

LA INVESTIGACION OPERATIVA COMO METODO DE TRABAJO EN GEOGRAFIA.

Por, José Ramón Díaz Alvarez (*)

RESUMEN: La investigación geográfica, para responder a buena parte de la problemática con la que debe enfrentarse en el momento actual (elección de la respuesta más adecuada a una demanda, posibilidad de predecir un comportamiento, -aunque solo sea con un carácter probabilístico-, etc.), precisa de un método coherente con la naturaleza de sus trabajos; la "investigación operativa" se presenta como tal método de trabajo; y, entre las fórmulas y modalidades de la investigación de operaciones, se resaltan por su idoneidad para el trabajo geográfico, el análisis dinámico, el análisis no paramétrico y el tratamiento informático. A continuación se efectúa una presentación conceptual y metodológica de cada una de las modalidades, destacando sus funciones y utilidad, puestas de manifiesto por la inserción de algunos ejemplos. Por último, se muestra una panorámica de las posibilidades que se ofrecen desde los Centros de Cálculo de las Universidades españolas y se analizan los contenidos del paquete de programas B.M.D.P.

SUMMARY: Geographical investigation-to respond to a large extent to what the problem implies at the present time (election of the most suitable answer to satisfy a demand; possibility to predict a behaviour- even if it were of probabilistic nature, etc.) requires a coherent method indicating the nature of the work; the "operative investigation" presents itself as such a method, and, among the formulas and modalities of operations, it stands out because of its suitability for geographical work, dynamic and non-parametric analysis and for its informatic treatment.

Then, a conceptual and methodological presentation of each modality is carried out, pointing up its functions and its usefulness by means of some examples.

Finally, the possibilities offered by the Calculus Centres of the Spanish Universities are unfolded in a panoramic insight and the Contents of the B.M.D.P. programs are analysed.

INTRODUCCION.

La "investigación operativa" o "investigación de operaciones" es la ciencia de la toma de decisiones, por tanto se trata de una disciplina que busca la elección de los caminos o métodos más óptimos para la resolución de un problema. La investigación y la aplicación de fórmulas contrarias a los factores adversos de un proceso se estudian con el propósito de establecer juicios cuantitativos para obtener una decisión óptima, es decir, la considerada como mejor, entre las múltiples que se podrían adoptar. En Geografía, donde la problemática específica incluye la toma de decisiones sobre las formas, de utilización del suelo, de tratar una superficie para acondicionarla al asentamiento humano, para protegerla de su degradación erosiva, o para obtener el mayor provecho económico-especulativo o social, de elegir el camino o el trayecto idóneo para el trazado de una autopista, o la ubicación más lógica de una industria, etc; la necesidad de la investigación operativa es incuestionable.

Puesto que por razones de idoneidad de tratamiento, la Geografía precisa de un método dinámico, la toma de decisiones deberá auxiliarse del *análisis dinámico*. Como, además, un buen número de componentes del hecho geográfico son atributos, y por tanto, no cuantificables, la toma de decisión se alcanzará, a veces, a través del *análisis no paramétrico*. Por último, dada la creciente importancia que está alcanzando una información, cada vez más precisa, el volumen de datos geográficos, cuantificables y no cuantificables, se multiplica, lo que dificulta su tratamiento; para obviar esta dificultad, las tomas de decisión basadas en grandes volúmenes de información han de realizarse con ayuda del *tratamiento informático*.

(*) Doctor en Geografía.

Así, aunque la investigación operativa puede ofrecer otras muchas manifestaciones metodológicas para ser utilizadas en la investigación geográfica, tienen una especial importancia las tres formas enunciadas: *el análisis dinámico, el análisis no paramétrico y el tratamiento informativo o la utilización del ordenador*. Tales formas pueden considerarse como las más claras posibilidades metodológicas en la misma frontera de la actual investigación geográfica.

1. EL ANALISIS DINAMICO.

1.a/ Concepto del análisis dinámico y valoración metodológica de la programación dinámica.

A grandes rasgos, la Geografía se ocupa y preocupa por dos aspectos fundamentales del análisis espacial: las variaciones de un determinado aspecto o conjunto de aspectos geográficos a lo largo del espacio, como una situación presente o intemporal; o bien, los diferentes estados que adopta un fenómeno sobre una unidad espacial de estudio a través del tiempo. Tanto la consideración espacial como la histórica tienen cabida dentro de la metodología geográfica y se encuadran dentro del estudio de las *distribuciones espaciales* y de las *series cronológicas*. Cualquiera de las dos vertientes del estudio de la fenomenología geográfica pueden tratarse en etapas, que cuando se individualizan en el proceso de toma de decisión para adoptar la posición óptima con respecto a un aspecto o conjunto de aspectos, se convierten en un modelo geográfico, cuyo análisis, momento a momento, ha de efectuarse mediante *la programación dinámica*.

La programación dinámica, componente metodológico esencial del análisis dinámico, es una técnica por medio de la cual se optimiza (maximizando o minimizando) una función de varias variables en varias etapas, de modo que en cada una de ellas sea preciso optimizar únicamente una de dichas variables.

Este proceso continuado y secuencial de optimizaciones, se alcanza, en virtud del principio de optimización de BELLMAN que dice que “el camino secuencial” (“política”) seguido en un proceso es óptimo sólo y cuando cada uno de los pasos intermedios (“subpolíticas”) son, asimismo, óptimos”; principio que tiene su dual en el establecido por BHAVANANI y CHEN que dice: “una política óptima tiene la propiedad de que cualesquiera que sean los estados y decisiones siguientes, las decisiones precedentes deben constituir una política óptima con respecto al estado que se sigue de la última decisión”.

1.b/ Presentación de un ejemplo sobre un supuesto de Geografía Aplicada a la que le es aplicable la programación dinámica.

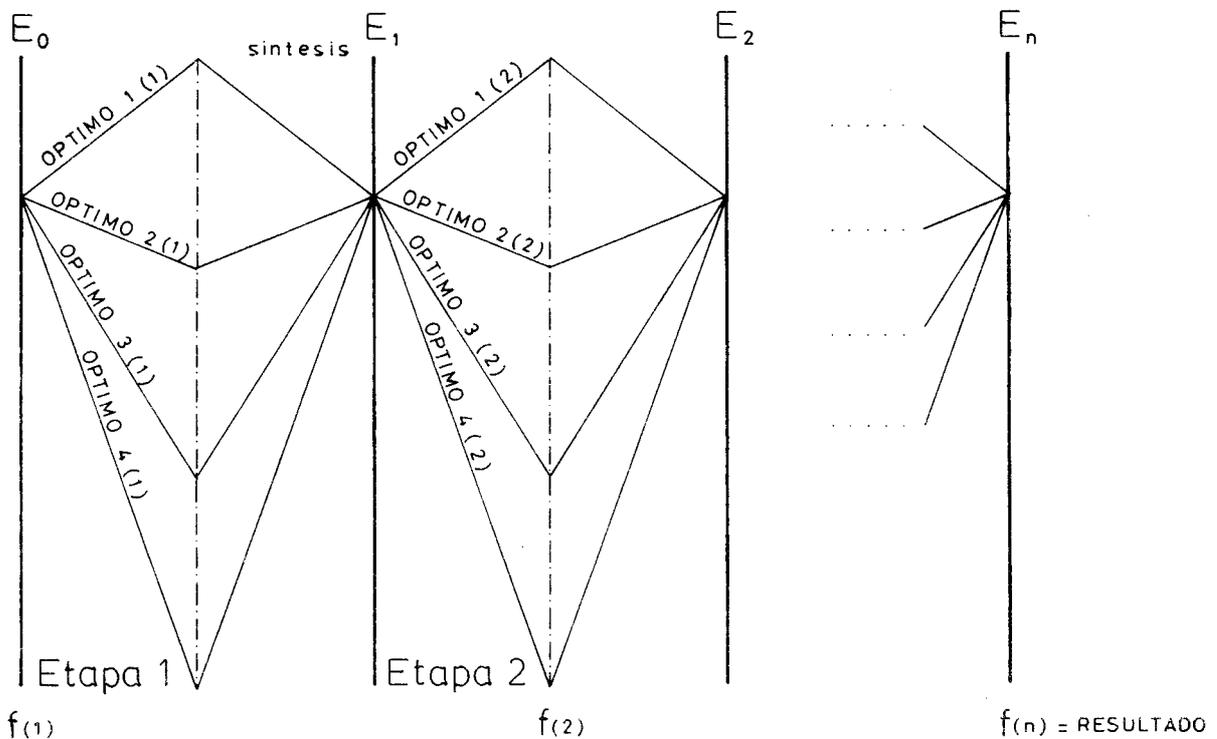
Por ejemplo, si tratamos de saber los salarios mínimos que se habrán de ofrecer en las áreas rurales para retener a un determinado número de trabajadores del campo en una determinada provincia o región para que sea posible un nivel de actividad en un horizonte previsto (1.985 — 1.990, pongamos por caso), precisaremos partir desde el horizonte buscado e investigar, año a año, cuales han sido las políticas óptimas en función de la variable salario (hipotética y modificable) de forma tal que cuando lleguemos al año de partida, el nº de trabajadores en el campo sea igual al nº real contabilizado. O bien, procediendo de forma directa, partiendo del nivel de ocupación en el momento de partida (año 1.980), calcular cuales serían las perspectivas de ocupación al año siguiente ofreciendo un salario determinado. Cualquiera de las dos opciones permite precisar el salario a ofrecer cada año, lo que representa la toma de decisión de ese paso en la secuencia.

En este sencillo ejemplo se presuponia que la decisión del agricultor, de permanecer en su puesto

de trabajo, estaba directa y únicamente ligada con el salario que le ofrecía; pero, en la realidad, tal decisión se ve influenciada por otra serie de variables como el número y calidad de los servicios sociales que se le ofrecen, la consideración social o el prestigio de su ocupación profesional, consideraciones familiares no cuantificables, etc. La existencia de varios condicionantes implica el que se hayan de tener en cuenta todos y cada uno de ellos en las diferentes etapas y, como consecuencia de ello, estarán presentes, implícitamente, en el resultado final. Esta nueva consideración presupone una fórmula más compleja de resolver el proceso, pues, ha de contemplar todas las variables condicionantes utilizadas independientemente en cada etapa o subpolítica, por lo que para cada estadio del proceso se obtendrán tantos valores óptimos como variables condicionantes estén implicadas en el mismo. Con el conjunto de los valores que configuran los posibles óptimos relativos de una subpolítica construiremos una función que es la *función óptima* de la etapa considerada. Así proseguiremos en las sucesivas etapas.

1.c/ Pasos para la resolución de un problema mediante la utilización de la metodología dinámica.

En la programación dinámica el método de operaciones comienza por un *análisis funcional del problema*, significado por una descomposición de la serie de estados, en los que se puede dividir el mismo, en varias etapas en las que solo sea necesario optimizar una función de una sola variable. Si la etapa comprende más de una variable a considerar en la toma de decisión se procedería individualmente con cada variable; cuando se hayan optimizado todas las variables de la etapa obtendremos una función combinación lineal de los óptimos de esta etapa que es, a su vez, la función de partida para la etapa siguiente. Gráficamente esto lo podríamos representar así:



cuya explicación es la siguiente: partimos de un estado inicial, supongamos que el del ejemplo enunciado en el caso anterior, (pero también podría ser el del estado atmosférico en un lugar determinado y en un momento preciso, al objeto de predecir el tiempo futuro; o bien, el volumen y la superficie de un suelo vegetal para analizar el impacto y los peligros que suponen la presencia de ciertas variables erosivas, etc). En el caso que presentábamos, el estado inicial vendría constituido por el nº de trabajadores empleados en el campo, en el año actual; las variables que influyen en su permanencia en el mismo trabajo podrían ser, salario, nº mínimo de habitantes del núcleo de residencia, nº de hijos mayores de 14 años (que comiencen estudios medios), edad del cabeza de familia inferior a los 45 años, etc; cada variable analizada alcanzaría un óptimo, es decir, retendría al nº deseado de trabajadores a un determinado nivel; la cuantificación del valor de la variable, que se efectuaría probabilísticamente, habría que precisarla mediante encuesta y análisis muestral. Así, por cada variable, configuraríamos una ecuación que representaría la cuantificación de la importancia de la variable en su poder de retención del trabajador del campo, y una combinación lineal de la presencia real de cada variable serviría para definir el atractivo o el poder de retención global del área analizada, representada por un nº determinado de trabajadores que permanecerían en su quehacer laboral. Este nº de trabajadores sería, a su vez, el estado de llegada de la primera etapa (E1) y el estado inicial para la segunda etapa. Así se procedería en los momentos siguientes hasta llegar al (En) estado final.

1.d/ Aplicación del análisis dinámico.

Bajo una primera aproximación, el análisis dinámico se muestra como un eficaz auxiliar en dos aspectos de la investigación geográfica:

— *En los procesos de toma de decisión, en general, eligiendo las alternativas más eficaces (es decir, las variables componentes del proceso a los niveles óptimos), o los costos mínimos si se trata de un proceso de componentes económicos (como en la elección de los diferentes y posibles tramos de una autopista entre dos ciudades precisadas de antemano).*

Hay que tener en cuenta que en los procesos de toma de decisión hay que presumir unos comportamientos probables ante la presencia de cada valor de la variable; pero, existe la posibilidad de que la decisión se haya de tomar ante un riesgo desconocido, es decir, desconociendo los niveles óptimos de cada variable y, entonces, debe seguirse el criterio de LAPLACE, de considerar equiprobables las influencias de cada variable.

— *En los procesos de predicción, en los que se intenta conocer una situación futura en función de un comportamiento, en las relaciones entre las variables, parecido o similar, al observado en el presente, presumiendo, asimismo, un comportamiento humano o natural previsible de una forma probabilística. Las posibilidades del proceso de predicción sirven para la planificación geográfica, detectando el resultado en una actuación continuada de una serie de variables, pudiendo analizar paso a paso las tendencias y los valores que van tomando las variables dependientes cuyos estados finales se investigan; pero, también sirven para descubrir cuales son las variables independientes que han de modificarse y en la cuantía en la que debe hacerse para llegar a un estado final, considerado como apetecible. Este segundo aspecto es de una importancia fundamental en Geografía Aplicada ya que ofrece el camino de la intervención geográfica en el acondicionamiento científico del espacio, permitiendo al hombre la posibilidad de abandonar su posición de observador pasivo ante las componentes activas de las transformaciones geográficas, para convertirse en una componente racional de tales transformaciones.*

III/ EL ANALISIS NO PARAMETRICO.

II.a/ La naturaleza del análisis no paramétrico.

Generalmente, los modelos de análisis utilizados en Geografía, que se auxilian de la Estadística o de las técnicas matemáticas, se refieren casi exclusivamente a la cuantificación de variables o al tratamiento de magnitudes susceptibles de medida por comparación con una unidad tipo (superficies, n° de habitantes, precipitaciones, distancias), o bien por relación con otros aspectos directamente medibles (densidades, potenciales, índices en general). Hemos visto que uno de los aspectos más interesantes de estas nuevas formas metodológicas para la investigación geográfica es que posibilitan la predicción de comportamientos futuros a través de la inferencia estadística. Pero, la inferencia estadística, se puede efectuar gracias a la información facilitada por un gran número de acontecimientos, o por la observación repetida de un mismo acontecimiento en circunstancias similares.

Los acontecimientos geográficos vienen configurados por la naturaleza de la fenomenología geográfica que los determina, y como los hechos geográficos se componen de aspectos cuantificables (que podemos representar por variables que adoptan diferentes valores) y de aspectos cualitativos no cuantificables (que denominamos atributos y que adoptan diferentes modalidades), será necesario buscar técnicas diferentes para el tratamiento de los distintos aspectos. El tratamiento de las variables se efectúa mediante técnicas estadísticas de inferencia que denominamos paramétricas, mientras que el de los atributos lo efectuamos a través de *las técnicas no paramétricas*. El análisis no paramétrico será la ordenación de las técnicas no paramétricas en favor de su sistematización, de cara a la mejor utilización de las mismas en la investigación de los atributos o componentes no cuantificables de los fenómenos geográficos. Es erróneo el pensar que no existen formas de medida de los atributos; pero, sus escalas de medida solo tienen definidas las relaciones de equivalencia (como en la escala nominal) y la relación de mayor a menor (en la escala ordinal). Estas escalas de medida son aplicables, únicamente, a clasificaciones en rangos o a categorías como positivo o negativo, más o meno, mejor o peor.

II.b/ Tipos de escala de medida del análisis no paramétrico.

SIEGEL señala cuatro niveles de medida con propiedades diferenciadas, que se pueden agrupar en escalas de nivel inferior (escalas nominal y ordinal) y en escalas de nivel superior (de intervalos y de relaciones). Ya habíamos señalado cuales eran las características que habían de reunir las escalas de medida en el análisis no paramétrico, ahora analizaremos parcialmente cada una de ellas.

La *escala nominal* representa el nivel más bajo de medición, y en ella los números o símbolos se utilizan como identificadores de los objetos; así, cuando hablamos de un modelado glaciar o un modelado cárstico solo podremos establecer con ellas la relación de equivalencia; la asignación de números o símbolos debe ser tal que, a cada clase, se le de uno que le identifique y diferencie del resto de las clases. Si consideramos como clase "1" el modelado cárstico y como clase "2" el modelado glaciar, cualquier otro modelado podría estar incluido en la clase "1", en la "2", o en ninguna de ellas. La relación de *equivalencia* que se da en esta escala goza de las propiedades *reflexiva, simétrica y transitiva!*

La escala nominal permite la clasificación y, los símbolos que identifican a las diferentes clases, pueden intercambiarse entre sí y continuarán cumpliéndose sus propiedades; por ello, las estadísticas descriptivas admisibles son las que no se alteran con el intercambio: *el modo, la frecuencia, el conteo*, etc, y se puede aplicar la prueba estadística no paramétrica del χ^2 (*Ji cuadrado*), para comprobar la hipótesis de trabajo. La medida de asociación más común es *el coeficiente de contingencia*.

La *escala ordinal* es una escala jerárquica o de rango, y por tanto la propiedad principal es la que establece las relaciones con otros objetos o hechos en términos de "mayor que", "mejor que", "preferible a", etc. Incorpora, pues, a la propiedad de equivalencia la de (mayor que) que es una relación que posee

el carácter de *irreflexiva, antisimétrica y transitiva*. Se suelen medir según la escala ordinal las ciudades, centros de comercio, líneas de comunicación, etc.

Las estadísticas más utilizadas en la escala ordinal son, la *mediana* para describir la tendencia central, *los coeficientes de correlación de Spearman o Kendall* para analizar la relación existente entre dos atributos, *la prueba de McNemar* para la significación de los cambios en una modalidad, o *la prueba de Wilcoxon* cuando se pueden ordenar significativamente las diferencias entre los pasos de una escala ordinal; y, por último, *los percentiles*.

Se debe recordar siempre que las propiedades de una escala ordinal no son isomórficas al sistema numérico conocido como aritmético y, por tanto, las probabilidades derivadas de la aplicación de pruebas paramétricas a las mismas, (como son las medias, las desviaciones estandares, etc), son, cuando menos, dudosas y poco válidas.

II.c/ Ideas generales sobre las más usuales formas estadísticas no paramétricas.

Cualquier tratado de estadística elemental suele presentar conjuntamente tanto las formas paramétricas como las no paramétricas; como lo que pretendemos mostrar es la utilidad geográfica de las formas no paramétricas para que sean utilizadas convenientemente en el análisis cualitativo, vamos a dar unas ideas generales sobre las características y funcionalidad de estas formas.

El modo o la moda es la modalidad que tiene una mayor frecuencia en un conjunto de informaciones de carácter cualitativo. Por ejemplo, si estamos estudiando la geomorfología de las unidades montañosas de la provincia de Santander, podríamos haber detectado 50 complejos geomorfológicos; si la frecuencia del modelado cárstico es la más alta, diremos que este constituye la moda de los modelados santanderinos.

La frecuencia es el número de veces que se repite o se presenta un fenómeno. En la escala nominal la frecuencia de cada modalidad se halla efectuando el puntaje de todos los hechos, abriendo una clase por cada diferente atributo y denominándola por el nombre de dicho atributo. Al final del puntaje hemos hallado las frecuencias de cada atributo. Así, si estamos analizando los usos del suelo en Cataluña, podríamos diferenciar los espacios aislándolos de los que no tienen igual utilización; de esta forma podemos separar entre suelos agrarios (a), ganaderos (b), forestales (c), urbanos (d), superficies viales (e), superficies de ríos y aguas de pantanos o lagos (f) y superficies improductivas (g). Cada espacio diferenciado se pone en relación con una clase (a,b,c,d,e,f o g) y se van contando las áreas existentes dentro de cada clase. Así obtenemos las frecuencias de clase. No se entra aquí en la importancia y extensión de las áreas, sino en la existencia de las mismas.

La prueba χ^2 (Ji cuadrado) sirve para probar una hipótesis, en la que se presuponen unas determinadas frecuencias para unos hechos o una jerarquización precisa entre los mismos, en relación con lo observado. La fiabilidad del ajuste entre lo presumido y lo observado nos indica la bondad de la hipótesis. Como vale tanto para frecuencias como para rangos, puede considerarse incluida en el análisis no paramétrico.

El coeficiente de contingencia, que es una medida del grado de asociación o relación entre dos conjuntos de atributos de los que solo tenemos información clasificatoria o a lo sumo jerárquica. Su método procede mediante tablas de contingencia.

La mediana, es el valor medio de un conjunto de datos ordenados. Por ejemplo si ordenamos las poblaciones de España por rangos desde los núcleos más pequeños y con menos servicios hasta los mayores, el orden central (equidistante de los extremos), o la media aritmética de los valores centrales, es la mediana.

Los coeficientes de correlación de Spearman y Kendall son una medida de asociación que requiere que los datos sean medidos, al menos, en una escala ordinal, de forma tal que los individuos estudiados puedan colocarse en dos series ordenadas. El de Spearman es el más utilizado, pero el de Kendall es más potente, pues puede generalizarse a correlaciones parciales.

La prueba de McNemar es apropiada para hacer consideraciones en las que la pauta de diferenciación entre los atributos es el tiempo y, la medida de tal diferenciación, la consideración propia de cada persona; por tanto, se trata de una prueba utilizable en Geografía de la Percepción.

La prueba de Wilcoxon sirve para comparar dos series de datos referidos al mismo atributo y que se hallen ordenados de forma tal que clasifica las diferencias por orden de tamaño absoluto. Por ejemplo, si queremos analizar el parecido existente entre la utilización del suelo de Galicia y Cataluña podemos comparar cada atributo en la extensión relativa que ocupa en el total de la superficie de la región; el puntaje de las diferencias sirve para establecer unos rangos de 1 a n (desde el más pequeño al más grande) y los rangos servirán para jerarquizar y ordenar las diferencias. Así obtendremos unos rangos muy altos para las diferencias de dedicación industrial o urbana y ello nos permitirá afirmar cuales son las diferencias observables de mayor peso o más significativas.

Los percentiles, y más generalmente los cuantiles, dividen una distribución en partes que contienen un cierto porcentaje del efectivo total. La técnica de los cuantiles es útil para obtener, las épocas de máxima crecida de un río o de máxima precipitación en un área, o la distribución de los municipios según su rango por el nº y la calidad de los servicios que ofrecen, etc.

III. EL TRATAMIENTO INFORMATIVO.

III.a/ La organización de la información para el tratamiento informático.

En el estado actual de los conocimientos, la cantidad de información a manejar es de tal magnitud que, resulta prácticamente imposible alcanzar un adecuado uso de la misma. En Geografía, la mejor comprensión de la compleja y cambiante realidad espacial y social de los fenómenos que son objeto de su estudio, precisan de la consideración de un enorme volumen de información, pero el tratamiento de la misma no será fructífero sin una adecuada organización que solo es posible cuando utilizamos una metodología válida para la sistematización y el acceso preciso al dato o datos que configuran el conjunto de características definidas de un fenómeno.

Como la información puede referirse a aspectos cualitativos y cuantitativos las premisas previas a su organización serán diferentes según se trate de una u otra: cuando nos referimos a aspectos cualitativos hemos de establecer como hipótesis de trabajo que el número de estados o modalidad del atributo es finito, premisa que no será necesaria cuando tratemos de valorar una variable. Seleccionados los componentes de la información de una determinada investigación (variables y atributos), posteriormente hay que medir la intensidad de presencia de tales componentes en el fenómeno que se analiza aplicándoles las cuatro escalas de medida que señala SIEGEL.

El siguiente paso sería el de relacionar la información con el espacio que representa, es decir, la constitución de un Banco de Datos espaciales, que es realmente una clasificación homogeneizada y ordenada y que recibe el nombre de *matriz de información geográfica*, que es un depósito de datos en el que se dispone de una serie de celdas alineadas en filas y columnas en una trama que les pone en correspondencia entre sí. La asociación entre información y espacio al que se refiere la misma puede realizarse colocando en cada fila los resultados de las diferentes observaciones efectuadas sobre una determinada área y, en cada columna las variaciones espaciales de una determinada observación, o viceversa. Me-

diante este sistema se puede localizar, con precisión, la información a que queramos referirnos. Por ejemplo, supongámonos que queremos organizar la información que poseemos sobre los usos del suelo en cada una de la cincuenta provincias de España. La diferenciación que hacemos nos sirve para distinguir siete tipos, como señalábamos en el apartado anterior (suelos agrarios = X1, suelos ganaderos = X2, suelos forestales = X3, suelos urbanos = X4, superficies de viales = X5, superficies de ríos y aguas de pantanos o lagos = X6, y superficies de roca desnuda improductivas = X7). Cada provincia ocupará una superficie que será igual a la suma de las superficies ocupadas por cada tipo de suelos; es decir:

$$\text{(Alava)} P_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{1,5} + X_{1,6} + X_{1,7} = 3.047 \text{ Km}^2$$

$$\text{(Albacete)} P_2 = X_{2,1} + X_{2,2} + X_{2,3} + X_{2,4} + X_{2,5} + X_{2,6} + X_{2,7} = 14,858 \text{ Km}^2$$

(Provincia que ocupa la posición i-exima).

$$P_i = X_{i,1} + X_{i,2} + X_{i,3} + X_{i,4} + X_{i,5} + X_{i,6} + X_{i,7} = \text{Superficie de la provincia que ocupa la i-esima posición.}$$

$$\text{(Zaragoza)} P_{50} = X_{50,1} + X_{50,2} + X_{50,3} + X_{50,4} + X_{50,5} + X_{50,6} + X_{50,7} = 17.194 \text{ Km}^2.$$

en las que el primer subíndice nos define la provincia a la que se refiere la variable, el segundo subíndice (1,2,3,4,5,6,7) nos indica el tipo de suelo al que nos referimos, y X, es la variable que contiene el valor de la superficie que se dedica en esa provincia a ese determinado uso del suelo.

Si queremos analizar las variaciones de cada uso del suelo en el conjunto del solar español, solo tenemos que investigar la columna correspondiente (1,2,..., 7) al uso del suelo que estudiamos. Si lo que pretendemos es analizar las variaciones que experimentan las utilidades de los suelos en el solar de una determinada provincia, solo habrá que investigar la fila correspondiente a la provincia que pretendamos estudiar.

Si pretendemos hacer una valoración económica de cada uso del suelo (con carácter medio) y conocemos el montante de la renta total de cada provincia, solo tendremos que darle a cada variable un coeficiente igual al de la superficie que ocupa (en Km²) y colocar en el término independiente el valor de la renta provincial. Así podríamos ir, sucesivamente, obteniendo información complementaria a partir de los datos iniciales, lo que comporta una gran ventaja.

Organizada así la información estamos en condiciones de *tratarla* (investigar relaciones y dependencias entre variables, agruparlas, asociarlas y clasificarlas en razón de unos criterios previstos, etc), mediante *máquinas electrónicas*, es decir entramos en un campo de enormes posibilidades por la cantidad de información que puede tratar y por la velocidad del tratamiento. El problema del tratamiento se resuelve mediante la programación y esta depende del enfoque y de las funciones que queramos obtener a partir de la información primaria.

No obstante, el proceso de organización de la información no se detiene en lo estudiado, sino que es preciso tener una idea de la morfología de los datos para que estos puedan entrar en las máquinas. Existe una ciencia, *la informática*, que pretende instruir sobre esta necesidad metodológica; y más particularmente, se han realizado una serie de estudios que constituyen toda una teoría de la *información, codificación y lenguajes* que deben ser conocidos para emprender la tarea. Ello puede ser demasiado com-

plejo para el interesado en la investigación geográfica y le haría perder demasiado tiempo en su específico campo de estudio y, en este caso, lo importante es auxiliarse de un técnico en programación, pero siempre bajo el compromiso de organizar convenientemente la información. Hemos de aprender, pues, las instrucciones de entrada, las de salida y las formas de codificar los datos para que sean inteligibles por el ordenador.

III.b/ Posibilidades que encierra la utilización del ordenador.

No solo para la investigación geográfica, sino para cualquier tratamiento de información, la utilización del ordenador se hace tan imprescindible que, prácticamente, todas las universidades del mundo tienen un departamento de informática. En el caso español, los centros de cálculo de las Universidades, poseen unos programas standards cuyo paquete responde a las siglas B.M.D.P. (Biomedical Computer Programa) y que es una colección de 26 programas divididos en series diferentes según las características generales de los trabajos que ejecutan. Las series son:

1/ *La serie D*, corresponde a programas especializados en la descripción de datos y comprende como programas más útiles en Geografía los:

BMDP1D, para descripción simple de datos, incluyendo como estadísticos la media, la desviación típica, el coeficiente de variación, el valor más grande y más pequeño, el recorrido, el nº total de casos aceptables.

BMDP2D, para frecuencias variables, que incluye el conteo de los casos, el máximo y mínimo, el recorrido, el semiintervalo, la media, la mediana, la moda, la desviación típica; además, traza un histograma y construye una tabla con el número de casos de cada valor, otra con el % de cada valor y otra con la función de distribución.

BMDP5D, para el trazado de gráficas univariacionales que dibuja histogramas acumulativos con frecuencias y porcentajes en cada intervalo, la curva normal y la de la distribución acumulada; y señala las frecuencias totales, la media y la desviación típica.

BMDP6D, para el trazado de gráficas bivariacionales, que calcula las ecuaciones de las regresiones entre dos variables que varían de forma interdependiente y la ecuación de la regresión lineal simple con la media cuadrática residual, la correlación y el recuento de casos.

2/ *La serie F* que tiene un solo programa para tablas de contingencia, calculando como salidas los porcentajes sobre filas y columnas de la tabla de frecuencias originales y, opcionalmente, nos da los coeficientes de contingencias, el máximo valor de los mismos, el χ^2 y la máxima verosimilitud del mismo, las probabilidades de Fisher, momento de correlación, los coeficientes de correlación, la V de Cramer, la Q de Yule, etc.

3/ *La serie M*, corresponde a programas de análisis multivariante incluyendo los programas:

BMDP1M, para análisis cluster entre variables, que a partir de la matriz de datos de entrada calcula una matriz de distancias entre las columnas de las primeras; en realidad, sirve para establecer la separación existente entre un conjunto de variables entre sí, definidas sobre un mismo lugar.

BMDP3M, para análisis clúster por bloques, que sirve para agrupar la matriz de información primaria (tanto de datos como de casos) en submatrices de tipo jerárquico, sin solapamientos, por lo que podría servir para diferenciar subregiones según diferentes tipos de especializaciones: las áreas industriales, agrarias, comerciales, etc, de una región autonómica, por ejemplo.

BMDP4M, para análisis factorial, que sirve para agrupar variables, eliminando redundancias obteniendo los factores principales de entre las variables componentes de una matriz de información. Calcula además las correlaciones múltiples de las variables, la adecuación de la distribución de datos y el peso factorial local de cada comunidad.

4/ *La serie R*, que corresponde a programas de regresión, incluyendo los programas:

BMDPIR, para regresión lineal múltiple, obtiene la dependencia lineal de una variable con respecto a una serie de variables. Por ejemplo, si pensamos que el porcentaje de terreno cultivado de un municipio depende de la densidad de población, el rendimiento medio por Ha. en ptas., la renta media per cápita, el nivel de ocupación de su población en edad de trabajar, etc. Mediante este programa podemos hallar la recta de regresión (ecuación que liga a estas variables con la variable dependiente).

BMDP3R para regresión no lineal que obtiene un ajuste por mínimos cuadrados pesados mediante sucesivas iteraciones, siguiendo el método de Gauss-Newton. Sirve para denunciar dependencias no lineales entre dos variables o entre una variable y un grupo de otras variables que tienen relación de interdependencia con la buscada.

BMDP6R, para correlación parcial y regresión multivariable, que sirve para hallar las correlaciones parciales de un conjunto de variables eliminando los efectos lineales de un segundo conjunto de variables. También muestra las correlaciones múltiples y sus niveles de significación y efectúa dibujos bivariantes para las variables originales y los residuales.

5/ *La serie S*, de programas especiales, entre los que destacan *el BMDP3S* para estadística no paramétrica, incluyendo:

- el test de los signos.
- el test de Wilcoxon.
- el test de rangos de Kendall.
- el test de Spearman.
- el análisis de la varianza de Friedman.
- el coeficiente de concordancia de Kendall.
- el test U de Mann-Whitney.
- el análisis de Kruskal-Wallis.

Pero, esto, con ser mucho, solo representa una mínima parte de las posibilidades que encierra el ordenador. Mediante programación adecuada, es decir, suministrando el correspondiente "Software", podremos realizar cualquier tipo de operaciones aritméticas y lógicas que podamos imaginar.

Otra gran posibilidad, solo iniciada, es la de servir de archivo de una ingente cantidad de información, es decir, la de conformar *bancos de datos*. Un banco de datos puede almacenar, en un volumen físico muy reducido, miles de millones de datos, cuyo acceso a los mismos puede ser abierto y estar a disposición del usuario, que podrá disponer de ellos para continuar sus investigaciones y para obtener información elaborada a partir de la primaria. Para el geógrafo español, el Instituto Geográfico y Catastral (actual Instituto Geográfico Nacional) está elaborando la base informática para constituir su propio Banco de Datos.

Finalmente, la utilización del ordenador, está posibilitando la incorporación directa de la información obtenida teleméricamente y mediante fotografía y opticometría, desde satélites artificiales, lo que unido a su rápido tratamiento, permitirá utilizar, adecuadamente, los resultados de las investigaciones efectuadas, a la Geografía Aplicada, antes de que las condiciones de deterioro de una situación sean extremadamente dramáticas.

I/ BIBLIOGRAFIA GENERAL SOBRE LA TEORIA DE LA INVESTIGACION DE OPERACIONES.

- 1) CHACON, E.: "La investigación Operativa". Ibérico Europea de Ediciones. Madrid. 1.970.
- 2) CHURCHMANN, C.W., ACKOFF, R.L. y ARNOFF, E.L.: "Introduction to operations research" John Wiley and Sons; Inc. New York, 1.958.
- 3) DESBAZEILLE, G.: "Exercices et problèmes de recherche operationnelle". París, 1.964.
- 4) HANSSMANN, F.: "Operations research in production and inventory control" John Wiley and Sons, Inc. New York. 1.962.
- 5) HILLIER, F.S. & LIEBERMAN, G.J.: "Introduction to operations Research". New York, 1.967.
- 6) KAUFMANN, A.: "Métodos y modelos de la Investigación de Operaciones". Compañía Editorial Continental S.A. (C.E.C.S.A.), México, 1.979.
- 7) MILLER, D.W. and STARR, M.K.: "Executive decisions and operations research". Prentice-Hall, Inc. Englewood Cliffs. N.J. 1.960.
- 8) SAATY, T.L.: "Mathematical methods of operations research". McGraw-Hill Book Company, Inc. New York. 1.959.

II.a/ BIBLIOGRAFIA ESPECIFICA SOBRE LA TEORIA DE LA PROGRAMACION DINAMICA:

- 1) BELLMAN, R.E. y DREYFUS, S.E.: "Applied dynamic programming" Princenton University Press Princenton, 1.962.
- 2) HOWARD, R.: "Dynamic programming and Markov Processes". Technology Press. M.I.T.; Wiley, 1.960.
- 3) KAUFMANN, A. et CRUON, R.: "La programmation Dynamique". Dunod. Paris, 1.965.
- 4) Además de las obras citadas en el apartado anterior en los lugares 1,6 y 9.

II.b/ BIBLIOGRAFIA SOBRE LA APLICACION DEL ANALISIS DINAMICO A LA INVESTIGACION GEOGRAFICA:

- 1) BATTY, M.: "An Experimental Model of Urban Dynamics". Journal of Town and Country Planning, nº 43, 1.972. pags. 166-186.

2) BRUMMELL, A.C.: "A model of intraurban mobility". Economic Geography. Vol. 55, nº 4. October, 1.979.

3) CLARK, W.A.V. & SMITH, T.R.: "Modeling information use in a spatial context". A.A.A.G. Vol. 69. nº 4. December 1.979.

4) COLLINS, L.: "An Introduction to Markov Chain Analysis" Norwich, Eng. GEO Abstracts. 1.975.

5) CORDEY — HAYES, M.: "Dynamic Frameworks for spatial Models". Socio-Economic Planning Sciences. nº 6. 1.972. pags. 365-385.

6) DIAZ ALVAREZ, J.R.: "El análisis dinámico como método geográfico. Su aplicación a dos casos andaluces". Paralelo 37°. Revista de estudios geográficos. nº 3, 1.979. Colegio Universitario de Almería.

7) KLEITER, G.D.: "Dynamic decision Behaviour: Comments on Rapport's Paper" Wendt and Vlek, eds. en "Utility, Probability and Human Decion Making". Dordrecht. Reidel, 1.975.

8) MACKINNON, R.: "Dynamic Programming and Geographical Systems". Economic Geography, nº 46 (Supl. de 1.970). pags. 350-366.

9) SCOTT, A.: "The Elementary Theoretical Dynamics of a Competitive Locational System". Discussion Paper, nº 6. Depart. Geog. University of Toronto. 1.970.

10) VANCE, J.E.: "Labor-shed, employment field, and dynamic analysis in urban geography". Economic Geography. nº 36 189-220. 1.960.

III.a/ BIBLIOGRAFIA ESPECIFICA DE CARACTER TEORICO SOBRE EL ANALISIS NO PARAMETRICO.

1) BLUM, J.R. y FATTU, N.A.: "Non parametric methods". Rev. Educ. Res. nº 24. pags. 476-487, 1.954.

2) EDWARDS, A.L.: "Statistical methods for the behavioral sciences". Rinehart, New. York. 1.954.

3) KRUSKAL, W.H.: "A nonparametric test for the several sample problem". Ann. Math. Stat, nº 23. pags. 525-540. 1.952.

4) SAVAGE, L.J.: "Bibliography of non-parametric statistics and related topics". J. Amer Statist. Ass., nº 48. pags. 844-906. 1.953.

5) SCHEFFE, H.: "Statistical inference in the non-parametric case". Ann. Math. Stat, nº 14. pags 305-332. 1.943.

6) SIEGEL, S.: "Estadística no paramétrica aplicada a las ciencias de la conducta". Editorial Trillas. México, 1.978.

III. b/ BIBLIOGRAFIA SOBRE LA UTILIZACION DEL ANALISIS NO PARAMETRICO EN LA GEOGRAFIA.

- 1) GREGORY, S.: *Statistical methods and the geographer*. Londres. 1.963.
- 2) GROUPE CHADULE.: *Initiation aux méthodes statistiques en Géographie*. Masson et Cie. Collection Géographie. Paris. 1.974.
- 3) HAGGETT, P.: *Análisis locacional en la Geografía humana*. Editorial Gustavo Gili. Barcelona, 1.976.
- 4) HAGGETT, P.; CHORLEY, R.J.; STODDART, D.R.: *Scale standards in geographical research: a new measure of area magnitude*. Nature, nº 205. pags. 844-847, 1.965.
- 5) HAGOOD, M.J.: *Statistical methods for delineation of region applied to data on agriculture and population*. Social Forces, nº 21. pag. 288-297. 1.943.
- 6) McCARTY, H.H.: *Use of certain statistical procedures in geographical analysis*. AAAG. nº 46. 1.956.
- 7) SPATE, O.H.K.: *Quantity and quality in geography*. AAAG. nº 50. 1.960.

IV/ BIBLIOGRAFIA ESPECIFICA SOBRE LA UTILIZACION DEL ORDENADOR EN LA INVESTIGACION GEOGRAFICA.

- 1) BRASSEL, K.: *Manipulation Processes in Computer Cartography*. JEEE. Computer Society. Long Beach, CA. 1.978.
- 2) BRIBIESCA, E. y AVILES, R.: *Codificación en cadenas y técnicas de reducción de información para mapas y dibujos lineales*. Informe CCAL-74-7; IBM Latin American Scientific Center (México City) 1.974.
- 3) DAVIS, J. and McCULLAGH, M.J.: *Display and Analysis of spatial Data*. Wiley-Interscience. London. 1.975.
- 4) DOUGENIK, J.A. y SHEEHAN, D.E.: *Symap users reference manual*. Laboratory for Computer Graphics and Spatial Analysis. Harvard University. Cambridge, Mass. 1.975.

- 5) INSTITUTO DE GEOGRAFIA DE LA UNAM.: *Computer Mapping for Resource Analysis*. A Cogeo-data Conference. Kansas Geological Survey. University of Kansas. México. 1.978.
- 6) KAO, R.C.: *The use of computers in the processing and analysis of geographic information*. Geographical Review, nº 53, pags. 530-547. 1.963.
- 7) MARBLE, D.F.; BRASSEL, K.; CALKINS, H.W. & PEUQUET, D.J.: *Computer Software for Spatial Data Handling*. International Geographical Union Commission on Geographical Data Sensing and Processing. Ottawa, Ontario, Canada, 1.978.
- 8) McCULLAGH, M.J.: *A Data Mapping and Analysis System for Mini-Computers*. University of Kansas, 1.977.
- 9) McCULLAGH, M.J.: *The development of a mini-computer graphics system*. Computer applications. Vol. 5. nº 1 y 2. 1.978.
- 10) NORDBECK, S.: *Location of areal data for computer processing*. Lung Studies in Geography. Series C. General and Mathematical, nº 2. 1.962
- 11) PEUQUET, D.J.: *Raster data handling in geographic information systems*. Harvard papers on geographic information systems. Vol. 2. Laboratory for computer graphics and spatial analysis. Cambridge. Mass. 1.977.
- 12) TIMOFEEFF, N.P.: *The use and mis-use of computer graphics in environmental studies*. Conference on application of computer cartography. American Geographical Society. New York, 1.975.
- 13) TOBLER, V.R.: *Automation and cartography*. Geographical Review. nº 44. pags 536-544. 1.959.
- 14) WRIGLEY, N.: *Probability Surface Mapping*. An introduction with examples and Fortran Programs: concepts and techniques in modern Geography. nº 16. Geo Abstracts Ltd. Norwich. 1.977.