COLLANA DEL
DIPARTIMENTO DI ECONOMIA

SPATIO-TEMPORAL WEIGHTS WITH SIMULTANEOUS EFFECT FOR ENVIRONMENTAL DATA

Alessia Naccarato

ISSN 2279-6916 Working papers
(Dipartimento di Economia Università degli studi Roma Tre) (online)

Working Paper n° 157, 2012

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. l del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

REDAZIONE:
Dipartimento di Economia
Università degli Studi Roma Tre
Via Silvio D'Amico, 77 - 00145 Roma
Tel. 0039-06-57335655 fax 0039-06-57335771
E-mail: dip_eco@uniroma3.it

UNIVERSITÀ DEGLI STUDI ROMA TRE

DIPARTIMENTO DI ECONOMIA

SPATIO-TEMPORAL WEIGHTS WITH SIMULTANEOUS EFFECT FOR ENVIRONMENTAL DATA

Alessia Naccarato

Spatio-Temporal Weights with Simultaneous Effect for Environmental Data

Alessia Naccarato*
Department of Economics
Roma Tre University
a.naccarato@uniroma3.it

Abstract

In this work we present space-time models with simultaneous effect in which the relations between phenomenon measured in different sites are not determined on the base of an a priori structure, but are estimated in the model's estimation procedure.

This approach avoids superimposition of a priori space fixed structure so that the spatial weights are not bound at being symmetrical and have not to be constant over time. In this way, spatial weights will then take into account the intensity and the direction of influence that a single station, in a given time including the contemporaneous one, has on the whole spatial system and vice versa.

To reduce the increase in the number of parameters a procedure for constrained estimation is presented.

Jel Code: C310

Keywords: VAR Models, Space-Time Models, Simultaneous Effect, Constrained Maximum Likelihood Estimator.

1. Introduction

The interest in space-time data analysis is constantly increasing in many research fields such as environmental, economics and health sciences, physics and meteorology; examples include Gelfand et al., 1998, Kyriakidis and Journel, 1999, Cressie and Huang, 1999, Gneiting, 2002, Lee and Ghosh. 2008, Glasbey and Allcroft, 2008, Heuvelink and Griffith, 2010.

*I am grateful to Prof. L. Pieraccini and Prof. S. Terzi for valuable suggestions on this article and for their constructive comments.

For this reason recently, in statistical literature appeared many contributions about space-time modelling; in particular dynamic linear models, or state-space model, are frequently used to see a space-time series as a multidimensional time series, both in a non-Bayesian approach (Huang and Cressie, 1996; Mardia et al., 1998; Wikle and Cressie, 1999; Huang et al., 2007) and in a Bayesian one (Tonellato, 2005; Stroud et al., 2001; Banerjee et al., 2005a, 2005b). In particular spatial econometric models have recently been widely studied (Anselin, 1988, 1995, 2003).

In this work we deal with environmental studies and starting from the idea that a space-time series can be seen like a multiple time series, we face a subject not very frequently considered in statistical literature: a space-time model with simultaneous effect (Deutsch and Pfeifer, 1980, 1981; Terzi, 1995; Guagnano and Terzi, 2000, Di Giacinto, 2006).

Environmental data are generally characterized by a large number of observations over time in more than one locations and it is of great importance to build up a model able to catch the spatial effect (Haining et al., 2010).

Having for each one of S locations, T consecutive observations from a random variable Y are available the *current* observation Y in one location has to be expressed as a linear combination of both its own *lagged* values and *current* and *lagged* values in the remaining locations.

A first point that has to be stressed for the estimation of simultaneous spatial autoregressive model regards spatial weights. In a very well known paper (F. Bavaud, 1998) it is written: *"Since the pioneering works of Moran [1948], Geary [1954] Cliff and Ord [1973] [1981] and many others, the problem of properly specifying the spatial weight coefficients has largely been recognized as difficult and controversial. … this state of affairs is not surprising, for there is no such thing as "true", "universal" spatial weights: good candidates must reflect the properties of the particular phenomenon, properties which are bound to differ from field to field (Cliff and Ord 1973; Arora and Brown 1977; Cressie 1991)......"* . In 2003 Anseline wrote: "…*The standard taxonomy … is perhaps too simplicistic and leaves out other interesting possibilities for mechanism through which phenomena or actions at a given location affect actors and properties at other location.*" More recently Getis

(2009) underlines the importance to consider spatial weights matrices strongly related to some aspects of phenomena (social contacts, human energy expended, elapsed time and so on) that couldn't be represented by a contiguity structure, which implies that the spatial weights matrix is exogenous to the system.

According to this point of view we think that spatial weights have to be estimated in the model's estimation procedure because of their endogeneity. This will have three main consequences: first, there will be no superimposition of a fixed structure (such as the one in which spatial weights are considered function of the distance between locations); second, the spatial weights are not bound at being symmetrical as stated in Deng, 2008 and Cliff and Ord, 2009; third, spatial weights have not to be constant over time (in the sense that they may change according to the lag considered). Spatial weights will then take into account the intensity and the direction of influence that a single station, in a given time, has on the whole spatial system and vice versa.

The relevant feature of the proposed model is then the attempt to catch the spatial effect together with the temporal one, regardless of the hypothesis that a greater proximity between different places implies a higher spatial correlation. If distance between spatial locations has any effect on the diffusion of the phenomenon, such effect will be cached over by the model through the estimated spatial weights.

The second aspects presented here is the simultaneous dependence of Y in one location from the current values observed in the remaining ones that is made in order to catch the "purely" spatial effect which allows to take into account the values in other locations at "contemporary" time t to update the dependent variable at the location under exam.

It must be stressed that when only lagged values at time $t-k$ ($k = 1, 2, \dots$) are considered a relevant loss of information takes place. On the other hand it must be recognized that simultaneity of relations makes the estimation problem more difficult. While in a preceding paper an estimation procedure based on Two Stage Least Squares principle has been proposed (Naccarato, 2004), in this work we consider Maximum Likelihood Estimators of the whole set of parameters of the model.

The simultaneous dependence of the current value of Y in all locations can then be seen as a second relevant features of the model here presented.

The third point this work deals with is a consequence of direct estimation of spatial weights: that is the possible substantial increase in the number of parameters to be estimated.

In a very well known paper Martin and Oeppen (1975) wrote: "… *Forecasting procedures could be seriously deficient if this models were either inadequate or unnecessarily excessive in the use of parameters. … The specific aim is to derive aggregate models possessing maximum simplicity and the smallest number of parameters consonant with representational accuracy. That is the objective is to obtain adequate but parsimonious models (Tukey, 1961; Box and Jenkins, 1970). …"*.

While the two authors propose the use of autocorrelation function both in time and space, in this work we face the problem using identifying restrictions on space-time variability of the unknown parameters. In this way the structure of space-time dependence becomes a part of the estimation procedure. To this extent it must be stressed that the imposition of constraints on spatial weight – that have anyway to be estimated – has a lesser impact on the estimation of the model than a completely arbitrary choice of their values according to prior decisions.

This article is organized as follows. First we present (§ 2) the space-time model in its VAR form (including simultaneous spatial effect) in order to establish notation; in § 3 we derive MLE for space and spatio-temporal parameters in the case of unconstrained models, while in § 4 we introduce the constraining procedure of parameters estimation. In § 5 we deal with a specific case of non symmetrical spatial structure with an application to spatial diffusion of solar radiation in central Italy, to show how the proposed procedure works in determining the spatio-temporal structure of the data. Few word of conclusion (§ 6) end the work.

2. The space-time model

Let us consider a spatial system in which there are S places $I = 1, ..., S$ and K temporal lags $k = 1,2, ..., K$, i.e.

Let's indicate with the superscript the spatial reference and the subscript the time one, hence y_t^j will be the value of the dependent variable measured at time t in place j. The linear model for each one of the S station, is:

$$\begin{aligned} y_t^i = {} & \beta_1^i y_{t-1}^i + \beta_2^i y_{t-2}^i + \cdots + \beta_K^i y_{t-K}^i + \\ & + w_0^{j_1 i} y_t^{j_1} + w_0^{j_2 i} y_t^{j_2} + \ldots + w_0^{j_{S-1} i} y_t^{j_{S-1}} + \\ & + w_1^{j_1 i} y_{t-1}^{j_1} + w_1^{j_2 i} y_{t-1}^{j_2} + \ldots + w_1^{j_{S-1} i} y_{t-1}^{j_{S-1}} + \\ & + \ldots + \\ & + w_K^{j_1 i} y_{t-K}^{j_1} + w_K^{j_2 i} y_{t-K}^{j_2} + \ldots + w_K^{j_{S-1} i} y_{t-K}^{j_{S-1}} + \\ & + u_t^i \end{aligned} \tag{1}$$

In (1) for simplicity the hypothesis that the autoregressive component and the space-time one have the same maximum order of lag is done; anyway, the structure of the model would be the same even in the case the lag order of the two components is not the same.

In (1):

- the spatial component $y_t^{j_m}$ ($m = 1, 2, \ldots, S-1$; $j_m \neq i$) indicates the value of the variable Y measured at time t in one of the other sites of the system, hence its coefficient $w_0^{j_m i}$ represents the effect that the variable Y measured at station j_m at the time t has on the one measured in the station i at the same time;
- the autoregressive component y_{t-k}^i (k=1,...,K) indicates the value of the variable Y measured in the station S at time t-k; its coefficient β_{t-k}^i indicates the effect that the variable observed in k lags before in the station i has on the variable measured at time t in the same station;
- the spatio-temporal component $y_{t-k}^{j_m}$ ($m = 1, 2, \ldots, S-1$; $k = 1, 2, \ldots, K$) indicates the value of the variable Y measured in k previous instants in one of the station of the system different from the *i-th* and the coefficient $w_k^{j_m i}$ indicates its effect on the phenomenon recorded in the station i at time t;

- the error component u_t^i that represents the effect of all the events that have not been explicitly considered in the model.

With regards to the whole spatial system the model (1) can be written in the following way:

$$\begin{aligned}\underset{T\times S}{Y_{(0)}} = & \underset{T\times S}{Y_{(1)}}\underset{S\times S}{\beta_1} + \ldots + \underset{T\times S}{Y_{(K)}}\underset{S\times S}{\beta_K} + \\ & + \underset{T\times S}{Y_{(0)}}\underset{S\times S}{W_0} + \\ & + \underset{T\times S}{Y_{(1)}}\underset{S\times S}{W_1} + \ldots + \underset{T\times S}{Y_{(K)}}\underset{S\times S}{W_K} + \\ & + \underset{T\times S}{U_{(0)}}\end{aligned} \tag{2}$$

where T is the number of observations taking into account the order of the maximum lag .

In equation (2)

$$Y_{(0)} = \begin{bmatrix} y_1^1 & y_1^2 & \cdots & y_1^i & \cdots & y_1^S \\ y_2^1 & y_2^2 & \cdots & y_2^i & \cdots & y_2^S \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_t^1 & y_t^2 & \cdots & y_t^i & \cdots & y_t^S \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_T^1 & y_T^2 & \cdots & y_T^i & \cdots & y_T^S \end{bmatrix}$$

is the matrix of current values of variables Y in the space and $Y_{(1)}, Y_{(2)}, \ldots$ are similarly defined as matrices of the lagged variable of order 1, 2,… In the same way, let's define $U_{(0)}$ as the matrix of current error components:

$$U_{(0)} = \begin{bmatrix} u_1^1 & u_1^2 & \cdots & u_1^i & \cdots & u_1^S \\ u_2^1 & u_2^2 & \cdots & u_2^i & \cdots & u_2^S \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_t^1 & u_t^2 & \cdots & u_t^i & \cdots & u_t^S \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_T^1 & u_T^2 & \cdots & u_T^i & \cdots & u_T^S \end{bmatrix}.$$

Since the spatial contemporary effect is explicitly included in the model, the usual condition on the error component became:

1. $E(u_t^i) = 0 \qquad \forall\ t = 1,\ldots,T,\quad i = 1,\ldots,S$

2. $E\left(u_{t_1}^i \ u_{t_2}^i\right) = \begin{cases} \sigma^2 & \forall\ i = 1, \ldots, S, \ \ \forall\ t_1 = t_2 \\ 0 & \forall\ i = 1, \ldots, S, \ \ \forall\ t_1 \neq t_2 \end{cases}$

3. $E\left(u_t^i \ u_t^j\right) = 0 \ \forall\ i, \ j = 1, \ldots, S$

4. $E\left(u_t^i \ u_{t-k}^j\right) = 0 \qquad \forall\, i \neq j; i, j = 1, \ldots, S; \ t = 1, \ldots, T; \ k = 1, \ldots, K$

5. u_t are normally distributed;

6. the predetermined variables $Y_{(1)}, \ldots, Y_{(K)}$ are not correlated with the contemporary errors u_t, such assumption implies that $p\lim_{t \to \infty}\left(\frac{1}{T}\sum_{t=1}^{T} y_{t-k}^i \ u_t^i\right) = 0, \ \forall\ k = 1, \ \ldots, K; \ i = 1, \ \ldots, S$

7. the second order moments of the variables are finite.

The matrix of spatial coefficients $\underset{S \times S}{W_k}$, with $0 \leq k \leq K$ in (2), is defined as follows:

$$\underset{S \times S}{W_k} = \begin{bmatrix} 0 & w_k^{12} & \cdots & w_k^{1i} & \cdots & w_k^{1S} \\ w_k^{21} & 0 & \cdots & w_k^{2i} & \cdots & w_k^{2S} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_k^{i1} & w_k^{i2} & \cdots & 0 & \cdots & w_k^{iS} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_k^{S1} & w_k^{S2} & \cdots & w_k^{Si} & \cdots & 0 \end{bmatrix}$$

where the generic coefficient w_k^{ij} measures the effect that the phenomenon observed in the station i at time $t - k$ has on the station j at time t.

Finally the matrix:

$$\underset{S \times S}{\beta_k} = \begin{bmatrix} \beta_k^1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \beta_k^2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \beta_k^i & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \beta_k^S \end{bmatrix}$$

is the matrix of the coefficients of the autoregressive variables, where β_k^i, indicates the effect that the phenomenon recorded in the *i-th* station at time *t-k* has on Y observed in the same station but at time *t*.

Equation (2) in a more usual VAR form can also be written:

$$Y_{(0)}^{T} = W_0^T Y_{(0)}^T + \left(W_1^T + \beta_1^T\right) Y_{(1)}^T + \ldots + \left(W_K^T + \beta_K^T\right) Y_{(K)}^T + U_{(0)}^T \quad (3)$$

or

$$\left(I_S - W_0^T\right) Y_{(0)}^T = \left(W_1^T + \beta_1^T\right) Y_{(1)}^T + \ldots + \left(W_K^T + \beta_K^T\right) Y_{(K)}^T + U_{(0)}^T \quad (4)$$

and defining

$$\left(I_S - W_0^T\right) = A_0 \quad (5)$$

$$\left(W_k^T + \beta_k^T\right) = A_k \ , \ 1 \le k \le K$$

(4) becomes:

$$A_0 Y_{(0)}^T = A_1 Y_{(1)}^T + \ldots + A_K Y_{(K)}^T + U_{(0)}^T$$

So that it is also

$$Y_{(0)}^T = A_0^{-1} A_1 Y_{(1)}^T + \ldots + A_0^{-1} A_K Y_{(K)}^T + A_0^{-1} U_{(0)}^T \quad (6)$$

Furthermore we have to notice that the process y_t is stable (i. e. stationary) if

$$\det\left(I_S - A_0^{-1} A_1 z - A_0^{-1} A_2 z^2 - \ldots - A_0^{-1} A_K z^K\right) \neq 0 \ per \ |z| \le 1$$

z being the characteristic roots of the process. Then the inverse of the matrix A_0 must exists (Lutkepohl, 1991).

3. Maximum likelihood estimators

As it is well known when the model includes the simultaneous spatial component, the Ordinary Least Squares method lead to not consistent estimators of unknown parameters, because of correlation between contemporaneous observed errors in different locations, $E\left(u_t^i u_t^j\right) \neq 0$. Therefore, under the hypothesis that the process is normally distributed, maximum likelihood estimation can be considered to obtain consistent estimators. To simplify derivations, let us consider the model in its compact form:

$$A_0 Y = AZ + U \quad (7)$$

where we have first of all dropped out the sub script of Y and the matrices A and Z are defined as follows

$$\underset{SK\times T}{Z} = \begin{bmatrix} Y_{(1)}^T \\ Y_{(2)}^T \\ \vdots \\ Y_{(K)}^T \end{bmatrix}, \quad \underset{S\times SK}{A} = \begin{bmatrix} A_1 & \cdots & A_K \end{bmatrix}.$$

Making use of vec operator, the model becomes:

$$(I_T \otimes A_0)vecY = (Z^T \otimes I_S)vecA + vecU \tag{8}$$

where $E(vecUvecU^{'}) = (\Sigma_u \otimes I_T)$, with $\Sigma_u = \sigma^2 I_S$, so that it is

$$E(vecUvecU^{'}) = (\sigma^2 \otimes I_{S\times T})$$

The likelihood function $L(\sigma^2, A_0, A|Y)$ is then:

$$L = |A_0|^T (2\pi\sigma^2)^{-\frac{ST}{2}} \exp\left\{-(2\sigma^2)^T \left[(I_T \otimes A_0)vecY - (Z^T \otimes A_0)vecA\right]^T \left[(I_T \otimes A_0)vecY - (Z^T \otimes A_0)vecA\right]\right\}$$

and its logarithm:

$$\log L = c - \frac{ST}{2}\log\sigma^2 + T\log|A_0| - (2\sigma^2)^{-1}\left[(I_T \otimes A_0)vecY - (Z^T \otimes A_0)vecA\right]^T \left[(I_T \otimes A_0)vecY - (Z^T \otimes A_0)vecA\right] \tag{9}$$

where $c = -\frac{ST}{2}\log(2\pi)$.

Deriving the log-likelihood function with respect to A_0, A, σ^2 and equating to zero the partial derivatives, the so called normal equations are obtained. In particular we have

$$\frac{\partial \log L}{\partial vecA} = \frac{1}{\sigma^2}(Z^T \otimes I_S)vecU \quad \text{and} \quad \frac{\partial \log L}{\partial \sigma^2} = -\frac{ST}{2\sigma^2} + \frac{ST}{2(\sigma^2)^2} vecU^T vecU$$

$$\frac{\partial \log L}{\partial vecA_0} = Tvec(A_0^T)^{-1} - (\sigma^2)^{-1}(Y^T \otimes I_S)^T vecU \tag{10}$$

where the derivative with respect to A_0 gives raise to a non linear equation so that, to maximize the log-likelihood function an iterative algorithm is needed.

The final solution for the parameters estimates is then $vec\hat{A} = \left[(ZZ^T)^{-1} Z \otimes A_0\right]vecY$.

4. Constrained estimation

The procedure seen up to now is the starting point in the search of a parsimonious model which allows not only the estimation of spatial weights but also to find through constrained estimation the structure of spatial influence and the underlining effect of contiguity.

Going back to the model in compact form (7):

$$\underset{S\times S}{A_0}\underset{S\times T}{Y} = \underset{S\times(SK)}{A}\underset{(SK)\times T}{Z} + \underset{S\times T}{U}$$

and premultiplying both sides by A_0^{-1} as in (6) we have:

$$Y = A_0^{-1}AZ + A_0^{-1}U$$

and then applying the *vec* operator we obtain:

$$vecY = \left(Z^T \otimes I_S\right)vecA_0^{-1}A + \left(I_T \otimes A_0^{-1}\right)vecU \quad (11)$$

or in a more usual form

$$vecY^T = \left(I_S \otimes Z^T\right)vec\left(A_0^{-1}A\right)^T + \left(A_0^{-1} \otimes I_T\right)vecU^T \quad (12)$$

with variance and covariance matrix of the error component given by $\left(\Omega \otimes I_T\right)$ where $\underset{S\times S}{\Omega} = A_0^{-1}\Sigma_u\left(A_0^{-1}\right)^T$.

Let us now define the vector:

$$\eta = vec\left[\left(A_0^{-1}A_1, \ldots, A_0^{-1}A_K\right)^T\right] = vec\left(A_0^{-1}A\right)^T = vec\left\{\left(A_0^{-1}\left[A_1, \ldots, A_K\right]\right)^T\right\}$$

So that equation (12) becomes:

$$vecY^T = \left(I_S \otimes Z^T\right)\eta + \left(A_0^{-1} \otimes I_T\right)vecU^T \quad (13)$$

It has to be noticed that in equation (12) appears the vector of unknown parameters $\eta = vec\left(A_0^{-1}A\right)^T$, while actually we are interested in estimating separately the elements of the matrices $A_0, A_1, \ldots, A_K$ and hence also the constraints must be imposed on θ, defined as $\theta = vec\left[A_0^T, A_1^T, \ldots, A_K^T\right]$.

Therefore, it is necessary to find a transformation that would make possible to pass from the vector η to the θ one and subsequently to impose parametric restrictions of the form

$$\underset{(S^2K+S^2)\times 1}{\theta} = \underset{(S^2K+S^2)\times M}{R} \underset{M\times 1}{\Psi} + \underset{(S^2K+S^2)\times 1}{r} \tag{14}$$

Our aim is to obtain the estimation of the vector of parameters Ψ starting from the model (12). Recalling that it is

$$\frac{\partial h[g(\upsilon)]}{\partial \upsilon^T} = \frac{\partial h(\xi)}{\partial \xi^T}\frac{\partial g(\upsilon)}{\partial \upsilon^T} \tag{15}$$

where ξ and υ are two vectors respectively of dimensions $(m\times 1)$ e $(n\times 1)$; and $h(\xi)$ and $g(\upsilon)$ are two vectors of dimensions $(p\times 1)$ and $(m\times 1)$, with $\xi = g(\upsilon)$ (Lutkepohl, 1991, pag. 469), and applying (15) twice, the derivative of the likelihood function with respect to the vector of the unknown parameters Ψ^T becomes:

$$\frac{\partial \log L}{\partial \Psi^T} = \frac{\partial \log L}{\partial \eta^T}\frac{\partial \eta}{\partial \theta^T}\frac{\partial \theta}{\partial \Psi^T} \tag{16}$$

Evaluating separately the three derivatives that appear in equation (16); because of we have because of (14)

$$\frac{\partial \theta}{\partial \Psi^T} = R \tag{17}$$

and let it be

$$\frac{\partial \eta}{\partial \theta^T} = H\ , \tag{18}$$

Then H is the matrix that makes it possible to express the vector η as function of the vector θ, that is:

$$\eta = H\theta \tag{19}$$

Then, the reparametrization of the model will be:

$$vecY^T = (I_S \otimes Z^T)(HR\Psi + Hr) + (A_0^{-1} \otimes I_T)vecU^T$$

The likelihood function of the model according to the form expressed by the (12) is:

$$L = (2\pi)^{-\frac{ST}{2}} |\Omega|^{-\frac{T}{2}} \exp\left\{ -\frac{1}{2} \left[vecY^T - (I_S \otimes Z^T) vec(A_0^{-1}A)^T \right]^T (\Omega^{-1} \otimes I_T) \left[vecY^T - (I_S \otimes Z^T) vec(A_0^{-1}A)^T \right] \right\}$$

and deriving the log-likelihood function with respect to η^T we have:

$$\begin{aligned} \frac{\partial \log L}{\partial \eta^T} &= (vecY^T)^T (\Omega^{-1} \otimes I_T)(I_S \otimes Z^T) - \\ &- \left[vec(A_0^{-1}A)^T \right]^T (I_S \otimes Z)(\Omega^{-1} \otimes I_T)(I_S \otimes Z^T) \end{aligned} \tag{20}$$

Hence, we can evaluate the derivative of log-likelihood function with respect to Ψ^T which is:

$$\frac{\partial \log L}{\partial \Psi^T} = \frac{\partial \log L}{\partial \eta^T} HR \qquad \text{i.e.}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \Psi^T} &= (vecY^T)^T (\Omega^{-1} \otimes I_T)(I_S \otimes Z^T) HR - \\ &- \left[vec(A_0^{-1}A)^T \right]^T (I_S \otimes Z)(\Omega^{-1} \otimes I_T)(I_S \otimes Z^T) HR \end{aligned}$$

and

$$\begin{aligned} \hat{\Psi} &= \Psi + \left[R^T H^T (\Omega^{-1} \otimes ZZ^T) HR \right]^{-1} R^T H^T (\Omega^{-1} \otimes ZZ^T) Hr + \\ &+ \left[R^T H^T (\Omega^{-1} \otimes ZZ^T) HR \right]^{-1} R^T H^T (\Omega^{-1} \otimes I_{SK+1})(A_0^{-1} \otimes Z) vecU^T \end{aligned} \tag{21}$$

Equation (21) formalizes the constrained maximum-likelihood estimator. As for the unconstrained estimation case, to maximize the log-likelihood function an iterative algorithm is needed.

The constraints imposed on the parameters' estimates needs to be tested and the use of maximum likelihood method of estimation allows the following testing procedure.

Under the condition of stability of process y_t and taking into account equation (14), with $r(R) = M$, $\hat{\Psi}$ is consistent and

$$\sqrt{T}(\hat{\Psi} - \Psi) \xrightarrow{d} N\left(0, \left[R^T H^T \left(\Omega^{-1} \otimes \frac{ZZ^T}{T} \right) HR \right]^{-1} \right).$$

It is then possible to use the log-likelihood ratio test (LR) based on the comparison between the value of the log-likelihood function under constrained and unconstrained estimation.

Since for both constrained and unconstrained model the log-likelihood function to be maximized requires an iterative algorithm, then the test statistic will be approximated. In case large samples are available, the approximation can be neglected.

5. A model for non symmetrical spatial weights

Having information on a phenomenon's space and space-time variability, it is possible to include this knowledge in the model in form of constraints.

This procedure allows us to reduce the number of unknown parameters, without loosing the characteristic of non-symmetric spatial weights in the attempt to obtain a parsimonious model, even if frequently space-time models' applications are characterized by a very large number of observations.

In this paragraph we present a model for solar radiation as a function of sunshine duration (Angstrom's law), where with solar radiation we mean solar energy measured in a certain time instant by the instrument and with sunshine duration we mean time interval during which there is not cloudiness. The inclusion of one or more exogenous variables in the model is always possible.

There are two different kind of interest in the study of solar radiation. The first one is related to environmental problems. The level of solar radiation is strictly connected with the presence of pollutants which, beside worsening the quality of the air we breathe, cause rarefaction of the ozone reducing the ability of the stratosphere to hold the UV-C waves, which are the most dangerous for human health. Hence monitoring the level of solar radiation is important for checking a possible increase of this level in time due to many different phenomena that have to be kept under control.

The second kind of interest is from a statistical point of view. The level of solar radiation measured in a certain time on a single site, depends on the behaviour of those nearby being a function of cloud. Streams and humidity rate cause the movement of clouds and then the intensity of solar radiation measured in a given site. For this reason we need suitable models to understand and predict solar radiation which as many other phenomena has a spatial diffusion that is not dependent on the distance among sites.

Very recently a third kind of interest for this kind of models has been added which regards the evaluation of solar radiation for production of electricity through photovoltaic panels.

The proposed model explains solar radiation at time *t*, in a given site, as a linear function of past observations in the same site, of past and simultaneous observations in the remaining ones and of the simultaneous observations of the pure exogenous effect of sunshine duration.

5.1 Data description

The data used in our application are daily spatio-temporal series of solar radiation and sunshine duration concerning the fifteen Italian meteorological sites of Mediterranean Basin - Pisa San Giusto (01), Elba (02), Pianosa (03), Vigna di Valle (04), Roma Ciampino (05), Ponza (06), Napoli Capo di Chino (07), Capo Palinuro (08), Messina (09), Ustica (10), Trapani Birgi (11), Pantelleria (12), Cagliari Elmas (13), Capo Bellavista (14), Olbia Costa Smeralda (15) – during the period 1 January 1991 – 31 December 1998. Considering sunshine duration as variable the sites, whose behaviour is homogeneous with respect to this variable, have been chosen using cluster analysis and in particular the non-hierarchical algorithm of K-means fixing five groups. The method has been applied to the data collected in 1991 on 50 meteorological sites in Italy; the stability of the partitions during the years has been verified applying the algorithm to the data for every years of the series. Moreover, the fifteen sites obtained are under the influence of streams coming from the English Channel and the Gulf of Gascony which determine the main movements of the clouds – in the North – South direction – on the Tyrrhenian sea. For these reasons the fifteen sites define a climatic area.

First of all, we carried out the analysis of global and partial autocorrelations of solar radiation series in each site to underline a possible non-stationarity in mean of the phenomenon. For all the series considered, the estimate global autocorrelation function goes to zero very slowly and in a straight way; the partial autocorrelation is in practice one at lag 1 and almost zero elsewhere and the estimated inverse autocorrelation goes to $-\frac{1}{2}$ at lag 1.

In order to evaluate the importance of simultaneous effects, in our application we compare the results concerning two different model estimated: the first one (ST(K)) not includes a simultaneous effect, the second one (STSE(K)) involves it in the model equation. In the ST(K) model the non-stationarity in mean is eliminated transforming the variable using first order differentiation and the differentiated series showed an autoregressive scheme of second order (AR(2)) for all the sites. Spatio-temporal effects for lags greater than two were also taken into account but their inclusion in the models was not significant. Including in the model simultaneous spatial effect – using hence STSE(K) model – the data do not show non-stationarity in mean probably due to the fact that simultaneous spatial effect underline the long term trend effect. In this case we used the original series.

6. Results

As noted before, the main problem on the use of models proposed is the high number of parameters to be estimated.

In our application, besides the parameters relative to spatio-temporal effects, we had to consider even those of the pure exogenous component sunshine duration. To reduce the number of these parameters, we used a priori information about meteorological features of the climatic area. In particular, the fifteen sites are under continuous and predominant influence of the streams coming from the English Channel and the Gulf of Gascony, which cause the movements of clouds in the same directions for two or three days; hence we assumed that the spatio-temporal effect of the phenomenon is constant in the same period. This is equivalent to suppose that the extra-diagonal elements of the matrix A_2 are the same as those of the matrix A_1. To verify this hypothesis of constancy in time of the spatial relations we used the Likelihood Ratio Test and on the basis of the results of this test, we put all the constrains of constancy in time of the spatial relations among the sites. In Table 1a and 1b we show the estimated coefficient matrices for ST(2) wrote as

$$y_t = \alpha + C_0 x_t + \sum_{k=1}^{2} A_k y_{t-k} + u_t$$

where C_0 is the coefficient matrix of the exogenous variable.

In matrix A_2 (Tab. 1b) we reported in bold the parameters which remain constant between lag 1 and lag 2. As we can see the number of parameters to be estimated is reduced and the model fit well the data ($R^2 = 0.77$). To take into account the effect that the phenomenon measured on the whole network at time t has on a single site at the same time, we considered also the simultaneous observations of solar radiation. Hence we used the model STSE(2) in the form

$$y_t = \alpha + C_0 x_t + A_0 y_t + \sum_{k=1}^{2} A_k y_{t-k} + u_t$$

where the original series are not differentiated supposing that the simultaneous observations could underline the long term trend effect. The results of the Likelihood Ratio Test underline that, if is it possible to consider valid the hypothesis of constancy of the relations from lag 1 to lag 2, it is not convenient to constrain to this hypothesis the simultaneous coefficients. In Table 2a, 2b and 2c we show the estimated coefficient matrices obtained maximizing the log-likelihood function.

Again (Table 2c) in matrix A_2 we reported in bold numbers the parameters which remain constant between lag 1 and lag 2. Also in this case the number of parameters to be estimated is reduced and the fitting to the data is improved ($R^2 = 0.85$). It must be noted that at lag 2 there are spatial effect among farer sites that were not at lag 1 as those among sites 04-05 and sites 11-12 as shown in matrix A_2.

7. Conclusion

In this work we present a model that is very flexible and appropriate to understand the directional effects in the analysis of phenomena characterized by a non-symmetric variability both over space and time. In it we consider the presence of simultaneous effect of spatial locations on each other. The model defines a non symmetric relationship both for contemporaneous and for lagged components. As far as the presence of "pure" spatial weights is concerned it is worthwhile to stress the better performance (seen in § 5) of the model with contemporaneous non symmetric effect with respect to the one in which simultaneity it is not taken into account. The introduction of contemporary effect improves substantially the performance of the model.

To assume non-symmetry of relations between sites at every lag, however, involves a large number of parameters to be estimated even if it has to be stressed that this difficulty can be irrelevant if the number of available observations is large. But even in that case constraints on the unknown coefficients can be formulated on the basis of a priori information about the spatial and spatio-temporal diffusion of the process.

To face this problem, in our application we used the results of the cluster analysis as well as information about the movements of the clouds on the target area and their constancy over time. It is worth to note that the estimated spatio-temporal coefficients of the two models estimated are in accordance with the main movements of clouds in the Mediterranean Basin. In particular they capture the effects of the main movements of clouds in the NW-SE direction as those from Tuscany Archipelago and Sardinia to the South of Italy. This fact is evident looking at the matrices of spatio-temporal coefficients where, after ordering the meteorological sites from the North to the South, we can see that almost all the under-diagonal elements are zero. Moreover, the model proposed point out that the spatial effects may change as a function of the temporal lag considered; as noted in our application, it is not possible to hold the constancy constrain for the relations among all the sites and in particular, using simultaneous effects we can see that at lag 2 there are spatial effects among farther sites.

In our opinion, while such a procedure should requires the support of a priori information, it would always be appropriate to obtain a more parsimonious fit of the model. In this sense it is possible to think to a procedure that through testing arrives at the definition of a set of constant parameters.

A less stringent hypothesis about both autoregressive and space-time parameters could be perhaps to consider their evolution over time to be expressed in polynomial form.

Table 1a – Estimated coefficient for matrix A_1 *in ST(2) model*

A_1	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**
01	0.56	-0.14	0.39	0.74	0.71	0.29									
02	0.22	0.61		0.48	0.57	0.71	0.32	0.17							0.41
03		0.53	0.36		0.29	0.19	0.06	0.21	0.34					0.37	
04				0.49	0.69	0.27	0.74	0.69	-0.14	0.38					
05					0.73	0.64	0.81	0.58	0.33	0.47					
06						0.51	0.18	0.67	0.48	0.57	0.67				
07							0.36	0.83	0.67	0.48	0.58				
08								0.58	0.52		0.23				
09									0.63	0.31	0.28				
10										0.44	0.69	0.54			
11											0.27	0.78			
12											0.47	0.21			
13							0.12		0.25	0.70	0.75	0.69	0.39		
14						0.51		0.39	0.42	0.82	-0.29	0.48	0.19	0.41	
15				0.17	0.22	-0.33	-0.18			0.36	0.30			0.57	0.87

Table 1b – Estimated coefficient for matrix A_2 in ST(2) model

A_2	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**
01	0.03		***0.39***	***0.74***		***0.29***	0.14								
02		0.13		0.23	0.19	***0.71***	***0.32***	***0.17***							
03		***0.53***	0.07			***0.19***	***0.06***	***0.21***	0.07						
04				0.38		***0.27***	***0.74***	***0.69***	-0.32		0.01				
05					0.05	0.11	***0.81***	***0.58***	***0.33***	***0.47***					
06				0.14		0.23	0.21	***0.67***	***0.48***	0.29					
07							0.09	***0.83***	***0.67***	***0.48***	0.45	0.03			
08								0.12	***0.52***	0.10	0.23				
09									0.20	***0.31***	0.28				
10										0.28	0.19	0.15			

11										0.07	0.08	0.78			
12											0.09	0.11			
13										***0.70***	0.75	0.69	0.07		
14						0.02			***0.42***	***0.82***	0.27	0.33	***0.19***	0.21	
15						0.08	0.11			***0.36***	***0.30***			***0.57***	0.34

Table 2a – Estimated coefficient for matrix A_0 *in STSE(2) model*

A_0	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**
01	1	0.54	0.82	0.59	0.69	0.23									0.33
02	0.74	1	0.81	0.48	0.71	0.56	0.31	0.46						0.18	0.64
03	0.52	0.63	1	0.14	0.33	0.49	0.28	0.14						0.39	0.53
04	0.84	0.79	0.72	1	0.61	0.54	0.70			0.15				0.44	0.41
05	0.33	0.25	0.49	0.65	1	0.69	0.74	0.21	0.41	0.12				0.38	0.12
06				0.18	0.42	1	0.53	0.86	0.74	0.59					
07						0.89	1	0.81	0.52	0.62					
08								1	0.79						
09								0.83	1		0.73				
10										1	0.64	0.58			
11										0.75	1	0.59			
12											0.48	1			
13											0.71	0.75	1		
14							0.11	0.46	0.42	0.54	0.58	0.22		1	
15				0.25	0.36	0.78	0.65			0.69	0.52				1

Table 2b – Estimated coefficient for matrix A_1 *in STSE(2) model*

A_1	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**

01	0.28	0.12	0.31	0.57	0.67	0.11	0.25								
02	0.18	0.31		0.32	0.43	0.38	0.19	0.11							0.28
03	0.23	0.26	0.44			0.16	0.12	0.09	0.10					0.13	
04				0.13		0.23	0.61	0.52	-0.32	0.27	0.09				
05					0.81	0.29	0.66	0.56	0.29	0.31	0.11				
06						0.63	0.33	0.73	0.41	0.64	0.37				
07							0.29	0.59	0.61	0.72	0.62				
08								0.52	0.34		0.19				
09									0.65		0.23				
10										0.32	0.47	0.36			
11											0.15	0.48			
12											0.41	0.18			
13								0.13	0.11	0.55	0.68	0.64	0.27		
14						0.58		0.22	0.25	0.79	-0.27	0.37	0.29	0.23	
15				0.31	0.33	-0.24	-0.15			0.28	0.23			0.68	0.73

Table 2c – Estimated coefficient for matrix A_2 in STSE(2) model

A_2	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**
01	0.09			***0.57***	***0.67***	0.09	0.14								
02		0.14		0.39	***0.43***	***0.38***	***0.19***	***0.11***							
03	***0.23***	***0.26***	0.31				0.02	0.12	0.09	0.10					
04				0.02		***0.23***	***0.61***	***0.52***	***-0.32***	***0.27***	0.05	0.02			
05					0.12	***0.29***	***0.66***	***0.56***	***0.29***	***0.31***	0.07	0.06			
06					0.08	041	***0.33***	***0.73***	***0.41***	0.32					
07							0.07	***0.59***	***0.61***	0.51	0.58				
08								0.11	***0.34***	0.06	0.19				
09									0.35	0.05	***0.23***				
10										0.18	***0.47***	***0.36***			
11										0.01	0.06	***0.48***			

12											0.01	0.21			
13										***0.55***	***0.68***	0.64	0.14		
14						***0.58***		***0.22***	***0.25***	0.63	0.25	***0.37***	***0.29***	0.08	
15				***0.31***	***0.33***	0.09				***0.28***	***0.23***			0.47	0.52

References

Anselin, L., (1988), Spatial Econometrics: Methods and Models, Dordrecht, the Netherlands: Kluwer Academic.

Anselin, L., (1995), Local Indicators of Spatial Association – LISA, Geographical Analysis, 27, 93-115.

Anselin, L., (2003), Spatial Externalities, Spatial Multipliers, and Spatial Econometrics, International Regional Science Review, 26, 153-166.

Banerjee, S., Gelfand, A. E. and MacEachern, S., (2005a), Bayesian Nonparametric Spatial Modelling with Dirichlet Process Mixing, JASA, 100, 1021-1035.

Banerjee, S., Gelfand, A.E., and Gamerman, D., (2005b), Spatial Process Modelling for Univariate and Multivariate Dynamic Spatial data. Environmetrics , 16, 1-15.

Cliff, A. D., Ord, J. K., (2009),What Were We Thinking?, Geographical Analysis, 41, 351-363.

Cressie, N., Huang, H. C., (1999), Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions, Journal of the American Statistical Association, 94, n° 448, 1330-1340.

Deng, M., (2008), An Anisotropic Model for Spatial Processes, Geographical Analysis, 40, 26-51.

Deutsch S., J., Pfeifer, P., E., (1980), A Three-Stage Iterative Procedure for Space-Time Modeling, Technometrics, 22, n° 1,35-47.

Deutsch S., J., Pfeifer, P., E., (1981), Space-Time ARMA Modeling With Contemporaneously Correlated Innovations, Technometrics, 23, n° 4, 401-409.

Di Giacinto V., (2006), A Generalized Space-Time ARMA Model with an Application to Regional Unemployment Analysis in Italy, International Regional Science Review, 29, n° 2, 159-198.

Gelfand, A. E., Ghosh S. K., Knight J. R., Sirmans C. F., (1998), Spatio-Temporal Modeling of Residential Sales Data, Journal of Business and Economic Statistics, 16, 312-321.

Getis, A., (2009), Spatial Weights Matrices, Geographical Analysis, 41, 404-410.

Glasbey, C. A., Allcroft D. J., (2008), A Spatiotemporal Autoregressive Moving Average Model for Solar Radiation, Applied Statistics, 57, 343-355.

Gneiting, T., (2002), Nonseparable, Stationary Covariance Functions for Space-Time Data, Journal of the American Statistical Association, 97, n° 458, 590-600.

Guagnano, G., Terzi, S., (2000), The Generalized Space-Time Autoregressive Moving Average Model, Quaderni del Dipartimento di Studi Geoeconomici, Università degli Studi di Roma La Sapienza.

Kyriakidis, P.C., Journel, A.G., (1999), Geostatistical Space-Time Models, Mathematical Geology, 31, 651–684.

Haining, R. P., Kerry, R., Oliver, M. A., (2010), Geography, Spatial Data Analysis, and Geostatistics: An Overview, Geographical Analysis, 42, 7-31.

Heuvelink, G. B. M., Griffith, D. A., (2010), Space-Time Geostatistics for Geography: A Case Study of Radiation Monitoring Across Parts of Germany, Geographical Analysis, 42, 161-179.

Huang, H. C., Cressie, N., (1996), Spatio-Temporal Prediction of Snow Water Equivalent Using the Kalman Filter, Computational Statistics and Data Analysis, 22, 159-175.

Huang, H. C., Martinez, F., Mateu, J., Montes, F., (2007), Model Comparison and Selection for Stationary Space-Time Models, Computational Statistics & Data Analysis, 51, 4577-4596.

Lee H., Ghosh S. K., (2008), A Reparametrization Approach for Dynamic Space-Time Models, Institute of Statistics Mimeo Series # 2587.

Lutkepohl, H., (1991), Introduction to Multiple Time Series Analysis, Springer Verlag, Berlin.

Mardia K.V., Goodall C., Redfern E.J., and Alonso F.J., (1998), The Kriged Kalman Filter (with discussion), Test, 7, 217-252.

Martin R. L., Oeppen J. I., (1975), The identification of Regional Forecasting Models Using Space-Time Correlation Functions, Transaction of the Institute of British Geographers, 66, 95-118.

Naccarato, A., (2004), I Minimi Quadrati a Due Stadi per la Stima di Serie Spazio-temporali con Effetto Simultaneo, Quaderni di Statistica, 6, 49-69.

Naccarato, A., Lamberti, A., (2003), VAR Models for Spatio-temporal Structures: An Application to Environmental Data, New Developments in Classification and Data Analysis, Proceedings of the Meeting of the Classification and Data Analysis Group of the Italian Statistical Society, Springer, 243-250.

Stroud, J. R., Muller, P., Sansò, B., (2001), Dynamic Models for Spatiotemporal Data, J. Roy. Statist. Soc., Ser. B 63, 673-689.

Terzi, S., (1995), Maximum Likelihood Estimation of a Generalized STAR (p; l_p) Model, Journal of the Italian Statistical Society, 4, n° 3, 377-393.

Tonellato, S. F., (2005), Identifiability Conditions for Spatio-Temporal Baesyan Dynamic Linear Models, Metron, LXIII, 81-101.

Wikle, C. K. and Cressie, N., (1999), A Dimension-reduced Approach to Space-Time Kalman Filtering, Biometrika, 86, 815–829.