

## **Una Propuesta Metodológica para la Identificación y Evaluación de Clusters a partir de Tablas Input-Output. Una Aplicación para Andalucía**

ROBLES TEIGEIRO, LUIS(\*) Y RAMOS CARVAJAL, CARMEN(\*\*).

(\*) *Economía Aplicada. Política Económica. Universidad de Málaga. El Ejido, nº 6, 29071.* (\*\*) *Departamento de Economía Aplicada. Universidad de Oviedo. Avda. del Cristo s/n, 33071.*

Tfo. (\*)952131272; (\*\*)985105054 - E-mail: (\*)[teigeiro@uma.es](mailto:teigeiro@uma.es); (\*\*) [cramos@uniovi.es](mailto:cramos@uniovi.es)

### RESUMEN

En este trabajo se realiza una propuesta metodológica para la determinación de clusters sectoriales, la cual se basa en la aplicación de la Teoría de Grafos lo que facilita la distinción de los conglomerados existentes. Una vez que los cluster han sido determinados se procede a su análisis, aplicando un doble enfoque al considerar tanto el volumen como la densidad de los flujos establecidos entre los mismos, para lo cual se utilizan técnicas de extracción hipotética, así como medidas derivadas de la teoría estadística de la información. La propuesta anteriormente presentada se aplica a Andalucía, cuyos clusters productivos son determinados y analizados.

*Palabras Clave:* Cluster sectorial, Análisis Input-Output, Teoría de Grafos, Extracción hipotética, Economía andaluza.

### **A Methodological Proposal for Getting the Sectorial Clusters in Andalusia**

#### ABSTRACT

A suitable methodological proposal for getting sectorial clusters is determined in this paper. This proposal is based on the Graphs Theory, which as long as it simplifies the relations between the sectors, it lets us to take out the underlying economical structure and, therefore it makes to distinguish between the existing conglomerates lighter. Once clusters had been determined, an analysis will be done using a double approach, taking into account the quantity and the density of the flows established by the clusters, in order to do so it will be used hypothetical extraction techniques, but also measures which come from the Statistical Information Theory. This paper makes an empirical application about Andalusia in order to illustrate the shown proposal.

*Keywords:* Sectorial cluster, Input-Output Analysis, Graphs theory, Hipotetical Extraction Method, Andalusian Economy.

Clasificación JEL: C67, O18, R15.

Artículo recibido en julio de 2007 y aceptado para su publicación en noviembre de 2007.

Artículo disponible en versión electrónica en la página [www.revista-eea.net](http://www.revista-eea.net), ref.: e-25313.

## 1. INTRODUCCIÓN

Muchos autores han destacado la importancia de las relaciones entre los sectores de actividad, de los denominados *linkages*, para el crecimiento económico. Es por ello que el término *linkage* aparece en la definición de dos de los conceptos de mayor tradición en el análisis Input-Output, el sector clave y el *cluster*. En efecto, puede afirmarse que un sector clave es aquel que posee destacadas interrelaciones con el resto de las ramas tanto hacia adelante como hacia atrás, mientras que un *cluster*, en nuestro ámbito de análisis, no es sino un grupo de actividades con gran interdependencia entre si. De esta manera, ambos conceptos, siendo distintos, se encuentran estrechamente ligados.

Este trabajo centra su atención en los *clusters*, siendo su interés principal analizar el *estado de la cuestión* acerca de la identificación de aglomerados o *clusters* a partir de tablas Input-Output. Es por ello que realiza un considerable recorrido sobre los dos grupos de técnicas empleadas, unas a partir del análisis multivariante y otras basadas en la teoría de grafos o afines.

De la aplicación práctica realizada se llega, sin embargo, a la conclusión de que los métodos multivariantes cuentan con diversas y graves limitaciones, pudiendo criticarse o limitarse su aplicación, aunque estos procedimientos sean, no se olvide, el paradigma en la obtención de *clusters*. De esta manera, y aún aceptando que se han obtenido varios *cluster* muy similares con ambas técnicas, nos inclinamos por el uso de métodos de grafos, elaborando una nueva propuesta que denominamos de *triple restricción*.

Finalmente, se analiza la importancia de los *clusters* obtenidos mediante la aplicación de una doble óptica: la consideración tanto de la relevancia de los flujos establecidos como de la densidad de los mismos. El primer aspecto será determinado a partir de una de las técnicas clásicas en la metodología Input-Output como es la extracción hipotética y el segundo mediante indicadores derivados de la teoría estadística de la información.

## 2. CLUSTERS: ALGUNAS CONSIDERACIONES CONCEPTUALES

Los *pôle de développement* introducidos por Perroux en 1955 han recibido periódicas muestras de interés en los pasados años pero, como ocurre con casi todos los desarrollos teóricos, es difícil encontrar al creador del ahora denominado “*cluster*” ya que es posible remontarse, al menos, hasta Marshall. De hecho, el recorrido del concepto ya ha sido muy largo aunque su denominación haya ido variando a lo largo del tiempo de acuerdo a las preferencias de los distintos autores: *filière, pôle, agglomeration, complex...* En la actualidad el estudio de los *clusters* ha vuelto a ponerse de moda gracias a los trabajos de Porter, por una parte, y a los miembros de la *New Economic Geography*, como Krugman, por otra. Éste último, ha explicado la aparición

de *cluster* por la obtención de economías de escala, por el abaratamiento de los costes de transporte y de transacción en general y por la mayor movilidad de los factores de producción que se logra en su seno. Porter, por su parte, señala que la aglomeración ahorra costes de infraestructura y comunicaciones, facilita la difusión tecnológica y permite, por último, mejor acceso a los inputs y más fácil salida del output a los mercados. En general, como se ve, estos autores y otros ofrecen razones que obedecen tanto a causas de oferta, de producción, como de demanda, obteniéndose mejoras en ambos sentidos de manera simultánea sin que sea fácil su distinción<sup>1</sup>.

Hay también un amplio consenso sobre la positiva aportación de los *clusters* al crecimiento económico toda vez que se da por supuesto que, entre otras cosas, generan ventajas comparativas, mejoran la competencia e incrementan la productividad. Sin embargo, también hay quién opina que algunas de estas ventajas están por concretar dada que muchas de ellas son de difícil identificación y medición, así como tampoco falta quién opina que pueden facilitar la cartelización y la formación de oligopolios, además de otros inconvenientes de diversa índole.

Existe, por otra parte, un amplio y antiguo debate sobre el propio concepto de *cluster*. En las primeras acepciones, se le definió como una concentración o aglomeración geográfica de varios establecimientos que adquirirían ventajas comparativas con la colocalización y mutua interrelación. De esta manera, se requerían dos componentes:

**Proximidad.** Los establecimientos o empresas que componían un *cluster* se habían de encontrar geográficamente próximos aunque esta proximidad podía no querer decir vecindad real.

**Interdependencia.** El *cluster* requeriría, para ser algo más que una mera concentración de industrias, de mutua interrelación funcional, ya sea productiva ya de otra índole significativa. Debían existir flujos de transmisión de tecnología o conocimientos u otras formas de colaboración de carácter permanente. En ocasiones, esta cooperación podría incluso instrumentarse mediante instituciones de diversa índole: administrativas, educativas, de investigación... El *cluster* sería así un sistema complejo con una dinámica propia y no una mera yuxtaposición de empresas. Y, *localización e interrelación* serían sus elementos básicos, los que lo explican y definen. Sin embargo, y como se verá, no todos los autores los han podido o querido reunir de forma simultánea en sus trabajos, lo que ha dado lugar a una larga polémica.

Así, por ejemplo, en uno de los trabajos pioneros en este campo, Roepke et al (1974, p. 15) diferenciaban expresamente entre dos tipos de complejos: primero, un grupo de industrias muy interrelacionadas y, segundo, un grupo de empresas también interdependientes pero que se reunirían además físicamente en un espacio común, en un centro industrial o en vecindad dentro de una región. El primer tipo de complejos, añadían, podían estudiarse con unas tablas Input-Output (TIO) a través del análisis

---

<sup>1</sup> Una adecuada revisión de la literatura sobre cluster se encuentra en Hoen (2001,II).

factorial, por triangulación o a través de grafos. Sin embargo, los segundos complejos, los territoriales, definidos por Isard, Smolensky (1963) o Kolosovsky (1965), requerirían de instrumentos propios de la Economía Regional o de la Geografía Económica para su identificación y no serían objetivo de su particular análisis. Czamanski y Ablas (1979), por su parte, distinguieron expresamente entre *clusters* y complejos, definiendo **cluster** (p. 62) –su campo de estudio– como un conjunto de industrias –ramas– interconectadas por unos flujos de bienes y servicios más fuertes que los que también poseen con el resto de la economía. Su definición, añadían, estaría desprovista de cualquier connotación espacial. Por su parte, **complejo**, sería un *cluster* que posee además una similitud en sus pautas de localización. Por tanto, en el término complejo sí se enfatizaría en el aspecto espacial.

En cualquier caso, no puede extrañar que la irrupción de quienes comenzaron a emplear tablas Input-Output y técnicas factoriales fuera seguida de diversas críticas por parte de investigadores del campo regional (Harrigan, 1982, Lathan,<sup>2</sup> 1976, 1977), reivindicando la imposibilidad de estudiar *clusters* sin tener muy presentes los elementos espaciales. Sin embargo, en buena medida, el debate se produjo por una falta de comprensión hacia la distinta naturaleza de los estudios que se realizaban toda vez que el concepto de *cluster* del que se partía o al que se quería llegar era diferente para unos y otros.

Volviendo al concepto de *cluster*, Hoen (2001, I, p. 2) distingue desde un punto de vista teórico entre dos tipos atendiendo a su dimensión, *microcluster* si se refieren a empresas o establecimientos y *mesoclusters* si se analizan a escala sectorial. Acepta este autor que, hipotéticamente, sería muy interesante el análisis de los *microclusters* pero las fuentes estadísticas disponibles, añade, sólo permiten el estudio de los *mesoclusters*.

Añadiendo este factor, al que cabe denominar *escala*, a los ya citados, *localización e interrelación*, serían tres ya los elementos de los *clusters* a los que, necesariamente, deberían referirse los autores en aras a la claridad expositiva. Pero, por último, lentamente la literatura ha llegado a acuñar una distinción adicional entre *clusters* que nos parece, como se verá, del mayor interés: *cluster horizontales* y *clusters verticales*; los primeros comparten el mercado para sus productos y requieren similares inputs intermedios y primarios combinados también con parecida

---

<sup>2</sup> Lathan (1976, 1977) insistió también en el sentido de la excesiva complejidad de los métodos econométricos frente a la obviedad de sus resultados –“Las ramas que componen cada factor son aquellas que, de una u otra manera, uno siempre esperaría que se asociaran” (p. 47)- pero su trabajo *Needless Complexity in the Identification on Industrial Complexes*, fue bien replicado por Czamanski (1977): “Después de todo el análisis de componentes principales, al nivel con que se usa en estos estudios, se enseña habitualmente en los cursos introductorios de estadística” (p. 455).

tecnología. Los verticales, por su parte, unen establecimientos que se suministran inputs de manera sucesiva y encadenada, conformando cadenas de valor añadido (Rosenfeld 1997).

### 3. MÉTODO DE IDENTIFICACIÓN DE LOS *CLUSTERS*

El segundo de los problemas que siempre está presente en el estudio de los *clusters* es el método más adecuado para su identificación ya que se carece de uno suficientemente aceptado y estandarizado. Por una parte, este problema es lógico si se tiene en cuenta que su propio concepto ha sido objeto de permanente discusión. Pero por otra, y aún quedando establecido un concepto determinado, es posible discrepar en el método a utilizar toda vez que ello puede depender, en gran medida, del objetivo a alcanzar por parte del investigador.

Se han utilizado métodos muy diversos para la identificación de los *clusters*. En ausencia de información, hay quién ha debido recurrir a su experiencia y a los conocimientos propios o ajenos –opinión de expertos-. Pero, independientemente de estos métodos primeros, quizás pueda hablarse de dos vías o procedimientos cuantitativos para el establecimiento de los *clusters*. Uno primero que utiliza análisis multivariante que pueden ser diversos, pero generalmente derivados del análisis factorial y uno segundo basado, con mayor o menor intención, en el **análisis por medio de grafos y métodos afines**. Estos procedimientos utilizan además habitualmente como instrumento estadístico esencial unas tablas Input-Output. Los procedimientos cualitativos, además, se han empleado preferentemente para la identificación de *miniclusters*, con los elementos que hemos denominado localización e interrelación casi siempre presentes, mientras que los métodos cuantitativos, por su parte, han pretendido el estudio de *megaclusters*, y sin que el elemento de localización estuviese, necesariamente, presente. Por último, hay que tener presente que las técnicas multivariante – análisis factorial, *cluster* jerárquico... - son útiles para detectar, en principio, las relaciones “horizontales” entre las ramas, la semejanza de sus perfiles productivos y que, por consiguiente, pueden servir para identificar preferentemente lo que hemos denominado *clusters* horizontales. Por su parte, el análisis de grafos relaciona verticalmente a las ramas, poniendo de manifiesto relaciones de tipo comprador-vendedor y, por tanto, puede utilizarse para identificar *clusters* verticales.

Insistiendo, si se utilizan unas TIOs y métodos multivariantes se obtendrán, en principio, *megaclusters horizontales* que mostrarán ramas de similares perfiles productivos y sin que se tenga en cuenta el elemento espacial o de localización de los *clusters*. Esto no evita, como es natural, que la matriz original sobre la que se trabaje pueda transformarse en la medida suficiente para que revele relaciones verticales o que, una vez establecidos los *clusters* pueda estudiarse su localización y conocer su grado de vecindad y aglomeración, pero este no sería el resultado normal o general. Por

su parte, TIOs y análisis grafos conducen a *megacusters* verticales sin que tampoco esté presente, en principio, el elemento localización. Claramente está el elemento de interrelación, coincidiendo además con lo que la literatura Input-Output denomina *linkages*, relaciones de compra venta entre las ramas. De nuevo, nada evita que, a posteriori, pueda analizarse la ubicación de los *clusters*.

Es posible señalar por último, que aunque no cabe duda de que la identificación de *megacusters* compuestos por ramas es útil por diversos motivos: para facilitar estudios sectoriales, para posibilitar la comparación internacional, para apoyar la toma de decisiones de la política económica... conviene ser conscientes de que también cuenta con diversas limitaciones. Así, el análisis Input-Output no permite revelar el tipo de interrelación entre los establecimientos del *cluster* ni las formas de colaboración establecidas más allá de las estrictamente productivas (Doeringer y Terkla 1995, p. 228). Como tampoco, ya se ha dicho, localiza a los *cluster* en el espacio.

### 3.1. Métodos Multivariantes

Los procedimientos factoriales son hoy dominantes entre quienes identifican *clusters* por medio de tablas IO. Sin embargo, y aceptando su utilidad, hay que insistir en que tienen todo su sentido en la determinación de lo que antes hemos denominado *clusters* horizontales, siendo su uso más cuestionable si lo que se pretende es observar y analizar relaciones verticales. Esto ha sido reconocido, con mayor o menor claridad o con mayor o menor énfasis por gran número de autores<sup>3</sup> aunque aún queden equívocos y reticencias. Dos extensos documentos de reciente publicación realizan una excelente síntesis de la evolución y situación de las técnicas IO en relación a

---

3 Oosterhaven et al (2001) afirman que el método estándar para la identificación de clusters –el análisis factorial o de componentes principales– es inútil para medir los linkages toda vez que: “Usa la semejanza de sus compras y ventas intermedias para agrupar a las industrias en un cluster... Este método, por supuesto, responde a interesantes cuestiones... Sin embargo, no contesta acerca de cuáles están más estrechamente vinculadas a otras” (p. 812). O’HUallachain (1984) también mantuvo una opinión muy clara: “La evidencia empírica confirma las críticas recientes a los componentes principales como un medio para detectar relaciones verticales, pero demuestra su eficacia para la identificación de relaciones complementarias (complementary relationships). Las agrupaciones obtenidas no son de grupos de empresas relacionadas verticalmente sino de empresas con perfiles similares en requerimientos de inputs y outputs” (p. 421). Por último, incluso en trabajos pioneros y de enorme influencia como los de Czamanski (1977) es ya posible encontrar referencias a que el método de AF agrupa a las ramas sobre la base de la semejanza de sus perfiles de flujos: “En los estudios basados en análisis multivariante (componentes o factores) el criterio fue incluso más general ya que las industrias fueron clasificadas como miembros de una agrupación sobre la base de la semejanza de sus perfiles de flujos con las otras ramas de la economía” (p. 455).

la estimación de *clusters* por procedimientos factoriales y es, por tanto, ciertamente cómodo remitir al interesado allí ya que son de fácil obtención a través de la red<sup>4</sup>.

La aplicación del análisis factorial (AF) a las TIOs ha requerido habitualmente de los siguientes pasos o tareas sucesivas:

1. Selección y agregación de las ramas de las TIOs. En efecto, y este es un problema general y no sólo del AF a la hora de elaborar clusters, a muchos autores les ha parecido necesaria la selección de las ramas a estudiar, eliminando o agregando aquellas de carácter general (comunicaciones, transportes, comercio ...) que aparecerían una vez tras otras en todos los clusters. Ramas que poseen relaciones con todo el resto pero con ninguna otra de manera especial<sup>5</sup>. Esto ha llevado, además, a primar los estudios de ramas manufactureras sobre las de servicios, lo cual es lógico si se piensa que la variedad de inputs, la complejidad tecnológica y los flujos exteriores son mucho más importantes en las manufacturas. Pero esta elección ha conllevado también el marginar o no tener en absoluto en cuenta buena parte de la economía de las regiones o países estudiados. De esta manera puede decirse, sin ambages, que las técnicas multivariantes son, hasta el momento, técnicas para el análisis de la industria manufacturera.

2. La estandarización o normalización de la matriz, método que pretende corregir el problema de la escala o diferente tamaño de las ramas. Para ello, se suele elaborar una matriz de coeficientes o bien una de ligazones. Se entiende por coeficiente

técnico a la ratio  $a_{ij} = \frac{z_{ij}}{X_j}$  y por coeficiente de distribución,  $b_{ij} = \frac{z_{ij}}{X_i}$  donde  $X_j$  y  $X_i$

son, respectivamente, la producción efectiva de la rama  $j$  y el producto distribuido de la rama  $i$ . Por otra parte,  $z_{ij}$  es el uso que la rama  $j$  hace de inputs intermedios de la rama  $i$ . La ligazón de compras o por columnas es similar al coeficiente técnico:

$$a_{ij}^* = \frac{z_{ij}}{\sum_{j=1}^n z_{ij}}, \text{ mientras que la ligazón de ventas queda definida: } b_{ij}^* = \frac{z_{ij}}{\sum_{i=1}^n z_{ij}}.$$

3. La construcción de una matriz input donde se aplicará la técnica factorial. Para este fin, suele realizarse una matriz de correlaciones que, en el caso más común Czamanski (1971) y Bermang y Fesser (2000), es simétrica y se formará eligiendo, en cada caso, el valor máximo de las cuatro matrices de correlaciones -C- realizadas a partir de las dos matrices de ligazones tal que:  $c_{ij} = c_{ji} = \max [ C(a^*_{.i}, a^*_{.j}), C(b^*_{.i}, b^*_{.j}), C(a^*_{.j}, b^*_{.j}), C(b^*_{.j}, a^*_{.j}) ]$ . En otras ocasiones -Roepke et al (1974)- esta

4 Estos son Rey y Mattheis (2000): Identifying Regional Industrial Clusters in California y Lainesse y Poussart (2005) Méthode de repérage des filières industrielles sur le territoire québécois basée sur les tableaux d'entrées-sorties.

5 Puede verse sobre este particular de la exclusión de ramas Lainesse y Poussart (2005), capítulo 6.

matriz input se ha realizado a partir de la mera suma de los valores absolutos de la matriz intermedia  $z^*_{ij} = z^*_{ji} = z_{ij} + z_{ji}$ . O bien -Bergman et al (1996)- se ha utilizado una matriz efectuada por los valores máximos de las cuatro ligazones que unen a dos ramas –o eligiendo el valor máximo de los coeficientes de Streit-:  $c_{ij} = c_{ji} = \max [a^*_{ij}, a^*_{j,i}, b^*_{i,j}, b^*_{j,i}]$ .

Con la construcción de este tipo de matrices *inputs* se consiguen, también, varios objetivos que no encontramos declarados o especificados en los trabajos revisados. En primer lugar se palía el problema de la cantidad de ceros y valores muy pequeños que tiene cualquier tabla IO y que hace muy difícil –y lo hemos corroborado con nuestra experiencia personal- aplicar sin más técnicas factoriales a las matrices de coeficientes o ligazones. De hecho, y como prueba de cuanto decimos, en un solo caso -Roepke et al (1974)- se afirma haber conseguido extraer factores sin utilizar una matriz *input* del tipo descrito (aunque ellos también agregan en uno de sus ensayos como acaba de indicarse en este mismo punto).

En segundo lugar, y como se habrá observado, la utilización de esta matriz input lleva, de hecho, a abandonar la búsqueda de clusters exclusivamente horizontales al combinarse relaciones de filas y columnas indistintamente. Además de ello, la casi segura posibilidad de encontrarse con agrupaciones de ramas que, si bien pueden llegar a ser semejantes en su estructuras productivas- industrias extractivas con servicios destinados a la venta, por ejemplo-, no son fáciles de asociar o aceptar en un cluster, queda adecuadamente superada.

Por último, no puede extrañar que la construcción de esta matriz haya sido siempre polémica ya que, en efecto, el gran problema es que las transformaciones propuestas son de tal magnitud que es difícil de encontrar sentido económico en estas matrices “inputs” reelaboradas. Esta opinión fue apuntada, otra vez, hace ya muchos años -Roepke et al (1974)-: “Czamanski comienza con matrices de coeficientes de correlación que relacionan las industrias a través de ligazones... sin embargo, la consecuencia es que esta matriz no está basada en relaciones claramente identificables, resultando difícil reconocer clusters coherentes” (p. 26).

4. Extracción de los factores, rotación –generalmente varimax- y asignación de las ramas a los *clusters*. De manera habitual, los programas estadísticos al uso presentan los resultados obtenidos en una lista de factores extraídos dispuestos de forma decreciente en función del porcentaje de varianza explicado. Es también común en trabajos ya publicados –Feser y Bergman (2000)- que un número reducido de componentes o factores expliquen un alto porcentaje de la varianza. Cada rama de actividad posee una correlación con cada grupo extraído lo que permite su asignación, dado el caso, al *cluster* con el que tenga mayor vinculación pero también, de manera simultánea, con algún otro con quien tenga una relación destacada. Suele establecerse una relación mínima para poder efectuar estas vinculaciones con las debidas garantías.

### 3.2. Métodos de Grafos y afines

Junto a trabajos fácilmente reconocibles dentro del ámbito de la teoría de grafos [Campbell (1975), Slater (1977), Morillas (1983), Bon (1989), y Aroche (2001)], otros autores han utilizado técnicas que podríamos considerar afines a los mismos. Distingamos brevemente algunas propuestas.

Aunque no siempre se indica expresamente puede considerarse que los métodos **máximo o máximos** son una derivación del análisis grafos. El método del máximo sólo tiene en cuenta la relación más importante de cada rama. Para su elaboración, se suele partir de la matriz **Z** de relaciones intermedias localizando su mayor valor -dejando aparte la diagonal o autoconsumos- y vinculando a las dos ramas responsables de él. Estas dos ramas conformarán el primero de los *clusters* provisionales. Se busca el siguiente máximo, se vinculan las ramas... y así sucesivamente... hasta que se determine un número de *cluster* que se juzgue razonable o que ya estuviese prefijado de una manera exógena. El procedimiento cuenta con la ventaja de su sencillez, de ser intuitivo, pero con el inconveniente principal de que puede perder mucha información sobre las relaciones secundarias entre las ramas, que pueden tener también notable importancia. Un segundo inconveniente puede venir cuando el máximo lo es en una matriz determinada, por ejemplo en la **Z** de transacciones absolutas, pero no lo es en otras, como en la matriz **A** de coeficientes técnicos, o en la **B** de coeficientes de distribución. Surge entonces la duda de qué máximos elegir o sobre qué matriz trabajar. Para evitar este problema, se han hecho diferentes propuestas todas ellas caracterizadas por el añadido de restricciones o filtros que garanticen que individual -en una matriz- o colectivamente -en varias simultáneamente- el máximo sea significativo. Se habla entonces de una técnica de *máximos con restricciones*.

Peeters et al (2000) en un trabajo para la OCDE proponen un procedimiento de **máximos con restricciones** que puede resumirse en los términos que siguen: se elaboran *clusters* hacia atrás -*backwards chains*, BC-, hacia adelante -*forward chains*, FC- y *megaclusters* o suma de los anteriores. Los *clusters* se basan en las matrices de ligazones ya definidas. La elaboración de cada uno de los *clusters* exige, a su vez, de dos fases sucesivas. Si se empieza por los BC, se filtrará en primer lugar la matriz de ligazones de compras de manera que todos sus elementos sean mayores a un filtro prefijado. A continuación, se establecen por columnas los máximos de cada rama transformando la matriz de ligazones original en otra binaria donde los máximos tendrán el valor 1, mientras que el resto de los elementos serán 0. Esta matriz, ya sería útil para elaborar *cluster* hacia atrás, pero de cara a garantizar que los máximos tengan la suficiente entidad se impone una segunda condición en la segunda etapa. En efecto, se propone ahora comprobar que la ligazón horizontal correspondiente a cada uno de los máximos ya elegidos sea también superior a un determinado filtro. Esta segunda condición no supondrá más que la eliminación o no consideración de algunos de los máximos anteriores. La elaboración de los FC es similar a la descrita,

si bien los filtros pueden variar ya que, habitualmente, las ligazones y coeficientes horizontales tienen un valor medio mayor –los autores proponen 0,20 y 0,05 respectivamente-. Los *megaclusters* serán por último, y sin más, la suma de los dos *clusters* anteriores.

Los métodos de **triangularización** y **descomposición** tienen su antecedente en los trabajos de Simpon and Tsukui (1965) sobre triangulación de la tabla y determinación de la “estructura fundamental”. Se trata de dividir la matriz intermedia  $\mathbf{Z}$  en grupos de filas y columnas que no tengan relación, o que sea muy débil, con el resto. Para ello, los flujos  $z_{ij}$  que no superen un cierto filtro se transforman en ceros, reordenando y denominando posteriormente los bloques resultantes. Esta reordenación suele hacerse reubicando las ramas de manera que los *clusters* se sitúen bien en un ángulo de la matriz, bien en el entorno de la diagonal, proceso lógico toda vez que, ya de antemano, en la tabla IO las ramas más afines intercambian precisamente allí. El método de diagonalización puede considerarse un tipo especial de triangulación, pero también, tiene evidentes conexiones con los métodos de máximos con restricciones.

Hoen<sup>6</sup> (2002, p. 136) ha desarrollado un proceso de descomposición muy sencillo de acuerdo a los siguientes pasos: se elige un nivel de significación  $\alpha$ . Se seleccionan los valores que pertenecen al  $\alpha\%$  de los mayores valores de  $\mathbf{Z}$  y se transforman el resto de valores en 0. Es decir,

$$\begin{cases} z_{ij} & \text{si y sólo si } z_{ij} \geq k, \forall k = x_{1-\alpha} \\ 0 & \text{en otro caso} \end{cases}$$

donde  $x_{1-\alpha}$  representa el centil  $1-\alpha$ .

Se reubican, por último, las ramas de forma que compongan bloques o *clusters*. El autor ve claras ventajas en su técnica (Hoen 2000, p. 6): “Puesto que los flujos intermedios que no cumplen la restricción son transformados en cero, los correspondientes coeficientes técnicos, de distribución y de la inversa de Leontief (en gran medida) también son transformados en cero. Así, todas las matrices son diagonalizables y, por ende, el método lleva a la obtención de los mismos *clusters* sin importar los datos utilizados de partida (p. 138, 139)”. Además de ello, su método no suministra enormes *clusters*, como tampoco *clusters* muy pequeños, como sí ocurre con otras técnicas. En tercer lugar, si bien el número de *clusters* puede variar en función del nivel del filtro establecido, no puede ocurrir que una rama cambie de *cluster* por ello. Si se baja el nivel, pueden aparecer los mismos o más *clusters* de tamaño inferior, y si se sube se mantendrán los mismos *clusters* o se agregarán algunos de ellos, pero en ningún caso ocurrirá el denominado *cluster*

<sup>6</sup> Hoen (2002) ofrece un algoritmo para realizar la reordenación en el apéndice de su trabajo.

*switching*. Por último, el método le permite obtener *cluster* similares a lo largo del tiempo –al menos en Holanda–, resultado que no se alcanza con otras técnicas y que es más congruente con la presunción de que los *clusters* son estables en el tiempo.

Sin embargo, en nuestra opinión, puede apuntarse que el método es bastante simple y drástico ya que es operativo y cuenta con las ventajas apuntadas sólo si se utiliza un filtro muy exigente, lo que conduce a una notable pérdida de información. De esta forma, y aunque el autor no indica el filtro elegido, calculamos replicando su trabajo sobre la misma TIO de Holanda, que éste debe estar próximo al 1%, porcentaje muy restrictivo como se ve. La elección de un filtro menos exigente es posible, pero entonces el número de *clusters* se multiplica así como el tamaño medio de los mismos, contradiciendo varias de las ventajas expuestas.

#### 4. UNA PROPUESTA ALTERNATIVA: MÉTODO DE TRIPLE RESTRICCIÓN

De la lectura del apartado anterior, de la réplica de los trabajos citados y de la reflexión y propia experiencia, se llega a la conclusión de que pueden existir diversos procedimientos para la elaboración de *cluster* a través de grafos manteniendo un mayor nivel de exigencia que lo requerido por el método del máximo, el de Peeters et al o del de Hoen ya citados. Nosotros, en particular, hemos elaborado uno que parte de dos decisiones previas. La primera consiste en trabajar con ligazones y no con coeficientes. En realidad, la elección no es muy importante dada la enorme semejanza entre ambos pero, en cualquier caso, las ligazones se ven menos influenciadas por los cambios de valoración de las tablas y menos alterados también por la presencia del valor añadido, la demanda final en el denominador de los coeficientes. En segundo lugar, se ha pretendido que las ligazones a considerar tuviesen siempre el adecuado respaldo de números absolutos en la matriz intermedia  $Z$ , números de suficiente importancia. En efecto, al dividir las transacciones interindustriales sobre el total de consumos intermedios –concepto de ligazones– puede ocurrir que el cociente sea muy elevado pero que, al mismo tiempo, el numerador y, sobre todo, el denominador sean irrelevantes, lo cual debería evitarse. Para salvar este problema, basta con filtrar toda la matriz intermedia con un nivel suficientemente alto como para eliminar los números pequeños pero respetando, al tiempo, los restantes. De hecho, a través de este procedimiento se eliminan muchos elementos, porque son muchos también los pequeños números que contiene una tabla, pero este efecto es intrascendente para los objetivos que se pretenden, garantizando sin embargo el buscado respaldo a las ligazones. Finalmente, el método, que toma ideas de los trabajos de Bergman y Fesser es un método de *triple restricción* y puede describirse como sigue:

La primera etapa del método propuesto consiste en construir una matriz de adyacencia asociada al dígrafo<sup>7</sup>, esto es una matriz cuadrada y binaria, cuyos elementos no nulos se determinarán de forma que el arco establecido entre cada par de supere el 10% de su promedio. Se trata, por tanto de una matriz de adyacencia construida de una manera más restrictiva que las habitualmente definidas, en las que sólo es necesario que esté garantizada la existencia de un arco sea cual sea su valor. La matriz de adyacencia, denotada como  $\mathbf{D}=\{d_{ij}\}$ , así obtenida está constituida por elementos que tienen la siguiente forma

$$d_{ij} = \begin{cases} 1 & \text{si y solo si } z_{ij} \geq 0.1 \frac{\sum_i \sum_j z_{ij}}{n} \\ 0 & \text{en otro caso} \end{cases}$$

Una vez definida y determinada esta matriz, se calculará a partir de ella las correspondientes ligazones, las cuales, en términos de la teoría de grafos, pueden considerarse como proporciones de los semigrados interiores y exteriores. El semigrado exterior asociado al sector  $i$ -ésimo  $d^+(S_i) = \sum_j d_{ij}$  es una medida de integración en compras, por lo tanto, la ligazón en compras anteriormente señalada puede escribirse como  $a_{ij}^* = \frac{d_{ij}}{d^+(S_i)}$ . Esto es, lo que aporta cada celda al total de relaciones de compra establecida. Análogamente, la medida de integración en ventas, semigrado interior, puede expresarse como  $d^-(S_i) = \sum_j d_{ji}$  y, en este caso, la ligazón correspondiente será  $b_{ij}^* = \frac{d_{ij}}{d^-(S_i)}$ , lo que aporta la celda al total de relaciones en ventas. Con estas ligazones, se establece un nuevo filtro tanto en compras como en ventas, con lo cual se habrá dado lugar a una nueva matriz de adyacencia  $\mathbf{D}^1=\{d_{ij}^1\}$  de la siguiente manera:

$$d_{ij}^1 = \begin{cases} 1 & \text{si y solo si } a_{ij}^* \text{ y } b_{ij}^* \geq 0.1 \\ 0 & \text{en otro caso} \end{cases}$$

7 Un grafo orientado es el par  $G=(V,U)$  donde  $V$  es un conjunto finito de vértices, nodos o polos y  $U$  representa un conjunto de arcos que pueden establecerse entre los nodos, una TIO puede considerarse como un grafo orientado o dígrafo, donde los elementos de  $V$  son los sectores mientras que los arcos podrían interpretarse como los flujos establecidos entre ellos, siendo relevante la orientación de los mismos.

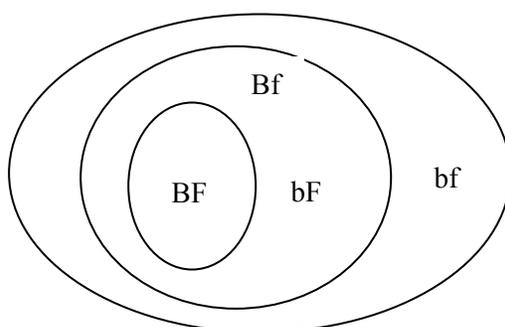
Realizado este filtrado, se clasifican las actividades según su grado de conectividad distinguiendo las que están fuertemente conexas, de las que presentan una conectividad más débil, tanto en compras como en ventas. Así podemos establecer la siguiente tipificación en la que se es preciso distinguir tanto la orientación de los flujos (compras o ventas) y mayor o menor nivel de conexión (mayúscula o minúscula, respectivamente). La clasificación establecida aparece recogida en el cuadro nº 1:

**Cuadro Nº 1: Clasificación de actividades según su grado de conectividad**

Clasificación	$a_{ij}^* > 0.2$	$0.1 \leq a_{ij}^* \leq 0.2$
$b_{ij}^* > 0.2$	BF	bF
$0.1 \leq b_{ij}^* \leq 0.2$	Bf	bf

Es decir, las actividades más fuertemente conexas en compras y ventas se denotarán con BF, las fuertemente conexas en compras y más débilmente conectadas en ventas con Bf y así sucesivamente. A partir de lo anteriormente expuesto podríamos representar el núcleo del *cluster* de acuerdo al gráfico nº 1:

**Gráfico Nº 1: Representación de las actividades según su grado de conectividad**



El procedimiento así recogido ha proporcionado clusters consistentes, pero ello no impide reconocer que la elección de cualquier otro filtro menos severo llevaría a la inclusión de nuevas ramas en los clusters como, incluso, a la aparición de algún otro agrupamiento nuevo. Por ello, y de nuevo en relación al gráfico adjunto, quizás se deberían concebir los clusters a modo de círculos concéntricos con un núcleo indiscutible que puede elaborarse, por ejemplo, con un método de máximo garantizado o respaldado como el propuesto pero sin que por ello se rechace la posibilidad de, a partir de éste, poder establecer otros agrupamientos que partirían o se desprenderían del núcleo central. Este planteamiento quedaría justo en las antípodas de lo propuesto por Hoen: clusters elementales, indiscutibles y estables en el tiempo; pero nos parece un planteamiento mucho más realista, mucho más acorde con la siempre compleja realidad.

## 5. IDENTIFICACIÓN DE LOS *CLUSTERS* EN LAS TIOAN2000

Expuestos los métodos de obtención *clusters*, en este apartado se procede a su aplicación a las últimas tablas Input-Output realizadas en Andalucía correspondientes al ejercicio 2000. Andalucía es una de las regiones españolas con mayor tradición en la elaboración de tablas Input-Output, ya que se dispone de ellas en intervalos regulares desde el año 1975. Todas estas tablas se han estimado por procedimientos directos, confeccionándose bien por universidades andaluzas bien por el Instituto de Estadística de Andalucía en colaboración asimismo con la universidad. Andalucía es una región relativamente grande -87.597 km<sup>2</sup> -, con unos ocho millones de habitantes (Instituto de Estadística de Andalucía) y una renta per capita que se estima en el 77% de la media europea (UE-27), según el INE.

### 5.1. Obtención de *clusters* por técnicas multivariantes

Para la estimación de *clusters* bajo técnicas multivariantes hemos replicado y utilizado la práctica totalidad de las técnicas recogidas en la literatura que se ha citado haciéndolo, además, no sólo con la tabla de Andalucía correspondiente al año 2000, sino también con tablas anteriores (1995 y 1990) e, incluso, con las tablas correspondientes a otros ámbitos (tablas españolas y tablas de otras regiones españolas). Se trataba de ensayar y de adquirir una cierta experiencia que permitiera opinar con el fundamento debido.

En líneas generales, no se han logrado *clusters* de interés cuando se ha trabajado directamente con coeficientes o ligazones. Y es que un mero análisis de correlaciones entre filas y columnas de la tabla muestra que son muy pocas las ramas que muestren perfiles semejantes siendo, por el contrario, norma común la diversidad. Pero tampoco el uso de la técnica estándar, Czamanski-Bergman y Fesser, ha proporcionado resultados de interés, por lo que nos hemos visto obligados a utilizar otra matriz input que acumulase efectos hacia adelante y atrás y que presentara coeficientes de correlación más elevados y pertinentes. Finalmente, los mejores resultados se alcanzaron mediante una matriz suma de ligazones de ventas y compras de cada rama, de manera que  $\mathbf{M} = \mathbf{A}^* + \mathbf{B}^*$ , donde  $\mathbf{A}^* = \{a_{ij}^*\}$  y  $\mathbf{B}^* = \{b_{ij}^*\}$ . Este método no es totalmente original ya que antes algunos autores han sumado las matrices de coeficientes técnicos y de distribución. Y también, aceptamos que la mera suma de ligazones o coeficientes sigue siendo discutible, pero en cualquier caso nos parece más sensata que la elección de coeficientes de correlación máximos entre filas y columnas de las ramas. La suma de ligazones, acumula las principales relaciones de filas y columnas de las ramas con lo que, al menos, es una matriz comprensible y con significado económico.

El uso de esta matriz ha permitido observar que determinados *clusters* o asociaciones industriales se repiten en el análisis de cualquier país o región desarrollada siempre que cuente con el suficiente tamaño o entidad. Y así, la experiencia muestra que, de manera

habitual, se obtiene un *cluster* agroalimentario, caracterizado por su relativa independencia con el resto de *clusters* y ramas junto a otros tres, energía, construcción y metal-mecánico. Estos dos últimos, según los casos, pueden asociarse en un solo *cluster*. En también posible encontrar junto a estos bloques descritos otro formado por los grandes servicios intermedios, así como un indeterminado número de ramas asociadas en parejas o tríos, *miniclusters*, tales como textil y confección, saneamiento público y privado, etc. Pero desde luego si hay que destacar dos *clusters* que repetidamente aparecen son el agroalimentario y el de construcción, ambos formados por ramas primarias en el sentido de Colin Clark, con estructuras de costes relativamente sencillas y, sobre todo, semejantes.

El cuadro nº 2 recoge las principales cargas de los componentes que sindicán a las ramas, pero varios son los hechos a considerar sobre la aplicación realizada:

- Se han eliminado aquellas ramas que suministran a un gran número de ramas pero no mantienen fuertes vínculos con ninguna en particular, dificultando una lectura acertada de los vínculos sectoriales: electricidad, comercio al por mayor y al por menor ... y hogares que emplean personal doméstico –este último caso por su escasa cuantía y ausencia de inputs-.

- Se han identificado seis *clusters* (agroalimentario, energía, construcción, metal-mecánico, actividades de transporte y otros servicios) así como un diverso número de parejas y tríos o *miniclusters*.

- El *cluster* agroalimentario se encuentra muy desglosado ya que al bloque dominante se le une una serie de ramas básicamente asociadas en parejas a través de un auténtico entramado de segundas cargas factoriales. En conjunto, quince ramas que recorren toda la cadena alimenticia: desde el origen –agricultura, ganadería y pesca- pasando por su industria manufacturera –productos cárnicos, lácteos, bebidas, fabricación y conservación de productos, ... - hasta llegar a los servicios representados por la actividad de cafeterías, bares y restaurantes ...

- En segundo lugar, aparece una aglomeración energética y química constituida por cinco sectores. Al bloque inicial claramente conectado –energía, química básica, gas y agua - se les une otros productos químicos a través de sus lazos con la producción y distribución de gas.

- Un *cluster* de actividad constructora que gira entorno a sus materiales y servicios afines. Está constituido por 12 ramas distribuidas en dos grandes bloques, dos conjuntos con destacadas conexiones internas pero también ligados entre sí vía segundas y terceras cargas factoriales.

- Un *cluster* metal-mecánico compuesto por extracción de minerales metálicos, fabricación de productos metálicos, maquinaria y, otras industrias manufactureras y reciclaje.

- Una agrupación con fabricación de equipo de transporte y otras ramas más o menos afines, conectada con segundas cargas con el *cluster* anterior. Se trata de un agregado de más difícil definición como entidad propia y que podría encontrar su ubicación final bien con el metal-mecánico bien con el *cluster* de la energía.

▪Por último, emerge un gran grupo misceláneo de servicios intermedios y finales, mercado o no mercado: telecomunicaciones, intermediación financiera, actividades informáticas...

▪La mayoría de las actividades económicas se encuentran interrelacionadas y de ahí la necesidad de introducir límites subjetivos de interrelación, máxime cuando el AF/ACP puede relacionar a un mismo sector con diferentes agrupaciones. Pero este hecho, en nuestra opinión, no debe ser visto como un inconveniente sino como un adecuado reflejo de la realidad.

**Cuadro N° 2: Componentes principales. Matriz de componentes rotados**

	1	2	3	4	5	6	7	8	9	10	12	15	16	18	19	21
19 Cerveza y bebidas no alcohólicas		0,9														
17 Otros productos alimenticios. Tabaco		0,9														
10 Industria cárnica		0,8											0,4			
11 Conservación de pescados		0,8														
18 Elaboración de vinos y alcoholes		0,8										0,6				
6 Pesca		0,7						0,4								
14 Industrias lácteas		0,7											0,6			
1 Cultivos de hortalizas y frutas		0,5							0,5					0,5		
13 Fabricación de grasas y aceites		0,4											0,9			
12 Conservación de frutas y hortalizas		0,6											0,8			
2 Cultivos de vid y olivo									0,6							
3 Otros cultivos y servicios agrarios									0,5				0,3			
16 Alimentación animal													0,9			
15 Fabricación de productos de molinería													0,5	0,5		
4 Producción ganadera y caza															0,8	
55 Cafeterías, bares y restaurantes		0,4						0,3								0,5
68 Actividades jurídicas, de contabilidad, etc.		0,9														
61 Intermediación financiera		0,8														
76 Educación de mercado		0,8														
60 Correos y telecomunicaciones		0,7					0,4									
54 Comercio al por menor		0,7														
69 Servicios técnicos de arquitectura e ingeniería		0,7	0,6													
63 Auxiliares a la intermediación financiera		0,7														
66 Actividades informáticas		0,6														0,3
71 Servicios de investigación y seguridad		0,6						0,3		0,6						
31 Fabricación de productos cerámicos			0,9													
30 Fabricación de cemento, cal, yeso			0,9													
50 Preparación e instalación de obras		0,3	0,9													
32 Industrias del vidrio y de la piedra			0,9													
65 Alquiler de maquinaria		0,4	0,7	0,3				0,3								
9 Extracción de minerales no metálicos			0,6	0,3				0,5								
23 Industria de la madera y del corcho			0,6													
42 Fabricación de otro material de transporte			0,9													
59 Actividades anexas a los transportes			0,9													
26 Refino de petróleo y tratamiento de residuos nucleares			0,9													
51 Comercio de vehículos y carburantes		0,4	0,4	0,7												
52 Mantenimiento y reparación de vehículos			0,7													
8 Extracción de minerales metálicos				0,9												
44 Otras industrias manufactureras				0,9												
45 Reciclaje				0,9											0,5	
35 Industria de la maquinaria y equipo mecánico				0,8	0,3										0,4	
34 Fabricación de productos metálicos			0,6	0,7												

## 5.2. Obtención de *clusters* por el método de triple restricción

La aplicación del método de triple restricción propuesto ha permitido identificar los siguientes agrupamientos –ver gráfico nº 2 -:

Un *cluster* agroalimentario amplio donde no obstante quedan fuera las ramas de fabricación de grasas y aceites, vino, cerveza y pesca. Incluye, sin embargo, restaurantes. Tiene relaciones a través de la química básica –que incluye fertilizantes y pesticidas- con el *cluster* energía y química.

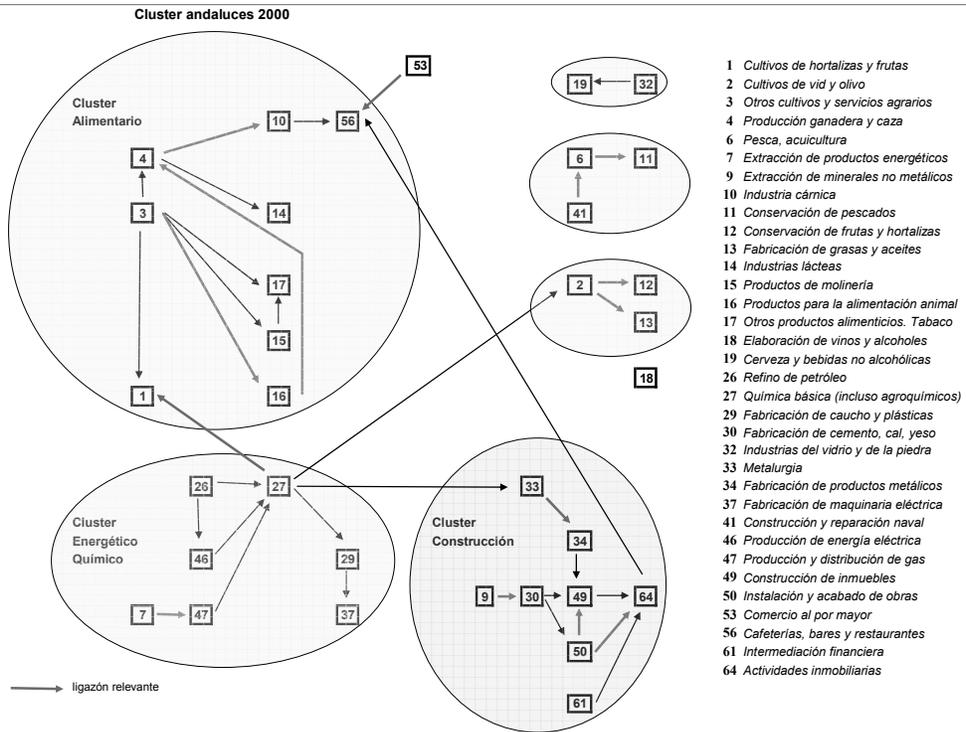
Se observa en el *cluster* el importante papel como emisor que hace la rama 3: otros cultivos y servicios agrarios y el papel como emisor de primeras materias lácteas y cárnicas y como receptor de piensos y de otros cultivos y servicios agrarios- que desarrolla la rama 4 Producción ganadera. Por último, también la rama 16: productos para alimentación animal, adquiere un importante volumen de productos agrarios que reenvía meramente transformados pero en un volumen igualmente importante a la ganadería.

Un *cluster químico-energético* que, como se ha indicado, mantiene relaciones con el agroalimentario, pero que también lo hace con el de la construcción y de nuevo a través de la química básica. Esta última rama hace así de auténtica bisagra entre los tres *cluster* que quedan interconectados.

Un *cluster de la construcción* que, se ha dicho, mantiene relaciones con el químico-energético, pero también con el agroalimentario a través de una relación entre restaurantes y la rama administración pública.

Un *cluster de la madera* con rama de cabecera en la silvicultura que comprende todos sus derivados, de muebles a papel, publicidad, artes gráficas...

Se obtiene, por último, varios *clusters* de menor tamaño, incluso simples parejas de ramas, de carácter independiente: pesca y conservación de pescado, textiles, sanidad, asistencia social ...

Gráfico nº 2: *Clusters a partir del método de triple restricción*

### 5.3. Comparación de los *clusters* obtenidos por ambas metodologías

Al comparar los *clusters* identificados por los métodos de componentes principales y de grafos, encontramos coincidencias y disidencias toda vez que estaremos comparando *clusters* verticales con otros que son de difícil clasificación, posiblemente híbridos o mixtos, un tanto verticales, un tanto horizontales. Pero, en cualquier caso, puede destacarse un buen número de coincidencias que muestran que ambas técnicas pueden aproximarse en función, sobre todo, de la matriz input que se utilice en el análisis multivariante.

De esta manera, si con grafos se obtienen tres grandes *clusters*: **agroalimentario**, **químico-energético** y **construcción**, los mismos grandes *clusters* se obtienen con el análisis de componentes si bien, como es lógico, no están compuestos exactamente las mismas ramas. Sin embargo, y como mayor disidencia, los grafos no permiten extraer dos *cluster* que si se obtienen con componentes, el metal-mecánico y el de servicios. El metal-mecánico aún podría obtenerse con los grafos si se rebajara el filtro empleado, pero se observa entonces que se asociaría fuertemente con el químico-energético. Pero el *cluster* de servicios, si bien lo hemos podido establecer con grafos en otros ámbitos -con las tablas españolas por ejemplo- no se obtiene con la tabla andaluza.

No debe escaparse con todo lo descrito que los resultados a obtener quedan siempre muy condicionado a los objetivos o legítimos intereses del investigador y que, por ello, pueden acentuarse las semejanzas o pueden destacarse las disidencias entre ambos procedimientos que no deben ser considerados contradictorios, sino complementarios. En cualquier caso, y dada la nítida naturaleza vertical de los *cluster* obtenidos por grafos, nos parecen más pertinentes y seguiremos trabajando con ellos en los epígrafes que restan.

## 6. ANÁLISIS DE LOS *CLUSTERS*

Una vez identificados los principales *clusters* de la economía andaluza, se procederá a la cuantificación de su importancia, para lo cual se considera no sólo el peso del output generado por el *cluster* sino también su distribución. Al analizar una agrupación de sectores es posible distinguir entre los flujos que se establecen con otras ramas del propio *cluster* (*intra cluster*) y los que se constituyen con el resto de sectores de la economía (*inter cluster*). Así se puede llegar a conocer en cuál de estos dos ámbitos las relaciones intersectoriales son más fuertes. Por otra parte, es preciso tener en cuenta otro factor de interés como es la densidad de dichos flujos. Un conglomerado puede generar el mismo output tanto si establece vínculos con un gran número de ramas (elevada densidad) o con un número muy reducido de ellas (flujos poco densos). Será clave aquella agrupación que genere un mayor nivel de output, preferiblemente con ramas del resto de la economía, y presente un entramado muy denso de relaciones.

### 6.1. Caracterización de los *clusters* a partir de los eslabonamientos

Según lo anteriormente señalado comenzaremos por analizar la importancia de un *cluster* a partir de su caracterización como clave aplicando técnicas de análisis estructural habituales en un marco Input-Output. Dicho análisis surgió con los trabajos de Hirschman (1958), Chenery y Watanabe (1958) y Rasmussen (1956), principalmente. Posteriormente, fue extendido con otras metodologías, una de las cuales es la extracción hipotética, iniciada por Strassert (1968) y continuada por autores como Cella (1984), Sonis et al (1995) y Dietzenbacher y Van der Linden (1997), entre otros. En líneas generales, el análisis estructural se centra en caracterizar los sectores económicos; una forma bastante habitual de efectuar tal caracterización consiste en determinar sus eslabonamientos hacia atrás (BL) y hacia adelante (FL) y, a partir de ellos, establecer una clasificación con respecto a su media en el cuadro nº 3:

**Cuadro Nº 3: Clasificación de los sectores según los eslabonamientos hacia atrás y hacia adelante**

	BL < Promedio	BL > Promedio
FL < Promedio	Sectores independientes	Sectores impulsores de economía
FL > Promedio	Sectores base o estratégicos	Sectores claves

Los sectores base presentan unos eslabonamientos hacia atrás menores que la media y hacia adelante por encima del mismo. Las ramas con fuerte arrastre hacia atrás o impulsores demandan inputs de otros sectores, por tanto, destacan debido al estímulo que generan en la producción de bienes intermedios. Los denominados sectores clave presentan encadenamientos hacia atrás y hacia adelante por encima de la media, se caracterizan por los elevados niveles de oferta y demanda que generan en el resto de sectores. Los sectores independientes provocan un menor impacto en la economía dado que su desarrollo no afecta en demasía ni a los sectores a los que demandan ni a los que emplean sus productos.

En términos generales y a partir de la anterior clasificación, impulsor si el BL supera la media y estratégico si es el FL quien está sobre ella. Por otra parte y, como ya se ha señalado, existen diferentes formas de cuantificar los eslabonamientos; una de las cuales es la denominada extracción hipotética, técnica que se empleará en este trabajo.

Aunque existen algunos antecedentes puede decirse que la metodología de extracción se perfila definitivamente con el trabajo de Strassert (1968). Este autor propuso cuantificar el efecto que experimentaría una economía si se extrajera, hipotéticamente, una rama (o conjunto de ellas), para lo cual sugirió su eliminación completa (fila(s) y columna(s)). Definió el encadenamiento asociado a un sector como el cambio que genera sobre el output una vez que dicho sector es extraído, por lo tanto, cuanto mayor sea la variación de la producción mayor será su peso en la economía. Es decir, el eslabonamiento total asociado al sector  $k$ -ésimo sería:

$$L(k) = \sum_{i=1, i \neq k}^n [x_i - \bar{x}_i(k)] \quad (1)$$

Donde  $L(k)$  representa el encadenamiento total de la rama analizada,  $x_i$  es su output antes de la extracción y  $\bar{x}_i(k)$ , el output después de la misma.

$$\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{y} \quad (2)$$

El modelo de Leontief señala que donde  $\mathbf{x}$  representa el output total,  $(\mathbf{I} - \mathbf{A})^{-1}$  es la inversa de Leontief e  $\mathbf{y}$  la demanda final. Por lo tanto, puede escribirse

$$\bar{\mathbf{x}}(k) = \left[ \mathbf{I} - \bar{\mathbf{A}}(k) \right]^{-1} \bar{\mathbf{y}}(k) \quad (3)$$

Siendo  $\mathbf{A}(k)$  la matriz de coeficientes técnicos de dimensión  $(n-1)(n-1)$ , ya que se ha eliminado la fila y columna del sector  $k$ -ésimo;  $\bar{\mathbf{x}}(k)$  e  $\bar{\mathbf{y}}(k)$ , representarán los vectores de output y demanda final, respectivamente, ambos de dimensión  $(n-1)$ . Se puede observar que  $\bar{\mathbf{x}}(k)$  es menor que  $\mathbf{x}$ , esto es,  $\bar{x}_i(k) < x_i, \forall i=1, 2, \dots, k-1, k+1, \dots, n$ . La suma de las diferencias entre los elementos  $x_i$  y  $\bar{x}_i(k)$  se considera como la medida del encadenamiento total del sector analizado. El BL podría obtenerse, por ejemplo, aplicando la metodología de Rasmussen y el FL se determina de manera indirecta por la diferencia entre el eslabonamiento total y el BL.

Cella (1984) retoma la idea de Strassert y propone determinar encadenamiento total como la suma de los eslabonamientos hacia atrás y hacia adelante. Asume que

el sector que se extrae no compra ni vende productos intermedios y, de esta forma, el encadenamiento total asociado al k-ésimo sector se obtiene como la divergencia entre el output de la economía completa y el obtenido tras la extracción. La metodología de Cella permite separar el intercambio propio del sector del resto de ramas, además de posibilitar la determinación de los eslabonamientos hacia atrás y hacia adelante de una forma directa. Sin embargo, los encadenamientos así obtenidos no son simétricos, por lo que la comparación de esta técnica con el resto se hace difícil.

Algo después, Sonis et al (1995) propusieron la metodología del “*pure linkage*”, la cual puede ser considerada como una variante de la extracción, ya que consiste en separar el sector analizado del resto de la economía, en lugar de extraerlo. En otras palabras, la economía se descompone en dos bloques, por un lado se considera la rama estudiada y, por otro, el resto de ellas.

Por último, Dietzenbacher y Van der Linden (1997), han planteado la obtención de los eslabonamientos a partir de la metodología de Cella, pero empleando el modelo de Leontief para el cálculo del BL y el de Ghosh para la determinación del FL.

La operatoria de los métodos de extracción hipotética se basa, en términos generales, en la utilización de matrices particionadas, de esta forma,  $\mathbf{A}_{ij}$  representan submatrices de coeficientes técnicos; en concreto,  $\mathbf{A}_{ij}$  recoge las relaciones dentro del bloque de sectores que se desea extraer,  $\mathbf{A}_{jr}$  y  $\mathbf{A}_{rj}$  muestran los intercambios de las ramas analizadas con el resto de la economía y  $\mathbf{A}_{rr}$  refleja el comportamiento del resto de ramas. Representando el modelo de Leontief de acuerdo a lo anteriormente señalado, se tiene

$$\begin{bmatrix} \mathbf{x}_j \\ \mathbf{x}_r \end{bmatrix} = \left[ \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} - \begin{bmatrix} \mathbf{A}_j & \mathbf{A}_j \\ \mathbf{A}_j & \mathbf{A}_r \end{bmatrix} \right]^{-1} \begin{bmatrix} \mathbf{y}_j \\ \mathbf{y}_r \end{bmatrix} \quad (4)$$

Donde  $\mathbf{x}_j$  hace referencia al output total del sector que se extrae y  $\mathbf{x}_r$  al del resto de los sectores,  $\mathbf{y}_j$  e  $\mathbf{y}_r$  representan la demanda final de ambos bloques. Aplicando las fórmulas habituales para determinar la inversa de una matriz particionada, se obtiene:

$$\begin{bmatrix} \mathbf{x}_j \\ \mathbf{x}_r \end{bmatrix} = \begin{bmatrix} \mathbf{H} & \mathbf{H}\mathbf{A}_{jr}\mathbf{G}_{rr} \\ \mathbf{G}_{rr}\mathbf{A}_{rj}\mathbf{H} & \mathbf{G}_{rr}(\mathbf{I} + \mathbf{A}_{rj}\mathbf{H}\mathbf{A}_{jr}\mathbf{G}_{rr}) \end{bmatrix} \begin{bmatrix} \mathbf{y}_j \\ \mathbf{y}_r \end{bmatrix} \quad (5)$$

Con  $\mathbf{G}_{rr} = (\mathbf{I} - \mathbf{A}_{rr})^{-1}$  y  $\mathbf{H} = [(\mathbf{I} - \mathbf{A}_{jj}) - \mathbf{A}_{jr}\mathbf{G}_{rr}\mathbf{A}_{rj}]^{-1}$ .

Para obtener el BL, Dietzenbacher y Van der Linden consideran que el sector que se analiza no compra inputs ni a si mismo ni al resto de ramas ( $\mathbf{A}_{jj} = \mathbf{A}_{rj} = \mathbf{0}$ ).

$$\bar{\mathbf{x}}(j) = \begin{bmatrix} \bar{\mathbf{x}}_j \\ \bar{\mathbf{x}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{A}_j \\ \mathbf{0} & \mathbf{A}_r \end{bmatrix} \begin{bmatrix} \bar{\mathbf{x}}_j \\ \bar{\mathbf{x}}_r \end{bmatrix} + \begin{bmatrix} \mathbf{y}_j \\ \mathbf{y}_r \end{bmatrix} \quad (6)$$

Operando convenientemente en la ecuación anterior, se llega a

$$\bar{\mathbf{x}}(j) = \begin{bmatrix} \bar{\mathbf{x}}_j \\ \bar{\mathbf{x}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{A}_j\mathbf{G}_r \\ \mathbf{0} & \mathbf{G}_r \end{bmatrix} \begin{bmatrix} \mathbf{y}_j \\ \mathbf{y}_r \end{bmatrix} \quad (7)$$

Y dado que  $d(j) = i' [x - \bar{x}(j)] y$  se obtiene:

$$d(j) = \begin{bmatrix} i'_j & i'_r \end{bmatrix} \begin{bmatrix} H-I & (H-I)A_{jr}G_{rr} \\ G_{rr}A_{ij}H & G_{rr}A_{ij}HA_{jr}G_{rr} \end{bmatrix} \begin{bmatrix} y_j \\ y_r \end{bmatrix} \quad (8)$$

De donde se deriva la expresión siguiente:

$$d(j) = [i'_j(H-I) + i'_r(G_{rr}A_{ij}H)]y_j + [i'_j(H-I)A_{jr}G_{rr} + i'_r(G_{rr}A_{ij}HA_{jr}G_{rr})]y_r \quad (9)$$

La fórmula de cálculo del eslabonamiento hacia atrás en términos relativos es:

$$BL_j^{D-VDL} = 100 \left( \frac{d(j)}{x_j} \right)$$

Con la finalidad de cuantificar el  $E_i^{D-VDL}$ , Dietzenbacher y Van der Linden parten del supuesto que el sector  $i$ -ésimo no se interrelaciona con el resto, lo cual da lugar a que la fila correspondiente de la matriz de distribución ( $B$ ) sea cero, es decir, la rama no vende *inputs* ni a si misma ni al resto de la economía ( $B_{ii} = B_{ir} = 0$ ).

$$\bar{x}'(i) = \begin{bmatrix} \bar{x}'_i & \bar{x}'_r \end{bmatrix} \begin{bmatrix} 0 & 0 \\ B_{ri} & B_{rr} \end{bmatrix} + [v_i \quad v_r] \quad (10)$$

Donde  $\bar{x}'(i)$  es el output total transpuesto de la rama  $i$ -ésima  $v_j$ , y  $v_r$  representan el valor añadido de los sectores extraídos y del resto de la economía, respectivamente.

El eslabonamiento hacia adelante se obtiene a partir de la expresión  $d(i) = v[x' - \bar{x}'(i)]$ , que formulada matricialmente toma la forma que sigue

$$d(i) = \begin{bmatrix} v_i & v_r \end{bmatrix} \begin{bmatrix} \hat{H}-I & \hat{H}B_{rr}W_{rr} \\ W_{rr}B_{ri}(\hat{H}-I) & W_{rr}B_{ri}\hat{H}B_{ir}W_{rr} \end{bmatrix} \begin{bmatrix} i_i \\ i_r \end{bmatrix} \quad (11)$$

donde  $W_r = (I - B_r)^{-1}$ . O análogamente

$$d(i) = v_i [(\hat{H}-I)i_i + (\hat{H}B_{ir}W_{rr})i_r] + v_r [W_{rr}B_{ri}(\hat{H}-I)i_i + W_{rr}B_{ri}\hat{H}B_{ir}W_{rr}i_r]$$

Finalmente, el FL relativo se calcula como  $E_i^{D-VDL} = 100 \left( \frac{d(i)}{x_i} \right)$ .

Comenzaremos a caracterizar los *clusters* a partir del cálculo de los encadenamientos utilizando el método de extracción propuesto por Dietzenbacher y Van der Linden. Se empleará este planteamiento por considerarlo más adecuado que otros comentados: la extracción no es tan taxativa como la que propone Strassert y posibilita determinar los encadenamientos simétricos de forma directa y emplea la matriz de Ghosh en la cuantificación del FL.

Como ya se ha señalado, la importancia del *cluster* dependerá, al menos en parte, del valor que alcancen los eslabonamientos hacia atrás y hacia adelante. La obtención de los primeros se efectuará a partir de la matriz de coeficientes técnicos

y la de los segundos mediante la matriz de distribución. Para efectuar la extracción se reordenará la matriz de flujos de manera que los sectores que forman parte del mismo *cluster* se encuentren situados consecutivamente. A partir de dicha matriz se determinarán los coeficientes técnicos y se obtendrá una nueva matriz particionada que se denotará por **A**:

$$\begin{bmatrix} \mathbf{A}_{jj} & \mathbf{A}_{jr} \\ \mathbf{A}_{rj} & \mathbf{A}_{rr} \end{bmatrix}$$

$\mathbf{A}_{jj}$  representará las relaciones de los sectores que forman parte del *cluster* con ellos mismos, mientras que  $\mathbf{A}_{jr}$  y  $\mathbf{A}_{rj}$  mostrarán los vínculos de las ramas del conglomerado con las demás y finalmente  $\mathbf{A}_{rr}$  recogerá las relaciones de los sectores fuera del *cluster*.

A continuación se procederá a determinar los eslabonamientos asociados al *cluster* analizado empleando la técnica anteriormente señalada. Con la finalidad de facilitar la clasificación y continuando la analogía con la metodología de Rasmussen, se han normalizado los encadenamientos de acuerdo a su media, de esta forma los indicadores que estén por encima (debajo) de la unidad son los que se encuentran sobre (bajo) su media.

Los resultados derivados de la aplicación de la técnica señalada se recogen en la tabla número nº 4. Es preciso notar que la clasificación que se muestra sólo considera los valores de los eslabonamientos y es, por tanto, preliminar.

**Cuadro nº 4: Clasificación de los *clusters***

<i>Cluster</i>	$BL_j$	$FL_j$	Clasificación
Agroalimentario	1,1738	0,5981	Impulsor
Químico-energético	0,6616	0,9754	Independiente
Construcción	1,0624	1,0292	Clave
Madera	1,1019	1,3970	Clave

Como se puede apreciar, el *cluster* agroalimentario se muestra como impulsor de la economía andaluza, dado que su BL supera la unidad y su FL está bajo ella. Se caracterizan, por tanto, por “arrastrar” a otros ramas estimulando su producción. El conglomerado químico-energético puede ser calificado como independiente al presentar ambos eslabonamientos por debajo de la media. Las agrupaciones formadas por sectores relacionados con la construcción y con la madera, aparecen como *claves* en la economía andaluza, es decir, propician el crecimiento económico al demandar bienes intermedios de otros sectores y a su vez ofrecer sus bienes al resto de la economía. Aunque también es cierto que el *cluster* madera presenta una caracterización como clave más “sólida” que la del conglomerado de la construcción, ya que los valores de este último se encuentran rayando la unidad.

## 6.2. Descomposición de los encadenamientos *intra e inter cluster*

Los encadenamientos anteriormente calculados pueden considerarse como globales, ya que se ha analizado cada conglomerado como un bloque único, sin embargo, se pueden diferenciar dentro de cada uno de ellos, por un lado, los flujos establecidos dentro del propio *cluster*, es decir, compras o ventas entre los sectores que forman el conglomerado y, por otro, los intercambios realizados con el resto de ramas de la economía (*inter cluster*). Con el objetivo de profundizar en este análisis se efectuará una descomposición de los hacia atrás y hacia adelante *linkages* propuestos por Dietzenbacher y Van der Linden en estas dos componentes. Para llevar a cabo dicha descomposición nos apoyaremos en los trabajos de Miyazawa (1971), Sonis, Hewings y Miyazawa (1997) y Hewings et al (1999).

Comenzaremos por referirnos al eslabonamiento hacia atrás. La matriz  $\mathbf{A}$  de coeficientes técnicos puede separarse en dos submatrices de la siguiente forma

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{jj} & \mathbf{A}_{jr} \\ \mathbf{A}_{rj} & \mathbf{A}_{rr} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{A}_{jr} \\ \mathbf{0} & \mathbf{A}_{rr} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_{jj} & \mathbf{0} \\ \mathbf{A}_{rj} & \mathbf{0} \end{bmatrix} = \mathbf{A}_j + \mathbf{A}_r \quad (12)$$

Se puede observar que la primera matriz corresponde al supuesto de extracción considerado por Dietzenbacher y Van der Linden: las compras del *cluster* son nulas tanto dentro del propio *cluster* como con el resto de la economía. Nótese que el eslabonamiento hacia atrás puede expresarse como  $d(j) = i'[(\mathbf{I} - \mathbf{A})^{-1} - (\mathbf{I} - \mathbf{A}_j)^{-1}]y$ .

Existen diferentes maneras de efectuar dicha descomposición, nosotros emplearemos la aditiva por considerar que se adecua mejor a los fines que perseguimos.

Si descomponemos aditivamente la Leontief inversa matriz (Sonis, Hewings y Miyazawa (1997) y Hewings et al (1999)) se tiene  $\mathbf{L} = \mathbf{L}_j + (\mathbf{M}_L - \mathbf{I})\mathbf{L}_r$ . Donde  $\mathbf{L}$  es la matriz inversa de Leontief (completa),  $\mathbf{M}_L = \mathbf{L}(\mathbf{I} - \mathbf{A}_j) = (\mathbf{I} - \mathbf{L}_j\mathbf{A}_j)^{-1}$  y  $\mathbf{L}_j = (\mathbf{I} - \mathbf{A}_j)^{-1}$ . Por lo tanto,  $d(j) = i'(\mathbf{L} - \mathbf{L}_j)y$ , entonces el *linkage* podría descomponerse como  $(\mathbf{M}_L - \mathbf{I})\mathbf{L}_r$ , esto es,

$$d(j) = \begin{bmatrix} i'_j & i'_r \end{bmatrix} \left( \begin{bmatrix} \mathbf{D}_j & \mathbf{0} \\ \mathbf{D}_r\mathbf{A}_{rj}\mathbf{G}_{jj} & \mathbf{I} \end{bmatrix} - \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \right) \begin{bmatrix} \mathbf{I} & \mathbf{A}_{jr}\mathbf{G}_{rr} \\ \mathbf{0} & \mathbf{G}_{rr} \end{bmatrix} \begin{bmatrix} y_j \\ y_r \end{bmatrix}$$

Si descomponemos nuevamente la primera matriz, tenemos

$$d(j) = \begin{bmatrix} i'_j & i'_r \end{bmatrix} \left( \begin{bmatrix} \mathbf{D}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_r \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{A}_{rj}\mathbf{G}_{jj} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \right) \begin{bmatrix} \mathbf{I} & \mathbf{A}_{jr}\mathbf{G}_{rr} \\ \mathbf{0} & \mathbf{G}_{rr} \end{bmatrix} \begin{bmatrix} y_j \\ y_r \end{bmatrix}$$

Se puede observar que la primera matriz está formada por los denominados multiplicadores extendidos de Leontief, es decir,  $\mathbf{D}_j = [\mathbf{I} - \mathbf{A}_{jj} - \mathbf{A}_{jr}\mathbf{G}_{rr}\mathbf{A}_{rj}]^{-1}$  y  $\mathbf{D}_r = [\mathbf{I} - \mathbf{A}_{rr} - \mathbf{A}_{rj}\mathbf{G}_{jj}\mathbf{A}_{jr}]^{-1}$ , los cuales son las inversas de los complementos de Schur ( $\mathbf{S}_j = \mathbf{A}_{jj} + \mathbf{A}_{jr}\mathbf{G}_{rr}\mathbf{A}_{rj}$  y  $\mathbf{S}_r = \mathbf{A}_{rr} + \mathbf{A}_{rj}\mathbf{G}_{jj}\mathbf{A}_{jr}$ ) y pueden ser interpretados como efectos de retroalimentación de la economía. Además es posible, a su vez, descomponerlos en el producto de los multiplicadores internos y externos de Miyazawa. Los multiplicadores internos se

representan por  $G_{jj}$  y  $G_{rr}$ , el primero recoge la propagación de los efectos dentro del cluster y el segundo, en el resto de la economía. Mientras que los externos recogen las relaciones entre los clusters, es decir,  $\Delta_j = [I - G_{jj} A_{jr} G_{rr} A_{rj}]^{-1}$  y  $\Delta_r = [I - G_{rr} A_{rj} G_{jj} A_{jr}]^{-1}$ . Por lo tanto, los multiplicadores extendidos tomarán la forma siguiente:  $D_j = \Delta_j G_{jj}$  y  $D_r = \Delta_r G_{rr}$ .

A partir de la descomposición efectuada, tendríamos de dos nuevos índices que proporcionan información que facilita la caracterización de los conglomerados: un indicador del peso de los flujos internos, medidos a través del multiplicadores internos de Miyazawa, y otro de la importancia de las relaciones entre clusters, definido a partir del multiplicador externo.

El anterior planteamiento puede extenderse al FL, ahora los multiplicadores internos de Leontief tomarían la expresión  $W_i = (I - B_i)^{-1}$  y  $W_r = (I - B_r)^{-1}$  y los externos  $T_i = [I - W_i B_r W_r B_i]^{-1}$  y  $T_r = [I - W_r B_i W_i B_r]^{-1}$ .

Los resultados obtenidos se presentan en el cuadro n° 5:

**Cuadro N° 5: Descomposición de los eslabonamientos**

	BL <i>intra cluster</i>	FL <i>intra cluster</i>	BL <i>inter cluster</i>	FL <i>inter cluster</i>
Agroalimentario	1,1365	1,0742	0,9990	1,0001
Químico-energético	0,9497	0,8960	0,9927	1,0049
Construcción	0,9773	1,1227	1,0121	1,0190
Madera	0,9365	0,9071	1,0030	1,0028

Por lo que se refiere a los BL y FL internos, puede apreciarse que los valores más elevados (por encima de la media) aparecen en el *cluster* agroalimentario y en el de la construcción, En lo que respecta a las compras y ventas de cada uno de los *clusters* determinados con el resto de la economía, se observa que, de nuevo, construcción y madera presentan eslabonamientos sobre el media. Estos resultados corroboran los anteriormente obtenidos, consideramos *clusters* claves aquellos que muestran encadenamientos elevados debidos al peso que tienen sus relaciones con el resto.

### 6.3. Caracterización del *cluster* a partir de la densidad de sus relaciones

Hasta aquí sólo hemos considerado la importancia del output del *cluster* extraído, es decir, un conglomerado será más importante de acuerdo al peso que tiene su producción, pero también puede ser interesante analizar la distribución de los flujos de cada *cluster*. Es decir, un *cluster* puede ser caracterizado como *clave* de acuerdo a su output, pero, sin embargo, la red que constituyen sus flujos sea poco densa, esto es, se establecen vínculos entre un pequeño número de sectores y no con la mayoría (totalidad) de los mismos. Para analizar la distribución de los

flujos sectoriales algunos autores emplean el coeficiente de Pearson, sin embargo, esta medida es propiamente de dispersión existiendo indicadores específicos que permiten considerar la distribución de una variable, como por ejemplo los índices de concentración. Dichos índices deben verificar un conjunto de propiedades, las cuales, según Encaoua y Jacquemin (1980) pueden resumirse en los dos axiomas que siguen:

a) La transferencia de una parte de producción de una empresa a otra de mayor tamaño no debe disminuir la concentración.

b) Si se unen dos o más empresas la concentración no disminuye.

Aunque la primera característica es verificada por el coeficiente de Pearson, no ocurre lo mismo con la segunda. Además, y en otro orden de cosas, dicho indicador no está acotado, con lo cual se dificulta su interpretación. Para analizar la distribución de los flujos se pueden emplear medidas derivadas de la teoría estadística de la información, una de ellas es el índice cuadrático. Consideremos  $\xi$  una variable aleatoria discreta con distribución de probabilidad  $P=(p_1, p_2...p_n)$ , se puede establecer la siguiente definición de entropía de grado  $\beta$  ( $\beta \neq 1$ ),  $\beta > 0$ :

$$H^\beta(\xi) = H^\beta(p_1, \dots, p_n) = \frac{1}{2^{1-\beta} - 1} \sum_i (p_i^\beta - p_i) \quad (13)$$

Siendo  $\beta$  el parámetro que caracteriza a esa familia; para el valor concreto  $\beta=2$  surge la denominada entropía cuadrática. Así pues, se define la entropía cuadrática a partir de la expresión

$$H^2(\xi) = 2 \sum_i p_i(1 - p_i) = 2 \left[ 1 - \sum_i p_i^2 \right] \quad (14)$$

Esta medida está acotada entre cero y  $2 \left( 1 - \frac{1}{n} \right)$ , alcanzando su menor valor en el caso de que la distribución sea degenerada, esto es, cuando son nulas n-1 probabilidades y la restante toma el valor uno, y el máximo cuando las probabilidades son todas ellas coincidentes.

Dado que nuestro interés se centra en la determinación de la incertidumbre asociada a una tabla Input-Output, consideraremos una variable aleatoria bidimensional que presenta una probabilidad conjunta  $p_{ij}$ . La medida de entropía bidimensional

$H^2(\xi, \eta) = \left[ 1 - \sum_i \sum_j p_{ij}^2 \right]$  presenta análogo comportamiento que la unidimensional

antes señalada, permitiendo, por tanto, cuantificar la densidad de las relaciones establecidas por el cluster. A medida que este índice se aproxime a su cota superior más diversificado estará el conglomerado estudiado, ya que más uniformemente repartidos están sus transacciones interindustriales. Y viceversa si se aproxima a cero, en este caso un cluster presentará una estructura productiva más concentrada. En la práctica,

en lugar de emplear la entropía se trabaja con el índice cuadrático de concentración, que es la mitad de la entropía cuadrática, es decir,

$$I^2(\xi, \eta) = \left[ 1 - \sum_i \sum_j p_{ij}^2 \right] \quad (15)$$

A partir de la expresión anterior se ha procedido a determinar la densidad de las relaciones tanto en compras como en ventas global, *intra* y *inter cluster*. Como puede apreciarse, para aplicar estos conceptos de teoría de la información a una tabla Input-Output se necesita un sistema de probabilidades<sup>8</sup> que puede establecerse al dividir cada celda de los flujos interindustriales entre el total de los mismos, esto es,

$$p_{ij} = \frac{z_j}{Z} \quad (16)$$

donde  $z_{ij}$  son los consumos intermedios y  $Z = \sum_i \sum_j z_{ij}$  es el total de dichos consumos.

Se han expresado dichas densidades en términos relativos con respecto a su cota superior, para efectuar una comparación más nítida, al estar constituidos los *clusters* por distinto número de sectores.

**Cuadro n° 6: Densidad global de las compras y ventas de los *clusters***

	Densidad compras	Densidad ventas
Agroalimentario	0,9369	0,9112
Químico-energético	0,9706	0,9822
Construcción	0,9746	0,9427
Madera	0,9721	0,9632

Se puede observar que los valores son en general elevados, bastante próximos a sus cotas superiores. El *cluster* agroalimentario es el que presenta unas compras y ventas más concentradas, es decir, menos densas. El conglomerado de la construcción muestra la mayor densidad en compras, seguido de cerca por el de la madera. Si nos fijamos en las ventas, podemos observar que las agrupaciones químico-energética y madera son las menos concentradas. Si consideramos, al igual que Rasmussen, que un sector clave es aquel que tiene elevados eslabonamientos y alta dispersión de sus efectos, podríamos caracterizar de esta forma al conglomerado de la madera, dado que verifica ambas condiciones.

Análogamente puede también determinarse la densidad de las relaciones dentro y entre cluster y de esa forma se tendrá una visión pormenorizada de los mismos. Los resultados de aplicar el índice cuadrático a los flujos dentro y entre cluster aparecen recogidos en el cuadro n° 7:

<sup>8</sup> Este aspecto está recogido, por ejemplo, en Tilanus y Theil (1965).

**Cuadro N° 7: Descomposición de la densidad de las compras y ventas**

	Densidad compras internas	Densidad ventas internas	Densidad compras externas	Densidad ventas externas
Agroalimentario	0,8568	0,8978	0,9839	0,9324
Químico-energético	0,9106	0,9042	0,9805	0,9812
Construcción	0,9081	0,9016	0,9731	0,9804
Madera	0,8218	0,8304	0,9850	0,9705

Se aprecia que, como en el caso anterior, las densidades son elevadas, alcanzándose los valores más bajos en las compras y ventas dentro del *cluster* de la madera, es decir, las compras y ventas se encontrarían menos uniformemente repartidas; un comportamiento similar presenta el *cluster* agroalimentario. Además, las relaciones establecidas por cada *cluster* con el resto de la economía se encuentran menos concentradas que en el interior del mismo. El *cluster* químico-energético es el que presenta un mayor valor del índice tanto en sus relaciones internas como externas, seguido de la construcción y madera, que también muestran elevados valores del índice y, por tanto, menores niveles de concentración en los flujos de compras y ventas.

En la caracterización de un *cluster* como clave parece más relevante tomar en cuenta el peso y la densidad de las relaciones externas frente a las internas, esto es, considerar preferentemente los flujos establecidos con el resto de la economía, por ello y de la consideración conjunta de los valores de los encadenamientos y de la densidad de dichos flujos, se desprende que los conglomerados construcción y madera pueden ser considerados como claves para la economía andaluza.

## 7. CONCLUSIONES

En este trabajo se ha efectuado una evaluación de las distintas técnicas empleadas en la obtención de *clusters* a partir de tablas Input-Output. Para ello, se han revisado e incluso replicado -siempre que ello ha sido posible- los trabajos recogidos en la literatura existente, así como se ha efectuado un nuevo ensayo sobre las tablas de una de las mayores regiones españolas, Andalucía.

Dos son las técnicas o familias de técnicas habitualmente empleadas, unas basadas en el análisis multivariante y otras en la teoría de grafos o métodos afines. Las primeras son, en principio, útiles en la obtención de los *clusters* que hemos denominado *horizontales*, mientras que los grafos son más útiles para la elaboración de *clusters* basados en los eslabonamientos clásicos o relaciones *verticales*. Sin embargo, la literatura existente sobre métodos multivariantes, aplica la técnica sobre matrices *inputs* de correlación de coeficientes que poseen una naturaleza cuanto menos discutible, perdiéndose así el sentido horizontal que pudieran tener los *clusters* resultantes. La no consideración, además, de diversas ramas muy relacionadas con las demás dado

que aglutinarían grandes *clusters* parece un gran inconveniente adicional que reduce apreciablemente el ámbito de la aplicación multivariante.

Por otra parte, la elaboración de *clusters* a partir de técnicas de grafos ha venido siendo quizás poco exigente a la vez que ha sido también criticada por ofrecer resultados no únicos. Sin embargo, un método de triple restricción como el que se ha diseñado y se presenta, parece contar con los requisitos adecuados para la obtención de *clusters* consistentes. Con todo, los resultados no son únicos pues, de acuerdo a nuestra apreciación, si bien cabe identificar en unas tablas IO un núcleo de *clusters* indiscutibles –*core cluster*– el resto será resultado de las condiciones o restricciones fijadas, pudiendo variar en la misma medida en que éstas lo hagan. Este hecho, lejos de resultar un inconveniente, lo observamos como un reflejo adecuado de la compleja realidad.

Cuatro grandes *clusters* se han obtenido finalmente con el método de grafos propuesto: el *cluster* agroalimentario, el químico-energético, el de la construcción y el de la madera, definidas estas agrupaciones, el paso siguiente ha consistido en su análisis y caracterización. Para ello se siguió el procedimiento planteado por Rasmussen (1956), definiendo como claves aquellos conglomerados que tienen unos eslabonamientos por encima de la media y a la vez presentan alta densidad en sus relaciones y vínculos de compras y ventas. Para determinar los encadenamientos se ha aplicado el método de extracción hipotética de Dietzenbacher y Van der Linden (1997), dado que se consideró más adecuado que las otras técnicas. Asimismo, se estudió la distribución de los efectos a través del índice cuadrático. Con el objetivo de profundizar en el análisis se han separado las relaciones establecidas dentro de cada *cluster* y las establecidas entre las distintas agrupaciones, aunque se consideran más relevantes para la economía estas últimas. Del empleo de las técnicas anteriormente señaladas, se desprende que los conglomerados construcción y madera pueden ser considerados como claves para la economía andaluza.

## 8. REFERENCIAS BIBLIOGRÁFICAS

- AROQUE-REYES, F. (2001) The question of identifying industrial complexes revisited: a qualitative perspective, en Lahr, M.L. y Dietzenbacher (eds), *Input-Output Analysis: Frontiers and Extensions*, (New York, E. Palgrave Publishers Ltd).
- BERGMAN E. M. Y LEHNER, P. (1995) Industrial cluster formation in European regions: U.S. cluster templates and Austrian evidence, SRE-Discussion 65, (Wien, Wirtschaftsuniversität Wien), <http://www.wu-wien.ac.at/inst/sre>.
- BERGMAN, E. Y FESER, E. J. (2000) National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis, *Regional Studies*, Vol. 34, I, pp.1-19.
- BERGMAN, E. Y FESER, E. J., SWEENEY, S. (1996) Targeting North Carolina manufacturing: understanding the state 's economy through industrial cluster analysis (North Carolina, Institute for Economic Development, University of North Carolina at Chapel Hill).

- BLIN, J Y COHEN, C. (1977) Technological Similarity and Aggregation in Input-Output Systems: A Cluster-Analytic Approach, *Review of Economics and Statistics*, 59, pp. 82-91.
- BON, R. (1989) Qualitative Input-Output analysis, *Frontiers in Input-Output Analysis*, (New York, Oxford University Press).
- CAMPBELL, J. (1975) Application of Graph Theoretic Analysis to Interindustry Relationships, *Regional Science & Urban Economics*, 5, pp. 91-106.
- CELLA, G. (1984) The Input-Output Measurement of Interindustry Linkages, *Oxford Bulletin of Economics and Statistics*, 46, pp. 73-84.
- CHENERY, H. B. & WATANABE, T. (1958) International comparison of the structure of production, *Econometrica*, vol. XXVI, nº 26, pp. 487-521.
- CZAMANSKI S. Y ABLAS, L. A. (1979) Identification of industrial clusters and complexes: a comparison of methods and findings, *Urban Studies*, 16, pp. 61-80.
- CZAMANSKI, S. (1971) Some empirical evidence of the strengths of linkages between groups of related industries in urban-regional complexes, *Papers, Regional Science Association*, 27, pp. 137-150 .
- CZAMANSKI, S. (1977) Needless complexity in the identification on industrial complexes: A comment, *Journal of Regional Science*, 17, 3, 455-457.
- DIETZENBACHER, E., VAN DER LINDEN, J. Y STEENGE, J. (1993) The Regional Extraction Method: EC Input-Output Comparisons, *Economic Systems Research*, 5, pp.185-206.
- DIETZENBACHER, E., VAN DER LINDEN, J. (1997) Sectoral and Spatial Linkages in the EC Production Structure, *Journal of Regional Science*, 37, pp. 235-257.
- DOERINGER, P.B. Y TERKLA, D.G. (1995) Business strategy and cross-industry clusters, *Economic Development Quaterly*, 9, pp. 225-37.
- ENCAOJA, D. Y JACQUEMIN A. (1980) Degree of monopoly, indices of concentration and threat of entry, *International Economic Review* 21(1), pp.87-105.
- FESER, E. Y SWEENEY, S (2003) Spatially binding linkages in manufacturing product chains.,. R. McNaughton and M. Green (eds.) *Global Competition and Local Networks* (New York, Ashgate). <http://www.planning.unc.edu/pdf/FeserPUBS/Chapter%205%20-20Feser%20and%20Sweeney.pdf>
- HARRIGAN, FRANK J. (1982) The relationship between industrial and geographical linkages: A case study of the United Kingdom, *Journal of Regional Science*, 22, 1, pp. 19-31.
- HEWINGS, G.J, SONIS, M, (1999) Hierarchies of Regional Sub-Structures and their Multipliers within Input-Output Systems: Miyazawa Revisited, *Hitotsubashi Journal of Economics*, 34, pp. 33-44.
- HIRSCHMAN, A. O. (1958) *The Strategy of Economic Development*, (Connecticut, Yale University Press, New Haven)

- HOEN, A. (2000) Three variations on identifying cluster, National Innovation Systems, Workshops and Meetings of the Focus Group on Clusters, 8-9 May 2000, Utrecht, OCDE.
- HOEN, A. (2001, I) An international comparison of national clusters, 41st Congress of the European Regional Science Association, August 29 - September 1, Zagreb, Croatia.
- HOEN, A. (2001, II) Clusters: Determinants and effects, CPB memorandum, Netherland Bureau for Economic Policy Analysis.
- HOEN, A. (2002) Identifying Linkages with Cluster Based Methodology, Economic Systems Research, 14, 2, pp. 131-146.
- ISARD, W., SCHOOLER, E. W. Y VIETORISZ. T. (1959) Industrial Complex Analysis and Regional Development (New York, John Wiley).
- ISARD W. Y SMOLENSKY, E. (1963) Application of Input-Output Techniques to Regional Science, in T. Barna, (ed), Structural Interdependence and Economic Development (London, Macmillan).
- KOLOSOVSKY ,N. (1965) anticipated Isard's contribution: ver R. E. Lonsdale The Soviet Concept of the Territorial Production Complex. Slavic Review Vol. 24 pp. 466- 78.
- LAINESSE, L. Y POUSSART, B. (2005) Méthode de repérage des filières industrielles sur le territoire québécois basée sur les tableaux d'entrées-sorties, Institut de la statistique du Québec, <http://www.stat.gouv.qc.ca/publications/savoir/pdf2005/Introduction.pdf>.
- LATHAM, W.R. (1977) Needless complexity in the identification of industrial complexes: a reply, Journal of Regional Science, 17, 3, pp. 459-461.
- LATHAM, WILLIAM R. (1976) Needless Complexity in the Identification on Industrial Complexes, Journal of Regional Science, 16,1, pp. 45-55.
- MESNARD, L. DE (2001) On boolean topological methods of structural analysis, en Lahr, M.L. y Dietzenbacher (eds) en Input- Output Analysis, Foundations and Extensions, (New York, Palgrave Publishers Ltd.).
- MIYAZAWA, K. (1971) An analysis of the interdependence between service and goods-producing sectors, Hitotsubashi Journal of Economics, 12, pp. 10-21.
- MORILLAS, A. (1983) La teoría de grafos en el análisis Input-Output. La estructura productiva andaluza, (Málaga, Universidad de Málaga).
- O'HUALLACHAIN, B. (1984) The identification of Industrial Complexes, Annals of the Association of American Geographers, 74, pp.420-6.
- OOSTERHAVEN, J., EDING, G. Y STELDER, D. (2001) Cluster, linkages and interregional spillovers: methodology and policy implications for the two Duch mainports and the rural north, Regional Studies, 35, 9, pp. 809-822.

- PEETERS, L., TIRI, M. Y BERWERT, A. (2001) Identification of techno-economic clusters using Input-Output data: Application to Flanders and Switzerland. *Innovative Clusters. Drivers of National Innovation Systems*, OECD, pp. 251-272.
- RASMUSSEN, P. (1956), *Relaciones Intersectoriales* (Madrid, Aguilar) .
- REY, S.J. Y MATTHEIS, D. J. (2000) Identifying Regional Industrial Clusters in California, California Center for Border and Regional Economic Studies, [http://www.ccbres.sdsu.edu/Community\\_Outreach/workshops/workshop11/cluster\\_analisis.PDF](http://www.ccbres.sdsu.edu/Community_Outreach/workshops/workshop11/cluster_analisis.PDF).
- ROEPKE H., ADAMS, D Y WISEMAN, R. (1974) A new approach to the identification of industrial complexes using Input-Output data, *Journal of Regional Science*, 14, pp.15-29.
- ROSENFELD, S.A. (1997) Bringing business clusters into the mainstream of economic development, *European Planning Studies*, 5, 1, pp. 3-23.
- SIMPSON, D, Y TSUKUI, J. (1965) The fundamental structure of Input-Output tables, an international comparison, *Review of Economics and Statistics*, 48, 4, pp. 434-446.
- SLATER, P.B. (1977,I) Needless complexity in the identification of industrial complexes: a reply, *Journal of Regional Science*, 17, 3, pp.459-461.
- SLATER, P.B. (1977,II) The Determination of Groups of Functionally Integrated Industries in the United States Using a 1967 Interindustry Flow Table, *Empirical Economics*, Vol. 2, 1, pp. 1-9.
- SONIS, M., GUILHOTO, J., HEWINGS, G. Y MARTINS, E. (1995) Linkages, Key Sectors, and Structural Change: Some New Perspectives. *The Developing Economics*, 33, 3, pp. 233-270.
- SONIS, M, HEWINGS, G.J Y MIYAZAWA, K. (1997) Synergetic Interactions within the pair-wise Hierarchy of Economic Linkages Sub-Systems, *Hitotsubashi Journal of Economics*, 38, pp. 183-199.
- STRASSERT, G. (1968) Zur Bestimmung strategischer Sektoren mit Hilfe von Input-Output Modellen. *Jahrbücher für Nationalökonomie und Statistik*, 182, 3, pp. 211-215.
- TILANUS, C.B. AND THEIL H. (1965) The Information approach to the evaluation of Input-Output forecast. *Econometrica*, Vol. 32, n° 4, pp. 847-862.
- TIRI, M., MOREAU, R., PEETERS L. (2000) Identification of clusters in Switzerland, OCDE.