

Actas del V Coloquio de Geografía Cuantitativa
Universidad de Zaragoza
1992, Zaragoza

INTEGRACION DE LA ESTRUCTURA S.I.G Y EL ANALISIS FACTORIAL. APLICACION A LA MODELIZACION DE PROCESOS EROSIVOS.

María Jesús PERLES ROSELLO

Departamento de Geografía. Universidad de Málaga

El análisis factorial constituye uno de los más eficaces procedimientos estadísticos para la organización de la información. De igual modo, la capacidad de análisis espacial que aportan los sistemas de información geográfica, amplían considerablemente las posibilidades de estrategias metodológicas de acercamiento a procesos que por su naturaleza multivariable y expresión espacial analítica, como es el caso de los procesos erosivos, han sido abordados tradicionalmente a una escala muy detallada, mediante estudios puntuales.

Esta escala de aproximación aporta una caracterización muy específica de los elementos y factores que intervienen en el proceso erosivo, así como de las interrelaciones entre ellos establecidas, como consecuencia de la profundización que la reducción del área de estudios posibilita. Sin embargo, la consecuencia inmediata de esta reducción radica en las dificultades de generalización de los resultados obtenidos, por su carácter puntual.

En este punto, las soluciones al problema de la generalización siguen dos vías: la extrapolación de resultados puntuales obtenidos mediante muestreo, o el uso de una escala poco detallada de trabajo, partiendo a priori de información ya generalizada por tratarse de valores medios.

En este contexto, nuestro trabajo propugna un uso combinado de las ventajas aportadas por el análisis S.I.G y el factorial para el manejo de la información a una escala media, intentando aproximar los conceptos de

precisión en el conocimiento de las variables y de operatividad en la generalización.

OBJETIVOS

La intención general que subyace al presente estudio es la de establecer una aproximación metodológica que permita, mediante una estrategia operativa a una escala media, la caracterización de las variables que inciden en la erosión, considerando principalmente los siguientes aspectos:

- Estimación de la incidencia relativa de cada una de las variables en el proceso, analizandolas simultaneamente, como componentes de un sistema.
- Análisis de las posibles variaciones que esta distribución del peso de las variables pueda experimentar en función de su observación en distintos ámbitos espaciales.

Esta procedimiento nos permitiría establecer las claves de ponderación de las distintas variables a la hora de valorar el proceso erosivo como la conjunción de todas ellas. Será útil igualmente como mecanismo de comprobación de la incidencia de variables no directamente relacionadas con la erosión, para su consideración en el proceso, o por contra, su eliminación definitiva.

De otro lado, la conexión de la información con su posición geográfica nos permitirá deducir las condiciones que determinan los cambios en la distribución del peso de las variables, efectuando ensayos en geosistemas de características diferenciadas. Una vez observadas las pautas que rigen estas variaciones, habremos obtenido una clave fundamentada y no apriorística para la extrapolación y predicción del comportamiento de las variables en distintos ámbitos espaciales.

Bajo estas premisas generales de trabajo, enmarcadas en un proceso de investigación más general, esta comunicación pretende, como objetivo específico, mostrar los resultados de un primer ensayo del procedimiento; se hará mayor hincapié en los aspectos relacionados con los problemas de integración de la estructura resultante del análisis espacial S.I.G. con la del análisis factorial, aportandose las vías de

solución adoptadas, y las modificaciones orientadas a la implementación de las potencialidades de ambos sistemas.

METODOLOGIA

El primer paso en la elaboración del trabajo lo constituyó la selección de las variables a considerar para el análisis factorial. No es objeto de esta comunicación el abundar en las pautas de selección y recopilación de la información referente a estas variables, por lo que su enumeración se ha obviado.

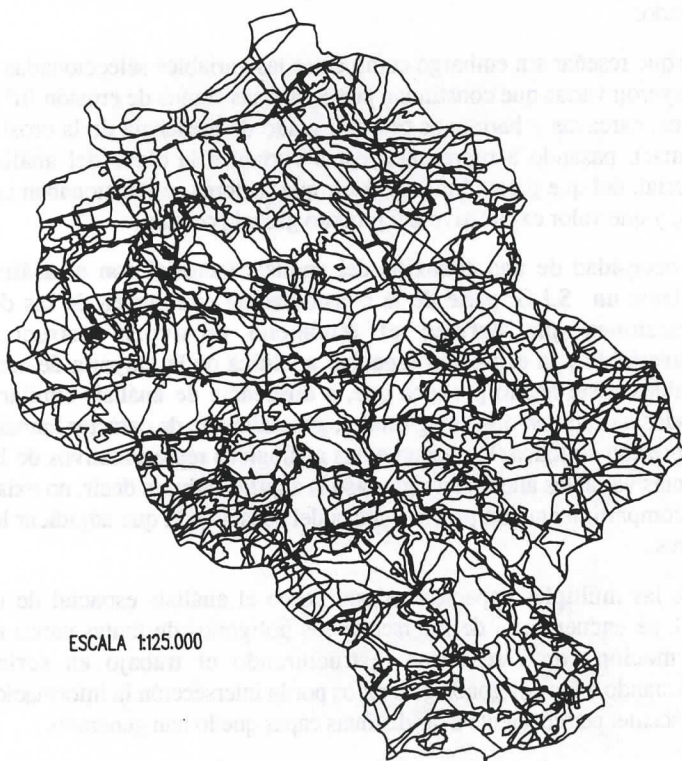
Hay que reseñar sin embargo como entre las variables seleccionadas se incluyeron varias que constituían por si mismas signos de erosión (nº de surcos, cárcavas y barrancos por km, grado de presencia de la erosión laminar), pasando a representar estas variables la clave del análisis factorial, del que pretendimos deducir qué factores se relacionaban con ellas, y que valor explicativo del proceso global poseían.

La necesidad de una conexión del análisis factorial con el análisis mediante un S.I.G. parte de la consideración simultánea de las dos dimensiones que definen el fenómeno erosivo: su estructura multivariable y la expresión espacial analítica de la mayoría de estas variables. Este hecho provoca que, a diferencia de análisis similares llevados a cabo con otras magnitudes, generalmente de carácter humano y expresión sintética, los documentos análogos representativos de las distintas variables alteren en cada caso la zonificación, es decir, no existe una compartimentación previa y única del espacio a la que adjudicar los valores.

Entre las múltiples capacidades que posee el análisis espacial de un S.I.G. se encuentra la de intersectar los polígonos de tantas capas de información como se desee (estructurando el trabajo en serie), adicionando a los polígonos generados por la intersección la información subyacente, perteneciente a las distintas capas que lo han generado.

Fue esta prestación la que se utilizó como solución al carácter analítico de la zonificación antes citado. El primer paso lo constituyó la digitización (S.I.G. vectorial), de la cartografía correspondiente a cada variable, capas de información que tras su procesamiento y adjudicación de atributos, pasaron a constituir coberturas (información espacial+ base de datos).

MAPA N.1:
RESULTADO FINAL DE LA INTERSECCION DE COBERTURAS.
(Cuenca del Rio de la Cueva)



A continuación se procedió a la intersección en serie del total de estas coberturas, obteniéndose de este modo una cobertura final en la que quedaban recogidas todas las posibles variaciones en la combinación de los atributos de las variables (Ver mapa nº1) Estos valores pasarían a constituir la base del análisis factorial.

Este procedimiento aporta las siguientes ventajas:

Se soluciona el problema de la expresión analítica de las variables sin necesidad de acudir a la definición de unidades a priori para homogeneizar la adjudicación de datos, ni a la adjudicación de valores medios a estas unidades.

Se produce un aumento del número de casos a observar, por la multiplicación de los polígonos. Este hecho que en caso de no contar con instrumentos potentes de estructuración de la información, constituiría un obstáculo a la aprehensión de la información, beneficia sin embargo la fiabilidad del análisis factorial.

Por último, hay que reseñar la idoneidad, a nivel operativo, de la estructuración de la información de los dos sistemas que se pretenden combinar. Así, el sistema de información geográfico, genera un archivo de base de datos con la ordenación de casos y variables requeridas por el factorial, y fácilmente transferible.

Reseñadas las ventajas deducidas de la complementación de los dos procedimientos de análisis, se recogen a continuación los principales problemas planteados y las soluciones adoptadas:

Abundando en el concepto de superposición, al intersectar capas cartográficas con arcos coincidentes en algunos tramos, pudimos observar la creación de concatenaciones de "falsos polígonos" de pequeñísimas proporciones que no se ajustaban a elementos cartográficos reales. Estos polígonos surgen como consecuencia del porcentaje de error introducido en el proceso de digitalización de los contornos, efecto del error humano en algunos casos, del propio proceso de calibración de la tableta digitalizadora, o de la resolución del scanner, si se utiliza este procedimiento. Por esta razón, dos capas de dibujo de trazos supuestamente idénticos, no necesariamente se funden en un mismo trazado al superponerlas, generándose así los pequeños polígonos antes citados.

La solución más primaria a este problema viene dada por la relación entre la escala de digitalización y la de representación, que, siendo rigurosa, eliminará por omisión perceptiva los polígonos defectuosos. Esta solución sin embargo no resulta satisfactoria en este procedimiento, puesto que los desajustes pasan a constituir casos erróneos en el análisis factorial.

Otra vía de solución pasa por la modificación de la distancia tolerada como error para el propio sistema de información. Esta solución se estimará pertinente en los casos en que la configuración del dibujo pueda asumir sin grandes costos las pérdidas en resolución que el proceso comporta.

La solución adoptada en el presente trabajo utiliza una de las herramientas prestadas por el propio sistema de información, como es la capacidad de generalización a partir de un atributo determinado. De este modo, se pueden generar coberturas secundarias a partir de la fusión de polígonos de una primera, asegurándose de este modo la conjunción de las capas en la superposición. Con este procedimiento se han generado, por ejemplo, la capa de información sobre suelos desde la de litología, pudiendo añadirse desde la estructura primaria, las diferenciaciones independientes a este trazado, mediante la transferencia del producto de la fusión, a través de un formato convertible, al programa de dibujo, para pasar a su definitivo procesamiento tras haber completado el trazado.

Por último hay que reseñar un último problema, que si bien no se consideró relevante en el caso elegido como ejemplo, plantea un obstáculo en la optimización de la integración de los sistemas. Se trata de la confusión que puede generar el hecho de que, partiendo de que cada uno de los polígonos finales constituyen una observación para el análisis factorial, la superficie que cada caso represente sea dispar, valorándose de igual modo una combinación que se presente en un espacio de ínfima entidad espacial, que otro que pueda constituir la práctica totalidad del territorio de análisis.

Como solución a este problema se ha considerado la posibilidad de transferir la cobertura final, producto de intersecciones vectoriales, a un formato teselar, lo que generaría una estructura de unidades homogéneas, constituyendo casos de idéntico peso para el análisis. Si bien el formato en celdas podría haber sido adoptado desde el principio del análisis, no se consideró oportuno, además de por las dificultades de representación y por

la carga que el volumen de información conlleva en este tipo de formato, por la inadecuación de su concepción (necesidad de establecer valores medios en las zonas de transición) a un proceso de acumulación de fronteras entre polígonos como el que se pretendía, ya que se hubiera producido una acumulación de valores medios.

PRIMEROS RESULTADOS

Como área de ensayo del procedimiento descrito se ha elegido

la cuenca del río de la Cueva, en el sector oriental de la provincia de Málaga. La elección de este área como piloto radica en la heterogeneidad de geosistemas que congrega.

El obstáculo fundamental encontrado a la hora de analizar los resultados es el alto nivel de imprecisión introducido como consecuencia de la no disponibilidad, en el momento de la elaboración, de la información precisa referente a las variables relativas a suelos y precipitaciones. Esto provocó la necesidad de hacer uso de fuentes concebidas a una escala mucho menos detallada que la presente en la mayoría de las variables, lo que produjo una repetición de valores iguales para un alto número de casos. En cualquier caso, si bien el valor desde un punto de vista geomorfológico de los resultados precisa de correcciones importantes para ser fiables, nos son útiles para observar el procedimiento y sus potencialidades como instrumento.

Se recogió y cartografió información sobre un total de veinte variables, que una vez digitizadas, fueron superpuestas seriadamente obteniéndose un total de 2730 polígonos/casos diferenciados. Hay que señalar como no existe una correspondencia entre el número de variables observadas y el de coberturas, puesto que se procuró incluir como atributos de una misma cobertura las variables que presentaran una misma zonificación o una compartimentación deducible de una primera mediante los métodos de fusión anteriormente comentados. Con ello se consiguió un aumento considerable de la operatividad de la intersección.

El resultado del análisis de componentes principales aplicado sobre valores estandarizados, se resolvió en un total de seis factores como explicativos del proceso erosivo, sin que el valor del análisis como sintetizador de la información fuera totalmente satisfactorio, puesto que la varianza acumulada explicada por estos seis factores representó solo

un 65.3%. Este hecho nos remite de nuevo al problema de la escasa precisión de la información utilizada en esta primera aproximación.

Fué interesante, sin embargo, la asociación en un factor, como se esperaba, de las variables más relacionadas con los signos de erosión, y así el factor 2, con un porcentaje de varianza del 14.1%, relacionó el número de cárcavas y surcos por km de una manera directa aunque no muy alta con la pendiente, indirecta y no muy alta con la profundidad del suelo y el porcentaje de materia orgánica, y con un carácter directo y muy alto con el porcentaje de cobertera de la vegetación. Conocer los pesos de las variables citadas en el factor constituirá una gran ayuda a la hora de establecer un orden en la ponderación de las distintas variables para valorar aspectos como la fragilidad o la susceptibilidad a la erosión. La variación en los pesos de mismos en circunstancias diversas, podrá analizarse, mediante métodos supervisados, concretando a priori los geosistemas susceptibles de ensayo diferenciado, o mediante un análisis cluster.

CONCLUSIONES

Salvando las distorsiones derivadas de la calidad de la información para algunas variables en el momento de la elaboración del trabajo, consideramos que, perfilando la resolución de algunas fuentes, el procedimiento metodológico posibilita unos resultados efectivos para el objetivo de caracterización y valoración del peso de las componentes del proceso erosivo, así como de sus variaciones circunstanciales.

Por otra parte, la experiencia ha aportado conocimientos acerca de la posibilidad del uso del análisis S.I.G. no sólo como una base de datos útil para la gestión de información, sino como un instrumento intermedio en el análisis, potenciador de una posibilidad más entre las vías de la investigación.

BIBLIOGRAFIA

CEBRIAN DE MIGUEL, J.A. (1987): "Una categorización de las principales funciones de establecimiento y explotación de SIGs" **Anales de Geografía de la Universidad Complutense**. Madrid.

CEBRIAN, J.A.- MARK, D. (1986): "Sistemas de Información Geográfica. Funciones y Estructuras de Datos". **Estudios Geográficos**. Nº 184