# SPATIO-TEMPORAL AUTOREGRESSIVE MODELS FOR U.S. UNEMPLOYMENT RATE

Xavier de Luna and Marc G. Genton

## ABSTRACT

*We analyze spatio-temporal data on U.S. unemployment rates. For this purpose, we present a family of models designed for the analysis and time-forward prediction of spatio-temporal econometric data. Our model is aimed at applications with spatially sparse but temporally rich data, i.e. for observations collected at few spatial regions, but at many regular time intervals. The family of models utilized does not make spatial stationarity assumptions and consists in a vector autoregressive (VAR) specification, where there are as many time series as spatial regions. A model building strategy is used that takes into account the spatial dependence structure of the data. Model building may be performed either by displaying sample partial correlation functions, or automatically with an information criterion. Monthly data on unemployment rates in the nine census divisions of the U.S. are analyzed. We show with a residual analysis that our autoregressive model captures the dependence structure of the data better than with univariate time series modeling.*

**Spatial and Spatiotemporal Econometrics**
**Advances in Econometrics, Volume 18, 279–294**

**ISSN: 0731-9053/doi:10.1016/S0731-9053(04)18009-2**

## 1. INTRODUCTION

In this article, we analyze spatio-temporal data on U.S. unemployment rates. Previous studies have often focused on purely temporal models of a single measure in the study of unemployment: the U.S. civilian unemployment rate, seasonally adjusted, see e.g. Montgomery et al. (1998), Proietti (2003). We, instead, develop a spatio-temporal model for monthly U.S. unemployment rates observed in the nine census divisions of the United States between January 1980 and May 2002. Figure 1 presents a map of the census regions and divisions of the U.S. The nine regions are New England (NE), Middle Atlantic (MA), South Atlantic (SA), East North Central (ENC), East South Central (ESC), West North Central (WNC), West South Central (WSC), Mountain (M), and Pacific (P). These data have two important characteristics. First, the geographical locations of the nine divisions form a spatial lattice (see, e.g. Haining, 1990). This means that the unemployment rate in a given

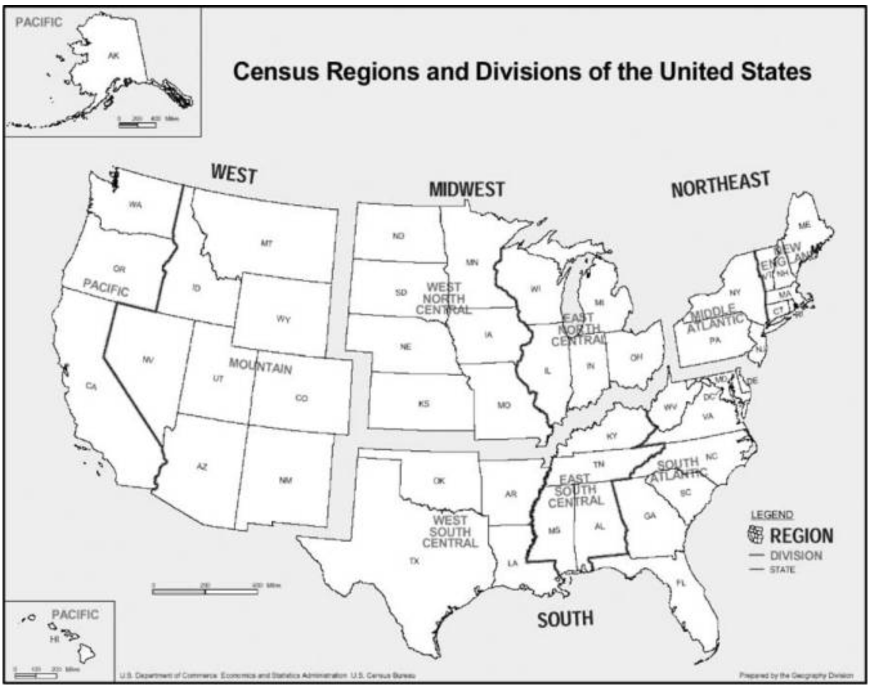


*Fig. 1.* Map of the Census Regions and Divisions of the U.S., Reproduced with Permission from the U.S. Census Bureau.

division might be spatially correlated with the rates in neighboring regions. This spatial information can be used in our model to improve the forecasts. Note that, unlike environmental applications with monitoring stations, spatial interpolation is not of interest in our context. Second, the spatio-temporal data are sparse in space but rich in time, that is, we have only nine spatial regions, but many observations at regular time intervals. This type of data can also be found in environmental studies, see de Luna and Genton (2003) for applications to wind speed data and carbon monoxide atmospheric concentrations.

The article is set up as follows. Section 2 describes a family of autoregressive models with spatial structure, first introduced by de Luna and Genton (2003) in the context of environmental applications. We argue that spatial stationarity assumptions are not necessary and a different model can be build for each of the nine divisions. We briefly discuss estimation and inference for these models and also provide two main approaches to deterministic trend modeling. We then describe our model building strategy which is based on a spatio-temporal ordering of the nine divisions. Section 3 presents two analyses of the U.S. unemployment rate data: one is based on univariate time series modeling, whereas the second uses the spatial information across divisions. We perform a residual analysis on the two approaches and show that our autoregressive model with spatial structure captures the dependence structure of the data better. We conclude in Section 4.

## 2. AUTOREGRESSIVE MODELS WITH SPATIAL STRUCTURE

In this section we present models that are specifically designed for the analysis of spatial lattice data evolving in time. Our purpose is to provide time-forward predictions at given spatial regions of the lattice, based on a minimum of assumptions.

### *2.1. The Model*

The model we consider is a vector autoregressive (VAR) model commonly used in multivariate time series analysis (e.g. Lütkepohl, 1991). It is usually applied in situations where several variables are observed at the same time. We, however, use the VAR model for a single variable, but observed at several spatial regions of a lattice at the same time. Specifically, we consider observations $z(\mathbf{s}_i, t)$ collected at $\mathbf{s}_i$, $i = 1, \ldots, N$ spatial regions of a lattice and $t = 1, \ldots, T$ times. A time-forward predictive model for $z_t = (z(\mathbf{s}_1, t), \ldots, z(\mathbf{s}_N, t))'$ is

$$z_t - \boldsymbol{\beta} = \sum_{i=1}^{p} R_i(z_{t-i} - \boldsymbol{\beta}) + \boldsymbol{\varepsilon}_t, \tag{1}$$

a VAR($p$) model with spatial lattice structure, where $p$ denotes the order of the autoregression in time. The parameter $\boldsymbol{\beta} = (\beta(\mathbf{s}_1), \ldots, \beta(\mathbf{s}_N))'$ is a vector of spatial effects representing a spatial trend among the regions of the lattice. The $N$-dimensional process $\boldsymbol{\varepsilon}_t$ is white noise, that is, $E(\boldsymbol{\varepsilon}_t) = 0$, $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \Sigma_\varepsilon$, and $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_u') = 0$ for $u \neq t$. The $N \times N$ parameter matrices $R_i$, $i = 1, \ldots, p$, are unknown and need to be estimated from the data, as well as the unknown $N \times N$ matrix $\Sigma_{\boldsymbol{\varepsilon}}$ of spatial covariances. The estimation of those parameters can be carried out with maximum likelihood (with distributional assumptions, typically Gaussian), with least squares or with moments estimators of Yule-Walker type, see, e.g. Lütkepohl (1991, Chap. 3). Note that the order $p$ is also unknown and its identification in the space-time context will be discussed in the next section.

We assume that iterations of the deterministic dynamic system defined by model (1) converge towards a constant. This stability property implies that the process $z_t$ is time stationary, that is $E(z_t) = \boldsymbol{\beta}$, for all $t$, and $\text{cov}(z_t, z_{t-\tau}) = \Gamma_z(\tau)$, a function of $\tau$ only, for all $t$ and $\tau = 0, 1, 2, \ldots$. The covariance matrix $\Gamma_z(\tau)$ can then be computed from the parameter matrices $R_1, \ldots, R_p$ and from $\Sigma_{\boldsymbol{\varepsilon}}$, see Lütkepohl (1991). Note that the matrix $\Sigma_{\boldsymbol{\varepsilon}}$ can represent nonstationary spatial covariances. Indeed, we do not assume any spatial stationarity for model (1), such as that spatial correlations depend only on the distance between stations (isotropy). Such assumptions have often been made in the spatio-temporal literature, for instance to develop space-time ARMA models, or STARMA, see Pfeifer and Deutsch (1980), Stoffer (1986). For the applications to space-time lattice data, spatial stationarity is an over-restrictive assumption which is, moreover, difficult to check in practice.

Because model (1) is assumed to be stationary in time, we need to remove deterministic temporal trends from the data. This can be achieved by differencing the observations with respect to time, as is commonly done in classical time series analysis. If the result depends only on the spatial regions, then it can be modeled by the spatial trend parameter $\boldsymbol{\beta}$ in (1). This will happen as soon as the spatio-temporal trend is a polynomial function in time with coefficients that may depend on the spatial location. In practice, this is typically the case, at least approximately. Note also that seasonal effects, such as monthly records, can be removed by taking differences in time. Another approach consists in modeling a deterministic spatio-temporal trend by means of a weighted sum of known basis functions, such as polynomials and/or sine/cosine periodic functions. The weights are then estimated by regression. Further discussions on the topic of spatio-temporal trend modeling can be found in the review article by Kyriakidis and Journel (1999).

## 2.2. *Model Building*

In this section, we present a model building strategy for the VAR model (1). Indeed, the spatial lattice data impose a specific structure on the unknown parameter matrices $R_1, \ldots, R_p$. Our strategy consists in identifying zero entries in the $R_i$'s matrices as well as the order $p$ of the autoregression. To this aim, we model each spatial region $\mathbf{s}$ of the lattice separately, thus avoiding any assumption of spatial stationarity. This can be seen as a complex covariate selection problem, where, for the response $z(\mathbf{s}, t)$, the available predictors are the time-lagged values at all spatial regions of the lattice, i.e. $z(\mathbf{s}_i, t-j), i = 1, \ldots, N$ and $j = 1, \ldots$. We make use of the spatio-temporal structure to define an ordering of the predictors, thus allowing to introduce them sequentially in our model, as is commonly performed in univariate time series analysis. A natural ordering of the spatial regions around $\mathbf{s}$, is given by the sequence of regions sorted in ascending order with respect to their distance to $\mathbf{s}$. In Section 3.2 we define such a distance based ordering for the U.S. census divisions. For instance, for the region ENC the ordering obtained is ENC, MA, WNC, ESC, SA, NE, WSC, M and P; see Table 1.

Other orderings can be considered as well. For instance, orderings based on the length of common border between two spatial regions (see Haining, 1990) or orders motivated by dynamical/physical knowledge about the underlying process, see de Luna and Genton (2003) for an application to wind speeds in Ireland.

From an ordering of the spatial regions one obtains a spatio-temporal ordering by considering first the predictors $z(\mathbf{s}_i, t-j)$ at lag one ($j = 1$) in the spatially defined

***Table 1.*** Matrix of "Distance" Used in the Spatio-Temporal Ordering.

| | NE | MA | ENC | WNC | SA | ESC | WSC | M | P |
|---|---|---|---|---|---|---|---|---|---|
| NE | 0 | 1 | 2 | 5 | 3 | 4 | 6 | 7 | 8 |
| MA | 1 | 0 | 2 | 5 | 3 | 4 | 6 | 7 | 8 |
| ENC | 5 | 1 | 0 | 2 | 4 | 3 | 6 | 7 | 8 |
| WNC | 7 | 5 | 1 | 0 | 6 | 4 | 2 | 3 | 8 |
| SA | 3 | 2 | 4 | 6 | 0 | 1 | 5 | 7 | 8 |
| ESC | 6 | 4 | 3 | 5 | 1 | 0 | 2 | 7 | 8 |
| WSC | 7 | 6 | 4 | 2 | 5 | 1 | 0 | 3 | 8 |
| M | 7 | 6 | 4 | 1 | 8 | 5 | 2 | 0 | 3 |
| P | 7 | 6 | 4 | 2 | 8 | 5 | 3 | 1 | 0 |

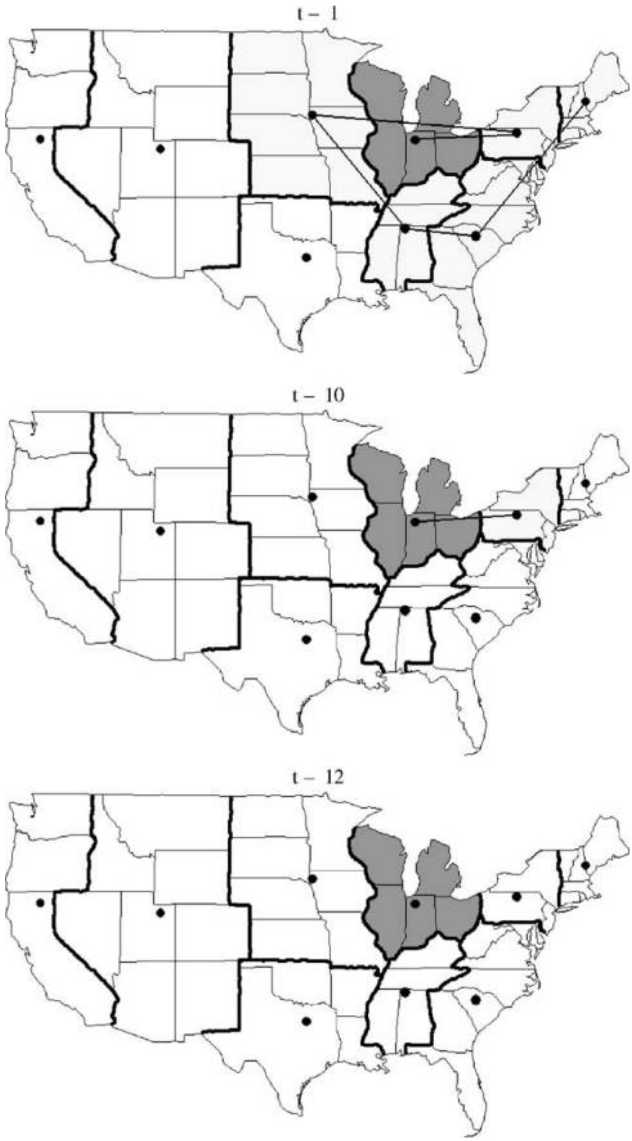


*Fig*. *2*. A Schematic Representation of the Spatio-Temporal Ordering of the Neighbors to the Spatial Division ENC for Time $t-1$, $t-10$, $t-12$. *Note:* The centroid of each census region is represented by a black disc. The segments connecting those discs represent the sequence of regions sorted in ascending order with respect to their distance to ENC (Table 1). Only the predictors with predictive power for ENC unemployment at time $t$ are connected.

order, then lag 2 ($j = 2$), etc. With a given spatio-temporal ordering, we can enter our predictors sequentially in our model and stop whenever the partial correlation between $z(\mathbf{s}, t)$ and a new predictor is zero, conditionally on all other predictors already in the model. This means that we are defining partial autocorrelations (PCF) along our ordering in space and time, see de Luna and Genton (2003) for details. This approach can be automated by using model selection criteria such as AIC (Akaike, 1974) or BIC (Schwarz, 1978). Figure 2 shows which neighbors (shaded states) are included at lag $t - 1$, $t - 10$ and $t - 12$ in the model built (see Section 3.2) to predict ENC unemployment at time $t$. The segments connecting the centroids (black discs) of the divisions highlight the ordering of the regions described above.

## 3. UNEMPLOYMENT RATES IN THE U.S.

In this section we build predictive models for unemployment rates observed monthly between January 1980 and May 2002 in the nine census divisions of the United States. Each division consists in a collection of states as shown in Fig. 1. The data was downloaded from the web site of the Bureau of Labor Statistics (http://www.bls.gov). The unemployment rates are plotted for each division in Fig. 3.

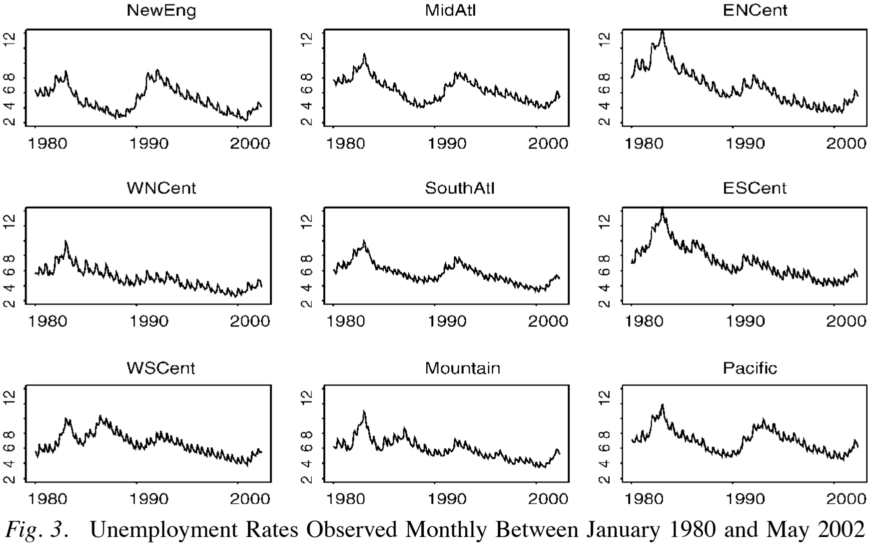


*Fig. 3.* Unemployment Rates Observed Monthly Between January 1980 and May 2002 for the Nine Census Divisions of the U.S.

### *3.1. Univariate Modeling*

We start with a purely time series data analysis where each division is considered separately with the purpose to fit a univariate autoregressive model. From Fig. 3 we see that unemployment rates show a long term cyclic (trend-cycle) behavior. We also observe the well known monthly seasonality.

We begin by removing the trend-cycle with the difference operator of order one. Note here that it is not customary to take first differences of unemployment series, because the long term cyclic behavior is often of interest to economists. In particular, the existence of a natural rate of unemployment around which the unemployment series fluctuates is often studied and discussed (e.g. Lucas, 1973; Phelps, 1994). In this paper, however, we focus on models geared towards the purpose of providing short term predictions. Thus, it is natural to analyze monthly changes in unemployment rates (first differences) from which the (hypothetical) natural rate has been removed.

The detrended series are shown in Fig. 4. From these plots and the autocorrelation functions displayed in Fig. 5 it is clear that we have a twelve month seasonal component. Thus, a seasonal differencing is carried out. The resulting series and their autocorrelation functions are displayed in Figs 6–8, respectively. The two autocorrelation functions (plain and partial) indicate that it is suitable to fit AR models to the different time series.

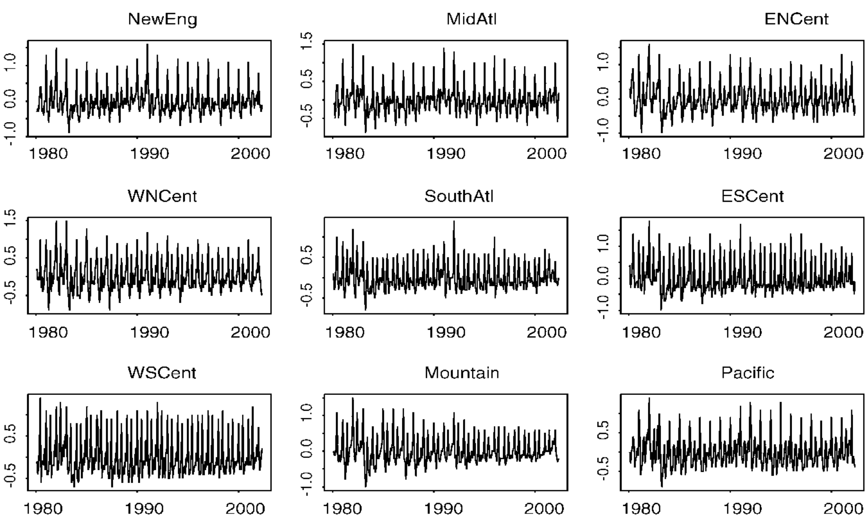


*Fig. 4.* First Differences of the Unemployment Series Showed in Fig. 3.

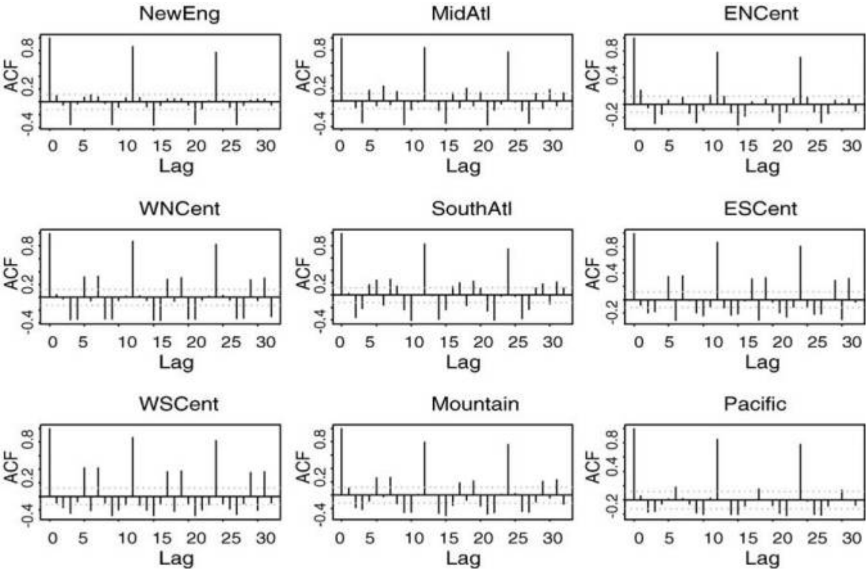


*Fig. 5.* Autocorrelation Functions of the Detrended Unemployment Series of Fig. 4.

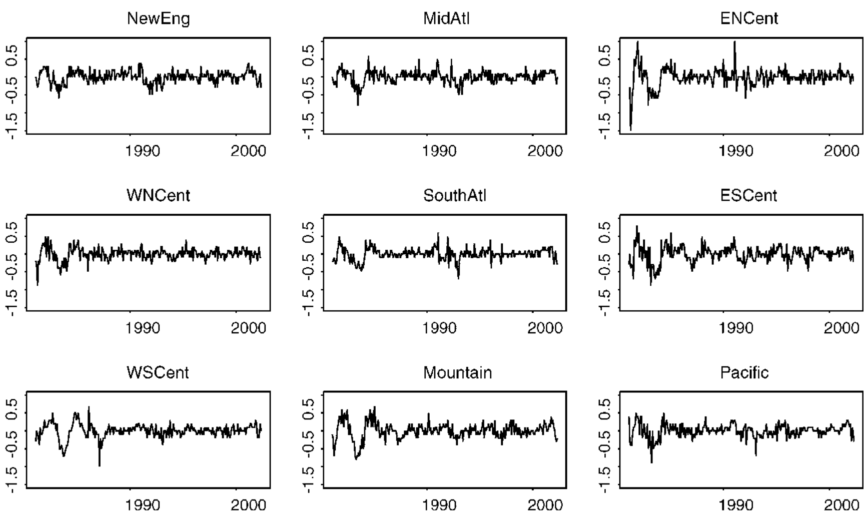


*Fig. 6.* Twelve Month Differences of the Detrended Unemployment Series in Order to Eliminate the Seasonal Component.

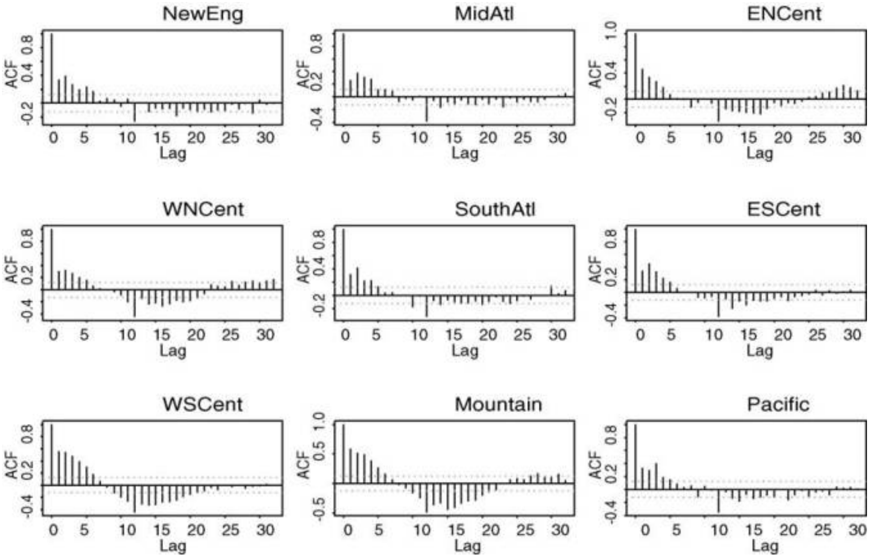


*Fig. 7.* Autocorrelation Functions of the Detrended and Deseasonalized Series of Fig. 6.

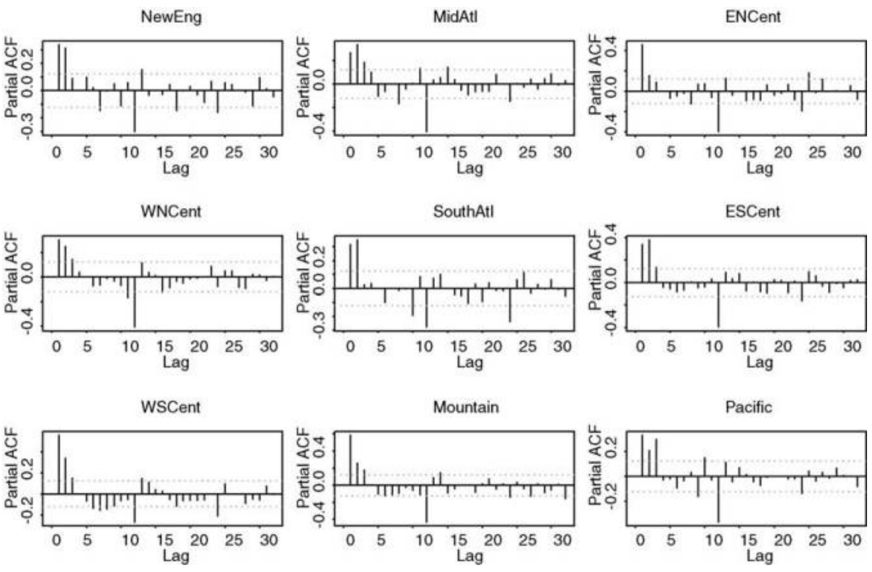


*Fig. 8.* Partial Autocorrelation Functions of the Detrended and Deseasonalized Series of Fig. 6.

### *3.2. Models with Spatial Structure*

The analysis is here geared towards building a VAR model with spatial structure as was described in Section 2.2. For this purpose the data must be stationary in the time dimension. From the previous univariate analysis we take that the nine time series must be detrended and deseasonalized. This is done by differencing as described above.

Model building is then performed on a division basis. That is for each division, a predictive model is built independently of the models considered for the other divisions. As mentioned earlier this is an essential component of our modeling approach since it allows us to avoid spatial stationarity assumptions.

We give the details of model building for the East North Central (ENC) division. A spatio-temporal ordering of the divisions is used to have a hierarchy in which to introduce covariates in the regression model. Such an ordering is most naturally constructed by considering geographical locations of the divisions, thereby anticipating that neighboring divisions will have a larger predictive power for ENC than far away regions. However, because divisions are not punctual locations but regions with an area, it is not clear how to define geographical proximity. On the other hand, a very precise ordering is not necessary since spatio-temporal neighbors with predictive power will eventually be included in the model. In other words, the ordering is used to help identifying zeros in the final VAR model, but without the ambition to identify them all.

We, therefore, take an arbitrary approach and define an ordering by considering first neighbors with a direct "frontier" to the ENC division, starting from East and proceeding anti-clockwise. Neighbors without direct frontier are then considered in approximate increasing distance, see Table 1 where each row represents a spatial ordering (from 0 to 9).

Table 2 displays the partial correlation function (PCF) defined with respect to the ordering described above. The table also indicates which partial correlations are significantly different from zero at the one percent level. This allows us to build a model with the relevant covariates. In this case, we should include at lag one the divisions ENC, MA, WNC, ESC, SA, and NE, because the latter is the last one being significant. Moreover, we should include ENC at lag 12. We consider the remaining significant partial correlations as spurious.

A weakness of the above approach is that we are performing many tests simultaneously and do not have control over the overall size of the test. However, as mentioned in Section 2.2 model building may be performed automatically by using an information criterion such as AIC or BIC. BIC (with max lag 15) selects ENC, MA, WNC, ESC, SA, NE at lag one, ENC, MA at lag ten, and ENC at lag twelve. This is in accordance with the above PCF analysis and is illustrated

*Table 2.* PCF for the ENC Division for Lag 1–25.

| Lag | ENC | MA | WNC | ESC | SA | NE | WSC | M | P |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.463** | 0.046 | 0.196** | 0.143 | 0.113 | 0.210** | −0.002 | 0.158 | 0.076 |
| 2 | 0.045 | 0.042 | 0.085 | 0.010 | 0.096 | −0.053 | −0.060 | 0.141 | −0.077 |
| 3 | 0.033 | −0.035 | 0.090 | −0.015 | 0.000 | −0.043 | 0.031 | 0.105 | 0.070 |
| 4 | −0.094 | 0.058 | 0.022 | 0.005 | 0.065 | 0.120 | −0.037 | −0.069 | 0.041 |
| 5 | −0.142 | 0.037 | 0.048 | −0.044 | 0.046 | 0.119 | −0.080 | 0.055 | 0.065 |
| 6 | −0.121 | 0.003 | 0.036 | −0.085 | −0.142 | −0.015 | −0.236** | −0.034 | −0.024 |
| 7 | 0.018 | 0.070 | 0.166 | 0.122 | −0.082 | 0.064 | −0.140 | −0.018 | −0.105 |
| 8 | −0.098 | 0.053 | 0.056 | 0.090 | 0.078 | −0.129 | −0.076 | −0.040 | −0.062 |
| 9 | −0.001 | 0.131 | 0.044 | 0.022 | 0.050 | 0.177 | −0.018 | −0.074 | 0.004 |
| 10 | 0.110 | −0.261** | −0.077 | 0.012 | 0.030 | 0.091 | −0.189 | −0.099 | 0.225** |
| 11 | 0.007 | 0.059 | 0.078 | −0.063 | −0.011 | −0.007 | −0.141 | 0.089 | −0.070 |
| 12 | −0.419** | 0.101 | −0.094 | 0.059 | −0.022 | −0.067 | 0.095 | 0.061 | 0.053 |
| 13 | 0.096 | 0.064 | 0.187 | 0.128 | −0.131 | 0.099 | −0.027 | −0.089 | 0.043 |
| 14 | −0.084 | −0.062 | 0.112 | −0.046 | 0.212 | −0.040 | 0.137 | 0.020 | 0.210 |
| 15 | 0.039 | 0.007 | 0.023 | −0.149 | 0.182 | −0.049 | 0.016 | −0.091 | −0.122 |
| 16 | −0.095 | −0.053 | 0.006 | 0.064 | 0.040 | 0.122 | −0.026 | 0.221 | −0.118 |
| 17 | −0.012 | 0.149 | 0.003 | −0.128 | 0.022 | 0.106 | −0.140 | −0.004 | 0.024 |
| 18 | −0.049 | 0.034 | −0.056 | 0.147 | 0.162 | 0.169 | −0.027 | 0.095 | −0.041 |
| 19 | −0.116 | −0.058 | −0.025 | 0.063 | −0.063 | −0.130 | −0.153 | 0.088 | −0.059 |
| 20 | 0.030 | 0.141 | 0.061 | 0.128 | −0.111 | 0.125 | −0.130 | 0.227 | 0.148 |
| 21 | −0.167 | 0.200 | −0.326** | 0.031 | 0.072 | −0.278 | 0.157 | 0.070 | 0.007 |
| 22 | 0.246 | 0.110 | −0.144 | 0.347** | −0.085 | −0.240 | 0.140 | −0.318 | 0.082 |
| 23 | −0.035 | 0.105 | −0.048 | −0.033 | 0.120 | −0.194 | −0.040 | 0.081 | 0.249 |
| 24 | 0.217 | −0.236 | −0.139 | 0.290 | −0.427** | −0.327 | −0.123 | −0.416** | 0.002 |
| 25 | 0.059 | −0.344 | −0.365 | 0.001 | −0.291 | 0.076 | 0.216 | −0.707** | 0.247 |

*Note:* ** indicate correlations significantly different from zero at the one percent level.

***Table 3.*** Order (Chosen by BIC) of the Autoregressive Models Fitted for Each Division.

| Division | NE | MA | ENC | WNC | SA | ESC | WSC | M | P |
|---|---|---|---|---|---|---|---|---|---|
| Order | 13 | 15 | 13 | 12 | 14 | 13 | 14 | 14 | 12 |

graphically in Fig. 2. In this automatic procedure we have chosen to include ENC at all lags.

### *3.3. Residual Analysis*

A residual analysis is carried out to investigate the relevance of the fitted models. In particular we look at the residuals obtained from a univariate modeling of the nine divisions, and the residuals obtained from a VAR modeling based on a spatio-temporal ordering.

The univariate modeling of the detrended and deseasonalized time series is performed by fitting autoregressive models separately to the nine divisions. The order of each autoregression is chosen with BIC and the parameters are fitted by least squares. These orders are given in Table 3.

***Table 4.*** Structure of the VAR Model Fitted.

| Lag | NE | MA | ENC | WNC | SA | ESC | WSC | M | P |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 3 | 6 | 4 | 9 | 2 | 4 | 4 | 6 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 10 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 12 | 1 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 1 |
| 13 | 1 | 2 | 0 | 2 | 1 | 1 | 1 | 1 | 1 |
| 14 | 3 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| 15 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |

*Note:* For each division the models are selected with the strategy of Section 2.2 combined with BIC.

***Table 5.*** Lag 1 Correlations of the Residuals from the Univariate Autoregressive Models as Described in Table 3.

| | NE | MA | ENC | WNC | SA | ESC | WSC | M | P |
|---|---|---|---|---|---|---|---|---|---|
| NE | −0.027 | 0.049 | 0.203 | 0.107 | 0.069 | 0.066 | 0.070 | −0.075 | 0.226 |
| MA | 0.136 | 0.016 | 0.115 | 0.038 | 0.077 | 0.108 | −0.024 | 0.022 | 0.140 |
| ENC | 0.220 | 0.071 | 0.050 | 0.073 | 0.132 | 0.167 | 0.009 | −0.002 | 0.047 |
| WNC | 0.037 | −0.055 | 0.022 | 0.030 | 0.051 | 0.112 | 0.011 | 0.136 | 0.095 |
| SA | 0.134 | 0.039 | 0.000 | −0.003 | −0.017 | 0.087 | 0.014 | 0.083 | 0.185 |
| ESC | 0.080 | 0.078 | 0.118 | 0.033 | 0.081 | 0.053 | −0.039 | −0.019 | 0.107 |
| WSC | 0.109 | 0.102 | 0.093 | −0.013 | 0.199 | 0.148 | −0.056 | 0.089 | 0.076 |
| M | 0.115 | 0.093 | 0.083 | 0.012 | 0.045 | 0.039 | 0.096 | 0.028 | 0.140 |
| P | 0.119 | 0.202 | 0.140 | 0.098 | 0.270 | 0.263 | 0.144 | 0.198 | 0.051 |

BIC is also used to specify a VAR model with spatial structure by using the model building strategy of Section 2.2. Model building is hence performed separately for each division, and parameters are estimated with least squares. The obtained models are described in Table 4. For each time lag and for each division, Table 4 reports the number of selected neighboring divisions.

We, hence, have two sets of residuals obtained from Table 3 and from Table 4 respectively. We can look at their spatio-temporal correlation structure to see whether there is some linear dependence left. Only the correlations at lag one are presented in Tables 5 and 6. Correlations at other lags are low for both models. We see that the residuals from the univariate modeling show spatio-temporal correlations at lag one, while the residuals from the VAR model have very low

***Table 6.*** Lag 1 Correlations of the Residuals from the VAR Model Described in Table 4.

| | NE | MA | ENC | WNC | SA | ESC | WSC | M | P |
|---|---|---|---|---|---|---|---|---|---|
| NE | 0.014 | 0.017 | −0.027 | 0.047 | 0.028 | 0.006 | −0.012 | −0.119 | 0.153 |
| MA | −0.032 | −0.051 | −0.073 | −0.064 | −0.052 | 0.054 | −0.060 | −0.007 | 0.133 |
| ENC | 0.055 | 0.028 | 0.047 | 0.009 | −0.067 | 0.030 | −0.085 | −0.039 | 0.034 |
| WNC | 0.036 | −0.059 | 0.026 | 0.000 | −0.054 | 0.096 | −0.047 | −0.026 | 0.044 |
| SA | −0.034 | 0.003 | −0.012 | −0.054 | −0.047 | 0.018 | −0.048 | −0.087 | 0.021 |
| ESC | 0.016 | −0.002 | 0.076 | −0.003 | −0.062 | 0.007 | −0.045 | −0.044 | 0.091 |
| WSC | −0.021 | 0.046 | 0.050 | −0.031 | 0.067 | −0.009 | −0.034 | −0.043 | 0.036 |
| M | 0.036 | 0.112 | 0.041 | −0.002 | −0.002 | −0.011 | −0.008 | 0.005 | −0.026 |
| P | 0.051 | 0.103 | −0.003 | 0.051 | 0.049 | 0.022 | 0.007 | 0.057 | 0.039 |

correlations. This indicates that the VAR model with spatial structure captures the dependence structure of the data better, as it was expected.

## 4. CONCLUSION

In this article, we have proposed a VAR model with spatial structure to analyze monthly U.S. unemployment rates observed in the nine census divisions of the United States. These spatio-temporal data are sparse in space but rich in time because we have only nine spatial regions but many observations at regular time intervals. This type of data is fairly common in econometric studies. We have shown that, unlike many other spatio-temporal modeling approaches, we do not have to assume any spatial stationarity or isotropy. We have described a model building strategy based on a spatio-temporal ordering of the nine census divisions. This ordering allows us to enter predictors sequentially in our model and identify interesting ones using partial correlation functions or model selection criteria. We have shown that our VAR model with spatial structure captures the dependence structure of the U.S. unemployment rate data better than univariate autoregressive time series models. An interesting feature of our model is its simplicity in implementation resulting from the linearity of our model and its nestedness with respect to a spatio-temporal hierarchy. We have used the software *Splus* for our computations, but any software performing linear regressions would be suitable to implement our modeling approach. The latter is obviously applicable to the prediction of other macroeconomic variables observed on different regions/countries.

Several possible generalizations of the models introduced can be identified. For instance, it is straightforward to take into account other variables measured at the same regions by including them as explanatory variables within the VAR framework. Moreover, instead of looking for a single optimal model, one may use model averaging procedures (see, e.g. Hoeting et al., 1999). Model averaging may be performed not only over different models based on a given ordering of the spatial regions, but also over models based on different such orderings. The model building strategy is also open to further generalizations, for example shrinkage techniques such as a ridge or Lasso penalty could be used to select the predictor variables.

## ACKNOWLEDGMENTS

We thank the Editors and an anonymous referee for helpful comments on this manuscript.

# REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

de Luna, X., & Genton, M. G. (2003). Predictive spatio-temporal models for spatially sparse environmental data. *Institute of Statistics Mimeo Series* #2546.

Haining, R. (1990). *Spatial data analysis in the social and environmental sciences*. Cambridge: Cambridge University Press.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial (with discussion). *Statistical Science*, *19*, 382–417.

Kyriakidis, P. C., & Journel, A. G. (1999). Geostatistical space-time models: A review. *Mathematical Geology*, *31*, 651–684.

Lucas, R. E., Jr. (1973). Some international evidence on output-inflation trade-offs. *American Economic Review*, *63*, 326–334.

Lütkepohl, H. (1991). *Introduction to multiple time series analysis*. Berlin: Springer-Verlag.

Montgomery, A. L., Zarnowitz, V., Tsay, R. S., & Tiao, G. C. (1998). Forecasting the US unemployment rate. *Journal of the American Statistical Association*, *93*, 478–493.

Pfeifer, P. E., & Deutsch, S. J. (1980). A three-stage iterative procedure for space-time modeling. *Technometrics*, *22*, 35–47.

Phelps, E.S. (1994). *Structural slumps: The modern equilibrium theory of unemployment, interest and assests*. Cambridge: Harward University Press.

Proietti, T. (2003). Forecasting the US unemployment rate. *Computational Statistics and Data Analysis*, *42*, 451–476.

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

Stoffer, D. S. (1986). Estimation and identification of space-time ARMAX models in the presence of missing data. *Journal of the American Statistical Association*, *81*, 762–772.