EL DESEMPLEO EN TÉRMINOS DE LOCALIZACIÓN. UNA ALICACIÓN A LAS PROVINCIAS ESPAÑOLAS.

López Hernández, Fernando A.
E-mail: Fernando.Lopez@upct.es
Universidad Politécnica de Cartagena
Palacios Sánchez, Maria de los Angeles.
E-mail: Mangeles.palacios@upct.es
Universidad Politécnica de Cartagena.
Ruiz Marín, Manuel.
E-mail: manuel.ruiz@upct.es
Universidad Politécnica de Cartagena.

Palabras Clave: Autocorrelación espacial, Modelos espaciales de Regresión, Tasa de Paro.

Área Temática: Métodos Cuantitativos.

Resumen:

En el trabajo de López y Palacios (2.000) *Distintos Modelos de Dependencia Espacial. Análisis de Autocorrelación* (ASEPEL) se presentaron tres modelos de dependencia espacial ligado a tres variables económicas observadas sobre el territorio nacional (considerando la división provincial). Las tres variables, Indice de Precios al Consumo, Renta Familiar Disponible y Tasa de Paro eran explicadas mediante un modelo de regresión en términos únicamente de la localización (latitud y longitud) obteniendo un superficie tendencia que explica en mayor o menor medida cada uno de estos fenómenos. Como se probó en el trabajo anteriormente citado, sólo en el caso de la Tasa de Paro se detectaba la presencia de autocorrelación espacial en los residuos del modelo de superficie tendencia.

Es objetivo de este trabajo el continuar con el análisis de esta variable, manteniendo el criterio inicial de explicarla únicamente en términos de localización. Cómo solución al problema de dependencia espacial en los residuos se analizarán diversos modelos de regresión espacial (Anselin 1.995) que permiten absorber esta dependencia a la vez que aumentar su capacidad explicativa de los modelos presentados. Queda patente con estos modelos la importancia de la distribución espacial de una variable a la vez que prueba la importancia de la posición relativa como un factor más que añade un importante potencial explicativo de estos fenómenos de carácter socio económicos.

1. La importancia del espacio.

Es incuestionable la importancia del espacio dentro del estudio estadístico de variables de carácter socio económico. Al analizar variables estadísticas en las que cada observación está ligada a una localización espacial (por ejemplo, Tasa de paro provincial, Renta Familiar Disponible, Indice de Precios al Consumo, etc.) no siempre se encuentra una distribución aleatoria de los valores de la variable en las distintas localizaciones, es más son frecuentes las ocasiones en las que la obtención de un valor en una determinada localización, viene en parte determinado por los valores observados en localizaciones próximas.

Este problema, de la autocorrelación espacial, ha sido profundamente analizado en el caso de las Series Temporales, donde ha recibido especial interés el estudio de la estructura de dependencia entre las observaciones de una sola variable (Box y Jenkins (1.970)). No parece razonable, y así está aceptado universalmente, adoptar el supuesto, en el caso de una sucesión de observaciones en el tiempo, que el valor que tome una variable en un instante de tiempo "t" es independiente de los valores que tomó la variable en los periodos que la han precedido. En este contexto los modelos autorregesivos temporales han alcanzado un elevado grado de desarrollo y han sido objeto de importantes avances teóricos.

Trasladar los resultados obtenidos en los procesos temporales a los espaciales no es una tarea simple. En el caso de la estadística aplicada a datos geo-referenciados el problema que se plantea es la pérdida de la condición de independencia de las observaciones tomadas en un área determinada. Así, en este tipo de modelos se supone que si se encuentra una determinada observación en un punto del área en estudio, es más fácil, o inversamente más difícil, encontrar observaciones semejantes en puntos próximos a éste que en puntos alejados. Esto es, en la localización de las observaciones hay latente una compleja información de la que no debe prescindirse a la hora de establecer estructuras matemáticas que describan el proceso. La complejidad del problema en el caso espacial aumenta con respecto a los procesos temporales debido a que las estructuras de interdependencias tienen carácter bidimensional y bidireccional frente a las estructuras unidimensionales y unidireccionales que se plantean en el tiempo.

Históricamente este problema, la dependencia espacial de variables geográficamente referenciadas, ha sido abordado por investigadores de distintas especialidades (Cliff y Ord, 1.981; Haining, 1.990; Ripley, 1.981; Cressie, 1991; Anselin 1.995) aportando sus contribuciones desde diversas ópticas. Estos investigadores han abordado diferentes aspectos de esta teoría estadística que van desde la elaboración de contrastes que permiten la detección de esta propiedad de dependencia en los datos, hasta la elaboración de complejos modelos que capten y cuantifiquen esta estructura de asociación espacial.

Las diferentes técnicas que permiten el desarrollo de esta teoría están adquiriendo un importante impulso en los últimos años, en parte fruto de dos factores que inciden positivamente en su desarrollo. Por un lado los Sistemas de Información Geográfica y el desarrollo de tecnologías informáticas que permiten un tratamiento más simple de la información desde el punto de vista espacial. Por otro lado, la disponibilidad de bases de datos que recogen la información con su referencia espacial, que permite a un mayor grupo de investigadores el tratamiento de este tipo de datos.

En el presente trabajo se pretende aplicar este tipo de técnicas espaciales para ofrecer un enfoque distinto a un problema clásico ampliamente tratado por los investigadores: la presentación de un modelo estadístico que explique el comportamiento de la variable tasa de paro. En el caso que nos ocupa, se supondrá que la única información que se dispone el la localización espacial de la variable, en nuestro caso las coordenadas de latitud y longitud de la capitalidad de la provincia al la que se refiera el dato.

El análisis de esta variable como un proceso estocástico espacial permite introducir dos niveles de variación, un primer nivel que se corresponde con la tendencia del proceso y que habitualmente se conoce como variación a gran escala, y un segundo nivel o variación a pequeña escala que se corresponde con la covarianza del proceso.

La idea que subyace en este trabajo es la de explicar ambos niveles de variación únicamente en términos de la localización de la variable, haciendo así patente la importancia del espacio como un factor más con una importante capacidad explicativa de cualquier fenómeno de carácter social o económico. Aplicaremos este análisis al caso de la distribución provincial de la tasa de paro en las provincias a nivel nacional.

En primer lugar, la tendencia del proceso se modelará mediante una superficie tendencia. Esta superficie tendencia explicará la variación Norte-Sur de la tasa de paro a nivel provincial. Así se introducirá la localización de la variable como un factor exógeno que explique el fenómeno. La superficie elegida es el plano ya que a la vez que capta suficientemente bien este efecto el un modelo simple. Formalmente se escribirá:

$$Z_i = \theta_0 + \theta_1 X_i + \theta_2 Y_i + e_i$$

Donde el subíndice indica cada una de las provincias, Z el la variable tasa de paro y X, Y las coordenadas asociadas a cada provincia.

En segundo lugar, se analizarán los residuos del modelo de regresión resultante (e_i) con la idea de contrastar la hipótesis de independencia que debe ser verificada por estos errores. En el caso de que esta hipótesis falle, como así ocurrirá, se pretende encontrar alguna estructura de dependencia espacial que modele esta dependencia, para lo cual se presentarán tres estructuras diferentes de variación a pequeña escala. Estos nuevos modelos permitirán a la vez que obtener errores independientes, aumentar la capacidad explicativa de los modelos.

Estructura del Trabajo.

En primer lugar (apartado 2.) se presentan de forma esquemática, los conceptos teóricos básicos para el desarrollo del trabajo. En el siguiente apartado (apartado 3.), se enumeran distintos aspectos que deben tenerse presente en el análisis: Variable a analizar, superficie en estudio, estructura de vecindades y ponderaciones de la estructura de conexiones. Con esta información se detecta la presencia de dependencia espacial en la variable (apartado 3.1) y se intenta modelar a través de un modelo de regresión mediante una superficie tendencia (apartado 3.2). Los residuos de estos modelos de regresión siguen presentando cierto grado de asocición espacial (apartado 3.3) que se modela mediante tres estructura diferentes de modelos de regresión con errores autocorrelados (apartado 4). Finalmente se exponen las conclusiones del análisis (apartado 5)

2. Conceptos Teóricos y modelos espaciales de regresión.

De una forma muy general en este apartado se recogerán los conceptos teóricos más importantes que se utilizarán en el desarrollo posterior del trabajo. Estos conceptos pueden verse de forma detallada en Cressie 1.993 o Anselin 1.995.

Los datos espaciales son medidas u observaciones que tienen asociada una localización específica. Las observaciones pueden ser de tipo discreto o continuo. La localización que tiene asociada cada observación puede ser un punto cualquiera de una determinada superficie, o bien estar asociada a un área de una superficie sobre la que se ha realizado una partición. Esta partición puede ser regular o irregular, y habitualmente se conoce como lattice[1] o retículo.

Formalmente, cualquier proceso espacial puede caracterizarse mediante un simple proceso estocástico (Cressie, 1993) . Sea $s \in R^d$ una posición genérica localizada en el espacio d-dimensional y sea Y(s) un valor aleatorio localizado en s. Consideremos ahora a "s" como un índice que varía sobre el conjunto $D \in R^d$, generamos entonces un proceso aleatorio

$$\{ Z(s) ; s \in D \}$$

cuyas realizaciones se escribirán como $\{z(s): s \in D\}$.

Bajo el supuesto de que el proceso mantiene cierta estructura de dependencia espacial, el estudio de este tipo de procesos presenta estructuras matemáticas capaces de describir las relaciones internas de dependencia de un conjunto de observaciones referenciadas geográficamente.

En el caso que nos ocupa supondremos que el proceso está indexado en un conjunto D finito sobre el que se ha establecido cierta estructura de conexiones, que analíticamente puede escribirse como:

$$D = \{(i;x_i,y_i) : i = 1, \ldots,n\}$$

Donde por (x_i,y_i) se denotan las coordenadas de latitud y longitud de un punto significativo de la i-ésima área.

La estructura de vecindades quedará recogida mediante una serie de conjuntos de la forma

$$N_j = \{k ; k \text{ es vecino de } j\}.$$

Que establece en base a criterios del investigador el conjunto de zonas que se relacionan o están conectadas con una determinada

1 Matemáticamente, una lattice viene definida por un conjunto de vértices y fronteras. Estas determinan un conjunto de zonas que podemos indexar y donde se establecerán criterios para determinar el grado de relación o vecindad de unas zonas con otras.

En estas n zonas en que queda dividida la superficie en estudio se tendrá definida una estructura de vecindades definida a través de una matriz cuadrada $W=\{w_{ij}\}$ de orden nxn. Esta matriz $W=\{w_{ij}\}$ determina si dos zonas son o no vecinas, de tal forma que si $w_{ij} = 0$ la zona i no es vecina de la zona j. En caso de que las dos zonas sean vecinas w_{ij} será distinto de 0 y este coeficiente se utilizará para medir la intensidad de la relación entre las dos zonas vecinas.

$$W = \begin{bmatrix} 0 & w_{12} & w_{13} & . & w_{1n} \\ w_{21} & 0 & w_{23} & . & w_{2n} \\ w_{31} & w_{32} & 0 & . & w_{3n} \\ . & . & . & . & . \\ w_{n1} & w_{n2} & w_{n3} & . & 0 \end{bmatrix}$$

Esta matriz no tiene porqué ser simétrica pudiendo plantearse relaciones unidireccionales o bidireccionales de diferente intensidad.

Si sobre esta superficie hemos observado cierta característica Z_i. La propiedad de dependencia espacial puede expresarse como una propiedad de Markov de un proceso de difusión espacial:

$$P(Z_i=y_i \mid Z_j \text{ con } j = 1, \ldots , n) = P(Z_i=z_i \mid Z_j \text{ con } j\in N_j)$$

De una forma muy general, como ya se mencionó en el anterior apartado, este tipo de datos asociados a referencias espaciales se suelen descomponer atendiendo al siguiente esquema:

Dato = variación a gran escala + variación a pequeña escala.

La componente de variación a gran escala suele ser el resultado de una tendencia global que puede ser modelada atendiendo o no a criterios espaciales (ajuste a un plano de regresión o a variables independientes). Esta primera componente se suele corresponder con la variación no estocástica del proceso.

La variación a pequeña escala es el residuo resultante de eliminar la tendencia. Una vez eliminada ésta podemos encontrar patrones de comportamiento que obedezcan al entorno en el que se encuentran las observaciones. Esta variación espacial tiene una estructura de dependencia estocástica y puede también modelarse atendiendo a criterios espaciales.

Restringiremos el análisis al caso de los procesos gaussianos de tal forma que el proceso puede escribirse como

$$Y \equiv N(\mu,V)$$

Se introducen distintas parametrizaciones de la tendencia del proceso μ y de la matriz de covarianzas V que se corresponden con diversas estructuras de dependencia espacial

En primer lugar, para el caso que nos ocupa, la tendencia del proceso se determinará mediante un modelo de regresión que utilice como variables exógenas las coordenadas geográficas asociadas a cada zona, así

$$\mu = A\,\theta$$

eligiendo para nuestra aplicación el modelo de un plano tendencia

$$A = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & x_n & y_n \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

que resulta

$$\mu = \theta_0 + \theta_1 X + \theta_2 Y$$

Manteniendo este mismo modelo para la tendencia del proceso se plantearan diversas alternativas sobre la estructura de covarianzas V distintas formas de recoger la dependencia espacial.

Se plantean en este trabajo tres estructuras de covariación conocidas en la literatura especializada como: modelo condicional autorregresivo (CAR), modelo simultaneo autorregresivo (SAR) y modelo de medias móviles (MA). Estos tres modelos asumen la normalidad multivariante de los datos con las siguientes matrices de dispersión:

CAR: $V = (I - \rho\, W)^{-1}\sigma^2$

SAR: $V = [\,(I - \rho\, W)^T (I - \rho\, W)]^{-1}\sigma^2$

MA: $V = (I + \rho\, W)\,(I + \rho W)^T\sigma^2$

Donde ρ y σ^2 son parámetros que deben estimarse en el modelo, y W es la matriz de ponderaciones antes indicada que establece la estructura de conexiones.

Tanto los modelos CAR como SAR se corresponden con los procesos autorregresivos de las series temporales. Los residuos del modelo SAR está correlacionados con los valores del entorno, resultando la estimación mínimo cuadrática inconsistente (Cressie, 1.993, pag. 408). El modelo CAR no tiene este problema, pero requiere que la matriz W sea simétrica. El modelo MA tiene su análogo en las series temporales con los modelos de medias móviles. Para más información en Cliff y Ord (1.981); Cressie (1.993) o Haining (1.990)

3. Modelos de superficie tendencia y tasa de paro.

Siguiendo la línea de razonamiento del trabajo de López y Palacios (2.000) los distintos aspectos de este análisis son los siguientes:

En primer lugar, la variable que se analizará será la tasa de paro a nivel provincial[2], como ya se comprobó en el trabajo antes mencionado, esta variable posee un alto grado de dependencia espacial que no es totalmente absorbido por la tendencia del proceso (en el este caso un plano de regresión) de tal forma que los residuos del modelo mantienen un importante grado de asociación espacial, lo que permite plantear los distintos modelos de regresión espacial antes señalados. El análisis que se realizó en el trabajo antes referenciado se ocupaba de analizar este modelo sólo para un año, pero para este trabajo se realizará sobre los años 1.990, 1.995 y 2.000, mostrando así la evolución temporal de la estructura de dependencia espacial.

En segundo lugar, la superficie a estudio es la Península Ibérica (sin tener en cuenta las islas). Cada una de las zonas en que se divide esta superficie se corresponde con las divisiones provinciales, en total tendremos 47 zonas. Para localizar espacialmente cada provincia se ha asignado las coordenadas de latitud y longitud correspondiente a la capital de la provincia por ser el punto más significativo.

En tercer lugar, las Vecindades: Los modelos de especificación de la dependencia espacial son sensibles a la elección del criterio de vecindades. Con el fin de obtener unos resultados fiables siempre es deseable plantear diversos escenarios, pero atendiendo a los resultados obtenidos el López y Palacios (2.000) se considerará únicamente un criterio para determinar las vecindades de cada zona que es el siguiente

---

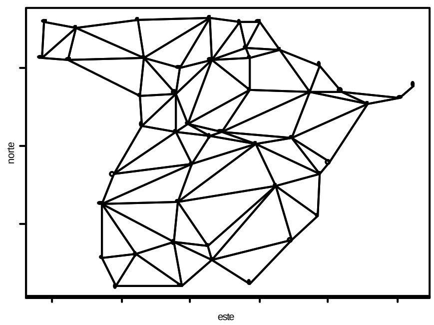
2 Media anual. Fuente http://www.ine.es . Banco de Datos TEMPUS

Dos provincias son vecinas si tienen frontera en común. La intensidad de esta relación es constante (independiente por ejemplo de la longitud de la frontera) y la misma para todas ellas. El siguiente gráfico presenta las provincias vecinas (en tono gris) a una dada (en color más oscuro).

Con este criterio la matriz de vecindades que obtenemos la nombraremos W y queda definida como:

$$w_{ij} = \begin{cases} 1 & \text{si las provincias tienen frontera común.} \\ 0 & \text{si las provincias no tienen frontera común.} \end{cases}$$

La estructura de conexiones que se genera mediante este criterio de vecindad, se puede presentar gráficamente observando la red de interrelaciones que liga las distintas provincias.

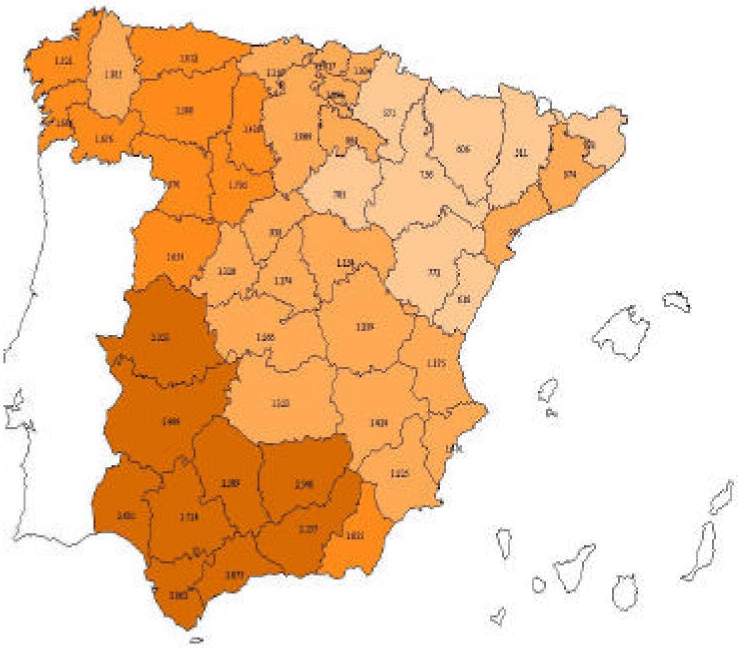


3.1. Estructura de dependencia espacial en la variable. Detección.

La presunción de dependencia espacial de la variable

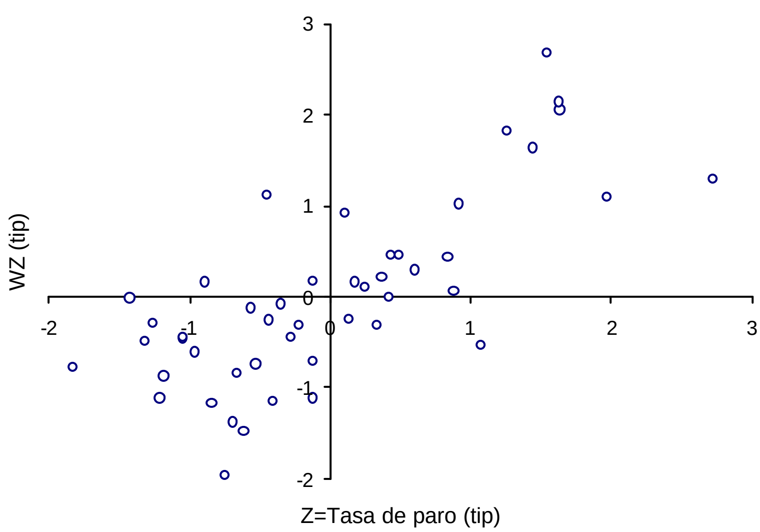
Z = Tasa de paro a nivel provincial

queda de manifiesto en el siguiente gráfico en el que se muestra la incidencia del desempleo en las distintas provincias de la península. La división en los cuatro niveles de color se ha realizado atendiendo a los correspondientes intervalos intercuartílicos de la variable, de tal forma, que el color más suave los valores se corresponde con valores de la variable que se encuentran en el primer intervalo intercuartilico y de color más fuerte los correspondientes a los siguientes intervalos intercuartílicos. Sólo se realiza el gráfico para la variable en el año 2.000, los otros son semejantes.

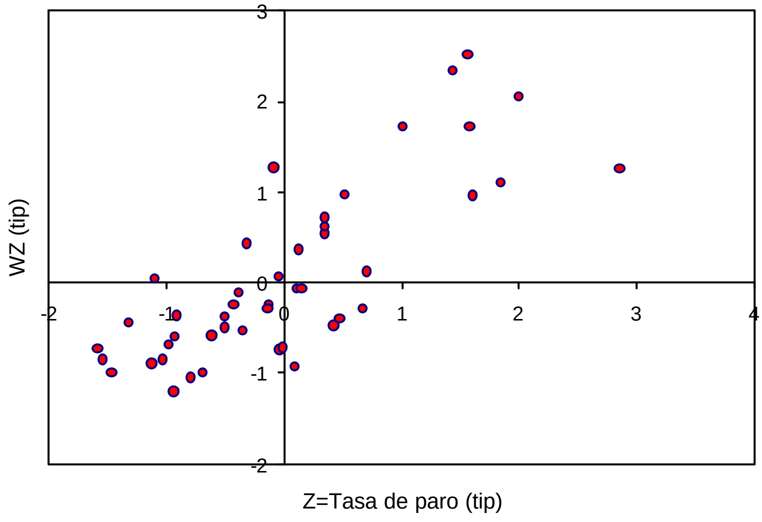


Otra forma de observar la fuerte presencia de dependencia espacial en la variable puede observarse en los siguientes gráficos en los que se muestran los llamados Scatterplot de Moran:

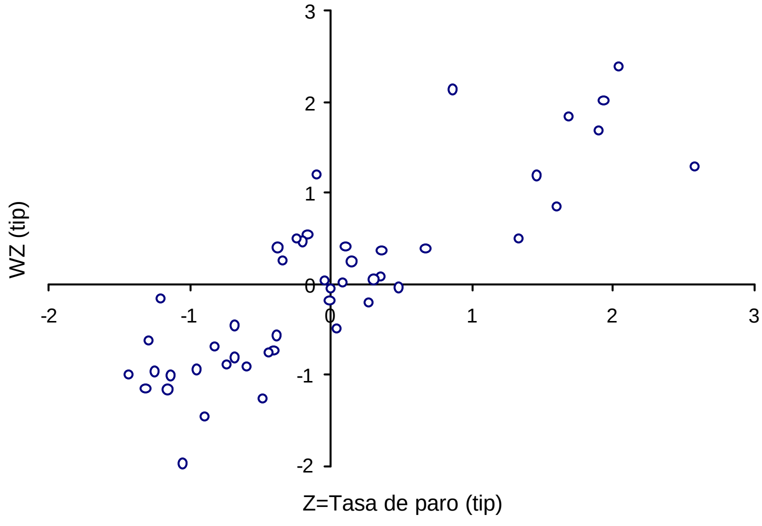
Scatterplot de Moran. Tasa de Paro 1.990



Scatterplot de Moran. Tasa de Paro 1.995



Scatterplot de Moran. Tasa de Paro 2.000



El índice de Moran, como medida cuantitativa del grado de asociación espacial de la variable presenta los siguientes resultados:

Test de Moran sobre la tasa de paro Z

	1.990	1.995	2.000
I Moran	0,555	0,611	0,699
Media E[I]	-0,056	-0,056	-0,056
Std I	0,0897	0,0897	0,0897
I Moran tipificado	**6,437***	**7,052***	**8,008***

*** Presencia significativa de correlación espacial con a =0,01**

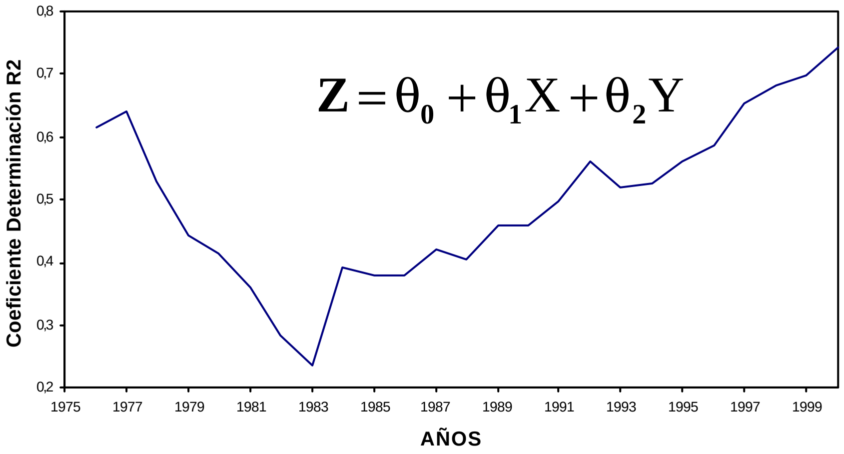
Resultados obtenidos con S+Spatialstats

En todos los casos se obtiene una fuerte dependencia espacial de la variable, como ya se sospechaba por los gráficos antes presentados. Esta dependencia espacial puede ser fruto de la tendencia espacial del fenómeno, y quizás pueda ser modelada introduciendo un modelo de regresión que explique la variación de la tasa de paro en términos de las coordenadas de longitud y latitud de un punto localizado en cada provincia.

3.2. Tendencia del proceso. ¿Se absorberá toda la dependencia espacial?

Una modelización de la tendencia del proceso mediante un plano de regresión indica una importante capacidad explicativa de la localización. El incremento del coeficiente de determinación en este tipo de modelos es llamativo y queda reflejado en la siguiente gráfica:

Ajuste a un Plano de Regresión



En el caso que nos ocupa y restringiéndonos a los años antes indicados los resultados del modelo de regresión aparecen en la siguiente tabla:

Resultados de la regresión.

	MCO		
	1.990	**1.995**	**2.000**
Constante (q_0)	82,522* (4,951)	112,969* (5,030)	79,407* (3,245)
Coor. Lat. (q_1)	-0,925* (0,277)	-0,917* (0,281)	-1,457* (0,181)
Cood. Long. (q_2)	-1,743* (0,345)	-2,348* (0,351)	-1,739* (0,226)
Ajuste, R^2	0,460	0,563	0,742
F (sign.)	63,341 (0,000)	28,281(0,000)	18,760(0,000)

* significativo con $\alpha = 0,01$

3.3. Análisis de los residuos. Test de dependencia espacial.

Como se puede observar, existe una fuerte tendencia Norte-Sur que explica en una parte importante la variación espacial del desempleo en la península. La pegunta que ahora surge es si mediante estos modelos se recoge toda la dependencia espacial presente en la variable o si queda parte sin explicar. En este sentido, el siguiente objetivo se debe centrar en el análisis de los residuos y la posible detección de autocorrelación espacial en éstos. Aunque en la literatura se dispone de una amplía batería de test para detectar la dependencia espacial en los residuos (Anselin 1.995) en este trabajo será suficiente con aplicar el test de Moran cuyos resultados se ofrecen en la siguiente tabla.

Test de Moran sobre los residuos del Modelo $Y = Aq + e$

	1.990	**1.995**	**2.000**
I Moran	1,208	0,9449	1,159
Media E[I]	-0,056	-0,056	-0,056
Std I	0,084	0,084	0,084
I Moran tipificado	**15,023***	**11,900***	**14,444***

*** Presencia significativa de correlación espacial con a =0,01**

Resultados obtenidos con la "toolbox" de LeSage bajo Matlab.

La presencia de autocorrelación espacial en los residuos de un modelo de regresión implica propiedades no deseables en la regresión MCO[3], de tal forma que deben plantease modelos alternativos.

4. Modelos espaciales de regresión con errores autocorrelados.

En esta sección se presentan los resultados correspondientes a los distintos modelos de regresión espacial que fueron expuestos en el apartado 2. Los resultados correspondientes a lo coeficientes del modelos pueden verse en la tabla adjunta (tabla 1)

En los tres años que han sido considerados en este trabajo 1.990, 1.995 y 2.000, puede apreciarse como cada uno de los modelos de regresión espacial capta mediante una estructura autorregresiva diferente, la dependencia espacial que aún subyacía en los residuos del modelo de regresión obtenidos por MCO.

La estimación de los parámetros que determinan cada uno de estos modelos se realiza por máxima verosimilitud, el proceso de inferencia es relativamente complejo y el software disponible para la obtención de éstos resultados es escaso. En este caso se ha utilizado un módulo que debe ser añadido al software estadístico S-PLUS conocido como S+SpatialStat específico para el análisis de datos espaciales.

En los todos los casos analizados, estos modelos mejoran de forma sustancial, la capacidad explicativa del modelo clásico, a la vez que modifican los coeficientes que ponderan cada una de las variables explicativas del modelo. Introducen también un coeficiente de dependencia espacial (ρ) cuyo valor es un indicador del grado de dependencia espacial de los residuos del modelo, y cuya contribución pueden entenderse como una corrección de los valores previstos para la tasa de desempleo en una provincia, en función de la tasa de desempleo de las provincias limítrofes.

Si bien es cierto que en los tres años considerados, el modelo que ofrece una mejor estimación es el modelo CAR, ya que ofrece un valor máximo para la función de verosimilitud (mínimo para la función de verosimilitud negativa que aparece en la tabla), también es cierto que las diferencias son mínimas, lo que no permite establecer una clara diferenciación cuantitativa entre los tres modelos espaciales. En este sentido, quizás el

3 La estimación MCO de modelos de regresión con residuos autocorrelados espacialmente produce estimaciones ineficientes, que conduce a sobreestimación de los parámetros del modelos (β) en comparación con la obtenida por MCG supuesto conocido el valor de ρ. Mas información Anselin (1.995).

criterio más adecuado sea seleccionar el modelo que desde el punto de vista interpretativo sea más adecuado.

5. Conclusiones.

Se han descrito la variación espacial del desempleo a nivel provincial haciendo uso exclusivamente de la localización de la variable. Lógicamente, los modelos que pueden obtenerse utilizando las coordenadas de localización de cada provincia como factores exógenos de un modelo de regresión son obviamente modelos con una capacidad explicativa restringida, sobre todo a la hora de explicar variables de carácter socioeconómico como es el caso. No obstante cabe destacar las siguientes conclusiones:

En Primer lugar la importancia del espacio como factor capaz de aportar respuestas a la variabilidad de una variable. En nuestro caso, sólo mediante una superficie tendencia se han podido obtener, modelos de regresión con un coeficiente de determinación creciente en el tiempo que alcanza un valor relativamente alto. Si se toma como punto de partida el año 1.983 con un coeficiente de determinación R^2 = 0,23 el crecimiento de este indicador es monótonamente constante para terminar en un valor R^2 = 0,74 para el año 2.000.

De esta primera conclusión pueden deducirse dos importantes cuestiones:

Por un lado la homogeneización del desempleo a nivel provincial: es indudable que de forma constante en el tiempo se está produciendo un proceso de homogeneización espacial, de tal forma que, aún manteniendo una tendencia Norte-Sur, la semejanza entre la tasa de desempleo en una provincia y sus limítrofes aumenta.

Por otra parte, la capacidad explicativa de la superficie tendencia está por encima de lo esperado, sobre todo en los últimos años. La tendencia del proceso tiene un peso específico más importante año tras año.

En segundo lugar, una vez obtenido el modelo de regresión deben ser examinados los residuos en busca de asociación espacial. La presencia de esta propiedad en los datos produce los efectos ya conocidos y antes mencionados sobre la estimación por mínimos cuadrados.

Este es el caso de esta aplicación, en la que de forma patente la tendencia del proceso no ha sido capaz de absorber toda la dependencia espacial en la variable y deben plantearse modelos alternativos que recojan esta estructura en el modelo. Algunos de estos modelos son presentados en este trabajo (otros no mejoran los presentados).

Por último cabría preguntarse sí la presencia de autocorrelación espacial en los residuos del modelo es debida al error en la estructura de superficie tendencia elegida y queda la pregunta de sí con una superficie tendencia de grado superior (por ejemplo un paraboloide seguirían quedando residuos correlacionados).

Bibliografía.

Anselin, L (1.988). *Spatial econometrics: Methods and Models* Kluwer Academic Publishers

Box G.E.P. y Jenkins G.M. (1.970). *Time Series Analisis, Forecasting and Contro*l (Holden-Day, San Francisco).

Cliff, A. D. y Ord, J. K. (1.981). *Spatial processes: Models and Applications* Pion Limited, London.

Cressie, N. (1.993). *Statistics for spatial data* John Wiley, New York.

Getis, A. y Ord J.K. (1.992) "The analysis of spatial association by use of distance statistis" *Geographical Analysis*, 24: 189-199.

Griffith D.A. (1.988) *Advanced Spatial Statistics* Kluwer

Haining R. (1.990). *Spatial data analysis in the social and environmental sciences* Cambridge University Press

López y Palacios (2.000) Distintos Modelos de Dependencia Espacial. Análisis de Autocorrelación. *XIV Reunión Asepelt España, Oviedo 22-24, Junio 2.000.*

Moran, P.A.P. (1.948) "The interpretation of stastistical maps" *Journal of the Royal Statistical Society B*, v.10, 243-251.

Ord, J.K. y Getis, A (1.995) "Local spatial autocorrelation statistics: distributional issues and an application" *Geographical Analysis*, 27: 286-296.

Ripley, B.D. (1.981). *Spatial Statistics* Wiley, New York.

Tabla 1: Resultados de los modelos de regresión.

	1.990				1.995				2.000			
	MCO	Máx. Verosimilitud			MCO	Máx. Verosimilitud			MCO	Máx. Verosimilitud		
		SAR	MA	CAR		SAR	MA	CAR		SAR	MA	CAR
Constante (q_0)	82,522	72,583	74,299	75,064	112,969	104,318	106,641	105,053	79,407	71,448	73,354	72,312
	(4,951)	(19,966)	(18,654)	(19,528)	(5,030)	(19,506)	(18,324)	(19,350)	(3,245)	(13,367)	(12,386)	(13,290)
Coor, Lat, (q_1)	-0,925	-0,871	-0,909	-0,8467	-0,917	-0,830	-0,876	-0,811	-1,457	-1,364	-1,418	-1,343
	(0,277)	(0,3611)	(0,353)	(0,347)	(0,281)	(0,3581)	(0,348)	(0,350)	(0,181)	(0,240)	(0,234)	(0,232)
Cood, Long, (q_2)	-1,743	-1,489	-1,538	-1,5456	-2,348	-2,125	-2,188	-2,139	-1,739	-1,524	-1,1581	-1,537
	(0,345)	(0,492)	(0,459)	(0,481)	(0,351)	(0,4806)	(0,451)	(0,476)	(0,226)	(0,329)	(0,305)	(0,327)
Coef, Dep, Espacial,(*)	----	0,102	0,117	0,1543	----	0,088	0,094	0,141	----	0,1078	0,123	0,165
		(0,032)	(0,038)	(0,042)		(0,035)	(0,030)	(0,032)		(0,031)	(0,032)	(0,038)
Neg Ln L		-160,9	-161,0	-161,1		-162,7	-162,8	-162,9		-141,0	-141,1	-141,1
Error est. residual	4.951	4,487	4,777	4,452	5.030	4,709	4,914	4,665	3.245	2,927	3,136	2,878
Ajuste R^2	0,460	0,556	0,543	0,563	0,563	0,616	0,590	0,6236	0,742	0,7904	0,782	0,7973

(*) Intervalo de variación de r (-0,3644 ; 0,1862)