

XXXI
REUNIÓN
ESTUDIOS
REGIONALES

El Estado Autonómico a debate: eficiencia, eficacia y solidaridad

Alcalá de Henares, 17-18 de noviembre de 2.005

UN ALGORITMO EVOLUTIVO PARA LA DELIMITACIÓN DE MERCADOS
LOCALES DE TRABAJO

JOSÉ MANUEL CASADO DÍAZ
INSTITUTO DE ECONOMÍA INTERNACIONAL Y DEPARTAMENTO DE ANÁLISIS ECONÓMICO APLICADO
UNIVERSIDAD DE ALICANTE
jmcasado@ua.es

FRANCISCO FLÓREZ REVUELTA
DEPARTAMENTO DE TECNOLOGÍA INFORMÁTICA Y COMPUTACIÓN
UNIVERSIDAD DE ALICANTE
florez@dtic.ua.es

LUCAS MARTÍNEZ BERNABEU
INSTITUTO DE ECONOMÍA INTERNACIONAL
UNIVERSIDAD DE ALICANTE
Lucas.Martinez@ua.es

RESUMEN (MÁXIMO: 150 PALABRAS)

En esta comunicación se propone un enfoque evolutivo para tratar el problema de delimitación de áreas funcionales relevantes para el análisis de los mercados de trabajo en ámbitos locales y urbanos. El objetivo es la delimitación del territorio en el que la oferta y la demanda de trabajo se encuentran en los ámbitos urbanos mediante la identificación de áreas (a) cuyas fronteras son cruzadas sólo excepcionalmente en el curso de los desplazamientos residencia-trabajo y (b) dentro de las cuales se produce un alto nivel de interacción. Se utilizará para ello un nuevo método basado en computación evolutiva que agrupa municipios en regiones funcionales mediante la maximización de una función que pondera los flujos intramercados sujeta a una serie de restricciones entre las que destaca la superación de un nivel dado de autonomía y de tamaño mínimo. El algoritmo propuesto se basa, por tanto, en una combinación de selección, recombinación y mutación y permite evolucionar hacia una solución mejorada del problema de regionalización partiendo de (a) un mapa de regiones obtenido a partir de alguno de los métodos existentes o (b) una solución inicial de carácter aleatorio que es reevaluada a lo largo del proceso.

PALABRAS CLAVE (DE TRES A CINCO): algoritmo evolutivo, movilidad residencia-trabajo (*commuting*), mercados locales de trabajo, áreas funcionales

CÓDIGOS JEL: R23, J61, C45, C61, C69

UN ALGORITMO EVOLUTIVO PARA LA DELIMITACIÓN DE MERCADOS LOCALES DE TRABAJO¹

1. Introducción

En las últimas décadas ha ido ganando terreno la idea de que el análisis de los aspectos espaciales del mercado de trabajo es fundamental para una mejor comprensión de los fenómenos que en él se desarrollan. Entre las razones que explican este hecho destacan (Fischer, 1986) la constatación de que los mercados de trabajo no son en absoluto homogéneos tal y como sugieren los enfoques neoclásicos, sino que se encuentran segmentados en sus diversas dimensiones: sectorialmente, por ocupación y también territorialmente, entre otras. En este último sentido, en la mayor parte de los países se observan variaciones territoriales sustanciales en cuanto al funcionamiento de los mercados de trabajo y sus principales indicadores, como la tasa de paro o la calidad de los empleos disponibles, algo especialmente relevante cuando las cifras globales (nacionales) de desempleo evolucionan favorablemente. Es sencillo constatar, por ejemplo, que la distribución espacial de la pérdida de empleo asociada a la reestructuración de las actividades productivas suele ser muy distinta de la asociada con la creación de puestos de trabajo. Por último, es indudable que existe un efecto de fricción sobre el ajuste de los mercados de trabajo que se asocia al territorio, a la distancia, en términos, por ejemplo, de información y coordinación entre agentes económicos.

La relevancia de los aspectos espaciales ha hecho necesario abordar el estudio del mercado de trabajo sobre ámbitos geográficos más desagregados. Así, en el caso español, los análisis empíricos referidos a la elaboración de modelos sobre determinación salarial, capital humano o migraciones, entre otros, han incorporado en sus análisis las comunidades autónomas o las provincias como si se tratara de genuinos mercados de trabajo. Cuando el ámbito elegido ha sido el local, los análisis se han referido a divisiones administrativas de mayor o menor relevancia, como los municipios o, en su caso las comarcas. En muy raras ocasiones, sin embargo, se han empleado procedimientos formales con el fin de identificar áreas funcionales que puedan considerarse como verdaderos mercados locales de trabajo (MLT) y que puedan servir como referencia para la obtención y explotación de

¹ Este trabajo es fruto del proyecto "Movilidad cotidiana y mercados de vivienda y trabajo en las áreas urbanas españolas", del Ministerio de Educación y Ciencia (Plan Nacional de I+D+i (BEC 2003-02391), y ha recibido fondos procedentes de los programas FEDER y FSE de la UE.

estadísticas y para un uso académico o en el ámbito de las políticas públicas. Tal y como señala Eurostat (1992), podría darse así la paradoja de que disciplinas que utilizan métodos cada vez más sofisticados, vean socavada la fiabilidad de sus análisis por utilizar áreas geográficas poco apropiadas². A ello se añade los sesgos inevitables cuando se realizan comparaciones internacionales entre áreas de carácter administrativo que no han sido delimitadas siguiendo criterios uniformes. De hecho, la necesidad de disponer de unidades territoriales coherentes y comparables para la obtención y publicación de estadísticas ha sido uno de los principales estímulos que ha impulsado la delimitación de este tipo de áreas en otros países de la OCDE, entre los cuales el caso español constituye una excepción (OCDE, 2002). Esta circunstancia podría estar, sin embargo, cambiando en la actualidad, cuando parece estar abriéndose la posibilidad de colaboración entre el *Instituto Nacional de Estadística* y sus homólogos autonómicos con el fin de obtener estimaciones con un grado de desagregación mayor que el provincial para diversas variables procedentes de la *Encuesta de Población Activa*, lo que otorga una gran relevancia a la reflexión sobre esta cuestión.

La generalización de las delimitaciones de este tipo de áreas por parte de las administraciones públicas de numerosos países ha venido asociada con usos muy variados. Así, por ejemplo, las denominadas *Travel-to-Work Areas* británicas han sido utilizadas desde la década de 1950 para el cálculo de tasas locales de desempleo, y más adelante como referencia para la distribución de ayudas regionales y para la delimitación de regiones objetivo 2 de la Unión Europea. También constituyen la unidad oficial para la identificación de áreas susceptibles de recibir ayudas nacionales o procedentes de la Unión Europea en los casos alemán, finlandés, francés e italiano (OCDE, 2002). Además, en este último caso han servido como unidad de referencia para la identificación de los denominados *distritos industriales* (Sforzi, 1987 y 1995; ISTAT, 1997). En Dinamarca han servido para la reorganización de los gobiernos locales (Andersen, 2002) con el argumento de que quienes se desplazan a trabajar a otro municipio deben contribuir a su financiación, lo que requiere que los límites municipales representen áreas funcionales, de tal forma que pocas personas las crucen en el curso de sus

² Ello ocurriría, por ejemplo, en el caso de que una determinada política de desarrollo fuera implementada en un territorio tras un diagnóstico basado en cifras municipales de registrados en las oficinas de desempleo. Dicha política podría concluir que es necesario estimular la actividad económica en municipios de carácter residencial (en los que la población desocupada podría ser mayor) y dejara de lado los municipios en los que se concentran los puestos de trabajo. Un diagnóstico adecuado y una actuación eficiente exigiría, probablemente, la consideración conjunta de ambos municipios como integrantes de un único MLT que debería servir de referencia territorial para ambos.

desplazamientos residencia-trabajo. Por último, cabe señalar que también se ha dado a los MLT un uso más *normativo*, al servir como instrumento para la planificación. Así, a partir del diagnóstico de una determinada realidad funcional como punto de partida, se han diseñado políticas con el fin de mitigar, por ejemplo, los problemas medioambientales o sociales asociados a unos desplazamientos demasiado largos (véase, por ejemplo, Cervero, 1995, y Van der Laan y Schalke, 2001).

En este artículo se propone un método innovador para la identificación de MLT que reúnan los requisitos necesarios que las hagan útiles para la producción de estadísticas, para análisis de tipo académico y para el diseño, implantación y evaluación de políticas públicas que otorguen relevancia a los aspectos espaciales del mercado de trabajo. El resto del documento se estructura como sigue: en el apartado siguiente se repasan brevemente algunos de los procedimientos oficiales utilizados en otros países. Esta sección concluye con la observación de que dichos métodos abordan sólo parcialmente, y de manera indirecta, este problema como la maximización de una función objetivo. En el artículo proponemos su resolución mediante el uso de un algoritmo evolutivo tras constatar que el problema de agrupación propuesto no es abordable mediante procedimientos exhaustivos de programación matemática. El apartado 3 plantea formalmente el problema y describe el algoritmo utilizado (entre otros aspectos, la función objetivo, las restricciones y los operadores que lo constituyen). En el apartado 4 se muestran algunos resultados experimentales aplicados a la Comunidad Valenciana utilizando para ello datos sobre movilidad residencia-trabajo procedentes del Censo de Población de 2001. El trabajo concluye con algunas consideraciones finales y la descripción de algunas de las vías en las que se sigue desarrollando este proyecto.

2. La delimitación de mercados locales de trabajo (MLT)

Los mercados locales de trabajo suelen ser definidos como áreas de intersección entre (a) la región delimitada por la distancia máxima a la que un grupo determinado de trabajadores estuviera dispuesto a desplazarse desde sus lugares de residencia para aceptar un puesto de trabajo y (b) la región dentro de la cual un grupo significativo de empleadores reclutan habitualmente sus trabajadores. Dado que la aplicación estricta de este tipo de conceptos haría necesario disponer de mapas de preferencias individuales, en la práctica se toman como manifestación de tales preferencias los patrones de movilidad residencia-trabajo (el denominado

commuting) que pueden obtenerse de forma exhaustiva a partir de datos censales³. En la delimitación de estos mercados dos conceptos son fundamentales: coherencia interna e independencia frente a otras áreas. Así, idealmente un mercado debe (i) estar formado por áreas que mantengan fuertes vínculos entre sí y, además, (ii) las fronteras entre mercados deben ser poco permeables a los flujos residencia-trabajo. A estos criterios suelen añadirse otros cuya motivación suele estar relacionada con su uso estadístico o de políticas públicas, como la incorporación de un tamaño mínimo en términos de población ocupada residente, la contigüidad geográfica de las regiones delimitadas, la ausencia de solapamientos entre ellas y la consideración exhaustiva de todo el territorio analizado, de forma que todas las áreas que lo integran formen parte de una de las regiones identificadas, y sólo de una de ellas.

Tal y como se ha señalado en la sección anterior, son muchos los países de la OCDE en los que se delimitan MLT utilizando procedimientos que difieren notablemente entre sí. Una clasificación frecuente (Casado y Coombes, 2005) distingue entre *métodos "multicriterio"* y *métodos jerárquicos*. Los métodos *jerárquicos* se caracterizan por la consideración de una sola regla de clasificación que se aplica una y otra vez sobre los datos de movilidad residencia-trabajo dando lugar a una agregación iterativa de las unidades de base que continúa hasta que se satisface una condición predeterminada. Es el caso de los procedimientos aplicados en Francia (INSÉE, 1998) y alguno de los utilizados en EEUU (Tolbert y Sizer, 1996). La diferencia fundamental entre los distintos métodos que pueden incluirse en esta categoría es la formulación concreta del índice que se usa para la medición de la intensidad de la relación entre pares de áreas/mercados⁴. Una vez que dicha función ha sido definida, la mayor parte de estos métodos procede mediante la fusión iterativa de los pares para los cuales se obtienen los mayores valores. La forma en que estos métodos operan otorga un papel central a la condición de integración *intramercado* frente a la referida a la separación entre mercados, ya

³ Este enfoque difiere del utilizado en otra familia de métodos en los que la delimitación de áreas funcionales se basa en condiciones de homogeneidad (y no en la interacción) entre las áreas integrantes de la región en relación con uno o varios atributos, y de heterogeneidad entre las regiones delimitadas (véase Duque, 2004, para una revisión reciente de estos trabajos).

⁴ En el caso estadounidense el índice de relación toma la forma $\frac{T_{ij} + T_{ji}}{\min(LF_i, LF_j)}$, mientras que en Francia

las funciones utilizadas son $\frac{T_{ij}}{\sum_k T_{ik}}$ y $\frac{T_{ij} + T_{ji}}{\sum_k T_{ik}}$, donde LF_i es la población activa de la zona i y T_{ij} son los

flujos cotidianos de trabajadores entre la zona i y la zona j .

que los procesos no incluyen una condición relativa a la superación de un cierto valor umbral en esta variable.

Una crítica frecuente a los ejercicios de tipo *jerárquico* es que su estricta simplicidad hace que las agregaciones iniciales no se evalúen de nuevo en otras fases del proceso, lo que podría dar lugar a asignaciones sub-óptimas⁵. Además, este tipo de procedimientos suele incluir restricciones de contigüidad para las cuales puede aplicarse un razonamiento análogo al de la nota 4. Otra familia de métodos aplica, por el contrario, diferentes criterios y reglas de decisión a lo largo de las sucesivas fases del algoritmo de regionalización, cuya complejidad varía mucho de unos ejercicios a otros. Son procedimientos con un fuerte contenido teórico y que a menudo parten de una concepción sobre cuáles son las variables explicativas de la realidad funcional analizada. Un ejemplo canónico de este tipo de ejercicios es el de la ya citada identificación de las *Travel-to-Work Areas* británicas, que se realiza a partir del método propuesto por Coombes *et al.* (1986)⁶. Se trata de un procedimiento en el que el logro de unos ciertos niveles de autonomía y tamaño mínimo son imprescindibles para la aceptación del conjunto final de áreas, aunque estos requisitos son incorporados paulatinamente en el procedimiento de forma que no constriñen las asignaciones iniciales. Gran parte del proceso es guiado por los valores obtenidos para la función:

$$\frac{T_{ij}^2}{\sum_{k=1}^n T_{ik} \times \sum_{k=1}^n T_{kj}} + \frac{T_{ji}^2}{\sum_{k=1}^n T_{jk} \times \sum_{k=1}^n T_{ki}}$$

Donde T_{ij} son los flujos residencia-trabajo entre la zona i y la zona j .

Pese a que ninguno de los procedimientos empleados por los países de la OCDE para la delimitación de MLT está planteado como un problema de optimización de una función objetivo, todos ellos, tanto los de carácter jerárquico como los

⁵ Así, por ejemplo, dos áreas podrían estar bastante relacionadas inicialmente, en términos bilaterales, por sus flujos principales de trabajadores, lo que daría lugar a su inclusión en una misma región. Sin embargo puede que tras la delimitación del mapa final de mercados, alguna de las áreas inicialmente agregadas tengan más relación con una región distinta de la suya que agrupe a varias áreas con las cuales mantiene fuertes vínculos que, agregados superan al que decidió inicialmente su adscripción (un flujo que podría superar sólo de forma muy marginal al siguiente en relevancia, sobre todo en entornos multicéntricos o en áreas ubicadas en los límites de las zonas de influencia de distintos centros de empleo). Los métodos "multicriterio" más sofisticados, como el empleado para la identificación de los mercados locales de trabajo británicos (las denominadas *travel-to-work areas*) incorporan diversos pasos que hacen que la solución final sea mucho menos dependiente de la delimitación inicial de áreas centrales de lo que podría pensarse tras un análisis superficial de los mismos.

⁶ Este procedimiento ha sido aplicado también, con ligeras variaciones, para la delimitación de áreas oficiales, *sistemi locali del lavoro*, en Italia (ISTAT-IRPET, 1989; ISTAT, 1997). También ha sido ensayado en otros países como España (Casado, 2000), Dinamarca (Andersen, 2002), Nueva Zelanda (Papps y Newell, 2002) y Australia (Watts, 2004).

procedimientos *multicriterio*, incorporan implícitamente esta idea al basarse en la agregación de áreas que maximizan una cierta medida de la interacción entre ellas. La solución propuesta en este artículo pasa por abordar directamente el problema de maximización de la cohesión interna de las regiones formadas en términos de movilidad residencia-trabajo sujeta al cumplimiento de una serie de restricciones entre las que destaca la superación de un cierto umbral de autonomía y tamaño mínimo, con el objetivo de delimitar tantas áreas como sea posible. Todo ello sin incorporar requisitos de contigüidad geográfica y sin utilizar la variable distancia sacando partido, por tanto, a la naturaleza de la variable básica, la movilidad residencia-trabajo que, por su propio carácter territorial, es capaz de dar lugar a soluciones viables. Se trata de un procedimiento inspirado en la propuesta de Coombes *et al.* (1986) y ONS (1998). Se pretende dar solución, por tanto, a un problema de agrupación (*clustering*) a partir de la información que proporciona una matriz de flujos residencia-trabajo entre municipios. Este tipo de problemas es muy complejo, incluso para un número reducido de unidades de base, y ha sido frecuentemente clasificado como *NP-duro* (Murthy y Chowdhury, 1996; Sarkar *et al.*, 1997, Cucchiara, 1998; Demiriz *et al.*, 1999); es decir, no es factible computacionalmente encontrar la solución óptima de una forma exhaustiva, por lo que se hace necesario utilizar métodos alternativos de optimización. Este artículo propone abordar el problema descrito mediante utilizando técnicas de computación evolutiva⁷. La aplicación de este enfoque⁸ permitirá mejorar sustancialmente la calidad de las regionalizaciones propuestas en relación con los métodos oficiales empleados actualmente aunque, tal y como se señala en el apartado siguiente, no pueda asegurarse la consecución de la mejor de las soluciones posibles (algo difícilmente contrastable, en cualquier caso, en ámbitos territoriales amplios dada la imposibilidad de dar solución al problema de optimización de forma exhaustiva).

⁷ Pese a su relativa novedad existe ya un gran volumen de trabajos en el ámbito de la Economía que utilizan algoritmos evolutivos para la resolución problemas en campos muy diversos. En una recopilación reciente de artículos en la que, además, se relaciona gran parte de estos trabajos, Chen y Kuo (2002) señalan que los ámbitos en los que existen más trabajos son la predicción financiera, las estrategias de comercio, y la simulación de mercados financieros artificiales (donde destaca el *Modelo Cobweb*). Entre las aplicaciones más relacionadas con la Geografía destacan, sobre todo, los diversos trabajos de Openshaw que, sin embargo, no han concluido con la utilización de este tipo de instrumentos para la delimitación de áreas funcionales.

⁸ Aunque en este caso se ha partido de una función de adecuación directamente inspirada en el método de Coombes *et al.* (1986) se trata, en cualquier caso, de un problema complejo también porque dadas varias particiones posibles del territorio existen numerosas medidas internas para estimar su calidad (Halkidi *et al.*, 2000), lo que complica la formulación de una función objetivo que pueda dirigir el proceso de agrupación (Jin 2003). Además debe considerarse el problema del *ruido* que caracteriza a los datos de entrada, bastante probable dado el origen censal de los datos utilizados.

3. Los algoritmos evolutivos como instrumentos para la delimitación de MLT

3.1. Una nota sobre algoritmos evolutivos

Como es sabido, los algoritmos evolutivos son un conjunto de paradigmas de procedimientos estocásticos de búsqueda global y optimización inspirados en los mecanismos biológicos de selección natural (Angeline, 1993; Bäck, 1996 y Freitas, 2002). Entre ellos destacan los algoritmos genéticos, las estrategias evolutivas, la programación evolutiva (véase Tomassini, 1995) y la programación genética (véase, por ejemplo, Angeline, 1996). Entre las razones que explican la generalización de su uso para la resolución de problemas en un gran número de disciplinas se encuentra su capacidad para encontrar soluciones óptimas de carácter más general que las alcanzadas por otros métodos más tradicionales, basados en heurística o *backtracking*, entre otros, especialmente en espacios no lineales y de elevada dimensión. Por otro lado, los algoritmos evolutivos permiten mayores grados de libertad en la definición de la función objetivo, que no necesita ser continua, convexa, diferenciable o siquiera tener una forma explícitamente cerrada (Goldberg, 1989; Mitchell, 1996).

La estructura genérica de un algoritmo evolutivo es la siguiente:

- | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><i>Paso 1:</i> Crear la población inicial.</p> <p><i>Paso 2:</i> Evaluar la adecuación de cada individuo de la población inicial.</p> <p><i>Paso 3:</i> Obtener población intermedia aplicando el operador de selección.</p> <p><i>Paso 4:</i> Generar nuevos individuos, aplicando los operadores de reproducción a la población intermedia, y añadirlos a la población actual.</p> <p><i>Paso 5:</i> Obtener población intermedia aplicando el operador de selección.</p> <p><i>Paso 6:</i> Generar nuevos individuos aplicando los operadores de mutación a la población intermedia, y añadirlos a la población actual.</p> <p><i>Paso 7:</i> Evaluar la adecuación de cada individuo de la población actual.</p> <p><i>Paso 8:</i> Volver al paso 2 si no se cumple la condición de terminación, o devolver como solución el mejor individuo.</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Un algoritmo evolutivo consiste, por tanto, en un proceso iterativo que, basándose en una *función de adecuación* (función objetivo) que evalúa el grado de idoneidad de las posibles soluciones, trabaja como sigue: A partir de una población inicial de *genes* (soluciones del problema), escoge mediante cierto esquema de selección un subconjunto de dicha población, sobre el que aplica una serie de *operadores de cruce* (reproducción) y *mutación*, obteniendo nuevos individuos que se incorporan a la población, la cual es evaluada de nuevo para decidir cuáles de entre ellos pasan a la siguiente *generación* (iteración). Los *genes* suelen representarse como vectores

cuyas celdas codifican características de las posibles soluciones. En su formulación estándar los *operadores de cruce*, mediante la combinación aleatoria de las características de diferentes individuos (normalmente dos, llamados *padres*) en un nuevo individuo (*prole*), son asimilables a una técnica de búsqueda global o exploración (grandes saltos en el espacio de búsqueda). Por su parte, los operadores de mutación, que cambian de forma aleatoria una o más de las características de un individuo, implementan la búsqueda local o explotación (*pequeños saltos*). De esta forma el algoritmo emula el *proceso de selección natural (evolución competitiva)*⁹.

Una de las críticas que suele hacerse a la utilización de estos métodos (Rudolph, 1994; Tomassini, 1995; y Szpiro, 2002) se refiere a la convergencia determinista de los algoritmos evolutivos, dado que no está asegurado formalmente que se alcance un estado estacionario habiendo encontrado la solución óptima. No debe olvidarse, sin embargo, que dicha seguridad sólo estaría garantizada en el caso de una resolución del problema a través de programación matemática mediante un algoritmo exhaustivo que en entornos como el planteado no es factible. Existen, en cualquier caso, mecanismos que pueden paliar este tipo de inconvenientes. Así, por ejemplo, en los apartados siguientes se argumenta que el algoritmo propuesto se sirve de un procedimiento de selección aleatoria uniforme que, pese a ser uno de los métodos más lentos, dado que la selección no se ve influida por la idoneidad de los individuos, impide el estancamiento en un óptimo local por endogamia, y ofrece mayores garantías de que, permitiendo un tiempo de procesamiento suficientemente largo, puedan alcanzarse soluciones extremadamente cercanas al óptimo global (Miller *et al.*, 1996). En este mismo sentido, también el efecto de los operadores de mutación y cruce es determinante (Goldberg, 1998; Goldberg, 1999; y Sinha y Goldberg 2001). Tal y como se ha señalado ya, los operadores de mutación tradicionales, al introducir ligeros cambios en los genes, toman el papel de un algoritmo de optimización local o explotación, mientras que los operadores de cruce realizan la función de un algoritmo de búsqueda global o exploración, al asociarse con grandes saltos en el espacio de búsqueda. Por último, recientemente se han propuesto procedimientos que, mediante la aplicación de métodos

⁹ Entre las referencias que permiten seguir el desarrollo de esta familia de métodos y sus distintas ramificaciones se encuentran Holland (1975), Goldberg (1989), Davis (1991), Michalaevicz (1994) y Goldberg (1998). En Coello (2002) puede encontrarse una revisión bastante sistemática de las técnicas más comúnmente utilizadas por estos métodos. Por último, en el ámbito de la Economía, Hernández (2004) y Fisher y Leung (1998), revisan las principales características de estos métodos antes de aplicarlos a la resolución de dos problemas concretos.

estadísticos tras sucesivas ejecuciones independientes del algoritmo, permiten generar reglas de decisión en torno a cuándo se ha alcanzado el óptimo global (Szpiro, 2002) así como otros algoritmos evolutivos paralelos en los que varias poblaciones independientes (islas) evolucionan por separado salvo por intercambios de individuos que se efectúan periódica o azarosamente, siguiendo o no alguna estrategia de migración que defina entre qué *islas* se realiza (Tomassini, 1995; Nowostawski y Poli, 1999, y Talbi 2002).

3.2. Planteamiento del problema

Sea $A = \{A_1, A_2, \dots, A_n\}$ un conjunto de áreas. El objetivo es obtener el conjunto de regiones $R = \{R_1, R_2, \dots, R_m\}$ tal que $\bigcup_{i=1}^m R_i = A$ y $R_i \cap R_j = \emptyset, \forall i, j \in [1, m], i \neq j$, que maximice la función¹⁰

$$\sum_{\forall i \in A} (f(i) - p(i)) \quad (1)$$

donde

$$f(i) = \frac{T_{\{i\}, R(i)-\{i\}}^2}{T_{\{i\}, A} \times T_{A, R(i)-\{i\}}} + \frac{T_{R(i)-\{i\}, \{i\}}^2}{T_{R(i)-\{i\}, A} \times T_{A, \{i\}}} \quad (2)$$

y

$$p(i) = \max_{\forall R_j \in R - R(i)} \left(\frac{T_{\{i\}, R_j}^2}{T_{\{i\}, A} \times T_{A, R_j}} + \frac{T_{R_j, \{i\}}^2}{T_{R_j, A} \times T_{A, \{i\}}} \right) \quad (3)$$

siendo $R(i)$ la región a la que pertenece el área i , y

$$T_{R_s, R_t} = \sum_{\forall i \in R_s} \sum_{\forall j \in R_t} T_{ij} \quad (4)$$

donde T_{ij} es el número de ocupados que diariamente se desplaza para realizar su trabajo desde el municipio i hasta el municipio j .

Asimismo, cada uno de las regiones $R_i \in R$ debe cumplir una serie de restricciones:

$$(a) \quad \min \left(\frac{T_{R_i, R_i}}{T_{R_i, A}}, \frac{T_{R_i, R_i}}{T_{A, R_i}} \right) \geq \beta_1 \quad (5)$$

$$(b) \quad T_{R_i, A} \geq \beta_4 \quad (6)$$

¹⁰ Esta función incluye tanto el índice de interacción propuesto por Coombes et al. (1986) como una función de penalización construida a partir de dicho índice. (2) se refiere a la interacción de un área con el resto de la región a la que pertenece, mientras que (3) toma el máximo valor de entre índices de interacción entre dicha área y el resto de regiones (para una discusión sobre el uso de funciones de penalización en computación evolutiva véase, por ejemplo, Coello, 2002).

$$(c) \quad \min \left(\frac{T_{R_i, R_i}}{T_{R_i, A}}; \frac{T_{R_i, R_i}}{T_{A, R_i}} \right) \geq a + b \times T_{R_i, A} \quad (7)$$

$$a = \beta_2 - b\beta_4$$

$$b = \frac{\beta_2 - \beta_1}{\beta_4 - \beta_3}$$

Las restricciones (6) a (7) sirven para flexibilizar la introducción de un tamaño y una autonomía mínima¹¹, y articulan una variación del *trade-off* entre ambos criterios propuesto por Coombes *et al.* (1986) siguiendo la formulación de Casado (2000). Así, a pesar de que el mínimo de tamaño es β_4 y el mínimo de autonomía es β_2 , en el caso de MLT muy poblados, cuyo número de ocupados residentes supera el umbral β_3 se acepta un nivel de autonomía algo inferior, del β_1 .

3.3. Representación

Cada uno de los individuos de la población representa una posible solución; es decir, una regionalización del conjunto de áreas pertenecientes a A. Su representación se realiza mediante un vector de tamaño n, donde cada componente del vector se corresponde con un área de A, tomando el valor del identificador de la región a la que pertenece dicha área.

Individuo									
1	2	1	3	2	1	3	2	3	4
$R_1 = \{a_1, a_3, a_6\}$					$R_3 = \{a_4, a_7, a_9\}$				
$R_2 = \{a_2, a_5, a_8\}$					$R_4 = \{a_{10}\}$				

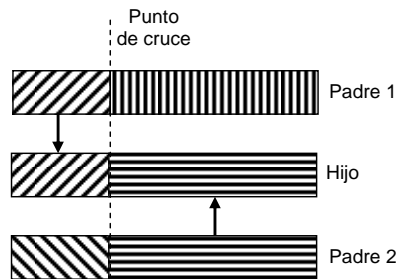
3.4. Selección

La selección de los individuos que experimentan una recombinación o una mutación se realiza de forma uniforme. En este método los padres son elegidos con la misma probabilidad (independientemente de su función de adecuación). La selección competitiva (que sí emplea la evaluación de los individuos) se realiza al final de cada iteración, y permite escoger el conjunto que formará la próxima generación (los mejores 100 individuos obtenidos).

¹¹ La autonomía se define como (i) la proporción de ocupados residentes en el área/región que realiza sus tareas laborales dentro de sus límites –autonomía de oferta y (ii) la proporción de puestos de trabajo del área/región que son ocupados por residentes en la misma zona –autonomía de demanda. El valor crítico utilizado es el mínimo de ambos, tal y como recoge la ecuación (5).

3.5. Operadores de cruce

- **Cruce1:** Se selecciona aleatoriamente un punto de cruce. La descendencia se forma tomando la parte inicial de uno de los padres y la parte final del otro. Este es el operador de cruce usualmente empleado en algoritmos genéticos, sin embargo, puede no existir una relación entre la numeración de las regiones de ambos padres, lo que da lugar a descendientes que no siguen estrictamente la regionalización de los padres.



Padre 1	1	2	1	3	2	1	3	2	3	4
---------	---	---	---	----------	---	---	----------	---	---	---

Hijo	1	2	1	3	2	3	4	2	4	5	Área de cruce = 4
------	---	---	---	----------	---	---	----------	---	---	---	-------------------

Padre 2	1	2	3	4	2	3	4	2	4	5
---------	---	---	---	----------	---	---	----------	---	---	---

Para evitar discrepancias en la numeración de las regiones de ambos padres se han desarrollado dos nuevos operadores de cruce:

- **Cruce2:** Se escoge al azar un identificador de región perteneciente a uno de los padres (padre 1). A la descendencia pasan las áreas de padre 1 con un identificador menor o igual a la región de cruce. Para el resto de áreas, se asigna la región de padre 2, siempre que no estuvieran en una región junto a un área que ya esté en el hijo. En ese caso, se asigna el grupo del padre.

Padre 1	1	2	1	3	2	1	3	2	3	4
---------	---	---	---	---	---	---	---	---	---	---

Hijo	1	2	1	4	2	1	4	2	4	5	Región de cruce = 2
------	---	---	---	---	---	----------	---	---	---	---	---------------------

Padre 2	1	2	3	4	2	3	4	2	4	5
---------	---	---	---	---	---	---	---	---	---	---

- **Cruce3:** Se escoge un área de cruce al azar (padre 1). Para las áreas previas a la de cruce, se le asigna el valor del padre 1. A partir del área de cruce, se toman los valores del padre 2, siempre que no estuvieran en una región junto a un área que ya esté en el hijo. En ese caso, se asigna el grupo del padre.

Padre 1	1	2	1	3	2	1	3	2	3	4
---------	---	---	---	---	---	---	---	---	---	---

Hijo	1	2	1	3	2	1	3	2	3	5	Área de cruce = 4
------	---	---	---	---	---	---	----------	---	----------	---	-------------------

Padre 2	1	2	3	4	2	3	4	2	4	5
---------	---	---	---	---	---	---	---	---	---	---

3.6. Operadores de mutación

Se ha diseñado un conjunto extenso de operadores de mutación, algunos de los cuales están realizados específicamente para el problema a resolver, con el objetivo de intentar acelerar la obtención de individuos con una buena función de adecuación.

- **Mutación1:** Este operador es el empleado genéricamente en computación evolutiva, mutando un gen y asignándole una región al azar distinta de aquella a la que pertenecía. La única diferencia radica en que en lugar de mutar un único gen, se muta un número de genes determinado aleatoriamente.

Padre	1	2	1	3	2	1	3	2	3	4
-------	---	---	---	---	---	---	---	---	---	---

Nº de mutaciones = 3

Hijo	1	2	3	3	2	2	3	2	3	1
------	---	---	----------	---	---	----------	---	---	---	----------

- **Mutación2:** Este método es similar al anterior, salvo que en lugar de escoger al azar las regiones que se asignan a las áreas mutadas, se las agrupa con sus regiones óptimas, entendidas como las regiones distintas a la que pertenecen con las que su índice de interacción es mayor:

$$R'(i) = \underset{\forall R_j \in R-R(i)}{\operatorname{argmax}} \left(\frac{T_{\{i\},R_j}^2}{T_{\{i\},A} \times T_{A,R_j}} + \frac{T_{R_j,\{i\}}^2}{T_{R_j,A} \times T_{A,\{i\}}} \right) \quad (8)$$

- Mutación3: Este operador es aplicable únicamente si se dispone de las situaciones geográficas de los municipios o de las distancias entre éstos. Es similar a los anteriores. Las áreas mutadas son asignadas a las regiones a la que están más próximas, calculando la distancia entre el área y el centroide de la región:
- Mutación4: En este caso, se seleccionan dos regiones escogidas aleatoriamente.
- Mutación5: Se selecciona una región al azar. Se asigna cada área que la forma a su región óptima. Como resultado de esta mutación, el nuevo individuo constará de una región menos que su padre.
- Mutación6: Este operador divide una región formando dos regiones a partir de ésta. El proceso es el siguiente:
 1. Se selecciona aleatoriamente una región R_i que cumpla que $T_{R_i,A} > 2\beta_4$ y $T_{R_i,A} - T_{\text{foco}(R_i),A} > \beta_4$ donde $\text{foco}(R_i) = \underset{\forall a \in R_i}{\text{argmax}} (T_{\{a\},A} + T_{A,\{a\}})$.
 2. Se selecciona aleatoriamente una zona perteneciente a R_i asignándola a la región R_i' .
 3. Se selecciona aleatoriamente otra zona perteneciente a R_i asignándola a la región R_i'' .
 4. El resto de zonas pertenecientes a R_i son tomadas en un orden aleatorio, asignándolas a la región R_i' o R_i'' con la que tenga mayor índice de interacción.
- Mutación7: Esta operación disgrega de una región de un tamaño grande aquellas zonas que tienen menos interacción con el resto de la región. El procedimiento seguido es:
 1. Se selecciona aleatoriamente una región R_i que cumpla que $T_{R_i,A} > 2\beta_4$ y $T_{R_i,A} - T_{\text{foco}(R_i),A} > \beta_4$.
 2. Se selecciona la zona a disgregar, obtenida como:

$$s = \underset{\forall j \in R_i}{\text{argmin}} \left(\frac{T_{\{j\},R_i-\{j\}}^2}{T_{\{j\},A} \times T_{A,R_i-\{j\}}} + \frac{T_{R_i-\{j\},\{j\}}^2}{T_{R_i-\{j\},A} \times T_{A,\{j\}}} \right) \quad (9)$$
 3. Si $T_{R_i-\{s\},A} > \beta_4$, el área s es asignada a su región óptima y se vuelve otra vez al paso 2. Si no se cumple la condición, finaliza la mutación.

- Mutación8: Este operador intercambia áreas entre dos regiones. Se escoge un área aleatoriamente asignándola a su región óptima. Un área al azar de esa región es traspasada a la región original.

- Mutación9: Esta mutación es similar a Mutación2 en el sentido de que las áreas son asignadas a su región óptima. La diferencia estriba en que en lugar de trasladar una única área, se transfieren un grupo de áreas fuertemente cohesionado. El proceso es el siguiente:
 1. Se selecciona aleatoriamente un área i .
 2. Se obtienen las k áreas pertenecientes a $R(i)$ con las que i tiene mayor índice de interacción. k es obtenido al azar.
 3. i y las k áreas seleccionadas son asignadas a la región óptima de i .

- Mutación10: Esta operación pretende redistribuir todas las áreas pertenecientes a varias regiones que tienen interacción. El procedimiento a seguir es:
 1. Se escoge al azar un número $k \geq 2$ de regiones a mutar.
 2. Se selecciona una región R_i al azar.
 3. Se seleccionan las $k - 1$ regiones con las que R_i tiene mayor interacción.
 4. Todas las áreas de esas regiones dejan de estar asignadas.
 5. Se seleccionan k áreas al azar de entre ellas. Estas áreas van a actuar como semilla de esas nuevas regiones.
 6. El resto de áreas van siendo tomadas al azar asignándolas a su región óptima de entre las k nuevas regiones.

- Mutación11: Esta operación es muy similar al anterior, variando únicamente el modo de conformar las nuevas regiones. El procedimiento seguido es:
 1. Se escoge al azar un número $k \geq 2$ de regiones a mutar.
 2. Se selecciona una región R_i al azar.
 3. Se seleccionan las $k - 1$ regiones con las que R_i tiene mayor interacción.
 4. Todas las áreas de esas regiones dejan de estar asignadas.
 5. Para cada una de esas k regiones:
 - i. Se seleccionan un área al azar como semilla.

- ii. Se van tomando áreas con las que la región en formación tiene mayor índice de interacción. Este proceso se detiene cuando el tamaño de la región es mayor que β_4 .
6. Las áreas que continúen sin asignar son añadidas a su región óptima de entre esas k regiones.

4. Resultados experimentales

La propuesta evolutiva presentada ha sido empleada para la delimitación de mercados locales de trabajo de la Comunidad Valenciana utilizando datos sobre la movilidad residencia-trabajo de los ocupados procedentes del Censo de Población de 2001¹². La información permite construir una matriz 541×541 en la que cada celda recoge el número de trabajadores (T_{ij}) que se desplaza para realizar sus tareas laborales entre el municipio indicado por la fila, i , y el municipio señalado por la columna, j (la diagonal principal recoge el volumen de trabajadores que trabajan en el municipio de residencia, T_{ii}).

Los parámetros empleados en la ejecución del algoritmo evolutivo han sido los siguientes:

- Tamaño de la población: 100 individuos
- Número de cruces por generación: 30, con 10 de cada uno de los tres modos de cruce planteados
- Número de mutaciones: 100. Se han empleado todas las formas de mutación excepto el tipo Mutación3 ya que no se ha deseado introducir información sobre distancia entre áreas a la hora de calcular la delimitación.
- Como condición de finalización del proceso evolutivo se ha establecido que si en 1000 generaciones no varía el mejor individuo se asume que se ha alcanzado la convergencia. Este individuo es la solución al problema, es decir, la delimitación en mercados locales de trabajo de un conjunto de áreas dado.

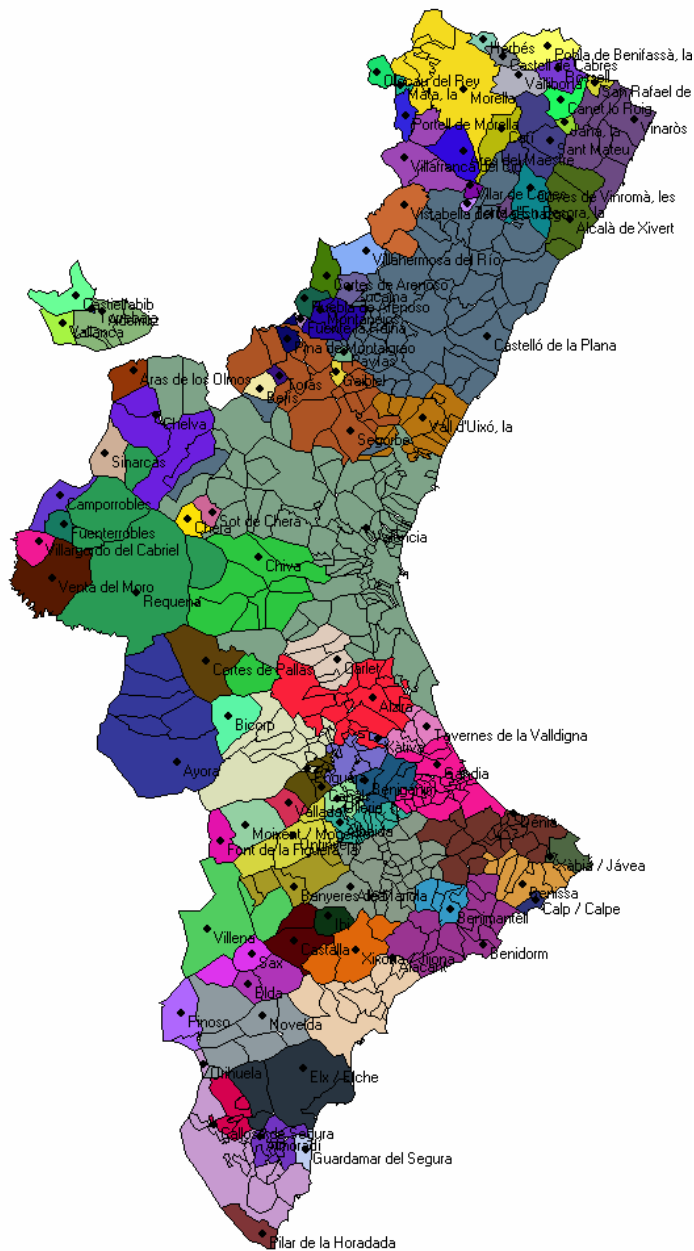
¹² Los datos han sido facilitados por el *Instituto Nacional de Estadística*. Dada el problema que se pretende resolver no se incluyen los individuos cuyo desplazamiento tiene su origen o destino en otras comunidades autónomas o en el extranjero, ni tampoco quienes declaran no tener un lugar fijo de trabajo. Los datos de ocupados incluyen también a quienes compatibilizan dicha actividad con los estudios.

La población inicial es generada aleatoriamente, salvo por un individuo en el que todos sus genes tienen el mismo valor; es decir, todas las áreas están unidas en una única región. De esta forma se consigue partir de una solución válida (*i.e.* que cumple las restricciones de autonomía y tamaño), y reducir así el tiempo empleado para la obtención de una primera solución válida.

Los individuos utilizados en cada uno de los cruces o mutaciones son seleccionados de entre todos los ya existentes, tanto los de la generación previa como los ya creados en la generación actual. Una vez instituida toda la generación y calculada la función de adecuación de todos los individuos, se realiza la selección de los mejores individuos de entre todos los existentes. Es decir, no se eliminan los individuos de la generación anterior, sino que éstos entran también en el proceso de selección y pueden, por tanto, mantenerse entre la población.

En el mapa 1 se presenta el resultado de la delimitación de la CV empleando los siguientes parámetros para las restricciones (5) (6) (7) relativas a la autonomía y al tamaño de las regiones: $\beta_1=0.70$, $\beta_2=0.75$, $\beta_3=20000$ y $\beta_4=3500$. Como función de adecuación se ha empleado la presentada en el apartado 3.

β_4 iguales a cero. Los valores mínimos de autonomía permitidos son $\beta_1=0.75$ y $\beta_2=0.75$. El resultado se presenta en el mapa 2.



Mapa 2. Delimitación de la CV sin restricciones de tamaño

Uno de los criterios que guían la delimitación de MLT es el la delimitación de tantas áreas como sea posible; esto es, entre dos regionalizaciones distintas se prefiere aquella que consta de un número mayor de MLT que cumplen los requisitos estadísticos prefijados. Con el fin de potenciar la relevancia de este criterio se ha

ensayado el funcionamiento del algoritmo diseñado sustituyendo la función de adecuación inicialmente propuesta por la función (10):

$$\left(\sum_{i \in A} (f(i) - p(i)) \right) \times \text{card}(R) \quad (10)$$

donde $\text{card}(R)$ es el número de regiones del individuo. Esta función es similar a la de los ejemplos previos, sin embargo, potencia la división de regiones, premiando a los individuos (regionalizaciones) con un número mayor de regiones (MLT). El resultado con restricciones de tamaño se presenta en el mapa 3 y sin restricciones en el mapa 4.

Los cuatro mapas mostrados en esta sección reflejan de forma muy coherente, y análoga a los resultados de algunos de los otros métodos oficiales existentes, la realidad funcional del territorio de la Comunidad Valenciana. Así, destaca el mayor tamaño medio de los MLT de gran parte de las provincias de Castellón y Valencia, especialmente en el caso de los dos grandes mercados centrados en ambas ciudades y, frente a ellos, unos mercados de tamaño mucho más homogéneo que cubren gran parte de la provincia de Alicante, que se caracteriza por la existencia de un mayor número de ciudades de tamaño intermedio que sirven como focos a unos mercados que, además, son bastante robustos frente a cambios en los distintos parámetros empleados. Los mapas en los que la restricción de tamaño no opera constan, obviamente, de un número sustancialmente mayor de mercados, entre los que se incluyen un gran número de mercados formados por un único municipio en las zonas más rurales del norte de la provincia de Castellón y el interior de la provincia de Valencia. La incorporación del número de MLT en la función de adecuación (10) logra, por su parte, su objetivo, al aumentar el número de MLT independientes delimitados. Por último, cabe destacar el éxito del algoritmo en la identificación de mercados constituidos por municipios contiguos a pesar de que esta condición o ha sido incorporada en el proceso, como tampoco lo ha sido la variable distancia, a pesar de que uno de los operadores de mutación diseñado (que no se ha utilizado) incorporaba su uso.

de regiones (Eurostat, 1992) incluyen la ausencia de solapamientos entre los MLT y el cubrimiento exhaustivo del territorio analizado. Cada uno de los MLT está formado por una o más áreas (municipios, por ejemplo) y se define por (a) un alto nivel de integración interna –una gran interacción entre las áreas que lo componen– y (b) un elevado nivel de autonomía –es decir, un bajo nivel de interacción con el resto de MLT; en ambos casos la condición se formula en términos de movilidad residencia-trabajo¹³. Además, suelen incorporarse otros criterios, como la contigüidad geográfica entre todas las áreas que constituyen un MLT, la preferencia por el detalle (por la maximización del número de MLT independientes) y, en algunos casos, pese a existir un acuerdo generalizado sobre las características ideales de estos mercados, existen enormes diferencias entre los procedimientos empleados por las administraciones públicas en los diversos países de la OCDE (2002). Prácticamente todos los métodos tienen en común el papel central que otorgan a la definición de un índice de interacción entre las áreas que guía el proceso (Casado y Coombes, 2005). En algún caso, como el que inspira el procedimiento propuesto (Coombes *et al.*, 1986), existen requisitos explícitos que controlan la superación de un umbral de autonomía y tamaño. Sin embargo, en ninguno de ellos el problema se plantea como la maximización de una función objetivo sujeta a un conjunto de restricciones. Ello se debe, probablemente, a que la magnitud del problema de *clustering* que debería resolverse lo hace inabordable mediante técnicas tradicionales.

En este trabajo se propone un enfoque sustancialmente diferente. El problema es abordado aquí mediante la utilización de un algoritmo evolutivo diseñado de manera específica tanto en lo que se refiere a función de adecuación (que se expresa como una función de interacción a la que se incorpora una de penalización) como a los operadores de cruce y mutación, que han sido desarrollados *ad hoc* para contribuir a la solución del problema. Los resultados obtenidos en la fase experimental muestran la capacidad del algoritmo para lograr regionalizaciones que cumplen con los requisitos establecidos y aumentan de forma notable el número de MLT independientes identificados. Al igual que el modelo *tradicional* que se toma como referencia (Coombes *et al.*, 1986), el algoritmo incorpora una gran flexibilidad en cuanto a la determinación de umbrales de autonomía deseados, así como la posibilidad o no de incorporar una restricción de tamaño (algo

¹³ Se trata, por tanto, de una regionalización basada en un flujo, en una variable de interacción, lo que abre numerosas posibilidades para su aplicación en otros ámbitos, como el análisis de los flujos migratorios, de comunicaciones o de bienes, por ejemplo.

especialmente relevante en el caso de las políticas públicas y el diseño de áreas para su uso estadístico) y el establecimiento de una relación de *trade-off* entre ambos criterios. Por otro lado, el procedimiento diseñado se ha mostrado muy eficiente en la identificación de MLT continuos geográficamente sin que haya sido necesario incorporar requisitos explícitos de contigüidad y, sin la utilización de información sobre la distancia geográfica/temporal entre las áreas, algo que la literatura especializada ha señalado como posible origen de soluciones sub-óptimas.

El diseño de operadores específicos abre la puerta a la experimentación posterior con *Grouping Genetic Algorithms* (Lorena y Furtado, 2001). Se trata de una variante de los algoritmos genéticos que aborda problemas relativamente análogos al descrito aquí, aunque no se han realizado aplicaciones en este ámbito. Los GGAs se caracterizan por una representación distinta de los genes, en forma de árbol, y algunos resultados recientes señalan que podrían converger más rápidamente (Brown y Sumichrast, 2005), aunque a cambio exigen un mejor conocimiento del problema tratado para el diseño de operadores de mutación y cruce especializados.

En relación con la función de adecuación, cabe señalar la dificultad de identificar los criterios más relevantes para la obtención de MLT apropiados, así como la ordenación de los mismos de acuerdo con su importancia (integración, autonomía, tamaño, contigüidad, compacidad, número de áreas identificadas, entre otros). La investigación desarrollada en este trabajo continúa en dos sentidos: por un lado, mediante el desarrollo de una función que unifique el mayor número posible de criterios para su uso en cualquier tipo de algoritmo de optimización y/o agrupación, incluyendo los que no permiten múltiples funciones. En segundo lugar, mediante el diseño de un algoritmo *multiobjetivo*, no basado en una medida de adecuación unificada como la anterior, sino en varias medidas ortogonales de la calidad de las agrupaciones (Handl y Knowles, 2004). De esta forma quedaría en manos del algoritmo el encontrar un equilibrio entre cada una de las medidas adecuado para la resolución del problema, evitando la dificultad de emplear funciones sesgadas a causa de una mala interpretación en el diseño.

Por otro lado, en relación con la optimización del método, se está trabajando en una técnica de mejora de la población inicial (*population seeding*) para acelerar la convergencia del método (Sastry, 2001) con el convencimiento de que esta técnica podría usarse también para incrementar la fiabilidad de la búsqueda global, introduciendo en la población individuos de diseño específico que se encuentren

dispersos por todo el espacio de soluciones. Además, se está llevando a cabo el análisis de la convergencia del algoritmo según diferentes esquemas de selección de los operadores de mutación y cruce que se aplican en cada iteración, con la intención de agotar la exploración global en etapas tempranas de la ejecución (localizando todos los posibles óptimos locales) para acotar mejor las búsquedas locales (véase Lobo y Goldberg, 1997; Goldberg y Voessner, 1999; Sinha y Goldberg 2001). Por último, en cuanto a la utilización de poblaciones independientes (islas) en un algoritmo paralelo, se están desarrollando diferentes esquemas de migración para su estudio y aplicación.

Referencias.

- Andersen, A.K. (2002) Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark?, *Regional Studies* 36 (9), 833-44.
- Angeline, P.J. (1993) Evolutionary Algorithms and Emergent Intelligence. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.
- Angeline, P.J. (1996) Genetic Programming's Continued Evolution. *Advances in Genetic Programming* 2(1), 1-20.
- Bäck, T. (1996) Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. New York: Oxford University Press.
- Brown, E.C., Sumichrast, R.T. (2005) Evaluating Performance Advantages of Grouping Genetic Algorithms. *Engineering Applications of Artificial Intelligence*, Vol 18, pp. 1-12.
- Calégari, P., Coray, G., Hertz, A., Kobler, D., Kuonen, P. (1999) A Taxonomy of Evolutionary Algorithms in Combinatorial Optimization. *Journal of Heuristics* 5, 145-158.
- Casado-Díaz, J.M. (2000) Local Labour Market Areas in Spain: A Case Study, *Regional Studies* 34 (9), 843-56.
- Casado-Díaz, J.M. and Coombes, M.G. (2005) The Delineation of 21st Century Local Labour Market Areas (LLMAs). Proceedings of the 8th Nectar Conference, Las Palmas C.G.
- Cervero, R. (1995) Planned Communities, Self-containment and Commuting: A Cross-national Perspective, *Urban Studies* 32(7), 1135-1161.
- Coello, C.A. (2002) Theoretical and Numerical Constraint-Handling Techniques used with Evolutionary Algorithms: A Survey of the State of the Art. *Computer Methods in Applied Mechanics and Engineering*, 191(11-12), 1245-1287.
- Coombes, M.G.; Green, A.E. and Openshaw, S. (1986) An efficient algorithm to generate official statistical reporting areas: the case of the 1984 Travel-to-Work Areas revision in Britain, *Journal of the Operational Research Society* 37, 943-53.
- Cucchiara, R. (1998) Genetic algorithms for clustering in machine vision. *Machine Vision and Applications*, 11, 1-6.
- Chen S.-H., Kuo, T.-W. (2002) Evolutionary Computation in Economics and Finance: A Bibliography. In: Chen (Ed.), *Evolutionary Computation in Economics and Finance*, Studies in Fuzziness and Soft Computing, Vol. 100. Heidelberg: Physica-Verlag, 419-455.
- Davis, L. (1991) (Ed.) *Handbook Genetic Algorithm*. New York: Van Nostrand Reinhold.
- Demiriz, A., Bennett, K.P., Embrechts, M.J. (1999) Semi-supervised Clustering Using Genetic Algorithms. In: *Artificial Neural Networks in Engineering (ANNIE-99)*, 809-814.
- Duque, J.C. (2004) Design of Homogeneous Territorial Units. A Methodological Proposal and Applications, Unpublished Ph Dissertation, Department of Econometrics, Statistics and Spanish Economy, Universitat de Barcelona, Spain.
- EUROSTAT and Coombes, M.G. (1992) *Étude sur les zones d'emploi*. Document E/LOC/20. Luxembourg, Office for Official Publications of the European Communities.
- Fischer, M.M. (1986) Why spatial labour market research?. *Environment and Planning A* 18, 1417-1420.
- Fischer, M.M., Leung, Y. (1998) A Genetic-Algorithms Based Evolutionary Computational Neural Network for Modelling Spatial Interaction Data. *The Annals of Regional Science* 32, 437-458.
- Freitas, A.A. (2002) A survey of evolutionary algorithms for data mining and knowledge discovery. In: Ghosh, A.; Tsutsui, S. (Eds.) *Advances in evolutionary computation*. Berlin: Springer-Verlag.
- Goldberg, D.E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, M.A.: Addison Wesley.
- Goldberg, D.E. (1998) The Design of Innovation: Lessons from Genetic Algorithms, Lessons for the Real World. IlliGAL Report No. 98004. Urbana, IL: University of Illinois at Urbana-Champaign.
- Goldberg, D.E., Voessner, S. (1999) Optimizing Global-Local Search Hybrids. Proceedings of the Genetic and Evolutionary Computation Conference - 1999, 220-228.

- Halkidi, M., Vazirgiannis, M., Batistakis, I. (2000) Quality scheme assessment in the clustering process. In Proceedings of the Fourth European Conference on Principles of Data Mining and Knowledge Discovery, volume 1910 of LNCS, pages 265–267. Heidelberg, Germany: Springer-Verlag.
- Handl, J., Knowles, J. (2004) Evolutionary Multiobjective Clustering. PPSN VIII, LNCS 3242, 1081-1091.
- Hernández López, M. (2004) Predicción económica con algoritmos genéticos: operadores genéticos versus matriz de transición. *Estadística Española*, Vol 46, No. 157, 389-407.
- Holland, J.H. (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbor: The University Michigan Press.
- INSEE (1998) *Les zonages: enjeux et méthodes*, INSEE Méthodes 83, Paris.
- ISTAT (1997) *I sistemi locali del lavoro 1991*, Rome, Istat.
- ISTAT-IRPET (1989) *I mercati locali del lavoro in Italia*, Milan, Angeli.
- Jin, Y. (2005) A Comprehensive Survey of Fitness Approximation in Evolutionary Computation. *Soft Computing* 9, 3-12.
- Khuri, S., Walters, T., Sugono, Y. (2000) A Grouping Genetic Algorithm for Coloring The Edges of Graphs. Proceedings of the 2000 ACM symposium on Applied computing - 2000.
- Laan, L van der and Schalke, R (2001) Reality versus Policy: The Delineation and Testing of Local Labour Market and Spatial Policy Areas, *European Planning Studies*, 9(2), 201-21.
- Lobo, F., Goldberg, D.E. (1997) Decision making in a hybrid genetic algorithm. Proceedings of the 1997 IEEE International Conference on Evolutionary Computation, 121-125.
- Lorena, L.A.N., Furtado, J.C. (2001) Constructive Genetic Algorithm for Clustering Problems. *Evolutionary Computation*, 9(3), 309-327.
- Michalaevicz, Z. (1994) *Genetic Algorithms + Data Structures = Evolution Programs*. Berlín: Springer-Verlag.
- Miller, B.L. and Goldberg, D.E. (1996) Genetic Algorithms, Selection Schemes, and the Varying Effects of Noise. *Evolutionary Computation*, 4(2), 113-131.
- Mitchell, M. (1996) *An Introduction of Genetic Algorithms*. Cambridge: MIT Press.
- Murthy, C.A., Chowdhury, N. (1996) In search of optimal clusters using genetic algorithms. *Pattern Recognition Letters*, 17, 825-832.
- Nowostawski, M., Poli, R. (1999) *Parallel Genetic Algorithm Taxonomy*. KES'99.
- OECD (2002) *Redefining Territories. The Functional Regions*, Paris, OECD Publications.
- Openshaw, S and Rao, L (1995) Algorithms for re-engineering 1991 Census geography *Environment & Planning A* 27, 425-446
- Papps, K. and Newell, J.O. (2002) Identifying Functional Labour Market Areas in New Zealand: A Reconnaissance Study Using Travel-to-Work Data. Discussion Paper n°. 443. Bonn, Institute for the Study of Labor (IZA) (<http://www.iza.org>). Previously published as Occasional paper 2001/6 by the Labour Market Policy Group, New Zealand Department of Labour (<http://www.mera.co.nz/projects/NZLLMA/index.htm>).
- Rudolph, G. (1994) Convergence analysis of canonical genetic algorithms. *IEEE Transactions Neural Networks*, Special Issue on Evolutional Computing, 5(1), 96-101.
- Sarkar, M., Yegnanarayana, B., Khemani, D. (1997). A clustering algorithm using an evolutionary programming-based approach. *Pattern Recognition Letters*, 18, 975-986.
- Sastry, K. (2001) Efficient Cluster Optimization Using Extended Compact Genetic Algorithm with Seeded Population. IlliGAL Report No. 2001018). Urbana, IL: University of Illinois at Urbana-Champaign.
- Sforzi, F (1987) L'identificazione spaziale in G Becattini (Ed) *Mercato e forze locali: il distretto industriale*, Bologna, Il Mulino, 143-167.
- Sforzi, F (1995) Sistemi locali di impresa e cambiamento industriale in Italia *Geotema* 2, 42-54.
- Sinha, A., Goldberg, D.E. (2001) Verification and Extension of the Theory of Global-Local Hybrids. IlliGAL Report No. 2001010. Urbana, IL: University of Illinois at Urbana-Champaign.
- Szpiro, G.G. (2002) Tinkering with Genetic Algorithms: Forecasting and Data Mining in Finance and Economics. In S.H. Chen (Ed.), *Evolutionary Computation in Economics and Finance*, Studies in Fuzziness and Soft Computing, Vol. 100, Heidelberg: Physica-Verlag, 273-285.
- Talbi, E.-G. (2002) A Taxonomy of Hybrid Metaheuristics. *Journal of Heuristics*, 8, 541-564.
- Tolbert, C.M. and Sizer Killian, M. (1996) US Commuting Zones and Labor Market Areas: 1990 Update, ERS Staff Paper AGES-9614, Rural Economy Division, Economic Research Service, US Department of Agriculture.
- Tomassini, M. (1995) A Survey of Genetic Algorithms. In: D. Stauer (Ed.), *Annual Reviews of Computational Physics*, World Scientific, Vol 3, 87-118.
- Watts, M. (2004) Local Labour Markets in New South Wales: Fact or Fiction?, Centre of Full Employment and Equity, WP 04-12, The University of Newcastle, Australia.