LÓPEZ, F.A. (*) and C. CHASCO (**)

(*) Department of Quantitative and Computing Methods. Universidad Politécnica de Cartagena. e-mail: Fernando.Lopez@upct.es
(**) Department of Applied Economics. Universidad Autónoma de Madrid. e-mail: coro.chasco@uam.es

ABSTRACT

The purpose of this article is to analyse the dynamic trend of spatial dependence. Firstly, we show some of the commonly used exploratory spatial data analysis (ESDA) techniques and we also propose other new ones, the exploratory space-time data analysis (ESTDA) that evaluates the instantaneity of spatial dependence. We also propose the space-time correlogram as an instrument for a better specification of spatial lag models. Some applications with economic data for Spanish provinces shed some light upon these issues.

Key words: Spatial dependence, spatial diffusion, ESDA, correlogram, Spanish provinces

JEL Classification: C21, C33, C51, C53, D14, O18

1. INTRODUCTION

The purpose of this article is to analyse the dynamic trend of spatial dependence. In effect, spatial dependence has been usually defined as a spatial effect, which is related to the spatial interaction existing between geographic locations and takes place in a particular moment of time. In other words, spatial dependence consists of the contemporary coincidence of value similarity with locational similarity and can be formally expressed, for a same moment of time, as a spatial autoregressive model, in which a variable $y$ is a function of its spatial lag $Wy$ (a weighted average of the value of $y$ in the neighboring locations, or spatial lag)\(^1\). But in most socio-economic phenomena this “coincidence in values-locations” (Anselin, 2001A) is usually a final effect of some cause happened in the past, which has been spread through geographic space during a certain period of time.

---

\(^1\) $W$ matrix is the so-called spatial weight matrix, which has been profusely defined in the literature (e.g., Cliff and Ord 1975, 1981; Anselin 1988).
Spatial dependence is produced by two causes: on the one hand, some kind of biasing due to the use of spatial data (e.g. the non-correspondence between the proper spatial scale of a variable and the scale at which the data is collected) and on the other hand, the existence of spatial interaction, spatial spillovers and spatial hierarchies in a phenomenon. When the former is present in the errors of a spatial regression, the spatial-error model is considered the best specification. But when spatial interaction effects are present in the endogenous variable of a spatial regression model, the spatial-lag or “simultaneous model of spatial dependence” (Anselin, 2001B) has been frequently mentioned in the literature as the solution. Nevertheless, as the spatial dependence caused by spatial interaction processes takes place not only in contemporary but also in time-wise contexts, we also propose the use of another specification, the space-time dependence regression models. This kind of models expresses better the effects due to spatial interaction as spatial diffusion phenomena, which is not only “horizontal” – simultaneous- but also time-wise.

Recently, Anselin (2001A) has shown a brief taxonomy for panel data models, with different kind of spatial dependence structure for the endogenous variable (space, time and space-time), which is called pure space-recursive, time-space recursive, time-space simultaneous and time-space dynamic models. Space-time dependence has also been specified in either some theoretical frameworks (Baltagi et al., 2003; Pace et al., 2000) or panel data applications (Case, 1991; Elhorst, 2001; Yilmaz et al., 2002; Baltagi and Li, 2003; Mobley, 2003).

In this article, we analyse the spatial dependence structure in regression models allowing for not only horizontal or vertical interactions but also space-time ones. In social contexts, it is no doubt that a shock produced in a certain location (e.g. an income growth) will be probably diffused over its neighboring locations during a certain period of time. But there are some questions that should be answered: “which proportion of this shock will be translated to the surroundings?”, “how long does it take until the diffusion process completely end up?” or “how many locations will be affected by this shock?”. We can only quantify these three items with space-time lag regression models, as it will be shown next.

The paper proceeds as follows: in the next chapter, we show some of the commonly used exploratory spatial data analysis (ESDA) techniques and we also propose other new ones, the exploratory space-time data analysis (ESTDA) that evaluate the instantaneity of spatial dependence. In section 3, the space-time correlogram is proposed as an instrument for the identification of space-time dependence models. It is
illustrated with some examples for economic series of Spanish provinces is the period 1986-2002. Some summary conclusions and the references close this reflection.

2. EXPLORATORY SPACE-TIME DATA ANALYSIS (ESTDA)

Before analysing the space-time confirmatory process, some ESDA tools are shown to illustrate space-time processes. Multivariate ESDA is still in an initial stage and the contributions in this area are scarce. Therefore, it is our aim to do a brief description of the available multivariate ESDA tools that can be useful for space-time statistical analysis. First, we briefly present (without going into further details) some bivariate ESDA tools to specifically apply them to analyse a same variable in two different moments of time. We finish this section proposing new indicators and exploratory tools for the analysis of space-time processes, which have been defined as “exploratory space-time data analysis” or ESTDA.

Our goal is twofold: firstly, we contribute towards obtaining appropriate indicators to evaluate the dynamic diffusion of spatial dependence and secondly, we develop some statistics for the dissociation of contemporary –simultaneous- and non-contemporary – lagged- spatial dependence present in a wide-range of socio-economic phenomena.

2.1. Multivariate spatial correlation

The multivariate spatial correlation coefficient quantifies the spatial dependence degree existent between two variables \( Y_k, Y_l \) in a same location \( i \) (Wartenberg, 1985). It is defined as follows:

\[
m_{ki} = z'_{ki} W^s z_i
\]

where: \( z_k = \frac{Y_k - \mu_k}{\sigma_k} \) and \( z_i = \frac{Y_i - \mu_i}{\sigma_i} \) have been standardized such that the mean is zero and standard deviation equals one. \( W^s \) is a doubly standardized (or, stochastic) spatial weights matrix, being \( W \) the familiar spatial weight matrix that defines the neighborhood interactions existent in a spatial sample (Anselin, 1988).

This concept of multivariate spatial correlation thus centers on the extent to which values for one variable \( z_k \) observed at a given location show a systematic (more than likely under spatial randomness) association with another variable \( z_l \) observed at the neighboring locations.

Note that this multivariate spatial correlation can be considered in addition to or instead of the usual (non-spatial) correlation between the two variables at the same location, provided that each location is associated to its neighboring observations.
2.1.1. Bivariate Moran spatial autocorrelation statistic

Our focus is on the linear association between a variable \( z_k \) at a location \( i \) \((z_{ik})\), and the corresponding “spatial lag” for the other variable, \([Wz_i]\). In this context, the usual singly-standardized (row-standardized) form of the spatial weights matrix can be used, which yields an interpretation of the spatial lag as an “average”\(^2\) of neighboring values. Also, the cross-product statistic (1) can be re-scaled by dividing by the sum of squares for the first variable. This yields a multivariate counterpart of a Moran-like spatial autocorrelation statistic as (Anselin et al. 2002):

\[
I_{ij} = \frac{z_i'Wz_j}{z_i'^2 \sum_k z_k^2} \tag{2}
\]

or:

\[
I_{ij} = \frac{z_i'Wz_j}{n} \tag{3}
\]

with \( n \) as the number of observations. Since the \( z \) variables are standardized, the sum of squares used in the denominator of (2) is constant and equal to \( n \), irrespective of whether \( z_k \) or \( z_{ik} \) are used.

The significance of this multivariate spatial correlation can be assessed in the usual fashion by means of a randomisation (or permutation) approach. In this, the observed values for one of the variables are randomly reallocated to locations and the statistic is recomputed for each such random pattern.

2.1.2. Bivariate Moran scatterplot

As it is well known, the univariate Moran’s I statistic can be visualized in the so-called Moran scatterplot. If the vertical axis represent the spatial lag, \([Wz_i]\), and the horizontal axis represents the original variable, \(z_{ik}\) (using the variables in standardized form), Moran’s I value coincides with the slope of the regression line of \( Wz_i \) on \( z_k \).

A bivariate generalization of Moran scatterplot corresponds with a scatterplot with the \( z_i \) spatial lag, \( Wz_i \), on the vertical axis and the variable \( z_k \) on the horizontal axis (using the variables in standardized form). The slope of the regression line in this scatterplot is equal to the value of the expression (2). In addition, it is also possible to examine each individual location as associated with the four quadrants of the scatterplot, which are the four types of multivariate spatial association.

\(^2\) It corresponds with an average but it is not a mean in a strict sense.
2.2. Space-time autocorrelation

2.2.1. Moran space-time autocorrelation statistic

Instead of being completely different, variables \( z_l \) and \( z_k \) could be the same variable observed in two instants of time, \( t \) and \( t-k \) – with the only limitation of avoiding future values explaining past ones. In this case, the bivariate Moran’s I computes the relationship between the spatial lag, \( Wz_t \), on instant \( t \) and the original variable, \( z \), on instant \( t-k \) (\( k \) order time lag). Therefore, this statistic quantifies the influence that a change in a spatial variable \( z \), operated in the past (instant \( t-k \)) in an individual location \( i \) (\( z_{t-k} \)) exert over its neighborhood in the present time (\( Wz_t \)). So, it is possible to define the Moran space-time autocorrelation statistic as follows:

\[
I_{t-k} = \frac{z'_{t-k}Wz_t}{z_{t-k}z_{t-k}}
\]  

(4)

where, as in the last case, the denominator can be substituted by \( n \) as this variable \( z \) is also standardized. The value adopted by this index, as in (2), corresponds with the slope in the regression line of \( Wz_t \) on \( z_{t-k} \). Note that for \( k=0 \), this statistic (4) coincides with the familiar univariate Moran’s I that from now on, we denote as \( I_t \).

Since the Moran’s space-time autocorrelation coefficient equals to the slope of the regression of \( Wz_{t-k} \) on \( z_t \), it is possible to connect this statistic with the standard Pearson correlation coefficient between these two variables:

\[
Corr(z_{t-k},Wz_t) = r_{z_{t-k},Wz_t} = \frac{Cov(z_{t-k},Wz_t)}{\sqrt{Var(z_{t-k})\sqrt{Var(Wz_t)}}} = \frac{1}{n} \frac{z'_{t-k}Wz_t}{\sqrt{Var(Wz_t)}}
\]  

(5)

or:

\[
r_{z_{t-k},Wz_t} = \frac{I_{t-k}}{\sqrt{Var(Wz_t)}}
\]  

(6)

So the Moran space-time autocorrelation statistic can also be expressed as:

\[
I_{t-k} = r_{z_{t-k},Wz_t} \sqrt{Var(Wz_t)}
\]  

(7)

2.2.2. Graphic tools

When considering both dimensions, some ESTDA tools can be defined to visualize and analyse the space-time structure of a distribution. That is the case of the space-time
Moran scatterplot, Moran space-time autocorrelation function (MSTAF), space-time Moran scatterplot matrix, space-time Moran surface plot, Moran’s I line graph and spatial dependence diffusion graph.

- **Space-time Moran scatterplot**

  In parallel with the bivariate Moran scatterplot, the space-time Moran scatterplot corresponds with a scatterplot with the $z_t$ spatial lag, $Wz_t$, on the vertical axis and the variable $z_{t-k}$ on the horizontal axis (using the variables in standardized form). The slope of the regression is equal to the value of the expression (4). In this case, it is also possible to examine each individual location as associated with the four quadrants of the scatterplot, which are the four types of space-time spatial association.

  ![Figure 1: Space-time Moran scatterplot](source)

  In Figure 1, two different cases are shown. In the left hand side, the space-time Moran scatterplot shows on the horizontal axis the employment rate of the 50 Spanish provinces in 1998 ($E_{98}$) and on the vertical axis, the correspondent spatial lag in 2002 ($W_{E02}$), considering $W$ as a row-standardized contiguity matrix (2 provinces are neighbors if they share a common border). As it can be seen, there is a high connection between employment rates variable in 1998 and its spatial lag 4 years later, as it is shown by the Moran space-time autocorrelation statistic ($I_{98,02}=0.6969; p-value=0.001$). But in the right graph we have represented a different situation of non-space-time autocorrelation between population for Spanish provinces in 1986 and the corresponding spatial lag in 2002 ($I_{86,02}=-0.071; p-value=0.427$).

- **Moran space-time autocorrelation function (MSTAF)**
The MSTAF is the result of plotting the values of the Moran space-time autocorrelation statistic (4), adopted by a variable in a certain period of time. It is a coordinate graph in which the Moran coefficient values are plotted on the vertical axis and the time lags on the horizontal one. The first value corresponds to the contemporaneous case, $k=0$, which is the univariate Moran’s I ($I_t$), whereas the other ones are proper Moran space-time autocorrelation coefficients ($I_{t-k,t}$). This graph visualizes the influence that a change in a spatial variable $z$, operated in the past (period from $t$ to $t-k$) in an individual location $i$ ($z_{t-k}$) exert over its neighborhood in the present time ($Wz_t$). Inference is necessary to evaluate the significance of $I_{t-k,t}$ values and, as a result, the existence or absence of spatial autocorrelation, either a contemporaneous or non-contemporaneous one.

**Figure 2: Moran space-time autocorrelation function (MSTAF)**

**Employment rate, 2002**

**Population, 2002**

![Graph showing MSTAF for employment rate and population](image)

Note: ? is 5% significance level, permutation approach (999 permutations). Source: Self-elaboration.

In Figure 2, we have represented this MSTAF function for two variables: employment rate and population of the 50 Spanish provinces in year 2002. In the horizontal axis, we have represented 16 time lags (period 1986-2002), for both series and the initial moment, lag 0 (year 2002). As it can be appreciated, there is a clear evidence of non-contemporaneous spatial dependence in the first function, as all the values exhibit significant high time-lagged values\(^3\), though it shows a decreasing trend from the 4\(^{th}\) lag to the end of the period.

Regarding to the population MSTAF function, it shows very low –almost constant-negative values, which are not significant throughout all the period. So we can conclude that there is no spatial autocorrelation –either contemporaneous or lagged ones- in the population function of Spanish provinces during 1986-2002.

\(^3\) All inference computations was done with the permutation approach and 999 permutations.
• **Space-time Moran scatterplot matrix**

It is a graph matrix in which each cell represents a Moran scatterplot: the diagonal cells include the univariate Moran scatterplot for different time lags and the non-diagonal cells contain space-time Moran scatterplots. Note that it is an upper-triangular matrix because of the limitation of avoiding future values explaining past ones.

In Figure 3, we have represented the space-time Moran scatterplot matrix for employment rate of the Spanish provinces in 1986, 1998 and 2002. The univariate Moran scatterplot for each year is on the second-diagonal cells, so it is possible to analyse the evolution of spatial dependence in this period of time. The corresponding space-time Moran scatterplots are shown in the lower-diagonal cells, so it is also possible to know how past values of per capita income in a location have affected to present values of its spatial lag. Note, for instance, that the spatial dependence degree between employment rate in 1998 (\(E_{98}\)) and its spatial lags in 2002 (\(W_{E02}\)) –represented in the first row of the matrix- is a bit higher in 1998 (\(I_{86,98}=0.6969\)) than in its own period of time (\(I_{02}=0.6554\)), suggesting some kind of diffusion process from each location \(i\) towards its neighbors in this period of time.

This scatterplot matrix also allows evaluating the evolution of each spatial unit in time detecting the evolution of spatial dependence in time (second-diagonal graphs) or changes of position in the quadrants in some provinces. For example in a further analysis, the province of Guipúzcoa (highlighted in Figure 3) was a second quadrant province in the scatterplot of employment lag variable in 1998 on employment in 1986 (lower employment in 1986, higher employment neighbors in 1998), but it has changed its position to the first quadrant in the scatterplot of employment lag variable in 2002 on employment in 1998 (higher employment in 1998, higher employment neighbors in 2002). This could be a consequence of a virtuous spatial diffusion in employment rates from Guipúzcoa neighbors from 1986 to 2002, which has translated this province from the second to the first quadrant in the space-time Moran scatterplot matrix.
Space-time Moran surface plot

It is a continuous representation of the Moran space-time autocorrelation statistics for the whole period of time. Therefore, a $k+1$ MSTAF are represented in the same graph. As in the space-time Moran scatterplot matrix, the diagonal values include the univariate Moran scatterplot for different time lags and only the upper-triangular values are represented to avoid future values explaining past ones.

In Figure 4, we have constructed the space-time Moran surface plots for the 17 employment rate series of the Spanish provinces during the period 1986-2002. As it can been seen, the higher values ($l_{t-k,t} \geq 0.48$) are more or less concentrated in –or nearby- the diagonal line, pointing out the existence of a strong contemporaneous -or quasi-contemporaneous- spatial dependence effect in all the series. But it is also seen that...
practically all the variables also contains some other higher “peaks” in the middle/end of their corresponding distributions, what should be interpreted as a non-contemporary spatial dependence effect.

**Figure 4: Space-time Moran surface plot of employment of Spanish provinces, 1986-2002**

A further analysis of the diagonal graphs in the space-time Moran scatterplot matrix leads to another useful representation: the Moran’s I line graph that allows showing the evolution of spatial dependence in a period of time. Other authors have already used this plot to explore the dynamics of contemporaneous spatial dependence in a period of time (Rey and Montouri, 1999).

In Figure 5, we have represented this coefficient trend for two provincial variables from 1986 to 2002: employment rate and population. As it can be seen, contemporaneous spatial dependence is very significant and constantly high in all the employment rate distributions (especially in 1999 and 2000). Nevertheless, all the population variables have non-significant low Moran’s I values during the whole period. The economic interpretation of these results points out a different behavior of both distributions throughout the Spanish provinces in 1986-2002: while provinces with relatively high (low) employment rate tend to be located nearby other provinces with high (low) employment more often than would be expected due to random chance, population distributions are not clustered at all.
Figure 5: Moran’s I line graph

Employment rate | Population

Note: ? is 1% significance level, permutation approach (999 permutations). Source: Self-elaboration.

- Spatial dependence diffusion graph

It represents in a same graph different scatterplots, corresponding to each spatial location \( i \), with \( Wz_i \), on the vertical axis and the variable \( z_{i,k} \) on the horizontal axis (using the variables in standardized form). It constitutes another useful view for analysing time evolution of spatial dependence.

In Figure 6, we have presented the evolution of spatial dependence in 4 Spanish provinces during 1986-2002. In the case of Madrid, it starts in the second quadrant (lower employment vs neighbors with higher rates) though stay the rest of the period in the first one (high-high employment). Note that Madrid also improves in terms of employment rate more than its neighbours in the end of the period, as detected by the bisectrix plotting. In the case of Girona, though it has clearly improved over the rest of Spanish provinces, its surroundings remains without changes. Murcia has also progressed and also its neighbors. Finally, Granada is a lagged province as it remains in the third quadrant and also its surroundings.

Figure 6: Spatial dependence diffusion graph, employment rate period 1986-2002

Source: Self-elaboration
2.3. Moran space-time partial autocorrelation coefficients

It is no doubt that the already shown spatial dependence measures include different sources of dependence that are difficult to separate. Formally:

\[ Cov(z_i, z_j) \neq 0 \]  \hspace{1cm} (8)

where sub-indexes \( i, j \) are different spatial locations and \( t, s \) are different instants of time. Therefore, we consider the following types of dependencies:

a) On the one hand, there is a dependence in expression (8) that is the result of time evolution:

\[ Cov(z_i, z_{i+k}) \neq 0 \quad ; \quad \forall i = j \]  \hspace{1cm} (9)

This expression affirms that (for \( s=t-k \)) the value of \( z \) variable in period \( t \) is more or less related to \( t-k \). This assertion is more correct for lower values of \( k \).

b) On the other hand, there is a dependence in expression (8) that is the result of spatial interactions:

\[ Cov(z_i, z_j) \neq 0 \quad ; \quad \forall t = s \]  \hspace{1cm} (10)

This second type of dependence –spatial dependence- can be produced by two sources:

b.1) **Simultaneous or contemporaneous dependence** constitutes the usual definition of spatial dependence in the literature and it is the consequence of an instant –very quick- spatial diffusion of a phenomenon in geographic space. It can be connected or be the consequence of a lack of concordance between a spatial observation and the region in which the phenomenon is analysed.

b.2) **Lagged or non-contemporaneous dependence** is the result of a slower diffusion of a phenomenon towards the surrounding space. This kind of dependence is due to the usual interchange flows existing between neighbor areas, which requires of a certain period of time to be tested.

Although it is very difficult to divide spatial dependence into its two dimensions, it is worth trying to compute them separately in order to correctly specify a spatial process that exhibits spatial dependence. That is why one of the aims of this article is to show a new
range of ESTDA tools that allow justifying the inclusion of both kind of spatial lags, contemporaneous \((Wy_t)\) and time-lagged \((Wy_{t-k})\) ones, to explain \(y_i\) in a spatial regression.

Some coefficients can be defined to evaluate the inclusion of a space-time lag term in a spatial regression. The basic underlying idea consists of eliminating the influence of one of the dimensions in order to compute separately contemporaneous and non-contemporaneous dependence. For that purpose, we substitute in (7) the space-time correlation coefficient by a partial correlation one.

A first expression computes the correlation between variable \(z\) in period \(t-k\) and its spatial lag \(Wz\) in period \(t\) removing the influence of \(z\) in in period \(t\). It can be defined as Moran space-time partial autocorrelation statistic:

\[
I_{t-k,t}^P = \text{Corr}(z_{t-k}, Wz_t | z_t) \sqrt{\text{Var}(Wz_t)} \quad ; \quad k = 1, 2, \ldots
\]

(11)

where \(\text{Corr}(z_{t-k}, Wz_t | z_t)\) is the partial correlation coefficient of variables \(z_{t-k}\) and \(Wz\) after eliminating the correlation from \(z_t\):

\[
\text{Corr}(z_{t-k}, Wz_t | z_t) = \frac{r_{z_{t-k}, Wz_t} - r_{z_{t-k}, z_t} \cdot r_{Wz_t, z_t}}{\sqrt{1 - r_{z_{t-k}, z_t}^2} \cdot \sqrt{1 - r_{Wz_t, z_t}^2}},
\]

being \(r\) the standard Pearson correlation coefficient.

This indicator removes contemporaneous spatial dependence from the relationship between variables \(z_{t-k}\) and \(Wz_t\). If the pattern of spatial dependence is one that can be captured by a contemporaneous spatial autoregression, then the partial autocorrelation will be close to zero. On the contrary, if the process is one that can be captured by a non-contemporaneous spatial dependence, then \(I_{t-k,t}^P\) will be significantly different from zero.

The complementary expression consists of computing contemporaneous—or instant- spatial dependence after removing lagged spatial dependence by the means of an index that can be defined as partial Moran’s I:

\[
I_{t-t}^P = \text{Corr}(z_t, Wz_t | z_{t-k}) \sqrt{\text{Var}(Wz_t)} \quad ; \quad k = 1, 2, \ldots
\]

(12)

where \(\text{Corr}(z_t, Wz_t | Wz_{t-k})\) is the partial correlation coefficient of variables \(z_t\) and \(Wz\) after eliminating the correlation from \(Wz_{t-k}\):

\[
\text{Corr}(z_t, Wz_t | Wz_{t-k}) = \frac{r_{z_t, Wz_t} - r_{z_t, Wz_{t-k}} \cdot r_{Wz_t, Wz_{t-k}}}{\sqrt{1 - r_{z_t, Wz_{t-k}}^2} \cdot \sqrt{1 - r_{Wz_t, Wz_{t-k}}^2}}.
\]
This indicator removes lagged spatial dependence from the contemporaneous relationship between variables $z_t$ and $Wz_t$. If the pattern of spatial dependence is one that can be captured by a lagged spatial autoregression, then the partial Moran’s I will be close to zero. On the contrary, if the process is one that can be captured by a contemporaneous spatial dependence, then $I_t^R$ will be significantly different from zero.

**Figure 7: Moran space-time partial autocorrelation functions**

![Image of Moran space-time partial autocorrelation functions]

Source: Self-elaboration.

The results obtained for these indexes for provincial employment rate and population distributions in period 2002 are in Figure 7. The bold line corresponds to $I_{t+k}$ (Moran’s space-time partial correlation coefficient) whereas the thin line is the one of $I_t^R$ (partial Moran’s I). The interpretation of the results is as follows. In the case of employment rate, the $I_{t+k}$ function has higher values than the $I_t^R$ one from lags 2 to 5. And there is an inflexion in lag 6 (1996), from which the partial Moran’s I are higher. So we can conclude that there is a non-contemporaneous spatial dependence compound in the explanation of employment rate of Spanish provinces in the years 1997-2000 that gradually declines to almost disappear in the end of the period. Especially relevant is the diffusion of provincial employment rates in 2000 and 1998 on their surroundings in 2002, whereas the impact of employment in 1986 is almost zero.

In the case of population, there is no evidence of neither contemporaneous nor non-contemporaneous spatial effect, as both $I_{t+k}$ and $I_t^R$ values are constantly very close to zero. And if any non-significant spatial autocorrelation could exist, this would be instant as $I_t^R$ values are always above than $I_{t+k}$ ones.

**3. IDENTIFICATION OF SPACE-TIME REGRESSION MODELS**
The joint representation of the Moran’s space-time autocorrelation function (MSTAF) in combination with the Moran space-time partial autocorrelation functions (MSTPF) leads to the **space-time correlogram**, which is useful to identify space-time autocorrelation processes. Therefore, this correlogram is a two-graph representation of three functions: a total autocorrelation function \( I_{t-k, t} \) and two partial ones \( I_{t-k, t}^p \) and \( I_{t-k, t}^n \).

The identification process should be tackled as follows:

- Firstly, the Moran space-time autocorrelation function (MSTAF) indicates the existence (or non-existence) of dynamic spatial dependence. If the MSTAF values are significant (using the regular inference process) we can conclude that there is spatial and temporal dependence in the corresponding distribution, and viceversa. The MSTAF trend is an indicator of the diffusion speed of the spatial distribution:
  - A decreasing trend in this function points out a **quicker space-time diffusion** of the spatial distribution, as correlations between \( z_{t-k} \) and \( Wz_t \) are lower for higher values of \( k \).
  - An increasing trend in MSTAF is a symptom of a **lower space-time diffusion** of the spatial distribution, as past values \( z_{t-k} \) will have more influence on the present spatial lagged ones \( (Wz_t) \).
  - A lack of trend –with low values- in MSTAF indicates that there is no **space-time diffusion** in the spatial distribution.

- Secondly, the Moran space-time partial autocorrelation functions (MSTPF) are the instrument to determine whether the existent spatial dependence is only instant or lagged or both at the same time.
  - **Contemporaneous or instant spatial dependence** is present in a variable if only the partial Moran’s I has significant values. In this case, only the present values of variable \( z \) (\( z_t \)) can explain its present spatial lag (\( Wz_t \)). In a spatial regression, if an endogenous variable \( y_t \) exhibits significant MSTAF and partial Moran’s I values, the spatial autocorrelation detected by this correlogram can be completely captured by a contemporaneous spatial lag of \( y_t \) (\( Wy_t \)) as an explicative variable in the model.

\[
\begin{align*}
y_t = \alpha + \rho Wy_t + \epsilon_t
\end{align*}
\]

\( \rho \) is the spatial parameter to estimate and \( \epsilon \) the error term.
- **Non-contemporaneous or lagged spatial dependence** is present in a variable if only the Moran space-time partial autocorrelation function has significant values. In effect, past values of variable \( z_t (z_{t-k}) \) completely explains its present spatial lag \( (Wz_t) \), so \( z_t \) and \( Wz_t \) are not correlated. In this case, the existence of spatial dependence in an endogenous variable \( y_t \), detected by the correlogram, can be completely captured by a space-time lag of \( y_t (Wy_{t-k}) \) as an explicative–exogenous–variable in the model.

\[
y_t = \alpha + \rho Wy_{t-k} + \epsilon_t, \tag{14}
\]

The most explicative time lagged variable \( (y_{t-k}) \) should be the one with highest value in the MSTAF. In practice, the most explicative time lagged variable \( (y_{t-k}) \) is not so straightforward to find. Several models should be estimated to determine the best space-time lag from the group of most relevant values in MSTAF.

If these significant values of the Moran space-time autocorrelation function are closer to the present moment (lag 0), it can be said that the correspondent variable presents a quick spatial diffusion process. And on the contrary, if the significant values are concentrated farther than the first time lags, the variable will have a slower spatial diffusion process.

- **Both contemporaneous and non-contemporaneous spatial dependence** are present in a variable if both partial functions have high and significant values for the same periods of time. In this case, not only present but past values of variable \( z_t \) \( (z_{t-k}) \) can completely explain its present spatial lag \( (Wz_t) \). Therefore, the existence of spatial dependence in an endogenous variable \( y_t \), detected by the correlogram, can be completely captured by both a spatial lag and a space-time lag of \( y_t (Wy_{t}, Wy_{t-k}) \) as explicative–exogenous–variables in the model.

\[
y_t = \alpha + \rho_1 Wy_t + \rho_2 Wy_{t-k} + \epsilon_t, \tag{15}
\]

\( \rho_1, \rho_2 \) are spatial parameters to estimate. In this case, the most explicative time lagged variable \( (y_{t-k}) \) should also be determined from the group of most significant values in both partial functions.

In Figure 8, the MSTAF of employment rate in 2002 has highly significant and also decreasing values throughout the period 1986-2002, so space-time dependence with a
quicker diffusion is expected in this variable. On the contrary, the MSTAF of the second variable—population in 2002—has constant (no trend) low and non-significant values. That is why a complete absence of spatial dependence—either instant or lagged—is expected here.

Regarding the MSTPF plots, provincial employment rate in 2002 exhibits both contemporaneous and non-contemporaneous spatial dependence in different sub-periods of time. In effect, on the one hand the Moran space-time partial autocorrelation function is higher than the partial Moran’s I from lags 2 to 5 (1997-2000), pointing out the domain of non-contemporary spatial dependence over instant spatial dependence in this period. But the opposite effect takes place from lags 6-16 (1996-1986) and also in lag 1 (2001), where instant spatial dependence is stronger than lagged one. In this second sub-period, lags 6-10 (1996-1993) have with higher values in both MSTPF (above 0.2), which could be indicative of instant and non-instant spatial dependence in this period of time.

Therefore, the existence of spatial dependence in employment rate in 2002 could be captured by on the one hand a space-time lag variable (E00, E99, E98 or E97) or on the other hand, both an instant spatial lag (E02) and one of the space-time lags with jointly higher values (E96, E95, E94, E93 or E92). The first option is a non-contemporary spatial dependence situation and the second one is both contemporary and non-contemporary spatial dependence. The non-contemporary (lagged) spatial dependence model (14) can be estimated by OLS, as the spatial-lag (Wy_{t-k}) is not correlated with the error term (e_t). But the space-time dependence model (15) is a spatial lag model, in which the main variable (y_t) is explained by its own instant spatial lag (Wy_t) and a exogenous space-time lag (Wy_{t-k}). As instant spatial lag is already correlated with the error term (e_t), OLS are no longer valid and a ML method must be used.
The election of the best specification must be determined after the corresponding estimation of these models (for instance, the model with best measure of fit – e.g. the less sum of squared residuals). As it is shown in Table 1, in the case of employment rate the best specification corresponds with model (14), a non-contemporary spatial model in which the only exogenous variable of employment rate in 2002 is its corresponding space-time lag in 1998.

\[ \hat{E}_{02} = 0.97 \cdot WE_{98} \]  

(16)

Therefore, it can also be concluded that the spatial diffusion process of this variable in 2002, at the provincial level in Spain, is strong (with a high coefficient) and quite quick, as it lasts 4 years.

The MSTAF values of the second variable, population, have no significant values. Therefore, no MSTPF analysis should be carried out, as no spatial autocorrelation is present at all. In any case, all the partial Moran’s I values are higher than the Moran’s I significance level.
space-time partial autocorrelation coefficients indicating a predominance of contemporary spatial dependence over non-contemporary one.

Some other variables have also been analysed and their correspondent space-time correlograms are shown in the Figure in the Annex. We have also estimated all the possible space-time models in each case to proof our conclusions: on the one hand, the non-contemporaneous space-time model (14) for each of the 16 the time lags considered in this exercise and on the other hand, the mixed contemporaneous and non-contemporaneous space-time model (15). The results are also in the Annex. As it can be seen, there is one variable with a very quick spatial diffusion effect (a quasi-contemporaneous space-time dependence): price index (2 years). But there are also two other variables with a slower spatial diffusion effect: per capita bank deposits and per capita registered cars (12 years). A special case is the one of per capita telephone lines, in which a mixed contemporaneous and non-contemporaneous spatial dependence is the best option, leading to a model (15) specification:

$$\hat{T}02 = 0.38 \cdot W02 + 0.69 \cdot W93$$  (17)

In this last case, we see that provincial per capita telephone lines variable in 2002 should be explained by both instant and non-instant spatial dependence, exhibiting a diffusion of 9 years. Non-instant or lagged spatial dependence variable \(W93\) exerts a stronger effect –almost the double- on this variable than instant spatial variable \(W02\), but both are necessary to explain this distribution.

6. CONCLUSIONS

The main aim of this paper is analysing the dynamics of spatial dependence. The inclusion of time dimension in spatial econometric models should make us think about the causes of spatial autocorrelation. If spatial dependence is produced by spatial interaction processes and information interchange flows between neighbor locations, it seems logical that these interactions need of some time to take place. Our proposal is making a differentiation between two types of spatial dependence: instant or contemporaneous and lagged or non-contemporaneous. The first one is the consequence of a very quick diffusion of the process over the neighboring locations. But the second one implies that a shock in a certain location needs of several periods of time to take place and be tested over its neighborhood. It is not easy to separate both types of spatial dependence but they must be both present very frequently when specifying a spatial dependence model.

In spatial econometrics it is usual to find some kind of spatial dependence, which is
always expressed as a contemporaneous one, when the correct specification should be only lagged spatial dependence of both instant and lagged one. It is especially true in the case of economic variables, like income or employment, which require several periods of time to spread over the geographic space. In this cases, the specification of spatial dependence as instant or contemporaneous could lead us to wrong conclusions. That is why we propose to evaluate both instant and lagged spatial dependence in the spatial lag model.

In the first parte of this paper, we have shown some tools to explore space-time dependence in a panel data. These instruments are still in an initial stage, though some of them have already been shown in previous papers (bivariate spatial autocorrelation statistic, space-time Moran scatterplot, Moran’s I line graph). We offer some other graph tools that complement these ones, which are the case of Moran space-time autocorrelation function, space-time surface plot, and spatial dependence diffusion graph. And we also include other numeric tools that allow to differentiate between both types of spatial dependence, which are based on the partial correlation concept: Moran space-time partial autocorrelation and Partial Moran I. All these indicators are useful to evaluate the dynamics of spatial dependence in a distribution.

A deepful insight of the ESTDA tools lead us to present the space-time correlogram, which can be a useful instrument to identify the existence of the different types of space-time dependence. In the second part of this paper, we illustrate the process for the identification of different types of spatial dependence in some variables, with the help of the space-time correlogram, as a way of testing our conclusions. Therefore, as expected, there are some variables –as the index price- with a quick speed of spatial diffusion, whereas there are other ones with a lower speed, as per capita bank deposits and per capita registered cars. There is also a variable –per capita telephone lines- with both instant and lagged spatial dependence.

6. REFERENCES


Annex: Space-time correlograms

<table>
<thead>
<tr>
<th>Variable</th>
<th>MSTAF</th>
<th>MSTPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered cars</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td>Price index</td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
<tr>
<td>Telephone lines</td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
<tr>
<td>Bank deposits</td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

Note: ? is 5% Moran’s I (MSTAF) significance level. Source: Self-elaboration.
<table>
<thead>
<tr>
<th>Time lag</th>
<th>Spatial lag</th>
<th>Employment rate</th>
<th>pc Registered cars</th>
<th>Price index</th>
<th>pc Telephone lines</th>
<th>pc Bank deposits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coef. SSR</td>
<td>Coef. SSR</td>
<td>Coef. SSR</td>
<td>Coef. SSR</td>
<td>Coef. SSR</td>
</tr>
<tr>
<td>0 2002</td>
<td></td>
<td>0.77</td>
<td>0.55</td>
<td>0.34</td>
<td>0.85</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15.99</td>
<td>35.95</td>
<td>42.82</td>
<td>13.590</td>
<td>28.783</td>
</tr>
<tr>
<td>1 2001</td>
<td>2002</td>
<td>0.94</td>
<td>0.71</td>
<td>0.93</td>
<td>0.91</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.077</td>
<td>34.53</td>
<td>43.422</td>
<td>12.425</td>
<td>12.625</td>
</tr>
<tr>
<td>2 2000</td>
<td>2002</td>
<td>0.96</td>
<td>0.90</td>
<td>0.76</td>
<td>0.61</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16.279</td>
<td>36.43</td>
<td>41.658</td>
<td>13.040</td>
<td>13.877</td>
</tr>
<tr>
<td>3 1999</td>
<td>2002</td>
<td>0.96</td>
<td>0.90</td>
<td>0.74</td>
<td>0.63</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16.170</td>
<td>35.82</td>
<td>42.777</td>
<td>13.040</td>
<td>13.040</td>
</tr>
<tr>
<td>4 1998</td>
<td>2002</td>
<td>0.97</td>
<td>0.90</td>
<td>0.73</td>
<td>0.61</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15.686</td>
<td>35.82</td>
<td>44.985</td>
<td>13.187</td>
<td>27.383</td>
</tr>
<tr>
<td>5 1997</td>
<td>2002</td>
<td>0.97</td>
<td>0.84</td>
<td>0.69</td>
<td>0.53</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16.907</td>
<td>37.09</td>
<td>45.426</td>
<td>13.092</td>
<td>27.727</td>
</tr>
<tr>
<td>6 1996</td>
<td>2002</td>
<td>0.97</td>
<td>0.75</td>
<td>0.70</td>
<td>0.53</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.081</td>
<td>37.97</td>
<td>45.426</td>
<td>13.092</td>
<td>27.727</td>
</tr>
<tr>
<td>7 1995</td>
<td>2002</td>
<td>0.97</td>
<td>0.71</td>
<td>0.67</td>
<td>0.51</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.596</td>
<td>38.45</td>
<td>45.426</td>
<td>13.092</td>
<td>27.727</td>
</tr>
<tr>
<td>8 1994</td>
<td>2002</td>
<td>0.97</td>
<td>0.72</td>
<td>0.66</td>
<td>0.45</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.170</td>
<td>39.06</td>
<td>45.426</td>
<td>13.092</td>
<td>27.727</td>
</tr>
<tr>
<td>9 1993</td>
<td>2002</td>
<td>0.97</td>
<td>0.58</td>
<td>0.66</td>
<td>0.44</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.559</td>
<td>39.06</td>
<td>45.426</td>
<td>13.092</td>
<td>27.727</td>
</tr>
<tr>
<td>10 1992</td>
<td>2002</td>
<td>0.96</td>
<td>0.59</td>
<td>0.77</td>
<td>0.59</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.130</td>
<td>36.76</td>
<td>43.048</td>
<td>19.185</td>
<td>25.657</td>
</tr>
<tr>
<td>11 1991</td>
<td>2002</td>
<td>0.90</td>
<td>0.50</td>
<td>0.79</td>
<td>0.70</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23.181</td>
<td>34.20</td>
<td>46.135</td>
<td>19.107</td>
<td>26.189</td>
</tr>
<tr>
<td>12 1990</td>
<td>2002</td>
<td>0.92</td>
<td>0.50</td>
<td>0.81</td>
<td>0.77</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.380</td>
<td>33.40</td>
<td>46.135</td>
<td>19.107</td>
<td>26.189</td>
</tr>
<tr>
<td>13 1989</td>
<td>2002</td>
<td>0.96</td>
<td>0.50</td>
<td>0.81</td>
<td>0.77</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.480</td>
<td>33.40</td>
<td>46.135</td>
<td>19.107</td>
<td>26.189</td>
</tr>
<tr>
<td>14 1988</td>
<td>2002</td>
<td>0.95</td>
<td>0.50</td>
<td>0.81</td>
<td>0.77</td>
<td>(30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21.242</td>
<td>33.40</td>
<td>46.135</td>
<td>19.107</td>
<td>26.189</td>
</tr>
<tr>
<td>15 1987</td>
<td>2002</td>
<td>0.93</td>
<td>0.50</td>
<td>0.81</td>
<td>0.77</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21.984</td>
<td>33.40</td>
<td>46.135</td>
<td>19.107</td>
<td>26.189</td>
</tr>
<tr>
<td>16 1986</td>
<td>2002</td>
<td>0.89</td>
<td>0.50</td>
<td>0.81</td>
<td>0.77</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25.110</td>
<td>33.40</td>
<td>46.135</td>
<td>19.107</td>
<td>26.189</td>
</tr>
</tbody>
</table>

Note: In black, 5% significance level; black and overwritten, 1% significance level; grey, final solution. SSR: Sum of Squared Residuals (not computed with non-significant coefficients).