

Análisis exploratorio de estructuras temporales desde la óptica de tablas múltiples. Una aplicación

LANDALUCE CALVO, M. ISABEL
Departamento de Economía Aplicada
Facultad de Ciencias Económicas y Empresariales
Universidad de Burgos (España)
Correo electrónico: iland@ubu.es

RESUMEN

En este trabajo se presenta un enfoque alternativo a los gráficos habituales de series temporales para su análisis comparativo exploratorio. Al disponer de varias series temporales medidas en diversos contextos, sugerimos su estructura en tablas múltiples para estudiar de forma adecuada sus similitudes y diferencias. El Análisis Factorial Múltiple (AFM) es la técnica seleccionada como la más idónea para alcanzar los objetivos planteados, a través de sus medidas numéricas y gráficas. Este enfoque del AFM supone una novedad, ya que en otros estudios el tiempo ha sido utilizado como criterio de estructuración de los grupos y no como unidad estadística como se plantea en este artículo. Una aplicación empírica a partir de datos mensuales, de los últimos 10 años, relativos a tres indicadores estrechamente relacionados con el turismo, desde el punto de vista de la demanda, y medidos en siete territorios nacionales, permite evidenciar la versatilidad y potencialidad de la propuesta.

Palabras claves: series temporales; análisis factorial; tablas múltiples.

Clasificación JEL: C38.

MSC2010: 62–07.

Exploratory Analysis of Temporal Structures from the Multiple Tables Optic. An Application

ABSTRACT

In this work we present an alternative approach to the usual graphics of time series for its comparative exploratory analysis. By having several time series measures in various contexts, we suggest its structure in multiple tables to study adequately their similarities and differences. The Multiple Factorial Analysis (MFA) is the technique that is selected as the most suitable for achieving the stated objectives, through their numerical measures and graphs. This approach of AFM is a novelty, since in other studies the time has been used as a criterion for structuring groups and not as statistical unit as discussed in this article. An empirical application from monthly data, of the last 10 years, relating to three indicators closely related to tourism, from the point of view of demand, and measured in seven national territories, makes evident the versatility and potential of the proposal.

Keywords: time series; factorial analysis; multiple tables.

JEL classification: C38.

MSC2010: 62-07.



1.- INTRODUCCIÓN. OBJETIVOS

La importancia del estudio de las series temporales deriva de la abundancia de este tipo de información en las ciencias, en general, y la economía, en particular. Así, en la literatura especializada se pueden encontrar muy diversos métodos diseñados para el análisis de fenómenos que permiten observar variaciones a lo largo del tiempo con fines explicativos y predictivos. El tratamiento directo de este tipo de información, dada su naturaleza dinámica, no está exento de dificultades. Si, además, el objetivo se amplía al estudio comparativo de varios fenómenos (indicadores) medidos en diferentes contextos (geográficos, por ejemplo) la complicación es aún mayor, al disponer de datos tridimensionales (véase Figura 1). Por ello, resulta conveniente proceder previamente a la descripción exhaustiva de todas las posibles regularidades y de patrones subyacentes en este tipo de información.

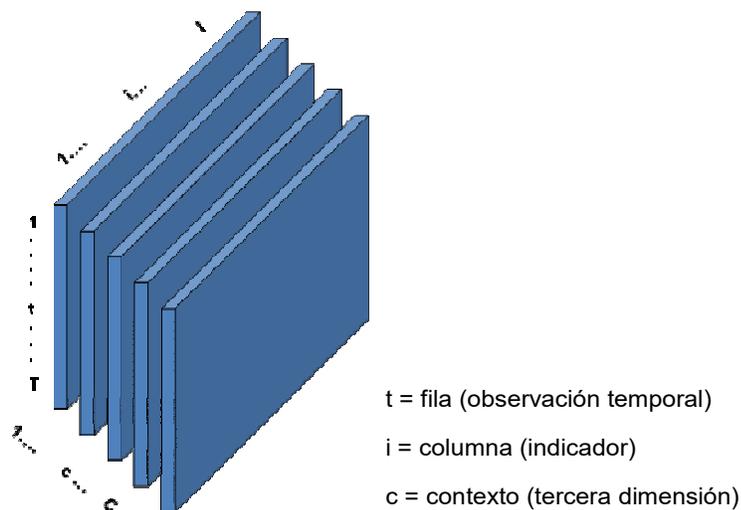


Figura 1: Estructura de la información. Tabla múltiple: conjunto de datos de tres dimensiones

La exploración es una etapa siempre previa y necesaria en el campo del análisis de datos, ya que proporciona un profundo conocimiento de la realidad estudiada. En ocasiones, en esta etapa previa se descubren regularidades que sirven de fundamento para la predicción del comportamiento de los fenómenos, incluso antes de que se entienda la razón o causa que justifique esas regularidades. En el caso concreto que nos ocupa, al tener las series estructuradas en diversas tablas, disponer de medidas de relación (semejanzas y diferencias) entre los diferentes conjuntos de datos, permitirá tomar decisiones sobre la información finalmente analizada y la aplicación de otras técnicas de análisis para posteriores objetivos con mayores garantías de éxito.

La propuesta desarrollada en este trabajo se focaliza en esta etapa de análisis descriptivo y comparativo de series temporales. Concretamente, en el uso de una técnica de análisis factorial exploratorio adaptada al tratamiento de datos susceptibles de ser agrupados en varios conjuntos de variables: Análisis Factorial Múltiple (AFM). La consideración de las tres dimensiones de los datos (fenómeno/s a analizar, observaciones temporales y diversos contextos) enriquece los objetivos de cualquier estudio.

La finalidad de esta investigación es poner de manifiesto la versatilidad de los planos factoriales del AFM para el análisis comparativo de un conjunto de series temporales. El análisis propuesto no pretende ser un sustituto del análisis visual tradicional, que siempre será útil en las investigaciones económicas, sino una herramienta complementaria. El AFM, además de ofrecer al investigador imágenes simplificadas en las que se pueden proyectar de forma conjunta unidades temporales, variables y contextos en los que se miden las anteriores, proporciona instrumentos numéricos que ayudan a cuantificar e interpretar las semejanzas y las diferencias entre las trayectorias temporales estudiadas. El investigador, a partir de estos resultados, puede, entre otras, tomar decisiones de seguir adelante con el total de la información de partida o eliminar alguno de los indicadores o incluso alguno de los contextos analizados, al presentar estructuras peculiares que merecen su estudio por separado.

La idoneidad de la propuesta de este artículo va a quedar justificada a través del análisis, con carácter meramente ilustrativo, de la siguiente información: tres indicadores estadísticos del ámbito del turismo (desde el punto de vista de la demanda), medidos en siete ámbitos geográficos a nivel nacional durante los últimos 10 años (enero 2006 – septiembre 2015). Se ha trabajado con datos mensuales correspondientes a un período lo suficientemente amplio como para reflejar diferentes fenómenos. Por un lado, hitos importantes en la evolución económica de nuestra economía (final de una larga etapa de crecimiento, depresión y recuperación incipiente) y, por otro lado, diferentes componentes de las series temporales (tendencia y componente estacional). Concretamente, los indicadores seleccionados son: Entradas de turistas según comunidad autónoma de destino principal (correspondiente a la Encuesta de Movimientos Turísticos en Fronteras, FRONTUR), Gasto medio diario de los turistas según destino principal y Estancia media de los turistas según destino principal (correspondientes a la Encuesta de Gasto Turístico, EGATUR). Además, las fuentes

secundarias consultadas proporcionan esta información desagregada para seis comunidades autónomas de forma individual y agregada para el resto de comunidades autónomas, lo que ha permitido la estructuración de la información en una tabla múltiple integrada por siete subtablas: Andalucía, Baleares, Canarias, Cataluña, Comunidad de Madrid, Comunidad Valenciana y resto de comunidades autónomas.

Este artículo está integrado por cinco secciones, además de esta primera que ha servido para situar la metodología propuesta en el preciso contexto. En la sección 2 se realiza una breve presentación de los principios del Análisis Factorial Múltiple. La idoneidad de esta técnica, dados los objetivos perseguidos, se evidencia en la sección 3, en la que se recogen los resultados de la aplicación empírica. Las conclusiones alcanzadas y las futuras líneas de investigación (sección 4), así como las referencias bibliográficas consultadas para su realización (sección 5) completan el trabajo.

2.- ANÁLISIS EXPLORATORIO DE ESTRUCTURAS TEMPORALES MÚLTIPLES. METODOLOGÍA

En el extenso campo de investigación del Análisis de Datos existen numerosas y diversas técnicas desarrolladas para el estudio exploratorio de tablas susceptibles de ser estructuradas en subtablas (también denominadas tablas múltiples, tablas de tres dimensiones, tablas de tres entradas...). Centrándonos en la Escuela Francesa de Análisis de Datos, se pueden mencionar métodos como STATIS, STATIS Dual, Doble Análisis de Componentes Principales (DACP), Análisis Factorial Múltiple (AFM), (Dazy y Le Barzic, 1996) o análisis simultáneo de tablas de contingencia (Goitisoló, 2001). La estructuración en subtablas puede estar asociada tanto a la dimensión de las columnas (indicadores) como a la dimensión de las filas (observaciones). Las investigaciones en las que se dispone de información de esta naturaleza múltiple tienen objetivos más ambiciosos, ya que no se limitan a la búsqueda de relaciones entre indicadores y tipologías de las observaciones, sino que se amplían al análisis comparativo de las realidades presentes en el seno de cada una de las tablas. Esta riqueza interpretativa, junto con la gran casuística de datos, ha animado a muchos investigadores a desarrollar metodologías apropiadas a estos objetivos. Sin embargo, nuestra experiencia como estudiosos y usuarios de varias de esas técnicas, nos conduce a proponer el Análisis Factorial Múltiple como la más apropiada para este trabajo. A lo largo del artículo quedará constancia del por qué de esta elección.

El Análisis Factorial Múltiple (AFM) es una metodología exploratoria con una gran versatilidad en el tratamiento de información de tres dimensiones. Son numerosos los trabajos que así lo reflejan, tanto desde el enfoque del desarrollo teórico como desde el enfoque de la potencialidad empírica en diversos campos científicos. Algunos de estos artículos aparecen referenciados en Fernández *et al.* (2013). Todos ellos ponen de manifiesto que el AFM con el tiempo y uso ha pasado de ser una técnica de análisis de tablas múltiples a constituir una filosofía de análisis comparativo, tanto gráfico como a través de medidas numéricas, de diferentes conjuntos de datos, sea cual sea su naturaleza y su estructura. Este trabajo se enmarca en el contexto de aplicaciones del AFM.

El AFM ya ha sido utilizado con éxito en el estudio de tablas múltiples en las que el tiempo define una de las tres dimensiones consideradas (Dazy y Le Barzic., 1996; García Lautre *et al.*, 2003; Abascal *et al.*, 2004, 2006). Sin embargo, en todos estos trabajos el tiempo es el criterio de estructuración de los grupos, esto es, cada una de las tablas está asociada a una unidad temporal distinta. Por tanto, no se exploran propiamente series temporales en ninguna de las investigaciones referenciadas. En este trabajo, el tiempo constituye la dimensión correspondiente a las observaciones (filas) en todas las tablas consideradas. En este aspecto es en el que se focaliza la principal aportación de esta investigación: la construcción y análisis de una tabla múltiple yuxtaponiendo distintos conjuntos de series temporales.

El objetivo nuclear de este artículo es presentar la riqueza interpretativa del AFM como novedosa y complementaria alternativa en este contexto de análisis exploratorio comparativo de datos dinámicos. Por ello, en esta sección se va a realizar una presentación de la técnica a través de unas breves pinceladas, incidiendo en la terminología propia del análisis de tablas múltiples y en una selección de instrumentos gráficos y numéricos que ayudan a alcanzar los objetivos planteados. El desarrollo detallado de sus principios metodológicos pueden encontrarse, entre otros trabajos, en Escofier y Pagès (1986, 1992, 1994) y Landaluce (1995).

El AFM analiza la *estructura común* de las distintas tablas de datos objeto de estudio, poniendo de manifiesto aquellas que tienen un comportamiento diferente al resto, señalando, además, qué observaciones e indicadores son los responsables del mismo.

La técnica se puede resumir en los siguientes puntos:

A) *La terminología del AFM* es la propia de los métodos de análisis de datos estructurados como tabla múltiple:

- *Nube parcial*. Es la nube de observaciones caracterizadas por un grupo de indicadores (variables). Se obtienen tantas *nubes parciales* como grupos de variables se consideren.
- *Nube global*. Es la nube de observaciones caracterizadas por todas las variables de la tabla global, resultando, por tanto, de la unión de todas las nubes parciales.
- *Puntos parciales*. El AFM proporciona una representación de todas las nubes parciales superpuesta a la representación de los puntos medios. Se proyectan tales nubes como elementos ilustrativos en los planos factoriales del análisis global.
- *Puntos medios o globales*. El AFM proporciona la representación de los elementos de la nube global en los planos factoriales obtenidos en el análisis global. Estos puntos reflejan la posición media de la observación teniendo en cuenta todos los grupos de variables considerados.
- *Inercia Total* es la variabilidad de la nube global respecto a su centro de gravedad. Se descompone de forma aditiva en Inercia Intra e Inercia Inter.
- *Inercia Intra* es la inercia de los puntos parciales respecto a su punto medio en la nube global, es decir, la inercia de las nubes correspondientes a la imagen de cada observación desde los diferentes puntos de vista, grupos, estudiados respecto a los centros de gravedad de cada observación.
- *Inercia Inter* es la inercia de los centros de gravedad de cada nube parcial respecto al centro de gravedad de la global.

B) *El AFM se basa en la metodología del Análisis de Componentes Principales (ACP) y actúa en 2 etapas:*

- 1°. Realiza un ACP de cada grupo por separado, denominados *análisis parciales*. Cada grupo tiene asociada una subtabla X_c , de $t=1, \dots, T$ observaciones (en filas) por $i=1, \dots, I$ indicadores (en columnas); y se parte de $c=1, \dots, C$ grupos diferentes (véase Figura 1). De estos análisis se

estudian los valores propios, que ponen de manifiesto la dimensionalidad de la estructura interna de cada uno de ellos.

- 2º. Posteriormente, realiza el análisis simultáneo de los grupos, previamente ponderados, que se denomina *análisis global*. Esto es, analiza la *tabla múltiple X* que surge al yuxtaponer de forma horizontal los diferentes grupos de indicadores considerados. La ponderación seleccionada es el inverso del primer valor propio (λ_1^c) obtenido en los análisis separados de cada subtabla.

$$X = \left[\frac{1}{\sqrt{\lambda_1^1}} X_1, \frac{1}{\sqrt{\lambda_1^2}} X_2, \dots, \frac{1}{\sqrt{\lambda_1^c}} X_c, \dots, \frac{1}{\sqrt{\lambda_1^c}} X_c \right]$$

Esta ponderación mantiene la estructura de cada tabla, debido a que todas las variables han recibido la misma ponderación, pero consigue equilibrar la influencia de los grupos al primer factor, ya que la inercia máxima de cada una de las nubes de observaciones, definida por los diferentes grupos, vale 1 en cualquier dirección.

Los valores propios obtenidos en el análisis global proporcionan una primera información sobre las relaciones existentes entre los grupos analizados. El valor máximo que puede tomar el primer valor propio coincide con el número de grupos activos analizados. Se alcanza cuando la dirección global de mayor inercia es común a todos los grupos. El valor alcanzado por el resto de los valores propios depende de la estructura interna de cada una de las tablas.

La existencia de factores comunes puede ser detectada, igualmente, a través del cálculo de las correlaciones entre los factores globales y los factores del mismo orden correspondientes a cada uno de los grupos analizados, obtenidos en los respectivos análisis parciales. Cuando la correlación es fuerte, el factor global traduce una estructura que está presente en todos los grupos. Si la correlación es alta para uno solo de estos grupos, se tratará de un factor específico de dicho grupo. Esta fase de la interpretación es importante ya que el interés del análisis simultáneo de varios grupos de variables radica en el estudio de las comunalidades que éstos mantengan entre sí. En caso de observar la presencia de un grupo que está únicamente correlacionado fuertemente con

factores que le son específicos, se procederá a la eliminación del mismo como elemento activo del análisis, debido a las diferencias respecto al resto de los grupos.

El AFM proporciona, además, medidas globales de relación entre los grupos, basadas en el coeficiente de correlación vectorial entre matrices RV de Escoufier (Robert y Escoufier, 1976). Este coeficiente se obtiene a partir de los coeficientes de correlación lineal entre dos variables cualesquiera. Su valor está comprendido entre 0 (no existe relación entre los indicadores de los contextos considerados) y 1 (las estructuras de los grupos comparados son iguales).

C) El AFM proporciona las siguientes *representaciones en los planos factoriales del análisis global*:

- Representación, en un mismo espacio, de las observaciones caracterizadas por el conjunto de indicadores (punto medio) y por cada uno de los grupos (puntos parciales). La proximidad o distanciamiento entre los puntos correspondientes a la misma observación refleja la similitud o diferencia, respectivamente, de su comportamiento en los diferentes contextos.
- Representación de los indicadores ponderados de todos los grupos (contextos) considerados. Este gráfico permite observar el posicionamiento de los puntos variable de los diferentes contextos. El análisis de estas posiciones permitirá estudiar la relación entre el mismo indicador medido en los distintos contextos considerados y, por tanto, proporciona otra visión de las similitudes y/o diferencias entre estos últimos.
- Representación de los grupos de indicadores, cada uno de ellos representado a través de un solo punto. La coordenada de un grupo sobre un factor es la inercia acumulada del grupo sobre el eje del AFM. Indica los grupos que han determinado en mayor medida los factores y, en consecuencia, el número de factores comunes a los diversos contextos comparados. Además, el distanciamiento de un punto de los demás indica que el grupo correspondiente es el más alejado del compromiso, lo que permite detectar qué contexto mantiene una estructura más alejada de la común y, junto con otras medidas (comentadas anteriormente), valorar si merece ser estudiado por separado.

Las evidencias sobre la idoneidad y versatilidad de esta metodología en el análisis exploratorio comparativo de varias series temporales medidas en diferentes contextos y estructuradas como tabla múltiple se presentan, a partir de una aplicación empírica, en la siguiente sección.

3.- ANÁLISIS EXPLORATORIO DE ESTRUCTURAS TEMPORALES MÚLTIPLES. UNA APLICACIÓN

En esta sección se detallan e interpretan los resultados obtenidos al aplicar la técnica del AFM a datos reales correspondientes al sector turístico. La elección de esta información está apoyada en las siguientes razones: por un lado, la importancia y actualidad de los estudios relacionados con este sector. Y, por otro lado, la fiabilidad y la amplitud de los datos relacionados con el mismo, disponibles a través de la página web del Instituto de Estudios Turísticos (y del Instituto Nacional de Estadística). El programa estadístico utilizado para la obtención de resultados ha sido *Système Pour L'Analyse des Données* (SPAD, versión 5.5).

El turismo se ha confirmado como la locomotora de nuestra economía. Según las últimas proyecciones de la Alianza por la excelencia Turística (Exceltur-2015), el turismo mantendrá su capacidad motora sobre la economía española, tanto en términos de actividad económica como en generación de empleo en 2015; con una perspectiva que supera en 0,5 p.p. la previsión del 3,1% del PIB (estimada recientemente por el Banco de España y el Fondo Monetario Internacional para el conjunto de la economía).

En este trabajo, con carácter meramente ilustrativo, se han seleccionado tres indicadores estadísticos del ámbito del turismo desde el punto de vista de la demanda. Concretamente, se va a trabajar con las variables: Entradas de turistas según comunidad autónoma de destino principal (correspondiente a la Encuesta de Movimientos Turísticos en Fronteras, FRONTUR), Gasto medio de los turistas según destino principal y Estancia media de los turistas según destino principal (correspondientes a la Encuesta de Gasto Turístico, EGATUR). En las encuestas mencionadas, estos indicadores están medidos para siete contextos geográficos: Andalucía, Baleares, Canarias, Cataluña, Comunidad de Madrid, Comunidad Valenciana y resto de comunidades autónomas (CCAA). El período elegido, datos mensuales de los últimos 10 años (enero 2006 - septiembre 2015), recoge hitos importantes en la evolución

económica (final de una larga etapa de crecimiento, depresión y recuperación incipiente).

La información analizada está estructurada, por tanto, en tres dimensiones y constituye una tabla múltiple. Siguiendo la Figura 1, consta de siete subtablas X_c , $c=1, \dots, 7$, cada una de ellas asociada a uno de los ámbitos geográficos antes mencionados, respectivamente. Estas matrices de datos tienen en filas las observaciones mensuales, $t=enero\ 2006, \dots, septiembre\ 2015$, y en columnas los tres indicadores seleccionados, cuya nomenclatura resumida que será utilizada en el resto del artículo es $i=Turistas, Gasto\ y\ Estancia$, respectivamente.

El objetivo es obtener gráfica y numéricamente un estudio detallado de las similitudes y de las diferencias que los siete contextos geográficos presentan en lo que respecta a la relación entre los tres indicadores medidos y su evolución temporal en el periodo analizado. El análisis exploratorio desde la óptica de tabla múltiple y a través del AFM va a permitir contestar, entre otras, a las siguientes preguntas: ¿la trayectoria turística, desde el punto de vista de la demanda, ha sido similar en los siete territorios considerados? ¿Cuáles mantienen las mayores similitudes? ¿Y las mayores diferencias? ¿A qué indicadores se deben estas semejanzas/diferencias? ¿En qué unidades temporales se han producido? Además, esta metodología también va a reflejar la existencia de tendencias y de componentes estacionales.

Un estudio visual, comparativo, basado en el gráfico habitual de las series temporales supondría una ardua tarea en el caso que nos ocupa, ya que disponemos de 21 series (tres para cada uno de los 7 contextos espaciales) y las conclusiones alcanzadas tendrían un carácter demasiado generalista. Como se va a poder comprobar, la nueva óptica que se propone en este trabajo, el análisis factorial exploratorio de una tabla múltiple, proporciona una visión más detallada y, por ello, más completa de la estructura de los datos, desde un punto de vista tanto global como parcial. Esta novedosa alternativa que permite el AFM no se basa únicamente en instrumentos gráficos, ya que ofrece numerosas medidas numéricas que permiten cuantificar y, por tanto, justificar todas y cada una de las afirmaciones realizadas al finalizar el estudio.

A continuación se presentan e interpretan los principales resultados obtenidos:

- Plano principal de los grupos de variables (véase Figura 2): esta herramienta visual del AFM, en el que se proyecta cada grupo de variables como un punto, es de gran utilidad, ya que resume en un gráfico las similitudes y diferencias entre los grupos estudiados. A partir de la proximidad o alejamiento de los distintos puntos, se puede tomar la decisión de analizar alguno (o varios) de los grupos de manera aislada (eliminandolo de este análisis simultáneo). En el caso que nos ocupa, se comprueba que hay varios núcleos geográficos con comportamientos peculiares: Baleares, Madrid y Canarias. El resto de territorios considerados se proyectan con mayores proximidades, por lo que presentan una realidad más parecida en lo que a la relación y evolución de los indicadores de turismo se refiere y que habrá que poner de manifiesto a través del resto de medidas. Dado el carácter ilustrativo de este estudio empírico no se va a eliminar ninguno de los grupos considerados, con el objetivo de poner de manifiesto la utilidad de la metodología seleccionada para reflejar y medir las diferencias que estos mantienen frente al comportamiento medio.

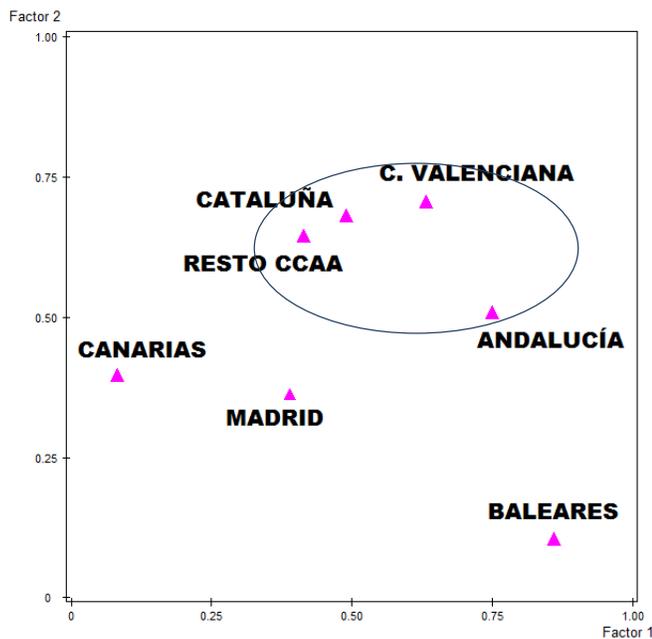


Figura 2: Plano principal de los grupos.

El AFM proporciona herramientas numéricas, asociadas a este gráfico, que miden y, por tanto, confirman estos resultados: las coordenadas y las contribuciones de los grupos en los diferentes ejes del análisis global (véase Tabla 1) y los coeficientes RV de relación entre grupos (véase Tabla 2).

	<i>Coordenadas</i>					<i>Contribuciones</i>				
	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F4</i>	<i>F5</i>
GRUPO 1 - ANDALUCÍA	0,75	0,51	0,08	0,01	0,02	20,7	15,0	7,10	1,00	3,60
GRUPO 2 - BALEARES	0,86	0,11	0,10	0,10	0,05	23,7	3,10	8,40	3,90	9,60
GRUPO 3 - CANARIAS	0,08	0,40	0,45	0,19	0,01	2,30	11,7	38,2	28,8	2,50
GRUPO 4 - CATALUÑA	0,49	0,68	0,09	0,04	0,03	13,5	20,0	7,40	6,50	6,20
GRUPO 5 - C. VALENCIANA	0,63	0,71	0,09	0,04	0,01	17,5	20,7	8,10	6,20	2,20
GRUPO 6 - MADRID	0,39	0,36	0,15	0,18	0,23	10,8	10,6	12,9	27,2	44,7
GRUPO 7 - RESTO CCAA	0,42	0,65	0,21	0,18	0,16	11,5	18,9	17,8	26,5	31,2

Tabla 1: *Coordenadas y contribuciones de los grupos en los cinco primeros factores del análisis global.*

Las mayores coordenadas y contribuciones sobre uno o varios de los ejes indican los grupos que han determinado en mayor medida dichos factores. La proximidad/distancia entre estos valores para varios grupos es reflejo de sus semejanzas/diferencias. Así, por ejemplo, destacar, por un lado, la proximidad entre los valores de Cataluña y Resto de CCAA (confirmando su cercanía en el plano anterior); y, por otro lado, la distancia que presenta Baleares (confirmando su posición alejada de los demás en el plano anterior por su mayor contribución al primer factor como consecuencia de su estructura interna más fuerte y una realidad turística reflejada prácticamente en una sola dimensión, tanto del análisis global como del análisis parcial; esto último se comprueba con los datos de la Tabla 3).

Además, a partir de la Tabla 1 ya se puede concluir que el análisis se va a centrar en los dos primeros factores, en los que las contribuciones, aún habiendo diferencias entre los grupos, son destacables en general; mientras que los factores a partir del tercero reflejan comportamientos particulares, correspondientes a alguno de los grupos (por ejemplo, destaca el peso de Baleares en el tercer factor), por lo que van a ser obviados para el resto de resultados.

Los coeficientes RV (véase Tabla 2), medida de relación global entre grupos, ponen de manifiesto que se están analizando simultáneamente siete grupos con similitudes, pero también con realidades particulares relativas a la relación y a la evolución de los indicadores del turismo seleccionados. Este estudio exploratorio va a vislumbrar las semejanzas y las diferencias más

destacables. Así, se confirma los comportamientos más alejados del comportamiento medio tanto de la Comunidad de Canarias (con los menores valores), como de la Comunidad de Madrid (con coeficientes también muy bajos en general).

	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>	<i>G5</i>	<i>G6</i>	<i>G7</i>
GRUPO 1 - ANDALUCÍA	1,00						
GRUPO 2 - BALEARES	0,51	1,00					
GRUPO 3 - CANARIAS	0,21	0,10	1,00				
GRUPO 4 - CATALUÑA	0,50	0,49	0,21	1,00			
GRUPO 5 - C. VALENCIANA	0,65	0,48	0,27	0,60	1,00		
GRUPO 6 - MADRID	0,44	0,24	0,11	0,15	0,33	1,00	
GRUPO 7 - RESTO CCAA	0,39	0,34	0,10	0,59	0,40	0,22	1,00

Tabla 2: *Coefficientes RV de relación entre los grupos.*

Una vez analizados estos resultados del análisis global y siendo conscientes de la decisión de no prescindir de ninguno de los grupos/territorios considerados inicialmente (y de las consecuencias que de ella se derivan), se va a proceder a interpretar el resto de medidas.

- El AFM comienza con el análisis de cada tabla por separado. En la Tabla 3 se recoge de forma resumida la información de estos Análisis de Componentes Principales (ACP) previos, a través de los primeros valores propios. Éstos ponen de manifiesto un resultado común a todos los grupos: la existencia de un primer eje preponderante que refleja en torno al 60% de su inercia. Hay que recordar que cada uno de los grupos está integrado por tres variables: los tres indicadores turísticos seleccionados. Como ya se ha mencionado en la sección anterior, el inverso del primer valor propio va a ser la ponderación seleccionada para equilibrar la influencia de los distintos grupos. En este caso, la ponderación es muy similar. Mencionar el mayor valor que presenta la Comunidad de Baleares, con un 70% de información recogida en el primer factor, lo cual indica que es la comunidad autónoma con estructura interna más fuerte, esto es, con mayor relación lineal entre sus indicadores (resultado ya comentado y que se reflejará más adelante a través de otras medidas).

	<i>Valores propios</i>			<i>Porcentaje de inercia</i>		
	<i>F1</i>	<i>F2</i>	<i>F3</i>	<i>F1</i>	<i>F2</i>	<i>F3</i>
GRUPO 1 - ANDALUCÍA	1,819	0,853	0,328	60,6	28,4	10,9
GRUPO 2 - BALEARES	2,101	0,650	0,250	70,0	21,7	8,30
GRUPO 3 - CANARIAS	1,840	0,739	0,421	61,3	24,6	14,0
GRUPO 4 - CATALUÑA	1,957	0,743	0,300	65,2	24,8	10,0
GRUPO 5 - C. VALENCIANA	1,724	0,993	0,283	57,5	33,1	9,40
GRUPO 6 - MADRID	1,853	0,914	0,233	61,8	30,5	7,80
GRUPO 7 - RESTO CCAA	1,840	0,739	0,421	54,7	36,3	9,10

Tabla 3: *Valores propios y porcentajes de inercia de los análisis parciales.*

- El AFM continua con el análisis de la tabla global resultante de la yuxtaposición de las 21 series temporales estructuradas en siete tablas asociadas a cada uno de los territorios explorados (previamente ponderadas). Los valores propios y la razón entre la inercia inter y la inercia total (véanse las Tablas 4 y 5, respectivamente) son las medidas que hay que interpretar inicialmente en esta nueva etapa. Ambas ponen de relieve la importancia similar de los dos primeros ejes factoriales, reflejando cada uno de ellos un porcentaje de información muy similar, próximo al 30%, respectivamente. Estos resultados confirman la existencia de similitudes entre los siete núcleos geográficos estudiados, pero también de notables diferencias entre los mismos. Señalar que, en el caso de que se tratara de grupos muy semejantes, el primer valor propio del análisis global alcanzaría un valor próximo a 7 y los ejes que reflejarían estas similitudes tendrían una razón inercia inter/inercia total próxima a 1.

El resto de los ejes, a partir del tercero, recogen, como no podía ser de otra manera, inercia residual que corresponderá a las peculiaridades de los grupos con un comportamiento con mayores divergencias respecto al comportamiento medio (como los casos de Canarias y Madrid, ya mencionados). Por ello, confirmamos que los siguientes resultados se van a centrar en estos dos primeros factores, esto es, en el plano principal.

	<i>Valor propio</i>	<i>Porcentaje de inercia</i>
1	3,6223	31,69
2	3,4101	29,84
3	1,1684	10,22
4	0,6747	5,90

Tabla 4: *Valores propios del análisis global.*

<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 1</i>	<i>Factor 2</i>
0,64	0,54	0,24	0,14

Tabla 5: Razón inercia inter/inercia total en los cuatro primeros factores.

- El plano principal (véase Figura 3) del análisis global en el que aparecen proyectadas las 21 series temporales, representada cada una por un punto, refleja los siguientes resultados:
 - La evolución del indicador “entradas de turistas” queda recogido en el eje horizontal, con un comportamiento muy similar en seis de los territorios estudiados (siendo Canarias la excepción). Concretamente, los cuadrantes 1 y 4 pondrán de manifiesto los meses en los que este indicador toma valores superiores a la media del periodo analizado, mientras que en la parte negativa (cuadrantes 2 y 3) se proyectarán los meses con el menor número de turistas. Esto es, la evolución del número de turistas en Andalucía, Baleares, Cataluña, Comunidad de Madrid, Comunidad Valenciana y resto de comunidades autónomas tiene claramente un comportamiento estacional con picos y valles. Canarias, por su parte, no mantiene este comportamiento estacional y su proyección más centrada es reflejo de esta particularidad. Señalar que Madrid, con una menor coordenada en este factor, también mantiene un comportamiento menos estacional que el resto de contextos geográficos.
 - El eje vertical, segundo factor, contrapone los indicadores “estancia media” (parte positiva – cuadrantes 1 y 2) y “gasto medio” (parte negativa – cuadrantes 3 y 4): se trata de dos indicadores turísticos con evoluciones distintas, contrarias, y sin un componente estacional destacable en el periodo analizado. Este comportamiento corresponde a todas las comunidades autónomas analizadas, con la excepción de Baleares. Esta comunidad insular presenta una evolución del gasto similar a la del indicador “entrada de turistas”, esto es, con el mismo comportamiento estacional. La relación positiva entre los indicadores de “entrada de turistas” y “gasto medio” es la razón por la que la Comunidad Balear tiene una estructura interna más fuerte. Este resultado queda reflejado en su mayor contribución a la formación del primer eje (véase Tabla 1) y en su mayor primer valor propio (véase Tabla 3), ya comentados anteriormente.

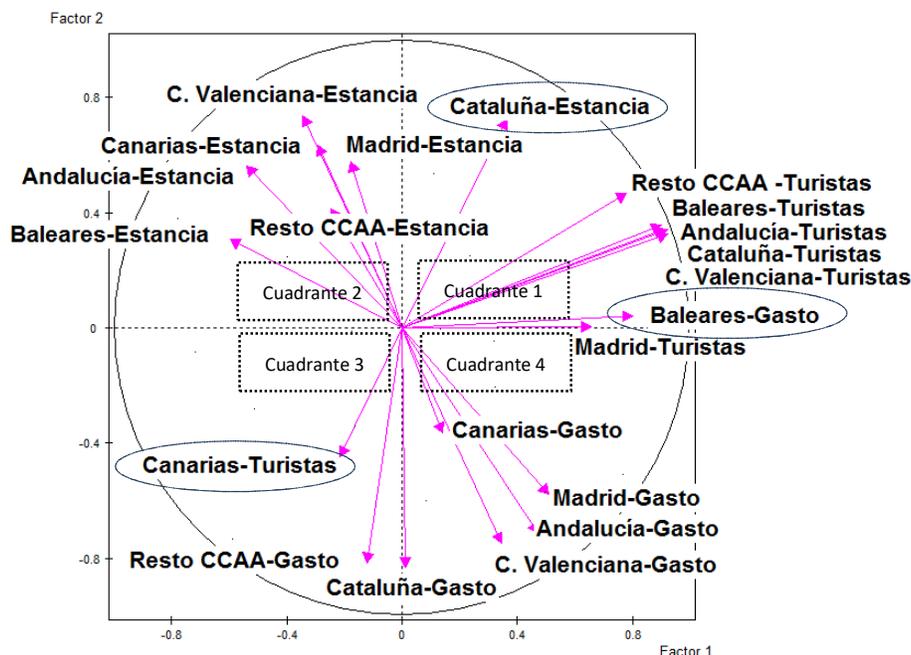


Figura 3: Plano principal del AFM global: los indicadores.

- El plano principal (véase Figura 4) del análisis global en el que aparecen proyectados los denominados “puntos medios” ayuda a completar las conclusiones alcanzadas con el plano anterior. En este caso, los puntos reflejan la posición media de las observaciones temporales (enero 2006 – septiembre 2015) teniendo en cuenta conjuntamente las siete tablas analizadas. Así, se puede concluir:
 - La parte positiva del primer factor (cuadrantes 1 y 4) permite visualizar los meses de mayor afluencia de turistas: mayo, junio, julio, agosto y septiembre. Además, se puede comprobar que esta realidad es más notable en los últimos años (crecimiento cronológico de las coordenadas), confirmando que en el sector turístico la senda del crecimiento económico se ha consolidado. En la parte negativa de este factor horizontal (cuadrantes 2 y 3), quedan reflejados los meses en los que la afluencia de turistas es menor: diciembre, enero y febrero, principalmente.
 - En lo que respecta al eje vertical, segundo factor (con una importancia similar al anterior), la trayectoria temporal ascendente desde la parte positiva del mismo (observaciones de los primeros años de las series) hacia la parte negativa (observaciones de los últimos años de las series)

refleja tanto la tendencia positiva del gasto medio en el periodo estudiado como la tendencia decreciente de la estancia media.

- Estos comportamientos corresponden a casi todas las unidades geográficas consideradas. No obstante, existen comportamientos peculiares que también quedan reflejados. Así, la posición del indicador de estancia media en Cataluña en el primer cuadrante refleja que, para la citada comunidad autónoma, esta variable sí que tiene un componente estacional (con ciertas similitudes respecto al comportamiento del número de turistas) con valores superiores a la media en los meses de verano.

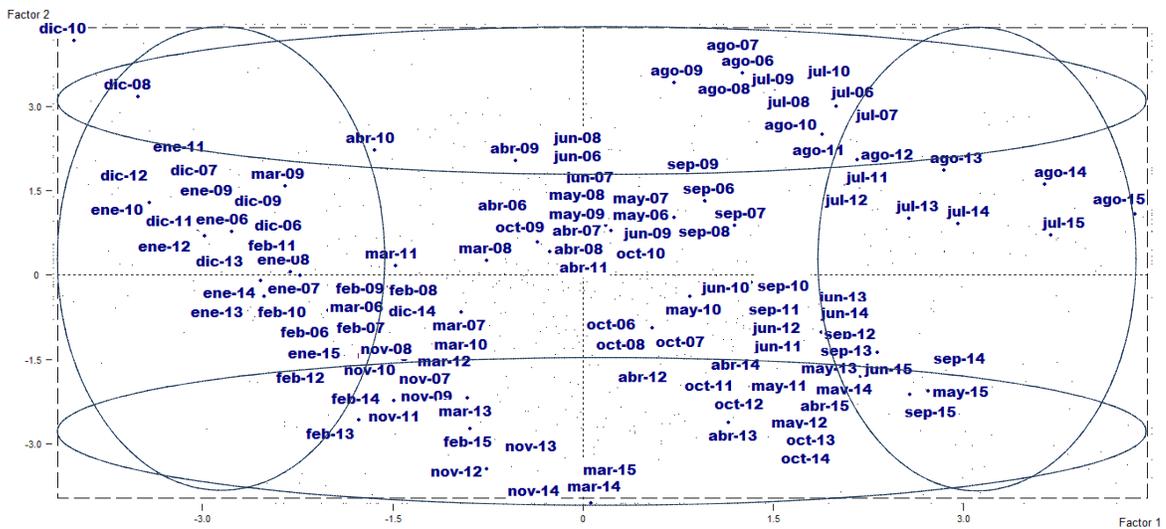


Figura 4: Plano principal del AFM global- puntos medios: observaciones temporales.

- El AFM también proporciona un gráfico en el que se pueden proyectar junto a los puntos medios los puntos parciales, observaciones temporales asociadas a cada grupo/región. De este gráfico merecen ser analizados con más detalle aquéllos puntos con mayor inercia intra, esto es, aquéllos que presentan cierta diferencia entre el comportamiento parcial, asociado a un grupo concreto, y su comportamiento medio. En esta aplicación la observación temporal con mayor inercia intra corresponde al mes de enero del año 2010. Para analizar el por qué de este resultado y dado el carácter ilustrativo de la aplicación, en la Figura 5 se han proyectado únicamente los ocho puntos asociados a este momento de tiempo. El análisis de sus posiciones a lo largo del primer factor, y teniendo en cuenta la

interpretación que del mismo se ha hecho, permite concluir que los responsables de esta inercia son los puntos asociados a las dos comunidades autónomas insulares. El mes de enero, como ya se ha comentado, es uno de los meses con menor número de turistas, en general, para todos los años considerados y para todas las zonas estudiadas. No obstante, esta realidad es más notable en Baleares, con el peor comportamiento, mientras que Canarias alcanza en esta observación temporal un número de turistas más próximo al valor medio (volvemos a insistir aquí en la ausencia de comportamiento estacional del indicador en cuestión en esta comunidad). El lector puede comprobar las importantes posibilidades que este plano ofrece y ser consciente de que hay que centrar los esfuerzos en estudiar detalladamente aquellas observaciones con inercia intra destacable, ya que la inercia intra reducida solo confirma las similitudes ya estudiadas anteriormente.

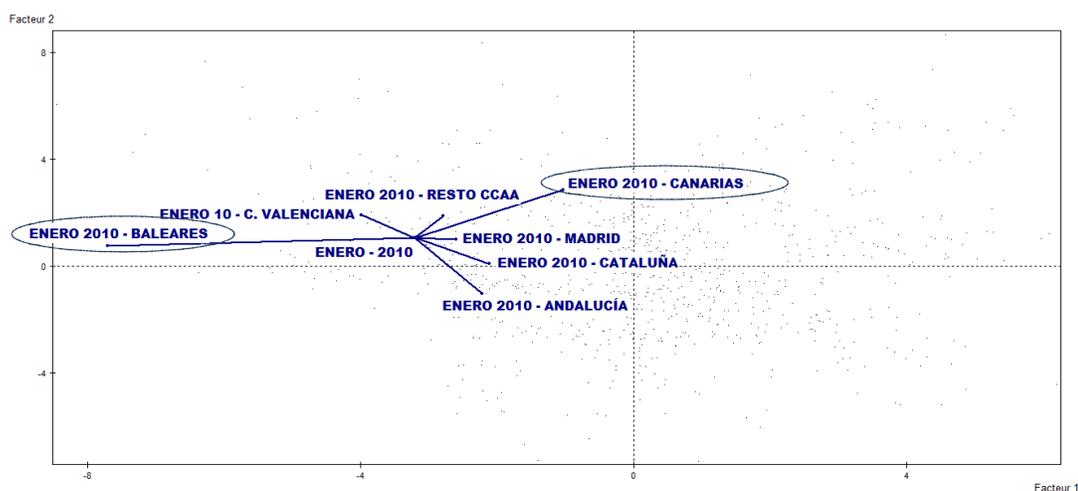


Figura 5: Plano principal del AFM global- punto medio y puntos parciales: extracto para la observación correspondiente a enero del año 2010.

4.- CONCLUSIONES Y NUEVAS LÍNEAS DE INVESTIGACIÓN

En este trabajo se ha puesto de manifiesto la versatilidad del enfoque de análisis factorial de tablas múltiples para explorar, en términos comparativos, numerosas series temporales yuxtapuestas y estructuradas en diferentes grupos definidos atendiendo a diversos criterios (geográficos, por ejemplo, como el utilizado en esta investigación).

Concretamente, ha quedado patente, a través de una aplicación, la idoneidad del Análisis Factorial Múltiple, como técnica alternativa y complementaria del habitual análisis gráfico usado para la descripción de las trayectorias y sus relaciones. El AFM, además de ofrecer al investigador imágenes simplificadas en las que se pueden proyectar de forma conjunta unidades temporales, variables y contextos territoriales en los que se miden las anteriores, proporciona instrumentos numéricos que ayudan a cuantificar e interpretar las semejanzas y las diferencias entre las trayectorias temporales estudiadas en los diferentes territorios considerados. Las ventajas más destacables de este nuevo enfoque son:

- Las herramientas visuales que proporciona la técnica del AFM se caracterizan por aportar de manera simplificada una gran riqueza interpretativa, dados los objetivos planteados.
- Uno de los gráficos (Figura 2: plano principal de los grupos) ofrece una clara imagen de las similitudes o divergencias entre los grupos definidos, en el que cada subtabla correspondiente está representada por un solo punto, a partir del cual se pueden tomar decisiones sobre la eliminación de algunos de los contextos comparados debido a su comportamiento diferente al comportamiento medio.
- En otro de los gráficos (Figura 3: plano principal del AFM global: los indicadores), se proyectan todas las series temporales analizadas a partir del cual se alcanzan las primeras conclusiones, no sólo relativas a las relaciones entre ellas y la existencia de un posible componente estacional, sino que también se pueden identificar cuál o cuáles de los indicadores son responsables en mayor medida de las diferencias y/o semejanzas entre los grupos.
- La proyección de las unidades temporales observadas (Figura 4: plano principal del AFM global – puntos medios: observaciones temporales) completa el gráfico anterior, imagen que, además de confirmar algunas de las conclusiones obtenidas y concretarlas en períodos, permite ampliar los resultados con la visión de tendencias en las trayectorias.
- En la interpretación de los gráficos referenciados no existe el riesgo de caer en subjetividades, ya que la metodología propuesta proporciona una serie de indicadores numéricos que permiten cuantificar las proximidades o alejamientos

de las diferentes proyecciones, matizando en qué contextos espaciales se producen y qué serie/s concreta/s es/son la/s responsable/s de los comportamientos más o menos homogéneos.

El estudio empírico realizado ha permitido a través de resultados concretos, con carácter meramente ilustrativo, dejar constancia de todas estas ventajas. El ámbito y las dimensiones de la información inicial han resultado suficientes para evidenciar la potencialidad de la propuesta. Se han seleccionado tres indicadores estadísticos relacionados con el turismo (desde el punto de vista de la demanda), debido a la importancia cada vez mayor que este sector está teniendo en la economía española y en su recuperación. Se ha trabajado con datos mensuales correspondientes a un período lo suficientemente amplio como para reflejar diferentes fenómenos: hitos importantes en la evolución económica de nuestra economía y diferentes componentes de las series temporales. Además, las fuentes secundarias consultadas proporcionan esta información desagregada para seis comunidades autónomas de forma individual y agregada para el resto de comunidades autónomas, lo que ha permitido la estructuración de la información en una tabla múltiple integrada por siete subtablas. El AFM de esta información ha permitido dar respuesta a las preguntas que se formulaban al inicio de la sección 3 del trabajo: ¿la trayectoria turística, desde el punto de vista de la demanda, ha sido similar en los siete territorios considerados? ¿Cuáles mantienen las mayores similitudes? ¿Y las mayores diferencias? ¿A qué indicadores se deben estas semejanzas/diferencias? ¿En qué unidades temporales se han producido?

Para finalizar el trabajo, señalar que las posibilidades de aplicación de esta propuesta de análisis comparativo de series temporales son diversas, entre las que merecen mención:

- El análisis de series temporales se suele asociar a información de carácter cuantitativo. Sin embargo, la propuesta de este trabajo se puede aplicar a datos de distinta naturaleza. Como se ha comentado en la sección 2, en la presentación de la metodología, el AFM permite el estudio de variables nominales (e incluso de datos mixtos, siempre que las variables que integran cada grupo sean de la misma naturaleza).
- Existen otras técnicas que permiten el análisis factorial de tablas estructuradas en grupos, con filosofía distinta, complementaria, a la del AFM. En este sentido,

una posible aplicación podría ser comparar los resultados de varias de estas técnicas a partir del mismo conjunto de series temporales, comprobando las ventajas y los inconvenientes de cada una de ellas en estudios de naturaleza dinámica. Esta línea de trabajo sería similar a la seguida en el libro de Dazy y Le Barzic (1996), en el que los conjuntos de datos también son de carácter dinámico, pero el tiempo está asociado a la estructuración de las tablas (se comparan varias tablas, cada una de ellas referida a un momento de tiempo).

Todo lo expuesto anteriormente permite afirmar que el análisis simultáneo de series temporales yuxtapuestas a través del análisis factorial de tablas múltiples es un campo de investigación aplicada novedoso y con diversas posibilidades que merecen ser exploradas.

REFERENCIAS

- Abascal, E.; García-Lautre, I. y Landaluce M.I. (2004). “Análisis de la evolución a través de encuestas. Trayectoria electoral de las comunidades autónomas españolas en el periodo 1977-2004”. *Metodología de Encuestas*, Vol. 6, Nº 2, pp. 147-162.
- Abascal, E.; García-Lautre, I. y Landaluce, M.I. (2006). “Multiple factor analysis of mixed tables of metric and categorical data”. En M. Greenacre y J. Blasius (eds.): *Multiple Correspondence Analysis and related Methods*. Chapman & Hall/CRC, pp 351-367.
- Dazy, F. y Le Barzic, J.F. (1996). “L’Analyse des Données Evolutives”. Technip. Paris.
- Escofier, B. y Pagès J. (1986). “Le traitement des variables qualitatives et tableaux mixtes par analyse factorielle multiple”. *Data Analysis and Informatics*, IV(2), pp. 179-191.
- Escofier, B. y Pagès, J. (1992). “Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación”. Servicio editorial de la Universidad de País Vasco.
- Escofier, B. y Pagès, J. (1994). “Multiple factor analysis (AFMULT package)”. *Computational Statistics & Data Analysis*, 18(1), pp. 121-140.

- Exceltur (2015). Alianza para la excelencia turística: *Informe Perspectivas Turísticas*, N° 54.
- Fernández, K.; Landaluce, M.I. y Modroño, J.I. (2013). “Nuevo procedimiento metodológico para el análisis exploratorio de una tabla estructurada en diversos conjuntos de individuos”. *Estadística Española*, Vol. 55, N° 182, pp. 305-322.
- García-Lautre, I. y Abascal, E. (2003). “Una metodología para el estudio de la evolución de variables latentes. Análisis de las infraestructuras de carreteras de las comunidades autónomas (1975-2000)”. *Estadística Española*, Vol. 45, N° 153, pp. 193-210.
- Goitisoló, B. (2001). “El análisis simultáneo. Propuesta y aplicación de un nuevo método de análisis factorial de tablas de contingencia”. Tesis doctoral. Universidad del País Vasco.
- Landaluce, M.I. (1995). “Estudio de la estructura de gasto medio de las Comunidades Autónomas españolas. Una aplicación del Análisis Factorial Multiple”. Tesis doctoral. Universidad del País Vasco.
- Robert, P. y Escoufier, Y. (1976). “A Unifying Tool for Linear Multivariate Statistical Methods: The RV-Coefficient”. *Applied Statistics*, 25(3), pp. 257-265.
- SPAD v5.5 (2000). “Système Portable d’Analyse des Données Numeriques”, CISIA, Montreuil, France.