Autologistic Spatial-Temporal Modeling

Ma. Andriena Ida B. del Ayre-Ofina
Industrial Technology Development Institute -
Department of Science and Technology

We postulate a combination of spatial-temporal and autologistic model in characterizing binary data collected over time and space. Using a second-order neighborhood system in defining the spatial component of the model, backfitting algorithm is used in estimating the model. As the incidence of success and failure responses becomes balanced, sensitivity and specificity increases. The predictive ability of the model is fairly robust to the spatial parameter but is significantly influenced by the temporal parameter. The bias of the estimate for the spatial parameter declines as it becomes dominant in the model. Furthermore, as autocorrelation becomes stronger, its estimate becomes less biased. The backfitting algorithm is also observed to converge fast in the estimation of the spatial-temporal autologistic model.

Keywords: *binary response, autologistic model, spatial-temporal model, backfitting*

1. Introduction

The last few decades have seen literature on spatial-temporal modeling. From models accounting only temporal dependencies or spatial dependencies, more integrated and hybrid models were proposed. The models now account for both spatial and temporal characteristics of the data. These models aptly characterize data collected from different spatial locations over a period of time, e.g., those data in ecology, biology, meteorology, etc. (Niu, 1998).

Spatial-temporal model takes into account spatial correlation and time dependence in addition to covariates (Fermin, 2001). On the other hand, autologistic model is used for predicting occurrence of certain characteristics while taking spatial correlation into account. Landagan (2005) noted that autologistic model was an early predecessor of spatial models. Autologistic models can be used to estimate the probability of success at a given site and predict the outcome at an unsampled site (Zhu et al., 2005).

We postulate in this paper a combination of spatial-temporal and autologistic model to characterize a binary response collected over space and time. The parameters of the model are then estimated through the backfitting algorithm. The autologistic spatial-temporal model as well as the backfitting algorithm was investigated through a simulation study.

The model postulated in this paper is similar to Landagan (2005), adding the feature of autologistic model. On the other hand, it generalizes the model in Fermin (2001) which is an autologistic modeling accounting for the spatial dependencies only. The use of the backfitting algorithm is motivated by Fermin (2001) and Landagan (2005) who observed the optimality of the algorithm.

Gumpertz et al. (1997), Huffer and Wu (1998), and Zhu et al. (2005) used autologistic models in characterizing varying phenomena. Gumpertz et al. (1997) postulated an autologistic model that takes into account temporal correlations and estimated all parameters simultaneously using the maximum pseudo-likelihood (MPL) method. Huffer and Wu (1998), on the other hand, used Markov chain Monte Carlo (MCMC) in estimating the parameters in an autologistic regression model for spatial binary data with covariates. Zhu et al. (2005) also used maximum pseudo-likelihood (MPL) for parameter estimation of a spatial-temporal autologistic regression model.

2. Spatial-Temporal Models

The autologistic model is a flexible model for predicting incidence or characteristics of events (Gumpertz et al., 1997). Earlier studies considered only the relationship between a binary response and potential explanatory variables. Recently, spatial correlation of the responses at neighboring locations on a regular or irregular lattice are taken into account. In autologistic model, the log odds of a characteristic in a particular quadrat are explained as a linear combination of characteristic incidence in neighboring quadrats and the other predictor variables (Gumpertzetal., 1997). Studies considered not only the spatial dependence of a binary data but also its temporal dependence. Thus, the so-called autologistic spatial-temporal model evolved. The autologistic spatial-temporal model captures the relationship between a binary response and potential explanatory variables, and makes possible for adjustments in both spatial dependence and temporal dependence (Zhu et al., 2005).

Murtaugh and Phillips (1998) studied the temporal feature of a statistical data, and developed a bivariate binary model for estimating the change in land cover from satellite images obtained at two different times. Two images, consisting of pixels, of Mexican emission data and uptake of atmospheric carbon dioxide were used. The binary classifications of each pixel as forest (1) or nonforest (0) cover at measured at two occasions (1986 and 1990) were modeled based from the condition of the true states of the pixel. Two types of classification errors were noted: a forested pixel was classified as nonforest, and an unforested pixel was classified as forest. To test for temporal correlation of pixel classification, the model was fitted to the dataset in which the true land cover was presumed from a reference dataset. Results indicated the existence of significant temporal correlation of classifications. They presented two model-based methods using data with complete information to adjust for temporal correlation in the classifications of nearby pixels that lack reference information.

HØst et al. (1995) separately modeled the spatial mean field, the spatial variance field, and the space-time residual fields to give a more detailed and more accurate representation of spatial interpolation errors when using repeated observations on a fixed monitoring network. The modeling approach was applied to monitoring data of sulfur dioxide concentration in air pollution from a European network for the period 1980-1986.

Gumpertz et al. (1997) applied autologistic model in disease incidence in neighboring quadrats of an agricultural field, based on soil characteristics. Thus, information about the degree of spread in different directions of the disease was obtained. Three models were included: Model 1 was a logistic regression model with two covariates ignored spatial correlation; Model 2 was a second-order autologistic model with two covariates, and; Model 3 was a pure second-order autologistic model without covariates. The three models were applied to disease epidemic in bell pepper. Parameters were estimated using a method called maximum pseudo-likelihood. Cross validation was also done to evaluate the predictive ability of the model, to examine the stability of parameter estimates, and to check for influential quadrats.

Autologistic regression models for binary responses exhibiting both spatial correlation and dependence on covariates were also explored by Huffer and Wu (1998). The model was applied to the distribution of about 180 plant species in the state of Florida in terms of nine (9) climate variables. They used Markov chain Monte Carlo (MCMC) to estimate the parameters in these models. The MCMC maximum likelihood estimates distributional behaviour was studied via simulation.

Autologistic and logistic model with spatial parameter with the assumption of autocorrelated errors, were further studied by Fermin (2001). The results indicated that the autologistic model is superior to the other model in terms of its ability to classify binary events. The binary response variable of “bloom” and “no bloom” were used to describe the occurrence of the red tide phenomenon. Significant improvement in the prediction of binary events was noted in the inclusion of the spatial parameter. Parameter estimation used the backfitting algorithm.

When data is composed of spatial and temporal features, it is better to use spatial-temporal statistical modeling than spatial only or temporal only aspect of the data. Spatio-temporal statistical analysis can provide benefits not possible from a spatial-only approach when spatial data are collected over time (Haas, 1995).

Zhu et al. (2005) developed an autologistic regression model which captured both spatial dependence and temporal dependence simultaneouslyusing a binary data on the incidence of southern pine beetle in the 100 counties of North Carolina. It was noted that throughout the model-selection steps, both components became the most significant variables. Maximum pseudo-likelihood was used in estimating the model parameters. Optimal prediction of future responses on the lattice was obtained via Gibbs sampler, a Markov chain Monte Carlo algorithm.

Another application of spatial-temporal statistical modeling and prediction is in atmospheric pollutant deposition processes discussed by Haas (1995). Particularly, he estimated a model of wet sulfate deposition, which captures the major effects of location, time, and season. The prediction method used was moving-cylinder spatial-temporal kriging (MCSTK), an extension of MWRRK (Moving-Window Regression Residual Kriging). Since MCSTK is computationally intensive, it is practical only if a high-end PC or workstation is available.

Cressie and Majure (1997) considered modeling of the large-scale variation and the small-scale (in space and time) variation in the (log) nitrate concentration in streams of the upper North Bosque watershed.

3. Methodology

Consider a cultivated agricultural field divided into a number of plots, usually close or adjacent to each other. A farmer plants a crop in each plot, for example, tomatoes. To protect the crops, the farmer applies pesticides. In a large agricultural field, the pesticides are usually sprayed, sometimes a few days before harvesting. In this so-called “pre-harvest period,” the number of days before harvest when pesticide was last applied, some amount of pesticide residue is left in the expected produce.

On the other hand, crops are harvested depending on its season, thus the so-called “seasonal crops.” In a year, different crops are harvested in different months. Thus, it can be said that every quarter of the year different crops are being harvested.

Suppose that we consider a binary response (presence/absence of pesticide residue in the crop within a plot), we aim to characterize the dynamics if the response variable over space and across time.

3.1 Spatial-temporal model for a binary response variable

Given a binary process $\{Y_{ijt} : i = 1,\ldots, n, j = 1,\ldots, m, \mathrm{t} \in Z\}$ observed in the spatial location i, j at time t. Consider Figure 1 below.

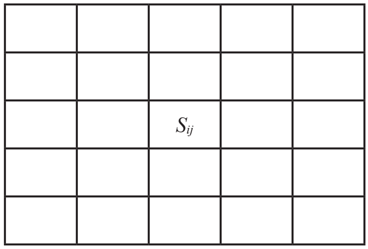

Figure 1. Spatial Partition of the Study Population

Observations in cell S_{ij} are compiled over a period of time, say $t = 1, 2,\ldots,$ T. Suppose Y_{ijt} is a binary variable, i.e.

$$Y_{ijt} = \begin{cases} 1, \text{ if characteristic A is present in site } S_{ij} \text{ at time t} \\ 0, \text{ otherwise} \end{cases}$$

Then we postulate the spatial-temporal autologistic model as follows:

$$Y_{ijt} = \frac{\exp(X_{ijt}\beta + w_{ijt}\gamma)}{1+\exp(X_{ijt}\beta + w_{ijt}\gamma)} + \varepsilon_{ijt} \qquad (1)$$

As an example, the response variable can be incidence of pesticide residue from location s_{ij} at time t. X_{ijt} denotes the set of covariates from location s_{ij} at time t. In the pesticide example, X can be the number of days before harvesting for the last pesticide application. W_{ijt} is the set of count/summary variables in the neighborhood system of location s_{ij} at time t. This can be taken as the sum of all second-order neighborhood values at a specific site S_{ij}. Let B_{ij}, A_{ij}, D_{ij1}, and D_{ij2} denote the second-order neighborhood sites at time t (Gumpertzet al., 1997), see the illustration below.

		Row		
		i+1	i	i-1
	j+1	D_{ij1}	B_{ij}	D_{ij2}
Quadrat	j	A_{ij}	S_{ij}	A_{ij}
	j-1	D_{ij2}	B_{ij}	D_{ij1}

Figure 2. Types of Neighbors at Site S_{ij}

where

$Bij = Y_{i,j-1,t} + Y_{i,j+1,t}$, the number of quadrats with the characteristic of interest present of the two adjacent quadrats within the same row;

$A_{ij} = Y_{i+1,j,t} + Y_{i-1,j,t}$, the number of quadrats of the two adjacent quadrats in neighboring rows with the characteristic present;

$D_{ij1} = Y_{i+1,j+1,t} + Y_{i-1,j-1,t}$, number of quadrats of the two diagonal quadrats in (1,1) and (-1,-1) direction with the characteristic present;

$D_{ij2} = Y_{i+1,j-1,t} + Y_{i-1,j+1,t}$, number of quadrats of the two diagonal quadrats in (1,-1) and (-1,1) direction with the characteristic present;

ε_{ijt} is the error term that assumes the process given by: $\varepsilon_{ijt} = \mu_{ij} + \upsilon_{ijt}$ where $\mu_{ij} \sim$ IID $(0, \sigma^2_{\mu})$; υ_{ijt} follows a stationary AR (1) $\upsilon_{ijt} = \rho\upsilon_{ijt,t-1} + a_{ijt}$ with $|\rho|<1$. $a_{ijt} \sim$ IID $(0, \sigma^2_{a})$; μ_{ij}'s are independent of the υ_{ijt}'s and $\upsilon_{ij0} \sim (0, \sigma^2_{a}/(1-\rho^2))$. The following assumptions by Landagan (2005) are also assumed: constant covariance (β) effect across locations and time; constant temporal effect (ρ) across locations, and; constant spatial effect (γ) across time.

3.2 Backfitting estimation of the autologistic spatial-temporal model

The Backfitting Algorithm is used in the estimation of the autologistic spatial-temporal model. Given equation (1), $Y_{ijt} = \dfrac{\exp(X_{ijt}\beta + w_{ijt}\gamma)}{1+\exp(X_{ijt}\beta + w_{ijt}\gamma)} + \varepsilon_{ijt}$, the logit is given by $\ln\left(\dfrac{Y_{ijt}}{1-Y_{ijt}}\right) = X_{ijt}\beta + w_{ijt}\gamma + \varepsilon_{ijt}$. Let $\ln\left(\dfrac{Y_{ijt}}{1-Y_{ijt}}\right) = Y^*_{ijt}$. Then, the model reduces to

$$Y^*_{ijt} = X_{ijt}\beta + w_{ijt}\gamma + \eta_{ijt} \qquad (2)$$

Considering the additive nature of Equation 2, the backfitting algorithm is implemented in the following steps to estimate the parameters.

Step 1: For each t, ignoring the term involving w_{ijt} and the autocorrelations,

1a.) Estimate β from $Y^*_{ijt} = X_{ijt}\beta + \psi_{ijt}$, there will be T such estimates of β, say $\hat{\beta}_t$, t = 1,.., T.

Compute the average of these estimates say, $\hat{\beta}^*$.

1b.) Compute the residuals $\hat{Y}_{ijt} = Y^*_{ijt} - X_{ijt}\hat{\beta}^*$

These residuals contain information on γ that was ignored initially.

Estimate γ from $\hat{Y}_{ijt} = w_{ijt}\gamma + \Upsilon_{ijt}$ for each t = 1,…,T

Compute the average of these T estimates of $\hat{\gamma}$ say $\hat{\gamma}^*$.

Step 2: Fix the location i, j. Compute the residual $e_{ijt} = \hat{Y}_{ijt} - w_{ijt}\hat{\gamma}^*$.

These new residuals contain information on ρ.

Estimate ρ from $e_{ijt} = \rho e_{ijt-1} + a_{ijt}$.

This results to nm estimates of ρ. Compute the average of these estimates $\hat{\rho}^*$.

Step 3: Iterate from Step 1 using the most recent parameter estimates every time.

From Step (1a) and (1b) new residuals based on $\hat{\rho}^*$ and $\hat{\gamma}^*$ will be computed instead of Y_{ijt}^*. Given this new estimate of $\hat{\beta}^*$, compute new residuals based on the new estimates of $\hat{\beta}^*$ and the existing estimate of $\hat{\rho}^*$. Proceed with the iteration of Step 2, computing new residuals based on the available new estimates of the parameters.

3.3 Simulation study

Models of binary data are typically used in predicting the likelihood of a success or failure given information on some covariates. Oftentimes, the predicted values are between 0 and 1, hence, one has to decide whether to take 0 or 1 as the likely outcome based on the predicted value. It is then possible to commit an error in the process of classifying success or failure based on the data. Sensitivity is defined as the proportion of correctly predicted success, while specificity is the proportion of correctly predicted failure.

To assess the performance of the model, sensitivity and specificity are analyzed using simulated data. We summarize the simulation procedure with the following:

Step 1. Predictor Variable (X): Generate the X's from a distribution.

Step 2. Error Terms (ε): Generate the error terms from N $(0, \sigma^2)$.

Step 3. Parameter Values (β's): Assume values for the parameters β_0 and β_1

Step 4. Response Variable (Y): Given the values for the β's, X's, and ε's, compute the response variable Y_{ijt}^0 from equation (1), this time ignoring w_{ijt} first. This will simulate the initial binary variables Y_{ijt}^0 values.

Step 5. Recode Y_{ijt}^0 to zero (0) or one (1) depending on the choice of incidence level. Choice of incidence level is arbitrary (usually 5% to 50%).Note that the cut-off value used in this step is dependent on the incidence level about the incidence of success and failure in the data. For example, if a 25% incidence is targeted, then the data is sorted in decreasing magnitude of Y and the top 25% are coded as one, the rest are coded as zero.

Step 6. Spatial Component (w_{ij}): $w_{ijt}\gamma = \gamma_1 B_{ij} + \gamma_2 A_{ij} + \gamma_3 D_{ij1} + \gamma_4 D_{ij2}$. The spatial component w_{ij} is the sum of all second-order neighborhood values at a specific site S_{ij}. The second-order neighborhood sites are B_{ij}, A_{ij}, D_{ij1}, and D_{ij2}. Where:

$B_{ij} = Y_{i,j\text{-}1,t} + Y_{i,j+1,t}$,

$A_{ij} = Y_{i+1,jt} + Y_{i\text{-}1,jt}$,

$D_{ij1} = Y_{i+1,j+1,t} + Y_{i\text{-}1,j\text{-}1,t}$, and

$D_{ij2} = Y_{i+1,j\text{-}1,t} + Y_{i\text{-}1,j+1,t}$.

The Y_{ijt0} generated in Step 4 and recoded in Step 5 were used for the assignment of the second-order neighborhood values Y_{ijt} at each site S_{ij}.

Step 7. Parameter Values (γ's): Assume values of the spatial parameters γ_1, γ_2, γ_3, and γ_4.

Step 8. Given the values for X's, ε's, β's, γ's, and w's generate another set of response variable denoted by Y_{ijt}. Note that the Y_{ijt}'s in Step 4 were used to simulate the neighborhood values while in this step, it accounts for both the logistic component and the neighborhood component of the model.

Step 9. Recode Y_{ijt} to zero or one using the cut-off used in Step 5.

Step 10. Finally, given Y_{ijt}'s, X's, and w's estimate the parameters for the model.

The goal in implementing steps 1 to 9 is to generate a binary variable that is related to X through the model in equation 1. The spatial and temporal structure for y_{ijt} and the associated error terms are also considered in the simulation.

3.4 Basis for the simulation conditions

The data simulated in this study benefits from a priori information about pesticides contained on the monographs of data and recommendations of the joint meeting of the FAO (Food and Agriculture Organization of the United Nations) Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Expert Group on Pesticide Residues for the year 1967, 1979 and 1984. In the monographs, crops were listed down together with initial pesticide residues (in ppm), typical pre-harvest period (in days), residues at pre-harvest indicated, and other related agronomic characteristics.

The following are the conditions considered in the simulation:

1. Predictor Variable (X): In this study the pre-harvest period, the number of days before harvest when pesticide was last applied was considered as the predictor variable (X). The average pre-harvest period is 10 days. Thus, X was simulated from a distribution with mean of 10.

2. Error Terms (ε): The error terms were simulated from the normal distribution with mean zero and variance equal to one, i.e., $\varepsilon \sim N(0, 1)$.

3. Parameter Values (β's and γ's): For the parameter β's, this study also used the value used by Huffer and Wu (1998) in their simulation study. They used $\beta_0 = 1$ and $\beta_1 = 2$. For $\beta_0 = 1$, however, usually this is the value of Y when X =

0. While with $\beta_1 = 2$, which produces strong covariate effects, they noted that the results were qualitatively similar when they run simulations using other values of β_1 (Huffer and Wu, 1998).For the spatial parameter, Huffer and Wu (1998) used a wide range of γ values, the results showed, however, that the estimation error (both bias and variance) and the skewness increase as the spatial interaction increases. They also examined negative values of γ, but in application, γ is generally positive. Thus, in this study, only two (2) values of γ will be used, $\gamma = 0.2$ and $\gamma = 0.8$, representing a small and a moderately large extent of spatial dependencies.

4. Spatial Component (w): The spatial component considered in the model is the second-order neighborhood system. Gumpertz et al. (1997) illustrated the first-order, second-order, and the third-order systems of neighbors of a standard rectangular lattice.

								*		
	*		*	*	*		*	*	*	
*	X	*	*	X	*	*	*	X	*	*
	*		*	*	*		*	*	*	
								*		
First-order			Second-order			Third-order				

Figure 3. Standard Systems of Neighbors

The second-order system includes the four diagonal neighbors in addition to the four adjacent quadrats in the set of neighbors, two within the row and two in adjacent row. Note, however, that it is not necessary that the lattice have any regular shape and it is not necessary to use the standard neighborhood systems. The requirement is to define a set of neighbors for each quadrat in a lattice. If quadrat i is a neighbor of quadrat j, then the converse is also true.

Spatial autocorrelation must be incorporated in the data with autologistic characteristic by conditioning the probability of the occurrence of the characteristic in quadrat when the characteristic is present in neighboring quadrats.

5. Temporal Component: This study assumed a first order autoregressive AR(1) process to account for temporal dependencies into the model. Wei (1990) explained that AR processes are useful in describing situations, which the present value of a time series depends on its preceding values plus a random shock.

Gumpertz et al. (1997) fitted their models to the inner 16x16 lattice of 256 quadrats. This is to accommodate adjacent quadrats and quadrats two spaces away. The simulation study done by Huffer and Wu (1998) used autologistic regression model on a complete 40x40 lattice; making the number of sites 1600. In this data simulation, a 5 x 5 lattice design was used. Thus, there were 25 simulated sites at 60 time points for the fifteen (15) years of quarterly data. This resulted in 1,500 cross-section time series data points.

For the second-order scheme a coding pattern as in Figure 4 was used. The values at the ⊗ sites are considered conditional on the values at the • sites (Bartlett, 1975).

•	⊗	•	⊗	•
•	•	•	•	•
•	⊗	•	⊗	•
•	•	•	•	•
•	⊗	•	⊗	•

Figure 4. 5x5 Second-Order Coding Pattern

After all the variables were simulated or generated, this was repeated one hundred times. Thus, 100 datasets were generated, each dataset has 36 different scenarios with 1500 observations per scenario. The simulation scenarios are summarized on Table 1. Scenarios were constructed by considering 5%, 25% and 50% incidence of samples possessing the characteristic of interest, that is, $Y_{ijt} = 1$, otherwise, $Y_{ijt} = 0$. Thus for 5%, 25% and 50% incidence this means that about 75, 375, and 750 respectively, out of 1500 observations have a value of Y =1.

4. Results and Discussion

Three incidence levels were considered for the samples where certain characteristic are present, i.e. 5%, 25%, and 50%. Two sets of values for neighborhood parameter (γ), and six sets of values for autocorrelation parameter (ρ), were used as inputs in the simulation. Thirty-six (36) scenarios were investigated to represent different values of the neighborhood and autocorrelation parameters, see Table 1.

Sensitivity and specificity of the model in detecting incidence of the characteristics of interest were examined as indicators of the predictive ability of the model. Sensitivity is the conditional probability that the model will correctly classify the "presence" event. On the other hand, specificity is the conditional probability that the model will correctly classify an "absence" event. High values of sensitivity and specificity are desirable but inherent relationship between the two types of errors would force the decision-maker to compromise. One may favor sensitivity at the expense of specificity or vice versa. Sensitivity and specificity will also be affected by the distribution of incidence of success and failure in the data.

Table 1. Simulation Scenarios

Incidence Level	Gamma	Rho	Scenario
5% incidence in the population	0.2	0.2 0.4 0.6 0.75 0.85 0.95	1 2 3 4 5 6
25% incidence in the population	0.2	0.2 0.4 0.6 0.75 0.85 0.95	7 8 9 10 11 12
50% incidence in the population	0.2	0.2 0.4 0.6 0.75 0.85 0.95	13 14 15 16 17 18
5% incidence in the population	0.8	0.2 0.4 0.6 0.75 0.85 0.95	19 20 21 22 23 24
25% incidence in the population	0.8	0.2 0.4 0.6 0.75 0.85 0.95	25 26 27 28 29 30
50% incidence in the population	0.8	0.2 0.4 0.6 0.75 0.85 0.95	31 32 33 34 35 36

4.1 Sensitivity and specificity

Table 2 summarizes specificity and sensitivity as the characteristic of interest becomes dominant. On the average, specificity of the autologistic spatial-temporal model decreases as the incidence level increases. On the other hand, sensitivity of the model increases as the incidence level increases. However, the unit increase in sensitivity rate is larger compared to the unit decrease of specificity rate. It is common in binary data response models to easily predict a case that is more commonly observed in the data. If more success cases are included in the data, then sensitivity is higher. While specificity is higher when there are more failure cases in the data.

Similar observation is noticed for the autologistic spatial-temporal model to correctly classify events. The correct classification of events by the model decreases as the incidence level in the data increases.

Table 2. Sensitivity and Specificity Rates by Incidence Level for 100 Replicates

Incidence Level	Statistics	Specificity	Sensitivity	Correct Classification
5%	Min Max Mean	94.95 99.37 96.7255	4.00 88.00 37.7765	90.40 98.80 93.7393
25%	Min Max Mean	76.62 96.53 85.7711	24.53 89.63 57.0669	63.74 94.80 78.6807
50%	Min Max Mean	53.07 92.00 72.7492	42.74 91.87 72.6364	48.26 92.53 72.6546

From Table 3, the behaviour of the model concerning specificity, sensitivity and the ability to correctly classify the events is the same between the two values of the neighborhood or spatial parameter used. The predictive ability of the autologistic spatial-temporal model is fairly robust to the magnitude of the spatial parameter.

Table 3. Sensitivity and Specificity Rates by Neighborhood (Spatial) Parameter for 100 Replicates

Neighborhood Parameter	Statistics	Specificity	Sensitivity	Correct Classification
0.2	Min Max Mean	53.07 99.37 85.0918	4.00 91.87 55.8463	52.87 98.80 81.6965
0.8	Min Max Mean	53.07 99.37 85.0721	4.00 91.87 55.8069	48.26 98.80 81.6866

Table 4 shows the behaviour of the model with respect to the autocorrelation (temporal) parameter. Both specificity and sensitivity increases as the autocorrelation parameter becomes stronger. Sensitivity and specificity benefits from the increasing temporal dependency in the data. Stronger autocorrelation parameter value implies the strength of information contained among the more recent observations hence, more accurate assessment of whether the case is likely to be a success or a failure. The role of the temporal dependencies in prediction of binary events in a spatial-temporal model is more important than the spatial dependencies. In addition, the percentage of correct classification becomes larger as the autocorrelation parameter increases.

Table 4. Sensitivity and Specificity Rates by Autocorrelation (Temporal) Parameter

Autocorrelation Parameter	Statistics	Specificity	Sensitivity	Correct Classification
0.2	Min Max Mean	53.07 95.79 76.5310	4.00 59.07 33.2154	48.26 95.00 71.3385
0.4	Min Max Mean	61.10 96.21 80.0010	10.67 65.33 41.8840	60.07 92.80 75.4277
0.6	Min Max Mean	68.76 96.77 83.6269	22.67 72.27 51.0399	65.53 94.33 79.9445
0.7	Min Max Mean	71.87 97.33 86.0409	30.67 77.23 58.0840	71.74 95.34 82.8674
0.85	Min Max Mean	80.53 98.11 90.2065	24.53 84.95 68.9430	79.34 96.80 87.8197
0.95	Min Max Mean	85.58 99.37 94.0854	60.00 91.87 81.7933	85.26 98.80 92.7514

4.2 Bias of parameters estimates

Percent bias is computed as follows:

$$\text{Percentage Bias} = \frac{\text{Actual-Estimate}}{\text{Actual}} \text{ x } 100\%.$$

This is used to assess the extent in which underestimation or overestimation observed in the course of backfitting estimation of the spatial-temporal model.

Table 5. Percentage Bias of Covariates β_0 and β_1 by Incidence Level

Incidence Level	β_0	β_1
5%	97.1748	50.0023
25%	85.6770	50.0135
50%	58.8782	50.0000

From Table 5, we see that there is an underestimation for both parameters β_0 and β_1 for the three incidence levels of the data. Notice that the percentage bias decreases for β_0 while the percentage bias stays almost the same for β_1 as the percentage of incidence level increases.

Table 6 shows that there is an overestimation for both values of the spatial parameter γ across the different incidence levels. Percentage biases for $\gamma = 0.2$ is much greater than the percentage biases for $\gamma = 0.8$. Nevertheless, in both spatial parameters, the percentage biases decreases as the incidence level increases. Thus, higher spatial parameter value has much lower percentage bias.

Table 6. Percentage Bias of Spatial Parameter (γ) by Incidence Level

Incidence Level	$\gamma = 0.2$	$\gamma = 0.8$
5%	-1490.08	-297.5273
25%	-781.2643	-120.2167
50%	-679.4346	-94.8562

With the same observation as Table 6, Table 7 also exhibits an overestimation of the actual value used for all values of the temporal or autocorrelation parameters ρ for the three incidence levels. Percentage Biases for $\rho = 0.2$ is the largest among the temporal parameters. It decreases as the autocorrelation parameter becomes stronger. However, the percentage biases increases as the incidence level increases. Again, as the temporal parameter ρ becomes stronger the value of the percentage bias becomes smaller.

Table 7. Percentage Bias of Temporal Parameter (ρ) by Incidence Level

Incidence Level	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$	$\rho = 0.7$	$\rho = 0.85$	$\rho = 0.95$
5%	-402.9155	-150.4213	-67.7349	-43.7706	-19.9965	-9.9633
25%	-407.9493	-154.9720	-69.4101	-44.5320	-20.5285	-14.2561
50%	-419.0633	-161.0154	-74.2778	-48.9404	-26.8316	-29.7371

4.3 Parameter estimates

The following section shows the summary of the parameter estimates upon convergence. Line graphs are also included to illustrate the behaviour of these estimates.

4.3.1 Covariate parameters

From Figure 5, notice that only two scenarios, scenario 3 and scenario 21, have a mean intercept value (0.81) closer to the initial value of the intercept equal to 1. In general, the intercept was not accurately estimated. This may be explained partly due to the simulation process. The parameters to be estimated were entered only at the initial steps of the simulation.

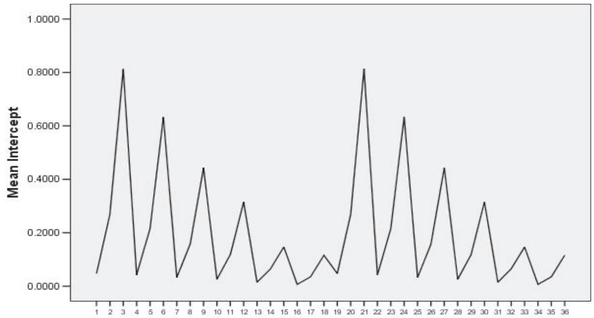

Figure 5. Behaviour of Intercept Estimates Upon Convergence

In Figure 6, for the 36 scenarios, the estimates for the X covariate parameter are all close to 1. Estimates of β_1 are all close to 1, this is one-half of the initial value of the β_1 parameter used, which is 2.

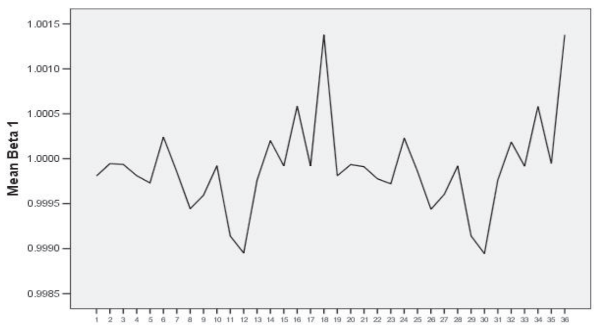

Figure 6. Behaviour of X Covariate Estimates Upon Convergence

4.3.2 Spatial parameter (γ)

All the 36 scenarios did not recover the assigned parameter values in the simulation. The values are much larger than the two initial spatial parameter values: 0.2 and 0.8. Figure 7 below, shows that the spatial parameter values for the first 18 scenarios are the same with the next 18 scenarios.

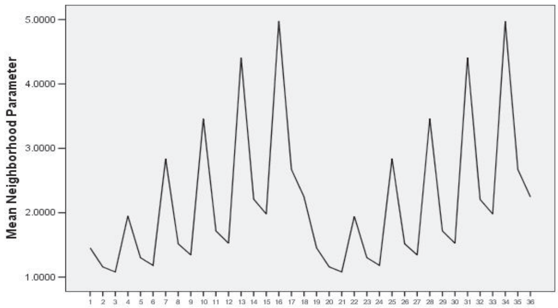

Figure 7. Behaviour of Neighborhood Parameter Estimates Upon Convergence

4.3.3 Temporal Parameter (ρ)

All the estimates have values close to 1 compared to the parameter value inputs. Similarly, the behaviour of the first 18 scenarios is the same with the next 18 scenarios. Again, all initial autocorrelation parameter values (0.2, 0.4, 0.6, 0.75, 0.85 and 0.95) were not recovered upon convergence.

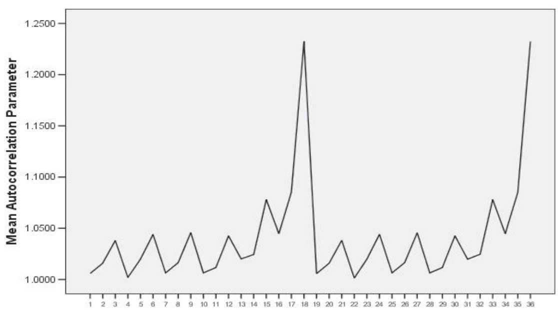

Figure 8. Behaviour of Autocorrelation Parameter Estimates Upon Convergence

4.4 Convergence of the algorithm

Usually, in an iterative process like the backfitting algorithm, convergence is a major concern that needs to be addressed. For as long as the model is truly additive, the backfitting algorithm is known to converge. In many cases, the algorithm usually converges after three (3) iterations. There are scenarios where it took 4,105 iterations before convergence. Tables 8, 9, and 10 summarize the percentage of replicates which converges for more than 3 iterations.

In Table 8, percentage of replicates that converges after more than three iterations between the covariates parameters are similar. With the highest value at 3.75% for the scenario with fewest incidence, while it is 0.50% for the 50% incidence level.

Table 8. Percentage of Replicates that Converges After More Than Three (3) Iterations Based on the Covariate by Incidence Level (N=100)

Incidence Level	β_0	β_1
5%	3.75%	3.75%
25%	1.42%	1.42%
50%	0.50%	0.50%

Except for the 50% incidence level, the percentage of replicates that converges after more than three iterations between the two spatial parameters used are slightly different in Table 9. For 5% incidence level, the percentage of replicates which converges after more than three iterations for $\gamma = 0.8$ is greater than the percentage of replicates which converges after more than three iterations for $\gamma = 0.2$. It is the other way around for the 25% incidence level, wherein the percentage of replicates which converges after more than three iterations for $\gamma = 0.2$ is greater than the percentage of replicates which converges after more than three iterations for $\gamma = 0.8$.

Table 9. Percentage of Replicates that Converges After More Than Three (3) Iterations Based on the Spatial Parameter γ by Incidence Level (N=100)

Incidence Level	$\gamma = 0.2$	$\gamma = 0.8$
5%	3.67%	3.83%
25%	2.67%	1.50%
50%	0.50%	0.50%

Table 10 shows the percentage of replicates that converges after more than three iterations among the temporal parameters. Except for $\rho = 0.85$, all the rest have decreasing percentage of replicates which converges after more than 3 iterations as the incidence level increases. With the highest percentage of replicates at 9% for $\rho = 0.4$ and lowest percentage of 0.5% for $\rho = 0.95$.

Table 10. Percentage of Replicates that Converges After More Than Three (3) Iterations Based on the Temporal Parameter (ρ) by Incidence Level (N=100)

Incidence Level	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$	$\rho = 0.7$	$\rho = 0.85$	$\rho = 0.95$
5%	1.0%	9.0%	2.0%	4.0%	2.5%	4.0%
25%	-	1.0%	1.0%	3.0%	3.0%	0.5%
50%	-	-	1.0%	2.0%	-	-
" - " all replicates converged after three (3) iterations						

5. Conclusions

We postulate a combination of spatial-temporal and autologistic model in modeling a binary data collected over time and space. A second-order neighborhood system is used in the definition of the spatial component of the model. The backfitting algorithm is then used in estimating the model.

On the average, specificity of the autologistic spatial-temporal model decreases as the incidence level increases. On the other hand, sensitivity of the model increases as the incidence level increases. Similar observation is noticed for the autologistic spatial-temporal model to correctly classify events. The correct classification rate of events by the model decreases as the incidence level increases.

It is common in binary data response models to easily predict a case that is more commonly observed in the data. If more success cases are included in the data, then sensitivity is higher. While specificity is higher when there are more failure cases in the data.

The behaviour of the model on specificity, sensitivity, and the ability to correctly classify the events are comparable for the two values of the spatial parameter used. Hence, the predictive ability of the autologistic spatial-temporal model is fairly robust to the magnitude of the spatial parameter.

However, the model is significantly influenced by the temporal parameter since the percentage of correct classification becomes larger as the autocorrelation parameter increases.

Sensitivity and specificity benefits from the increasing temporal dependency in the data. Stronger autocorrelation parameter value implies the strength of information contained among the more recent observations hence, more accurate assessment of whether the case is likely to be a success or a failure. The role of the temporal dependencies in prediction of binary events in a spatial-temporal model is more important than the spatial dependencies.

There is an underestimation for both parameters β_0 and β_1 for the three incidence levels. On the other hand, there is an overestimation for both values of the spatial parameter γ across the different incidence levels. Biases for $\gamma = 0.2$ are much greater than the biases for $\gamma = 0.8$. Nevertheless, in both spatial parameters, the biases decreases as the incidence level increases. Thus, higher value of spatial parameter have much lower bias. In addition, there is an overestimation of the actual value used for all values of the temporal or autocorrelation parameters ρ for the three incidence levels. Biases for $\rho = 0.2$ is the largest among temporal parameters. It decreases as the autocorrelation parameter becomes stronger. However, the biases increase as the incidence level increases. Again, as the temporal parameter ρ becomes stronger the value of the bias becomes smaller.

The bias of the estimate for the spatial parameter declines as it becomes dominant in the model. Furthermore, as autocorrelation becomes stronger, the estimate becomes less biased.

For the spatial parameter values, the estimates are much larger than the true values. While for the temporal parameters, all the estimates have values close to the true values compared to the temporal parameter value inputs.

The backfitting algorithm converges fast in the estimation of the spatial-temporal autologistic model. In many cases, the algorithm usually converges after three (3) iterations. The percentage of replicates that converges more than three iterations between the covariates parameters are similar.

REFERENCES

BARTLETT, M.S., 1975, *The Statistical Analysis of Spatial Pattern*, London: Chapman and Hall Ltd.

CRESSIE, N. and MAJURE, J.J., 1997, Spatio-temporal statistical modeling of livestock waste in streams, *Journal of the Agricultural, Biological, and Environmental Statistics* 2(1):24-47.

DENISON, D.G.T., MALLICK, B.K. and SMITH, A.F.M., 1998, Automatic bayesian curve fitting, *Journal of the Royal Statistical Society Series B* 60 (Part 2):333-350.

EUBANK, R.L., 1988, *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, Inc.

FAN, J. and JIANG, J., 2005, Nonparametric inferences for additive models, *Journal of the American Statistical Association* 100 (471):890-907.

FERMIN, B.S., 2001, Fitting logistic and autologistic models with autocorrelated errors using backfitting, M.S. Thesis, University of the Philippines School of Statistics, Diliman, Quezon City.

GUMPERTZ, M.L., GRAHAM, J.M. and RISTAINO, J.B., 1997, Autologistic model of spatial pattern of Phytophthora epidemic in bell pepper: Effects of soil variables on disease presence, *Journal of the Agricultural, Biological, and Environmental Statistics* 2 (2):131-156.

HAAS, T.C., 1995, Local prediction of a spatio-temporal process with an application to wet sulfate deposition, *Journal of the American Statistical Association* 90 (432):1189-1199.

HASTIE, T.J. and TIBSHIRANI, R.J., 1990, *Generalized Additive Models*, London: Chapman & Hall.

HOST,G., OMRE, H. and SWITZER, P., 1995, Spatial interpolation errors for monitoring data, *Journal of the American Statistical Association* 90 (431):853-861.

HUFFER, F.W., and WU, H., 1998, Markov chain monte carlo for autologistic regression models with application to the distribution of plant species, *Biometrics* 54:509-524.

LANDAGAN, O.Z., 2005, Estimation of a proposed generalized spatial-temporal model, M.S. Thesis, University of the Philippines School of Statistics, Diliman, Quezon City.

LUO, Z., 1998, Backfitting in smoothing spline ANOVA, *The Annals of Statistics*, 26(5):1733-1759.

MURTAUGH, P.A. and PHILLIPS, D.L., 1998, Temporal correlation of classifications in remote sensing, *Journal of the Agricultural, Biological, and Environmental Statistics* 3 (1):99-110.

NIU, X., 1998, Estimating integrals of stochastic processes using space-time data, *The Annals of Statistics* 26 (6):2246-2263.

WEI, W., 1990, *Time Series Analysis: Univariate and Multivariate Methods*, Redwood City, California: Addison-Wesley Publishing Company, Inc.

ZHU,J., HUANG, H. and WU, J., 2005, Modeling spatial-temporal binary data using Markov random fields, *Journal of the Agricultural, Biological, and Environmental Statistics* 10 (2):212-225.