SCHOOL OF STATISTICS
UNIVERSITY OF THE PHILIPPINES DILIMAN

WORKING PAPER SERIES

Nonparametric Hyphothesis Testing in a Spatial-Temporal Model: A Simulation Study

By

Jacqueline M.Guarte, Ph.D.
Erniel B. Barrios, Ph.D.

UPSS Working Paper No. 2010-06
June 2010

School of Statistics
Ramon Magsaysay Avenue
U.P. Diliman, Quezon City
Telefax: 928-08-81
Email: updstat@yahoo.com

ABSTRACT

Nonparametric procedures based on the bootstrap were developed for testing two assumptions in the course of estimating a spatial-temporal model: constant temporal effect across locations/spatial units and constant spatial effect over time. Simulation studies indicate that the proposed procedures can correctly identify the setting of parameters for reasonably sized data. Presence of spatial clustering can improve the sensitivity of the test under non-constant spatial effect over time. The test procedure for constant spatial effect over time is robust to model misspecification while that for constant temporal effect across spatial units is not robust to misspecification of the temporal model.

Key Words: spatial-temporal model, nonparametric bootstrap methods, bootstrap confidence interval, coverage probability, spatial clustering

1. INTRODUCTION

A spatial-temporal model was postulated and estimated by Landagan and Barrios (2007). They used the following additive model to properly characterize and understand the dynamics of agricultural production:

$$Y_{it} = X_{it}\beta + w_{it}\gamma + \varepsilon_{it}, \quad i = 1,\dots,N \text{ and } t = 1,\dots,T. \qquad (1)$$

The response variable for location i at time t, Y_{it}, is expressed as a linear function of X_{it}, the set of covariates for location i at time t; w_{it}, the set of variables in the neighborhood system of location i at time t; and ε_{it}, the error term. The estimation procedure is based on the backfitting algorithm that is embedded with the Cochranne-Orcutt procedure. In the estimation process, the error component is investigated assuming dependence structures but only the autoregressive behavior is explored. They proposed an estimation procedure with the following assumptions: constant covariate effect across locations and time; constant temporal effect across locations; and constant spatial effect over time. This study developed nonparametric procedures to verify the last two assumptions using the bootstrap.

The first procedure (Algorithm 1) is based on the model-based resampling in the time domain (Davison and Hinkley, 1997) or AR-sieve bootstrap (Bühlmann, 1997). It takes advantage of the different time series for each location representing the temporal variation in these locations. Bootstrapping these time series allows us to approximate the sampling distribution of the temporal parameter estimator to compare temporal effects across locations. This bootstrap procedure, however, only yields reliable bootstrap estimates of the temporal parameter if the AR(p) model fitted is chosen correctly (Davison and Hinkley, 1997). The second procedure (Algorithm 2) is based on the bootstrap method for regression analysis, resampling the cases (Chernick, 1999). It makes use of the different cross-sectional data for each time point to represent the spatial variation in these time points. Resampling the cases in regression analysis allows us to approximate the sampling distribution of the spatial parameter estimator to compare spatial effects over time. This bootstrap method is known to be less sensitive to model misspecification (Efron as cited by Chernick, 1999).

This study sets out to verify the performance of these two algorithms including robustness to model misspecification through simulation studies.

2. TESTING FOR CONSTANT TEMPORAL EFFECT ACROSS LOCATIONS

Given the time series in each location/spatial unit, we test the following hypotheses:

H_0 : All locations/spatial units have the same temporal effect.

H_1 : At least one location/spatial unit differs in temporal effect.

Algorithm 1:

Given the time series $y_{i1},..., y_{it},..., y_{iT}$ in each location i, assume AR(p) fits the time series, i.e.,

$$y_{it} = \phi_0 + \phi_1 y_{i,t-1} + ... + \phi_p y_{i,t-p} + \varepsilon_{it}, \quad t = 1,...,T \tag{2}$$

with $\varepsilon_{it} \sim NID(0, \sigma_\varepsilon^2)$.

1) Estimate the model (2) and predict y_{it} by

$$\hat{y}_{it} = \hat{\phi}_0 + \hat{\phi}_1 y_{i,t-1} + ... + \hat{\phi}_p y_{i,t-p}, \quad t = 1,...,m \tag{3}$$

where $\hat{\phi}_1,...,\hat{\phi}_p$ are maximum-likelihood estimates of the corresponding parameters. Let ϕ_p be the temporal parameter of interest. Conditioning on $(y_{i,t-1},..., y_{i,t-p})$, the empirical distribution of the centered residuals is

$$e_{it} \sim NID(0, MSE). \tag{4}$$

2) Generate *k* bootstrap samples from the distribution (4) say, $(e_{i0}^*,..., e_{im}^*)$ for each location i from a sample of size *m*.

3) Generate *k* time series for each location i, one for each bootstrap sample in step 2) using the estimated model (3). Each time series $\left(y_{i1}^{*},...,y_{im}^{*}\right)$ will be simulated as follows:

i. Initialize $y_{i0}^{*}, y_{i1}^{*},..., y_{i,t-1}^{*}.$

ii. Then $y_{it}^{*} = \hat{\phi}_0 + \hat{\phi}_1 y_{i,t-1}^{*} + ... + \hat{\phi}_p y_{i,t-p}^{*} + e_{it}^{*},\ t = 1,...,m.$

4) Estimate the AR(p) model used in step 1) for each of the simulated time series in step 3).

5) Compute the estimated standard error for each of the estimated temporal parameter $\hat{\phi}_p$ in step 1) using the corresponding *k* bootstrap temporal parameter estimates $\hat{\phi}_{pj}^{*}$ generated in step 4) if the bootstrapped sampling distribution of the estimated parameter is normal or approximately normal:

$$\hat{\sigma}_{\hat{\phi}_p}^{*} = \left[\frac{1}{k-1}\sum_{j=1}^{k}\left(\hat{\phi}_{pj}^{*} - \overline{\hat{\phi}}_p^{*}\right)^2\right]^{1/2}, \overline{\hat{\phi}}_p^{*} = k^{-1}\sum_{j=1}^{k}\hat{\phi}_{pj}^{*}. \tag{5}$$

6) Construct the 95%(99%) normal-approximation confidence intervals on each temporal parameter ϕ_p estimated in step 1) computing the limits as follows:

$$\hat{\phi}_p \pm z_{\frac{\alpha}{2}} \hat{\sigma}_{\hat{\phi}_p}^{*} \tag{6}$$

where $\hat{\phi}_p$ is the original point estimate from step 1) and $\hat{\sigma}_{\hat{\phi}_p}^{*}$ is its standard error estimate from step 5). Find the appropriate percentiles for the limits if the distribution of $\hat{\phi}_p^{*}$ cannot be assumed to be normal.

7) Compute the average of the temporal parameter estimates $\hat{\phi}_p$ in step 1).

8) Reject the null hypothesis that there is constant temporal effect across locations with 95% (99%) coverage probability if more than 5% (1%) of the constructed intervals fail to contain the average computed in step 7).

3. TESTING FOR CONSTANT SPATIAL EFFECT OVER TIME

Given the cross-sectional data in each time point, we test the following hypotheses:

H_0 : All time points have the same spatial effect.

H_1 : At least one time point differs in spatial effect.

Algorithm 2:

Given the cross-sectional data $y_{1t},\ldots,y_{Nt}$ in each time point t, without loss of generality, suppose the regression model

$$y_{it} = \underset{\sim}{x}_{it}\underset{\sim}{\beta}_t + \varepsilon_{it}, i = 1,\ldots,N \tag{7}$$

with $\varepsilon_{it} \sim NID\left(0,\sigma_{\varepsilon}^2\right)$, is adequate.

1) Estimate the model (7) and predict y_{it} by

$$\hat{y}_{it} = \underset{\sim}{x}_{it}\underset{\sim}{\hat{\beta}}_t, i = 1,\ldots,N \tag{8}$$

where $\underset{\sim}{\hat{\beta}}_t$ is a vector of maximum-likelihood estimates of the regression parameters. Now, without loss of generality, we identify the spatial parameter of interest, β_t, as associated with the lone predictor variable x_{it}.

2) Generate *k* bootstrap samples of N pairs of $\left(x_{it}, y_{it}\right)$ from the N pairs of observations on the predictor and response variables in each time point.

3) Estimate the regression model (7) for all the bootstrap samples in each time point.

4) If the bootstrapped sampling distribution of the estimated parameter is normal or approximately normal,

   i. Compute the estimated standard error (5) for the estimated spatial parameter of interest $\hat{\beta}_t$ in the model (8) using the corresponding *k* bootstrap estimates $\hat{\beta}^*_{tj}$ generated in step 3) in each time point.

   ii. Construct the 95% (99%) normal-approximation confidence intervals (6) on the spatial parameter of interest in 1) in each time point.

5) If the bootstrapped sampling distribution of the estimated parameter is not normal,

   i. Sort the *k* bootstrap estimates in either ascending or descending order.

   ii. Find the appropriate percentiles to construct the corresponding intervals.

6) Compute the average of the estimated spatial parameter of interest in step 1).

7) Reject the null hypothesis that there is constant spatial effect over time with 95% (99%) coverage probability if more than 5% (1%) of the constructed intervals fail to contain the average computed in step 6).

4. SIMULATION STUDIES

The performance of the two proposed test procedures were evaluated by simulating two (2) parameter conditions (constant and not constant parameter effect) and then verifying if actual behavior of the parameter is correctly identified under these conditions. The robustness of the proposed procedures to misspecification of the model was also verified.

4.1 PERFORMANCE UNDER CONSTANT TEMPORAL EFFECT ACROSS SPATIAL UNITS

Two (2) spatial-temporal data sets were generated on a response variable Y_{it} with N =100 stationary time series (corresponding to 100 spatial units) each with T =100 time points. Each time series in the first data set is generated from an AR(4) process with temporal parameter ϕ_4 =0.5 (securely stationary). Each time series in the second data is generated from an AR(4) process with ϕ_4 = 0.9 (nearly nonstationary). We note here that the data generating process used has no intercept term for the AR(p) model. Algorithm 1 was then applied to the two data sets using 95% coverage probability.

Data Set 1:

The AR(4) process was estimated across all spatial units using the model

$$y_{it} = \phi_4 y_{i,t-4} + \varepsilon_{it}, t = 1,...,100 \quad (9)$$

with ϕ_4 as the temporal parameter of interest whose true value is 0.5.The distribution of the original estimates is normal with mean 0.499 and median 0.508 (Figure 1). .Across the 100 spatial units, 5 bootstrap estimators are not normal, 26 are approximately normal, and 69 are normal (at the 5% level of significance) based on 200 resamples per spatial unit. The 5 non-normal distributions are those of original temporal parameter estimates belonging to the top 16% of the array.

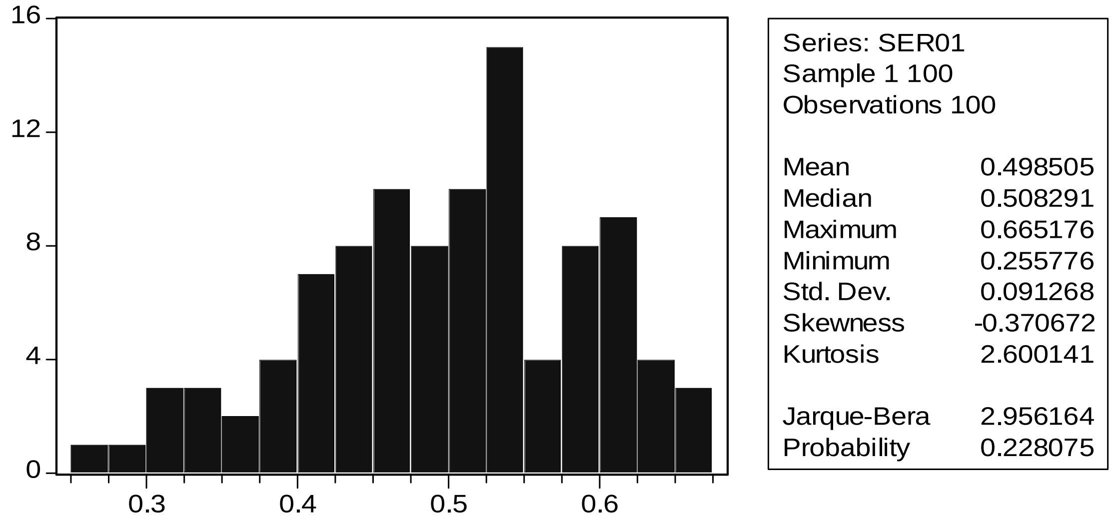

Figure 1. Histogram of the temporal parameter estimates across spatial units (N =100) with true parameter value 0.5.

The null hypothesis of constant temporal effect across spatial units is not rejected with 95% coverage probability (Table 1). Of the 100 confidence intervals constructed, three (3) failed to contain the mean while two (2) failed to contain the median (Table 2). The three spatial units that significantly differed from the rest are those with the highest (0.665176), lowest (0.255776), and second lowest (0.280414) original parameter estimates, respectively. Thus, even if the population of spatial units is homogeneous with respect to temporal effect, sampling variation (i.e., sampling from an AR process that has been occurring for an infinitely long time in each spatial unit) will induce some spatial units/time series to be different in the sample with respect to temporal effect.

Table 1. Result of comparison of temporal effects across spatial units (N =100) under constant temporal effect with true parameter value 0.5 using 95% coverage probability.

Criterion [Value]	Decision on H_0
Mean [0.498505]	Do not reject H_0 (3)
Median [0.508291]	Do not reject H_0 (2)

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

Table 2. Bootstrap confidence intervals for the identified three (3) different spatial units/time series in Table 1 (based on the mean with 95% coverage probability).

Spatial Unit No.	95% C.I.
42	(0.501708, 0.828644)
49	(0.058694, 0.452858)
65	(0.077068, 0.483760)

The test procedure was able to correctly identify the true situation and is also properly sized for this data set. That is, not rejecting the null hypothesis of constant temporal effect with 95% coverage probability actually captured the fact that not more than 5 spatial units differ in temporal effect, in this case, none.

Data Set 2:

The same model (9) was estimated for each of the 100 spatial units/time series with ϕ_4's true value set at 0.9 this time. The distribution of the original estimates is no longer normal with mean 0.889 and median 0.892 (Figure 2). Across the 100 spatial units, 77 bootstrap estimators are not normal while 23 are approximately normal based on 200 resamples per spatial unit at the 5% level of significance. The non-normal original temporal parameter estimates are spread out across the array.

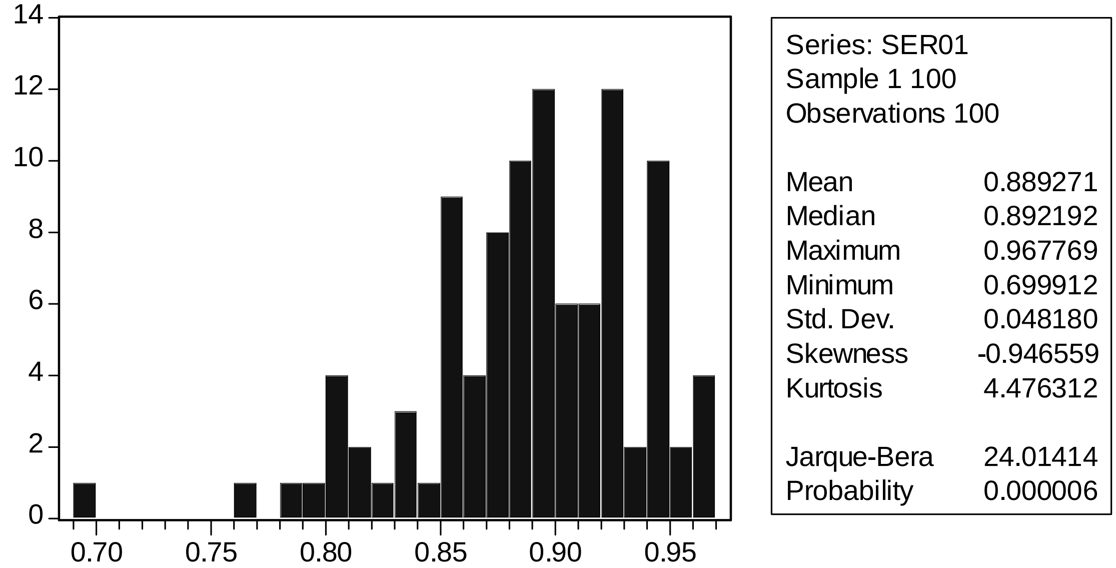

Figure 2. Histogram of the temporal parameter estimates across spatial units (N =100) with true parameter value 0.9.

The null hypothesis of constant temporal effect across spatial units is also not rejected with 95% coverage probability (Table 3). Of the 100 confidence intervals constructed, three (3) failed to contain the mean and the median, respectively (Table 4) and these are the same for both. The three spatial units that significantly differed from the rest are those with the second lowest (0.767117), highest (0.967769), and lowest (0.699912) original parameter estimates, respectively. Again, sampling variation leads to a few spatial units varying with respect to temporal effect in the sample. The test procedure was able to correctly identify the true setting and is also correctly sized for this data set.

Table 3. Result of comparison of temporal effects across spatial units (N =100) under constant temporal effect with true parameter value 0.9 using 95% coverage probability.

Criterion [Value]	Decision on H_0
Mean [0.889271]	Do not reject H_0 (3)
Median [0.892192]	Do not reject H_0 (3)

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

Table 4. Bootstrap confidence intervals for the identified three (3) different spatial units in Table 3 (based on the mean with 95% coverage probability).

Spatial Unit No.	95% C.I.
49	(0.501708, 0.828644)
51	(0.058694, 0.452858)
82	(0.077068, 0.483760)

Note: Spatial unit numbers are independent of those in Table 2.

4.2 PERFORMANCE UNDER NOT CONSTANT TEMPORAL EFFECT ACROSS SPATIAL UNITS

Seven (7) data sets on Y_{it} were first generated using Data Set 1 with 2, 3, 4, 5, 6, 7, and 8 temporal parameters different from the rest, respectively. Each case is a mixture of high and low parameter values relative to 0.5. Algorithm 1 was applied to these spatial-temporal data sets using 95% coverage probability. The results are presented in Table 5. Except for the first case (with only two different spatial units with different temporal parameters), the result of the test always correctly detect the non-constant temporal effect. The test is also correctly sized in each case.

Table 5. Results of comparison of temporal effects across spatial units (N =100) in seven cases with underlying non-constant temporal effect (2% to 8% different temporal effects) using 95% coverage probability.

Case No. and Description	Criterion[Value]	Decision on H_0	Power of the Test
1. 98 time series with $\phi_4 = 0.5$; 2 time series with $\phi_4 = 0.9$, 0.2	Mean[0.497180]	Do not reject H_0 (5)	1
	Median[0.505090]	Do not reject H_0 (5)	1
2. 97 time series with $\phi_4 = 0.5$; 3 time series with $\phi_4 = 0.9$, 0.8, 0.2,	Mean[0.501144]	Reject H_0 (6)	1
	Median[0.508291]	Reject H_0 (7)	1
3. 96 time series with $\phi_4 = 0.5$; 4 time series with $\phi_4 = 0.9$, 0.8, 0.3, and 0.2	Mean[0.500738]	Reject H_0 (6)	0.75
	Median[0.508291]	Reject H_0 (7)	0.75
4. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.9$, 0.8, 0.7, 0.3, and 0.2	Mean[0.504372]	Reject H_0 (7)	0.8
	Median[0.511097]	Reject H_0 (8)	0.8
5. 94 time series with $\phi_4 = 0.5$; 6 time series with $\phi_4 = 0.9$, 0.8, 0.7, 0.4, 0.3, and 0.2	Mean[0.504441]	Reject H_0 (7)	0.67
	Median[0.511097]	Reject H_0 (8)	0.67
6. 93 time series with $\phi_4 = 0.5$; 7 time series with $\phi_4 = 0.9$, 0.8, 0.7, 0.6, 0.4, 0.3, and 0.2	Mean[0.504070]	Reject H_0 (7)	0.57
	Median[0.511097]	Reject H_0 (8)	0.57
7. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.9$, 0.8, 0.7, 0.6, 0.4, 0.3, 0.2, and 0.1	Mean[0.499924]	Reject H_0 (8)	0.625
	Median[0.508291]	Reject H_0 (9)	0.625

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

The power of the test or the probability of correctly rejecting the null hypothesis (i.e., proportion of bootstrap confidence intervals correctly failing to contain the mean/median) decreased as alternative values of the temporal parameter were set closer to 0.5 (0.3, 0.4, and 0.6). The test failed to detect that these values are different from 0.5 (Table 6). As more parameter values were made different from 0.5, the median became more sensitive than the mean in detecting these alternative values. The power of the test tended to be high with the presence of more distant parameter values from 0.5.

Table 6. Bootstrap confidence intervals for the known eight different spatial units in Table 5 and for the most different spatial units with $\phi_4 = 0.5$.

True ϕ_4	Original $\hat{\phi}_4$ (Bootstrap S.E.)	95% C.I. (Case No. 7)
0.9	0.853643 (0.054616)	(0.746596, 0.960690)
0.8	0.805001 (0.068233)	(0.637057, 0.887179)
0.7	0.735941 (0.069082)	(0.600540, 0.871342)
0.6	0.577454 (0.087337)	(0.406273, 0.748635)
0.5	0.665176 (0.083402)	(0.501708, 0.828644)
0.5	0.255776 (0.100552)	(0.058694, 0.452858)
0.5	0.280414 (0.103748)	(0.077068, 0.483760)
0.5	0.319438 (0.096160)	(0.130964, 0.507912)
0.5	0.309932 (0.098060)	(0.117734, 0.502130)
0.4	0.404189 (0.092683)	(0.222530, 0.585848)
0.3	0.326173 (0.095997)	(0.138019, 0.514327)
0.2	0.161269 (0.104869)	(-0.044274, 0.366812)
0.1	0.119484 (0.107492)	(-0.091200, 0.330168)

Sixteen (16) data sets on Y_{it} were also generated using Data Set 1 with the first 5 spatial units having the same temporal parameter that differs from 0.5 (first 8 cases) and the first 8 spatial units having the same temporal parameter that differs from 0.5 (second 8 cases). The results are summarized in Tables 7 and 8.

Table 7. Results of comparison of temporal effects across spatial units (N=100, T=100) in eight cases with underlying non-constant temporal effect (5% different with same temporal effect) using 95% coverage probability.

Case No. and Description	Criterion[Value]	Decision on H_0	Power of the Test
1. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.9$	Mean[0.519451]	Reject H_0 (10)	1
	Median[0.515192]	Reject H_0 (10)	1
2. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.8$	Mean[0.513278]	Reject H_0 (9)	1
	Median[0.515192]	Reject H_0 (9)	0.8
3. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.7$	Mean[0.508724]	Reject H_0 (6)	0.4
	Median[0.515192]	Reject H_0 (7)	0.4
4. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.6$	Mean[0.503733]	Do not reject H_0 (4)	0.2
	Median[0.513234]	Do not reject H_0 (4)	0
5. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.4$	Mean[0.494514]	Do not reject H_0 (4)	0.2
	Median[0.501822]	Do not reject H_0 (3)	0.2
6. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.3$	Mean[0.485754]	Reject H_0 (8)	0.8
	Median[0.501822]	Reject H_0 (6)	0.8
7. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.2$	Mean[0.485194]	Reject H_0 (9)	1
	Median[0.501822]	Reject H_0 (7)	1
8. 95 time series with $\phi_4 = 0.5$; 5 time series with $\phi_4 = 0.1$	Mean[0.483893]	Reject H_0 (10)	1
	Median[0.501822]	Reject H_0 (7)	1

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

Table 8. Results of comparison of temporal effects across spatial units (N =100, T =100) in eight cases with underlying non-constant temporal effect (8% different with same temporal effect) using 95% coverage probability.

Case No. and Description	Criterion[Value]	Decision on H_0	Power of the Test
1. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.9$	Mean[0.529412]	Reject H_0 (13)	1
	Median[0.518704]	Reject H_0 (13)	1
2. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.8$	Mean[0.521904]	Reject H_0 (12)	0.875
	Median[0.518704]	Reject H_0 (12)	0.875
3. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.7$	Mean[0.515963]	Reject H_0 (10)	0.625
	Median[0.518704]	Reject H_0 (9)	0.5
4. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.6$	Mean[0.507578]	Do not reject H_0 (5)	0.25
	Median[0.515192]	Reject H_0 (6)	0.125
5. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.4$	Mean[0.491179]	Do not reject H_0 (4)	0.125
	Median[0.495656]	Do not reject H_0 (4)	0.125
6. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.3$	Mean[0.479288]	Reject H_0 (9)	0.625
	Median[0.495656]	Reject H_0 (8)	0.625
7. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.2$	Mean[0.478927]	Reject H_0 (10)	0.75
	Median[0.495656]	Reject H_0 (9)	0.75
8. 92 time series with $\phi_4 = 0.5$; 8 time series with $\phi_4 = 0.1$	Mean[0.473390]	Reject H_0 (11)	0.75
	Median[0.495656]	Reject H_0 (10)	0.875

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

With 5% of the spatial units in the population having an alternative temporal parameter value, the test procedure failed to detect the true state of not constant temporal effect across spatial units when the alternative parameters were close enough to 0.5 (cases 4 and 5). However, when this increased to 8% of the spatial units, the median was more sensitive compared to the mean in detecting the true state (Table 8); this time, the test procedure was able to differentiate 0.6 from 0.5. Only the alternative parameter value 0.4 remained indistinguishable from 0.5 using the test. The power of the test is highest for the farthest alternative parameter values from 0.5 and becomes closer to 0 as the alternative value gets closer to 0.5.

4.3 PERFORMANCE UNDER CONSTANT SPATIAL EFFECT OVER TIME

For this evaluation, T =100 data sets of N =100 pairs of observations on new values of Y_{it} and a predictor variable X_{it}, based on the response variable Y_{it} in Data Set 1 were generated. The values of Y_{it} were first translated in location and scale since these were relatively small for the purpose of using an exponential model. Values of the predictor variable X_{it} were generated based on the exponential model and using the translated values of Y_{it}. The corresponding new values of Y_{it}, denoted by y''_{it}, were then computed from the exponential model as follows:

$$y''_{it} = \beta_1 x_{it}^{\beta_2} + e_{it}, \text{ where } e_{it} \sim N\left(0, \sigma_e^2\right) \qquad (10)$$

In this study, $\beta_1 = 10, \beta_2 = 0.5$ (the spatial parameter of interest), and $\sigma_e = 100$. Spatial clustering was implemented by dividing the data on the translated values of Y_{it} across space ($i = 1, ..., N = 100$) into four (4) clusters.

Algorithm 2 was applied to the spatial-temporal data sets using 95% coverage probability. The exponential regression model (10) with β_1=10 was estimated for all time points. While the distribution of the spatial parameter estimates with no spatial clustering is normal, that with spatial clustering is positively skewed but with almost the same variance. Across time, however, the behavior of the bootstrap statistics is nearly the same for the two conditions. Without spatial clustering, 13% are normal, 29% are approximately normal, and 58% are not normal (at the 5% level of significance) based on 200 bootstrap samples per time point. With spatial clustering, 12% are normal, 30% are approximately normal, and 58% are not normal. These three groups, however, are not necessarily the same time points for the two conditions.

Spatial clustering has a slight effect on the magnitude of the spatial parameter estimate and an important effect on the variability of the bootstrap estimates (Table 9). While the spatial parameter estimates are only slightly different in the two conditions, the bootstrap estimates are substantially less variable under spatial clustering, yielding narrower bootstrap confidence intervals. The difference in the bootstrap standard error estimates (relative to that under no spatial clustering) stand at 33.0%, 25.6%, 15.2%, 20.0%, and 27.1% for the first 5 time points, respectively. Note that at t=27, where a too narrow 95% confidence interval resulted, the bootstrap standard error estimate under spatial clustering is 98.5% lower. This implies that the test procedure for constant spatial effect over time will be more sensitive in detecting changes in spatial effect across time (i.e., “structural change” in time) when there is spatial clustering.

Table 9. Effect of spatial clustering on the parameter estimates and bootstrap confidence intervals when $\beta_2 = 0.5$ for the first five time points and the 27th time point (a).

Time point t	Original $\hat{\beta}_2$ (Bootstrap S.E.) and 95% C.I.	
	With no spatial clustering	With spatial clustering
1	0.489790 (**0.083156**) (0.277786, 0.604716)	0.484435 (**0.055681**) (0.375300, 0.593570)
2	0.518016 (**0.075583**) (0.330897, 0.634137)	0.506864 (**0.056218**) (0.378610, 0.594134)
3	0.490315 (**0.068855**) (0.355359, 0.625271)	0.486452 (**0.058376**) (0.330765, 0.572304)
4	0.444948 (**0.085288**) (0.268765, 0.568174)	0.456975 (**0.068228**) (0.275187, 0.558107)
5	0.537602 (**0.061730**) (0.393354, 0.628543)	0.519371 (**0.044990**) (0.431191, 0.607551)
27	0.513319 (**0.062520**) (0.390780, 0.635858)	0.636074 (**0.000963**) (0.634187, 0.637961)

(a) With the lowest bootstrap standard error estimate under spatial clustering.

The null hypothesis of constant spatial effect over time is correctly not rejected with 95% coverage probability under both conditions (Table 10). However, one interval under spatial clustering incorrectly failed to contain the average. This is that of the 27th time point which has the highest spatial parameter estimate and also the lowest bootstrap standard error estimate.

Table 10. Result of comparison of spatial effects over time (T=100) with underlying constant spatial effect and true parameter value $\beta_2 = 0.5$ using 95% coverage probability.

Criterion [Value]	Decision on H_0
With no spatial clustering	
Mean [0.508173]	Do not reject H_0 (0)
Median [0.506551]	Do not reject H_0 (0)
With spatial clustering	
Mean [0.499758]	Do not reject H_0 (1)
Median [0.497902]	Do not reject H_0 (1)

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

4.4 PERFORMANCE UNDER NOT CONSTANT SPATIAL EFFECT OVER TIME

Fourteen (14) data sets based on sub-section 4.3 with 5% and 8% spatial parameters different from the rest, respectively, were constructed. The data for the first five and first eight time points, respectively, were changed to have the desired spatial parameter values other than 0.5. Algorithm 2 was applied to the spatial-temporal data sets using 95% coverage probability. The effect of spatial clustering on the magnitude of the parameter estimates and the variability of the bootstrap estimates remain the same as in the condition of constant spatial effect. In general, the bootstrap standard error estimate under spatial clustering is expected to be lower than that under no spatial clustering for all spatial parameter values using the said exponential model (10). It was observed that the bootstrap estimates decreased in variability as the spatial parameter value in the model decreased from 0.9 to 0.2 under the two conditions. Consequently, due to the relatively lower variability of the bootstrap estimates for 0.4, 0.3, and 0.2, the bootstrap confidence intervals were narrower compared to those of the larger parameter values.

The null hypothesis of constant spatial effect across time is incorrectly not rejected for all cases when there is no spatial clustering but correctly rejected for all cases (except for β_2=0.6, 0.7) when there is spatial clustering under the case of 5% time points with different spatial effects (Table 11). The power of the test is higher under spatial clustering when the alternative spatial parameters are 0.9, 0.8, and 0.7. The test has the same power under the two conditions for alternative parameter values 0.6, 0.4, 0.3, and 0.2.

With 8% of the time points having different spatial parameters from the rest (Table 12), the null hypothesis of constant spatial effect across time is now correctly rejected even without spatial clustering when the alternative parameter values are 0.9, 0.8, 0.4, 0.3, and 0.2. When there is spatial clustering, correct inference is observed for all cases except when the alternative value is 0.6. Thus, 0.6 remains to be identified as similar to 0.5 despite spatial clustering. The power of the test when there is spatial clustering remains higher than that with no spatial clustering for the alternative parameter values of 0.9, 0.8, 0.7, and 0.4. It is equal for the two conditions for the remaining alternative parameter values of 0.6, 0.3, and 0.2.

Table 11. Results of comparison of spatial effects over time (T=100) with and without spatial clustering in seven cases with underlying non-constant spatial effect (5% different with same spatial effect) using 95% coverage probability.

Case No. and Description	Criterion[Value]		Decision on H_0		Power of the Test	
	With no clustering	With clustering	With no clustering	With clustering	With no clustering	With clustering
1. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.9$	Mean [0.528018]	Mean [0.519391]	Do not reject H_0 (3)	Reject H_0 (6)	0.6	1
	Median [0.508085]	Median [0.498232]	Do not reject H_0 (4)	Reject H_0 (6)	0.8	1
2. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.8$	Mean [0.523057]	Mean [0.514483]	Do not reject H_0 (3)	Reject H_0 (6)	0.6	1
	Median [0.508085]	Median [0.498232]	Do not reject H_0 (3)	Reject H_0 (6)	0.6	1
3. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.7$	Mean [.518095]	Mean [0.509575]	Do not reject H_0 (1)	Do not reject H_0 (4)	0.2	0.6
	Median [.508085]	Median [0.498232]	Do not reject H_0 (1)	Do not reject H_0 (4)	0.2	0.6
4. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.6$	Mean [0.513134]	Mean [0.504666]	Do not reject H_0 (0)	Do not reject H_0 (1)	0	0
	Median [0.508085]	Median [0.498232]	Do not reject H_0 (0)	Do not reject H_0 (1)	0	0

5. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.4$	Mean [0.503211]	Mean [0.494850]	Do not reject H_0 (4)	Reject H_0 (6)	0.8	1
	Median [0.506106]	Median [0.497390]	Do not reject H_0 (5)	Reject H_0 (6)	1	1
6. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.3$	Mean [0.498250]	Mean [0.489942]	Do not reject H_0 (5)	Reject H_0 (6)	1	1
	Median [0.506106]	Median [0.497390]	Do not reject H_0 (5)	Reject H_0 (6)	1	1
7. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.2$	Mean [0.493289]	Mean [0.485034]	Do not reject H_0 (5)	Reject H_0 (6)	1	1
	Median [0.506106]	Median [0.497390]	Do not reject H_0 (5)	Reject H_0 (6)	1	1

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

Table 12. Results of comparison of spatial effects over time (T =100) with and without spatial clustering in seven cases with underlying non-constant spatial effect (8% different with same spatial effect) using 95% coverage probability.

Case No. and Description	Criterion[Value]		Decision on H_0		Power of the Test	
	With no clustering	With clustering	With no clustering	With clustering	With no clustering	With clustering
1. 92 time points with $\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.9$	Mean [0.540072]	Mean [0.531283]	Reject H_0 (6)	Reject H_0 (9)	0.75	1
	Median [0.510666]	Median [0.500058]	Reject H_0 (7)	Reject H_0 (9)	0.875	1
2. 92 time points with $\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.8$	Mean [0.532097]	Mean [0.523402]	Do not reject H_0 (5)	Reject H_0 (8)	0.625	0.875
	Median [0.510666]	Median [0.500058]	Reject H_0 (6)	Reject H_0 (9)	0.75	1
3. 92 time points with $\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.7$	Mean [0.524122]	Mean [0.515520]	Do not reject H_0 (1)	Reject H_0 (6)	0.125	0.625
	Median [0510666]	Median [0.500058]	Do not reject H_0 (1)	Reject H_0 (6)	0.125	0.625
4. 92 time points with	Mean [0.516147]	Mean [0.507639]	Do not reject H_0	Do not reject H_0	0	0

$\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.6$			(0)	(1)		
	Median [0.510666]	Median [0.500058]	Do not reject H_0 (0)	Do not reject H_0 (1)	0	0
5. 92 time points with $\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.4$	Mean [0.500198]	Mean [0.491877]	Do not reject H_0 (5)	Reject H_0 (8)	0.625	0.875
	Median [0.505988]	Median [0.496701]	Reject H_0 (7)	Reject H_0 (9)	0.875	1
6. 92 time points with $\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.3$	Mean [0.492223]	Mean [0.483996]	Reject H_0 (8)	Reject H_0 (9)	1	1
	Median [0.505988]	Median [0.496701]	Reject H_0 (8)	Reject H_0 (9)	1	1
7. 92 time points with $\beta_2 = 0.5$; 8 time points with $\beta_2 = 0.2$	Mean [0.484248]	Mean [0.476115]	Reject H_0 (8)	Reject H_0 (9)	1	1
	Median [0.505988]	Median [0.496701]	Reject H_0 (8)	Reject H_0 (9)	1	1

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

4.5 ROBUSTNESS OF THE TESTING PROCEDURES TO MISSPECIFICATION OF THE MODEL

In sub-section 4.1, Algorithm 1 was applied using the correct model AR(4) with model (9) estimated across all spatial units. In this section, Algorithm 1 was applied again to the same data but this time an AR(3) model was estimated instead using the same form of the model (9). It is very clear that the parameter estimates using AR(3) are not only very much different and far from the true parameter value but also very different among themselves compared to those using AR(4) (Table 13).

Table 13. Effect of model misspecification on the parameter estimates and bootstrap confidence intervals when $\phi_4 = 0.5$ for the first five spatial units.

Spatial Unit, i	Original $\hat{\phi}_4$ (Bootstrap S.E.) and 95% C.I.	
	Correct Model (AR(4))	Incorrect Model (AR(3))
1	0.602101 (**0.080156**) (0.444995, 0.759207)	-0.088264 (**0.098082**) (-0.280505, 0.103977)
2	0.545371 (**0.094011**) (0.361109, 0.729633)	0.180325 (**0.107258**) (-0.029901, 0.390551)
3	0.408617 (**0.104997**) (0.202823, 0.614411)	-0.285795 (**0.109820**) (-0.501042, -0.070548)
4	0.366756 (**0.096781**) (0.177065, 0.556447)	0.031241 (**0.103855**) (-0.172315, 0.234797)
5	0.372530 (**0.093881**) (0.188523, 0.556537)	0.151687 (**0.098329**) (-0.041038, 0.344412)

The null hypothesis of constant temporal effect across spatial units was incorrectly rejected when an AR(3) was used to estimate the model. Many (11%) spatial units were incorrectly identified as different from the rest as opposed to only a few (at most 3%) when an AR(4) was used. This result confirms observation on the importance of using the correct model for inference (Davison and Hinkley, 1997). Thus, the testing procedure is not robust to temporal model misspecification as expected.

In sub-section 4.3, Algorithm 2 was applied to the generated data ($N = T$ =100) in testing for constant spatial effect over time with no spatial clustering using the exponential model $y_{it} = \beta_1 x_{it}^{\beta_2} + e_{it}$. This time, Algorithm 2 is applied to the same data by specifying a linear model instead

$$y_{it} = \log \beta_1 + \beta_2 x_{it} + e_{it} \tag{11}$$

using the same value for the first regression coefficient (i.e., $\beta_1 = 10$).

While the parameter estimates using the linear model instead of the exponential model are very much different from the true parameter value (Table 14), they are clearly alike among themselves as opposed to that observed when using AR(3) instead of AR(4). Also, the bootstrap estimates using the linear model are highly variable compared to those using the correct model. This is indicative of the relatively wider 95% confidence intervals under this model compared to those under the correct model. Thus, not surprisingly, the null hypothesis of constant spatial effect over time is correctly not rejected even when the model used is incorrect (Table 15).

Table 14. Effect of model misspecification on the parameter estimates and bootstrap confidence intervals when $\beta_2 = 0.5$ for the first five time points.

Time Point, t	Original $\hat{\beta}_2$ (Bootstrap S.E.) and 95% C.I.	
	Correct Model (Exponential)	Incorrect Model (Linear)
1	0.489790 (**0.083156**) (0.277786, 0.604716)	1.642331 (**0.438216**) (0.783428, 2.501234)
2	0.518016 (**0.075583**) (0.330897, 0.634137)	1.869030 (**0.432247**) (1.021826, 2.716234)
3	0.490315 (**0.068855**) (0.355359, 0.625271)	1.653745(**0.474687**) (0.723358, 2.584132)
4	0.444948 (**0.085288**) (0.268765, 0.568174)	1.400871 (**0.424893**) (0.568081, 2.233661)
5	0.537602 (**0.061730**) (0.393354, 0.628543)	1.996479 (**0.335457**) (1.338983, 2.653975)

Table15. Result of comparison of spatial effects over time (T=100) when an exponential model is specified as linear under constant spatial effect with true parameter value 0.5 using 95% coverage probability.

Criterion [Value]		Decision on H_0	
Exponential	Linear	Exponential	Linear
Mean [0.508173]	Mean [1.776369]	Do not reject H_0 (0)	Do not reject H_0 (0)
Median [0.506551]	Median [1.781646]	Do not reject H_0 (0)	Do not reject H_0 (0)

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

Under not constant spatial effect with 5% different time points, the decisions on the null hypothesis for all cases (incorrectly not rejecting the null hypothesis) are the same using the two models except at β_2=0.9 based on the mean (Table 16). This is due to one interval for the common value that failed to contain the average since it is associated with the lowest point estimate. In this case, correct inference resulted using the incorrect model. Using the linear model also resulted to higher power of the test except at β_2=0.2, 0.3, 0.4 when equal power was attained.

Table 16. Results of comparison of spatial effects over time (T=100) with no spatial clustering when an exponential model is specified as linear in seven cases with underlying non-constant spatial effect (5% different with same spatial effect) using 95% coverage probability**.**

Case No. and Description	Criterion[Value]		Decision on H_0		Power of the Test	
	Exponential	Linear	Expo.	Linear	Expo.	Linear
1. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.9$	Mean [0.528018]	Mean [2.088816]	Do not reject H_0 (3)	Reject H_0 (6)	0.6	1
	Median [0.508085]	Median [1.789384]	Do not reject H_0 (4)	Do not reject H_0 (5)	0.8	1
2. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.8$	Mean [0.523057]	Mean [2.005405]	Do not reject H_0 (3)	Do not reject H_0 (5)	0.6	1
	Median [0.508085]	Median [1.789384]	Do not reject H_0 (3)	Do not reject H_0 (5)	0.6	1
3. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.7$	Mean [0.518095]	Mean [1.922657]	Do not reject H_0 (1)	Do not reject H_0 (4)	0.2	0.8
	Median [0.508085]	Median [1.789384]	Do not reject	Do not	0.2	1

			H_0 (1)	reject H_0 (5)		
4. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.6$	Mean [0.513134]	Mean [1.844411]	Do not reject H_0 (0)	Do not reject H_0 (1)	0	0.2
	Median [0.508085]	Median [1.789384]	Do not reject H_0 (0)	Do not reject H_0 (1)	0	0.2
5. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.4$	Mean [0.503211]	Mean [1.725736]	Do not reject H_0 (4)	Do not reject H_0 (5)	0.8	1
	Median [0.506106]	Median [1.773365]	Do not reject H_0 (5)	Do not reject H_0 (5)	1	1
6. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.3$	Mean [0.498250]	Mean [1.698292]	Do not reject H_0 (5)	Do not reject H_0 (5)	1	1
	Median [0.506106]	Median [1.773365]	Do not reject H_0 (5)	Do not reject H_0 (5)	1	1
7. 95 time points with $\beta_2 = 0.5$; 5 time points with $\beta_2 = 0.2$	Mean [0.493289]	Mean [1.691055]	Do not reject H_0 (5)	Do not reject H_0 (5)	1	1
	Median [0.506106]	Median [1.773365]	Do not reject H_0 (5)	Do not reject H_0 (5)	1	1

Note: Figures in parentheses are the number of bootstrap confidence intervals that failed to contain the mean/median.

Under not constant spatial effect with 8% different time points, using the linear model resulted to correct inference for all cases except at β_2=0.6. This is due to the more pronounced difference between the alternative parameter estimates and the common value estimates using the linear model. At β_2=0.7, 0.8, 0.9, the parameter estimates are very much higher than those at β_2=0.5 and at β_2=0.2, 0.3, 0.4, they are very much lower than those at β_2=0.5. On the other hand, there was incorrect inference at both β_2=0.6 and β_2=0.7 using the correct model. That is, the test procedure was more sensitive in detecting a difference in parameter value using the incorrect model. Also, misspecifying the model even yielded higher power of the test at alternative values below the common value. Thus, even under the condition of not constant spatial effect over time, the proposed test procedure remains robust to model misspecification.

5. CONCLUSIONS

The proposed procedures can correctly identify the actual behavior (constant and not constant parameter effect) of the temporal and spatial parameters, respectively. Power of the test (proportion of intervals constructed for the alternative parameter value that do not contain the common value) was higher (close to 1 or equal to 1) for distant alternative values and equal to zero or close to zero for alternative parameter values very close to the common value, for both test procedures. When spatial clustering was imposed, the proposed test procedure became more sensitive in detecting the presence of spatial units that can cause the rejection of the constant spatial effect over time assumption. Consequently, the power of the test increased under spatial clustering for alternatives distinctly different from the null.

The issue of model used to estimate temporal effect is critical for correct inference in testing for constant temporal effect across spatial units. In testing for constant spatial effect over time, the model used to estimate spatial effect across time is not critical. It is expected that using this procedure will not pause any problem at all. In the presence of substantial spatial clustering, the test can gain more power.

REFERENCES.

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

Chernick, M.R. (1999). *Bootstrap Methods: A Practitioner's Guide*. New York: John Wiley & Sons, Inc.

Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and their Application*. Cambridge:Cambridge University Press.

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