative technique yield robust estimates in the presence of atypical observations during outbreaks. In addition, their proposed method provides good model fit even for small sample sizes for short time series data.

Guarte and Barrios (2013) proposed a nonparametric hypothesis testing for spatial-temporal models. The proposed test is based on nonparametric bootstrap. Using the bootstrap confidence interval, this test was developed for testing two assumptions in a spatial-temporal model, namely, constant temporal effect across locations or spatial units, and constant spatial effect over time. For reasonably

sized data, their simulation study showed that their proposed procedure can correctly validate the two assumptions. More recently, Asaad and Barrios (2015) extended these tests based on nonparametric bootstrap to multivariate spatial-temporal models. They found that these test procedures are powerful and correctly sized using simulation studies.

In small area estimation, del Prado and Barrios (2016) postulated a spatial-temporal mixed model intended to describe the structure inherent in a rotating panel survey with nested random effects. They demonstrated that the use of backfitting algorithm in conjunction with bootstrap can produce reliable small area statistics using a rotating panel survey data even at the boundary conditions. They noted that the estimation of model parameters is simplified with the use of backfitting algorithm, whereas the bias of the estimates is reduced by employing bootstrap.

Another study that employed backfitting algorithm in the estimation procedure recently is Villejo, Barrios, and Lansangan (2016). They proposed a dynamic spatial-temporal model with constant covariate effect but with varying spatial effect over time and varying temporal effect across locations. The postulated model can be estimated using the backfitting algorithm embedded with forward search algorithm and bootstrap. They demonstrated its superiority over the linear model in terms of predictive ability. Moreover, their proposed estimation procedure yielded lower relative bias and standard errors for the spatial parameter estimates.

Inevitably, there is always room for improvement in the existing methods for spatial-temporal models. There are still cases wherein these methods do not work well and these things merit separate studies. The future studies on spatial-temporal modeling, with the help of computational statistics, will certainly continue as there are real-life phenomena concerning time and space that are still not yet explored, most especially in the age of big data.

REFERENCES

ASSAD, A., BARRIOS, E., 2015, Nonparametric bootstrap test in a multivariate spatial-temporal model: A simulation study, *The Philippine Statistician,* 64(2): 1-16.

ATKINSON, A., RIANI, M., 2001, A unified approach to outliers, influence and transformations in discriminant analysis, *Journal of Computational and Graphical Statistics*, 10(3):513-544.

ATKINSON, A., RIANI M., 2002, Forward search added-variable t-tests and the effect of masked outliers on model selection, *Biometrika*, 89(4): 939-946.

ATKINSON, A., RIANI, M., 2007, Building regression models with the forward search, *Journal of Computing and Information Technology – CIT*, 15(4): 287-294.

BAILEY, N., HOLLY, S., PESARAN, M. H., 2016, A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence, *Journal of Applied Econometrics,* 31:249-280.

BALTAGI, B., SONG, H., JUNG, S., KOH, W., 2003, Testing for serial correlation, spatial autocorrelation and random effects using panel data, *Journal of Econometrics*, 140(1):5-51.

BALTAGI, B., 1995, *Econometric Analysis of Panel Data.* Chichester: Wiley.

BARLOW, R., BARTHOLOMEW, D., BREMMER, J., BRUNK, H., 1972, *Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression*. New York: John Wiley & Sons.

BASTERO, R. BARRIOS, E., 2011, Robust estimation of a spatiotemporal model with structural change, *Communications in Statistics – Simulation and Computation*, 40(3):448-468.

BINDER, H., TUTZ, G., 2008, A Comparison of Methods for the Fitting of Generalized Additive Models, *Statistics and Computing*, 18:87-99.

BRADLEY, J.R., HOLAN, S. H., WILKE, C. K., 2015, Multivariate spatio-temporal models for high-dimensional areal data with application to longitudinal, *The Annals of Applied Statistics*, 9(4):1761-1791.

BUHLMANN, P., 1997, Sieve bootstrap for time series, *Bernoulli*, 3(2):123-148.

BUHLMANN, P., 2002, Bootstraps for time series, *Statistical Science*, 17(1):52-72.

BURDAKOV, O., GRIMVALL, A., HUSSIAN, M., 2004, A Generalized PAV Algorithm for Monotonic Regression in Several Variables. In *COMPSTAT 2004 Symposium Held in Prague, Czech Republic.* Physica-Verlag/Springer, 762-767.

CAMPANO, W., BARRIOS, E., 2011, Robust estimation of a time series model with structural change, *Journal of Statistical Computing and Simulation***,** 81(7):909-927**.**

CHEN, R., TSAY, R., 1993, Functional-Coefficient Autoregressive Models, *Journal of the American Statistical Association*, 88(421):298-308.

CHENG, G., 2009, Semiparametric additive isotonic regression, *Journal of Statistical Planning and Inference*, 139:1980-1991.

CRESSIE, N., HUANG, H., 1999, Classes of Nonseparable, Spatial-temporal Stationary Covariance Functions, *Journal of the American Statistical Asscociation*, 94(448): 1330-1340.

DAVISON A., HINKLEY, D., 1997, *Bootstrap Methods and Their Application.* Cambridge U.K.: Cambridge University Press.

DE CESARE, L., MYERS, D., POSA, D., 2001, Estimating and modeling space-time correlation structures, *Statistics and Probability Letters*, 51(1): 9-14.

DE IACO, S., MYERS, D., POSA, D., 2001, Space-time analysis using a general product-sum model, *Statistics and Probability Letters*, 52(1): 21-28.

DE IACO, S., MYERS, D., POSA, D., 2003, The linear coregionalization model and the product-sum space-time variogram, *Mathematical Geology,* 35(1):25-38.

DE LEEUW, J., HORNIK, K., MAIR, P., 2009, Isotone optimization in R: Pool-adjacent-violators algorithm (PAVA) and active set methods, *Journal of Statistical Software,* 32(5):1-24. Available at http://www.jstatsoft.org/.

DEL PRADO, D. G., BARRIOS, E., 2016, Small area estimation with spatiotemporal mixed model, *The Philippine Statistician,* 65(2):1-36.

DUMANJUG, C, BARRIOS, E., LANSANGAN, J., 2010, Bootstrap Procedures in a Spatial-Temporal Model, *Journal of Statistical Computation and Simulation*, 80(7): 809-822.

EFRON, B., 1979, Bootstrap Methods: Another Look at the Jackknife, *Annals of Statistics*, 7:1-26.

FRIEDMAN, J.H. AND STUETZLE, W., 1982, *Smoothing of scatterplots*, Dept. of Statistics Technical Report, Orion 3, Stanford University.

FUENTES, M., 2006, Testing for separability of spatial-temporal covariance functions, *Journal of Statistical Planning and Inference*, 136:447-466.

GNEITING, T., 2002, Nonseparable, stationary covariance functions for space-time data, *Journal of the American Statistical Association*, 97(458): 590-600.

GNEITING, T., GENTON, M., GUTTORP, P., 2007, Geostatistical space-time models, stationarity, separability and full symmetry, In Findenstaedt, B., Held, L., Isham, V., eds., *Statistics of Spatial-temporal Systems*, Champman & Hall/CRC Press Monographs in Statistics and Applied Probability 151&175.

GUARTE, J., BARRIOS, E., 2013, Nonparametric hypothesis testing in a spatial-temporal model: A simulation study, *Communications in Statistics – Simulation and Computation*, 42(1):153-170.

HANSON, D., PLEDGER, G., WRIGHT, F., 1973, On consistency in monotonic regression, *The Annals of Statistics*, 1(3):401-421.

HARTFIELD, M., GUNST, R., 2003, Identification of model components for a class of continuous spatiotemporal models, *Journal of Agricultural Biological and Environmental Statistics*, 8(1):105-121.

HASTIE, T. TIBSHIRANI, R., 1990, *Generalized Additive Models*. London: Chapman & HALL.

HERING, A., BELL, C., GENTON, M., 2009, Modeling spatio-temporal wildfire ignition point patterns, *Environmental and Ecological Statistics*, 16(2):225-250.

IOANNIDIS, E., 2005, Residual-based block bootstrap unit root testing in the presence of trend breaks, *Econometrics Journal*, 8(3):323-351.

KONIS, K., LAURINI, F., 2007, Fitting a forward search in linear regression, *Bulletin of the 56th Meeting of the International Statistical Institute*, Portugal: International Statistical Institute.

LANDAGAN, O, BARRIOS, E., 2007, An estimation procedure for a spatial-temporal model, *Statistics and Probability Letters*, 77(4): 401-406.

LEE, L., YUN, J., 2010, Estimation of spatial autoregressive panel data models with fixed effects, *Journal of Econometrics*, 154:165-185.

MA, C., 2003a, Families of spatial-temporal stationary covariance models, *Journal of Statistical Planning and Inference*, 116:489-501.

MA, C., 2003b, Nonstationary covariance functions that model space-time interactions, *Statistics and Probability Letters*, 61:411-419.

MACKINNON, J.G., 2006, Bootstrap methods in econometrics, *Economic Record,* 82 (Issue Supplement s1):S2-S18.

MAVRIDIS, D., MOUSTAKI, I., 2008, Detecting outliers in Factor Analysis using the Forward Search Algorithm. *Multivariate Behavioral Research*, 43:453-475.

MAVRIDIS, D., MOUSTAKI, I., 2009, The Forward Search Algorithm for Detecting Aberrant Response Patterns in Factor Analysis for Binary Data. *Journal of Computational and Graphical Statistics,* 18(4):1016-1034.

MCMURRY, T., POLITIS, D., 2008, Bootstrap confidence intervals in nonparametric regression with built-in bias correction, *Statistics and Probability Letters*, 78(15):2463-2469.

MOONEY, C., DUVAL, R., 1993, *Bootstrapping: A nonparametric approach to statistical inference***,** Sage University Paper series on Quantitative Applications in the Social Sciences, Series no. 07-095, Newbury Park, CA: Sage.

PAPARODITIS, E. POLITIS, D., 2003, Residual-based block bootstrap for unit root testing, *Econometrica*, 71(3):813-855.

PARK, J., 2003, Bootstrap unit root tests, *Econometrica*, 71(6):1845-1895.

POLITIS, D., ROMANO, J., 1994, The stationary bootstrap, *Journal of the American Statistical Association*, 89(428):1303-1313.

PORCU, E., MATEU, J., ZINI, A., PINI, R. 2007, Modeling spatial-temporal data: a new variogram and covariance structure proposal, *Statistics and Probability Letters*, 77(1):83-89.

RIANI, M. 2004, Extensions of the forward search to time series, *Studies in Nonlinear Dynamics and Econometrics*, 8(2): Article 2.

RIANI, M., ATKINSON, A., 2001, A unified approach to outliers, influence and transformations in discriminant analysis, *Journal of Computational and Graphical Statistics*, 10**:**513-544.

ROBERTSON, T., WRIGHT, F., DYKSTRA, R., 1988, *Order Restricted Statistical Inference*. New York: John Wiley & Sons.

SANTOS, L.A., BARRIOS, E., 2011, Small sample estimation in dynamic panel data models: A simulation study, *American Open Journal of Statistics,* 1, 58-73.

SANTOS, K., DE LA TORRE, J., BARRIOS, E., 2015, *Detecting aberrant examinees in cognitive diagnosis models*. Paper presented at the Annual Meeting of the Psychometric Society, Beijing, China.

SANTOS, K., DE LA TORRE, J., BARRIOS, E.,2016, *Using forward search algorithm for person fit analysis in general cognitive diagnosis models*. Paper presented at the International Conference on Computational Statistics, Oviedo, Spain.

STEIN, M., 2005, Space-time covariance functions, *Journal of the American Statistical Association*, 100(469):310-321.

TOMAN, A., 2014, Robust confirmatory factor analysis based on the forward search algorithm, *Statistical Papers,* 55:233-252.

VILLEJO, S. J., BARRIOS, E., LANSANGAN, J.R., 2016, Robust estimation of a dynamic spatio-temporal model with structural change. *Journal of Statistical Computation and Simulation,* 87(3), 505-518.

WALLER, L., GOODWIN, B., WILSON, M., OSFELD, R., MARSHALL, S., HAYES, E., 2007, Spatio-temporal patterns in country-level incidence and reporting of Lyme disease in the Northeastern United States, 1990-2000, *Environmental Ecological Statistics*, 14(1):83-100.

WESTERLUND, J., EDGERTON, D., 2007, A panel bootstrap cointegration test, *Economic Letters*, 97(3):185-190.

WIKLE, C. K., HOOTEN, M. B., 2010, A general science-based framework for dynamical spatio-temporal models. *Test,* 19, 417-451.

WRIGHT, S., 2010, An investigation of two nonparametric regression models for value-added assessment in education, *SAS White Paper*. Available at http://www.sas.com.

ZIJLSTRA, W., VAN DER ARK, L. A., SIJTSMA, K., 2011, Robust Mokken scale analysis by means of the forward search algoritm for outlier detection, *Multivariate Behavioral Research,* 46:58-89.