This is to certify that the
dissertation entitled

SPACE-TIME MODELING AND APPLICATION TO
EMERGING INFECTIOUS DISEASES

presented by

CHENG-YU LEE

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Forestry and EEBB program

Bryan K. Epperson
Major Professor's Signature

May 10 2005
Date

SPACE-TIME MODELING AND APPLICATION TO EMERGING INFECTIOUS DISEASES

By

Cheng-Yu Lee

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Forestry;
Ecology, Evolutionary Biology, and Behavior Program

2005

ABSTRACT

SPACE-TIME MODELING AND APPLICATION TO EMERGING INFECTIOUS DISEASES

By

Cheng-Yu Lee

A refined modeling framework for space-time analyses, specifically developed for general Space-Time Autoregressive Moving Average (STARMA) models, is proposed. To enhance accuracy and performance of these analyses, statistical tools and algorithms were extended from univariate case to space-time case, including space-time extensions of the Hannan-Rissanen algorithm, the bias-corrected Akaike information criterion, and the Bayesian information criterion. Methods for assessing statistical significance of model parameters are also presented. A general-purpose statistical software, called Integrated Environment for Analyzing STARMA models (IEAST), is developed for space-time analyses in this research. As an empirical example, the framework and these space-time modeling methods are then applied to investigate the spreading dynamics of West Nile virus (WNV) epidemic in crows and humans in the Detroit Metro area in 2002. Both datasets of dead crows and human cases fit

very closely to those expected from a purely STAR (Space-Time Autoregressive) process having low spatial and temporal orders. The use of the STARMA model allows estimation of the rate of spread of WNV at different spatial scales and thus characterization of the expected spatial and temporal scales. In addition, a space-time cross correlation analysis between crow and human cases is conducted. The result shows that there exists high cross correlation between dead crow and human cases at specific spatial and temporal lags. This evidence provides a foundation for the control of human WNV epidemics by using dead crows as a sensitive indicator variable. Statistical inferences from a biological point of view based on these analyses can be used to formulate the prevention and control policies for WNV. The determination of spatial-temporal autoregressive parameters using STARMA holds considerable promise for characterizing emerging infectious diseases.

To my parents and my wife

ACKNOWLEDGEMENTS

First, I would like to thank my dissertation advisor Professor Bryan K. Epperson for his enthusiasm and support during the years. It would not have been possible to complete this dissertation without Professor Epperson's valuable help. Further, Professors Joseph P. Messina, Kim T. Scribner, and Edward D. Walker are greatly acknowledged for their guidance and inspirational cooperation during my study.

I would also like to acknowledge data collation efforts of Dr. Erik Foster and Amy Williams. Dr. Jonathan S. Patterson (Diagnostic Center for Population and Animal Health, Michigan State University) performed immunohistochemistry analyses to confirm West Nile virus infection on a subsample of dead crows.

This study was supported by the Michigan Agricultural Experiment Station and the Center for Emerging Infectious Diseases, Michigan State University; and a cooperative agreement with the Centers for Disease Control and Prevention, US Public Health Services.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Space-Time Processes and Space-Time Dependency in Biology and Ecology

Spatio-temporal models have gained widespread popularity for the last decade. One major reason for this is an abundance of new challenging applications arising in the environmental sciences and epidemiology. Typical examples include forecasting of global climate change, infectious disease mapping, and their inter-relationship. Space-time datasets are usually very large and, therefore, require substantial computing power for modeling. The major improvement in computational power, especially personal computing, is another significant cause for the recent surge in using the models.

Biological and ecological data are often organized by units of time as well as by geographic locations. The processes that produce such data may have strong correlations not only in time but also in space. It is not always reasonable to study such

processes by considering space and time separately. With increasing accessibility and accuracy in remote sensing technology, large scale analyses (especially in space) of space-time data become possible. This trend highlights the importance and necessity of space-time analysis in various disciplines. In particular, some problems with high dependency in space and time, such as precipitation/temperature forecasting, pollution spreading, population succession of organisms, and epidemics, require space-time analysis and modeling to reveal or even forecast process dynamics.

In the natural world, biological or ecological systems are neither spatially nor temporally homogeneous. Rather, the processes in these systems are dynamically (and often systemically) self- or inter-correlated in space and in time. Analyses considering only time or only space may produce misleading results and not reveal the dynamical behavior of the system. In fact, processes such as epidemics, succession, competition, evolution, interactions, and population dynamics, assume that the elements of an ecosystem close to one another in space or in time are more likely to be affected by the same generating process. The lack of spatio-temporally explicit analytical framework is considered to be a major obstacle to understanding the fundamental mechanisms of such processes.

1.2 Statistical Modeling and Various Space-Time Models

In biological and ecological sciences, we usually have incomplete knowledge of the physical mechanisms responsible for the dynamics, and thus must resort to statistical estimation to represent the dynamic behavior of a system. Statistical modeling is an important procedure for analyzing stochastic systems because it allows us to connect the data to theoretical processes. Furthermore, in principle theoretical models can be used to understand, forecast, or control the behavior of a realistic ecosystem. Figure 1.1 shows how a model is used to represent the data-generating process. The data together with knowledge of the underlying process are used to identify a statistical model. While modeling, a dataset is used to fit this hypothetical model, following determination of some of the characteristic features (e.g. model types and orders) of the process. The parameters of the model are estimated during the fitting process. Next, adequacy of this statistical model can be evaluated. Finally, the statistical model can be used for forecasting, analysis, control, or other purposes. The objective is to find a "parsimonious" model with the smallest number of parameters needed to adequately fit the patterns in the data.

Many space-time models have been developed since the initial development of basic theory of space-time models in the 1970s. The following gives an introduction to some popular space-time models.

Kalman filtering has been applied in areas as diverse as aerospace, marine navigation, signal processing, demographic modeling, etc. Kalman filtering estimates the

state of a system from observations containing random errors. Due to high success, the Kalman filter quickly became an essential role in control theory. In the mid 1990s, Cressie suggested that a Kalman filter incorporating space and time can be a powerful tool for modeling space-time processes [17]. Two years later, Huang and Cressie showed the details of the space-time Kalman filter (STKF) [31]. The dynamic space-time structure is separable in space and time.

In the late 1990s, Mardia and Goodall et al. suggested an approach combining Krigging and the Kalman filter, called Krigged Kalman filter (KKF), to model space-time processes [38]. This model uses a time-varying linear combination of spatial fields to represent variations in space and time. The most important feature of this approach is that it reduces the number of parameters to a small number of key parameters. This can be a good method for forecasting; however, these parameters do not provide much information about the underlying mechanisms.

Vector Autoregressive Moving Average (VARMA) models (Lutkerpohl [36, 37]) are used to determine the dynamic interactions among multiple time series. VARMA models can be viewed as a subclass of the state-space models that are widely used in physics and engineering. VARMA processes are featured by their autoregressive and moving average orders. A p-th autoregressive order and q-th moving average order VARMA process can be expressed as

$$\mathbf{Z}_t = \sum_{k=1}^{p} \boldsymbol{\phi}_k \mathbf{Z}_{t-k} - \sum_{k=1}^{q} \boldsymbol{\theta}_k \boldsymbol{\varepsilon}_{t-k} + \boldsymbol{\varepsilon}_t \tag{1.1}$$

where $\mathbf{Z}(t) = [Z_1(t), Z_2(t), \cdots, Z_N(t)]'$ is a k-dimensional time series vector, $\boldsymbol{\varepsilon}(t)$ is

a random noise vector where $\boldsymbol{\epsilon}(t) = [\varepsilon_1(t), \varepsilon_2(t), \cdots, \varepsilon_N(t)]'$, $\boldsymbol{\phi}_k$ and $\boldsymbol{\theta}_k$ are unknown coefficient matrices to be estimated. As shown in Equation 1.1, no structured spatial dependence is defined. Instead, the spatial dependence structure is distributed among the two sets of parameters, $\boldsymbol{\phi}_k$ and $\boldsymbol{\theta}_k$. For most ecological and biological space-time variables, spatial correlation (dependence) structures usually can be regularly defined. Thus, the disadvantage of VARMA is that it does not provide a systematic structure of spatial dependence.

If processes or data reveal systematic dependencies between observations in neighborhoods, the multi-variate time series models can be refined [27]. Weight matrices can be used to reflect spatial correlation (or spatial dependence) for a given spatial configuration and can be incorporated into a VARMA model to result in Space-Time Autoregressive Moving Average (STARMA) models. In fact, STARMA models can be viewed as special cases of the VARMA models.

The extension of univariate ARMA models into the spatial-temporal domain results in a general class of models known as Space-Time AutoRegressive Moving Average (STARMA) models [12, 40]. STARMA models can be used to represent a wide range of stochastic processes that are space-time correlated. In 1980, Pfeifer and Deutsch culminated the collective efforts to extend the Box-Jenkins approach [9] for time series modeling to STARMA modeling [44, 48]. These studies also provided a computational basis for STARMA modeling and analyses.

1.3 Scope of the Work

Space-time ARMA modeling potentially has very wide applications. Its primary theory and basic modeling framework have been established since the mid 1970s. However, because of its computational complexity, paucity of appropriate algorithms and appropriate modeling framework, there have been few applications and space-time analyses in the literature. Thanks to modern computing power and the sophistication of statistical algorithms, it is now possible to represent explicitly spatio-temporal dynamics of natural processes in mathematical models and to realize the space-time model in computation.

In this thesis, a computationally efficient and widely applicable method is introduced. The difficulties encountered while applying the traditional modeling framework are reduced by using the proposed algorithms and method with the space-time extensions of the statistical criteria. The thesis will introduce basic STARMA theory, describe how the improvements can be achieved, and give examples of the statistical inferences that can be made using the proposed methods and STARMA modeling theory.

The main aim of this thesis is to introduce the improved methods, to establish general space-time ARMA models for empirical data and to show how to establish stochastic space-time models for biological or ecological processes. Because of the presence of noises and errors in natural processes and datasets, it is often not easy to identify and estimate a candidate model from a real dataset using existing methods. To solve these problems, several statistical tools and algorithms were developed in

this research for increasing model accuracy and improving computing performance.

The secondary purpose of this thesis is to make the stochastic space-time modeling theory more accessible to scientists. Since space-time modeling is a relatively unfamiliar field for most scientists, a refined STARMA modeling framework is provided for general purpose modeling. Furthermore, to support our analysis and to increase the accessibility to space-time modeling for scientists, we implemented a new method and general-purpose software called Integrated Environment for Analyzing STARMA models (or IEAST, please refer to `http://fried.for.msu.edu/~ieast`), developed explicitly for space-time analysis of processes with two-dimensional lattice data. This is the first general-purpose space-time analysis and STARMA modeling software. Since this work is one of the essential parts of the study, this software will be described and used for data analysis and modeling in this thesis.

Finally, a thorough application of STARMA modeling, including statistical space-time modeling and cross analysis of real data for the infectious disease, West Nile virus, is conducted. This application illustrates how the proposed methodology can be tailored to model the spread of infectious diseases. In addition to its potential use for short-term forecasting, this model class contributes to the understanding of the spatio-temporal evolution of disease spreading process, since it can be used to estimate how changes in spreading patterns in some specific locations are propagated to the remaining of the spatial locations.

1.4 Organization of the Thesis

This thesis consists of theoretical and application components, according to the following chapters:

- *Chapter 1 (this chapter) Introduction*: An overview of space-time processes and space-time dependence in biological and ecological sciences, space-time statistical modeling and various space-time models, and the coverage of the thesis is provided.

- *Chapter 2 Space-Time ARMA Models*: The space-time modeling in this research is based upon a well-defined model, Space-Time ARMA (STARMA). This chapter gives a historical background of the development of STARMA modeling, the general definition of STARMA, spatial correlation structures, limitations, assumptions, and conventional modeling method/statistical tools.

- *Chapter 3 Development of the Proposed Method for STARMA Modeling, and IEAST*: This chapter explores the details of both the Box-Jenkins modeling approach and the proposed modeling method. Statistical algorithms and tools for improving modeling performance are developed and introduced in this chapter. In addition, a software environment based on the developed method for modeling STARMA models, called IEAST, is implemented and introduced.

- *Chapter 4 Application to the Spreading of West Nile Virus in Detroit Metro Area*: This chapter is another important part of the thesis. Based on the two real datasets, human cases and dead crow data for WNV in Detroit, Michigan

in 2002, the dynamic behaviors of both datasets are analyzed and their space-time models are established, using the method and tools discussed in Chapter 3. Statistical inferences are given based on the analysis results. Moreover, the cross analysis between these two datasets were examined for statistical inferences on the space-time relationship between human and dead crow variables.

- *Conclusion*: Future research and limitations for this method are discussed in this chapter.

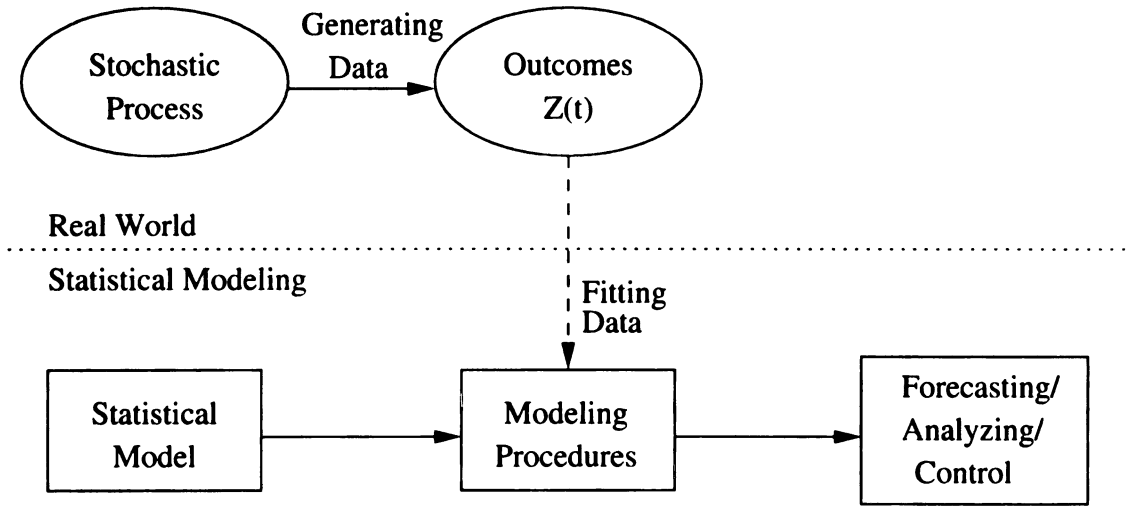

Figure 1.1. Relationship between data-generating and statistical modeling procedures

CHAPTER 2

Space-Time ARMA Models

2.1 Historical Background

Research in statistical/ecological models that describe the spatio-temporal evolution of a single variable or multi-variable relationships in space and time started in the mid 1970s. It has increased significantly during the last twenty years, because it is closely related to progress in computer technology and the prevalence of remote sensing for collecting large space-time datasets. Cliff and Ord were among the first to establish a model for interactions in space and time [14, 15], and since then several techniques have been developed corresponding to different inferential needs and data types.

The STARMA model class was first presented in the literature in the late 1970's by Bennett [6]. Since then, it has been applied to spatial time series data from a wide variety of disciplines such as river flow, traffic control, and spatial econometrics. A series of studies on the basic space-time modeling theories was completed during the beginning of 1980s. Aroian defined m-dimensional time series, derived properties, and

applied to the examples of space-time series [4]. In following papers, univariate time series autoregressive models [3], moving average models [55], and mixed models [42], i.e. autoregressive moving average models, were extended to m-dimensional cases (STARMA is a special case of m-dimensional time series). Furthermore, some of the statistical properties of these models were investigated by Pfefier and Deutsch [46, 47].

Recently, Giacomini and Granger pointed out that the STARMA class of models can be derived through a transformation of the Vector Autoregressive Moving Average (VARMA) models [27]. In fact, the transformation is a restriction related to the systematic spatial correlation structure as revealed by a set of weight matrices [27]. STARMA models can represent well most applications at large spatial scales, but for the applications involving only a few observations across space it may be too parsimonious.

For the related univariate ARMA modeling methodology, Box and Jenkins in 1970 approached the identification problem for time series by using temporal autocorrelation and partial autocorrelation analysis [9]. This method consists of identifying the characteristic behaviors of the correlograms for temporal autoregressive, moving average, and mixed processes. Martin and Oeppen then extended the identification procedure in the Box-Jenkins approach to the space-time case, also using space-time correlation functions [40]. At the same time, Bennett proposed an alternative procedure for model identification for the spatial time series [5]. Various aspects of identification have been addressed by Bennett [5] and Martin and Oeppen [40]. Pfeifer and Deutsch in 1980 provided a comparison of accuracy and precision of various estimation methods for the parameters in space-time autoregressive models, using computer

simulations [45].

The three-stage iterative model-building philosophy commonly referred to as the Box-Jenkins approach [9] for building univariate time series models has been adapted for use with STARMA models. Pfeifer and Deutsch (1980) [44, 45, 48] were among the first to develop space-time modeling techniques for lattice spaces in the context of STARMA models, and they illustrated the model-building details for the identification, estimation, and diagnostic checking of the STARMA model, using an iterative three-stage procedure.

The extension of univariate ARMA time series models into the space-time domain results in the general model class of STARMA models [12, 14]. These models apply to a single random variable observable at N fixed sites or locations in space at discrete points or periods of time, $t = 1, 2, ..., T$. They are of value for descriptive and forecasting purposes when the observed system exhibits spatial autocorrelation defined by Cliff and Ord [13]: 'If the presence of some quality in a county of a country makes it presence in neighboring counties more or less likely, we say that the phenomenon exhibits spatial autocorrelation.'

2.2 Model Definitions

STARMA is a space-time extension of the ARMA model [9]. The STARMA model class is characterized by linear dependence lagged in space as well as time. We summarize some of the results of Pfeifer [44], using essentially the same notation. STARMA expresses $Z_{i,t}$, the observation of the space-time random variable at site

i (i=1,2,...,N) and time t, as a weighted linear combination of past observations and random noise inputs, which may be lagged both in space and time. Let $\mathbf{Z}_t = [Z_{1,t}, Z_{2,t}, \cdots, Z_{N,t}]'$ be the $N \times 1$ vector of observations at time t, where N is the total number of sites in space.

The concept of a spatially lagged variable requires an ordering of the neighbors of each site. A definition of spatial order appropriate for regularly spaced systems is found in [7]. To formulate the spatial relations among sites, spatial weight matrices $\boldsymbol{W}^{(l)}$ are used. Assume weight matrix $\boldsymbol{W}^{(l)}$ has elements $w_{ij}^{(l)}$ that are the weighting contributions of site j to site i, and which are nonzero if and only if site i and j are *l-th* order neighbors in space. Then, the general STARMA model can be expressed in the following form:

$$\mathbf{Z}_t = \sum_{k=1}^{p}\sum_{l=0}^{r} \phi_{kl}\mathbf{W}^{(l)}\mathbf{Z}_{t-k} - \sum_{k=1}^{q}\sum_{l=0}^{s} \theta_{kl}\mathbf{W}^{(l)}\boldsymbol{\varepsilon}_{t-k} + \boldsymbol{\varepsilon}_t \tag{2.1}$$

where p and r are respectively the maximum autoregressive temporal and spatial orders, q and s are respectively the maximum moving average temporal and spatial orders, ϕ_{kl} and θ_{kl} are respectively the autoregressive and moving average parameters at temporal lag k and spatial lag l, $\mathbf{W}^{(l)}$ is the $N \times N$ weight matrix for spatial order l, and $\boldsymbol{\varepsilon}_t = [\varepsilon_{1,t}, \varepsilon_{2,t}, \cdots, \varepsilon_{N,t}]'$ is the random noise vector at time t.

The weights $w_{ij}^{(l)}$ should reflect an ordering or proximity of spatial neighbors. Figure 2.1 illustrates an example of a spatial order definition. The first order neighbors (corresponding to $\mathbf{W}^{(1)}$) are those that are closest to a given site (the central gray circles). The 2^{nd} order neighbors are farther away from than the 1^{st} order neighbors,

but closer than the 3^{rd} order neighbors.

There are three major model types (STAR, STMA, and mixed models) defined for general STARMA models. A process is said to be a Space-Time AutoRegressive process of temporal order p and spatial order r if $q = 0$ (named as STAR(p, r)). A STAR process can be expressed as

$$\mathbf{Z}_t = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} \mathbf{W}^{(l)} \mathbf{Z}_{t-k} + \boldsymbol{\varepsilon}_t. \tag{2.2}$$

A Space-Time Moving Average process is of temporal order q and spatial order s if $p = 0$ (named as STMA(q, s)). A STMA process can be expressed as

$$\mathbf{Z}_t = \boldsymbol{\varepsilon}_t - \sum_{k=1}^{q} \sum_{l=0}^{m_k} \theta_{kl} \mathbf{W}^{(l)} \boldsymbol{\varepsilon}_{t-k}. \tag{2.3}$$

The mixed model combines both autoregressive and moving average effects (if $p > 0$ and $q > 0$), and is named as Mixed(p, q, r, s) (its mathematical form is given by Equation 2.1). STAR and STMA model are popularly used in practice. Not only can many practical stochastic processes be simply attributed to either STAR or STMA, but also there exist primary statistical differences between them.

2.3 Spatial Correlation Structure and Weight Matrices

As shown in Equations 2.1-2.3, spatial correlation structure is specified by a set of weight matrices ($\mathbf{W}^{(l)}$). There are $max(r, s) + 1$ weight matrices, i.e. $\mathbf{W}^{(0)}, \mathbf{W}^{(1)}, \cdots, \mathbf{W}^{(max(r,s))}$, for correlation structures of each spatial lag for a Mixed(p, q, r, s) model; $r + 1$ weight matrices for STAR(p, r); $s + 1$ weight matrices for STMA(q, s). Specifically, $\mathbf{W}^{(0)}$ is a unit matrix $\mathbf{I}_N$.

If the total number of sites in space is N (i.e. the length of the observation vector $\mathbf{Z}_t$), the size of all weight matrices for the configuration is $N \times N$. The n-th row of the weight matrix for spatial lag l is the distribution of the relative contributions of the sites in the neighborhood of spatial lag l to the site n at some temporal lags. In other words, if the weights $w_{ij}^{(l)}$ are the elements in $\mathbf{W}^{(l)}$, it is nonzero only if sites i and site j are neighbors for spatial lag l with

$$\sum_{j=1}^{N} w_{ij}^{(l)} = 1 \quad \text{and} \quad w_{ij}^{(l)} \geq 0 \text{ for all } i, j, l.$$

Weights $w_{ij}^{(l)}$ for each spatial lag l can be selected to reflect physical connections and relationships among sites (or locations) of the system under study.

The weight matrices can be specified by model builders based on prior knowledge of the data observed, or can be assigned to be isotropic and generated systematically. Because weight matrices define the relationship between each site in a general sense, STARMA model can be easily adapted to various spatial dimensions (e.g.

one-dimensional application, as in modeling the dispersal of species along a river, or three-dimensional application, like modeling air conduction and circulation) or special spatial arrangements (e.g. anisotropic weighting for directional spatial correlation, like modeling spatial tendency of migration of a species).

2.4 Box-Jenkins Modeling Approach

The classic (1970s) time series analysis uses a Box-Jenkins approach which is a general procedure for modeling and forecasting stationary autoregressive and moving average processes. The main output from such an approach is a regression model explaining current values of the series in terms of past values. The coefficients in the model can then be used to forecast the series into the future.

This approach involves identifying an appropriate ARMA model, fitting it to the data, and using the fitted model for forecasting. One of the attractive features of Box-Jenkins approach for forecasting is that ARMA processes are a very rich class of possible models. It is usually possible to find a model which provides an adequate description for the data. The Box-Jenkins approach consists of iterative steps of *model identification*, *parameter estimation*, *diagnostic checking* and *forecasting* (as shown in Figure 2.2). Iterations of these steps are then used to find increasingly better solutions.

The method for STARMA modeling in our research is based on a space-time extension of the Box-Jenkins approach in [48]. In chapter 3, improvements and details of STARMA modeling methodology are described.

2.5 Space-time Autocorrelation and Partial Autocorrelation Functions

There are two questions that arise in modeling procedures. First, what type (STAR, STMA, or mixed) of the space-time model should be used? Secondly, what are the spatial and temporal orders of the model? Because there are so many potential candidates, it is advantageous to have a computationally efficient method for identifying the type and orders. As shown in Box and Jenkins [9], the popular identification method for time series is to use the estimated autocorrelation and partial autocorrelation functions. In an analogous way, the following statistical tools, space-time covariance, autocorrelation, and partial autocorrelation can provide a systematic method for revealing the characteristics (type and orders) of the underlying process for a given dataset.

The space-time covariance between l^{th} and k^{th} spatial order neighbors at time lag s can be defined as in [44]

$$\gamma_{lk}(s) = E_t\left[\frac{[\mathbf{W}^{(l)}\mathbf{Z}_t]'[\mathbf{W}^{(k)}\mathbf{Z}_{t+s}]}{N}\right]. \tag{2.4}$$

Based on the definition of space-time covariance in Equation 2.4, the Space-Time Autocorrelation Function (STACF) between l^{th} and k^{th} spatial order neighbors at time lag s can be defined as

$$\rho_{lk}(s) = \frac{\gamma_{lk}(s)}{\sqrt{\gamma_{ll}(0)\gamma_{kk}(0)}}. \tag{2.5}$$

Having the space-time autocorrelation function from Equation 2.5, to fully implement the identification stage in modeling procedure the corresponding partial autocorrelations have to be estimated. The Space-Time Partial Autocorrelation Function (STPACF) can be found in the following way. Given temporal lag m and spatial lag n, the best linear predictor of $\mathbf{Z}_t$ can be expressed as

$$\hat{\mathbf{Z}}(t)|_{m,n} = \sum_{k=1}^{m} \sum_{l=0}^{n} \phi_{kl} \mathbf{W}^{(l)} \mathbf{Z}_{t-k}. \tag{2.6}$$

For every m (=1...maximum temporal lag) and n (=0...maximum spatial lag), a set of coefficients $\xi_{mn} = \{\phi_{kl} | k = 1...m, l = 0...n\}$ can be found. For given temporal lag m and spatial lag n, the STPACF is the element ϕ_{mn} in the set ξ_{mn}. ξ_{mn} that is the solution of the space-time Yule-Walker equations for maximum temporal lag m and spatial lag n [40, 48]. In other words, the space-time partial correlation function can be found by successively fitting STAR(MaxT=m,MaxS=n) models for m =1, 2, 3, ... and for n =0, 1, 2, ... and picking out the estimates of the last coefficient from this sequence of models. Computationally, the partial autocorrelations can be obtained by solving the space-time Yule-Walker equation for each spatial order $0, 1, 2, ...$ and for each temporal order $1, 2, ...$ using the recursive method shown by Durbin [19].

2.6 Assumptions

For space-time modeling, there are some assumptions made, including spatial regularity [25], stationarity [47], invertibility [47], constant correlation structure, and

normality.

A spatial system has spatial regularity if the spatial structure is regularly defined over entire space, e.g. as in a lattice. Without spatial regularity, it is not meaningful to define spatial lag structures [25]. Thus, we generally will not obtain spatial stationarity without spatial regularity. In practice, it is extremely difficult to define a spatial lag structure so that STARMA models are stationary and correlation functions well defined without the assumption of regular distribution of locations.

The most important assumption in STARMA modeling is the validity of stationarity. A space-time process is called stationary if its statistical properties do not change with time or locations. However, it is often impractical to assume that the observation data is generated by a stationary process, especially in space. In the natural world, seasonality, trends, and clusters can cause non-stationarity. Generally, some preprocessing or transformations such as differencing, de-seasonalizing, de-trending, or logarithmic transformation must be performed in advance to obtain a stationary series before a STARMA model can be fitted. It makes no sense to try to fit a data generating from strongly non-stationary process to a stationary model like STARMA. Without stationarity or approximation of stationarity we cannot make detailed statistical inferences about the properties of the underlying processes. Furthermore, by ensuring the property of spatial replication, spatial stationarity allows us to gain more statistical power [25]. Using a stationary model to fit a process without stationarity is just like trying to characterize a system in which the characteristics of the system are constantly changing. There is no way to look inside or to characterize such a system from its inputs and outputs.

With the assumption of normality, only weak stationarity (second-order) is needed for STARMA modeling to achieve strict stationarity. Weak stationarity is defined as follows: Given a space-time process $\{\mathbf{Z}_{x,t} | \ t \in \mathbb{Z}, \ x \in \mathbb{R}^n\}$, it is weakly stationary if: 1) second moment of $\mathbf{Z}_{x,t}$ is finite for all t and x; 2) first moment of $\mathbf{Z}_{x,t}$ is independent of t and x; 3) cross moment of $\mathbf{Z}_{x,t}$ and $\mathbf{Z}_{x+s,t+h}$ depends only on s and h. Weak stationarity plus an assumption of normality are sufficient to produce strict stationarity. For a strictly stationary process, we have the same probability distribution for all time and locations. If the probability distribution is not independent of time and locations, it is difficult to define statistical properties of the process, e.g. level, variance, covariance, correlation, and even hypothesis test. Among these properties, the critical ones are correlations and hypothesis tests, which are necessary during STARMA modeling and general space-time analysis.

To be stationary and having a unique solution, as in the univariate case, the *causality* and *invertibility* conditions of the process must be satisfied. If $p = 0$ in Equation 2.1, the model is a pure STMA model (as in Equation 2.3) and $\mathbf{Z}_t$ is always stationary. When a STARMA model has p autoregressive terms ($p > 0$), $\mathbf{Z}_t$ is stationary if and only if all roots of

$$det \left[\mathbf{I} - \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} \mathbf{W}^{(l)} x^{-k} \right] = 0$$

are inside the unit circle, that is, $|x| < 1$. If $q = 0$ in Equation 2.1, the model is a pure STAR model (as in Equation 2.3) and $\mathbf{Z}_t$ is always invertible. When a STARMA

model has q moving average terms ($q > 0$), $\mathbf{Z}_t$ is invertible if and only if all roots of

$$det\left[\mathbf{I} - \sum_{k=1}^{q}\sum_{l=0}^{m_k}\theta_{kl}\mathbf{W}^{(l)}x^{-k}\right] = 0$$

are inside the unit circle, that is, $|x| < 1$ [48].

However, stationarity does not guarantee constant spatial correlation structure. Processes that can be modeled by STARMA are the stochastic processes whose properties, such as correlative structures, do not vary with location and time. Hence, we need to further assume that the spatial correlation structure is constant, not varying with either time or locations.

It is assumed that each observation of the space-time series has the same expectation function, standard deviation, and probability distribution function. We further assumed a noise error component, which is a sequence of uncorrelated random vectors with a constant distribution (Gaussian), constant variance, and zero mean. The random noise vector $\boldsymbol{\varepsilon}_t$ is normally distributed at time t and satisfies the following characteristics:

$$\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \sigma^2\mathbf{I}_N)$$

$$E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_{t+s}] = \begin{cases} \sigma^2\mathbf{I}_N, & s = 0 \\ \mathbf{0}, & \text{otherwise} \end{cases}$$

$$E[\mathbf{Z}_t\boldsymbol{\varepsilon}'_{t+s}] = \mathbf{0} \quad \text{if } s > 0.$$

This assumption of normality assures that weak stationarity is sufficient for a uniform probability distribution for all times and locations. Although technically the normality assumption is required, little to nothing is known about how strict

violations may affect inferences.

2.7 Finest Lag Structure and Selection of Spatial Correlation Structure

Although space-time ARMA appears to be a direct extension of time series ARMA, several properties of time series models do not apply to STARMA models. In particular, some important characteristics of STARMA models are only valid for the *finest lag structure* [30]. A finest spatial lag structure is a special case of correlation structures when, for each pair of site i and site j, there exists s such that the neighborhood for spatial lag s of site i contains only site j. Strictly speaking, properties such as stationarity, causality, invertibility, autocorrelation functions, and Yule-Walker equations, are only valid for finest lag structure [30]. Spatial correlation structure significantly affects the autocorrelation function, which may further influence the determination of model types and orders during model identification.

The following points should be kept in mind while defining spatial correlation structures. First, choosing a particular lag structure other than finest lag structure limits the generality of the models, so care must be taken in defining spatial correlation structure to ensure that useful models are not eliminated from consideration. Second, spatial correlation structure should be defined in a uniform manner across space to ensure that the STARMA models are stationary and the correlations are well-defined [30].

It may often be desirable to use a coarser lag structure in order to reduce the number of parameters that must be estimated. However, having a coarser lag structure restricts the generality of the class of STARMA models. Therefore, the choice of an appropriate definition of lag structure is an important first step in the process of selecting an adequate but parsimonious model. Knowledge about the process underlying the data is helpful in deciding on an appropriate lag structure or correlation structure. For instance, if the interactions among locations, for the process being studied, depend only on the distance not the direction (i.e. isotropic), then the lag structure like the following can be a reasonable choice.

$$\begin{bmatrix} & & & 3 & & & \\ & & 3 & 2 & 3 & & \\ & 3 & 2 & 1 & 2 & 3 & \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ & 3 & 2 & 1 & 2 & 3 & \\ & & 3 & 2 & 3 & & \\ & & & 3 & & & \end{bmatrix}$$

If both distance and direction influence the interactions among locations, the following definition (considering the distance and the direction in south, north, west, and east) might be desirable.

$$\begin{bmatrix} 11 & 11 & 11 & 11 & 11 & 11 & 10 \\ 12 & 7 & 7 & 7 & 7 & 6 & 10 \\ 12 & 8 & 3 & 3 & 2 & 6 & 10 \\ 12 & 8 & 4 & 0 & 2 & 6 & 10 \\ 12 & 8 & 4 & 1 & 1 & 6 & 10 \\ 12 & 8 & 4 & 5 & 5 & 5 & 10 \\ 12 & 8 & 9 & 9 & 9 & 9 & 9 \end{bmatrix}$$

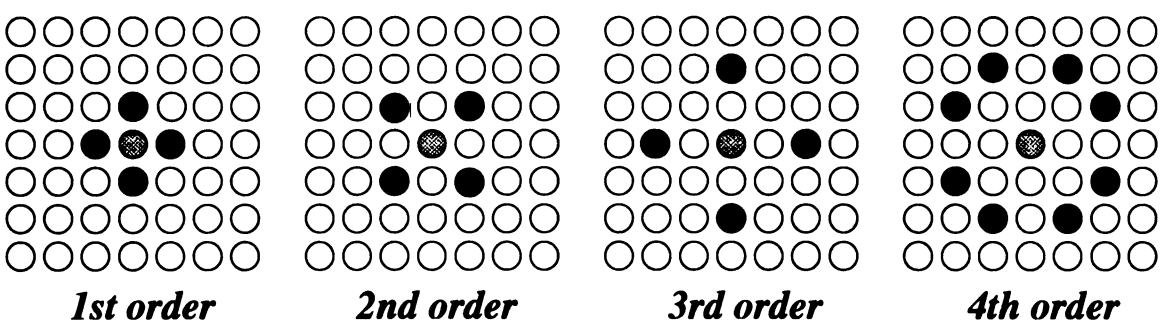

Figure 2.1. An example of the definition of spatial orders/lags in systems with two spatial dimensions.

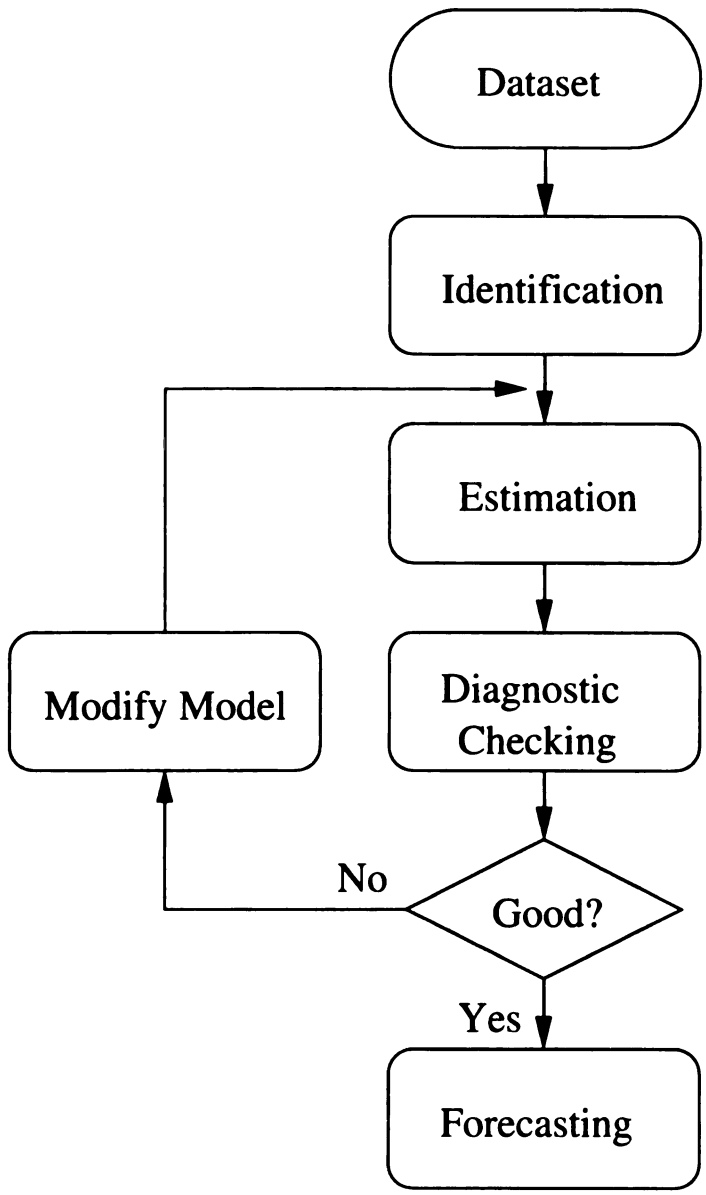

Figure 2.2. Box-Jenkins modeling approach.

CHAPTER 3

Development of the Proposed Method for STARMA Modeling, and IEAST

In STARMA modeling, there are some problems that have impeded applications since the theory was developed, including: no robust modeling method; no appropriate statistical tools/algorithms; and no flexible or general purpose software. This chapter describes the proposed STARMA modeling method with the related extension of statistical tools and algorithms. This chapter will provide an introduction to the conventional method as well as detailed description of the refined method. At the end of this chapter, the general-purpose software IEAST for STARMA modeling will also be introduced.

3.1 Specification of Spatial Correlation Structures and Derivation of Weight Matrices

The first step of STARMA modeling is the specification of spatial correlation structure (i.e. assignment of spatial weight matrices, or SWM for short). A systematic procedure used in this research to specify spatial correlation structure is introduced in this section.

The procedure for specifying spatial weight matrices consists of three steps: (1) specify Spatial Order Definition (SOD) matrix; (2) specify Spatial Relation Matrix (SRM); (3) based on (1) and (2), weight matrices (SWM) can therefore be determined and generated. In summary, the specification follows the sequence of $SOD \rightarrow SRM \rightarrow SWM$.

SOD is a matrix defining the spatial order of each 'related' neighboring cells relative to the central cell (i.e. the cell which is under consideration for spatial correlation with neighboring cells). The matrix size depends upon how large the correlated spatial neighboring area must be as needed. For example, if we define the first spatial order cells as the cells which are at a single "rook's move" away; the second spatial order cells as the cells which are at a single "bishop's move" away; and the third spatial order cells as the four cells which are exactly two cells away from the central cell in the direction of east, west, south, and north; ... and so on, up to fourth order, we can specify the SOD matrix as

$$SOD = \begin{bmatrix} - & 4 & 3 & 4 & - \\ 4 & 2 & 1 & 2 & 4 \\ 3 & 1 & 0 & 1 & 3 \\ 4 & 2 & 1 & 2 & 4 \\ - & 4 & 3 & 4 & - \end{bmatrix}. \tag{3.1}$$

The number of cells that can be defined for a specific spatial order is arbitrary. The central element of SOD represents spatial order zero, that is, it is the cell (under consideration) itself. In this example, four neighboring cells are defined in each of the first three orders of the central cell, but eight neighboring cells are defined in the fourth order. The definition of SOD needs not be symmetric, and its arrangement could depend on the spatial correlation structures of the applications.

Once the SOD is specified, the weighting distribution in a given spatial order can be defined. The weighting distributions of each spatial order are defined by a matrix called Spatial Relation Matrix (SRM). The sum of the weight distribution of every specific spatial order should be unity. The coefficients in SRM represent the importance (in the sense of spatial correlation to the central cell) of every neighboring cell in each specific spatial order. For uniform weights (isotropic), a SRM for the SOD defined above can be defined as

$$SRM = \begin{bmatrix} - & .125 & .25 & .125 & - \\ .125 & .25 & .25 & .25 & .125 \\ .25 & .25 & 0 & .25 & .25 \\ .125 & .25 & .25 & .25 & .125 \\ - & .125 & .25 & .125 & - \end{bmatrix}. \tag{3.2}$$

By comparing SOD and SRM, we can definitely associate the spatial order of any given cell with its corresponding weighting (in that spatial order) in the correlated neighborhood.

The above is an isotropic (directionless) example. We can also use a directional spatial correlation structure, i.e. anisotropic correlation structure. The following is an anisotropic example (a biased distribution towards the north in the first and second spatial orders) of SRM for the SOD given in Equation 3.1.

$$SRM = \begin{bmatrix} - & .125 & .25 & .125 & - \\ .125 & .35 & .40 & .35 & .125 \\ .25 & .25 & 0 & .25 & .25 \\ .125 & .15 & .10 & .15 & .125 \\ - & .125 & .25 & .125 & - \end{bmatrix} \tag{3.3}$$

Every two-dimensional spatial matrix can be rearranged into a linear vector (as the column vector $\mathbf{Z}_t$ in Equation 2.1) for mathematical manipulations and computations. For example, an observation in 4×4 spatial matrix can be rearranged into a linear vector $\mathbf{Z}$ as following.

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix} \rightarrow \begin{bmatrix} a \\ b \\ c \\ \vdots \\ \vdots \\ p \end{bmatrix} = \mathbf{Z} \tag{3.4}$$

Weight matrix $\mathbf{W}^{(l)}$ is actually a linear transformation matrix for spatial lag l (or spatial order l). Consider $\mathbf{W}^{(l)}\mathbf{Z}$: the observation vector $\mathbf{Z}$ is spatially-weighted and transformed into vector $\mathbf{Z}^{(l)}$ with the weighting distribution at spatial order l by multiplying $\mathbf{Z}$ with the weight matrix of spatial order l, $\mathbf{W}^{(l)}$. The sum of the transformed $\mathbf{Z}$ for all spatial orders can be written as $\sum_l \mathbf{Z}^{(l)}$.

$$\mathbf{W}^{(l)}\mathbf{Z} = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ .125 & .35 & .40 & \cdots & \cdots & .125 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \\ \vdots \\ \vdots \\ p \end{bmatrix} = \begin{bmatrix} a^* \\ b^* \\ c^* \\ \vdots \\ \vdots \\ p^* \end{bmatrix} = \mathbf{Z}^{(l)} \qquad (3.5)$$

In the above example, the n-th row of the weight matrix for spatial order l (i.e. $\mathbf{W}^{(l)}$) is a rearranged linear row vector of the relative contributions of the cells in the neighborhood of spatial order l to the cell n.

Hence, the weight matrices can be generated systematically in the following way. For the given SOD (Equation 3.1), SRM (Equation 3.3), and space dimension of 5×5, we can now construct the weight matrix row by row for the first spatial order weight matrix $\mathbf{W}^{(1)}$. Let us look at a simple example.

After removing the coefficients for the spatial orders other than the first order, the matrix SRM can be rewritten as

$$SRM = \begin{bmatrix} 0 & .40 & 0 \\ .25 & 0 & .25 \\ 0 & .10 & 0 \end{bmatrix}. \tag{3.6}$$

In order to use a weight matrix based on the above first order SRM to transform an observation (say vector $\mathbf{Z}$) for each cell (from the 1st to the 25th element) in a space with dimension of 5×5, we need to have a weight matrix $\mathbf{W}^{(1)}$ with 25 rows, and it can be obtained as follows.

To obtain the first row of $\mathbf{W}^{(1)}$, put the SRM into an empty (filled with zeros) 5×5 matrix by aligning the central cell (bold **0**s in the following equations) of SRM to the first cell of the 5×5 matrix. Everything outside the 5×5 matrix is discarded. Thus, we have

$$\text{Row 1} \rightarrow \begin{bmatrix} \mathbf{0} & .25 & 0 & 0 & 0 \\ .10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & .25 & 0 & 0 & 0 & .10 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

For the second row of $\mathbf{W}^{(1)}$, the SRM is shifted for a cell to the right of the first cell so that the central cell of SRM aligns with the second cell of the 5×5 matrix. Keep doing this until the 25th row and have

$$
\text{Row 2} \rightarrow \begin{bmatrix} .25 & \mathbf{0} & .25 & 0 & 0 \\ 0 & .10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} .25 & 0 & .25 & 0 & 0 & 0 & .10 & 0 & \cdots & 0 \end{bmatrix}
$$

$$
\text{Row 3} \rightarrow \begin{bmatrix} 0 & .25 & \mathbf{0} & .25 & 0 \\ 0 & 0 & .10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & .25 & 0 & .25 & 0 & 0 & 0 & .10 & \cdots & 0 \end{bmatrix}
$$

$$
\vdots \qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots
$$

$$
\text{Row 25} \rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .40 \\ 0 & 0 & 0 & .25 & \mathbf{0} \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & \cdots & 0 & .40 & 0 & 0 & 0 & .25 & 0 \end{bmatrix}.
$$

Integrating all these 25 rearranged row vectors above, the weight matrix $\mathbf{W}^{(1)}$ is thus obtained. The same procedure can be also applied to spatial order two, three, and so forth, to generate $\mathbf{W}^{(2)}$, $\mathbf{W}^{(3)}$, ... etc. Finally, the whole set of weight matrices $\mathbf{W}^{(l)}$ ($l = 1, 2, \cdots,$ max spatial order) for the given SOD and SRM can be generated. This procedure has be automated in IEAST.

3.2 Identification Stage

The first stage of the modeling procedure is the identification stage, in which the initial most probable underlying model type (STAR, STMA or Mixed) and its maximum spatial/temporal orders are determined. There are diagnostic differences in the STACF and STPACF among pure STAR, pure STMA, and mixed models. The model type is determined based on whether the curves of STACF or STPACF drop abruptly, or if both of them decay in an exponential manner along the time (or spatial) axis (e.g. Pfeifer and Deutsch [48]). In the following sections, we call the abrupt dropping behavior "cut-off" and the exponential decaying behavior "tail-off". Technically, we can distinguish these two behaviors by their first order derivatives of the autocorrelation functions. The derivatives of tailing-off functions generally are exponential shape, and the derivatives of cutting-off functions have abrupt change around the cut-off point. Conventionally, the method of identifying the process type was simply a qualitative comparison of observations to theoretical behaviors. In the following sections, I have developed some quantitative metrics and methods to set up a systematic identification process that significantly improves computational performance. Furthermore, the spatial and temporal order can be determined by investigating the lags (in space or in time) at which the STACF or STPACF cuts-off.

In a manner analogous to that of the univariate ARMA, the subclasses of STARMA family are characterized by their distinct STACF and STPACF behaviors. A pure STAR model exhibits a STACF that tails-off in both space and time, and STPACF that cuts-off after certain lags in space and time. In contrast, a pure

STMA model is characterized by a STACF that cuts-off after some temporal and spatial lags and a STPACF that tails-off spatially and temporally. A mixed model exhibits the spatial and temporal tailing-off of both the STACF and STPACF. Two examples shown in Figure 3.1 and 3.2 are to illustrate the cut-off/tail-off behaviors of STACF/STPACF of two typical simulated data. These data were generated by picking the last 100 generations from a total of 5000 simulated generations. The two-dimensional space was assumed to be 8×8, i.e. 64 cells in total. Uniform weight matrices (or isotropic spatial correlation structure) were used for all spatial orders (from first to fourth). The random noise vector was $\boldsymbol{\varepsilon}_t \sim N(0, \sigma^2 = 0.00125)$.

In summary, for STAR(2,1), the STPACF cuts-off at temporal lag two and at spatial lag one; STPACF tails-off. On the other hand, for STMA(1,1), the STACF cuts-off at temporal lag one and at spatial lag one and the envelope for STPACF tails-off. These two examples illustrated how the model type is selected and how the temporal/spatial orders are determined by observing the behavior of STACF and STPACF.

3.3 Estimation Stage

The second stage is to estimate the parameters in the model identified in terms of the model type and orders. In general, parameter estimation is to minimize the following sum of squared error function (or maximize the likelihood function) (Equation 3.7) to find a set of maximum likelihood estimates. The maximum likelihood function to

be optimized is

$$S(\hat{\boldsymbol{\beta}}) = \sum_{t=1}^{T} \left(\mathbf{Z}_t - \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \hat{\phi}_{kl} \mathbf{W}^{(l)} \mathbf{Z}_{t-k} + \sum_{k=1}^{q} \sum_{l=0}^{m_k} \hat{\theta}_{kl} \mathbf{W}^{(l)} \boldsymbol{\varepsilon}_{t-k} \right)^2 \tag{3.7}$$

where T is the number of observations in time, $\mathbf{Z}_t$ is the observation vector at time t, $\boldsymbol{\varepsilon}_t$ is the random error vector at time t, and $\hat{\boldsymbol{\beta}} = [\hat{\phi}_{10}, \hat{\phi}_{11}, \cdots, \hat{\phi}_{p\lambda}, \hat{\theta}_{10}, \hat{\theta}_{11}, \cdots, \hat{\theta}_{qm}]'$. This is a quadratic nonlinear optimization problem.

For linear models (i.e. STAR for STARMA case), these maximum likelihood estimates can be found by applying linear estimators. For example, in a model of STAR(maxT=2,maxS=1), all observations can be substituted into the model equation and arranged as in Equation 3.8.

$$\begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \\ \mathbf{Z}_3 \\ \vdots \\ \mathbf{Z}_T \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{Z}_1 & \mathbf{W}^{(1)}\mathbf{Z}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{Z}_2 & \mathbf{W}^{(1)}\mathbf{Z}_2 & \mathbf{Z}_1 & \mathbf{W}^{(1)}\mathbf{Z}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{Z}_{T-1} & \mathbf{W}^{(1)}\mathbf{Z}_{T-1} & \mathbf{Z}_{T-2} & \mathbf{W}^{(1)}\mathbf{Z}_{T-2} \end{bmatrix} \begin{bmatrix} \phi_{10} \\ \phi_{11} \\ \phi_{20} \\ \phi_{21} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \boldsymbol{\varepsilon}_3 \\ \vdots \\ \boldsymbol{\varepsilon}_T \end{bmatrix} \tag{3.8}$$

The equation above can be simply expressed in linear model form as $\mathbf{Z} = \mathbf{X}\boldsymbol{\Phi} + \mathbf{E}$, and the best linear unbiased estimator is $\hat{\boldsymbol{\Phi}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}$, where the observation vector $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \cdots, \mathbf{Z}_T]'$, model parameter vector $\boldsymbol{\Phi} = [\phi_{10}, \phi_{11}, \phi_{20}, \phi_{21}]'$, error vector $\mathbf{E} = [\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_T]'$, and $\mathbf{X}$ is the large matrix in Equation 3.8 [48].

Because of the nonlinear nature of STMA and mixed models, the linear estimator above is not appropriate. In this research, the parameters for nonlinear processes (i.e.

STMA and mixed) are estimated by maximizing Equation 3.7 and searching through a quadratic surface. A quadratic optimization algorithm, Marquardt's algorithm [39], is used to maximize the likelihood function (or minimizing sum of least squared errors) and to find the parameter estimates. Marquardt's algorithm is a Gauss-Newton based algorithm for least squares optimization. However, as one may commonly encounter in such optimization problems, it is important to locate an appropriate starting point for the optimization process. This is especially critical for multi-variate nonlinear optimization problems. During this research, it was found that the optimization process for parameter estimates, in most cases, either converges to local optima or does not converge. Thus, to avoid reaching a local maxima and to reduce the number of iterations needed during the optimization process of the maximum likelihood function, a preliminary stage, *pre-estimation*, is implemented to calculate an appropriate starting point for Marquardt's algorithm. In this study, I have extended and implemented the Hannan-Rissanen algorithm from univariate case [29] to space-time case for calculating the initial points for STMA and mixed processes as shown below.

The space-time extension of Hannan-Rissanen algorithm has three iterative steps.

1. A high order STAR model is fitted to the data using the space-time Yule-Walker equation. Then, we have the following approximate model

$$\mathbf{Z}_t^* = \sum_{k=1}^{u} \sum_{l=0}^{v} \hat{\eta}_{kl} \mathbf{W}^{(l)} \mathbf{Z}_{t-k}^* + \boldsymbol{\epsilon}_t^*$$

   where $\{\hat{\eta}_{kl} | k = 1 \cdots u, l = 0 \cdots v\}$ are the Yule-Walker estimates.

2. The estimated random noise vectors can be computed as

$$\hat{\boldsymbol{\varepsilon}}_t \equiv \boldsymbol{\varepsilon}_t^* = \mathbf{Z}_t^* - \sum_{k=1}^{u} \sum_{l=0}^{v} \hat{\eta}_{kl} \mathbf{W}^{(l)} \mathbf{Z}_{t-k}^*. \tag{3.9}$$

3. Once the estimated random noise vectors $\hat{\boldsymbol{\varepsilon}}_t, t = m+1, \cdots, T$ have been found from Equation 3.9, pre-estimates of the model parameters, $\hat{\boldsymbol{\alpha}} = [\hat{\boldsymbol{\varphi}}', \hat{\boldsymbol{\psi}}']'$ are determined by least squares linear regression of $\mathbf{Z}_t$ onto the space $\{\mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \cdots, \mathbf{Z}_{t-p}, \hat{\boldsymbol{\varepsilon}}_{t-1}, \hat{\boldsymbol{\varepsilon}}_{t-2}, \cdots, \hat{\boldsymbol{\varepsilon}}_{t-q}\}, t = m+1, \cdots, T$. By minimizing the sum of square errors

$$S(\hat{\boldsymbol{\alpha}}) = \sum_{t=m+1}^{T} \left(\mathbf{Z}_t - \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \hat{\varphi}_{kl} \mathbf{W}^{(l)} \mathbf{Z}_{t-k} + \sum_{k=1}^{q} \sum_{l=0}^{m_k} \hat{\psi}_{kl} \mathbf{W}^{(l)} \hat{\boldsymbol{\varepsilon}}_{t-k} \right)^2$$

with respect to $\hat{\boldsymbol{\alpha}}$, we can obtain the space-time extension of the Hannan-Rissanen estimator

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Omega}$$

where

$$\hat{\boldsymbol{\alpha}} = [\hat{\varphi}_{10}, \hat{\varphi}_{11}, \cdots, \hat{\varphi}_{kl}, \hat{\psi}_{10}, \hat{\psi}_{11}, \cdots, \hat{\psi}_{kl}]',$$

$$\boldsymbol{\Omega} = [\mathbf{Z}_{m+1}, \mathbf{Z}_{m+2}, \cdots, \mathbf{Z}_T]',$$

$$\mathbf{X} = \begin{bmatrix}
\mathbf{Z}_m & \mathbf{W}^{(1)}\mathbf{Z}_m & \cdots & \mathbf{W}^{(r)}\mathbf{Z}_m & \mathbf{Z}_{m-1} & \cdots & \mathbf{W}^{(r)}\mathbf{Z}_{m-1} \\
\mathbf{Z}_{m+1} & \mathbf{W}^{(1)}\mathbf{Z}_{m+1} & \cdots & \mathbf{W}^{(r)}\mathbf{Z}_{m+1} & \mathbf{Z}_m & \cdots & \mathbf{W}^{(r)}\mathbf{Z}_m \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\mathbf{Z}_{T-1} & \mathbf{W}^{(1)}\mathbf{Z}_{T-1} & \cdots & \mathbf{W}^{(r)}\mathbf{Z}_{T-1} & \mathbf{Z}_{T-2} & \cdots & \mathbf{W}^{(r)}\mathbf{Z}_{T-2}
\end{bmatrix}$$

$$\vdots \quad \vdots \quad \vdots$$

$$\vdots \quad \vdots \quad \vdots$$

$$\vdots \quad \vdots \quad \vdots$$

$$\begin{bmatrix}
\hat{\boldsymbol{\varepsilon}}_m & \mathbf{W}^{(1)}\hat{\boldsymbol{\varepsilon}}_m & \cdots & \mathbf{W}^{(r)}\hat{\boldsymbol{\varepsilon}}_m & \hat{\boldsymbol{\varepsilon}}_{m-1} & \cdots & \mathbf{W}^{(r)}\hat{\boldsymbol{\varepsilon}}_{m-1} & \cdots \\
\hat{\boldsymbol{\varepsilon}}_{m+1} & \mathbf{W}^{(1)}\hat{\boldsymbol{\varepsilon}}_{m+1} & \cdots & \mathbf{W}^{(r)}\hat{\boldsymbol{\varepsilon}}_{m+1} & \hat{\boldsymbol{\varepsilon}}_m & \cdots & \mathbf{W}^{(r)}\hat{\boldsymbol{\varepsilon}}_m & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\hat{\boldsymbol{\varepsilon}}_{T-1} & \mathbf{W}^{(1)}\hat{\boldsymbol{\varepsilon}}_{T-1} & \cdots & \mathbf{W}^{(r)}\hat{\boldsymbol{\varepsilon}}_{T-1} & \hat{\boldsymbol{\varepsilon}}_{T-2} & \cdots & \mathbf{W}^{(r)}\hat{\boldsymbol{\varepsilon}}_{T-2} & \cdots
\end{bmatrix}.$$

From simulation results, the space-time extension of Hannan-Rissanen algorithm is very efficient and accurate. The deviations of the parameter pre-estimates $\hat{\boldsymbol{\alpha}}$ from the correct values are generally lower than 10%. This algorithm greatly reduces the possibility of (Marquardt's algorithm) converging to a wrong local optima and hence improves the robustness of the modeling method.

Using the pre-estimate $\hat{\boldsymbol{\alpha}}$ as the initial point for the optimization, the final least squares estimates of the model, $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\Theta}}$, are obtained by finding the best $\hat{\boldsymbol{\beta}} = [\hat{\boldsymbol{\Phi}}', \hat{\boldsymbol{\Theta}}']'$ to minimize Equation 3.7. The maximum likelihood estimate of the residual variance $\hat{\sigma}^2$ is thus

$$\hat{\sigma}^2 = \frac{S(\hat{\boldsymbol{\Phi}}, \hat{\boldsymbol{\Theta}})}{TN} \tag{3.10}$$

where T is the number of points in time of the observations, N is the number of sites in each observation, $S(\hat{\mathbf{\Phi}}, \hat{\mathbf{\Theta}})$ is the residual's sum of square errors, $\hat{\mathbf{\Phi}}$ is a vector containing the estimates for parameters in autoregressive terms, and $\hat{\mathbf{\Theta}}$ is a vector containing the estimates for parameters in moving average terms.

3.4 Diagnostic Checking Stage

Once the model type, orders, and parameter estimates are determined, the acquired model can be used to forecast or to analyze the future behavior of the system. However, to prevent unacceptable forecasting errors, further diagnostic checks must be made to ensure the selected model is appropriate. The methods for checking the adequacy of the models include: residuals' autocorrelation analyses; residuals' randomness testing, residuals' variance checking [48, 49], and statistical significance testing of parameter estimates [48]. In this study, residuals' space-time autocorrelation analyses, significance testing of parameters, measures for the model's adequacy, and residuals' variance checking are used for diagnostic checking.

Regarding the residuals of the STACF and STPACF, inadequacy of a candidate model can be revealed by the presence of some significant correlation values. If a candidate model adequately represents the data, residuals should be white(uncorrelated in space and time) noise. In other words, all autocorrelations (in space and in time) at nonzero lags are equal to zero (or approximately equal to zero, in practice).

A candidate model is also checked for any unduly complexity, i.e. having statistically insignificant parameters. Thus, statistical significance of model parameters

is tested for this purpose using the hypotheses that a parameter of interest $\hat{\beta}_k$ is zero while keeping other parameters unrestricted. That is, let $\hat{\boldsymbol{\beta}}^* = [\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k = 0, \hat{\beta_{k+1}}, \cdots]'$. The test for this hypothesis is based on the statistic

$$\frac{[TN - K(\hat{\boldsymbol{\Phi}}) - K(\hat{\boldsymbol{\Theta}})][S(\hat{\boldsymbol{\beta}}^*) - S(\hat{\boldsymbol{\beta}})]}{S(\hat{\boldsymbol{\beta}})} \sim F_{1,TN-K(\hat{\boldsymbol{\Phi}})-K(\hat{\boldsymbol{\Theta}})} \tag{3.11}$$

where T is the number of points in time in the dataset, N is the total number of sites in the dataset, $S(\hat{\boldsymbol{\beta}}^*)$ is the sum of square residuals for parameter $\hat{\boldsymbol{\beta}}$ with the restriction $\hat{\beta}_k = 0$, $S(\hat{\boldsymbol{\beta}})$ is the sum of square residuals for parameter $\hat{\boldsymbol{\beta}}$ without restriction, $\hat{\boldsymbol{\Phi}}$ is a vector containing the estimated parameters for autoregressive terms in column vector form, $\hat{\boldsymbol{\Theta}}$ is a vector containing the estimated parameters for moving average terms in column vector form, $K(\hat{\boldsymbol{\Phi}})$ is the number of elements in $\hat{\boldsymbol{\Phi}}$, $K(\hat{\boldsymbol{\Theta}})$ is the number of elements in $\hat{\boldsymbol{\Theta}}$. Any estimated parameter shown to be statistically insignificant should be removed from the model, and the number of spatial or temporal orders may be reduced, giving a simpler model that is then considered as the next candidate model in an iterative modeling procedure.

The results of autocorrelation analysis may be obscure for model identification, and high order models are not necessarily desirable although the residuals' variances are small. As shown in Box and Jenkins [9], an autoregressive model with invertibility can expressed as an infinite sum of moving average terms (or with very high maximum orders); and a moving average model with causality can be expressed as an infinite sum of autoregressive terms (or with very high maximum orders). In practice, we need to consider another model type (e.g. STMA) if the selection of current model

type (e.g. STAR) causes very high orders in space or in time. To summarize the model fitness and the model parsimony, we need to introduce some control criteria to provide measures for evaluating candidate models, as well as to penalize the excessive use of extra parameters in the fitting of models.

I extended the space-time extension of bias-corrected Akaike's information criterion [1], referred to as the AICC, to be the criterion in determining the model type, orders and checking adequacy of parameters. It can be derived directly from Hurvich and Tsai [32] by using the likelihood function for STARMA [48]

$$L(\boldsymbol{\beta}, \sigma^2|\mathbf{Z}) = (2\pi\sigma^2)^{-\frac{TN}{2}} exp(-\frac{S(\boldsymbol{\beta})}{2\sigma^2}) \tag{3.12}$$

and obtained as

$$AICC_{STARMA} = TNlog(2\pi\hat{\sigma}^2) + \frac{S(\hat{\boldsymbol{\beta}})}{\hat{\sigma}^2} + \frac{TN[K(\hat{\boldsymbol{\Phi}}) + K(\hat{\boldsymbol{\Theta}})]}{TN - K(\hat{\boldsymbol{\Phi}}) - K(\hat{\boldsymbol{\Theta}}) - 1}. \tag{3.13}$$

In the definition of AICC, the last term is the penalty for high order models, which makes the value of AICC higher than that of lower order model. Thus, while determining an adequate model, we are searching for a model (given type, orders, and parameters) with a minimum value of AICC.

However, because AICC has a tendency of overfitting autoregressions [33, 51], a second criterion, Bayesian Information Criterion (BIC), is suggested to correct this

tendency [2, 50]. The BIC is also extended to the space-time case as

$$BIC_{STARMA} = TN log(2\pi\hat{\sigma}^2) + \frac{S(\hat{\boldsymbol{\beta}})}{\hat{\sigma}^2} + (K(\hat{\boldsymbol{\Phi}}) + K(\hat{\boldsymbol{\Theta}})) \cdot log(TN). \qquad (3.14)$$

We illustrate the performance of the model type selection with AICC and BIC, by using 150 Monte Carlo simulated datasets based on STAR, STMA and Mixed processes (50 datasets for each subclass). In this illustration, we used each of the four criteria: AICC only; BIC only; -AICC*BIC; and variance of residuals. The parameters of each simulated model were randomly chosen, with higher weights on the parameters having lower spatial or temporal orders. As expected, using AICC or variance as the only criterion has a bias toward indicating (overparameterized) mixed model (Table 3.1).

The AICC is useful when the candidates are only STAR and STMA (excluding mixed model). If we would like to select one out of these three subclasses, BIC performs much better than AICC. As is shown in Table 3.1, all of the 50 simulated STAR datasets were correctly classified as STAR, i.e. 100% accuracy. If the given simulated datasets are STMA and STARMA, the accuracies are 86% and 66%, respectively. Furthermore, we found that using the product of AICC and BIC as the criterion can provide even better performance (Table 3.1). The accuracy of model type determination for all subclasses can be as high as 78%. In summary, these methods provide a reliable quantitative measure for determining model choice, whereas previous research relied on subjective qualitative behavior of the STACF and STPACF.

With regard to variance checking, it has been shown by Pfeifer [44] that if a

Table 3.1. Accuracies of model type selection using variances, AICC, BIC, and -AICC*BIC based on 150 Monte Carlo simulated datasets.

Using variances of residuals

Simulated datasets based on/	Datasets identified as STAR	STMA	Mixed
STAR	**4%**	0%	96%
STMA	4%	**6%**	90%
Mixed	8%	2%	**90%**

Using AICC

Simulated datasets based on/	Datasets identified as STAR	STMA	Mixed
STAR	**16%**	0%	84%
STMA	4%	**6%**	90%
Mixed	8%	2%	**90%**

Using BIC

Simulated datasets based on/	Datasets identified as STAR	STMA	Mixed
STAR	**100%**	0%	0%
STMA	4%	**86%**	10%
Mixed	18%	16%	**66%**

*Using -AICC*BIC*

Simulated datasets based on/	Datasets identified as STAR	STMA	Mixed
STAR	**100%**	0%	0%
STMA	4%	**78%**	18%
Mixed	16%	4%	**80%**

process is pure white noise, then

$$variance(\hat{\rho}_{l0}(s)) \approx \frac{1}{N(T-s)} \tag{3.15}$$

where $\hat{\rho}_{l0}(s)$ is the sample STACF as defined in Equation 2.5. That is, if the estimated model appropriately represents the dataset, the variance of the sample STACF of residuals should satisfy Equation 3.15.

3.5 A Refined STARMA Modeling Procedure

Combining the criteria and algorithms developed, a STARMA modeling procedure was developed and used for analyses in this research, as shown in the flowchart in Figure 3.3. The intention of the development of this refined modeling procedure has two purposes. The first is to reduce the computing time for optimization in the estimation stage, and to increase reliability during parameter estimation, especially for nonlinear models (STMA and mixed). The second is to overcome the difficulty in model identification resulting from the ambiguous behaviors of some STACF/STPACF. These two advances greatly promote the feasibility and robustness for automating the STARMA modeling procedure computationally.

This refined procedure is similar to the three-stage iterative procedure in [48], in which STACF and STPACF are used to identify model types/orders and ordinary least square is used for parameter estimation. For our approach, in those cases where the model type cannot be qualitatively identified by STACF/STPACF, the

criteria AICC and BIC are used to iteratively determine the best candidate model. In parameter estimation, the space-time extension of Hannan-Rissanen algorithm is implemented to generate the starting point for optimization process. For the first diagnostic test, every parameter's statistical significance is evaluated, in order to decide whether the current numbers of maximum temporal and spatial orders need to be reduced, increased, or neither. Finally, the residuals' STACF/STPACF are calculated to ensure that no significant spatial/temporal correlation exists in the residuals. If there are correlations, another iteration is needed for further refining the candidate model.

3.6 Cross Correlations and Space-Time Regression Models

Cross correlation is a standard method for estimating the degree to which two series are correlated. It is robust to noise and used widely in many disciplines for feature/signal detection and image/data processing. To further extend the usefulness of the research, cross correlation analyses among the two space-time variables are conducted. Cross correlation analysis is implemented in IEAST and also applied to the analysis of the epidemics of West Nile virus. This is especially useful when investigating space-time relation between two variables. It does not necessarily give us the direction of causation flow, but, at the least, we can gain more information of how (and to what degree) variables are similar to each other.

Suppose we have two (correlated) space-time variables $\mathbf{Z}$ and $\mathbf{Y}$. $\mathbf{Z}$ is dependent on an independent variable $\mathbf{Y}$ based on the Space-Time Regression model [40]

$$\mathbf{Z}_t = \sum_{k=1}^{d} \sum_{l=0}^{c} \omega_{kl} \mathbf{W}^{(l)} \mathbf{Y}_{t-k} + \varepsilon_t, \tag{3.16}$$

where $\mathbf{Z}_t$ is expressed as a weighted sum of $\mathbf{Y}_t$ lagged in space and time, with added space-time noise ε_t. In simple words, the variable $\mathbf{Y}$ is the cause of the variable $\mathbf{Z}$. The cross correlation of $\mathbf{Z}$ and $\mathbf{Y}$ is defined in [40] and transformed into a format as below

$$\rho_{\mathbf{Z},\mathbf{Y}}(s,k) = \frac{E[\mathbf{Z}_t' \mathbf{W}^{(s)} \mathbf{Y}_{t-k}]}{\sqrt{E[\mathbf{Z}_t' \mathbf{Z}_t] \cdot E[(\mathbf{W}^{(s)} \mathbf{Y}_t)' \mathbf{W}^{(s)} \mathbf{Y}_t]}}. \tag{3.17}$$

where s is the spatial lag and k is the temporal lag between the two series.

In the cross correlation function, the two variables ($\mathbf{Z}$ and $\mathbf{Y}$) are ordered, that is, $\rho_{\mathbf{Z},\mathbf{Y}}(s,k) \neq \rho_{\mathbf{Y},\mathbf{Z}}(s,k)$. $\rho_{\mathbf{Z},\mathbf{Y}}(s,k)$ gives the cross correlation of the major variable $\mathbf{Z}$ and the secondary variable $\mathbf{Y}$, and shows the correlation values for various spatial and temporal lags ($\mathbf{Z}$ leading $\mathbf{Y}$ for some positive temporal lag). The differences or lags in space and time of the peak (of the cross correlation function) from zero are the distances that the two variables separate in space and time. In the analysis conducted for the spread of West Nile virus, these lags indicate the time delay and any spatial lags between the human cases and the dead crows findings. The cross correlation is based on an assumption of space-time regression model (Equation 3.16). The linear relationship between variable $\mathbf{Z}$ and $\mathbf{Y}$ is shown in Figure 3.4. It is evident that the

autocorrelation of $\mathbf{Y}$ also contributes to the cross correlation $\rho_{\mathbf{Z},\mathbf{Y}}(s,k)$ because

$$\begin{aligned}\rho_{\mathbf{Z},\mathbf{Y}} &= SpaceTimeCorr(\mathbf{Z},\mathbf{Y}) \\ &= SpaceTimeCorr(STR(\mathbf{Y}),\mathbf{Y}) \\ &= SpaceTimeCorr(\textstyle\sum_{k=1}^{d}\sum_{l=0}^{c}\omega_{kl}\mathbf{W}^{(l)}\mathbf{Y}_{t-k},\mathbf{Y}).\end{aligned}$$

The cross correlation can be spurious. When calculating the cross correlation of two variables, this will inflate the cross correlation values and further make it difficult to achieve a parsimonious model. One solution to this problem is to use *pre-whitening* to eliminate the autocorrelated components in both variables. To calculate $\rho_{\mathbf{Z},\mathbf{Y}}(s,k)$ with pre-whitening, we can first fit a STARMA model for the input data ($\mathbf{Y}$) to reduce sufficiently the residuals to white noise; then, filter the input data with this model to obtain the white noise residuals, $\mathbf{Y}^*$. Subsequently, we can filter the response data ($\mathbf{Z}$) with the same STARMA model, and finally, calculate the cross correlation between the filtered response $\mathbf{Z}^*$ and the filtered input $\mathbf{Y}^*$, that is, $\rho_{\mathbf{Z}^*,\mathbf{Y}^*}(s,k)$.

3.7 The Development of IEAST

The integrated statistical computing environment, IEAST (Integrated Environment for Analyzing STARMA models), was designed for analyzing and modeling stochastic processes based on general Space-Time AutoRegressive Moving Average (STARMA) models in two-dimensional gridded space. The current version only supports spatial dimension $N \times N$, that is, where the numbers of cells in the X and Y directions are equal. IEAST takes data files (**.dat**), spatial weight files (**.wet**), program files

(**.pgm**), and spatial/temporal configurations (in the *Setup* menu) as inputs, then according to program or user selection, outputs plots, text numerical analysis results, and model equations for STARMA analysis (see Figure 3.5).

This system was developed on RedHat Linux 9.0 using GNU OCTAVE v2.1.40 as the programming language, which is a high-level language primarily intended for numerical matrix computations. OCTAVE provides a convenient command line interface for numerically solving linear and nonlinear problems, and for performing other numerical operations. OCTAVE is largely compatible with the commercial software MATLAB® and free of charge. Because of source code compatibility, IEAST can be easily run under Windows/Unix/MacOS without modifications.

To provide flexibility, there are two user interfaces provided in IEAST: the **menu-driven mode** and the **interpreter mode** (i.e. IEAST programming mode). In the menu-driven mode, users can carry out the modeling or analyzing procedure by selecting among a series of hierarchical menu commands. Users must control all of the flow of modeling procedures by themselves. This is convenient, but inefficient in terms of time spent. With the programming mode, a simple and powerful IEAST programming language is provided. Highly integrated instructions are provided for users to compose STARMA modeling or space-time analysis procedures. Users can design an efficient and autonomous modeling procedure for specific applications by simply combining instructions of core functions and flow controls. It is commonly possible to implement an entire STARMA modeling procedure in less than 20 lines of code. This programmability is especially useful when the iterative procedure is necessary for analysis (as in the modeling procedure shown in Figure 3.3). Figure 3.3

is a suggested general modeling procedure for using IEAST.

This system is designed for two-dimensional STARMA analysis. However, by carefully designing the spatial weight matrices, this system can be adapted to one-dimensional or even three-dimensional systems. If the computer memory and computing power are sufficient, there are no limitations to the spatial dimensions of the dataset that can be explored. However, if the spatial dimension is too small, boundary effects would be significant and should be taken into consideration. After trading-off the computing power and practical needs, the largest spatial and temporal orders of STARMA models in this implementation are limited.

IEAST consists of three major parts: menu-driven user interface, interpreter, and core functions. Figure 3.6 illustrates the logical structure of IEAST. The core functions provide a complete set of functionalities for STARMA modeling and space-time analysis, and they are also the basis for the menu-driven interface and interpreter. The core functions are categorized as below:

- *System setup*: These functions are for configuring a given space-time system, and include loading space-time data, importing datasets from various formats, generating datasets from simulations, loading/setting up spatial correlation structures, and 2D/3D visualization of space-time datasets in time/frequency domains.

- *Preprocessing*: To generalize STARMA models, there are provided various transformations (or preprocessing) for datasets to be transformed into stationary ones before analysis. These functions include de-seasonalizing, de-trending,

differencing, subsequencing/resampling, smoothing, recovering missing data, pre-whitening with STARMA, and the Box-Cox transformation.

- *Correlation Analysis*: Space-time auto-correlation (single variable) and cross-correlation (between two variables) analyses, and plotting ability are provided.

- *Model Identification*: This contains functions for identifying model type and maximum spatial/temporal orders based on the behaviors of STACF and STPACF.

- *Parameter Estimation*: For linear processes (STAR), this function simply provides best linear estimators for model parameters. For nonlinear processes (STMA and mixed), two steps are needed. The first step calculates a starting point using the space-time extended Hannan-Rissanen algorithm. Then this point is used as an initial point for the multi-variate nonlinear optimization process for convergence of the estimates of parameters of models (Marquardt's algorithm). In addition, several mechanisms for estimating spatial correlation structures based on a given dataset and a model (with known type and orders) are provided.

- *Diagnostic Analysis*: Diagnostic functions provide various criteria for evaluating the adequacy of the obtained model. In our system, the residual's STACF and STPACF, variance analysis, AICC/BIC, and parameter significance analysis are used.

The functionalities of menu-driven interface for analysis are distributed in the

menu hierarchy. In this interface, users conduct step-by-step set up and analyze the dataset through the menus. follow their own modeling procedure. The first two levels of the menu system are shown in Figure 3.7. The current version of IEAST (version 1.30.01) has the following features:

- A true spatial and temporal analysis software
- Full configurability - spatial and temporal correlation structures can be specifically designed by users
- Interactive programming environment - a script interpreter with highly integrated instructions is provided to simplify programming efforts and to improve efficiency of analysis and modeling
- Estimation of spatial weighting structure - The spatial weighting structure can be retrieved and estimated from given datasets.
- Flexibility of defining spatial structure - a three-step definition for spatial structure makes IEAST very flexible in defining spatial structures, i.e. $SOD \rightarrow SRM \rightarrow SWM$
- Improved estimation algorithm - the introduction of space-time extension of Hannan-Rissanen algorithm greatly reduces the estimation time and possibility of converging to wrong results
- Improved diagnostic measures - the space-time AIC/BIC and statistical significance for each parameter provide good criteria for the selection of model type and orders, which is especially useful in the programming environment

- Cross correlation analysis - space-time cross correlation between two variables is provided

- 2D/3D plotting abilities - both data and analysis results can be plotted in two- or three-dimensional graphs

For further information and usage, please refer to the IEAST manual, which can be downloaded from the website `http://fried.for.msu.edu/~ieast`.

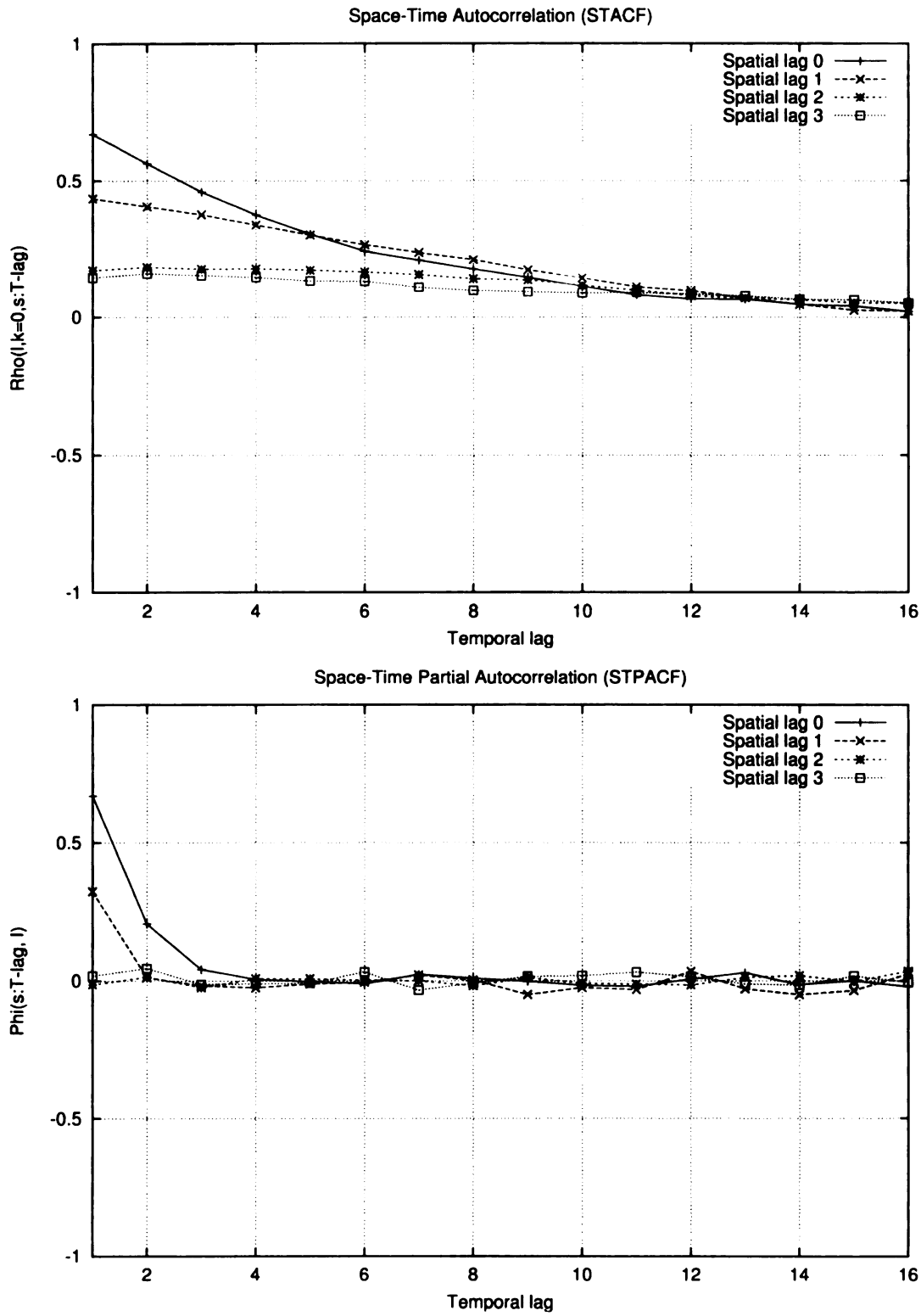

Figure 3.1. STACF and STPACF for a simulated pure STAR dataset with maximum temporal order 2 and maximum spatial order 1: $\mathbf{Z}_t = 0.5\mathbf{Z}_{t-1} + 0.3\mathbf{W}^{(1)}\mathbf{Z}_{t-1} + 0.15\mathbf{Z}_{t-2} + \boldsymbol{\varepsilon}_t$.

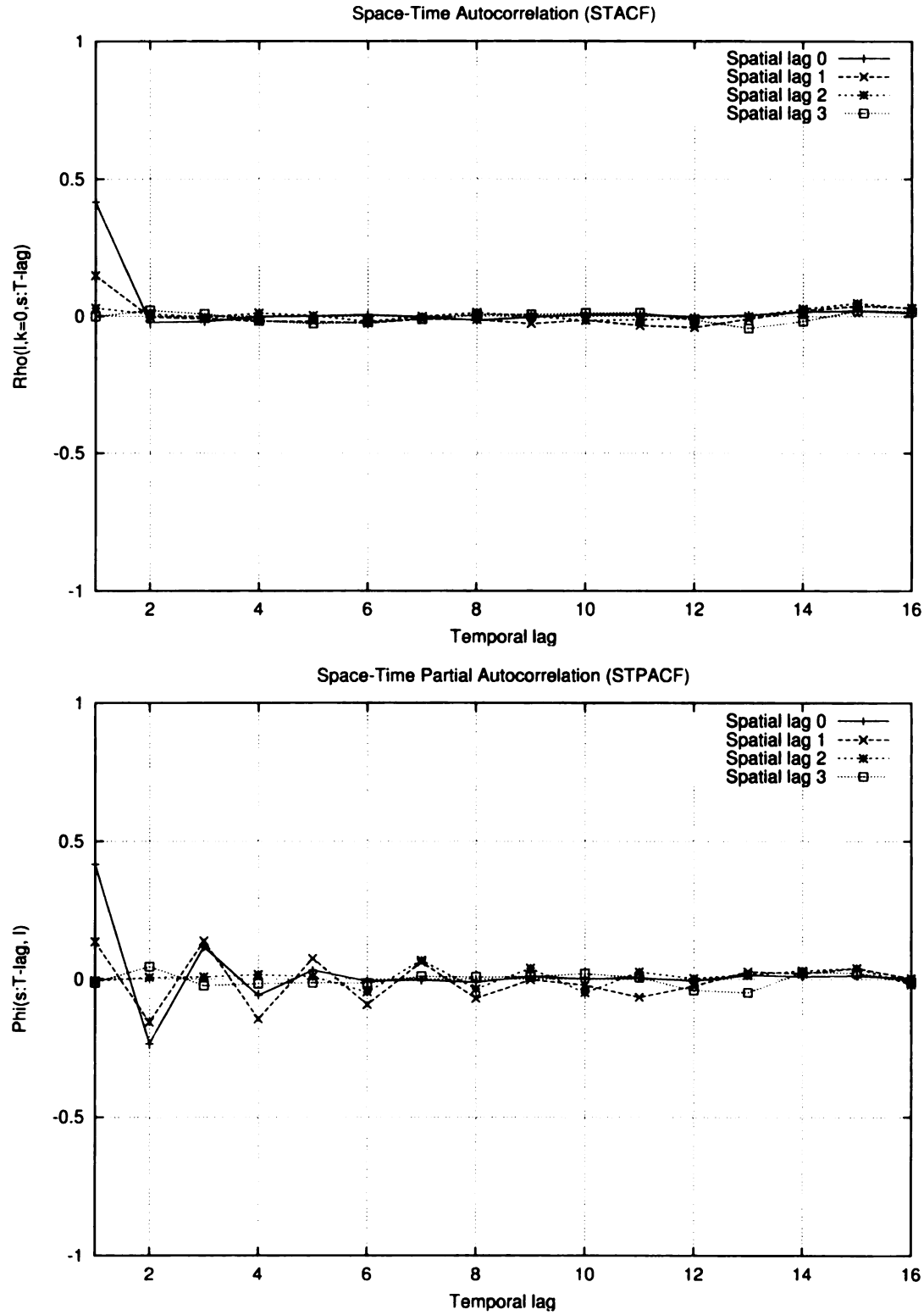

Figure 3.2. STACF and STPACF for a pure STMA dataset with maximum temporal order 1 and maximum spatial order 1: $\mathbf{Z}_t = \boldsymbol{\epsilon}_t - (-0.6)\boldsymbol{\epsilon}_{t-1} - (-0.4)\mathbf{W}^{(1)}\boldsymbol{\epsilon}_{t-1}$.

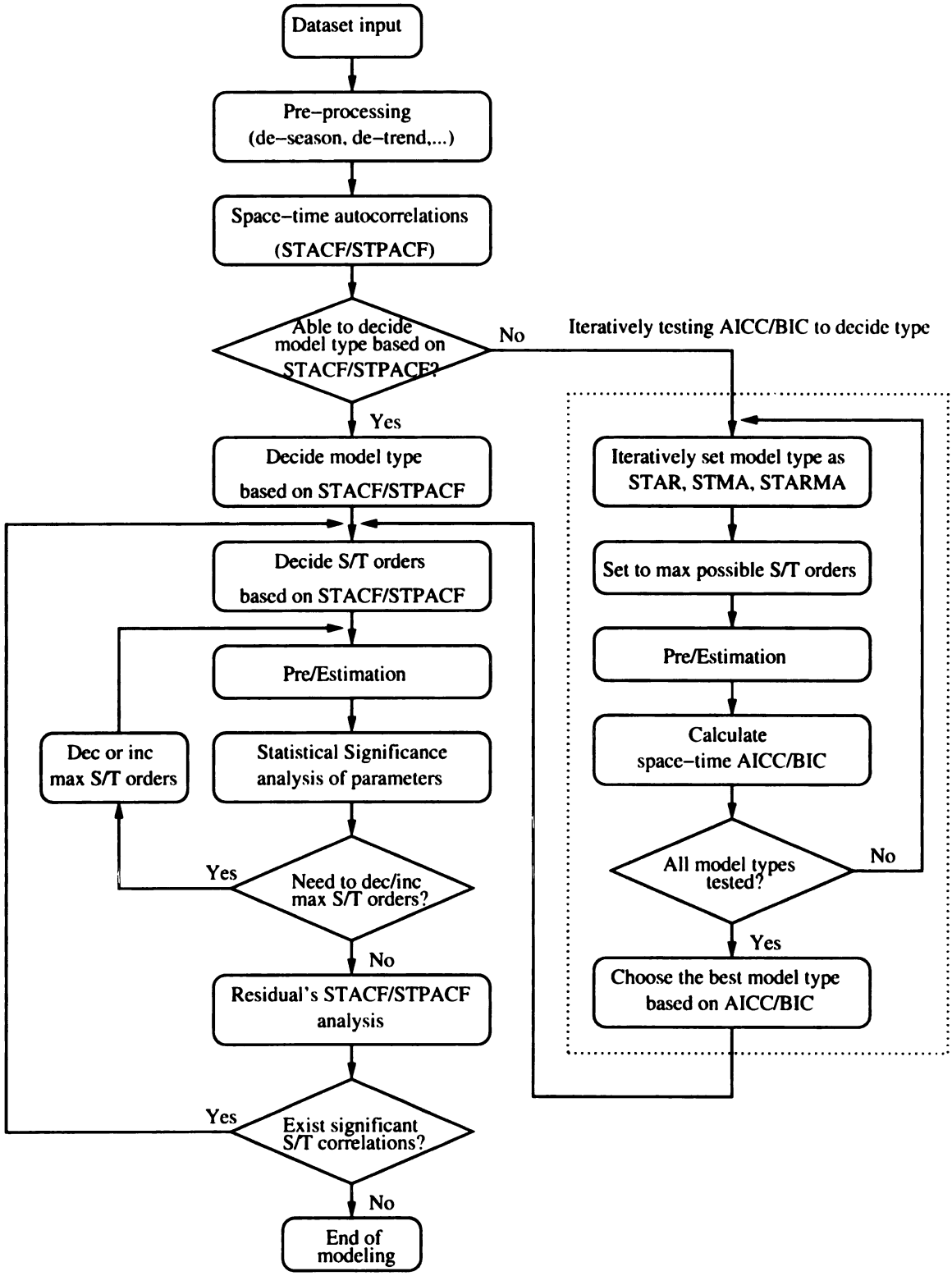

Figure 3.3. The refined STARMA modeling procedure. The area enclosed by the dotted line is for the case where model type cannot be decided from the STACF/STPACF. AICC/BIC of every possible model type is evaluated and compared, in order to select the best one.

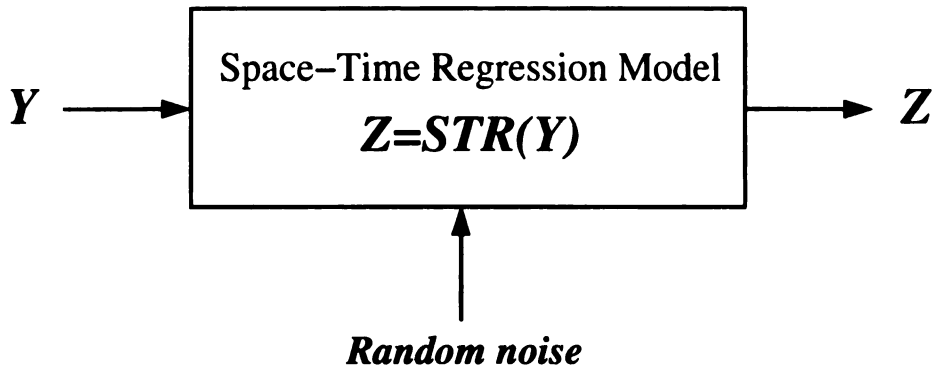

Figure 3.4. The linear relationship between space-time variable **Z** and independent variable **Y** under the assumption of space-time regression (STR) model.

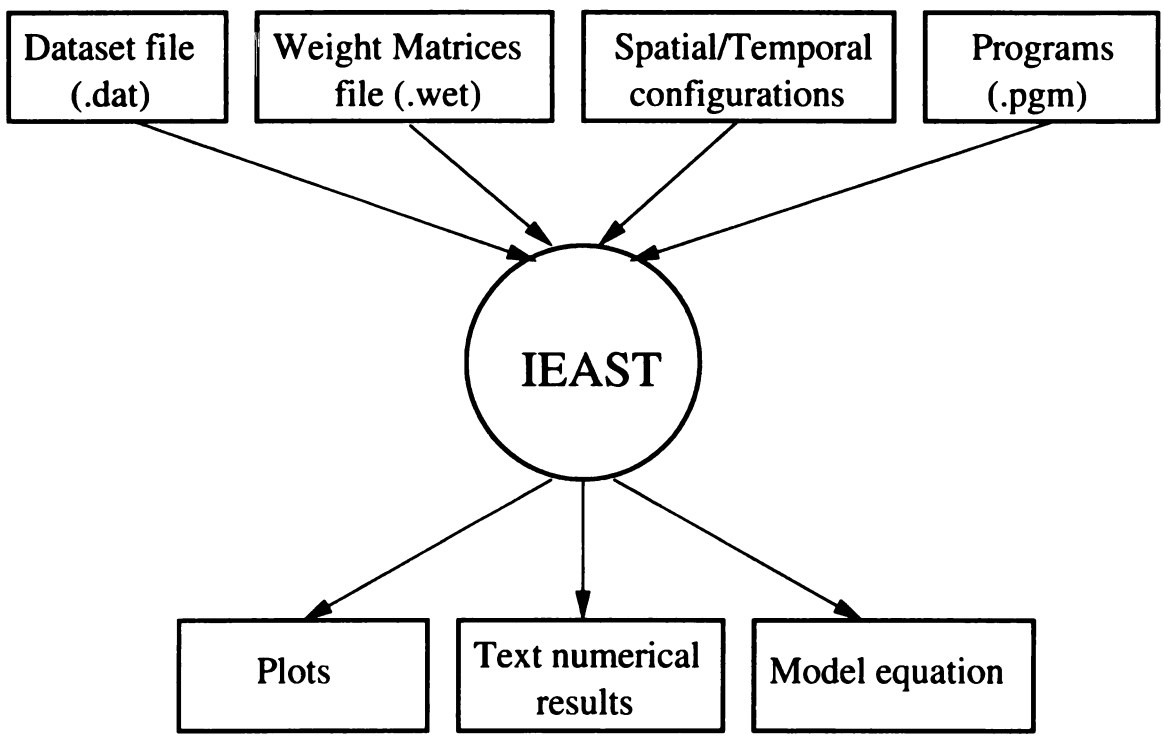

Figure 3.5. Inputs and outputs of IEAST.

Figure 3.6. System structure of the software IEAST. The gray blocks are the components of IEAST which consist of three major parts: menu-driven interface, interpreter, and core functions.

Figure 3.7. The first two levels of the menu system in IEAST

CHAPTER 4

Application to the Spreading of West Nile Virus in Detroit Metro Area

4.1 West Nile Virus

The West Nile Virus (WNV) was first detected in a woman with a mild fever in the West Nile District of Uganda in 1937. Since then WNV has been spreading to North Africa, Europe, West and Central Asia, and the Middle East. Before reaching the United States, outbreaks of the virus had occurred in Uganda (1937), Israel (1951-1957), France (1962), South Africa (1974), Algeria (1994), Romania (1996-1997), Czech Republic (1997), Congo (1998), and Russia (1999). In the United States, WNV was first recognized during an outbreak of arboviral encephalitis in New York City in September 1999 [10] (first time found in the Western Hemisphere). As of

2004, the virus had been found in 46 states with 264 human deaths [18].

Humans are infected by WNV through the bite of an infected mosquito. Mosquitoes become infected when they feed on infected birds that have high levels of West Nile viremia. These infected mosquitoes then become a vector of WNV and spread the virus when they feed on the blood of humans (and horses, etc). The virus causes fatal neurologic diseases in humans, horses, and a variety of birds. Wild birds and mosquitoes (*Culex* species) are considered the primary hosts and vectors, respectively [34, 53, 54]. Humans and horses are considered incidental hosts and infected during transmission cycles.

The outbreaks in North America are characterized by steep epidemic curves in American crow (*Corvus branchyrhynchus*) populations over time, with highly localized clusters of crow infections in space [20, 43, 52]. Mosquitoes of the genus *Culex* transmit WNV among these birds [34, 53], and crows succumb rapidly to infection within 4-8 days after exposure [35, 41]. Although other species of birds are susceptible to viral infection and may participate in viral amplification during an epizootic, corvid deaths of American crows precede other bird species, consistent with the idea that crows facilitate transmission [21]. In an outbreak in New York City in 1999, 89% of laboratory-confirmed WNV-infected birds were American crows [34]. The estimated death rate due to WNV infection was 68% within a marked crow population during an epizootic in Illinois in 2002 [58]. Thus, the absolute accumulation of dead crows (even if infection is not always laboratory-confirmed) is a measure of the spread of virus infection in the crow population, and has been identified as a sensitive indicator variable for an on-going epizootic in crows and associated heightened risk of infection

in humans [20, 52, 56].

Previous studies that predict the spread of WNV rely mainly on dead crow densities [21, 22, 23]. The dynamical behavior of the incidence of dead crows remains a poorly understood biological system with complex space-time interactions. Current understanding of the spreading dynamics of WNV infection based upon incidence of dead crows has several limitations, including: 1) there is no explicit space-time analysis or model of it [20]; 2) there are few statistical inferences provided; 3) dead crow densities are highly dependent on various factors other than WNV loads, e.g. crow population size; 4) these analyses are purely temporal. There is little research on methodologies that account for both spatial and temporal dynamics. Recently, an epidemiological early warning system called Dynamic Continuous-Area Space-Time (DYCAST) was implemented by Theophilides et al [52]. It is based on cluster analysis and uses the Knox test to retrieve the space-time interactions of dead crow sightings as indicators of an intense WNV amplification cycle. This system successfully predicted co-localization of 5 of 7 human WNV cases within 13 days of appearance of dead crow clusters in burroughs of New York City in 2001. DYCAST can identify clusters or areas of high risk for WNV infection, but does not determine the dynamics of the spread of the disease.

While STARMA modeling may or may not reveal all mechanistic components of a process, it may be useful for examining the space-time structure of its dynamics. In particular, STARMA modeling may place spatial-temporal scale constraints on epidemic infectious diseases. In this study, once either the model type and parameters or just the spatial correlation structures are obtained, the model can be used to char-

acterize the past and possibly even predict the future dynamics of spreading behavior and, most importantly, to provide information about the factors which govern the spreading behavior. Thus, we not only can recognize the areas of epidemiologic concerns in human health [28, 59], but also provide details on the space-time transmission dynamics of WNV.

Clusters of crow deaths and clusters of human cases of West Nile viral meningoencephalitis or West Nile fever tend to co-occur geographically [21, 52, 56]. In addition, accumulations of crow deaths generally precede human case notifications [22]. Many public health agencies have adopted dead crow surveillance as the primary tool for monitoring WNV activity in their jurisdictions.

Despite the confidence in accepting the validity of dead crow data (DCD) as proximate indicators of human risk of WNV infection, there exists no model of the space-time behavior of crow deaths and human cases during any single epidemic. More importantly, there has been no generalizable, explicit statistical correlation of the epidemic curves of dead crows and human cases simultaneously considered in space and time. If a positive correlation is found to exist with identifiable temporal and spatial lags between DCD and human case data, then one could reliably utilize the correlation coefficients as a predictive tool for human risk of infection in space-time, given the space-time dynamics of a dead crow epizootic curve. In this study, I not only attempt to model separately the behavior of crow epizootic and human epidemic curves during a West Nile virus outbreak using a space-time autoregressive moving average (STARMA) model, but also to investigate the interaction between these two processes using space-time cross-correlation analysis and a space-time re-

gressive (STR) model (Equation 3.16) with the intent of building a generalizable method for correlating the epidemic of infectious diseases to the indicator processes.

4.2 Description of the Datasets and Diagnoses

The dead crow data (DCD) was collected systematically before and during an outbreak of West Nile viral meningoencephalitis among humans in southeastern Michigan in the summer of 2002 [11]. Data consisted of public sightings of dead crows logged onto a website or reported by telephone call to a toll free number from April to October, 2002. The total crow death distribution (1,817 dead crow sightings) is shown in Figure 4.1. WNV infection was confirmed in a sub-sample of crows by immunohistochemical detection of viral protein of necropsied specimens [26]. The estimated rate of infection in dead crows was 70% (K. Signs and J. Patterson, unpublished). A total of 1807 dead corvid sightings were documented statewide. A plot of 1,514 dead crow sightings (the rectangular area enclosed by the thick line in Figure 4.1) were successfully classified by latitude and longitude and by date of sighting, spanning 28 weeks. These data were plotted within a rectangular area of 31.6×25.8 miles over a map of the metropolitan region of Detroit, in southeastern Michigan.

When determining an appropriate array of cell for analysis, boundary effect is an important concern. In IEAST, a mechanism is applied to the boundary cells when needed, that is, the weighting of boundary cells is normalized to compensate those cells outside the boundary. According to simulations, boundary effect can become significant when the array of cells is smaller than 8×8. In addition, to avoid

the probability distribution of the divided cells being close to Poisson distribution (violation to the assumption), configurations of smaller arrays are preferred. Thus, for purposes of analysis, this area was divided into a grid of 10×10 cells, each cell of 3.16×2.58 miles dimensions. Accordingly, the dataset was of spatial dimensions 10×10 and had 28 time periods. Figure 4.2 shows the space-time dynamics of the data. An average over time of the regulated data is shown in Figure 4.5 in which two foci and spatial trends can be identified.

Reports of human cases and samples of cerebral spinal fluid or blood serum for analysis of antibodies were obtained from clinicians and clinical laboratories in Michigan. Reports of patients presenting with viral encephalitis or meningitis, altered mental status including Guillain-Barre syndrome with atypical features, and acute febrile illness accompanied by malaise, anorexia, nausea, vomiting, and neurologic signs were screened for matching with the case definition for West Nile viral meningoencephalitis or West Nile fever. Data on address of residence and date of onset of disease were obtained from the case-patient or attending physician through telephone interviews. Antibodies to WNV and other arboviruses in sera from the case-patients were detected by IgM capture ELISA in either single serum specimens or in paired acute and convalescent sera and were classified as presumptive (single specimens with high titers) or confirmed (four-fold rise in antibody titer from acute to convalescent specimens). A total of 504 human cases were classified within the Detroit Metro area. A plot of 387 human cases spanning 13 weeks is chosen for analysis as shown in Figure 4.3. These data points are scattered in a rectangular area of 36.28 miles $\times$ 34.23 miles. The space-time dynamics are shown in Figure 4.4. Thus, the human

case data was of spatial dimensions 10×10 and had 13 time periods. An average over time of the regulated data is shown in Figure 4.5.

4.3 Assumptions and the Specification of Spatial Correlation Structure

Considering the real situations, several reasonable assumptions can be made for the dead crow and human case data in these analyses:

1. Crows died mainly of WNV.

2. The locations of human cases (or dead crows) reported were the same (or near) places where human (or crows) got infected.

3. The spatial correlation structures for these data were isotropic and invariant with time.

There are many ways to define spatial correlation structures for applications. Most of them are based on the Euclidean distances between cells. For the analyses in this chapter, the de-trended data is assumed to be isotropic (directionless) or uniform. A definition used by Pfeifer and Deutsch [48] was used. The two matrices used for specifying correlation structures, the spatial order matrix (SOM) and the spatial relation matrix (SRM), are defined in Equation 4.1 and 4.2, respectively. The matrix SOM defines the relative positions of spatial orders (or lags) up to sixth. For example, all elements in SOM with number 3 are the third order neighbors of the

central cell. The matrix SRM gives relative weights of every cell in each spatial order. As mentioned in Section 3.1, all weights in a given spatial order should sum up to unity. The spatial weight matrices $W^{(l)}$, which are six matrices (from spatial order one to six) with dimension 100×100, can be derived from the given SOM and SRM (as shown in Section 3.1), and are required in all calculations in STARMA modeling.

$$SOM = \begin{bmatrix} - & 6 & 5 & 5 & 5 & 6 & - \\ 6 & 5 & 4 & 3 & 4 & 5 & 6 \\ 5 & 4 & 2 & 1 & 2 & 4 & 5 \\ 5 & 3 & 1 & 0 & 1 & 3 & 5 \\ 5 & 4 & 2 & 1 & 2 & 4 & 5 \\ 6 & 5 & 4 & 3 & 4 & 5 & 6 \\ - & 6 & 5 & 5 & 5 & 6 & - \end{bmatrix} \tag{4.1}$$

$$SRM = \begin{bmatrix} - & .125 & .063 & .063 & .063 & .125 & - \\ .125 & .063 & .125 & .25 & .125 & .063 & .125 \\ .063 & .125 & .25 & .25 & .25 & .125 & .063 \\ .063 & .25 & .25 & 1 & .25 & .25 & .063 \\ .063 & .125 & .25 & .25 & .25 & .125 & .063 \\ .125 & .063 & .125 & .25 & .125 & .063 & .125 \\ - & .125 & .063 & .063 & .063 & .125 & - \end{bmatrix} \tag{4.2}$$

4.4 Space-Time Trends and De-Trending

The environmental conditions for disease spreading are generally not uniform over space and time, especially for large scales. For example, population of crows, popu-

lation of mosquitoes, and habitats for coexisting of crows and mosquitoes definitely affect the spreading behavior and are clustered over space. Many reasons (such as still waters for mosquitoes, accessibility of food for crows, etc) result in this non-uniformity. In this study, the non-uniformity was removed using space-time de-trending.

To make the space-time series stationary, the spatial and temporal trends were removed. Many polynomial surfaces for spatial trends were fitted and tested, but no substantial differences were found for the space-time autocorrelations of the de-trended data. Even the cut-off positions in space and in time remained constant. The only difference was that the STACF/STPACF tended to be lower when the orders of the spatial and the temporal trends in the polynomials were higher. This is to be expected, because the high orders removes autoregression. To avoid removing too much autoregression, the orders of the polynomials for the trend surfaces were selected to be as small as reasonable. For temporal trends, the averages over space were preferred for use because there is no appropriate smoothing function (e.g. polynomials) that can be well-fitted to the data (there existed a second high peak in Figure 4.6, which was probably caused by news announcements of dead crow cases of West-Nile virus at week 19). If a smooth function was fitted to such data and then used for de-trending, it could introduce spurious negative correlations in the de-trended data.

Because there were at least two spatial foci of incidence in the data, a fourth order polynomial trend surface combined with the temporal trend (average over space) was ultimately removed from the data. The fourth order polynomial spatial trend

$S_{trend}(x, y)$ is the least-square best-fit of Equation 4.3 to the data and is expressed as

$$S_{trend}(x, y) = \sum_{i=0}^{4} \sum_{j=0}^{i} a_{ij} x^j y^{i-j} \tag{4.3}$$

where a_{ij} are the least-square best-fit coefficients. The temporal trend is the average over two-dimensional space and can be written as

$$T_{trend}(t) = \frac{1}{N} \sum_{x} \sum_{y} Z_{(x,y),t}$$

where N is the number of sites, $Z_{(x,y),t}$ is the element of $\mathbf{Z}_t$ at location (x, y), and vector $\mathbf{Z}_t$ is the original space-time series data. In combination, the spatial and temporal trends were removed according to the equation

$$Z^*_{(x,y),t} = Z_{(x,y),t} - h \cdot S_{trend}(x, y) \cdot T_{trend}(t)$$

where h is a constant so that $h \cdot S_{trend}(x, y) \cdot T_{trend}(t)$ best-fit to $\mathbf{Z}_t$, $\mathbf{Z}^*_t$ is the de-trended $\mathbf{Z}_t$. After de-trending, the grand mean was also subtracted.

The spatial trend combining with temporal trends (as shown in Figure 4.6) and grand mean were removed from the DCD dataset before analysis and modeling.

4.5 Results for the Analysis and Modeling of the Dead Crow Dataset

In order to identify an appropriate model type for the DCD, space-time autocorrelations were calculated. The STACF and STPACF of the de-trended data for the given SOM and SRM (defined in Equation 4.1 and 4.2) are shown in Figure 4.7. They indicated that the model type was STAR, because the STACF tails-off and the STPACF cuts-off. Based on the temporal and spatial lags where the STPACF cuts-off, the maximum temporal and spatial orders were determined as three and four, respectively. Thus, the candidate of the underlying process of the data was STAR(maxT=3, maxS=4).

Based on the model type and orders found, the maximum likelihood estimates of the parameters and corresponding significance levels can be calculated, and these estimates are shown in Table 4.1. Substituting the estimates into Equation 2.1, we have the following STARMA model equation:

$$\begin{aligned}
\mathbf{Z}_t = \; & .26\mathbf{Z}_{t-1} && +.04\mathbf{Z}_{t-2} && -.02\mathbf{Z}_{t-3} \\
& +.36\mathbf{W}^{(1)}\mathbf{Z}_{t-1} && -.18\mathbf{W}^{(1)}\mathbf{Z}_{t-2} && -.11\mathbf{W}^{(1)}\mathbf{Z}_{t-3} \\
& +.10\mathbf{W}^{(2)}\mathbf{Z}_{t-1} && -.07\mathbf{W}^{(2)}\mathbf{Z}_{t-2} && +.02\mathbf{W}^{(2)}\mathbf{Z}_{t-3} \\
& -.09\mathbf{W}^{(3)}\mathbf{Z}_{t-1} && -.04\mathbf{W}^{(3)}\mathbf{Z}_{t-2} && -.02\mathbf{W}^{(3)}\mathbf{Z}_{t-3} \\
& +.04\mathbf{W}^{(4)}\mathbf{Z}_{t-1} && -.11\mathbf{W}^{(4)}\mathbf{Z}_{t-2} && -.03\mathbf{W}^{(4)}\mathbf{Z}_{t-3} \\
& +\boldsymbol{\varepsilon}_t.
\end{aligned} \tag{4.4}$$

Table 4.1. The estimates and significance levels of the model parameters (STAR(MaxT=3, MaxS=4)).

Estimates of parameters

	Spatial lag				
Temporal lag	S=0	S=1	S=2	S=3	S=4
T=1	.26	.36	.10	-.09	.04
T=2	.04	-.18	-.07	-.04	-.11
T=3	-.02	-.11	.02	-.02	-.03

Significance levels of parameters

	Spatial lag				
Temporal lag	S=0	S=1	S=2	S=3	S=4
T=1	.001	.001	.010	.100	.400
T=2	.040	.001	.040	.300	.010
T=3	.400	.010	.250	.900	.600

Two diagnostic checks were conducted for this analysis: significance test for the parameters and correlation analysis for the residuals. The significance test results are shown in Table 4.1. Several parameters at relatively high temporal or spatial lags were statistically insignificant. However, technically the maximum temporal and spatial orders from the candidate STAR(maxT=3, maxS=4) should not be reduced, because highly significant parameters were found at S=4/T=2 (spatial lag 4 and temporal lag 2) and S=1/T=3 (spatial lag 1 and temporal lag 3). The STACF/STPACF of residuals show no significant correlation among residuals (Figure 4.8). Therefore, the candidate model represented the underlying process of the DCD adequately, as it is well-fitted by the data and it is parsimonious.

In addition to the 10 × 10 cell configuration, the same area was divided into various arrays of cells up to 20 × 20, for analyses in different configurations. Based on the same spatial correlation structure (i.e. same SOM and SRM) and de-trending

methods, these modeling results were consistent. The maximum spatial order and the parameters of the model differed when different arrays of cells were used to model the same data as is expected, because smaller cell size is associated with higher maximum spatial order needed for the model. However, the physical meaning of the modeling results does not change (as shown in Table 4.2 and 4.3).

Three analyses using various arrays of cells: 10×10, 16×16, and 20×20 were performed. Only parameters that met the 0.01 significance level were entered into the final models. Based on the significance level, the maximum spatial orders for these three cases can be determined as fourth, sixth, and seventh, respectively. The maximum spatial order can be transformed into distance in the following way. For configuration 10×10, the cell size is 3.16×2.58 miles. According to the spatial lag definition in Equation 4.1, the spatial lag four is two cells away in X-direction and one cell away in Y-direction from the central cell (lag zero). Thus, the physical distance in the configuration 10×10 for spatial lag four can be calculated as the following.

$$\text{physical distance} = \sqrt{(2 \cdot 3.16)^2 + (1 \cdot 2.58)^2} = 6.83 \text{ miles.}$$

After transformations, the equivalent physical distances (measured between cell centers) for these maximum spatial orders were found to be quite uniform (6.83, 6.76, and 6.45 miles, respectively). This shows that the spatially correlated areas from these modeling results are consistently about 6.4 miles in radius, regardless of the cell size used.

Furthermore, additional analyses were conducted over a wide range of still other

Table 4.2. The maximum spatial lags and the corresponding distances between cell midpoints for that lag, for three models using different cell numbers (larger numbers, smaller areas).

Cell numbers	Estimated max S order	Equivalent real distances
10x10	4	6.83 miles
16x16	6	6.76 miles
20x20	7	6.45 miles

configurations, including: various datasets of sub-sampling of various parts of the entire area; and various other cell configurations and SOMs and SRMs. For each of these, the refined iterative modeling procedure and the de-trending method were used. In all cases (results not shown), the model that was converged to was consistent with the model reported.

The results show that the incidence of reported dead crows per week fitted a STAR model. A high degree of space-time autoregression was found, and no evidence of moving average influences. The latter result indicates that although there was both stochasticity and statistical noise in the observed values (the estimated variance is 1.451), these sources of variation were not directly or immediately shared spatially or temporally among cells. Rather, they were shared only through the autoregressive structure attributable to spatial spreading of WNV. Diagnostically, the envelope of the STPACF cuts to zero after temporal lag three, whereas the STACF tails off with increasing time lag (Figure 4.7). The robustness of this diagnostic check, especially so since the process is STAR, was also indicated in the extensive investigations using computer simulations. Moreover, the residuals are those expected from random (white) noise. There do not appear to be any particular reasons to expect STMA

spatially-shared stochastic inputs or shared random errors in the DCD.

The results also show that the process underlying the DCD has low spatial and temporal orders. Moreover, the autoregressive coefficients show a pattern that generally fits those expected for most biological processes. For example, consider the 10×10 array of cells and the spatial order matrix (SOM) of Equation 4.1 and the spatial relation matrix (SRM) of Equation 4.2. As could be expected, the autoregressive parameters are positive and large for close spatial proximities, i.e. small distances, and drop in value as distance increases (Table 4.1). Exceptions involving fairly large negative parameters for short distances at time lags two and three are discussed below. The results also showed that the temporal order of the process is nearly Markovian, since large and significant parameters are observed only for time lags of one and two weeks. It has been argued that the Markovian property should be common in biological processes [24]. In addition, the prominence of one- and two-week lagged effects is consistent with the short interval between infection and death [35], and with empirical evidence of intense localized epizootics [20, 52].

Using the same SOM and SRM, the best fitted model is STAR(maxT=3, maxS=4), and completely consistent results were found using various other arrays of cells. The maximum spatial lags consistently gave maximum distances of ca. 6.4 miles, in terms of distances between cell-centers. I also examined models with various other SOMs and SRMs, and again obtained completely consistent results. However, as noted in the results and further explained below, there was evidence that the models could be reduced further in size; the model might be reduced to STAR(maxT=2, maxS=2), consistent with the array of positive autoregressive effects.

Table 4.3. The maximum spatial lags for positive autoregressive effects and the corresponding distances between cell midpoints for that lag, for three models using different cell numbers.

Cell numbers	Estimated max S order	Equivalent real distances
10x10	2	4.07 miles
16x16	3	3.96 miles
20x20	4	3.41 miles

Focusing on positive autoregressive effects, it appears that the DCD operates over rather smaller distances, with maximum distances of direct effects of locations on one another of ca. four miles or less. However, the distances over which the virus load is primarily moving (spreading) may still be much smaller. There is a remarkable reduction of autoregressive effects for spatial lag 2 (0.10) compared to spatial lag 1 (0.36) for the 10×10 array of cells and SOM of Equation 4.1 and SRM of Equation 4.2. The main difference between the two is the length of the boundaries between cells with various spatial lags. This strongly suggests that the length of boundary largely determines the rates of spread of WNV among cells. If that is the case, it would indicate that WNV directly spreads mostly over very short distances across cell borders, less than a mile or possibly even a few hundred meters.

Crows can fly long distances, so the scale over which the DCD show large positive autoregressive effects may seem surprisingly small. However, the mosquito vector is active primarily during hours when the crows are roosting, and crows are territorial and show high fidelity to roosting sites [57]. Thus, crows may largely be being infected and transmitting the virus at their roosting sites, and thus apparently not spatially spreading the disease, although they would amplify the viral load locally. In addition,

the distances of flights between where dead crows are found and their roosting sites are superimposed on this process [58]. Although such movements do not spread the disease, they could influence the short term autoregression of DCD. While crows are a very good indicator of viral loads, they may not be largely responsible for spreading the disease spatially. Other animals or even the mosquito vector itself may be more responsible for the local spatial spread per se.

Two points should be reinforced. First, since crows are one of the most infected animals, they could be very important in amplifying the disease locally (say within a cell). Second, the results do not mean that the crows do not on occasion spread WNV over long distances, and perhaps even start a new local epidemic. It is unlikely that such low frequency events would be detected in the analyses. Nevertheless, the results do suggest that there is high probability that local epidemics could be contained, for example, through a concentrated eradication effort, by spraying mosquito repellant at mosquito breeding ponds, etc.

One specific feature of the autoregressive structure that deserves special examination is the rather large negative autoregressive parameters at time lag 2 for some spatial lags. This was observed in all models, and appears to be a real effect. It may be explained by the depletion of the crow population. Depletion of the crow population was extreme: during the course of the WNV season in 2002, up to 70 % of crows in Detroit died from WNV (K. Signs and J. Patterson, unpublished). Hence, for example, if a cell had unusually high DCD for a given time period (week), especially during the height of the WNV season, a significant portion of the crow population in that cell would have died, leaving fewer to be infected and die one to two weeks

later. Thus, the results suggest that there are actually two DCD autoregressive processes. The first is dependent on the local and nearby WNV loads, and the second is a weaker echo effect caused by depletion. The mixture of the two may make temporal lag one DCD autoregressive parameters all positive, but not all of the temporal lag two parameters.

4.6 Results for the Analysis and Modeling of the Human Case Dataset

The analyzing procedure followed for human case data was similar to that for DCD. The same spatial correlation structure (i.e. same SOM and SRM as in Equation 4.1 and 4.2) was used for spatial dimension 10×10, but a fourth order polynomial trend surface combined with the temporal trend (weekly average over space) (as shown in Figure 4.9) and the grand mean were removed from the human case data before STARMA modeling. The autocorrelations for the de-trended human case data are shown in Figure 4.10. It is not surprising that the autocorrelations fluctuate considerably since the number of cases of human case data is much smaller than that of DCD. However, the behavior of STACF/STPACF still revealed some clues about model type. The curves in the STACF gradually decay to zero as temporal lag increases. On the contrary, the curves in the STPACF drop quickly and cross axis at a short temporal lag, and then fluctuate and gradually diminish at higher temporal lags. This suggests that the pure STAR model is a good candidate for the de-trended data.

Regarding the maximum orders of the candidate model, we can start with by using six as the maximum for both spatial and temporal lags, because the STACF showed that there's no significant correlation for spatial and temporal lags greater than six. Although a STAR model with maximum spatial and temporal order six may be too complex for the data, we start from the candidate model STAR(MaxT=6,MaxS=6) and then iteratively refine the model to achieve a parsimonious one.

Given the candidate model STAR(MaxT=6,MaxS=6), the model parameter estimates and the significance level for each parameter can be found in the stages of estimation and diagnostic checking, respectively (Table 4.4). As in the analysis of DCD, if significance level is higher than 0.01, the corresponding parameter is classified as insignificant and then is removed from the candidate model. Once all parameters for a temporal lag (or a spatial lag) are removed, the maximum temporal or spatial order may be reduced. In the significance analysis listed in Table 4.4, all significance levels less than equal to 0.01 are underlined. The results show that the maximum temporal order four and spatial order six are sufficient for these significant parameters. Thus, the new candidate model is a reduced STAR(MaxT=4,MaxS=6).

Next, based on the new candidate model, the new parameter estimates and corresponding significance levels are calculated as in Table 4.5. Note that the values of retained parameters are very similar to those estimated using STAR(MaxT=6,MaxS=6). Based on significance tests, it is not possible to further reduce the maximum orders because the parameter with temporal lag four and spatial lag six is significant (significance level ≤ 0.01). Substituting the parameter estimates into the model equation for STAR(MaxT=4,MaxS=6), a mathematical equation for

Table 4.4. Initial estimates and significance levels of the model parameters (STAR(MaxT=6, MaxS=6)) for the human dataset.

Estimates of parameters

	Spatial lag						
Temporal lag	S=0	S=1	S=2	S=3	S=4	S=5	S=6
T=1	.26	.06	-.10	-.29	-.05	-.30	-.06
T=2	.12	.27	.13	-.12	-.11	-.22	-.11
T=3	.07	.10	-.15	.05	-.00	.06	-.01
T=4	.04	-.17	-.07	-.02	.16	.25	.11
T=5	-.01	-.10	-.04	.10	-.06	.11	.06
T=6	-.04	.08	.09	.03	-.03	-.19	-.09

Significance levels of parameters

	Spatial lag						
Temporal lag	S=0	S=1	S=2	S=3	S=4	S=5	S=6
T=1	.001	.700	.040	.001	.500	.100	.150
T=2	.010	.100	.040	.150	.300	.020	.020
T=3	.900	.400	.040	.150	.900	.400	.900
T=4	.900	.010	.900	.900	.020	.010	.010
T=5	.500	.300	.900	.080	.500	.300	.150
T=6	.900	.250	.150	.900	.600	.040	.040

Table 4.5. Final estimates and significance levels of the model parameters (STAR(MaxT=4, MaxS=6)) for the human dataset.

Estimates of parameters

	Spatial lag						
Temporal lag	S=0	S=1	S=2	S=3	S=4	S=5	S=6
T=1	.26	.06	-.11	-.30	-.07	-.30	-.06
T=2	.10	.27	.12	-.11	-.09	-.22	-.11
T=3	.05	.07	-.15	.06	.01	.06	-.01
T=4	.03	-.14	-.07	.03	.16	.21	.10

Significance levels of parameters

	Spatial lag						
Temporal lag	S=0	S=1	S=2	S=3	S=4	S=5	S=6
T=1	.001	.700	.040	.001	.400	.010	.150
T=2	.020	.100	.060	.150	.400	.040	.020
T=3	.900	.900	.040	.150	.700	.400	.900
T=4	.900	.040	.800	.400	.020	.020	.010

representing the underlying process of the human epidemics can be obtained as in Equation 4.5.

The autocorrelation analysis for the residuals shows no large correlations among spatial lags zero to six and temporal lags one to four. However, there do exist some correlations in higher temporal and spatial lags. The correlations in higher spatial and temporal orders (which are not included in the model) are not as significant as that in the lower orders (Table 4.5). Therefore, this model (Equation 4.5) can be accepted as the final model for representing the process of human epidemics. In summary, the human epidemic in Detroit Metro area is highly autocorrelated within spatial lag six (equivalent to 12.57-12.86 miles away in physical distance) and four weeks time period.

$$
\begin{aligned}
\mathbf{Z}_t = \; & .26\mathbf{Z}_{t-1} && +.10\mathbf{Z}_{t-2} && +.05\mathbf{Z}_{t-3} && +.03\mathbf{Z}_{t-4} \\
& +.06\mathbf{W}^{(1)}\mathbf{Z}_{t-1} && +.27\mathbf{W}^{(1)}\mathbf{Z}_{t-2} && +.07\mathbf{W}^{(1)}\mathbf{Z}_{t-3} && -.14\mathbf{W}^{(1)}\mathbf{Z}_{t-4} \\
& -.11\mathbf{W}^{(2)}\mathbf{Z}_{t-1} && +.12\mathbf{W}^{(2)}\mathbf{Z}_{t-2} && -.15\mathbf{W}^{(2)}\mathbf{Z}_{t-3} && -.07\mathbf{W}^{(2)}\mathbf{Z}_{t-4} \\
& -.30\mathbf{W}^{(3)}\mathbf{Z}_{t-1} && -.11\mathbf{W}^{(3)}\mathbf{Z}_{t-2} && +.06\mathbf{W}^{(3)}\mathbf{Z}_{t-3} && +.03\mathbf{W}^{(3)}\mathbf{Z}_{t-4} \\
& -.07\mathbf{W}^{(4)}\mathbf{Z}_{t-1} && -.09\mathbf{W}^{(4)}\mathbf{Z}_{t-2} && +.01\mathbf{W}^{(4)}\mathbf{Z}_{t-3} && +.16\mathbf{W}^{(4)}\mathbf{Z}_{t-4} \\
& -.30\mathbf{W}^{(5)}\mathbf{Z}_{t-1} && -.22\mathbf{W}^{(5)}\mathbf{Z}_{t-2} && +.06\mathbf{W}^{(5)}\mathbf{Z}_{t-3} && +.21\mathbf{W}^{(5)}\mathbf{Z}_{t-4} \\
& -.06\mathbf{W}^{(6)}\mathbf{Z}_{t-1} && -.11\mathbf{W}^{(6)}\mathbf{Z}_{t-2} && -.01\mathbf{W}^{(6)}\mathbf{Z}_{t-3} && +.10\mathbf{W}^{(6)}\mathbf{Z}_{t-4} \\
& +\boldsymbol{\varepsilon}_t
\end{aligned}
\tag{4.5}
$$

There are several points needed to be noted. The area defined for analysis for DCD is different from that for human case data. Nonetheless, this would not affect the final results because the spatial correlation structure used for both datasets is directionless or isotropic. The highly infective area (up to spatial lag 6, or about 12 miles) for human is larger than that for dead crows (up to spatial lag 4, or about 6 miles). This infers that the disease spreading of human cases is, spatially, faster than that of crow cases. It can be explained by the mobility and active area of human is much larger than that of crows. Because of this, the observation error in human case dataset is also larger than that in DCD.

From the parameter estimates, the significant parameters for DCD tend to be gathered around lower spatial orders, and tend to be dispersed in a wide area for human case data. In short spatial lags (one or larger) and temporal lags(two or larger), especially for spatial lag one at temporal two or higher, the parameters for

human case data is positive and significantly above zero. Contrarily, the parameters for DCD are negative and significantly below zero. Because human cases should not have evident depletion effect as compared to crow cases, this further justifies that the depletion effect among crow is the major reason causing these negative parameters around low spatial orders for DCD.

4.7 Summary of Analyses for the Dead Crow Dataset and the Human Case Dataset

As needed for the first step of the STARMA analyses, the space-time autocorrelation function (STACF) and the space-time partial autocorrelation function (STPACF) are shown in Figure 4.7 for the DCD and in Figure 4.10 for the human case data. For both datasets, the STACF values are very large at short spatial lags (distances) and temporal lags, and they tail-off as either temporal or spatial lag increases. The STPACFs of both become and stay near-zero (cut-off) at small spatial and temporal lags. These fundamental aspects of the both dataset are diagnostic in showing that both DCD and human cases behaved as a STAR process, and possessed no shared stochastic inputs. In other words, each space-time series behaved as a purely autoregressive process with added white noise (spatially and temporally random stochastic inputs). Moreover, both datasets have small spatial and temporal orders, the maximum possible values of which are determined by the lags at which the STPACF cuts to zero.

For the DCD, based on inspection of the STPACF, the maximum spatial lag is four and the maximum temporal lag is three. Estimates of the space-time autoregression coefficients for various spatial and temporal lags for the STAR(maxT=3, maxS=4) structure are shown in Table 4.1. Large and highly significant positive coefficients are only observed for (T=1, S=0), (T=1, S=1), (T=2, S=0). This indicates that most of the spread operates between nearest neighbor cells, and over a two week time period. It is also noteworthy that the coefficient value is much larger for rooks move (spatial lag S =1) than for bishops move (S=2) nearest neighboring cells. Similarly, the STPACF for the human case data indicate that the maximum spatial lag for positive autoregression is six and the maximum temporal lag is four. Estimates of the space-time autoregression coefficients for various spatial and temporal lags for the STAR(maxT=4, maxS=6) model are shown in Table 4.5.

Diagnostic checking for model adequacy is conducted not only by the significance test of parameters, but also by inspection of the STACF/STPACF of the residuals. The graphs of the STACF/STPACF of the residuals for the de-trended DCD, as shown in Figure 4.8 are flat throughout, closely resembling that of the random white noise. This is consistent with the fitted STAR model having captured all of the autoregression in the de-trended data [16]. The same was observed in the STACF and STPACF of the residuals of the fitted model for the human case data.

As noted before, I further examined both datasets with other model constructions. I varied cell sizes and used sub-areas of the entire dataset, and in all cases, for both datasets, the best fitted models were purely STAR models, and these had fully consistent distances and time scales. I also used different polynomials for spatial

de-trending. Again, the results were fully consistent, except that very large-order polynomials generally tended show reductions in autoregressive coefficients. This is expected, since over-de-trending generally removes some autoregression or autocorrelation [8].

4.8 Results for the Cross Analysis Between Human Case and Dead Crow Datasets

Dead crows are sentinels of WNV viral activity generally and locally, and there is a burgeoning effort to use dead crow reports as a local signal of an outbreak (Theophilides et al [52] with their Knox statistic and Watson et al [56] with their geostatistical analysis). However, there were no studies that considered spatial and temporal interactions simultaneously, nor that presented a cross-correlation analysis between dead crow reports and human cases. In addition, there have been no studies that incorporated specific spatial units and project the spread of WNV. Therefore, a few question arose: how predictive in space units and time units are DCD for human cases? What is the predictive power? If the two processes (dead crow epizootic curve and human case epidemic curve) are demonstrably space-time regressive, then a model can be developed in which spatial and temporal lags with specific model parameters are provided, showing the "strength" of the relationships at defined space-time units. In this study, the cross correlations between the DCD and the human case dataset are investigated using the method mentioned in Section 3.6. The results provide a

foundation for creating a space-time regression in which dead crows data are used as an indicator variable to derive the space-time dynamics of human epidemics.

A plot of 1,753 dead crow sightings was successfully classified by latitude and longitude and by date of sighting, spanning 28 weeks. The plot of 385 human cases of the same rectangular area covered as in previous dead crow sightings was classified and spanned 14 weeks. These data points and the area for the cross correlation analysis are shown in Figure 4.12 and Figure 4.13. These data were plotted within a nearly square area of 39.17 miles $\times$ 39.23 miles over a map of the metropolitan region of Detroit, Michigan. This area was again divided into a grid of 10 $\times$ 10 cells, each cell with 3.92 miles $\times$ 3.92 miles dimensions.

Because the human case data spanned from Julian week 30 to week 43 and DCD from week 16 to week 43, zeros was applied to human case data for the cells in time periods in which there were no cases reported. Thus, in space and time the DCD and human case data cover the same time period and physical area. Accordingly, the data set was of spatial dimensions 10 $\times$ 10 and had 28 time periods.

We also assumed that the spatial correlation structure of the data was isotropic and invariant with time. A higher order (up to ten) of spatial correlation structure is used for the cross analysis (the SOM and SRM are shown in Equation 4.6 and 4.7). The method in section 3.6 was used for calculating cross correlations, and no spatial and temporal de-trending was used.

$$SOM = \begin{bmatrix} - & - & - & 10 & 9 & 9 & 9 & 10 & - & - & - \\ - & 10 & 9 & 8 & 7 & 7 & 7 & 8 & 9 & 10 & - \\ - & 9 & 8 & 6 & 5 & 5 & 5 & 6 & 8 & 9 & - \\ 10 & 8 & 6 & 5 & 4 & 3 & 4 & 5 & 6 & 8 & 10 \\ 9 & 7 & 5 & 4 & 2 & 1 & 2 & 4 & 5 & 7 & 9 \\ 9 & 7 & 5 & 3 & 1 & 0 & 1 & 3 & 5 & 7 & 9 \\ 9 & 7 & 5 & 4 & 2 & 1 & 2 & 4 & 5 & 7 & 9 \\ 10 & 8 & 6 & 5 & 4 & 3 & 4 & 5 & 6 & 8 & 10 \\ - & 9 & 8 & 6 & 5 & 5 & 5 & 6 & 8 & 9 & - \\ - & 10 & 9 & 8 & 7 & 7 & 7 & 8 & 9 & 10 & - \\ - & - & - & 10 & 9 & 9 & 9 & 10 & - & - & - \end{bmatrix} \tag{4.6}$$

$$SRM = \begin{bmatrix} - & - & - & .083 & .050 & .050 & .050 & .083 & - & - & - \\ - & .083 & .050 & .083 & .083 & .083 & .083 & .083 & .050 & .083 & - \\ - & .050 & .083 & .125 & .063 & .063 & .063 & .125 & .083 & .050 & - \\ .083 & .083 & .125 & .063 & .125 & .250 & .125 & .063 & .125 & .083 & .083 \\ .050 & .083 & .063 & .125 & .250 & .250 & .250 & .125 & .063 & .083 & .050 \\ .050 & .083 & .063 & .250 & .250 & .000 & .250 & .250 & .063 & .083 & .050 \\ .050 & .083 & .063 & .125 & .250 & .250 & .250 & .125 & .063 & .083 & .050 \\ .083 & .083 & .125 & .063 & .125 & .250 & .125 & .063 & .125 & .083 & .083 \\ - & .050 & .083 & .125 & .063 & .063 & .063 & .125 & .083 & .050 & - \\ - & .083 & .050 & .083 & .083 & .083 & .083 & .083 & .050 & .083 & - \\ - & - & - & .083 & .050 & .050 & .050 & .083 & - & - & - \end{bmatrix} \tag{4.7}$$

As shown in the analysis of results for the DCD and the human cases, the two space-time process both behave as a STAR process and possessed no shared stochastic or statistical noise inputs. The space-time cross-correlations between the DCD and the

human case data are shown in Figure 4.14. The spatial lags span from 0 to 10, and the temporal lags from -12 to +12. The temporal lags in the cross correlation represent the number of weeks in which human cases reported lagged after dead crow cases reported. The cross correlation coefficients were very large (as high as 0.7). Across all spatial lags the peaks were such that DCD preceded the human cases by three weeks, and these cross correlation curves are perfectly aligned. Although this result does not provide a proof of causation flow between dead crows and human cases, the result do support the use of dead crow data as a space-time indicator variable for forecasting human epidemics. Moreover, the correlations drop smoothly and symmetrically for temporal lags deviating from the three week time lag and spatial lags increasing from 0 to 10. Generally the correlations also decrease smoothly with spatial distance. An exception is the slightly greater values for spatial lag one than spatial lag zero. This may infer that crows died in adjacent cells is even more correlated with the incidence of human cases than the crows that died in the same cell. However, this outcome may simply result from sampling effects. At least, we know that the cross correlations between human cases and dead crows for spatial lag one are as large as those for spatial zero. When temporal lag decreases to -8 or below, the correlations between these two datasets are negligible (less than 0.1). This result indicates that the crow epidemic that occurred eight weeks earlier no longer correlates with current human epidemic in space and time. When spatial lag increases up to 10 (equivalent to 21.1 miles), the cross correlations are reduced to as low as 0.2.

As mentioned in section 3.6, the autocorrelation of the DCD can spuriously contribute to cross correlations. To eliminate this effect, the cross correlation is conducted

with pre-whitening. Both the DCD and human case data are 'filtered' with the estimated crow model equation (Equation 4.4), and then calculate the cross correlation between the two filtered space-time series as usual. The result is shown in Figure 4.15. Generally, the cross correlation is lower, but there is no significant difference between Figure 4.14 and Figure 4.15 in terms of pattern. This provides compelling evidence that the high cross correlation between the human and the dead crow data is causal and not spurious.

4.9 Discussion

The STARMA framework of statistical space-time modeling appears to hold considerable promise for characterizing the space-time autoregressive structure and correlations for many biological processes, including many variables associated with the spread of zoonotic infectious diseases. The application of STARMA to dead crow data (DCD) as indicators of the environmental load of the WNV in metropolitan Detroit showed that DCD follows a STAR process having low spatial and temporal orders. This outcome is consistent over various definitions of spatial lag structures. The results indicate that the incidence of dead crow does not spread spatially very far very fast. Indeed, a focus on positive autoregressive effects and the contrast between rooks-move and bishops-move nearest-neighboring cells suggests that most of the spread is determined by the length of shared boundaries. Biological considerations suggest that although crows are a likely cause of local amplification of WNV, they are not spatially spreading the virus very far very fast. It is possible that other hosts

or even the mosquito vector itself may be more responsible for the spatial spread. In addition, there appear to be significant effects of depletion of the crow population on the space-time dead crow incidence.

The autocorrelation results for modeling the human case data appear more erratic than that for DCD (as shown in the STACF/STPACF for human case data in Figure 4.10), however it follows the same pattern as the DCD. It should also be kept in mind that the number of human cases is much smaller than that for DCD. Such paucity in the number of samples can introduce statistical noise during analysis. In space-time modeling, the condition of abundant samples in time and in space is critical.

It is important to discuss the de-trending over time and space. The fact that we de-trended temporally over the season means that the autoregressive parameters are not rates of spread, but rather relative rates given the overall increase or decrease in WNV. The former might be recovered by adding multipliers to the autoregressive (the variance of the errors term $\boldsymbol{\varepsilon}_t$ might also change over time) structures, depending on the exact week. More interesting is the rather striking spatial non-stationarity. Since these have been removed by polynomial regression, the analyses reveals nothing about why they occurred.

Finally, in the cross analysis, the two datasets revealed very high correlation which reaches the maximum values for each spatial lag at temporal lag -3. That is, the number of crows dying in neighboring cells three weeks earlier has the highest correlation with the number of human infections. This result supports the assumption in the epidemiological literature that crow epidemic is leading human epidemic by three weeks

and it strengthens the basis for using dead crow data as an indicator for forecasting human epidemics.

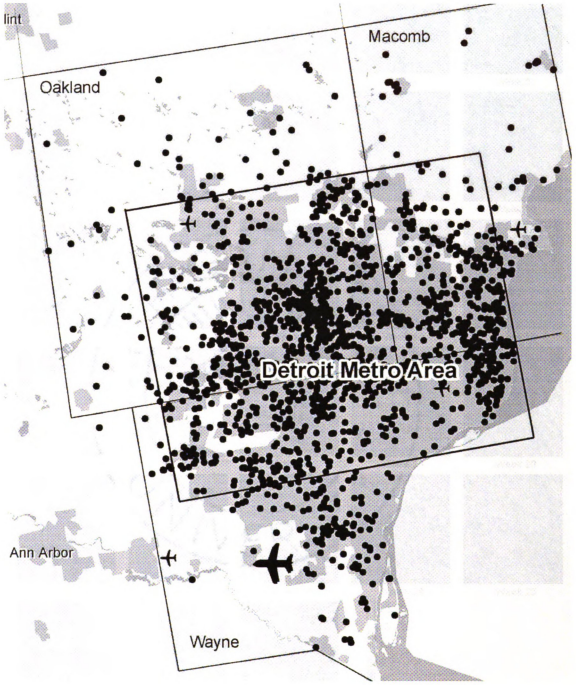

Figure 4.1. Dead crows reported in the Detroit Metro area. The rectangular area enclosed by thick line is the area retained after truncation.

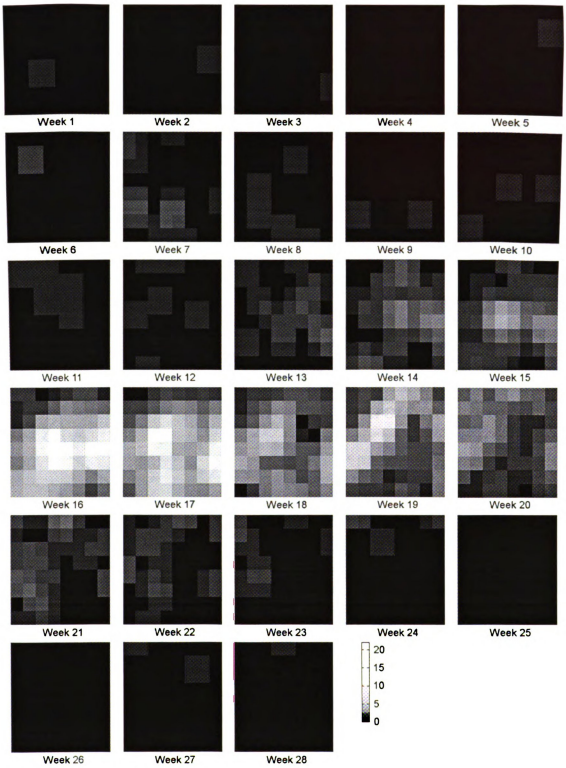

Figure 4.2. Dead crow dataset in the Detroit Metro area with 10×10 cells in space and spanning 28 weeks in time.

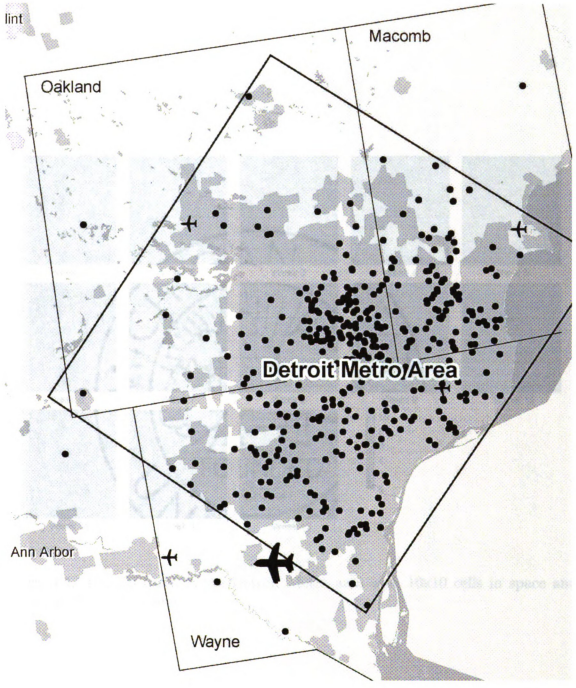

Figure 4.3. Human cases reported in the Detroit Metro area. The rectangular area enclosed by thick line is the area retained after truncation.

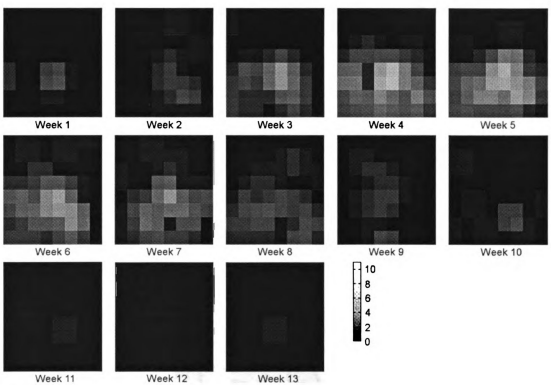

Figure 4.4. Human cases in the Detroit Metro area with 10x10 cells in space and spanning 13 weeks in time.

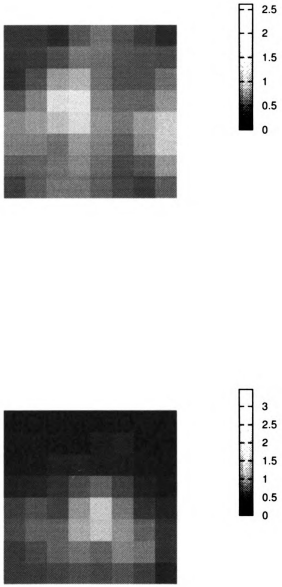

Figure 4.5. The spatial trends (averaging over time) of dead crow (upper) and human case data. The areas cover by these two datasets are not aligned geographically.

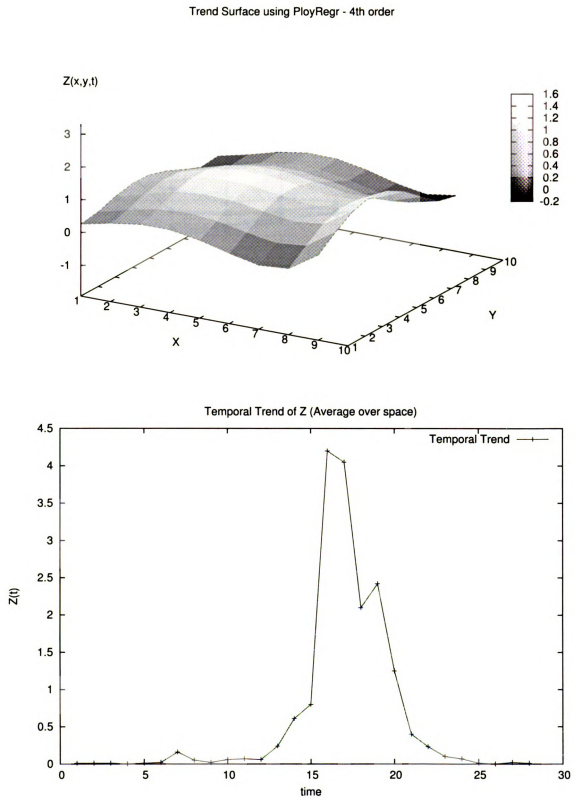

Figure 4.6. The spatial and temporal trends removed from the dead crow dataset.

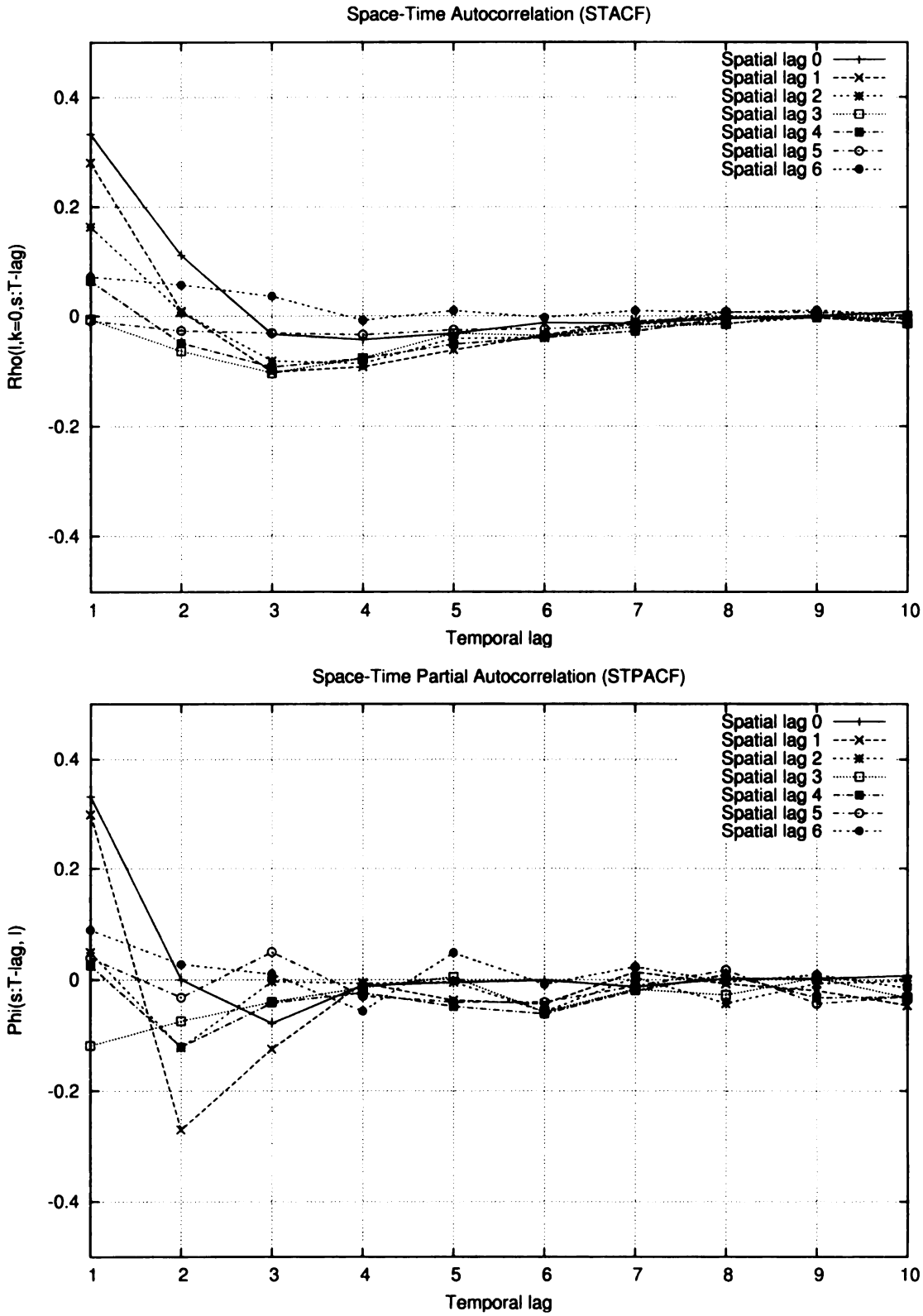

Figure 4.7. Space-time autocorrelation analysis of the de-trended dead crow dataset.

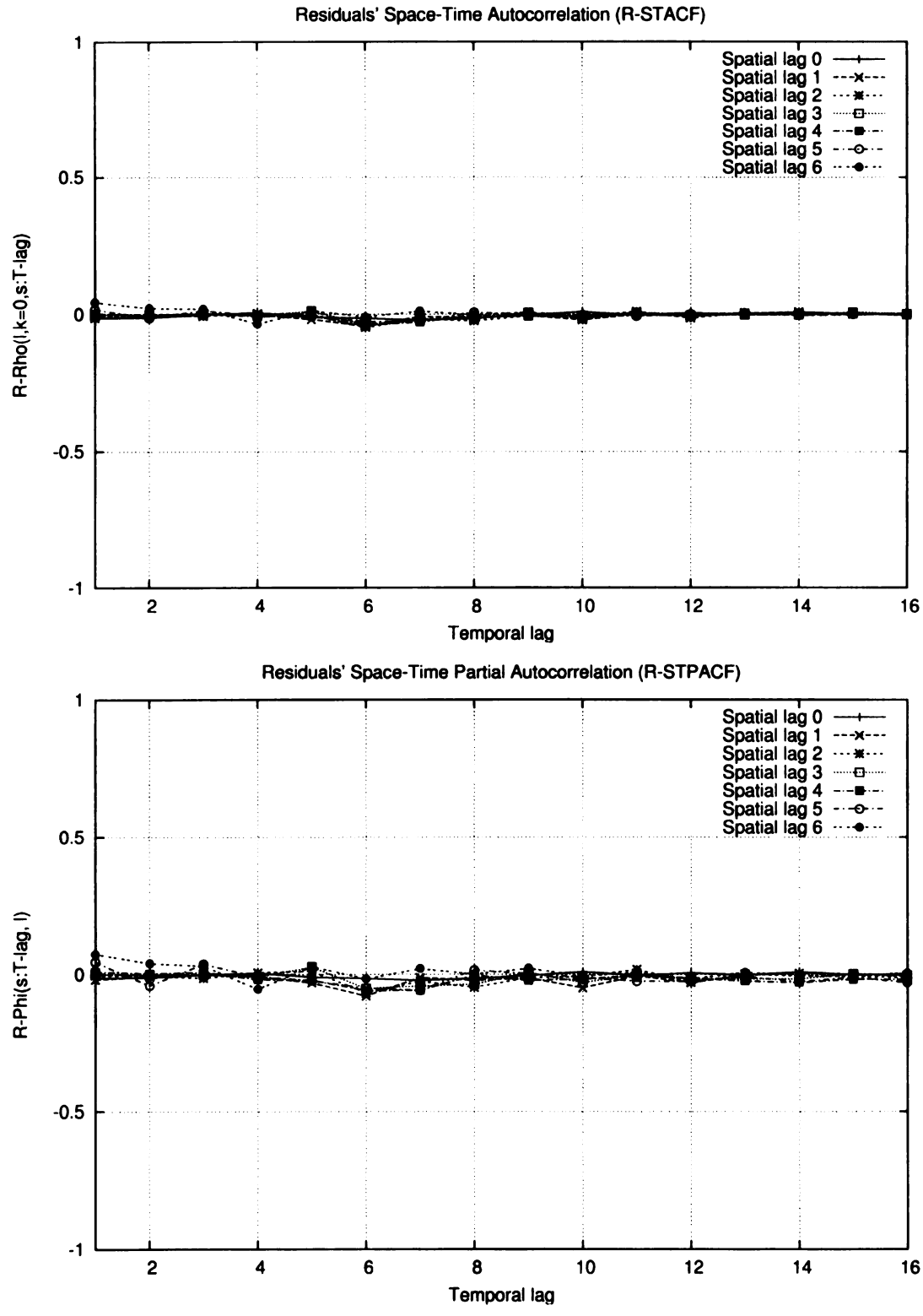

Figure 4.8. Space-time autocorrelation analysis for the modeling residuals of dead crow dataset

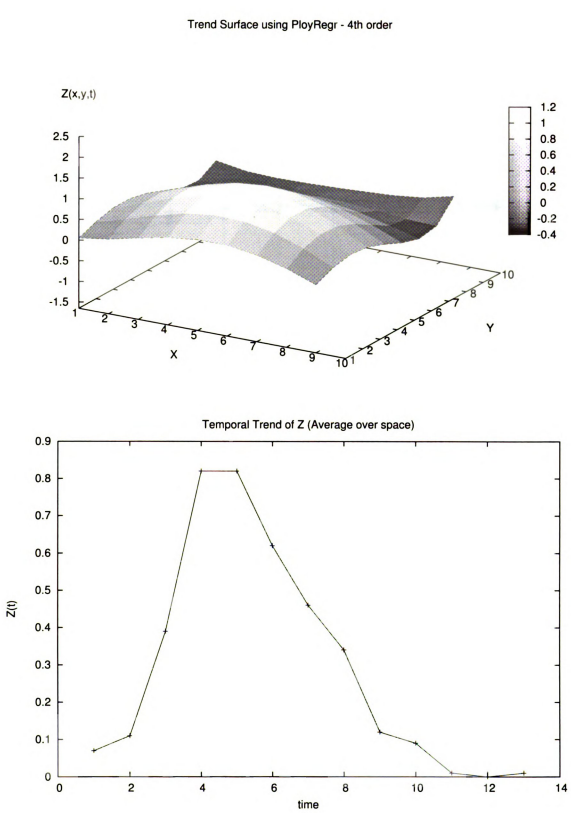

Figure 4.9. The spatial and temporal trends removed from the human case dataset.

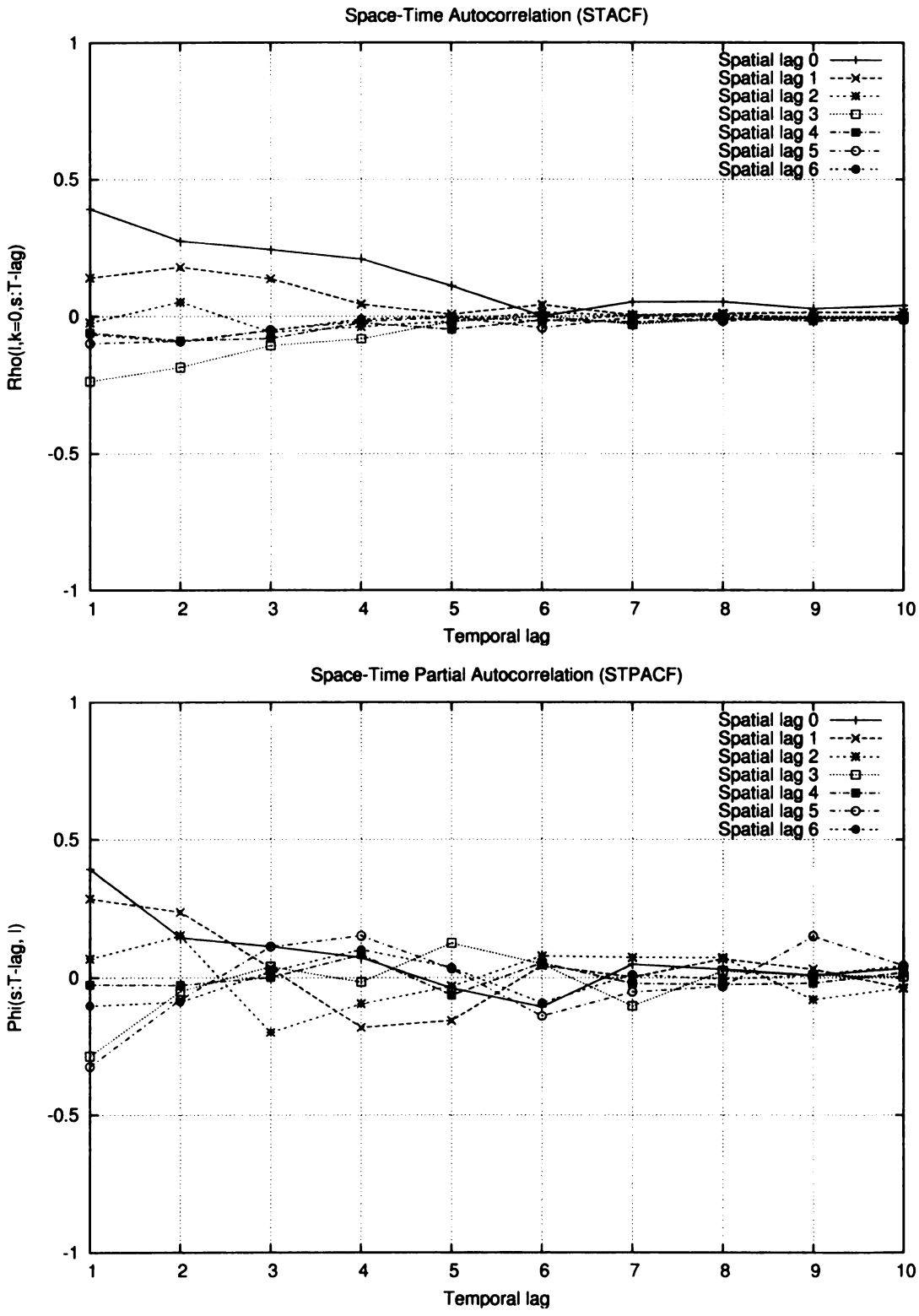

Figure 4.10. Space-time autocorrelation analysis for the de-trended human case dataset.

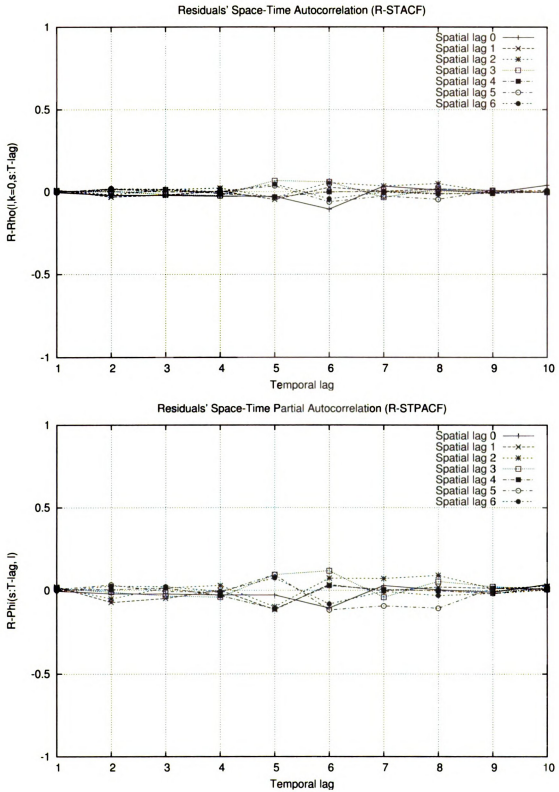

Figure 4.11. Space-time autocorrelation analysis for the modeling residuals of human case dataset

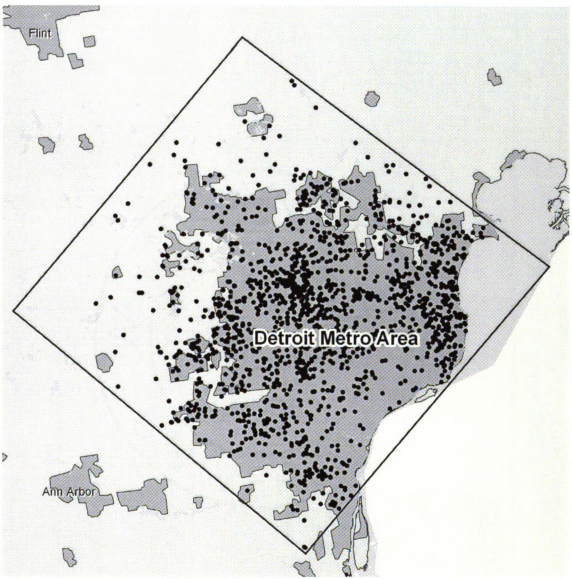

Figure 4.12. Dead crows reported and area for cross analysis.

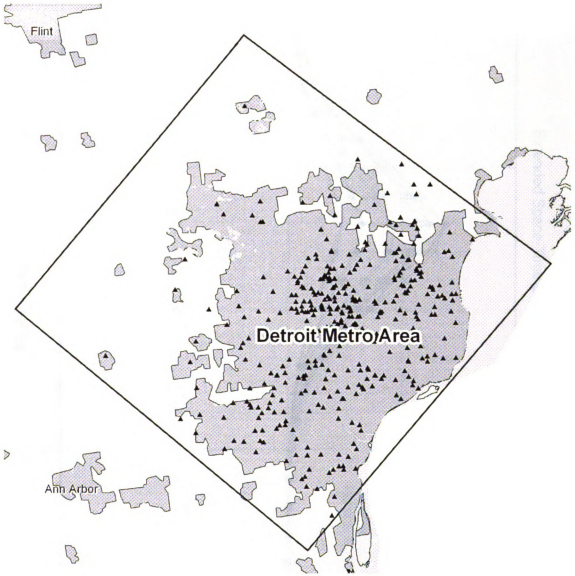

Figure 4.13. Human cases reported and area for cross analysis.

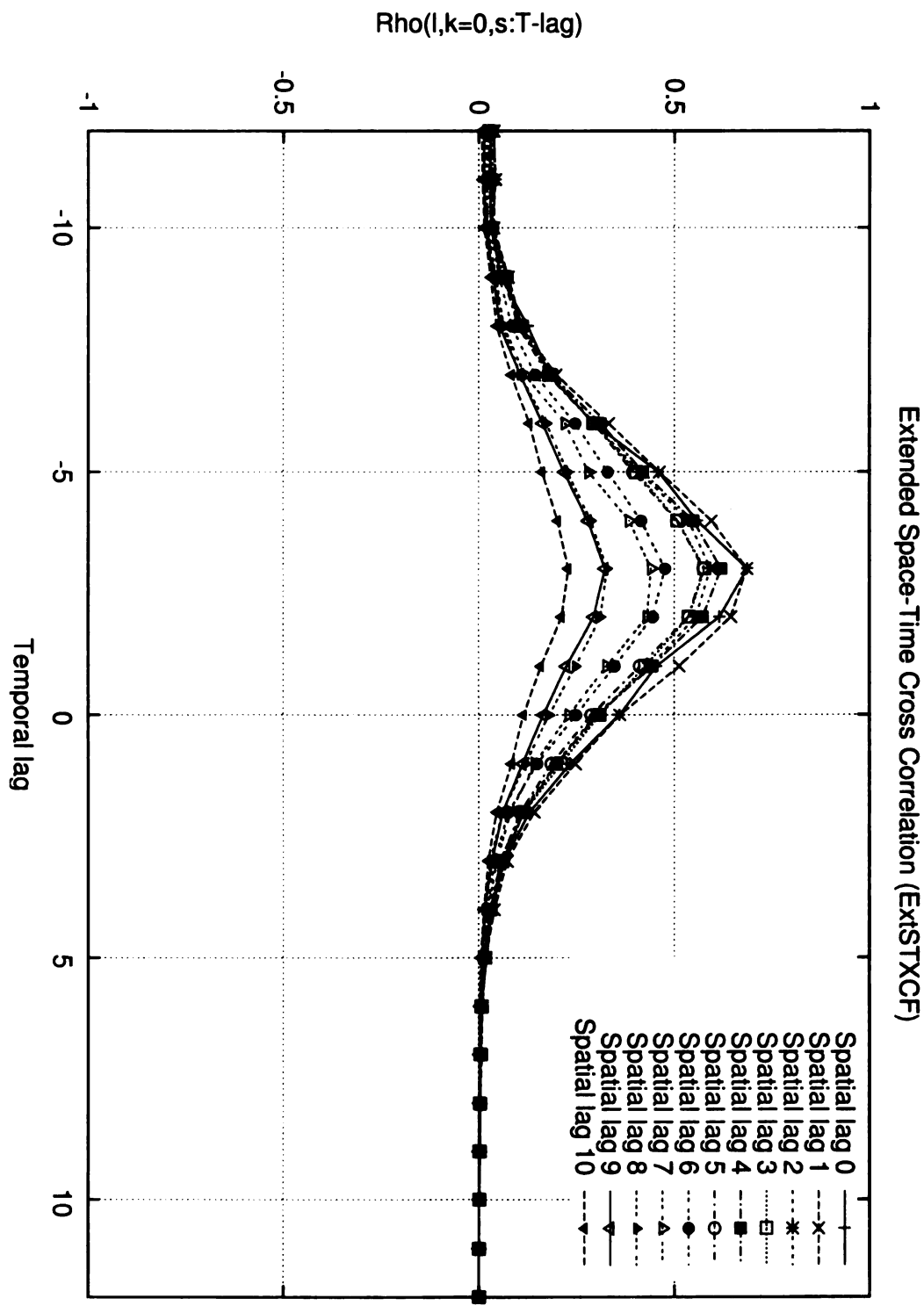

Figure 4.14. Space-time cross correlation between the human case data and the dead crow data.

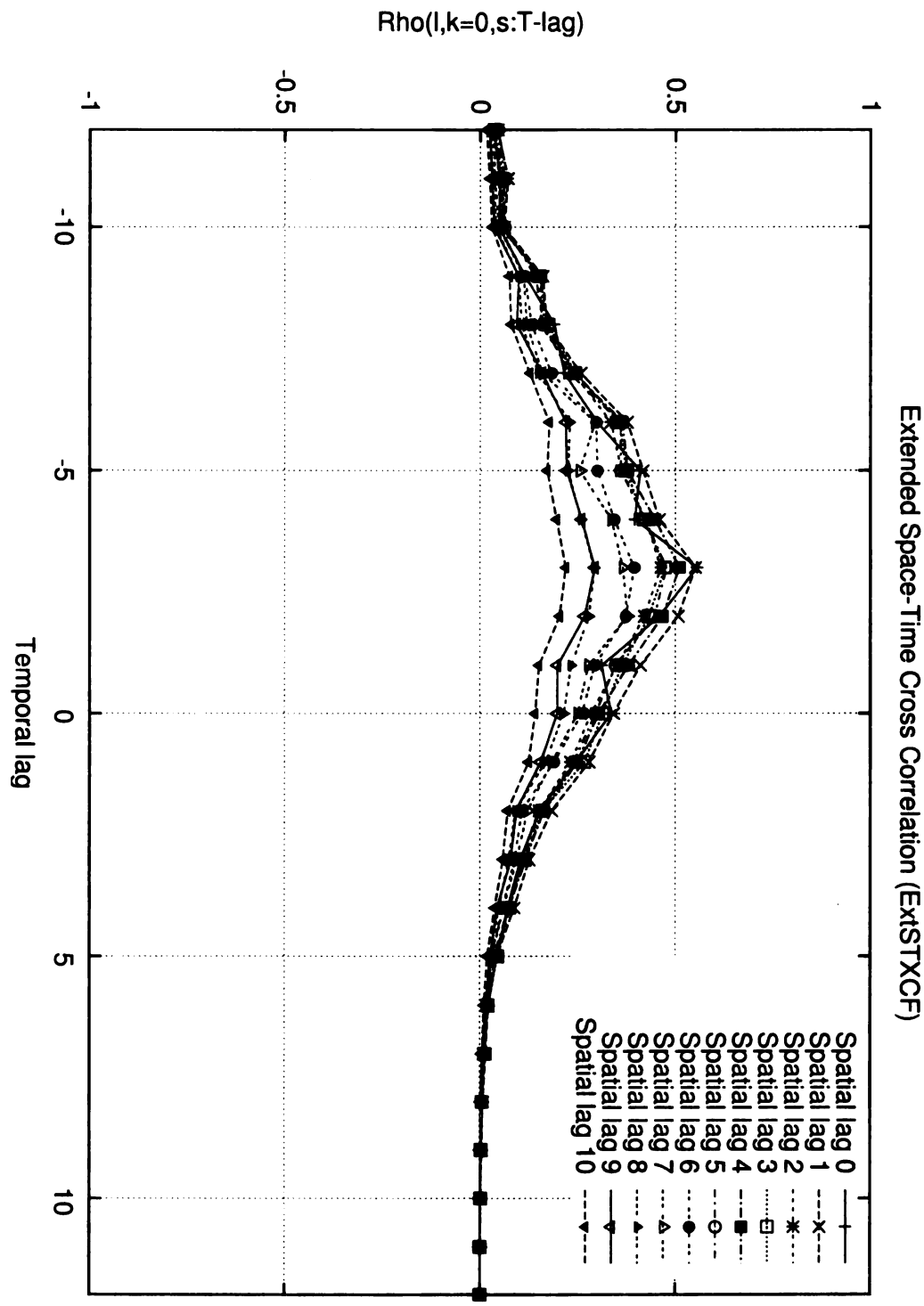

Figure 4.15. Space-time cross correlation between the filtered human case data and the filtered dead crow data.

CHAPTER 5

Conclusion

In this thesis, the application has demonstrated the effectiveness of STARMA modeling in analyzing and modeling space-time data, specifically in disease spreading. At the same time, it also provides a demonstration of using the refined modeling framework and related statistical tools/algorithms. If the programming language in IEAST is used, we can even automate the whole process of analysis and modeling, and increase efficiency and flexibility of computations.

However, there are some points or limitations to be kept in mind when modeling space-time data. First, the number of data points in space and time need to be 'sufficiently' large, or the random distribution (point pattern processes) of the given data can be quite different than the basic assumption of normal distribution in our STARMA model. In this research, several transformation methods (transforming to near normal distribution) have been applied to the DCD and human data to try to reduce the influence of the potential non-normality. Incidentally, after comparing the analysis results of transformed and non-transformed data, any differences are subtle.

This may result from the nature of autocorrelations, that is, the relative values or positions (in time or in space) are more important than the absolute values. Second, the STARMA model assumes that the value of the space-time variable for a given time and location is 'linearly' correlated with its past values in time and in neighboring locations. Although in reality there is no purely linear process, most of them are either close to or able to be transformed into linear process. If we know some process is highly non-linear, some methods have to be used to transform it into a linear one, or STARMA models are not appropriate for modeling such processes. Third, as mentioned in the basic assumptions, the correlation structure (in space and in time) of the process underlying the given data should be time-invariant. In fact, no model can represent a process with constantly changing characteristics. Fourth, as implied in the context, STARMA modeling using IEAST does not allow missing observations. For cells without samples or observations, its value simply is regarded as zero (e.g. no cases for the WNV modeling). If missing observations is an important issue, the dataset needs to be interpolated or extrapolated in advance with some other temporal or spatial methods, such as Krigging, smoothing, etc.

For future research, there are two directions for follow-up. First, anisotropic (directional) correlation structure can be an option while modeling. In my research, the spatial correlation structure (or weight matrices) was assumed and assigned to be uniform or directionless. Not every natural process is isotropic. It is especially interesting that an unknown spatial correlation structure can be estimated for a given space-time data. For example, with some epidemics, there may exist a significant tendency in some specific direction. Once the spatial correlation structure is retrieved

from the given data, it can be very useful in forecasting the spatial spreading tendency. In IEAST, there is a functionality for estimating spatial correlation structure based on a given dataset. However, for most cases, the number of parameters to be estimated for a correlation structure is too large and the computation needed intensive, involving too much computation to be applied to real data to get useful results. The work of developing algorithms or statistical tools to simplify this calculation is necessary. However, in general the model for forecasting human epidemic from dead crow data is very useful for prediction and forecast.

BIBLIOGRAPHY

[1] H. Akaike. Information theory and an extension of the maximum likelihood principle. In Petrov, B., F., and C., editors, *2nd International Symposium on Information Theory*, pages 267–281, Akademiai Kiado, Budapest, 1973.

[2] H. Akaike. Time series analysis and control through parametric models. In D. F. Findley, editor, *Applied Time Series Analysis*, pages 1-23, New York, 1978. Academic Press.

[3] L. A. Aroian. Time series in m-dimensions: autoregressive models. *Communications in Statistics - Simulation and Computation*, B9 (5):491-513, 1980.

[4] L. A. Aroian. Time series in m-dimensions definition, problems and prospects. *Communications in Statistics - Simulation and Computation*, B9 (5):453-465, 1980.

[5] R. J. Bennett. The representation and identification of spatio-temporal systems: an example of population diffusion in north-west england. *Transactions of the Institute of British Geographers*, 66:73-94, 1975.

[6] R. J. Bennett. *Spatial Time Series.* Pion Ltd, London, 1979.

[7] J. S. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society Series B*, 36:197-242, 1974.

[8] J.-P. Bocquet-Appel and R. R. Sokal. Spatial autocorrelation analysis of trend residuals in biological data. *Systematic Zoology*, 38, 4:331-341, 1989.

[9] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco, 1970.

[10] Centers for Disease Control and Prevention. Outbreak of West Nile-like viral encephalitis. *Morbidity Mortality Weekly Report*, 48:845-849, 1999.

[11] Centers for Disease Control and Prevention. Provisional surveillance summary of the West Nile virus epidemic - United States, January-November 2002. *Morbidity Mortality Weekly Report*, 51:1129-1133, 2002.

[12] A. D. Cliff, P. Haggett, J. K. Ord, K. A. Bassett, and R. B. Davies. *Elements of Spatial Structure: A Quantitative Approach.* Cambridge University Press, New York, 1975.

[13] A. D. Cliff and J. K. Ord. *Spatial autocorrelation.* London, 1973.

[14] A. D. Cliff and J. K. Ord. Model building and the analysis of spatial pattern in human geography. *Journal of the Royal Statistical Society Series B*, 37:297 348, 1975.

[15] A. D. Cliff and J. K. Ord. Space-time modelling with an application to regional forecasting. *Transactions of the Institute of British Geographers*, 64:119–128, 1975.

[16] A. D. Cliff and J. K. Ord. *Spatial processes: Models and applications.* Pion Ltd., London, 1981.

[17] N. Cressie. Comment on 'An approach to statistical spatial-temporal modeling of meterological fields' by M.S. Handcock and J.R. Wallis. *Journal of the American Statistical Association*, 89:379–382, 1994.

[18] Division of Vector-Borne Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention. West Nile virus statistics, surveillance, and control. 2004.

[19] J. Durbin. The fitting of time series models. *Review of the International Statistical Institute*, 28:233–243, 1960.

[20] M. Eidson, N. Komar, F. Sorhage, R. Melson, T. Talbot, and F. Mostashari. Crow deaths as a sentinel surveillance system for West Nile virus in the northeastern United States. *Emerging Infectious Diseases*, 7:615–620, 2001.

[21] M. Eidson, N. Komar, F. Sorhage, R. Melson, T. Talbot, and F. Mostashari. Crow deaths as a sentinel surveillance system for West Nile virus in the northeastern united states. *Emerging Infectious Diseases*, 7:615–620, 2001.

[22] M. Eidson, L. Kramer, W. Stone, Y. Hagiwara, K. Schmit, and New York State West Nile Virus Surveillance Team. Dead bird surveillance as an early warning system for West Nile virus. *Emerging Infectious Diseases*, 7:631–635, 2001.

[23] M. Eidson, M. Miller, and L. Kramer. Dead crow densities and human cases of West Nile virus, New York State. *Emerging Infectious Diseases*, 7:662–664, 2001.

[24] B. K. Epperson. Spatial and space-time correlations in ecological models. *Ecological Modelling*, 132:63- 76, 2000.

[25] B. K. Epperson. *Geographical Genetics.* Princeton University Press, New Jersey, 2003.

[26] S. Fitzgerald, J. Patterson, M. Kiupel, H. Simmons, S. Grimes, C. Sarver, R. Fulton, B. Steficek, T. Cooley, J. Massey, and J Sikarskie. Clinical and pathologic features of West Nile virus infection in native North American owls (Family Strigidae). *Avian Diseases*, 47(3):602 610, 2003.

[27] R. Giacomini and C. W. J. Granger. Aggregation of space-time processes. *Journal of Econometrics*, 118:7 26, 2001.

[28] S. C. Guptill, K. G. Julian, G. L. Campbell, S. D. Price, and A. A. Marfin. Early-season avian deaths from West Nile virus as warnings of human infection. *Emerging Infectious Diseases*, 9:483–484, 2003.

[29] E. J. Hannan and J. Rissanen. Recursive estimation of mixed auto-regressive moving average order. *Biometrika*, 69:81–94, 1982.

[30] P. M. Hooper and G. J. D. Hewing. Some properties of space-time processes. *Geographical Analysis*, 13:203–223, 1981.

[31] H. C. Huang and N. Cressie. Spatio-temporal prediction of snow water equivalent using the kalman filter. *Computational Statistics and Data Analysis*, 22:159 175, 1996.

[32] C. M. Hurvich and C.-L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76:297–307, 1989.

[33] R. H. Jones. Fitting autoregressions. *Journal of the American Statistical Association*, 70 (351):590 592, 1975.

[34] N. Komar. West Nile viral encephalitis. *Revue Scientifique et Technique*, 19:166 176, 2000.

[35] N. Komar, S. Langevin, S. Hinten, N. Nemeth, Edwards E., D. Hettler, B. Davis, R. Bowen, and Michel Bunning. Experimental infection of North American birds with the New York 1999 strain of West Nile viral. *Emerging Infectious Diseases*, 9:311-322, 2003.

[36] H. Lutkerpohl. *Forecasting Aggregated Vector ARMA Processes.* Springer-Verlag, Berlin, 1987.

[37] H. Lutkerpohl. *Introduction to Multiple Time Series Analysis.* Springer-Verlag, Berlin, 1993.

[38] K. V. Mardia, C. R. Goodall, E. Redfern, and F. Alonso. The kriged kalman filter. *Test*, 7:217–285, 1998.

[39] D. W. Marquardt. An algorithm for least squares estimation of nonlinear parameters. *Journal of the Society of Industrial and Applied Mathematics*, 11:431–441, 1963.

[40] R. L. Martin and J. E. Oeppen. The identification of regional forecasting models using space-time correlation functions. *Transactions of the Institute of British Geographers*, 66:95–118, 1975.

[41] R. G. McLean, S. R. Ubico, D. E. Docherty, W. R. Hansen, L. Sielo, and T. S. McNamera. West Nile virus transmission and ecology in birds. *Annals of the New York Academy of Sciences*, 951:54–57, 2001.

[42] C. A. Oprian, V. S. Taneja, D. A. Voss, and L. A. Aroian. General considerations and interrelationships between MA and AR models, time series in m dimensions, the ARMA model. *Communications in Statistics - Simulation and Computation*, B9 (5):515–532, 1980.

[43] L. Petersen and J. Roehrig. West Nile virus: a reemerging global pathogen. *Emerging Infectious Diseases*, 7(4):611–614, 2001.

[44] P. E. Pfeifer. *Spatial-Dynamic Modeling (Unpublished Ph.D. dissertation)*. Georgia Institute of Technology, Atlanta, Georgia, 1979.

[45] P. E. Pfeifer and S. J. Deutsch. A comparison of estimation procedures for the parameters of the STAR model. *Communications in Statistics - Simulation and Computation*, B9 (3):255–270, 1980.

[46] P. E. Pfeifer and S. J. Deutsch. Independence and sphericity test for the residuals of space-time ARMA models. *Communications in Statistics - Simulation and Computation*, B9 (5):533–549, 1980.

[47] P. E. Pfeifer and S. J. Deutsch. Stationarity and invertibility regions for low order STARMA models. *Communications in Statistics - Simulation and Computation*, B9 (5):551–562, 1980.

[48] P. E. Pfeifer and S. J. Deutsch. A three-stage iterative procedure for space-time modeling. *Technometrics*, 22 (1):35–47, 1980.

[49] P. E. Pfeifer and S. J. Deutsch. Variance of the sample space-time autocorrelation function. *Journal of the Royal Statistical Society Series B*, 43 (1):28–33, 1981.

[50] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6 (2):461–464, 1978.

[51] R. Shibata. Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*, 63:117–126, 1976.

[52] C. N. Theophilides, S. C. Ahearn, S. Grady, and M. Merlino. Identifying West Nile virus risk areas: The dynamic continuous-area space-time system. *American Journal of Epidemiology*, 157:843–854, 2003.

[53] M. J. Turell, M. O'Guinn, and J. Oliver. Potential for New York mosquitoes to transmit West Nile virus. *American Journal of Tropical Medicine and Hygiene*, 62:413–414, 2000.

[54] M. J. Turell, M. R. Sardelis, and D. J. Dohm. Potential North American vectors of West Nile virus. *Annals of the New York Academy of Sciences*, 951:317–324, 2001.

[55] D. A. Voss, C. A. Oprian, and L. A. Aroian. Moving average models - time series in m-dimensions. *Communications in Statistics - Simulation and Computation*, B9 (5):467–489, 1980.

[56] J. Watson, R. Jones, K. Gibbs, and W. Paul. Dead crow reports and location of human West Nile virus cases, Chicago, 2002. *Emerging Infectious Diseases*, 10(5):938–940, 2004.

[57] S. A. Yaremych, R. J. Novak, A. Raim, P. Mankin, and R. E. Warner. Home range and habitat use by American Crows in relation to West Nile virus transmission. *The Wilson Bulletin (in press)*, 2004a.

[58] S. A. Yaremych, R. E. Warner, P. Mankin, J. Brown, and R. J. Novak. West Nile virus and high death rate in American Crows. *Emerging Infectious Diseases*, 10(4):709–711, 2004b.

[59] S. A. Yaremych, R. E. Warner, M. T. Wyngaerde, A. M. Ringia, R. Lampman, and R. J. Novak. West Nile virus detection in American Crows. *Emerging Infectious Diseases*, 9:1319–1321, 2003.