

DISTINTOS MODELOS DE DEPENDENCIA ESPACIAL. ANÁLISIS DE AUTOCORRELACIÓN.

López Hernández, Fernando A.
Dpto. Ciencias Jurídicas, Métodos Cuantitativos e Informática.
Facultad de C.C. Económicas.
Universidad Politécnica de Cartagena.

Palacios Sánchez, M. Angeles.
Dpto. Ciencias Jurídicas, Métodos Cuantitativos e Informática.
Facultad de C.C. Económicas.
Universidad Politécnica de Cartagena.

La importancia del espacio dentro del estudio estadístico de variables económicas es incuestionable. En concreto el problema de la dependencia espacial, y en particular de la autocorrelación espacial, ha sido objeto de una gran cantidad de estudios (Cliff y Ord, 1981; Haining, 1990; Ripley, 1981; Cressie, 1993; Tiefelsdorf 2000)

El problema que se plantea es la pérdida de la condición de independencia de las observaciones tomadas en un área determinada. Hay latente en la localización de las observaciones una compleja información sobre las observaciones del entorno. La complejidad del problema aumenta debido a las interdependencias multidireccionales que se pueden plantear en el espacio.

Se presentan en este trabajo tres modelos de dependencia espacial que se ejemplarizan con tres variables económicas observadas sobre el territorio nacional (considerando la división provincial). Para la primera variable económica IPC encontramos el caso de independencia espacial, para la segunda variable, Renta Familiar Disponible se observa que la dependencia se pierde al eliminar la tendencia y por último, con la tasa de paro este fenómeno no se elimina totalmente al eliminar la tendencia espacial.

Palabras clave: Estadística Espacial, Autocorrelación Espacial, Índice de Moran.

E-mail: Fernando.Lopez@upct.es ; Mangeles.Palacios@upct.es

1. INTRODUCCION.

La importancia del espacio dentro del estudio estadístico de variables económicas es incuestionable. En concreto el problema de la dependencia espacial, y en particular de la autocorrelación espacial, ha sido objeto de una gran cantidad de estudios (Cliff y Ord, 1981; Ripley, 1981 ; Anselin 1988 ; Haining, 1990; Cressie 1993 ; Tiefelsdorf 2000).

La estadística presenta la mayor parte de sus resultados bajo el supuesto de independencia, cuando se viola esta condición, los resultados obtenidos ya no son válidos. Un caso típico de violación de la condición de independencia se presenta en el estudio de las series temporales. Los modelos de dependencia temporal han sido objeto de un profundo estudio y se conocen estructuras que expresan el comportamiento de las variables con respecto al tiempo. No ocurre lo mismo con aquellos modelos que intentan captar la dependencia de una variable en función de su "entorno". No es más importante la dimensión temporal del problema que la espacial, pero en el desarrollo científico no han recibido igual atención estos dos problemas, en buena parte debido a la complejidad que involucra las estructuras de dependencia espacial.

En el caso de la estadística espacial 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 encontramos una determinada observación en un punto de esta área es más fácil (o inversamente más difícil) encontrar observaciones semejantes en puntos próximos a éste que en puntos alejados. Es decir, hay latente en la localización de las observaciones una compleja información sobre las observaciones del entorno. La complejidad del problema aumenta debido a las interdependencias multidireccionales que se pueden plantear en el espacio.

Los datos espaciales son medidas u observaciones que tienen asociada una localización específica y de forma general podemos referirnos a ellos como datos georeferenciados. Estas 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 lattice¹ o retícula. En función de toda esta casuística se presentan diferentes modelos que llevan asociados diferentes tratamientos estadísticos.

2. EL MODELO ESPACIAL GENERAL.

Matemáticamente, cualquier proceso espacial puede caracterizarse mediante un simple proceso estocástico (Cressie, 1993) . Sea $s \in \mathbb{R}^d$ 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 $D \in \mathbb{R}^d$, generamos entonces un proceso aleatorio

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

cuyas realizaciones las escribiremos como $\{y(s); s \in D\}$.

Dependiendo de las características del conjunto D se diferencian tres tipos de datos:

Datos Geoestadísticos. D es un subconjunto fijo de \mathbb{R}^d que contiene un rectángulo d -dimensional con volumen positivo (en el plano D es una superficie). $Y(s)$ es un vector aleatorio localizado en $s \in D$.

Datos Lattice. D es un subconjunto fijo de \mathbb{R}^d (regular o irregular) formado por una cantidad contable de elementos. $Y(s)$ es un vector aleatorio localizado en $s \in D$. Asociado a D se presenta una estructura de vecindades.

Mapas de Puntos. D es un proceso puntual en \mathbb{R}^d o un subconjunto de \mathbb{R}^d . $Y(s)$ es un vector aleatorio localizado en $s \in D$.

De una forma muy general, los datos asociados a referencias espaciales se suelen descomponer atendiendo al siguiente esquema:

$$\text{Dato} = \text{variación a gran escala} + \text{variación a pequeña escala.}$$

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.

3. AUTOCORRELACION ESPACIAL.

Centraremos nuestro análisis en datos sobre retículas irregulares, esto es, en observaciones que supondremos asociadas a distintas áreas en las que se ha dividido una superficie. Cada área y por tanto a cada dato, le asociaremos unas coordenadas correspondientes a la latitud y longitud de un punto que sea representativo de la zona, en principio elegido a criterio del investigador.

Esta superficie debe dotarse de una estructura de vecindades que establezca las relaciones espaciales de los datos. De esta manera se define N_j como el conjunto formado por todas las zonas que son vecinas de la zona j , esto es $N_j = \{k ; k \text{ es vecino de } j\}$. Daremos criterios para determinar esta estructura de vecindades (por

¹ Matemáticamente, una lattice viene definida por un conjunto de vértices y fronteras. Estas determinan un conjunto de zonas que podemos indexar y sobre la que se define una estructura de vecindades.

ejemplo: dos zonas son vecinas si tienen frontera en común, o si se encuentran a menos de una determinada distancia)

Supondremos que tenemos una superficie que ha sido particionada en n zonas. En estas n zonas tendremos definida una estructura de vecindades definida a través de una matriz $W = \{w_{ij}\}$ de orden $n \times n$. 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 ($w_{ij} \neq 0$) este coeficiente mide la intensidad de la relación. Esta matriz no tiene porqué ser simétrica pudiendo plantearse relaciones unidireccionales o bidireccionales de diferente intensidad. Sobre esta superficie hemos observado cierta característica Y_i .

La propiedad de autocorrelación espacial puede expresarse como una propiedad de Markov de un proceso de difusión espacial:

$$P(Y_i = y_i | Y_j \text{ con } j = 1, \dots, n) = P(Y_i = y_i | Y_j \text{ con } j \in N_j)$$

El objetivo es determinar si la presencia de un valor concreto de una variable en un lugar del espacio hace más verosímil (o menos) la presencia de ese valor o valores semejantes en un entorno de ese lugar del espacio. Cuando la presencia de un valor en una zona facilita que ese valor ocurra en lugares próximos hablaremos de correlación espacial positiva. De otra manera, si la presencia de un valor en una zona del espacio hace menos verosímil que valores semejantes ocurran en su entorno, hablaremos de correlación espacial negativa. Y finalmente si no existe esta relación en ninguno de los sentidos que hemos señalado (positivo o negativo) supondremos independencia o, dicho de otra manera, autocorrelación espacial nula.

Una de las razones más importantes por las que se debe detectar la autocorrelación espacial en un conjunto de observaciones es por las consecuencias catastróficas que tiene sobre los procesos de inferencia en modelos de regresión (Estimación ineficiente de los coeficientes del modelo, estimación sesgada (subestimación) de la varianza residual, R^2 inflando, valores erróneos de t y F e Inconsistencia)

Con el fin de poner de manifiesto la presencia de autocorrelación espacial en un conjunto de datos se han presentado en la literatura diferentes estadísticos que son capaces de evaluar la ausencia o carencia de esta propiedad en un conjunto de observaciones. Ejemplos más relevantes de estos estadísticos son el debido a Geary (Geary 1.954), el de Moran (Moran 1.948), y uno más general dado por Knox (Knox 1.964) para procesos espacio temporales.

De entre todos ellos hemos seleccionado para nuestro análisis el debido a Moran. Este es el más utilizado con diferencia y se basa en el cálculo de los productos cruzados $z_i z_j$. Siendo $z_i = y_i - \bar{y}$. En este caso, una agrupación de valores próximos dará como resultado unos productos que tenderán a ser positivos mientras que en un modelo de alternancias los productos tenderán a ser negativos.

Un test basado en este principio se genera mediante el estadístico I de Moran, cuya expresión analítica es la siguiente:

$$I = \frac{n}{A} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} Z_i Z_j}{\sum_{i=1}^n Z_i^2}$$

Siendo

$$A = \sum_{i,j} w_{ij}$$

Este estadístico puede tomar valores positivos y negativos dependiendo del signo de la correlación espacial, de tal forma que si la correlación es positiva obtendremos valores de I positivos y viceversa, si la correlación espacial es negativa obtendremos valores del índice I negativos.

Este estadístico bajo la hipótesis de independencia de las observaciones sobre la superficie tiene una distribución asintótica normal y se pueden obtener sus dos primeros momentos². Así se construye un test que permite medir la significación de la autocorrelación espacial.

Un profundo estudio sobre este estadístico y sus momentos y su distribución exacta puede verse por ejemplo en Cliff y Ord 1.981 (Capítulos 1 y 2), Tiefelsdorf (2.000)

Este indicador espacial, el Índice de Moran, mide de forma global la presencia o ausencia de autocorrelación espacial en un conjunto de observaciones distribuidas sobre una superficie. Lógicamente este fenómeno de mimetismo entre zonas no tiene porqué ser uniforme en todo el espacio, así no todas las zonas ejercen la misma influencia en su entorno, y tampoco con el mismo grado. Con la idea de cuantificar la contribución que cada zona aporta al Índice de Moran surgen los indicadores de autocorrelación parciales (ver Anselin 1.995; Ord y Getis 1.995).

Estos indicadores alcanzan su máximo sentido, tanto a la hora de analizar las razones de la presencia de autocorrelación global, como en aquellas situaciones en las que globalmente no se puede afirmar la presencia de este fenómeno, pero que es de destacar su presencia en zonas determinadas.

4. APLICACIÓN EMPÍRICA.

Tras esta breve presentación teórica vamos a presentar una pequeña aplicación con el fin de ilustrar de forma empírica el procedimiento para detectar la presencia de autocorrelación espacial en un conjunto de datos que se encuentran distribuidos sobre una superficie particionada de forma irregular. Con esta idea vamos a realizar

² Los momentos serán distintos dependiendo de la distribución que se considere para Y: (1) Normalidad: supone que las $\{y_i\}$ son n extracciones independientes de una distribución Normal y (2) Aleatoriedad: El mapa observado es una realización de los n valores observados. En total podemos encontrarnos n! formas distintas de observar estos valores.

el análisis correspondiente a tres variables económicas que presentan distintos modelos de interrelación espacial.

En el primer caso, presentaremos una variable que no presenta síntomas de mimetismo espacial, de forma que la observación obtenida en una zona es independiente de lo que observamos en su entorno. Para este ejemplo se ha elegido la distribución provincial del Índice de Precios al Consumo (I.P.C.)

En un segundo caso, presentamos dos variables que si poseen esta propiedad de dependencia espacial, y que el Índice de Moran detecta de forma significativa su presencia en las observaciones. Ahora bien, esta característica puede presentarse a distintos niveles.

Como primer exponente de este segundo caso presentaremos un ejemplo en el que la dependencia espacial está presente en lo que se denomina variación a gran escala (también tendencia), de tal forma que al eliminar esta tendencia los residuos resultan ser independientes. Para este caso se ha elegido la distribución provincial de la Renta Familiar Disponible per cápita (RFD)

El segundo exponente de la presencia de autocorrelación espacial se manifiesta en la variación a pequeña escala. Una vez eliminada la tendencia espacial mediante el ajuste a un plano de regresión, los residuos siguen manifestando esta propiedad, aunque de una forma más débil. El ejemplo se presenta con la distribución provincial de la Tasa de Paro (TP).

Los distintos aspectos de este análisis son los siguientes:

La Superficie: 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.

Las Vecindades: Los indicadores de autocorrelación espacial son sensibles a la elección del criterio de vecindades. Con el fin de obtener unos resultados fiables siempre es deseable plantear diversos escenarios. Con esta idea hemos fijado tres criterios para determinar las vecindades de cada zona y con cada uno de estos criterios evaluaremos la presencia o ausencia de correlación espacial.

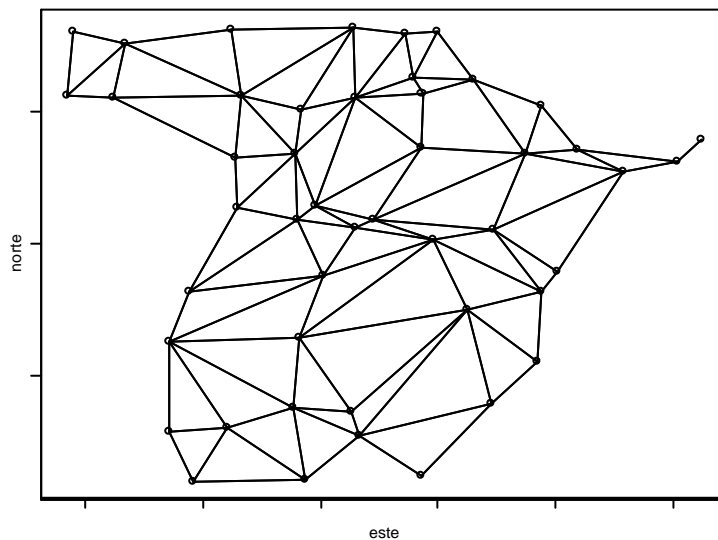
Criterio A: Dos provincias son vecinas si tienen frontera en común. La intensidad de esta relación es constante (independiente por ejemplo de la longitud de la frontera) y la misma para todas ellas.



Ilustración 1

Con este criterio la matriz de vecindades que obtenemos la nombraremos W1.

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.



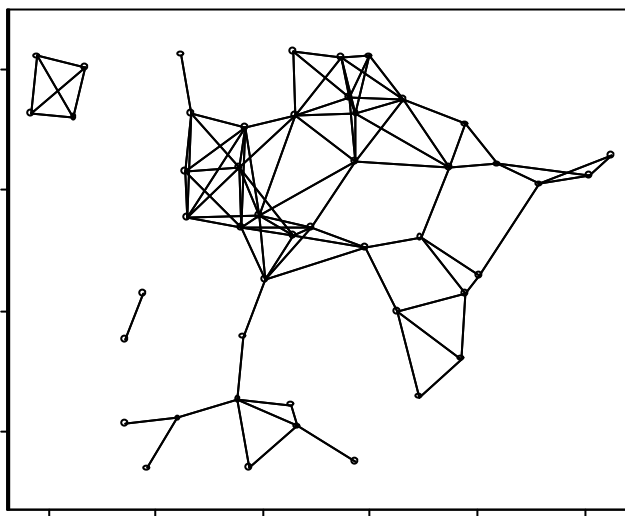
Criterio 2: Dos provincias son vecinas si las capitales de provincia se encuentran a menos de 200 Km. En este caso la intensidad de la relación entre las distintas zonas no es constante como en el caso anterior. Planteamos el modelo de tal forma que la intensidad de la relación sea inversamente proporcional a la distancia euclídea entre dos capitales de provincia.

Concretamente

$$w_{ij} = \begin{cases} 0 & \text{si } i, j \text{ no son vecinas} \\ \frac{1}{d_{ij}} & \text{si } i, j \text{ son vecinas} \end{cases}$$

A esta matriz de conexiones la llamaremos W2.

La estructura de vecindades que genera este criterio da lugar a la siguiente malla:



Criterio 3: Dos provincias serán vecinas si sus capitales de provincia se encuentran a menos de 200 Km. El modelo que determina la intensidad de la relación entre zonas vecinas lo seleccionamos con el fin de presentar una relación NO SIMETRICA. La relación existente entre dos zonas será inversamente proporcional a la distancia y directamente proporcional al cociente entre las poblaciones. La expresión de la matriz de conexiones es la siguiente:

$$w_{ij} = \begin{cases} 0 & \text{si } i, j \text{ no son vecinas} \\ \frac{1}{d_{ij}} \frac{P_j}{P_i} & \text{si } i, j \text{ son vecinas} \end{cases}$$

Con esta matriz establecemos el criterio por el cual las provincias más pobladas ejercen más influencia que las menos pobladas dentro de una misma vecindad. La matriz de vecindades en este caso la llamaremos W3.

Las Variables: Las variables que hemos seleccionado para este ejemplo son las siguientes:

a) El Índice de Precios al Consumo (IPC), base 1.992. Media anual. Fuente <http://www.ine.es> . Banco de datos TEMPUS.

b) La Tasa de Paro (TP). Media anual. Fuente <http://www.ine.es> . Banco de Datos TEMPUS.

c) La Renta Familiar Disponible neta per cápita (RFD) Fuente: BBV

Con el fin de situar todas las variables en el mismo plano temporal hemos seleccionado el año 1.993 por ser el último del que disponemos información a nivel provincial de la última variable RFD.

Los modelos: Como ya indicamos en un apartado anterior, podemos detectar dos niveles de variación en la distribución espacial de la variable observada. El primer nivel es una variación a gran escala, resultado de los cambios del proceso (en media) debido a la localización espacial (o quizás a otras variables explicativas). El segundo nivel se llama variación a pequeña escala y recoge la interacción entre las zonas vecinas.

$$Y_i = \mu_i + \delta$$

Donde Y_i es el proceso en la zona i ; μ_i es la media en la zona i , que puede ser constante o ajustarse a un modelo de regresión; $\delta \equiv N(0; \Sigma)$ donde Σ es la matriz de covarianzas que determina el modelo de variación espacial a pequeña escala.

Una de las razones por las que se advierte autocorrelación espacial en un conjunto de datos es por la presencia de una tendencia asociada a la localización espacial de las observaciones (en nuestro caso un plano de regresión). Esta tendencia es la desencadenante en muchas ocasiones del fenómeno de la autocorrelación espacial. Ahora bien, una vez eliminada esta tendencia debemos estudiar si en los residuos sigue quedando latente dependencia entre residuos colindantes o vecinos.

Para cada una de las variables que antes hemos enumerado obtendremos la tendencia del proceso mediante un plano de regresión. Una vez eliminada esta tendencia contrastaremos la presencia o ausencia de autocorrelación espacial en los residuos del modelo.

Análisis de la Tendencia.

En este apartado el objetivo es detectar la presencia de tendencia y evaluarla para cada una de las variables que estamos estudiando.

En un análisis gráfico previo obtenemos los siguientes resultados:

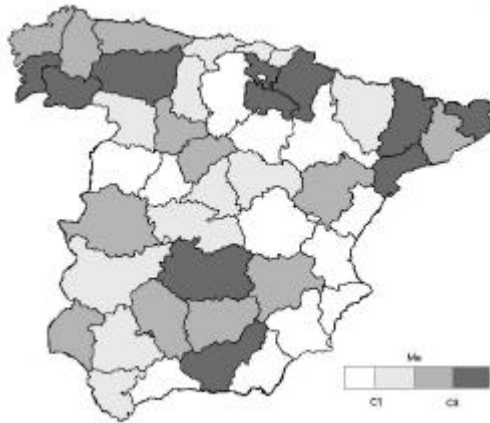


Ilustración 2: Cuartiles de IPC

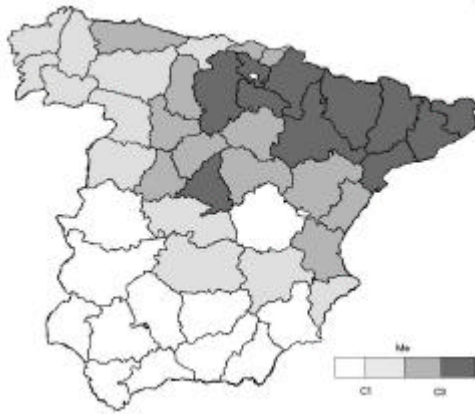


Ilustración 3: Cuartiles de la RFD



Ilustración 4: Cuartiles de TP

Hemos coloreado cada región en función de los valores que toma cada una de las variables. En cada caso hemos considerado dividida la distribución en cuatro partes iguales, coloreando de blanco las provincias con valores inferiores al primer cuartil, de gris claro las zonas comprendidas entre el primer cuartil y la mediana, con gris oscuro las provincias con valores comprendidos entre la mediana y el tercer cuartil y de color negro aquellas provincias con valores superiores al tercer cuartil.

Se desprende de forma automática la ausencia de tendencia en los datos correspondientes al IPC. Por el contrario, las otras dos variables presentan una marcada tendencia. En el caso de la RFD casi todas las observaciones superiores al tercer cuartil están localizadas en la zona norte mientras que los valores inferiores al primer cuartil se encuentran todos al sur. Para la TP la agrupación es muy semejante aunque de forma inversa a lo que ocurre con la RFD. Casi todos los valores superiores al tercer cuartil se encuentran agrupados al sur de la península y los valores inferiores al primer cuartil ocupan posiciones localizadas al norte.

Para cuantificar esta impresión calculamos el Índice de Moran para cada una de las situaciones de vecindad antes planteadas, obteniendo los siguientes resultados (podemos verlos en el anexo)

I.Moran	IPC	RFD	TP
W1	0,38	7,82*	6,65*
W2	0,44	6,28*	4,18*
W3	0,34	2,64*	1,69**

*Significativas para $\alpha = 0,05$

**Significativas para $\alpha = 0,01$

Con lo que queda de manifiesto la dependencia espacial de las observaciones.

Para determinar la superficie tendencia realizaremos un ajuste por mínimos cuadrados a un plano de regresión atendiendo al siguiente modelo.

$$Y_i = \beta_0 + \beta_{10}x_i + \beta_{01}y_i$$

siendo (x_i, y_i) las coordenadas cartesianas (latitud, longitud) de cada capital de provincia.

Los resultados correspondientes a las tres variables son los siguientes:

	I.P.C.	R.F.D.	T.P.
b₀	103,23	825153,12	1,08
b₁₀	-0,02	36236,61	-0,008
b₀₁	0,03	51226,32	-0,002
r	0,18	0,89	0,72

Atendiendo a los resultados obtenidos detectamos la presencia de tendencia mediante el ajuste a un plano de regresión sólo para las variables RFD y TP.

Para el IPC tenemos un modelo de tendencia constante; μ_i será constante e igual a la media del proceso. Gráficamente supondremos un plano tendencia paralelo a la superficie.

$$\mu_i = \mu = \text{media del proceso}$$

Para la variable RFD tenemos claramente una tendencia que se ajusta a un plano de regresión

$$\mu_i = 825153,12 + 36236,61 x_i + 51226,32 y_i$$

Para la última variable TP tenemos también una tendencia que se ajusta a un plano de regresión

$$\mu_i = 1,08 - 0,008 x_i - 0,002 y_i$$

Variación a pequeña escala.

Vamos a determinar los residuos de cada variable eliminando la tendencia según los modelos planteados en el apartado anterior.

Para obtener una imagen gráfica, al igual que para la variable original, realizamos gráficos equivalentes a los que presentamos en la sección anterior.



Ilustración 5: Residuos del IPC



Ilustración 6: Residuos de RFD



Ilustración 7: Residuos de TP

Una vez con los residuos de cada uno de los modelos realizaremos los contrastes de hipótesis correspondientes para determinar si sigue presentándose autocorrelación espacial.

Test de Autocorrelación Espacial.

a) Variable IPC. Con esta variable se presenta el primero de los tres modelos de dependencia espacial. Este es el caso de ausencia de autocorrelación espacial. La variable no presenta tendencia o, dicho de otra manera, la tendencia es constante (ajustamos a un plano horizontal de altura la media de los datos). Los residuos resultantes de eliminar esta tendencia se someten a un test de autocorrelación espacial basado en el estadístico de Moran. Los resultados son evidentes visto los datos localizados espacialmente: No existe autocorrelación espacial en las observaciones y estas se distribuyen espacialmente de forma independiente unos de otros. Esto es igualmente cierto con cualquiera de los tres criterios de vecindad que hemos elegido

Los resultados del test de Moran se presentan en la siguiente tabla:

	W 1	W 2	W 3
I Moran	0,013	0,022	0,048
Media E[I]	-0,021	-0,021	-0,021
Std I	0,091	0,105	0,206
I Moran tipificado	0,380	0,417	0,335

b) Variable RFD. Este segundo caso que presentamos se corresponde con un modelo donde los datos originales presentan una fuerte correlación espacial tal y como puede observarse en el gráfico 3.

Observando una agrupación importante de zonas con elevada RFD en la región catalana que disminuye de forma progresiva según avanzamos hacia el sur, Extremadura y Andalucía son las regiones con menor RFD.

Al determinar la tendencia, constatamos esta fuerte evolución espacial determinando un plano de regresión con un elevado coeficiente de correlación ($r = 0,88$). Ahora bien, una vez eliminada esta tendencia, los residuos del modelo que gráficamente hemos representado en la ilustración 6, aparecen distribuidos de forma independiente en el espacio.

Confirmación de esta afirmación son los resultados que obtenemos con el estadístico de Moran. Al igual que en el caso anterior para los tres modelos de vecindades encontramos valores semejantes.

En este caso toda la dependencia espacial de nuestras observaciones ha sido absorbida por la tendencia de la que es fruto.

	W 1	W 2	W 3
I Moran	0,089	0,090	0,444
Media E[I]	-0,021	-0,021	-0,021
Std I	0,090	0,102	0,203
I Moran tipificado	1,229	1,096	2,289*

c) **Variable TP.** El último de los modelos que presentamos. Analizada la tendencia espacial de los datos observamos una fuerte tendencia sur-norte. Eliminamos la tendencia y observamos los residuos del modelo Ilustración 7.

Como podemos observar sigue persistiendo una cierta agrupación espacial de valores semejantes en los residuos. Analizando estos mediante el estadístico de Moran obtenemos los siguientes resultados:

	W1	W2	W3
I Moran	0,258	0,304	0,387
Media E[I]	-0,021	-0,021	-0,021
Std I	0,091	0,102	0,203
I Moran tipificado	3,063*	3,190*	2,013*

* Presencia significativa de correlación espacial con $\alpha=0,05$

6. CONCLUSIONES

Hemos descrito, a través de tres ejemplos, modelos que estudian de forma conjunta las observaciones de una variable y su localización espacial. El estudio se centra en determinar la presencia de dependencia estadística entre estas dos variables: las observaciones y su localización espacial. De los tres ejemplos presentados, el primero se corresponde con el caso de independencia, la ausencia de autocorrelación espacial en las observaciones (bajo diferentes criterios de vecindades). En los otros dos casos constatamos una fuerte agrupación espacial de observaciones semejantes induciendo la presencia de autocorrelación espacial positiva. Estos dos modelos son distintos. Para el primero de ellos la autocorrelación se pierde una vez que eliminamos la tendencia espacial del modelo, ya que la localización de las observaciones a través de sus coordenadas absorbe totalmente la dependencia espacial. En el segundo de los casos esta autocorrelación persiste aunque de forma más débil ya que parte de la dependencia espacial se ha absorbido por la tendencia del proceso en base a las coordenadas.

El siguiente objetivo que debemos marcarnos es el de encontrar modelos matemáticos que modelen la matriz de covarianzas de los residuos. Se liga así este trabajo con los modelos de regresión espaciales y su problemática (Anselin 1988).

7. BIBLIOGRAFIA

- Anselin, L (1.988). *Spatial econometrics: Methods and Models* Kluwer Academic Publishers
- Anselin, L. (1.995). "Local Indicators of Spatial Associations-LISA." *Geographical Analysis*, vol.27:93-115
- 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.
- Geary, R.C. (1.954) "The contiguity ratio and statistical mapping" *The Incorporated Statistician* 5,115-145

Getis, A. y Ord J.K. (1.992) "The analysis of spatial association by use of distance statistics" *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

Knox E.G. (1964). "The detection of space-time interactions" *Applied Statistics* 13, 25-29

Moran, P.A.P. (1.948) "The interpretation of statistical 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.

Tiefelsdorf, M. (2.000). *Modelling Spatial Processes* Springer-Verlag.