

UNIVERSIDAD DE EXTREMADURA

FACULTAD DE CIENCIAS DE LA DOCUMENTACIÓN Y LA COMUNICACIÓN

DEPARTAMENTO DE INFORMACIÓN Y COMUNICACIÓN



**NUEVO MAPA CONCEPTUAL DE LA
INVESTIGACIÓN CIENTÍFICA EN CIENCIAS DE LA
ALIMENTACIÓN BASADO EN TÉCNICAS
CIENCIOMÉTRICAS (2003-2014): ESTUDIO DEL
CASO MUNDIAL Y ESPAÑOL**

TRABAJO DE FIN DE MÁSTER

Trabajo presentado por D. Jesús Blázquez Ruiz para la obtención del título de Máster Universitario en Investigación en Ciencias Sociales y Jurídicas, bajo la dirección del profesor D. Vicente Pablo Guerrero Bote

BADAJOS
2017

“Nuevo mapa conceptual de la investigación científica en ciencias de la alimentación basado en técnicas cuantitativas (2003-2014): Estudio del caso mundial y español”

Trabajo presentado por D. Jesús Blázquez Ruiz para la superación de la asignatura *Trabajo Fin de Máster* (Código 400763), del título de *Máster Universitario en Investigación en Ciencias Sociales y Jurídicas* (curso 2015/2016), bajo la dirección de D. Vicente Pablo Guerrero Bote, profesor del Departamento de Información y Comunicación de la Universidad de Extremadura.

El alumno

Vº Bº del Director

Fdo. Jesús Blázquez Ruiz.

Fdo. Vicente Pablo Guerrero Bote.

“Nuevo mapa conceptual de la investigación científica en ciencias de la alimentación basado en técnicas cuantitativas (2003-2014): Estudio del caso mundial y español”

Resumen

Se analiza un conjunto de palabras clave para revelar patrones de publicación en el campo de las Ciencias de la Alimentación en el periodo 2003-2014, con el objetivo de determinar la estructura temática del campo. Con este fin, los registros de la materia específica “Food Science” fueron recuperados primero de Scopus, para después procesar sus palabras clave resolviendo los problemas obvios de sinonimia y limitando el estudio a los más utilizados. Estas palabras clave se agruparon en comunidades temáticas o clústeres a través de la técnica cuantitativa conocida como, análisis de copalabras. El análisis reveló una estructura compuesta por 5 comunidades principales o clústeres que engloban a su vez a otros 18 específicos. Seguidamente se analizó la estructura de los clústeres, su producción, su impacto científico y su evolución temporal, en primer lugar para el caso mundial y en segundo lugar para el caso español. Los resultados muestran un claro aumento en la producción científica relacionada con las Ciencias de la Alimentación en ambos casos. Siendo el clúster número 5, correspondiente a los antioxidantes en alimentos, el que más atención recibe con un mayor impacto científico y con mayor crecimiento en el período, en los 2 casos estudiados.

Palabras clave: Ciencias de la Alimentación, Análisis temático, Estructura de las Ciencias de la Alimentación, Análisis de copalabras, Cuantitativa.

Abstract

It's analyzed a set of keywords to reveal patterns of publication in the field of Food Sciences in the period 2003-2014, with the objective of determining the field structure. To this end, the records of the specific subject "Food Science" were first recovered from Scopus, then processed their keywords by solving the obvious problems of synonymy and limiting the study to the most used ones. These keywords were grouped into thematic communities or clusters through the scientometric technique known as co-words analysis. The analysis revealed a structure composed of 5 main communities or

clusters that in turn encompasses 18 specific ones. Next, the structure of the clusters, their production, their scientific impact and their temporal evolution were analyzed, firstly for the world case and secondly for the Spanish case. The results show a clear increase in the scientific production related to the Food Sciences in both cases. Being cluster number 5, corresponding to the antioxidants in food, the one that receives the most attention with a greater scientific impact and with greater growth in the period, in the 2 cases studied.

Keywords: Food Science, Thematic Analysis, Food Science Structure, Co-word Analysis, Scientometrics.

Agradecimientos

Vicente, gracias por todo lo enseñado, tanto a lo largo de la carrera como finalmente en este trabajo de fin de máster. Sin duda, esto hubiera sido imposible sin contar con tu conocimiento y buen hacer. Gracias por tu paciencia, por haberme resuelto cualquier duda casi antes de que te la preguntara, y por haber estado siempre pendiente de que esto saliera adelante y saliera bien.

Mamá y Papá, gracias por ser tan insistentes y por empujarme a seguir cuando yo me dispersaba un poco.

Leonor, gracias por ayudarme con los gráficos y, sobre todo, por estar ahí.

A mis padres,

ÍNDICE GENERAL

Índice de figuras	8
Índice de tablas	10
1. INTRODUCCIÓN.....	13
1.1. Las ciencias de la alimentación.....	13
1.2. Evaluación de la ciencia	14
1.3. Justificación del estudio.....	18
1.4. Objetivos del estudio	18
2. METODOLOGÍA.....	20
2.1. Fuentes	20
2.2. Recuperación y tratamiento de datos	21
2.3. Análisis y visualización de datos	22
3. RESULTADOS	27
3.1. Estructura temática de las ciencias de la alimentación	27
3.2. Las ciencias de la alimentación en el mundo.....	29
3.3. Las ciencias de la alimentación en España	61
3.4. Comparativa del caso español y mundial a través de los resultados relativos de España con respecto al mundo	88
4. CONCLUSIONES.....	98
5. BIBLIOGRAFÍA	103
6. ANEXOS	107

Índice de figuras

Figura 1: Vista general del mapa de copalabras distinguiendo los 5 clústeres R1 (Fuente: Elaboración propia).	30
Figura 2: Vista de densidad del mapa general de copalabras. En rojo aparecen las zonas donde hay keywords que aparecen en muchos documentos (Fuente: Elaboración propia).	31
Figura 3: Porcentaje de producción de los clústeres R1, respecto al total de producción en FS (Fuente: Elaboración propia).	34
Figura 4: Mapa de copalabras coloreado esta vez en función del IN obtenido por los documentos etiquetados con las keywords del mapa (Fuente: Elaboración propia).	35
Figura 5: Zoom de la zona superior derecha del mapa de copalabras donde aparece el clúster 1 de R1 dedicado a la Composición de alimentos y nutrición (Fuente: Elaboración propia).....	37
Figura 6: Evolución temporal del IN en los clústeres R2 del clúster 1 de R1 (Fuente: Elaboración propia).	40
Figura 7: Zoom de la zona inferior izquierda del mapa de copalabras donde aparece el clúster 2 de R1 dedicado al Procesado y modificado de alimentos, y también el clúster 4 dedicado a la Conservación de alimentos y vida útil. (Fuente: Elaboración propia).	43
Figura 8: Evolución temporal del IN en los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).	46
Figura 9: Zoom de la zona central del mapa de copalabras donde aparece el clúster 3 de R1 dedicado a la Seguridad alimentaria. (Fuente: Elaboración propia).	48
Figura 10: Evolución temporal del IN en los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).	51
Figura 11: Evolución temporal del IN en los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).	54
Figura 12: Zoom de la zona superior izquierda del mapa de copalabras donde aparece el clúster 5 de R1 dedicado a los Antioxidantes en alimentos (Fuente: Elaboración propia).	56
Figura 13: Evolución temporal del IN en los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).	58
Figura 14: Porcentaje de los documentos excelentes en el 10% como “Research Guarantor” comparado con el IN de los 5 clústeres R1 (círculos discontinuos) y con los	

18 de R2 (círculos continuos). Los 3 círculos concéntricos de cada clúster se corresponden, respectivamente, con los indicadores Ndocc, %Exc, y %Exc1 (Fuente: Elaboración propia).	60
Figura 15: Vista general del mapa de copalabras de España distinguiendo los 5 clústeres R1 (Fuente: Elaboración propia)	63
Figura 16: Vista de densidad del mapa general de copalabras de España. En rojo aparecen las zonas donde hay keywords que aparecen en muchos documentos (Fuente: Elaboración propia).....	64
Figura 17: Porcentaje de producción de los clústeres R1 de España, respecto al total de producción en FS de España (Fuente: Elaboración propia).	67
Figura 18: Mapa de copalabras del caso español, coloreado esta vez en función del IN obtenido por los documentos etiquetados con las keywords del mapa (Fuente: Elaboración propia).	68
Figura 19: Evolución temporal del IN de España en los clústeres R2 del clúster 1 de R1 (Fuente: Elaboración propia).	71
Figura 20: Evolución temporal del IN de España en los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).	75
Figura 21: Evolución temporal del IN de España en los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).	79
Figura 22: Evolución temporal del IN de España en los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).	82
Figura 23: Evolución temporal del IN de España en los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).	85
Figura 24: Porcentaje de los documentos excelentes en el 10% como “Research Guarantor” comparado con el IN de los 5 clústeres R1 de España (círculos discontinuos) y con los 18 de R2 (círculos continuos). Los 3 círculos concéntricos de cada clúster se corresponden, respectivamente, con los indicadores Ndocc, %Exc, y %Exc1 (Fuente: Elaboración propia).	87
Figura 25: Vista general del mapa de copalabras donde los colores marcan la Citación Normalizada Relativa de España respecto al mundo (Fuente: Elaboración propia).	91
Figura 26: Evolución de la producción relativa de España respecto al mundo (Fuente: Elaboración propia).	92

Índice de tablas

Tabla 1: Estructura temática de las ciencias de la alimentación con sus 2 niveles de profundidad (Fuente: Elaboración propia).	28
Tabla 2: Tabla resumen de los 5 clústeres R1 (Fuente: Elaboración propia).	32
Tabla 3: Tabla resumen de los clústeres R2 del clúster 1 de R1 (Fuente: Elaboración propia).	38
Tabla 4: Periodos bursting de las keywords del clúster 1 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).	41
Tabla 5: Tabla resumen de los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).	44
Tabla 6: Periodos bursting de las keywords del clúster 2 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).	47
Tabla 7: Tabla resumen de los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).	49
Tabla 8: Periodos bursting de las keywords del clúster 3 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).	52
Tabla 9: Tabla resumen de los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).	53
Tabla 10: Periodos bursting de las keywords del clúster 4 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).	55
Tabla 11: Tabla resumen de los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).	57
Tabla 12: Periodos bursting de las keywords del clúster 5 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).	59
Tabla 13: Tabla resumen de los 5 clústeres R1 de España (Fuente: Elaboración propia).	66
Tabla 14: Tabla resumen de los clústeres R2 del clúster 1 de R1 de España (Fuente: Elaboración propia).	70
Tabla 15: Periodos bursting para el caso español de las keywords del clúster 1 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).	72
Tabla 16: Tabla resumen de los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).	74

Tabla 17: Periodos bursting para el caso español de las keywords del clúster 2 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).....	76
Tabla 18: Tabla resumen de los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).....	77
Tabla 19: Periodos bursting para el caso español de las keywords del clúster 3 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).....	80
Tabla 20: Tabla resumen de los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).....	81
Tabla 21: Periodos bursting para el caso español de las keywords del clúster 4 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).....	83
Tabla 22: Tabla resumen de los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).....	84
Tabla 23: Periodos bursting para el caso español de las keywords del clúster 5 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).....	86
Tabla 24: Tabla resumen de los clústeres R1 y R2 con los datos relativos de España respecto al mundo (Fuente: Elaboración propia).....	94

1. INTRODUCCIÓN

1. INTRODUCCIÓN

1.1. Las ciencias de la alimentación

Una de las prioridades de la ciencia es dar respuesta a las necesidades que se van generando en el mundo y en la sociedad. En la actualidad nos encontramos ante un problema de carácter primordial como es alimentar a todas las personas del planeta sabiendo que cada día somos más. Ante este problema, la ciencia ha respondido con las Ciencias de la Alimentación (FS¹ en adelante). La FS es la disciplina en que se estudia “*la estructura física, biológica y química de los alimentos; las causas del deterioro de los alimentos; y los conceptos subyacentes del procesado de alimentos*”² (IFT, 2016).

En 2015 éramos unos 7300 millones de personas, en 2030 seremos unos 8500 millones, en 2050 unos 9700 millones y en 2100 unos 11200 millones (Blackwell, 2015) o tal vez más, hasta unos 13200 millones señala Gerland et al. (2014). Teniendo en cuenta esto, nos damos cuenta que no se trata de una cuestión baladí el hecho de pensar en la FS como algo realmente necesario para el presente y el futuro. De hecho, en estos momentos, la FS está generando una gran cantidad de trabajos y estudios sobre su temática a nivel internacional.

Como demuestra el trabajo de Guerrero-Bote y Moya-Anegón, (2015) la producción científica en FS ha crecido a gran ritmo en la última década, incluso a un ritmo superior al resto de la producción científica mundial. Este hecho muestra lo relevante de esta disciplina en la actualidad y por tanto la necesidad de realizar trabajos que estudien este fenómeno y la disciplina FS en general.

Actualmente supone una gran importancia, y más supondrá en un futuro no muy lejano, todo lo que tiene que ver con la FS. Es por ello que se hace necesario organizar y orientar la investigación desarrollada en esta área, evitando así que los países que más trabajan en la actualidad en la disciplina impongan el rumbo a seguir, y determinen cuáles son los problemas prioritarios (Guerrero-Bote, Olmeda-Gómez, y Moya-Anegón, 2016). Además, toda esta investigación debe ser evaluada para conocer su impacto y relevancia

¹ Food Science.

² Traducción propia realizada a partir de la cita, “*the physical, biological, and chemical makeup of food; the causes of food deterioration; and the concepts underlying food processing*”, incluida en la web del IFT.

en el mundo y en la sociedad, y, en definitiva, para que el mundo académico pueda identificar el camino correcto a seguir, en este caso concreto, respecto a la FS.

Pero, ¿en qué consiste la evaluación de la ciencia?

1.2. Evaluación de la ciencia

Una vez que los trabajos científicos son publicados en alguna revista científica, aparecen recogidos junto con sus referencias en las distintas bases de datos bibliográficas facilitando su recuperación. De esta manera, las revistas y las bases de datos se convierten en protagonistas de la comunicación científica (Miguel, Chinchilla-Rodríguez, y Moya-Anegón, 2011). Otro aspecto importante que recogen estas bases de datos científicas son las palabras clave (keywords en adelante), que son las encargadas de representar el contenido de los trabajos de la manera más básica y sintética posible.

A partir de esas bases de datos bibliográficas se desarrolla la *cienciometría*³, que fue definida por primera vez por Nalimov y Mulchenko en 1971 como, “*métodos cuantitativos de investigación en el desarrollo de la ciencia como proceso informativo*”⁴.

Aunque bien es cierto, como todos sabemos, que no todos los trabajos publicados tienen el mismo valor. Este valor dado a un trabajo parte de la calidad del mismo y para analizar esa calidad, la *cienciometría* ha adoptado la idea de *impacto científico*, usándolo como indicador que nos muestra el impacto, propiamente dicho, que provoca la publicación de un trabajo entre la comunidad científica. Para calcular este impacto se utiliza la citación partiendo de la base de que, a pesar de sus diferentes motivaciones (Bornmann y Daniel, 2008), una cita supone un reconocimiento a un trabajo anterior (Moed, 2006). Dicha citación entre autores crea unos enlaces entre personas, publicaciones, ideas, instituciones, etc., que al final conforman una red susceptible de ser analizada cuantitativamente (Guerrero-Bote, 2015).

³ Scientometrics en inglés.

⁴ Traducción propia realizada a partir de la cita, “*quantitative methods of the research on the development of science as an informational process*”, incluida en el trabajo de Nalimov y Mulchenko (1971) titulado, “*Measurement of Science. Study of the Development of Science as an Information Proces*”.

Eugene Garfield puso en valor estas redes creadas a partir de enlaces y creó el Science Citation Index⁵ (SCI) como base de datos bibliográfica que incluye las referencias (Garfield, 1955).

El historiador Price (1963, 1965), también fue de los primeros que se dio cuenta de la importancia de estas redes de trabajos, científicos y autores, comenzando a analizar procesos cuantitativos. A través de él llegamos a la idea de “*ventaja acumulativa*”⁶ (Price, 1976), también llamada: conexión preferencial, muy parecido a lo que Merton (1968, 1988) definió como “Efecto Mateo”. Desde el principio, Price se dio cuenta de ciertos problemas que los cuantitativos tendrían que resolver, y que aún siguen sin resolverse, como pueden ser: el mapeo de los colegios invisibles⁷, la relación entre productividad y calidad, así como los hábitos de citación en los diferentes campos.

En torno a esta época en la que ciertos autores empezaron a definir conceptos muy útiles para la cuantimetría, también comenzó a utilizarse el análisis de citas en la Política Científica. Por poner un ejemplo, los datos extraídos del “Institute for Scientific Information” (ISI) fueron utilizados en 1972 en los “Science Indicators Reports”⁸ de Estados Unidos, y por la OCDE. También se desarrolló el Factor de Impacto que, pese a todas sus debilidades, a día de hoy sigue siendo utilizado (Garfield y Sher, 1963).

Aunque el actual papel de la cuantimetría ha sido criticado por muchos, creemos que es muy necesario. Concretamente, se ocupa de monitorizar, recopilar y evaluar la actividad científica basándose en bases de datos bibliográficas. Algo que es necesario ante la necesidad de transparencia y como ayuda a la toma de decisiones (Mingers y Leydesdorff, 2015).

A raíz de este nuevo papel, Paul Wouters, en su tesis doctoral (Wouters, 1999), describe como ven la cuantimetría los distintos agentes que intervienen en la actividad científica:

“Los cuantitativos sienten que están midiendo la ciencia, ya sea como “científicos de la ciencia” o como sociólogos. Para la gente de la política de la

⁵ Índice de citación de la ciencia (traducción propia).

⁶ Cumulative advantage en inglés según el propio Price.

⁷ Comunidad informal de científicos que trabajan en un mismo tema y que intercambian información, pero que no tienen que pertenecer a un mismo marco institucional vinculante (Crane, 1972).

⁸ Informes de indicadores científicos (traducción propia).

ciencia, la cienciometría es apenas una más de las fuentes de los instrumentos de la política. Los eruditos de los estudios científicos tienden a considerar la cienciometría como un simple método sin teoría. Por último, los científicos tienden a dividirse en dos grupos: opositores y partidarios. Esto también es cierto para los investigadores de las ciencias sociales y las humanidades. Los adversarios plantean todo tipo de argumentos contra la medición de la ciencia en general (por ejemplo, la naturaleza creativa no mensurable del descubrimiento científico) y el análisis de citas en particular (por ejemplo, la falta de significado de la citación). Los partidarios del análisis de citas tienden a considerar la valoración científica del proceso científico como un medio para mejorar la calidad de la investigación, a pesar de sus limitaciones.”⁹.

Este papel que está jugando la cienciometría hace también que no solo se limite a reflejar la realidad, si no que la está transformando en la medida que modifica los comportamientos de académicos e investigadores (Wouters, 2014).

La aplicación de la cienciometría en la FS no es muy abundante. Hinze y Grupp en 1996 realizan mapas de Biotecnología en FS utilizando términos controlados tanto de patentes como de publicaciones científicas (1985-1993) para realizar mapas temáticos. Analizan la producción concluyendo que los países menos desarrollados de la EU habían aumentado su producción en este campo.

Seetharam y Ravichandra (1999) comparan el incremento en la producción en FST (Food Science and Technology¹⁰) de su institución (CFTRI), con su país (India) y con el resto del mundo. Utilizan publicaciones científicas, patentes, tesis doctorales y normas del periodo 1950-1990. Terminan indicando que crece, aunque decrece el ritmo de crecimiento.

⁹ Traducción propia realizada a partir de la cita, “*Scientometricians feel they are measuring science, either as “scientists of science” or as sociologists. For science policy people, scientometrics is just one of many sources of policy instruments. Scholars in science studies tend to view scientometrics merely as a method without theory. Lastly, scientists tend to be divided into two groups: opponents and supporters. This is also true of researchers in the social sciences and the humanities. Adversaries raise all sorts of arguments against measuring science in general (e.g. the unmeasurable creative nature of scientific discovery) and citation analysis in particular (e.g. the lack of meaning of the citation). The proponents of citation analysis tend to see the scientometric scrutiny of the scientific process as a means of improving the quality of research, notwithstanding its limitations.*”, incluida en el trabajo de Wouters (1999) titulado “The citation culture”.

¹⁰ Ciencia y Tecnología de los Alimentos (traducción propia).

Alfaraz y Calviño en (2004) estudian la producción científica en Ciencia y Tecnología de los Alimentos entre 1990 y 2000 de los países Iberoamericanos (incluyendo España y Portugal). Concluyen que España tiene más de la mitad de los registros y que en la década tiene una tasa de crecimiento del 11% anual.

Zhou, Zhang, y Xu (2012) analizan los cambios que ha sufrido la industria cárnica china y los retos y oportunidades que tiene por delante en el mercado mundial.

Muscio y Nardone (2012) abordan las relaciones entre la industria y la universidad en lo relativo a FS en Italia.

Guerrero-Bote, Olmeda-Gómez y Moya-Anegón (2016), analizan las actividades de investigación en FS que se llevan a cabo por centros de investigación en España tal y como quedan reflejadas en las revistas científicas internacionales. Si nos fijamos en el trabajo de Guerrero-Bote y Moya-Anegón (2015), España, China e Italia se encuentran entre los 8 primeros países en lo que a producción científica en FS se refiere.

Dentro de la cienciometría, también se utiliza la información de estas grandes bases de datos para analizar temáticamente la investigación. Una de las técnicas más utilizadas para el análisis temático es el análisis de copalabras, el cual genera una red en la que las distintas keywords están unidas por enlaces que están ponderados por el número de “papers” en los que coocurren dichas palabras (Callon, Courtial, y Laville, 1991). Sobre esta red se aplican procedimientos para detectar los grupos de keywords más relacionados lo que lleva a la estructura temática de la investigación (Romo-Fernández, Guerrero-Bote, y Moya-Anegón, 2013).

Sabiendo que en España la FS es una disciplina muy estudiada, como se apuntan Guerrero-Bote y Moya-Anegón (2015), creemos también necesario realizar el estudio cienciométrico de este caso concreto. Se buscará identificar cómo se desarrolla la actividad científica investigadora en FS en España en relación con el mundo.

Pocos son los estudios que abordan el tema de la FS en nuestro país desde una perspectiva cienciométrica. Por citar un ejemplo:

Chinchilla-Rodríguez y Olmeda-Gómez (2010), a través de un análisis multinivel pretenden, “*caracterizar la generación de conocimiento visible internacionalmente en el*

campo de la agroalimentación en España durante los últimos quince años a partir del análisis de los resultados de investigación en forma de publicaciones científicas”.

1.3. Justificación del estudio

En el pasado solo los investigadores sénior conocían la estructura temática de una disciplina, normalmente la relativa a su campo de estudio. Pero esta estructura no era formal ni estaba recogida en ningún soporte, era una estructura subjetiva que el investigador se formaba mentalmente por el gran conocimiento que tenía de su disciplina. Esto adolecía de conservadurismo, sesgo y subjetividad (Bornmann, 2011; Irvine, Martin, Peacock, y Turner, 1985).

Por tanto creemos necesaria la elaboración de trabajos que se ocupen de revelar la estructura de los campos científicos de una forma objetiva y fácilmente asimilable por investigadores noveles y sénior. Esto permitirá establecer una estructura permanente de tal manera que toda la ciencia pueda servirse de ella.

1.4. Objetivos del estudio

Fruto de esta necesidad, el presente trabajo se marca como objetivo principal, establecer la estructura temática de FS a partir del análisis de las keywords presentes en los “papers” publicados en el área. Por otra parte también se plantean otros 2 objetivos específicos y de gran importancia: conocer la consistencia de la estructura creada a partir del análisis de distintos indicadores cuantitativos; y conocer cómo se comporta esta estructura a nivel global y local, analizando el caso mundial y español. Esto nos lleva a hacernos las siguientes preguntas:

- ¿Cuántas áreas y subáreas componen la estructura temática de FS?
- ¿Cómo se relacionan las 5 áreas temáticas generales entre sí en el caso mundial y español?
- ¿Cómo de cohesionadas internamente aparecen las 5 áreas en los 2 casos?
- ¿Qué temas son los más centrales y los más especializados en los 2 casos?
- ¿Cuál es el impacto científico de cada área y cómo evoluciona?, ¿Cuenta España con un impacto científico destacado en cada área de investigación?
- ¿Cómo se comportan las modas de las keywords en los 2 casos?

2. METODOLOGÍA

2. METODOLOGÍA

2.1. Fuentes

Como base para el desarrollo de este trabajo hemos utilizado una colección de keywords extraídas de la base de datos Scopus, la cual ha sido la fuente principal utilizada para el estudio y la que ha permitido obtener los resultados del mismo. Creada por Elsevier, Scopus es una de las bases de datos bibliográficas que mayor número de revistas científicas incluye. Ha sido utilizada en numerosos trabajos cuantitativos y ha sido objeto de numerosos estudios que intentan caracterizarla y analizarla. Por ejemplo, el de Leydesdorff, Moya-Anegón, y Guerrero-Bote (2010) donde se comparan las bases de datos de Scopus y la de ISI y concluye que las dos bases de datos hacen un buen trabajo a la hora de suministrar materiales de base para el mapeo de la ciencia y que sus diferencias principales se basan en la madurez y el tamaño. O también el de Falagas, Pitsouni, Malietzis, y Pappas (2008) que compara las bases de datos PubMed, Scopus, Web of Science(WoS) y Google Scholar y determina que Scopus cuenta con un espectro más amplio de revistas que PubMed y WoS y que su análisis de citas es más rápido e incluye más artículos que el análisis de citas de WoS¹¹.

Teniendo en cuenta todo lo anterior y las características propias del presente trabajo, la decisión final de trabajar con Scopus se ha debido al hecho de que esta base de datos incluye un mayor número de revistas científicas internacionales en general y a que “la cobertura de Scopus es más exhaustiva y el impacto de citación de las revistas es menos discriminatorio”¹² (López-Illescas, Moya-Anegón, y Moed, 2009).

En la base de datos de Scopus aparecen dos tipos de keywords para cada registro, unas, las “Author Keywords”, que son las asignadas por el autor del documento y, otras, las “Index Keywords”, que son las añadidas por indizadores profesionales. En este estudio se van a utilizar las “Author Keywords”, que son más numerosas y ofrecen una descripción más detallada de los documentos recuperados.

¹¹ En el caso de Google Scholar destaca sus insuficiencias en la búsqueda científica y el hecho de que mucha información acerca de su contenido sea desconocida.

¹² Traducción propia realizada a partir de la cita, “Scopus coverage is more comprehensive, and citation impact of journals is less discriminative”, incluida en el trabajo de López-Illescas, Moya-Anegón, y Moed (2009) titulado, “Comparing bibliometric country-by-country rankings derived from the Web of Science and Scopus: the effect of poorly cited journals in oncology”.

2.2. Recuperación y tratamiento de datos

Las “Author Keywords” descargadas provienen de los trabajos publicados en todas las revistas incluidas en la “specific subject area” (área temática específica) de Scopus, FS, en el periodo de tiempo 2003-2014, y se ha procedido a su análisis. En investigaciones similares la metodología a la hora de descargar las keywords puede ser diferente, por ejemplo, en el trabajo de Romo-Fernández, Guerrero-Bote, y Moya-Anegón (2013), en el cual se hace un análisis de palabras clave para revelar patrones en el campo de las energías renovables, solamente se descargan las palabras clave de una revista del área en cuestión.

Un total de 230.007 keywords diferentes, descargadas de 184.801 documentos citables o “papers” fueron importadas a una base de datos relacional creada ad hoc que hizo posible estudiarlas de una manera más exhaustiva y someterlas a varios procesos de normalización para evitar incongruencias.

Como novedad para la normalización de las keywords se ha recurrido a la Distancia de Levenshtein y a la Distancia de Damerau-Levenshtein, dos instrumentos que nos han permitido identificar keywords similares, en lo que a caracteres se refiere, para poder trabajar con ellas y normalizarlas de una manera más eficiente. La Distancia de Levenshtein (Levenshtein, 1966) se define como el mínimo número de operaciones requeridas para transformar una cadena de caracteres en otra. Se entiende por operación, bien una inserción, eliminación o la sustitución de un carácter. Por otra parte, la Distancia de Damerau-Levenshtein (Damerau, 1964) incluye, además de la inserción, la eliminación y la sustitución, la operación de transposición de dos caracteres, operación que en la distancia de Levenshtein cuenta como 2, una de eliminación y otra de inserción.

En concreto, el proceso de normalización ha consistido en:

1. Eliminar signos de puntuación carentes de significado, tales como las comillas.
2. Unificar singulares y plurales.
3. Agrupar parejas de keywords similares, basándonos en las distancias de Levenshtein y Damerau-Levenshtein,

- a. Aquellas que tenían una distancia pequeña y una gran parte en común fueron unificadas automáticamente.
- b. Aquellas que no cumplían la condición anterior fueron revisadas manualmente y unificadas si era pertinente. Se identificaron 7469 parejas con estas características.

2.3. Análisis y visualización de datos

Una vez normalizadas las keywords, las 230.007 iniciales fueron reducidas a un total de 215.409 que serían la base para generar una red de copalabras que permitiera agruparlas por clústeres¹³, los cuales darían lugar a los distintos campos y subcampos de la futura estructura temática de FS.

Las redes de copalabras están formadas por un conjunto de nodos, que son las keywords, y por un conjunto de enlaces que conectan unas keywords con otras y que están ponderados por el número de documentos en el que coocurren ambas keywords. Es decir, si un enlace entre dos keywords tiene un gran peso, eso significa que esas dos keywords aparecen de manera conjunta en un gran número de documentos. Como expresan Neff y Corley en su trabajo de 2009, “el análisis de copalabras se basa en la teoría de que los campos de investigación pueden ser caracterizados y analizados basándose en los patrones de keywords utilizadas en sus publicaciones”¹⁴.

Puesto que, a nivel estadístico, no es igual de relevante que una keyword aparezca 10 veces que 500, era necesario establecer un umbral mínimo de frecuencia de aparición de keywords considerado relevante. El estudio previo de la distribución de las keywords, permitió generar una red consistente, basada en keywords muy representativas y no en lo que podríamos considerar palabras vacías o no representativas de la realidad del estudio. Finalmente, y partiendo de las 215.409 keywords útiles para generar la red, establecimos el umbral mínimo de frecuencia de aparición de keywords en 300 documentos, siendo un

¹³ Palabra adaptada del inglés que en una de las acepciones del diccionario Clave es definida como, “técnica estadística por medio de la cual se forman grupos que tienen un cierto grado de homogeneidad al compartir, en distinta cuantía, una serie de características semejantes” (SM, 2017).

¹⁴ Traducción propia realizada a partir de la cita, “Co-word analysis is based on the theory that research fields can be characterized and analyzed based on patterns of keyword usage in publications.”, incluida en el trabajo de Neff y Corley (2009) titulado, “35 years and 160,000 articles: A bibliometric exploration of the evolution of ecology”.

total de 297 las que cumplían este requisito y las que definitivamente formarían la red de copalabras que daría lugar a la estructura temática de FS.

A la hora de representar las redes muchos son los algoritmos que se pueden utilizar para la construcción de mapas, arrojando cada uno de ellos un diseño o distribución final diferente (Layout en adelante). Uno de los más utilizados es el del Escalamiento Multidimensional (MDS) (Van-Eck, Waltman, Dekker, y Van Den Berg, 2010), el cual tiende a localizar los elementos en una estructura artificial circular que ofrece una imagen final de la estructura de la red alejada de la realidad. También encontramos el método VOS de Van-Eck y Waltman (2007) que sitúa los elementos más reconocidos o mejor relacionados en el centro del mapa, dejando en la periferia los que son menos reconocidos. Este método no impone ninguna estructura artificial, aunque como indica Van-Eck (2011), puede ser un poco decepcionante a nivel local, mientras que una técnica como el Linlog (Noack, 2007) parece dar resultados satisfactorios tanto a nivel global como local.

Algo parecido al análisis de clústeres en las redes son los métodos de detección de comunidades que tratan de determinar, valga la redundancia, comunidades de nodos que se relacionan más entre sí que con el resto. El método tradicional consiste en ir eliminando los enlaces de mayor intermediación (betweenness) (Newman y Girvan, 2004). Waltman, Van-Eck, y Noyons (2010) proponen “una variante ponderada y parametrizada de clustering basada en la modularidad”¹⁵ que implementan en el VOSviewer (Van-Eck y Waltman, 2009). Este método permite variar la resolución de las comunidades mediante un parámetro, permitiendo obtener clasificaciones de diferente granularidad.

Para hallar el Layout se ha utilizado el software VOSviewer de Van-Eck y Waltman en su última versión, que permite seleccionar los parámetros de atracción y repulsión, y el método de normalización. Se han empleado los parámetros propios del LinLog de Noack (2009).

Dado que había que crear un mapa de copalabras general que representase la distribución de los clústeres para el caso mundial y otro específico para el caso de España, la aparición de España como país colaborador de las investigaciones fue el factor determinante. Las 297 keywords que habían pasado el corte de frecuencia de aparición

¹⁵ Traducción propia realizada a partir de la cita, “a weighted and parameterized variant of modularity-based clustering” incluida en el trabajo de Waltman, Van-Eck, y Noyons (2010) titulado, “A unified approach to mapping and clustering of bibliometric networks”.

de 300, fueron distribuidas para el caso mundial sin tener en cuenta la procedencia. Para el caso español, estas mismas 297 keywords fueron distribuidas teniendo en cuenta que apareciesen en investigaciones en las que España era país colaborador.

Además, se ha utilizado el algoritmo de “bursting” (explosión), desarrollado por Jon Kleinberg en 2003, que detecta cuándo ciertos términos se ponen de moda en un discurso y cuándo dejan de estarlo. En el caso concreto de este trabajo detecta las keywords de moda en el periodo de tiempo 2003-2014 en el área temática FS. El algoritmo genera una tabla con los periodos de bursting de los términos más frecuentes, indicando la longitud (length), la intensidad (strength) y el intervalo de tiempo en que se produce el bursting.

Finalmente, también se han utilizado los siguientes indicadores de impacto científico:

- Ndocc: Número de documentos citables publicados en revistas científicas recogidas en la base de datos de Scopus. Este tipo de documentos son los que realmente tienen un aporte científico. Concretamente en Scopus son: “articles”, “reviews”, “conference papers” y “short surveys”. En adelante nos referiremos a Ndocc cuando se hable de producción científica y a %Ndocc (%) al hablar del porcentaje de producción con respecto al total de producción en FS.
- Citas por documento (CpD): Promedio de citas por documento. La citación depende en gran medida del tiempo que el documento haya tenido para ser citado, por esta razón no se evalúa la evolución de este indicador.
- Impacto normalizado (IN): Promedio de la citación normalizada recibida por cada documento, entendiéndose esta como la ratio entre la citación recibida por el documento y la citación promedio de los documentos del mismo tipo, año y categoría (Rehn y Kronman, 2008). Si el resultado de este promedio es 1, el IN está en la media, si es 1,2 es un 20% superior a la media y si es 0,7 un 30% inferior a la media.
- Tasa de variación porcentual del IN (%VarIN): Representa la diferencia entre el IN obtenido al inicio del periodo estudiado (2003) y el IN final en términos

de porcentaje. Si el resultado es positivo significará que el IN es superior, un tanto por ciento X, con respecto al principio del periodo; y viceversa.

- Percentil: Este indicador nos muestra en una escala de 0 a 100 donde se encuentran situados los documentos que en este trabajo se emplean, siendo los más próximos a 0 los mejor situados y los más cercanos a 100, los peores.
- % Excelencia10 (%Exc): Porcentaje de documentos que se encuentran entre el 10% (percentil 10) más citado de un año, tipo y categoría determinados (Bornmann, Moya-Anegón, y Leydesdorff, 2012). La media general de este indicador será pues del 10%, de modo que si un clúster obtiene un %Exc del 13%, esto significará que este clúster supera en un 30% la media de excelencia fijada por este indicador. Si se obtiene un valor por debajo de la media, ocurriría lo mismo solo que en sentido contrario.
- % Excelencia1 (%Exc1): Porcentaje de documentos que se encuentran entre el 1% (percentil 1) más citado de un año, tipo y categoría determinados. El funcionamiento de este indicador es igual al del %Exc solo que en este la media está en el 1%.
- % Documentos citados (%Citados): Porcentaje de documentos citados. Al igual que el indicador anterior depende en gran medida del tiempo que los documentos han tenido para ser citados, por esta razón ni en este indicador ni en los 2 anteriores se evalúa su evolución.
- Tasa de variación porcentual del % (%Var%): Funciona igual que el %VarIN, pero en este caso representa la diferencia entre el % inicial y el % final.
- Índice de especialización temática (IET): Este indicador fue originalmente propuesto por Frame (1977) para caracterizar el esfuerzo relativo que un país hacía en una disciplina concreta, a partir de dividir el porcentaje mundial de publicaciones del país en la disciplina, entre el porcentaje mundial de publicaciones del país en todas las disciplinas. En nuestro caso lo utilizamos para observar el esfuerzo de España en FS con respecto al mundo.

3. RESULTADOS

3. RESULTADOS

3.1. Estructura temática de las ciencias de la alimentación

Como se indica en la metodología, de la búsqueda retrospectiva realizada en Scopus en la “specific subject area”, “Food Science”, se extrajeron 230.007 “Author Keywords” diferentes (con un total de 918.588 ocurrencias) de un total de 184.801 “papers” publicados entre 2003-2014. Tras normalizarlas, fueron reducidas a 215.409, de las cuales sólo 297 cumplían el requisito de aparecer 300 veces o más. Partiendo del total de “papers” que formaban el conjunto original de datos, 110.994 fueron los que incluían una o más de esas 297 keywords, es decir, algo más del 60%.

Para generar la estructura jerárquica de clústeres o comunidades con las 297 keywords finales, probamos varios valores del parámetro de resolución. De todos los probados, se eligió el de 1,9 porque daba lugar a 18 clústeres de resolución 2 (R2 en adelante) de un tamaño aceptable. Posteriormente, el parámetro “min size cluster” (tamaño mínimo del clúster) fue fijado en 35, agrupando estos 18 clústeres en 5 clústeres generales o de resolución 1 (R1 en adelante)¹⁶.

Con la estructura jerárquica ya establecida, se consultó a un experto en FS para que, en primer lugar, validara esta división de clústeres y, en segundo lugar, para que nos ayudara en la labor de etiquetar cada uno de los clústeres en función de las keywords que contenía. El resultado final de este proceso nos permitió crear la Estructura Temática de las Ciencias de la Alimentación (Tabla 1), compuesta por unos niveles generales y específicos (clústeres R1 y R2), cumpliendo así unos de los objetivos de este trabajo.

¹⁶ Aunque se podía haber usado un valor más bajo en el parámetro de resolución para establecer los clústeres R1, decidimos no hacerlo así debido a que los clústeres devueltos no incluirían a los de R2, algo que nos impediría crear la estructura jerárquica buscada.

Tabla 1: Estructura temática de las ciencias de la alimentación con sus 2 niveles de profundidad
(Fuente: Elaboración propia).

Clústeres R1	Clústeres R2
1. Composición de alimentos y nutrición	1.1. Nutrición y metabolismo 1.2. Nutrientes y calidad 1.3. Composición de alimentos 1.4. Composición y calidad
2. Procesado y modificado de alimentos	2.1. Influencia del procesado en las características sensoriales de los alimentos 2.2. Métodos de procesado o tratamiento de alimentos 2.3. Utilización de microorganismos beneficiosos como cultivos iniciadores en el procesado de alimentos 2.4. Modificaciones de los alimentos durante el procesado y cómo determinarlas
3. Seguridad alimentaria	3.1. Microorganismos patógenos en la leche y en productos lácteos 3.2. Micotoxinas en cereales y métodos de detección de micotoxinas en alimentos 3.3. Otros contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas) 3.4. Agentes antimicrobianos utilizados en alimentos. Métodos de determinación de estos agentes antimicrobianos
4. Conservación de alimentos y vida útil	4.1. Incremento de vida útil en alimentos vegetales durante su conservación. Métodos y modificaciones de calidad 4.2. Proceso de fermentación en vino y cerveza como método de conservación 4.3. Parámetros que condicionan la conservación de alimentos 4.4. Métodos de análisis estadístico
5. Antioxidantes en alimentos	5.1. Antioxidantes y sus efectos 5.2. Antioxidantes vegetales

Cada uno de los clústeres R1 contiene un determinado número de keywords distribuidas entre el conjunto de clústeres R2 que forman parte de él¹⁷. En los mapas de copalabras que a continuación veremos, se han utilizado distintos colores para diferenciar los clústeres, concretamente, se ha usado el color azul para representar el clúster nº 1, el verde para el nº 2, el gris para el nº 3, el rojo para el nº 4 y el amarillo para el nº 5. Los colores de los clústeres R2 pertenecen a distintas tonalidades del color principal que se le ha asignado al clúster R1 en el cuál están incluidos.

En los siguientes apartados se observan los mapas resultantes para el caso mundial, en primer lugar, y para el caso español, en segundo lugar. Además de los mapas, diferentes gráficos y tablas, todo ello diferenciados para cada uno de los casos, complementarán los resultados extraídos del análisis de copalabras y mostrarán un primer estudio de esta estructura temática creada para las FS.

3.2. Las ciencias de la alimentación en el mundo

En la Figura 1 se muestra el mapa de copalabras del caso mundial con sus 5 clústeres R1. En este tipo de mapas, los clústeres se diferencian por colores, variando el tamaño de la etiqueta y el del círculo de sus nodos (keywords) según el número de documentos asociados a cada nodo. Se observa claramente que el clúster 5 de R1 de *antioxidantes en alimentos* (amarillo) es el que presenta mayor cohesión interna con todas sus keywords bien relacionadas entre sí, ocupando una zona bien definida del mapa, y sin solaparse con otros clústeres. Por otro lado, sabiendo que los clústeres con mayor centralidad son los que se encuentran más hacia el centro de la red, no podemos decir que ningún clúster destaque por su centralidad. De hecho, si observamos las keywords situadas más al centro, comprobamos que pertenecen a varios clústeres. Aunque sí se observa que los clústeres 3 y 4 ocupan zonas menos periféricas. Por ejemplo se observa como el clúster 3 (gris) de *seguridad alimentaria*, partiendo desde la zona inferior derecha se adentra en la zona central tocando o fundiéndose con el resto de clústeres. Esto tiene su lógica, ya que la seguridad alimentaria no puede ser un campo aislado sino que tiene que ver con toda la FS.

¹⁷ En el Anexo II se muestran los datos completos de dicha distribución.

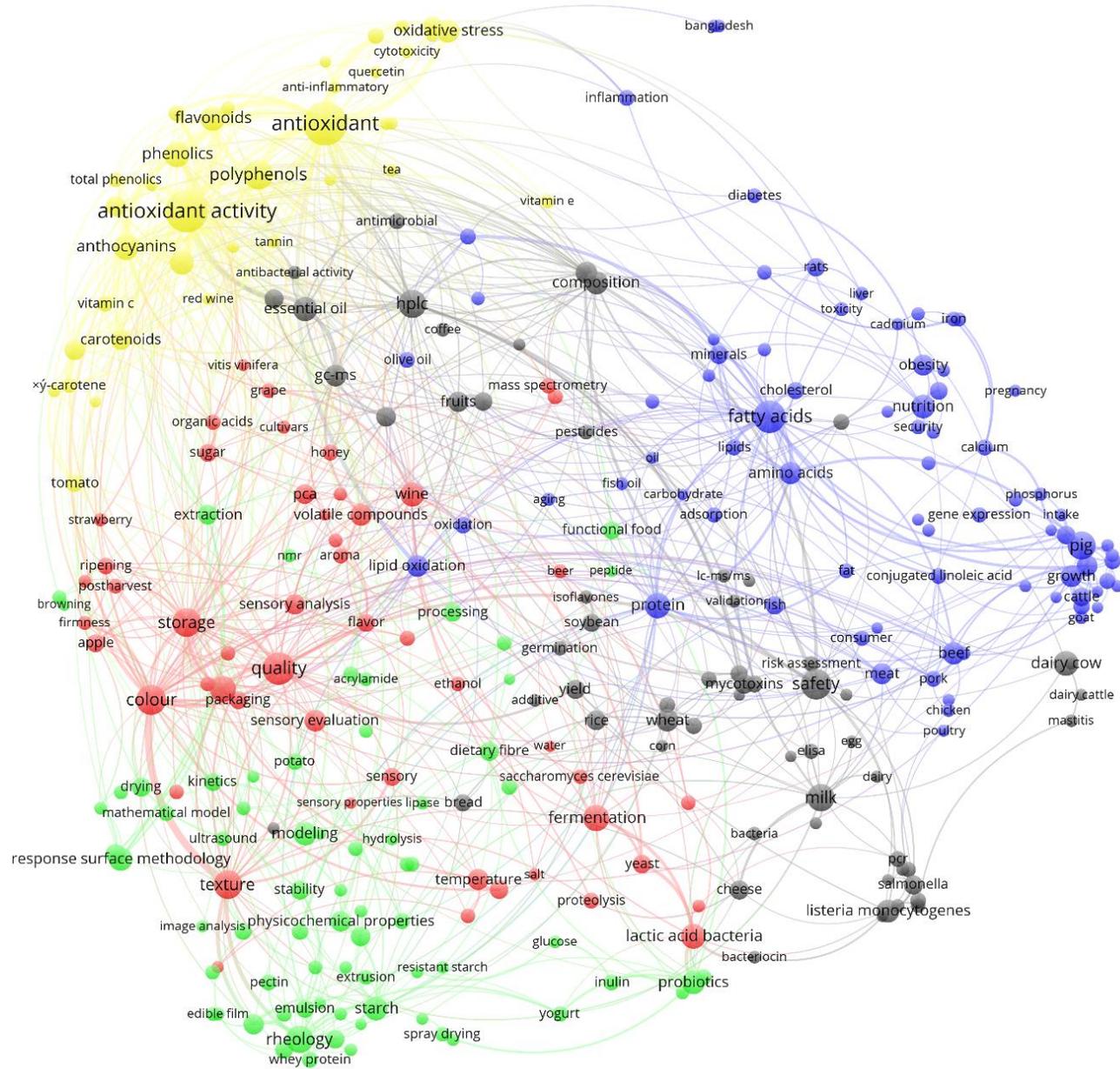


Figura 1: Vista general del mapa de copalabras distinguiendo los 5 clústeres R1 (Fuente: Elaboración propia).

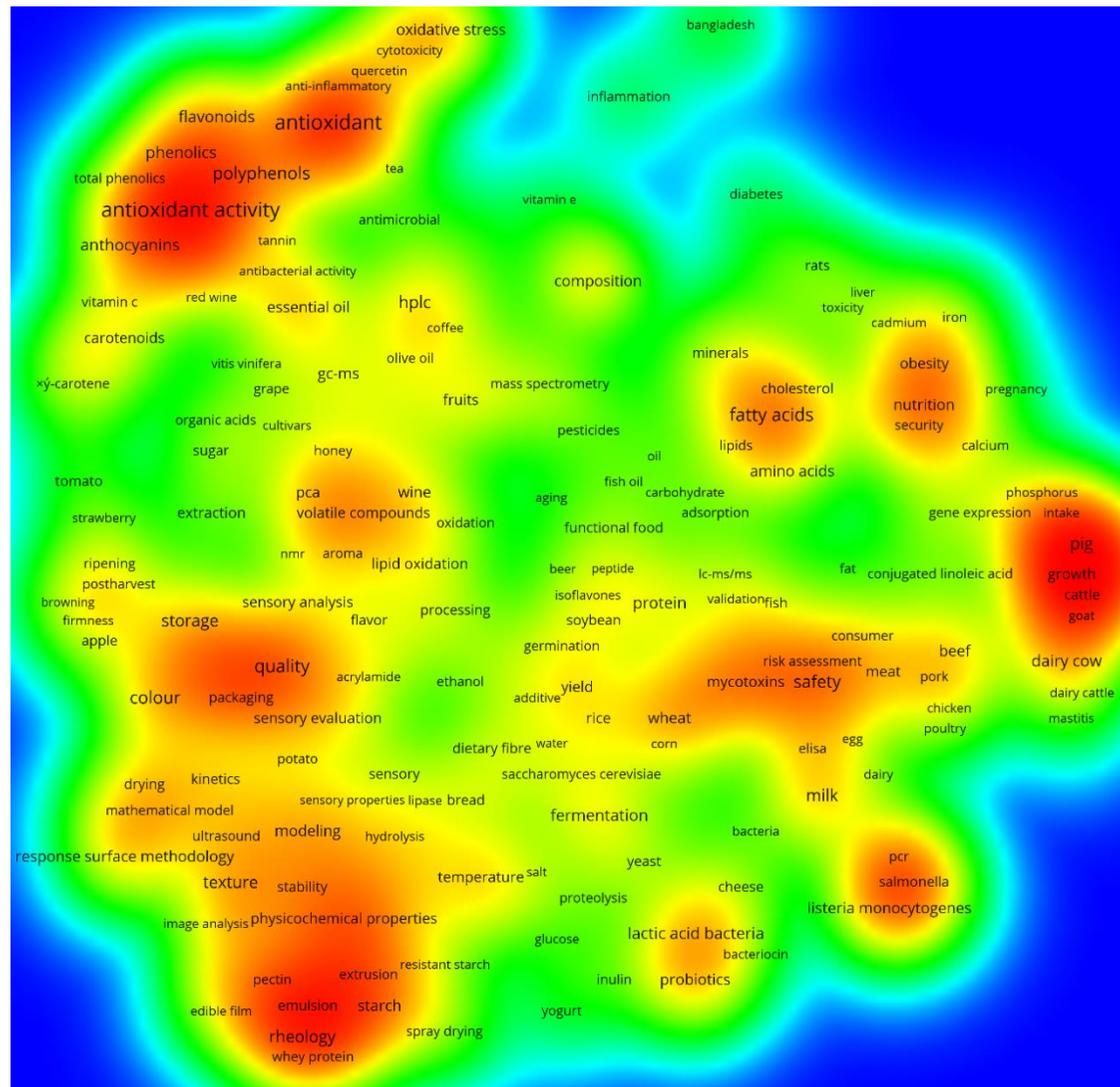


Figura 2: Vista de densidad del mapa general de copalabras. En rojo aparecen las zonas donde hay keywords que aparecen en muchos documentos (Fuente: Elaboración propia).

La Figura 2 corresponde al mapa de densidad y en él se muestran, a simple vista, 11 zonas destacadas en rojo. Estas zonas señalan gran actividad de keywords las cuales enlazan muchos trabajos y están muy relacionadas entre sí. Es decir, se puede decir que son frentes de investigación. Esto quiere decir que no todos los clústeres R2 van a tener una zona destacada, pudiendo adolecer de este modo de una falta de cohesión interna. De estas 11 zonas, 4 de ellas destacan especialmente. La primera de esas zonas destacadas estaría en el clúster R1 número 1, correspondiéndose principalmente con términos ganaderos: “pig”, “growth”, “cattle” y “goat”. La segunda de las zonas destacadas está en la parte inferior izquierda que corresponde al clúster R1 número 2, concretamente a los términos: “rheology”, “emulsion”, “whey protein”, “starch”, “pectin”, “extrusion” (clúster 2.1). Las otras dos zonas de actividad intensa están dentro del clúster R1 número 5. Una está en torno a “antioxidant activity” y la otra en torno a la keyword que agrupa más documentos que es “antioxidant” (clúster 5.1). Al hilo con lo destacado en la Figura 1, respecto a la centralidad, en este mapa de densidad tampoco se observa nada destacable y se confirma lo que comentado anteriormente.

Tabla 2: Tabla resumen de los 5 clústeres R1 (Fuente: Elaboración propia).

Clúster R1	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
1. Composición de alimentos y nutrición	33253	11,05	1,01	2,89	48,53	9,61	0,92	83,58	17,99	-22,27
2. Procesado y modificado de alimentos.	31170	12,42	1,19	0,55	44,22	12,38	1,17	86,22	16,87	-1,44
3. Seguridad alimentaria	31578	12,45	1,18	5,53	44,94	12,41	1,14	85,79	17,09	-10,22
4. Conservación de alimentos y vida útil	29154	11,65	1,06	-7,01	46,84	10,17	0,78	84,99	15,78	-18,83
5. Antioxidantes en alimentos	20901	15,99	1,65	-12,01	37,69	21,26	2,85	88,05	11,31	110,44

En cuanto a la Tabla 2 se observan los siguientes datos: el número de documentos que contiene cada clúster R1 (Ndocc), el número de citas por documento (CpD)¹⁸, el impacto normalizado (IN) y su tasa de variación porcentual (%VarIN), el percentil promedio en el que se encuentran los documentos según las citas (dentro de cada tipo documental y año de publicación. Percentil), el porcentaje de documentos excelentes

¹⁸ En los indicadores que usan la citación para su cálculo, la producción de 2014 no está incluida ya que consideramos que sus datos de citación no eran lo suficientemente estables.

dentro del 10% (%Exc) y del 1% (%Exc1) más citados, el porcentaje total de documentos citados al menos una vez (%Citados), y el porcentaje de documentos con respecto al total en FS (%) y su tasa de variación porcentual (%Var%).

Como se puede ver, todos los clústeres tienen un IN superior a la media, 1. Esto significa que la selección de keywords de mayor frecuencia nos ha llevado a la producción científica de mayor impacto científico y por tanto al “mainstream” del campo científico. Este IN viene a cristalizar en un %Exc también superior a la media (10%) salvo en el caso del clúster 1 dedicado a la *composición de alimentos y nutrición*. E igual ocurre con el %Exc1, aunque en este caso se suma a la excepción el clúster 4 encargado de la *conservación de alimentos y vida útil*.

Los documentos del clúster 5 son los que mayor impacto científico¹⁹ tienen de entre todos los clústeres R1, llegando a duplicar la media de %Exc y a casi triplicar la media de %Exc1. Todo esto se observa claramente en la Figura 4 y es indicativo de la especialización, cohesión, concreción e impacto de este clúster. Además, este clúster 5 también obtiene el mejor resultado en las citas por documento, sitúa a sus documentos en el percentil más bajo de los 5 clústeres, es el que consigue un porcentaje superior de documentos citados alguna vez y su tasa de variación porcentual de documentos con respecto al total en FS es la única positiva y además supera el 100% de crecimiento, es decir, su producción en el último año del periodo estudiado es más del doble que la del primer año del periodo. En cuanto al número de documentos y el porcentaje de estos con respecto al total en FS, es el clúster 1 el que mayores números obtiene; y el clúster 3 de *seguridad alimentaria*, el que mayor tasa de variación porcentual de impacto normalizado genera.

En el otro lado se encuentra el clúster 1 con los documentos que menos CpD reciben, el menor IN, los documentos ubicados en el percentil más alto, el menor %Exc, el menor %Citados y el %Var% que más decrece. El clúster 5 por su parte, y aunque es el que mejores resultados obtiene en muchos indicadores, también es el que peores resultados consigue en otros, concretamente en, Ndocc, %VarIN, y %. Por último es el clúster 4 el que menor %Exc1 obtiene.

¹⁹ En este trabajo, usamos 3 indicadores cuantitativos de impacto científico: IN, %Exc y %Exc1.

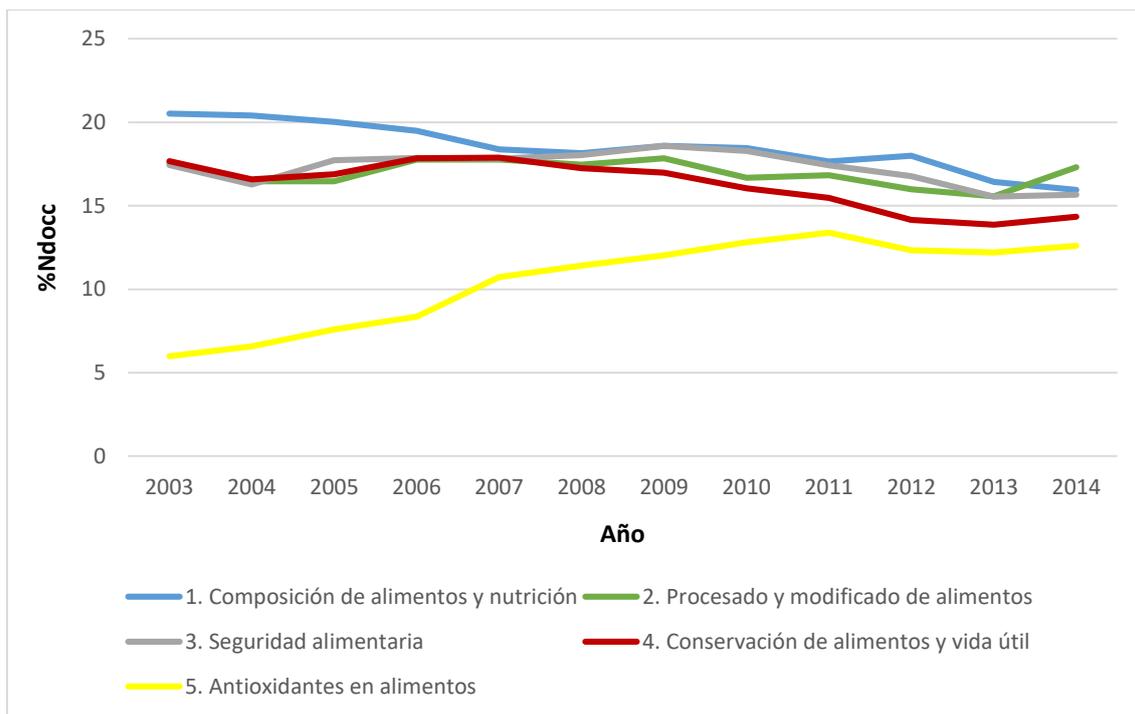


Figura 3: Porcentaje de producción de los clústeres R1, respecto al total de producción en FS (Fuente: Elaboración propia).

El gráfico anterior (Figura 3) muestra el porcentaje de producción de los distintos clústeres R1 respecto a la producción total en FS. Se observa que todos los clústeres, excepto el 5, dedicado a los *antioxidantes en alimentos*, sufren una ligera tendencia a la baja. Esto significa que durante este periodo de tiempo, los antioxidantes en los alimentos han ido captando cada vez más atención. De todas formas, la tendencia general a la baja es con respecto a la producción total en FS, porque en realidad todos los clústeres incrementan su producción (Anexo I²⁰), incluso todos la incrementan a un ritmo mayor que la producción científica mundial. Como se indicaba en la introducción, la producción científica en FS crece a mayor ritmo que la mundial (Guerrero-Bote y Moya-Anegón, 2015).

²⁰ El Anexo I muestra la producción absoluta de todos los clústeres R1 y R2.

La Figura 4 ya mencionada, representa la red de copalabras coloreada esta vez en función de la citación normalizada promedio obtenida por los documentos etiquetados con las keywords que aparecen en el mapa. Destaca la zona correspondiente al clúster 5 de R1, siendo la zona donde mayor concentración de documentos citados por encima de la media hay. En la Tabla 2 se encuentra el dato concreto de IN de cada clúster R1.

Destacan algunas keywords más por impacto pero de manera más dispersa, sin poder decir que todo el clúster al que pertenece destaque. Por ejemplo, está el grupo de las “mycotoxins” (“ochratoxin a”, “deoxnivalenol”, “cereals”, “aflatoxins”) en el clúster 3.2 de R2 (*micotoxinas en cereales y métodos de detección de micotoxinas en alimentos*). El resto de keywords que etiquetan trabajos de impacto destacado están más aisladas. El clúster 1 de R1 dedicado a la *composición de alimentos y nutrición*, tiene el impacto normalizado más bajo. Al clúster 5 le siguen en impacto el clúster 2 (*procesado y modificado de alimentos*) y el 3 (*seguridad alimentaria*), variando su orden cada año, y después el 4 dedicado a la *conservación de alimentos y vida útil*.

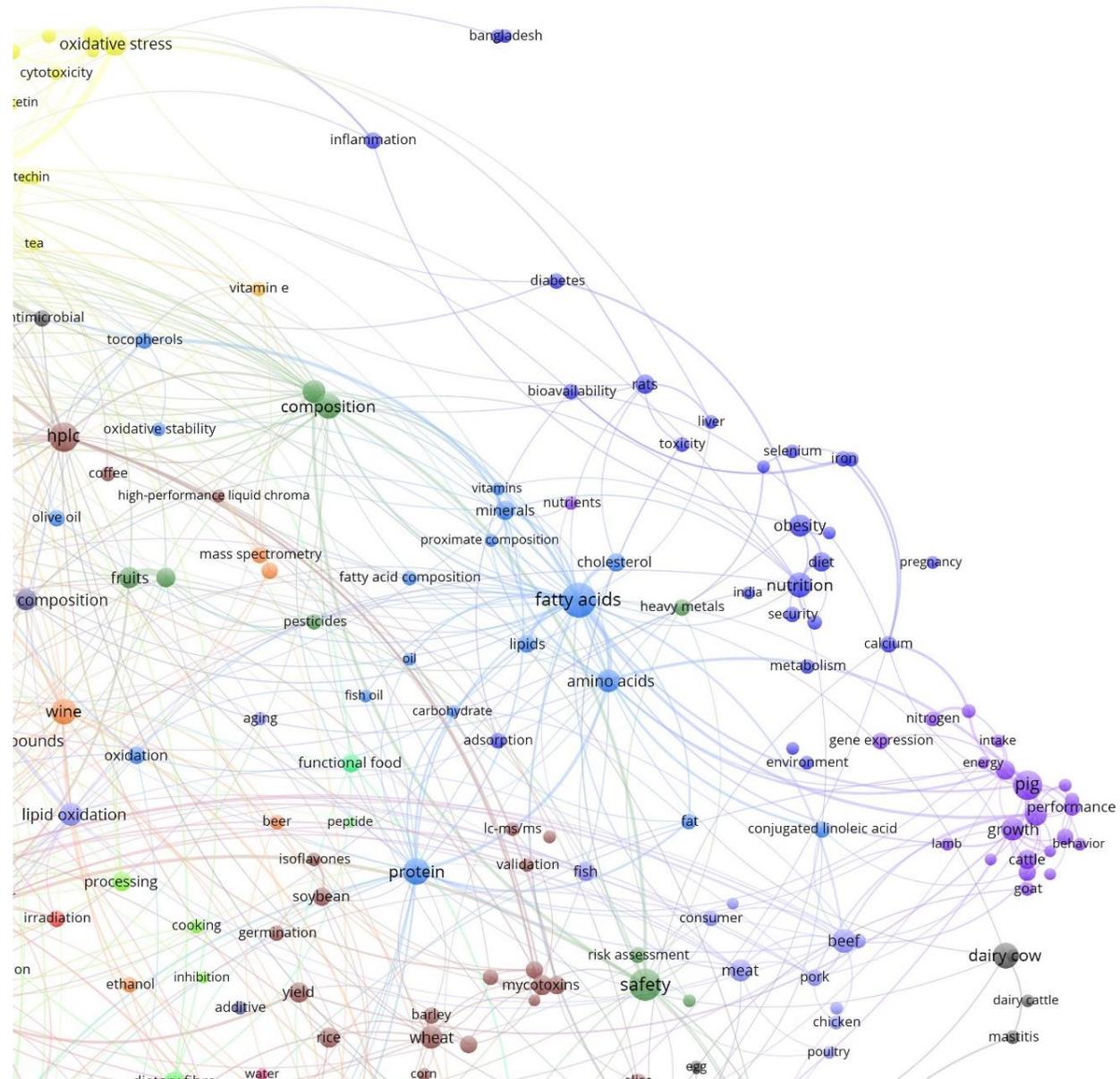


Figura 5: Zoom de la zona superior derecha del mapa de copalabras donde aparece el clúster 1 de R1 dedicado a la Composición de alimentos y nutrición (Fuente: Elaboración propia).

En la Figura 5 se muestra una vista ampliada de la zona superior derecha del mapa de copalabras donde se encuentra el clúster 1 de R1 correspondiente a la *composición de alimentos y nutrición*. Con distintas tonalidades de azul se distinguen sus clústeres R2. Aunque no es un clúster con gran cohesión, sí que se puede ver que sus 3 primeros clústeres R2 mantienen una cierta cohesión, pero no en cambio el cuarto relativo a *composición y calidad*. Si se toma como apoyo el mapa de densidad (Figura 2), se observa que en la vista en cuestión hay tres zonas destacadas: una que incluye las keywords “pig”, “growth”, “cattle” y “goat”, todas ellas pertenecientes al clúster 1.2 de R2, dedicado a los *nutrientes y calidad*; otra con las keywords “nutrition”, “obesity” y “security” relativas al clúster 1.1 de *nutrición y metabolismo*; y otra donde se leen las keywords “fatty acids”, “colesterol” y “lipids”, todas ellas del clúster 1.3 relativo a *composición de alimentos*. Y aunque algunas keywords del clúster 1.4 como “beef”, “pork”, y “chicken” aparecen situadas en una zona más o menos amarilla, no se puede decir que esa zona destaque por su densidad. Por otro lado, salvo el clúster 1.2 que se extiende por una zona muy limitada, los otros 3 clústeres tienen algunas de sus keywords alejadas de la zona donde la mayoría de sus keywords confluyen y se mezclan con las de otros clústeres. Tal es el caso de “inflammation” y “bangladesh” en el caso del clúster 1.1, algo que tiene cierta lógica puesto que no son términos específicos de FS; “tocopherols”, “oxidative stability”, “olive oil” y “oxidation” en el clúster 1.3, que se aproximan a la zona de los antioxidantes; o la de “lipid oxidation” en el clúster 1.4.

Como se puede ver en el Anexo II, de las keywords incluidas en el clúster 1 de R1, la que mayor frecuencia de aparición tiene es “fatty acids” (2331) y la que menos “stress” (300).

Tabla 3: Tabla resumen de los clústeres R2 del clúster 1 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
1.1. Nutrición y metabolismo	10677	11	1,07	5,95	49,26	10,56	1,48	81,34	5,78	18,11
1.2. Nutrientes y calidad	9930	10,4	0,91	3,52	49,97	7,61	0,52	83,87	5,37	-46,67
1.3. Composición de alimentos	10178	11,6	1,02	-10,10	47,69	9,86	0,67	84,52	5,51	-25,33
1.4. Composición y calidad	6225	12,6	1,13	14,67	45,32	11,48	1,10	86,36	3,37	-35,20

En la Tabla 3 se ve que de los 4 clústeres específicos del clúster 1, los mejores resultados se los reparten entre el clúster 1.1, dedicado a la *nutrición y el metabolismo*, y el 1.4, dedicado a la *composición y calidad*. Por su parte, el 1.1 es el que tiene el mayor número de documentos y por tanto el que mayor porcentaje de documentos incluye con respecto al total, junto con su tasa de variación porcentual, siendo la única positiva del conjunto, es decir, es el único clúster cuyo porcentaje de producción al final del periodo, supera al inicial, aunque como se explicaba antes, la producción absoluta sí que crece en todos los clústeres²¹, como se puede comprobar en el Anexo I. En cuanto al porcentaje de documentos excelentes dentro del 1% más citado, también es este clúster 1.1 el que mejor resultado consigue. Por lo tanto, en el resto de indicadores recogidos en esta tabla, el que mejores resultados obtiene es el clúster 1.4. Llama la atención el hecho de que el clúster 1.1, pese a no tener un dato especialmente destacable en impacto normalizado o en %Exc, sí que casi supera en un 50% el promedio de %Exc1, lo que significa que tiene una distribución de impacto muy sesgada con unos pocos trabajos de gran impacto (más de 50 trabajos con un impacto normalizado superior en más de 10 veces la media).

En el otro lado, es decir, en los resultados más bajos, el clúster 1.1 es el que menor porcentaje de documentos citados al menos una vez tiene, y no sólo con respecto a los incluidos dentro del clúster general 1, si no con respecto a todos los clústeres R2. Lo mismo le pasa al clúster 1.2, dedicado a *nutrientes y calidad*: es el que menos citas promedio recibe, el que menos impacto normalizado obtiene, el que coloca sus documentos en el percentil más alto de toda la lista, el que menor porcentaje de documentos excelentes tiene tanto en el 10 como en el 1 por ciento, y el que cuenta con la tasa de variación que más decrece de entre todos los clústeres R2. El clúster 1.4, dedicado a *composición y calidad*, es, ya entre los 4 clústeres R2 pertenecientes al clúster 1, el que menor número de documentos tiene y, por tanto, el que menor porcentaje de documentos tiene con respecto al total.

²¹ Aunque en muchos casos veamos que la tasa de variación porcentual del porcentaje de producción es negativa, esto no significa que el clúster en cuestión produzca cada año menos, de hecho todos producen cada año más (Anexo I); esto es así porque el porcentaje de producción es con respecto al total de FS.

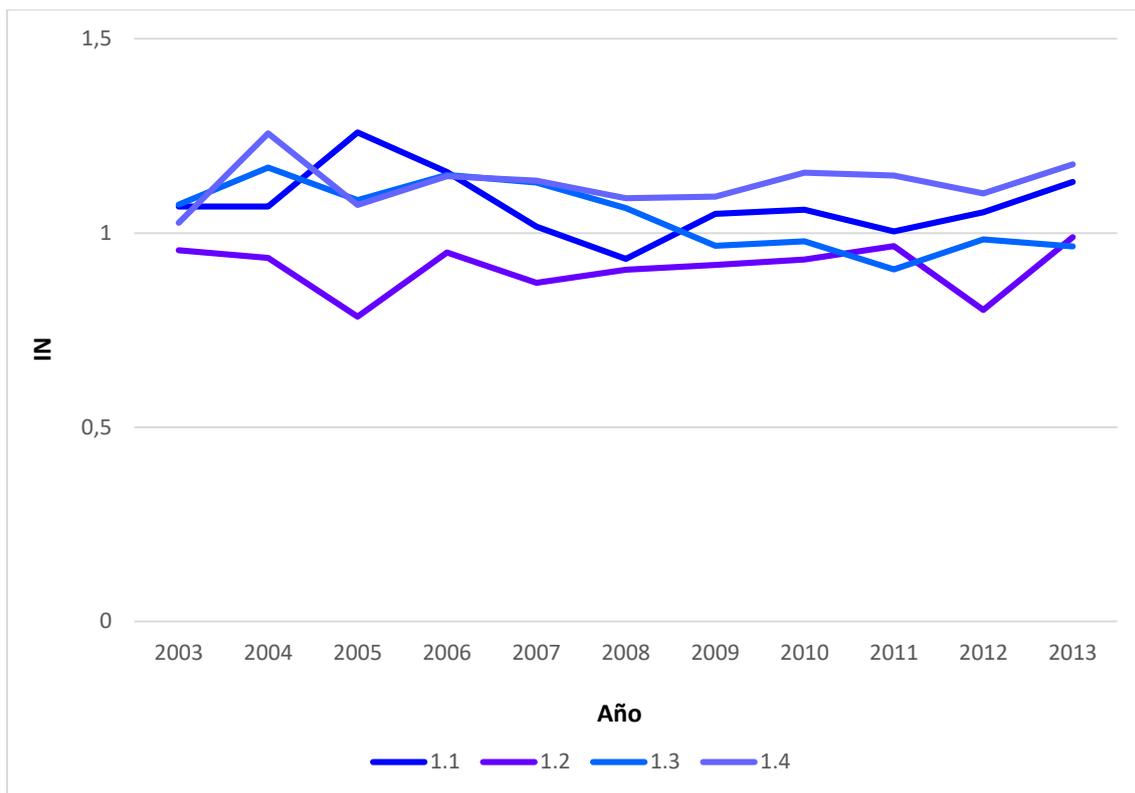


Figura 6: Evolución temporal del IN en los clústeres R2 del clúster 1 de R1 (Fuente: Elaboración propia).

En la Figura 6 se muestra la evolución del IN de cada clúster R2 incluido dentro del clúster general 1, en función de su citación normalizada promedio²². El único clúster que supera la media (1) durante todo el periodo estudiado, es el 1.4, algo que por una caída en 2008 no consigue también el 1.1, ya que en el resto de años está siempre por encima. Por el contrario, el clúster 1.2 se mantiene en todo momento por debajo de esta media y el 1.3 cae por debajo de ella en el año 2009 sin volver a ser capaz de superarla en el resto del periodo estudiado. Así todo, y de manera leve, excepto el clúster 1.3, todos ven aumentado su IN final con respecto al que tenían al inicio en el año 2003, siendo el 1.4 el que gana más y quedándose los 2 clústeres que no superan la media 1, muy cerca de ella.

²² A lo largo del trabajo habrá varios gráficos más como este, uno por cada clúster R2, y aunque lo que de ellos se puede inferir también puede hacerse a través de la tasa de variación porcentual del IN (% VarIN), hemos considerado incluir los gráficos por presentar esta información de una manera que nos es más fácil de asimilar. De esta manera, toda la explicación sobre la variación del IN se hará tomando como referencia estos gráficos. Además, en todos los gráficos de este tipo, observaremos que el último año del periodo es 2013, esto es debido a que, como se explicó antes, en los indicadores que utilizan la citación para su cálculo no hemos contabilizado la de 2014 por ser algo inestable todavía.

Tabla 4: Periodos bursting de las keywords del clúster 1 de R1, ordenados por los clústeres R2

(Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
inflammation	4	29,93	2012	-	1.1
security	3	27,20	2013	-	1.1
obesity	2	18,05	2014	-	1.1
sustainability	3	13,36	2013	-	1.1
pregnancy	1	4,60	2003	2003	1.1
children	1	3,10	2005	2005	1.1
cadmium	1	2,61	2004	2004	1.1
pig	4	81,08	2003	2006	1.2
sheep	2	29,16	2003	2004	1.2
growth	3	25,77	2003	2005	1.2
beef cattle	2	12,84	2003	2004	1.2
digestibility	1	12,50	2004	2004	1.2
carcass	2	11,98	2003	2004	1.2
goat	2	7,14	2004	2005	1.2
lamb	3	5,48	2003	2005	1.2
cattle	2	4,18	2003	2004	1.2
fat	3	16,73	2003	2005	1.3
conjugated linoleic acid	3	14,48	2003	2005	1.3
carbohydrate	1	4,57	2003	2003	1.3
olive oil	1	3,11	2004	2004	1.3
beef	2	28,72	2003	2004	1.4
tenderness	2	13,36	2003	2004	1.4
pork	2	12,42	2003	2004	1.4
chicken	2	2,76	2005	2006	1.4

Por último, en la Tabla 4 aparecen los datos relativos al bursting de este clúster 1 de R1. El clúster 1.1 tiene 7 keywords y 4 de ellas con una intensidad superior a 10 y con 3 periodos bursting similares: “inflammation” desde 2012 hasta la actualidad; “security” y “sustainability” desde 2013 hasta la actualidad; y “obesity” desde 2014 hasta la actualidad. En el caso del clúster 1.2 aparecen 9 keywords y 6 de ellas con una intensidad superior a 10. De estas 6, 5 comienzan sus periodos bursting en 2003, y la sexta, “digestibility”, comienza y acaba en 2004. Las 5 que comienzan en 2003 finalizan su

periodo bursting en distintos años: 3 de ellas (“sheep”, “beef cattle”, “carcass”) en 2004, otra en 2005 (“growth”), y la última en 2006 (“pig”). En el clúster 1.3 hay 4 keywords, 2 de ellas con una intensidad superior a 10 y con el mismo periodo de bursting, 2003-2005, son: “fat” y “conjugated linoleic acid”. Por último, en el clúster 1.4 se incluyen 4 keywords, 3 de ellas con el mismo periodo de bursting, 2003-2004 y que además son las 3 que superan la intensidad 10: “beef”, “tenderness” y “pork”.

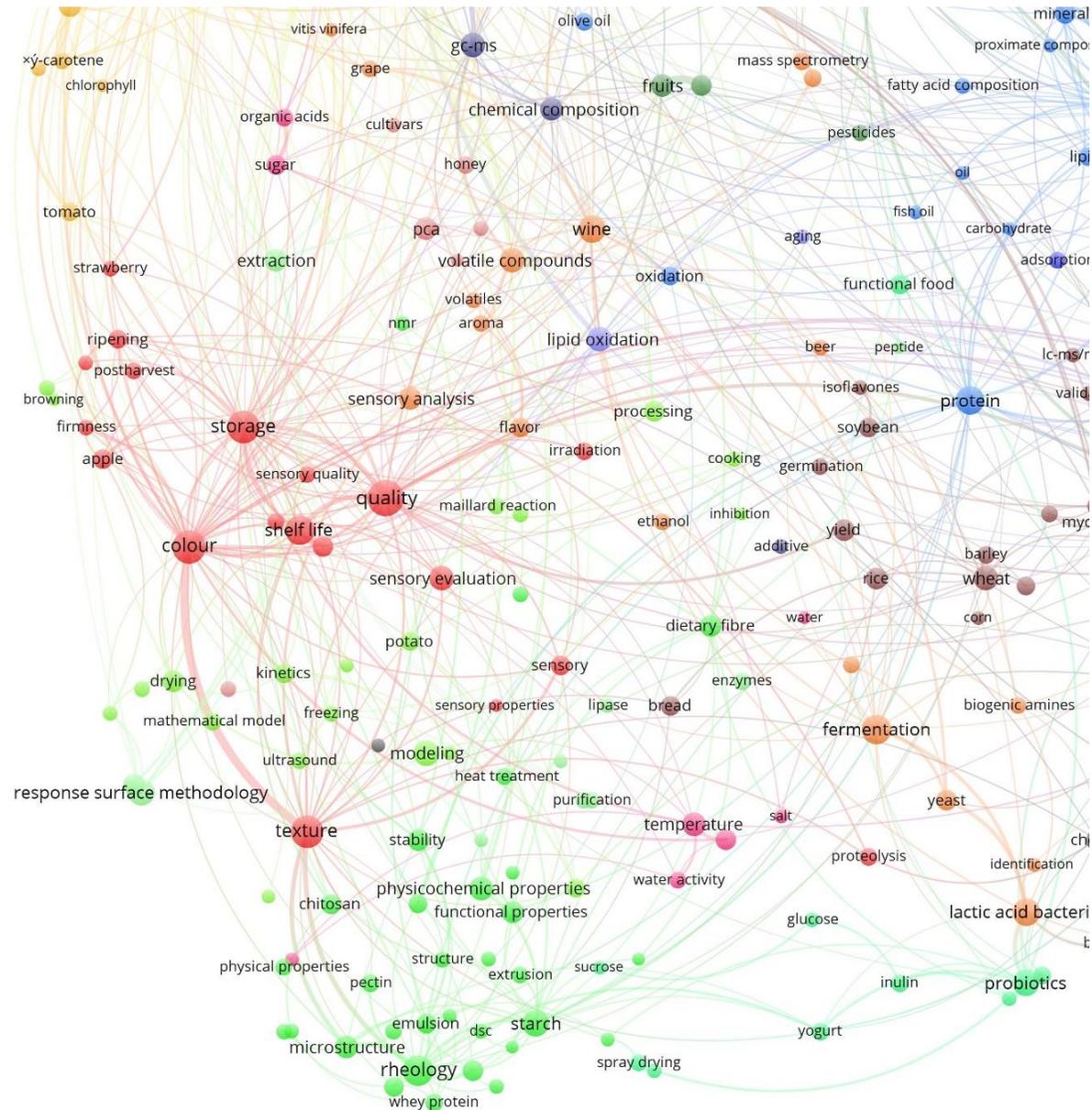


Figura 7: Zoom de la zona inferior izquierda del mapa de copalabras donde aparece el clúster 2 de R1 dedicado al Procesado y modificado de alimentos, y también el clúster 4 dedicado a la Conservación de alimentos y vida útil. (Fuente: Elaboración propia).

En la Figura 7 se observa una vista ampliada de la zona inferior izquierda del mapa donde se pueden observar 2 de los clústeres R1: el 2, dedicado al *procesado y modificado de alimentos*, y que distingue en distintas tonalidades de verde sus clústeres R2; y el 4, dedicado a la *conservación de alimentos y vida útil*, el cuál distingue sus clústeres R2 en distintas tonalidades de rojo y separa del resto del mapa al clúster 2. En primer lugar se analiza el clúster 2 detalladamente y posteriormente (después del 3) se analizará el 4.

En el caso del clúster 2 de R1, su clúster 2.1, dedicado a *la influencia del procesado en las características sensoriales de los alimentos*, aparece fuertemente cohesionado, mientras que el resto está muy disperso. De hecho, revisando la Figura 2, se puede ver una zona de alta densidad donde se sitúan las keywords: “rheology”, “starch”, “viscosity”, “emulsion”, etc. En esta Figura 2 hay otra zona de menor densidad en el clúster 2.3, encargado de los *Utilización de microorganismos beneficiosos como cultivos iniciadores en el procesado de alimentos*, una en torno a “probiotics”, “prebiotics” y “lactic acid bacteria” (aunque esta es del clúster 2.2). Mención especial para los clústeres 2.2 y 2.4 dedicados a los *métodos de procesado o tratamiento de alimentos*, y a las *modificaciones de los alimentos durante el procesado y cómo determinarlas* respectivamente, que se abren paso fundiéndose con el clúster 4 para tocar el clúster 3 de R1. Como se puede ver en el Anexo II, de las keywords de este clúster 2 de R1, la más frecuente en cuanto a aparición es “rheology” (1605) y la que menos “peptide” (303).

Tabla 5: Tabla resumen de los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
2.1. Influencia del procesado en las características sensoriales de los alimentos	15070	12,8	1,26	11,55	42,65	13,70	1,41	87,80	8,15	5,05
2.2. Métodos de procesado o tratamiento de alimentos	9184	12,5	1,10	-7,13	45,02	10,48	0,74	86,18	4,97	-28,61
2.3. Utilización de microorganismos beneficiosos como cultivos iniciadores en el procesado de alimentos	4680	13,6	1,31	-7,72	42,55	14,64	1,69	87,01	2,53	8,62
2.4. Modificaciones de los alimentos durante el procesado y cómo determinarlas	5220	11,3	1,11	-16,97	46,92	11,87	1,09	81,96	2,82	46,66

Observando ahora los datos de producción, citación y excelencia en la Tabla 5, se comprueba que el clúster 2.1 es el que mayor número de documentos tiene y por tanto el que mayor porcentaje de documentos incluye con respecto al total, además de ser el que posee el %VarIN que más crece y el mayor porcentaje de documentos citados al menos una vez. En los indicadores de impacto, el clúster 2.3 es el que mejores resultados obtiene seguido del 2.1 (aunque todos tienen un impacto normalizado promedio superior a 1), y también en CpD y en cuanto a colocar sus documentos en el percentil más bajo de los 4 clústeres R2 pertenecientes al clúster general 2. Respecto al %Var%, es el clúster 2.4 el que más la aumenta.

En cuanto a los resultados más bajos, el clúster 2.3 es el que menor número de documentos tiene y también el que menos documentos con respecto al total incluye. Por su parte, el clúster 2.2 obtiene los peores datos en los 3 indicadores de impacto, además de ser el único que tiene un %Var% negativo, siendo así el único de estos 4 clústeres que ve menguado su porcentaje de producción final con respecto a su porcentaje inicial. Por último, el clúster 2.4 es el que menor número de citas consigue, el que ve cómo su %VarIN es la que más decrece, el que sitúa sus documentos en el percentil más alto, y el que menor porcentaje de documentos citados al menos una vez acumula.



Figura 8: Evolución temporal del IN en los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).

En la Figura 8 se observa la evolución del IN de los clústeres R2 que comprenden el clúster general 2. En la mayoría de años, los 4 clústeres superan la media 1, pero el 2.2 cae por debajo de ella en 2011, y el 2.4 en 2004 y en 2013. En 2006 el clúster 2.3 obtiene el mayor impacto de entre estos 4 clústeres R2 representados en el gráfico, apreciable por un gran pico, superando en casi 7 décimas la media. Al final del periodo, tan solo el clúster 2.1 ve superado su impacto final con respecto a su inicial.

Tabla 6: Periodos bursting de las keywords del clúster 2 de R1, ordenados por los clústeres R2
(Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
cyclodextrin	2	30,94	2006	2007	2.1
encapsulation	3	14,31	2013		2.1
β -lactoglobulin	2	5,14	2003	2004	2.1
viscosity	1	4,25	2006	2006	2.1
gelatinization	1	3,83	2003	2003	2.1
ultrafiltration	1	2,95	2003	2003	2.1
gelatinization	1	6,74	2008	2008	2.1
mathematical model	2	17,21	2014		2.2
high pressure	4	10,76	2003	2006	2.2
kinetics	1	7,22	2003	2003	2.2
osmotic dehydration	1	3,89	2007	2007	2.2
modeling	1	3,61	2006	2006	2.2
ultrasound	2	3,05	2014		2.2
inhibition	1	2,60	2006	2006	2.2
glucose	2	6,38	2003	2004	2.3
enzymes	2	9,54	2003	2004	2.4
purification	2	6,23	2014		2.4
response surface methodology	2	3,33	2014		2.4
lipase	1	3,01	2003	2003	2.4

Finalmente se presentan los datos de bursting de este clúster 2 de R1 en la Tabla 6. Solo 2 de sus clústeres R2 tienen keywords con intensidad superior a 10. Es el caso del 2.1 y 2.2, con 7 keywords cada uno. En el primero, las keywords que superan el valor de intensidad 10 son, “cyclodextrin”, con un periodo bursting de 2006 a 2007, y “encapsulation”, con un periodo bursting de 2013 a la actualidad. En el segundo, “mathematical model” y “high pressure”. La primera comienza su bursting en 2014 y continúa hasta la actualidad, y la segunda va de 2003 a 2006. El clúster 2.3 sólo incluye una keyword, “glucose”, con una intensidad de 6,38 y con un periodo bursting de 2003 a 2004. El clúster 2.4 tiene cuatro keywords y una de ellas, “enzymes”, con un periodo bursting de 2003 a 2004, casi llega a una intensidad 10, se queda en 9,54.

En la Figura 9 aparece la zona central del mapa de copalabras y observamos en distintas mezclas de gris con otros colores los 4 clústeres R2 incluidos dentro del clúster 3 de R1 dedicado a la seguridad alimentaria. Este clúster se extiende diagonalmente por todo el mapa tocando la mayor parte de clústeres. Al no estar localizado en una zona concreta no podemos decir que tenga una gran coherencia como un todo, sin embargo, si nos fijamos nuevamente en la Figura 2 comprobamos que incluye 2 zonas de alta densidad, una primera en torno a las keywords “salmonella”, “listeria monocytogenes” y “pcr” y otra más extensa pero menos intensa en torno a “safety” y “mycotoxins” que se extiende por un lado sobre las keywords “wheat”, “corn” y “barley”, por otro sobre “milk” y “egg”, y por otro sobre “meat”, “pork”, “beef”, y “chicken”, pertenecientes ya al clúster 1 R1.

En el Anexo II aparecen las keywords que se encuentran dentro de cada uno de los clústeres R2 del clúster 3 de R1, dedicado a seguridad alimentaria. La keyword con mayor frecuencia es “safety” (2055), y “allergy” es la que menos frecuencia tiene (301).

Tabla 7: Tabla resumen de los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
3.1. Microorganismos patógenos en la leche y en productos lácteos	11148	13,2	1,24	4,12	42,18	13,23	1,02	87,73	6,03	-24,26
3.2. Micotoxinas en cereales y métodos de detección de micotoxinas en alimentos	11796	11	1,07	12,79	47,33	11,00	0,61	83,95	6,38	-3,51
3.3. Otros contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas)	6435	11,9	1,15	8,90	45,76	11,90	1,39	85,21	3,48	12,21
3.4. Agentes antimicrobianos utilizados en alimentos. Métodos de determinación de estos agentes antimicrobianos	4349	16	1,43	5,98	42,50	16,68	2,81	87,79	2,35	-1,73

En la Tabla 7 aparecen desglosados los datos de producción, citación y excelencia del clúster general 3, a través de sus clústeres R2. El clúster 3.1, dedicado a *microorganismos patógenos en la leche y en productos lácteos*, es el que coloca sus documentos en el percentil más bajo. El 3.2, dedicado a *micotoxinas en cereales y métodos de detección de micotoxinas en alimentos*, es el que más Ndocc tiene y lógicamente el que mayor porcentaje de documentos aporta al total, además, también es el que consigue aumentar más su %VarIN. En cuanto a CpD, IN, %Exc, %Exc1 y %Citados, el clúster 3.4, dedicado a *agentes antimicrobianos utilizados en alimentos...*, es el que mejores resultados obtiene de entre todos los clústeres que forman el clúster 3 de R1, y el segundo con mejor puntuación en comparación con todos los clústeres R2. Llama la atención en ese sentido que, aunque el 3.1 es el segundo en IN y en %Exc, sea adelantado ampliamente por el 3.3 (*otros contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas)*) en %Exc1, lo que supone que la distribución de impacto está más sesgada en este último. Sin embargo, el clúster 3.2, que es el último en IN aunque se mantiene por encima de la media 1, y pese a estar por encima también de la media 10 en %Exc, cae a 0,60 en el %Exc1, lo que significa que hay muy pocos “papers” destacados en el mismo. El clúster 3.3 es el único de los 4 clústeres R2, que aquí se analizan, que presenta un crecimiento positivo en lo que a su %Var% se refiere, siendo pues el único de los 4 que termina el periodo de tiempo estudiado (2003-2014) con un porcentaje de producción con respecto al total de FS, superior al inicial.

En lo referente a los peores datos de este clúster 3, el 3.4 es el que menos Ndocc y % presenta. El 3.2 el que menos CpD e IN consigue, el que coloca a sus documentos en el percentil más alto, y el que tiene los menores porcentajes de %Exc, %Exc1 y %Citados. Siendo el clúster 3.1 el que menor %VarIN, aunque sí que crece su porcentaje final con respecto al inicial, y %Var% tiene, en este caso siendo el que más decrece en esta tasa de variación.

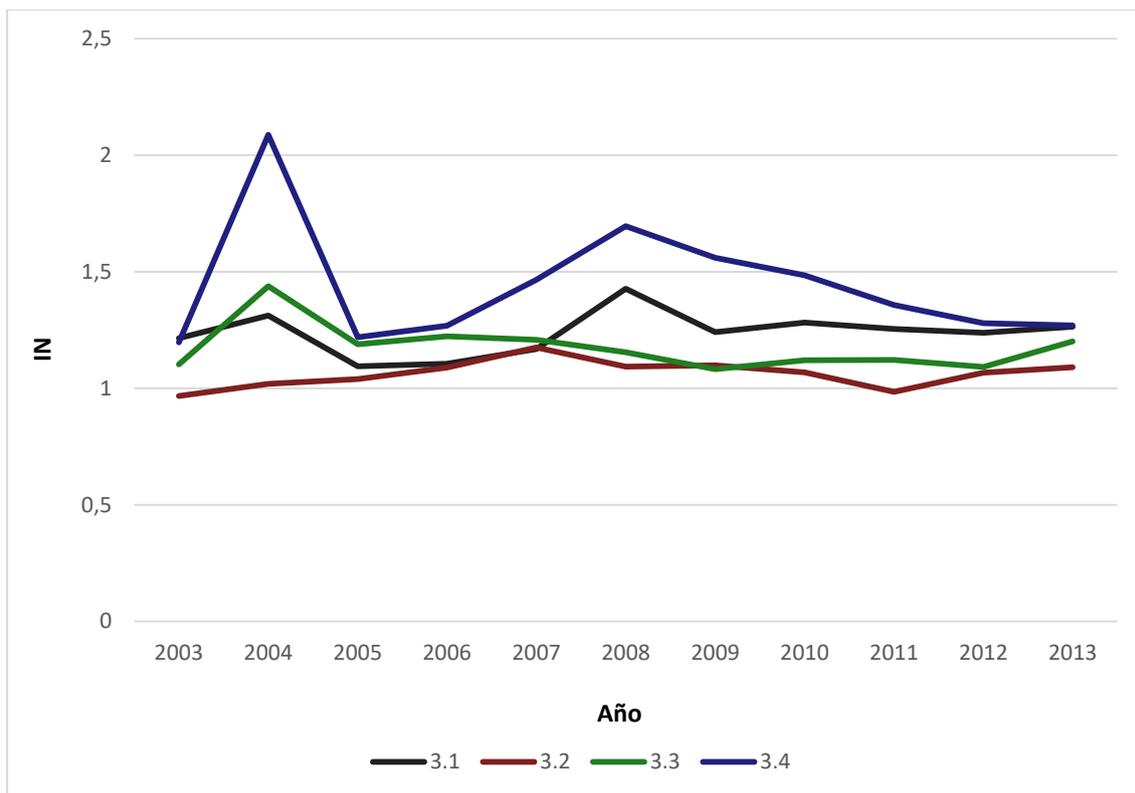


Figura 10: Evolución temporal del IN en los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).

Mirando ahora el gráfico que muestra la evolución del IN de cada clúster R2 (Figura 10) del clúster general 3, se observa que todos los clústeres superan la media de 1 durante todo el periodo estudiado, excepto el 3.2 que al inicio del periodo está por debajo de ella (2003) y en 2011 también. En 2004 el clúster 3.4 obtiene un significativo pico de impacto llegando a superar en más de un punto la media. También obtiene otro pico en 2008, algo que comparte junto con otros clústeres en esos años. Al terminar el periodo, todos los clústeres tienen un IN superior al que tenían al iniciarlo.

Tabla 8: Periodos bursting de las keywords del clúster 3 de R1, ordenados por los clústeres R2
(Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
salmonella	1	10.92	2012	2012	3.1
milk production	2	10.18	2003	2004	3.1
dairy cow	1	9.85	2004	2004	3.1
dairy cattle	1	6.35	2004	2004	3.1
mastitis	3	5.38	2003	2005	3.1
escherichia coli	1	3.37	2003	2003	3.1
dairy	3	3.36	2013		3.1
cheese	1	3.01	2003	2003	3.1
lc-ms/ms	3	24.44	2013		3.2
ochratoxin a	2	7.32	2005	2006	3.2
soybean	2	3.96	2003	2004	3.2
residues	1	2.79	2006	2006	3.2
composition	3	29.19	2009	2011	3.3
analysis	2	22.19	2010	2011	3.3

En la Tabla 8 se ven los datos bursting de 3 de los clústeres R2 del clúster 3 de R1, ya que el que falta, el 3.4, dedicado a los *agentes antimicrobianos utilizados en alimentos...*, no incluye ninguna keyword con un periodo bursting. El 3.1, referente a los *microorganismos patógenos en la leche y en productos lácteos*, cuenta con 8 keywords y 2 de ellas superan una intensidad 10: “salmonella” y “milk production”, con un periodo bursting en 2012 para la primera y uno que va de 2003 a 2004 para la segunda. El clúster 3.2, que trata el tema de las *micotoxinas en cereales y métodos de detección de micotoxinas en alimentos*, tiene 4 keywords y tan solo una de ellas supera una intensidad de 10, “Lc-ms/ms”, con un periodo bursting de 2013 a la actualidad. Por último, en el clúster 3.3, de otros *contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas)*, tenemos “composition” y “analysis” con periodos bursting de 2009 a 2011 y de 2010 a 2011 respectivamente, ambas superando la intensidad 10.

Volvemos ahora al mapa representado en la Figura 7 para analizar el clúster 4 de R1, dedicado a la *conservación de alimentos y vida útil*, el cual aparece junto al clúster 2 en esa vista zoom del mapa.

Los clústeres R2 incluidos en este clúster no tienen zonas muy confinadas sino que tiende a fundirse mucho con otros clústeres. Como se dijo anteriormente, hace de barrera entre el clúster 2 de R1 y el resto del mapa, aunque fundiéndose mucho con este último. Sin embargo, si miramos el mapa de densidad (Figura 2) sí que vemos un par de zonas de alta densidad. La primera por la zona central del clúster 4.1 (*incremento de vida útil en alimentos vegetales durante su conservación. Métodos y modificaciones de calidad*) en torno a las keywords “quality”, “shelf life”, “sensory quality” y “packing”. La segunda zona de densidad destacada está en torno a “volatile compounds” del clúster 4.3, dedicado al *proceso de fermentación en vino y cerveza como método de conservación*. Llama la atención la existencia de un clúster, el 4.4 (*métodos de análisis estadístico*) formado por keywords alusivas al análisis numérico.

En el Anexo II se muestran todas las keywords que se incluyen en este clúster 4, distribuidas por cada uno de sus clústeres R2. La keyword con la frecuencia de aparición más elevada es “quality” (2272), y la que tiene una frecuencia más baja es “water” (304).

Tabla 9: Tabla resumen de los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
4.1. Incremento de vida útil en alimentos vegetales durante su conservación. Métodos y modificaciones de calidad	14055	11,62	1,01	-6,26	47,87	9,82	0,67	84,78	7,61	-25,54
4.2. Proceso de fermentación en vino y cerveza como método de conservación	10617	12,03	1,12	-13,62	45,12	10,99	0,65	86,33	5,75	-23,22
4.3. Parámetros que condicionan la conservación de alimentos	3979	11,13	0,99	-17,77	48,22	8,48	0,81	83,55	2,15	-33,18
4.4. Métodos de análisis estadístico	3080	12,24	1,24	22,34	44,09	13,42	1,68	84,85	1,67	53,45

En la Tabla 9 se presentan los datos de producción, citación y excelencia de los 4 clústeres que forman parte del clúster general 4. El clúster 4.1 es el que mayor Ndocc acumula y también el que mayor % aporta al total. El 4.2 es el que mayor porcentaje de documentos citados alguna vez tiene. En el resto de indicadores, incluidos los de impacto,

el clúster mejor situado es el 4.4, algo que hace suponer que los trabajos que utilizan estas metodologías obtienen un impacto científico promedio superior. Además es el único de los 4 clústeres R2 del clúster general 4 que tiene las 2 tasas de variación porcentual estudiadas en positivo, siendo la de % VarIN la que más crece de todos los clústeres R2.

Los peores resultados en esta serie de indicadores los obtiene: el 4.2 para el %Exc1; el 4.3 para CpD, IN, % VarIN, ubicación en percentil, %Exc, %Citados y % Var%. En el caso Ndocc y %, el 4.4 es el peor situado tanto en el conjunto de clústeres R2 pertenecientes al clúster 4 de R1, como en el conjunto de los 18 clústeres R2 que se analizan en este trabajo.

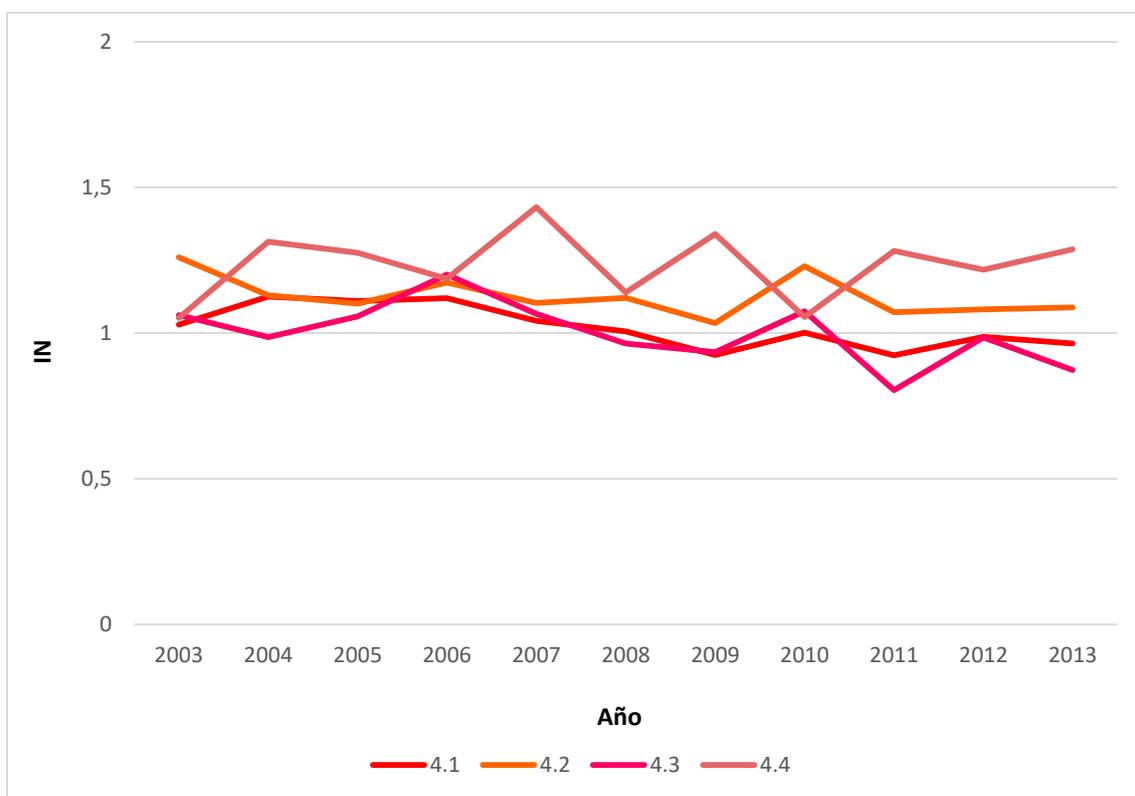


Figura 11: Evolución temporal del IN en los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).

La Figura 11 nos muestra la evolución del IN de cada clúster R2 de este clúster general 4. El clúster 4.2 y el 4.4, superan la media 1 en todo momento. Pero no ocurre lo mismo con el 4.1 y el 4.3, que antes de 2008 superaban esta media de 1, pero en este año caen por debajo de ella, recuperándose un poco en 2010, pero volviendo a quedarse por debajo en 2011. El clúster 4.1 no recupera y sigue al final del periodo por debajo de la

media, al contrario que el 4.3 que sí se recupera en 2014 llegando a un IN de 1,19. Es en 2007 cuando el clúster 4.4 obtiene el mayor impacto del conjunto de clústeres R2 de este gráfico. Únicamente el clúster 4.4 consigue tener un IN, al final del periodo, superior al del inicio.

Tabla 10: Periodos bursting de las keywords del clúster 4 de R1, ordenados por los clústeres R2
(Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
proteolysis	2	26,60	2003	2004	4.1
irradiation	2	15,43	2003	2004	4.1
proteolysis	1	38,12	2006	2006	4.1
storage	2	9,12	2003	2004	4.1
ripening	2	8,77	2005	2006	4.1
firmness	1	7,08	2006	2006	4.1
irradiation	1	20,58	2007	2007	4.1
apple	1	5,01	2003	2003	4.1
aroma	1	7,14	2003	2003	4.2
grape	1	5,14	2006	2006	4.2
volatiles	1	4,08	2003	2003	4.2
yeast	1	2,58	2003	2003	4.2
ph	1	14,76	2003	2003	4.3
moisture content	1	3,64	2007	2007	4.3
water	1	3,08	2004	2004	4.3
pca	2	5,81	2014		4.4

Pasando ya al bursting de este clúster 4 de R1 (Tabla 10), observamos que el clúster 4.1 incluye 8 keywords de las que 2 tienen una intensidad superior a 10 y la peculiaridad de contar con 2 periodos bursting distintos. La primera, “proteolysis”, tiene un periodo que va de 2003 a 2004 y otro que empieza y acaba en 2006. La segunda, “irradiation” también tiene su primer periodo bursting de 2003 a 2004 y el segundo empieza y acaba en 2007. El clúster 4.2 incluye 4 keywords, ninguna con intensidad superior a 10, al igual que el 4.4, que sólo tiene una keyword pero tampoco supera la barrera 10 de intensidad. Por último, el clúster 4.3 cuenta con 3 keywords y “ph” es la única con una intensidad superior a 10 (14,76). Su periodo bursting se produce en 2003.

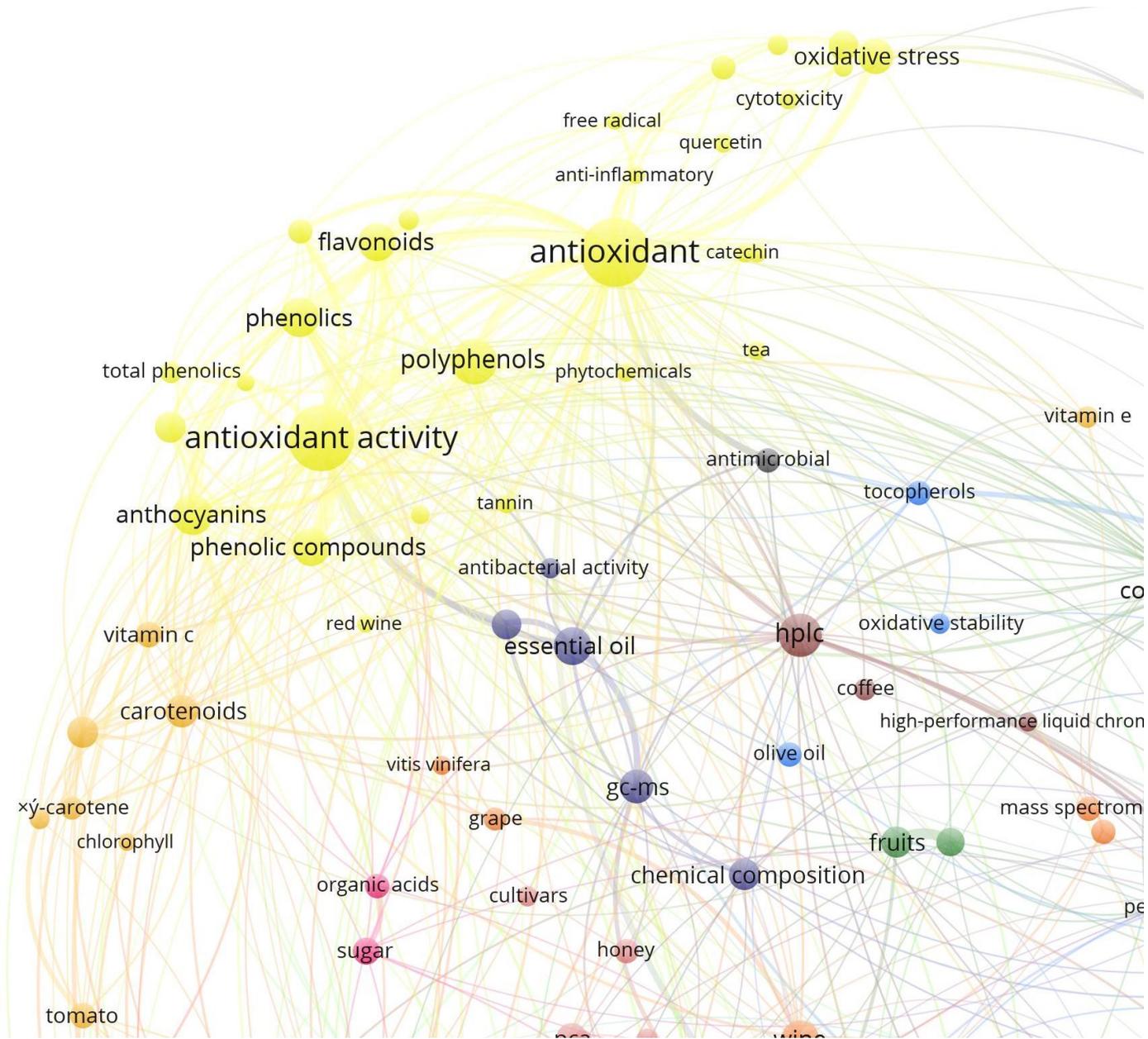


Figura 12: Zoom de la zona superior izquierda del mapa de copalabras donde aparece el clúster 5 de R1 dedicado a los Antioxidantes en alimentos (Fuente: Elaboración propia).

En la Figura 12 se presenta la zona superior derecha del mapa donde se pueden observar en distintas tonalidades de amarillo los 2 clústeres R2 incluidos dentro del clúster 5 de R1, dedicado a los *antioxidantes en alimentos*. Los documentos etiquetados con las keywords de este clúster son los que mayor impacto científico tienen. Además, si volvemos atrás y observamos la Tabla 2 también son los que consiguen un mayor porcentaje en excelencia, llegando a duplicar la media de %Exc y a casi triplicar la media de %Exc1. Esto se observa claramente en la Figura 4.

De forma general, el clúster 5 ocupa una zona muy bien delimitada del mapa, lo que es indicativo de su gran cohesión interna. Sin embargo, si se comprueba lo que ocurre a nivel de los clústeres R2 que lo forman, se observa que todo lo dicho solo es aplicable al clúster 5.1, dedicado a los *antioxidantes y sus efectos*, pero no así al 5.2, dedicado a los *antioxidantes vegetales*, el cual se extiende por una zona más amplia, con algunas keywords más alejadas, y no tiene ninguna zona de alta densidad. Si se regresa al mapa de densidades en la Figura 2, se ve que las 2 zonas de mayor densidad que comprenden a este clúster 5, pertenecen ambas al clúster 5.1: una en torno a las keywords “antioxidant”, “anti-inflammatory”, “quercetin”, etc., y otra en torno a “antioxidant activity”, “phenolics”, “polyphenols”, etc, subrayando esto lo que se acaba de comentar con respecto a los 2 clústeres R2 de este clúster 5.

En el Anexo II aparecen todas las keywords que conforman el clúster 5 de R1, siendo “antioxidant” la de mayor frecuencia (4227), ya no sólo de este clúster si no de las 297 empleadas para elaborar todo este estudio; y “chlorophyll”, con 305 apariciones, la keyword de menor frecuencia.

Tabla 11: Tabla resumen de los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
5.1. Antioxidantes y sus efectos	18047	16,68	1,74	-23	36,48	22,88	3,22	88,33	9,77	161,02
5.2. Antioxidantes vegetales	3912	13,87	1,30	9,12	41,98	15,35	1,37	87,71	2,12	-12,78

El clúster con la mayor cantidad de documentos de entre todos los clústeres R2 incluidos en este trabajo, es el 5.1, con un total de 18047. Además, este mismo clúster, es el que mayor CpD consigue, también es el que tiene mayor IN, el que tiene más documentos excelentes dentro del 10% y del 1%, teniendo además el resto de sus

documentos en el percentil más bajo con respecto a los demás clústeres R2, así como el mayor porcentaje de documentos citados al menos una vez, y el mayor porcentaje de documentos con respecto al total, así como el mejor %Var%. Solo en el %VarIN no consigue superar ni al otro clúster R2 que forma el clúster general 5, ni por supuesto ser el mejor de los 18 clústeres R2, el cual sería, como se dijo antes el clúster 4.4; de hecho su %VarIN es negativo, es decir, su IN inicial es superior al final. Por todo ello, se puede decir que el clúster 5.1 es el que mejores valores acumula en todas las categorías de la Tabla 11 y en general de entre los 18 clústeres R2 analizados en diferentes tablas a lo largo de este trabajo, exceptuando el indicador %VarIN. Sabiendo esto y que solo hay otro clúster más de R2, el 5.2, dedicado a los *antioxidantes vegetales*, se puede afirmar que este segundo es el que peores valores obtiene en la Tabla 11, exceptuando nuevamente el %VarIN, pero sin embargo, obtiene unos buenos resultados de impacto en relación a los otros 16 clústeres R2, de hecho es el segundo tras el 5.1 en cuanto a citación y en cuanto a posicionar sus documentos en un percentil más bajo.

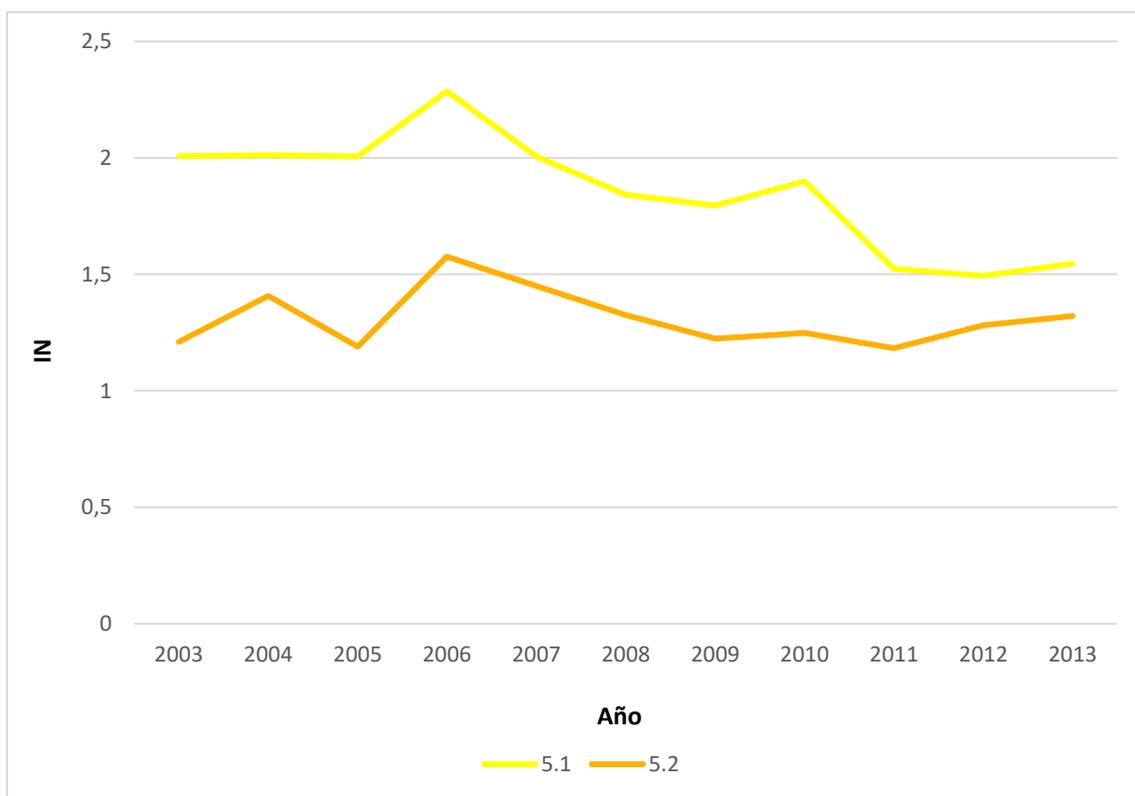


Figura 13: Evolución temporal del IN en los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).

En el gráfico de evolución del IN (Figura 13) se observa en primer lugar que ambos clústeres se mantienen por encima de la media 1 en todo momento. Por otro lado el clúster 5.1, aunque siempre supera al 5.2, presenta una clara tendencia a la baja estando al final del periodo 5 décimas por debajo del IN inicial, mientras que el 5.2 acaba con impacto similar al inicial aunque algo superior. Los 2 clústeres muestran un pico de impacto en 2006, donde ambos consiguen su mayor impacto.

Tabla 12: Periodos bursting de las keywords del clúster 5 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
oxidative stress	3	37,81	2013		5.1
bioactive compounds	3	24,07	2013		5.1
anti-inflammatory	4	14,70	2012		5.1
tannin	1	5,86	2004	2004	5.1
tannin	1	9,70	2006	2006	5.1
vitamin e	2	6,67	2003	2004	5.2

Por último, en la Tabla 12, vemos que el clúster 5 únicamente tiene su primer clúster R2 con keywords con intensidad superior a 10, estas son: “oxidative stress”, con un periodo bursting de 2013 a la actualidad; “bioactive compounds”, también con un bursting de 2013 a la actualidad; y “anti-inflammatory” con un bursting de 2012 a la actualidad. Además tiene la keyword “tannin” que cuenta con dos periodos bursting distintos, pero en ninguno de los 2 supera la intensidad 10: uno en 2004 y otro en 2006 que por 3 décimas no llega al 10 en intensidad. El clúster 5.2 sólo tiene una keyword, “vitamin e”, con un periodo bursting de 2003 a 2004, que no supera la intensidad 10.

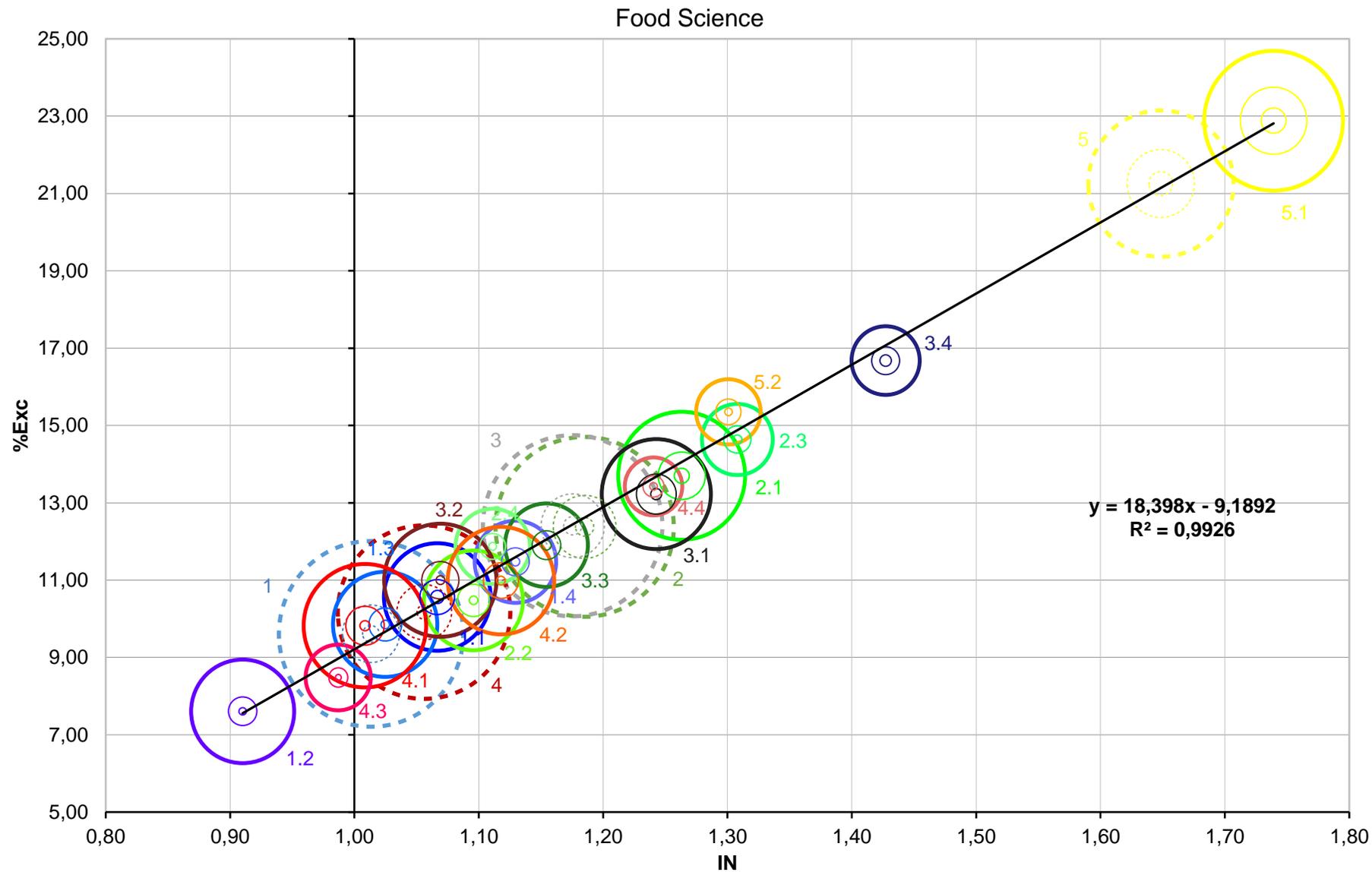


Figura 14: Porcentaje de los documentos excelentes en el 10% como “Research Guarantor” comparado con el IN de los 5 clústeres R1 (círculos discontinuos) y con los 18 de R2 (círculos continuos). Los 3 círculos concéntricos de cada clúster se corresponden, respectivamente, con los indicadores Ndocc, %Exc, y %Exc1 (Fuente: *Elaboración propia*).

En la Figura 14 se muestra el gráfico de dispersión en el que se compara el porcentaje de documentos de excelencia 10 (%Exc) con el impacto normalizado (IN) de los 5 clústeres R1 y los 18 clústeres R2.

Los círculos concéntricos representan tres de los parámetros estudiados: el número de documentos de cada clúster, el número de documentos excelentes dentro del 10 por ciento más citado, y el número de documentos excelentes dentro del 1 por ciento más citado, respectivamente.

De un simple vistazo se observa que existe una correlación fuerte positiva entre las dos variables, ya que, en primer lugar, todos los clústeres están concentrados en torno a la línea de tendencia, y en segundo lugar, a medida que aumenta el valor del eje Y, aumenta el del X y viceversa. Es decir, mientras mayor sea el impacto normalizado de un clúster, mayor será el número de documentos excelentes dentro del 10 por ciento y viceversa. Aparte, esta correlación también la confirma de manera más formal el valor del coeficiente de correlación de Pearson (R2) que es igual a 0,9926 y que por poco no llega a ser una correlación positiva perfecta con una dependencia total directa entre las 2 variables. Esto hubiera pasado si R2 fuera igual a 1.

Por otra parte, como ya se ha comentado en otras ocasiones, se ve claramente que la gran mayoría de clústeres superan la media 1 de IN, quedándose únicamente por debajo los clústeres 1.2 y 4.3. Igual pasa con la media 10 de %Exc, la mayoría la superan, pero no todos: el clúster el 1.2, el 1.3, el 4.1, y el 4.3, no la superan. También se observa destacado el clúster 5 de R1 y su clúster 5.1 de R2, como tantas otras veces.

3.3. Las ciencias de la alimentación en España

El mapa de copalabras de la Figura 15 presenta la disposición de los 5 clústeres R1 para el caso de España. Como se dijo antes, en este tipo de mapas, los clústeres se diferencian por colores (clúster nº 1 azul; nº 2 verde; nº 3 gris; nº 4 rojo; y nº 5 amarillo).

Al buscar cuál es el clúster R1 con mayor cohesión, se observa que es el dedicado a los *antioxidantes en alimentos* (el número 5), el que cuenta con una mayor cohesión interna entre sus keywords. Este clúster se encuentra en una única zona del mapa, siendo esta una característica esencial que permite determinar que es el que mayor cohesión

interna presenta entre sus keywords, las cuales están bien relacionadas entre sí y apenas presentan dispersión y solapamiento con las keywords de otros clústeres.

El clúster 4 de R1 dedicado a la *conservación de alimentos y vida útil*, también presenta una muy buena cohesión interna pero en este caso se observa una división vertical en el mapa, a través del mismo, que establece 2 zonas del mismo bien cohesionadas internamente por separado, correspondiéndose cada una con un clúster R2 del mismo, concretamente, la zona más a la izquierda se corresponde con el número 4.2, dedicado al *proceso de fermentación en vino y cerveza como método de conservación*; y la zona más al centro se corresponde con el número 4.1, dedicado al *incremento de vida útil en alimentos vegetales durante su conservación. Métodos y modificaciones de calidad*. Aparte de estos 2 clústeres de R2, los otros 2 que conforman la *conservación de alimentos y vida útil* se encuentran en la zona izquierda del mapa pero mostrando una mayor dispersión entre sus keywords. En el mapa de densidad (Figura 16) se observa muy bien la división vertical comentada.

En cuanto a la centralidad, a simple vista se observa que es el clúster número 4 el que mayor grado de centralidad presenta. Como se comentaba antes, este clúster aparece dividido en 2 zonas, pero no cabe duda que una de ellas, la relativa al *incremento de vida útil en alimentos vegetales durante la conservación. Métodos y modificaciones de calidad* (clúster 4.1), ocupa el centro del mapa. Además, si nuevamente se echa un vistazo al mapa de densidad, se puede comprobar que esta zona está muy destacada en el centro del mismo.

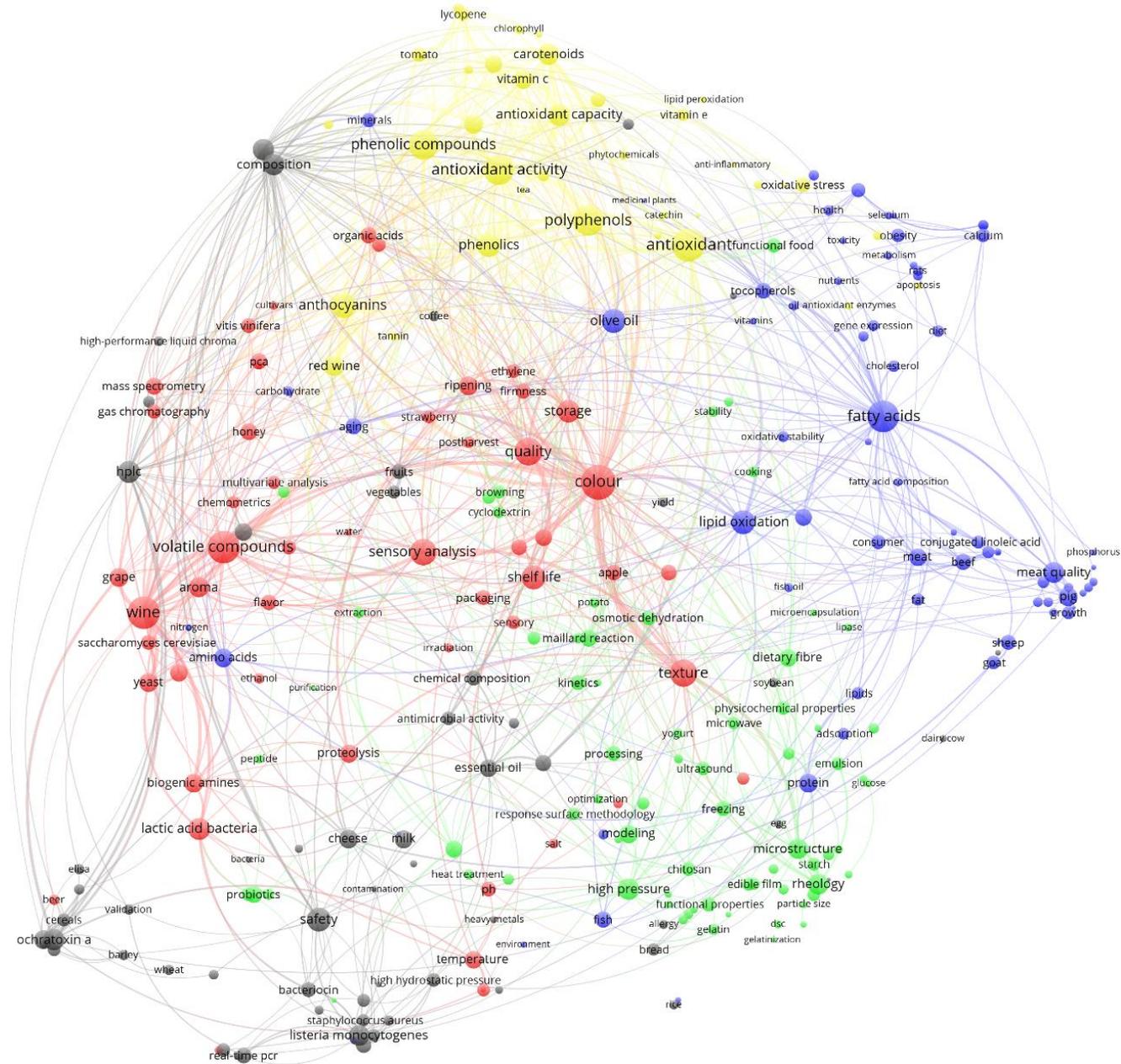


Figura 15: Vista general del mapa de copalabras de España distinguiendo los 5 clústeres R1 (Fuente: Elaboración propia)

La Figura 16 se corresponde con el mapa de densidad de España. En este mapa, como ya se apuntaba antes, se pueden ver de manera muy clara los clústeres con mayor cohesión si se atiende a las zonas con mayor concentración de keywords en rojo.

En este mapa aparecen destacadas especialmente 8 zonas y con menor intensidad otras 6. Si se da por hecho que cada zona destacada se corresponde con un clúster R2 y sabiendo que en total tenemos 18 de estos clústeres, en total tenemos 14 clústeres R2 representados en el mapa, quedando 4 sin representar, lo que indica que estos 4 no presentan gran cohesión interna entre sus palabras clave.

En la parte superior del mapa aparece una de las 8 zonas más destacadas y que se corresponde, principalmente, con el clúster 5 de R1, incluyendo keywords de sus 2 clústeres de R2 como: “antioxidant”, “polyphenols”, “tea”, “catechin”, etc. del 5.1; o “carotenoids”, “vitamin c”, “chlorophyll”, etc. En la parte izquierda del mapa se encuentran mayoritariamente keywords propias del clúster 4.2 como, “volatile compounds”, “aroma”, “wine”, “grape”, etc.; y del 4.4, “chemometrics”, y “multivariate analysis”. La parte central está claramente dominada por keywords pertenecientes al clúster 4.1: “quality”, “colour”, “storage”, “shelf life”, etc., aunque no todas pertenecen a este clúster. En la parte derecha hay 2 zonas: una más arriba con 2 keywords (“fatty acids” y “cholesterol”) que pertenecen al clúster 1.3; y otra más abajo que incluye, principalmente, keywords de varios clústeres R2 del clúster general 1: “pig”, “growth”, “meat quality” del 1.2; “conjugated linoleic acid” del 1.3; y “beef” y “meat” del 1.4. En la parte inferior del mapa, la parte destacada en rojo situada más a la derecha comprende keywords de los clústeres R2 del clúster 2 de R1, principalmente del 2.1 (“rheology”, “starch”, “microstructure”...) y del 2.2 (“modeling”, “high pressure”...), aunque también incluye keywords de otros clústeres. La zona central de esta parte inferior, en color rojo intenso tiene 2 keywords, “staphylococcus aureus” y “listeria monocytogenes”, ambas del clúster 3.1. Por último, la zona de la parte izquierda inferior muestra 6 keywords, todas ellas pertenecientes al clúster 3.2, menos “beer” que pertenece al 4.2. (Véase nota al pie número 17).

Tabla 13: Tabla resumen de los 5 clústeres R1 de España (Fuente: Elaboración propia).

Clúster R1	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
1. Composición de alimentos y nutrición	2176	14,67	1,35	25,67	37,94	14,29	0,74	91,45	16,87	-13,75
2. Procesado y modificado de alimentos.	2061	15,74	1,52	26,21	35,09	17,90	1,31	93,89	15,98	-6,53
3. Seguridad alimentaria	1980	15,66	1,63	37,34	33,13	19,34	1,82	93,79	15,35	18,823
4. Conservación de alimentos y vida útil	3057	15,70	1,40	25,68	36,26	15,96	0,75	93,75	23,71	-29,47
5. Antioxidantes en alimentos	1615	18,65	2,00	44,23	28,26	27,93	2,91	94,80	12,52	138,36

En la Tabla 13 se muestra el resumen de los 5 clústeres R1 para el caso español a través de los indicadores que ya conocemos del caso anterior.

En esta tabla, se observa que los 5 clústeres están por encima de la media 1 en cuanto a IN (esto también se observa muy bien en la Figura 18). Otro dato importante, que además está en estrecha relación con el de IN, es que el porcentaje de documentos excelentes que se encuentra dentro del 10% de los documentos más citados (%Exc), supera en los 5 clústeres la media 10. No ocurre igual con los documentos excelentes dentro del 1% (%Exc1) en el cual, los clústeres 1 y 4, dedicados a la *composición de alimentos y nutrición* y a la *conservación de alimentos y vida útil*, respectivamente, se quedan por debajo de la media 1.

El clúster 5 es el que mayor porcentaje de documentos excelentes tiene, superando en casi 8 puntos al segundo con mayor porcentaje en %Exc, que es el 3, correspondiente a la *seguridad alimentaria*, y en más de 1 punto al segundo con mayor porcentaje en %Exc1, que también vuelve a ser el clúster 3. Esto indica que el clúster número 5 es el que mayor impacto científico tiene de los 5 clústeres de R1 y, además, que cuenta con una gran cohesión interna y especialización. Si nos fijamos en el resto de indicadores de la tabla, comprobamos que este clúster es el que mejores datos obtiene en todos ellos, excepto en Ndocc y, como es lógico, en %, que es el que los tiene más bajos.

En el otro lado se encuentra el clúster número 1, dedicado a la *composición de alimentos y nutrición*, con los datos más bajos en todos los indicadores de la tabla, excepto en el Ndocc y en el %, aunque en estos 2, tampoco es el que más tiene. El que mayores resultados acumula en estos 2 últimos indicadores es el clúster 4.

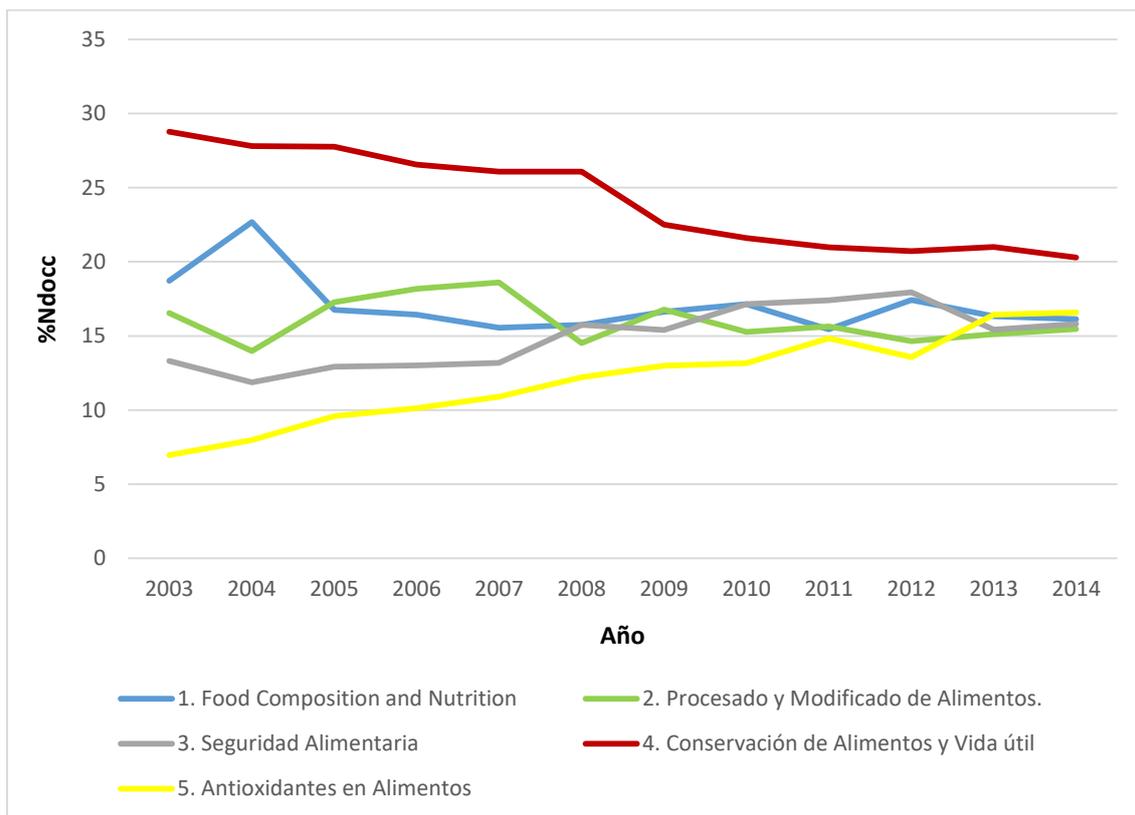


Figura 17: Porcentaje de producción de los clústeres R1 de España, respecto al total de producción en FS de España (Fuente: Elaboración propia).

En la Figura 17 se observa la evolución en el porcentaje de producción de los 5 clústeres generales con respecto al total de la FS en España. En 2 de los clústeres, el número 3 de *seguridad alimentaria* y el número 5 de *antioxidantes en alimentos*, se ve una tendencia al alza mientras que en los otros 3 se observa a la baja, siendo la del clúster 4 la caída más pronunciada. Esto significa que durante el periodo de tiempo estudiado, la *seguridad alimentaria* y los *antioxidantes en alimentos* han captado con el tiempo más atención en los estudios españoles que el resto de clústeres. Pero, al igual que ocurre en el caso de la FS en el mundo, aunque en 3 de los 5 clústeres de R1 se observe una tendencia a la baja, esto no significa que su producción decrezca, ya que este gráfico está hecho en función del total de producción de FS en España, es más, la producción de los 5 clústeres aumenta cada año, algo que se puede comprobar en el Anexo III.

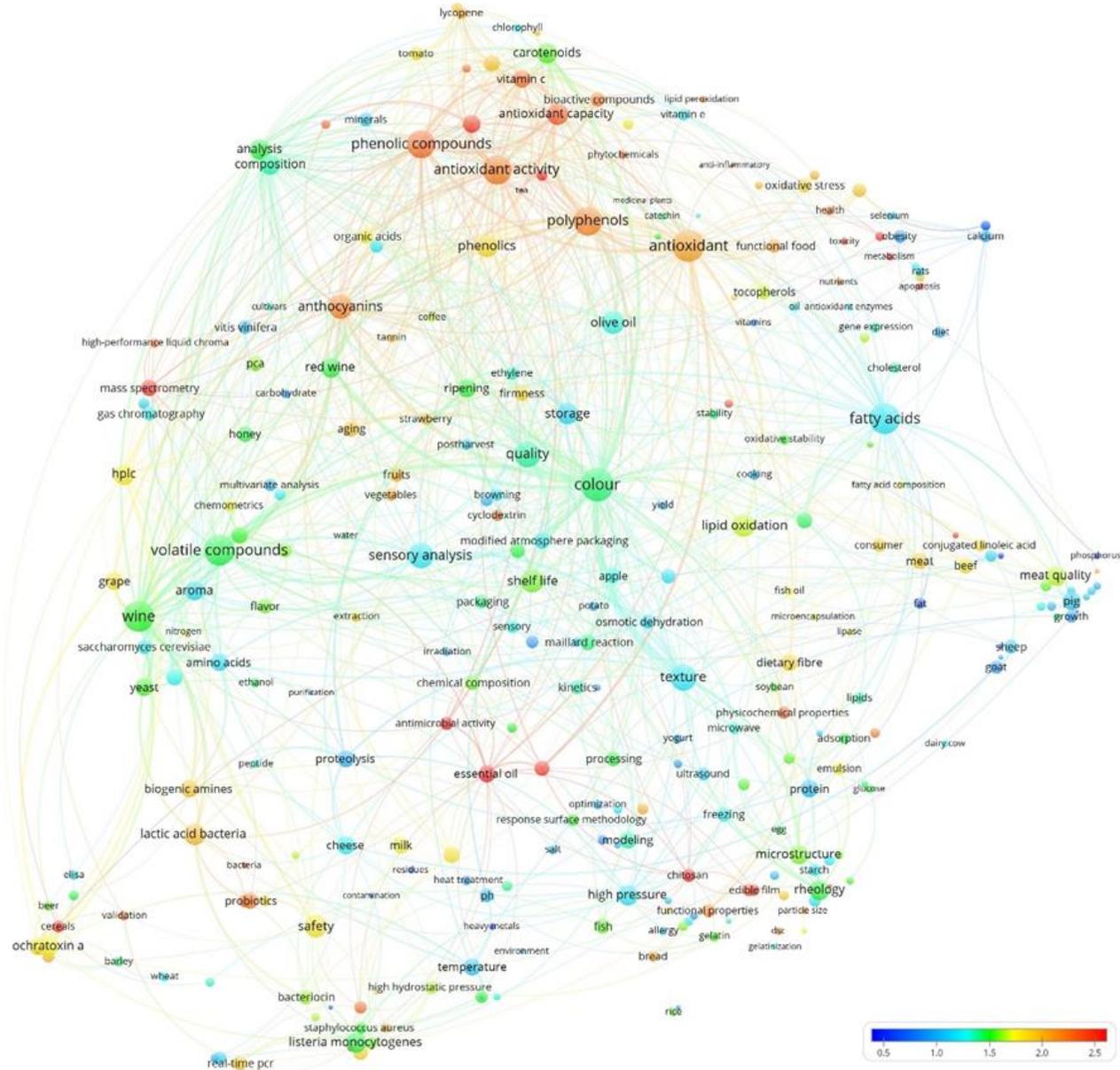


Figura 18: Mapa de copalabras del caso español, coloreado esta vez en función del IN obtenido por los documentos etiquetados con las keywords del mapa (Fuente: Elaboración propia).

En este nuevo mapa (Figura 18) aparece representada la red de copalabras coloreada en función de la citación normalizada promedio que obtienen los documentos etiquetados con las keywords del mapa. A simple vista se observa que es la zona superior, correspondiente a los *antioxidantes en alimentos*, la que más destacada aparece en cuanto a intensidad de color mostrando ser la zona del conjunto donde hay mayor concentración de documentos muy citados. Destaca también la zona inferior izquierda con más tonos amarillos, anaranjados y rojizos, correspondiente sobre todo a la *seguridad alimentaria*.

Aparte hay más keywords dispersas por el mapa con una gran citación asociada a sus documentos y, de forma general, son muy pocas las keywords que presentan un tono azul oscuro, que son las que se identifican con los documentos menos citados cayendo por debajo de la media 1. Algunas de las keywords más destacadas son: “essential oil” del clúster R2 3.4, “mass spectrometry” del 4.2, “toxicity” del 1.1, etc.

Todos estos datos nos permiten confirmar que las keywords utilizadas para la detección de los clústeres poseen una gran frecuencia de aparición, estando presentes en los documentos y temas con mayor impacto científico. En la Tabla 13 se pueden ver los datos concretos de citación, siendo en cuanto a IN el clúster 5 el que más acumula, seguido del 3, del 2, del 4 y en último lugar el clúster número 1.

Pasando ya a comentar el mapa general de copalabras (Figura 15), el clúster 1 de R1 se encuentra mayoritariamente en la zona derecha del mapa de copalabras y en primer lugar se puede decir que no es un clúster que destaque especialmente por su cohesión, ya que algunas de sus keywords se relacionan con las de otros clústeres y aparecen muy dispersas y distanciadas por el mapa en relación a la zona principal de concentración de las keywords con las que comparten clúster, tal es el caso de keywords como: “amino acids”, “olive oil” o “minerals”, las 3 del clúster 1.3. Si nos fijamos en el mapa de densidad (Figura 16) se observa un foco intenso en torno a keywords del clúster 1.2 (“pig”, “growth”, etc.) que va menguando según se acerca a keywords del 1.4 (“beef”, “meat”, etc.). También está bien destacada la keyword “fatty acids”, perteneciente al clúster 1.3. La zona donde hay más keywords pertenecientes al clúster 1.1, tiene un color amarillo rojizo que realmente no puede decirse que sea foco de gran intensidad, aunque también es cierto que en mayor o menor medida, los 4 clústeres R2 del clúster 1 de R1 se pueden identificar en el mapa de densidad.

Tabla 14: Tabla resumen de los clústeres R2 del clúster 1 de R1 de España (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	%VarIN	Percentil	%Exc	%Exc1	%Citados	%	%Var%
1.1. Nutrición y metabolismo	424	15,25	1,44	29,92	39,98	14,86	1,42	88,68	3,29	78,43
1.2. Nutrientes y calidad	550	14,35	1,33	17,43	38,40	13,82	0,91	90,55	4,26	-35,06
1.3. Composición de alimentos	922	14,07	1,27	26,20	38,46	13,34	0,11	92,41	7,15	-13,35
1.4. Composición y calidad	527	15,85	1,50	60,53	33,93	16,89	0,95	93,55	4,09	-32,90

Centrando la atención ahora en los datos de producción, citación y excelencia presentados en la Tabla 14, de los 4 clústeres R2 del clúster general 1, se comprueba que es el clúster 1.3 dedicado a la *composición de alimentos* el que mayor Ndocc tiene así como mayor %. En el resto de indicadores, excepto en %Exc1 y en %Var%, es el clúster 1.4, dedicado a la *composición y calidad*, el que mejores resultados acumula. En el %Exc1 el clúster 1.4 no llega a superar la media 1 y es el clúster 1.1 de *nutrición y metabolismo* el único de los 4 que la supera, siendo así el que mejor puntuación obtiene en este indicador. En cuanto al %Var%, vuelve a ser el 1.1 el que mejor resultado obtiene y además con diferencia del resto, siendo el único que lo tiene positivo y demostrando así que es el único cuyo porcentaje de producción con respecto al total de FS en España final, es superior al inicial.

En el otro lado, es el clúster 1.1 el que menor producción acumula y por tanto también es el que menor porcentaje de documentos con respecto al total tiene, además, también es el clúster que sitúa sus documentos en el percentil más alto, siendo además, el que más alto los sitúa de entre todos los clústeres R2, y el que menor porcentaje de documentos citados alguna vez consigue, también de entre todos los clústeres R2. En cuanto a las citas por documento, es el clúster 1.3 el que obtiene peor resultado, tanto entre los clústeres R2 de este clúster general 1, como entre todo el conjunto de clústeres R2; y también es el peor en IN, en %Exc, aunque supera la media 10, y en %Exc1. Por último, en las 2 tasas de variación (%VarIN y %Var%) el clúster 1.2 es el que peores datos presenta.

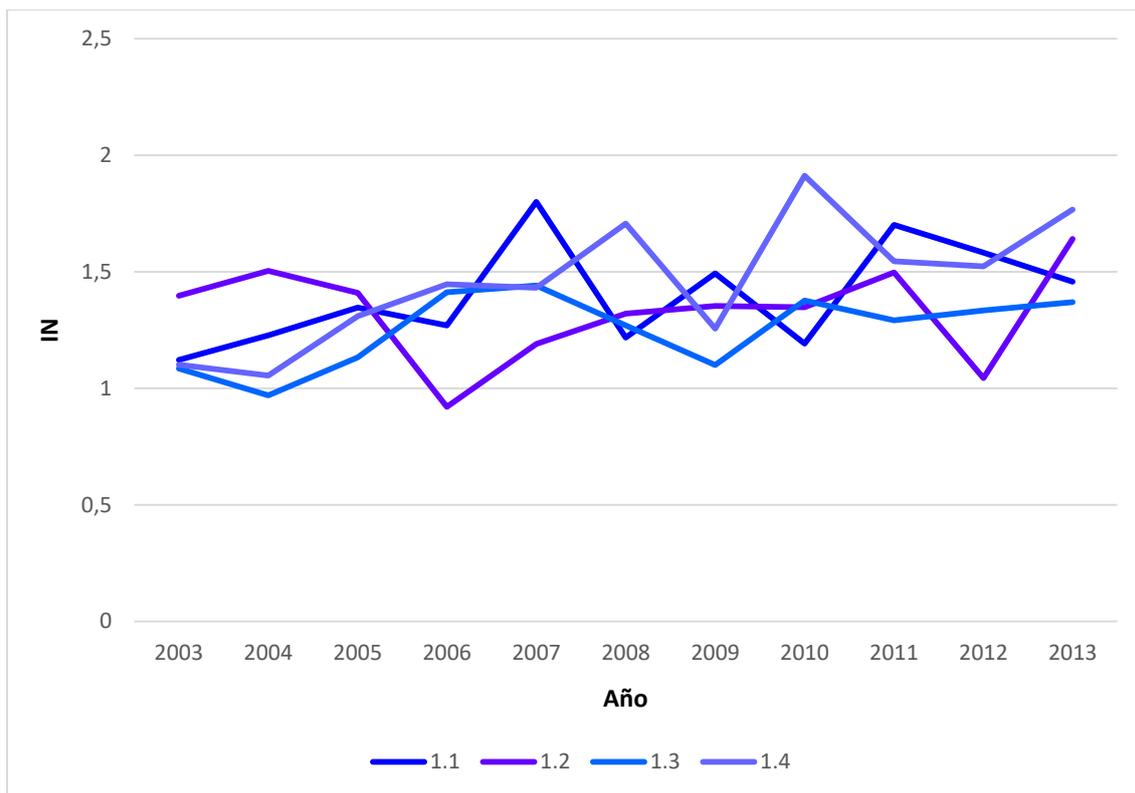


Figura 19: Evolución temporal del IN de España en los clústeres R2 del clúster 1 de R1 (Fuente: Elaboración propia).

En la Figura 19 se muestra el gráfico de la evolución temporal del IN para los 4 clústeres R2 del clúster R1 número 1. Se observa que en el año 2004, el clúster 1.3 cae por debajo de la media 1, y en 2006 el clúster 1.2 también cae por debajo de ella, siendo los únicos 2 momentos en que esto ocurre, ya que en el resto del periodo estos 2 clústeres se mantienen por encima y el 1.1 y el 1.4 superan la media en todos los años. En lo referente a la evolución del IN como tal, se ve que todos los clústeres terminan el periodo con un IN superior al del inicio confirmando así la tímida pero existente tendencia ascendente que se observa.

Tabla 15: Periodos bursting para el caso español de las keywords del clúster 1 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
zinc	4	4,50	2003	2006	1.1
obesity	3	4,53	2013		1.1
calcium	4	4,79	2003	2006	1.1
adsorption	5	3,86	2005	2009	1.1
inflammation	3	5,81	2013		1.1
cadmium	4	2,76	2005	2008	1.1
gene expression	3	2,57	2013		1.2
goat	2	4,90	2009	2010	1.2
pig	3	3,42	2004	2006	1.2
meat quality	3	3,59	2003	2005	1.2
performance	3	2,57	2010	2012	1.2
oxidation	1	3,24	2004	2004	1.3
conjugated linoleic acid	3	4,62	2008	2010	1.3
protein	5	3,04	2003	2007	1.3
olive oil	1	4,37	2004	2004	1.3
lipids	2	4,75	2003	2004	1.3
fish oil	3	3,41	2013		1.3
aging	2	3,23	2003	2004	1.4
meat	1	5,12	2003	2003	1.4
consumer behaviour	4	5,08	2012		1.4

Se observan ahora los datos bursting de este primer clúster R1 del caso español recogidos en la Tabla 15. Mientras que en el caso mundial la media de intensidad para el bursting era de 10, en este caso de España la media está en 4, por ello, a lo largo de las diferentes tablas creadas para cada conjunto de clústeres R2 en las que se analizan los periodos bursting, cuando se hable del hecho de superar o no la intensidad, se sabrá que se hace referencia a esta media de 4.

En primer lugar, se observa que los 4 clústeres R2 aquí analizados cuentan con alguna keyword que supera la intensidad 4. El clúster 1.1 de *nutrición y metabolismo*, tiene 4 de sus 6 keywords que superan la intensidad 4: “zinc” (4,50) con un periodo

bursting que dura 4 años, de 2003 a 2006; “obesity” (4,53) la cual comienza su periodo en 2013 y en la actualidad sigue sin apagarse; “calcium” (4,79), de 2003 a 2006; y por último “inflammation” (5,81) que comenzó su periodo de moda en 2013 y aún sigue encendido. El clúster 1.2, dedicado a los *nutrientes* y *calidad*, también tiene 6 keywords con periodos bursting, pero en este caso solo una de ellas supera la intensidad 4, es el caso de “goat” (4,90) con un periodo que dura 2 años, de 2009 a 2010. El clúster 1.3, encargado de la *composición de los alimentos*, vuelve a incluir 6 keywords con periodos bursting y la mitad de ellas superan la intensidad 4: “conjugated linoleic acid” (4,62) tiene un periodo bursting de 3 años (2008-2010); “olive oil” (4,37) está de moda solo en 2004, año en que comienza y termina; “lipids” (4,75) con un periodo bursting que va de 2003 a 2004. Finalmente, en el clúster 1.4 de *composición* y *calidad*, 2 de sus 3 keywords superan la media 4 y además son las que mayor intensidad consiguen de entre estos 4 clústeres R2 del clúster 1 de R1: “meat” con 5,12 es la que mayor intensidad tiene, consiguiéndolo en un único año, 2003; la segunda es “consumer behaviour” (5,08), la cual comenzó su periodo bursting en 2012 y aún sigue activa.

Volviendo a la Figura 15, a simple vista se observa que el clúster 2 de R1, dedicado al *procesado y modificado de alimentos*, no es un clúster con demasiada cohesión ya que tiene dispersas keywords de sus clústeres R2 por casi todo el mapa, aunque también es cierto que la gran mayoría de ellas se encuentran ubicadas en la zona inferior derecha del mapa. Además, observando el mapa de densidad (Figura 16) se comprueba que justo en esta zona hay un foco de densidad bien destacado y que está formado en su mayoría por keywords del clúster 2.1 (*influencia del procesado en las características sensoriales de los alimentos*) tales como “rheology”, “starch”, “edible film”, etc. Por todo ello, salvo el clúster 2.1, los otros 3 no son discernibles en el mapa de densidad, indicando la poca cohesión del clúster en general y confirmando la dispersión de sus keywords antes comentada.

Tabla 16: Tabla resumen de los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
2.1. Influencia del procesado en las características sensoriales de los alimentos	876	16,10	1,69	89,08	33,40	22,15	1,71	94,29	6,79	27,61
2.2. Métodos de procesado o tratamiento de alimentos	855	14,81	1,31	-1,89	37,06	13,33	0,58	93,45	6,63	-30,38
2.3. Utilización de microorganismos beneficiosos como cultivos iniciadores en el procesado de alimentos	291	17,66	1,73	-2,78	32,57	20,27	2,41	94,85	2,26	39,75
2.4. Modificaciones de los alimentos durante el procesado y cómo determinarlas	241	14,44	1,31	-3,24	38,31	13,69	0,41	91,29	1,87	-22,73

Como se puede comprobar en la Tabla 16 el clúster 2.1 es el que mayor Ndocc tiene y por consecuencia también el que mayor %, pero además, también cuenta con la mayor tasa de variación de IN, que a su vez es la única positiva del conjunto, y con el %Exc más alto. En el resto de indicadores es el clúster 2.3, de *Utilización de microorganismos beneficiosos como cultivos iniciadores en el procesado de alimentos*, el que obtiene los mejores resultados, siendo junto con el 2.1 los que tienen un %Var% positivo.

En lo tocante a los peores resultados, es el clúster 2.4, dedicado a las *modificaciones de los alimentos durante el procesado y cómo determinarlas*, el que cuenta con ellos en Ndocc y en %, en CpD, en IN junto con el 2.2 ya que en ambos casos el resultado es de 1,31, en % VarIN, en Percentil situando sus documentos en el más alto de todos, en %Exc1, sin llegar a superar la media 1 que tampoco es superada por el clúster 2.2, y en %Citados. En los indicadores de %Exc y %Var%, es el clúster 2.2 el que tiene peores resultados.

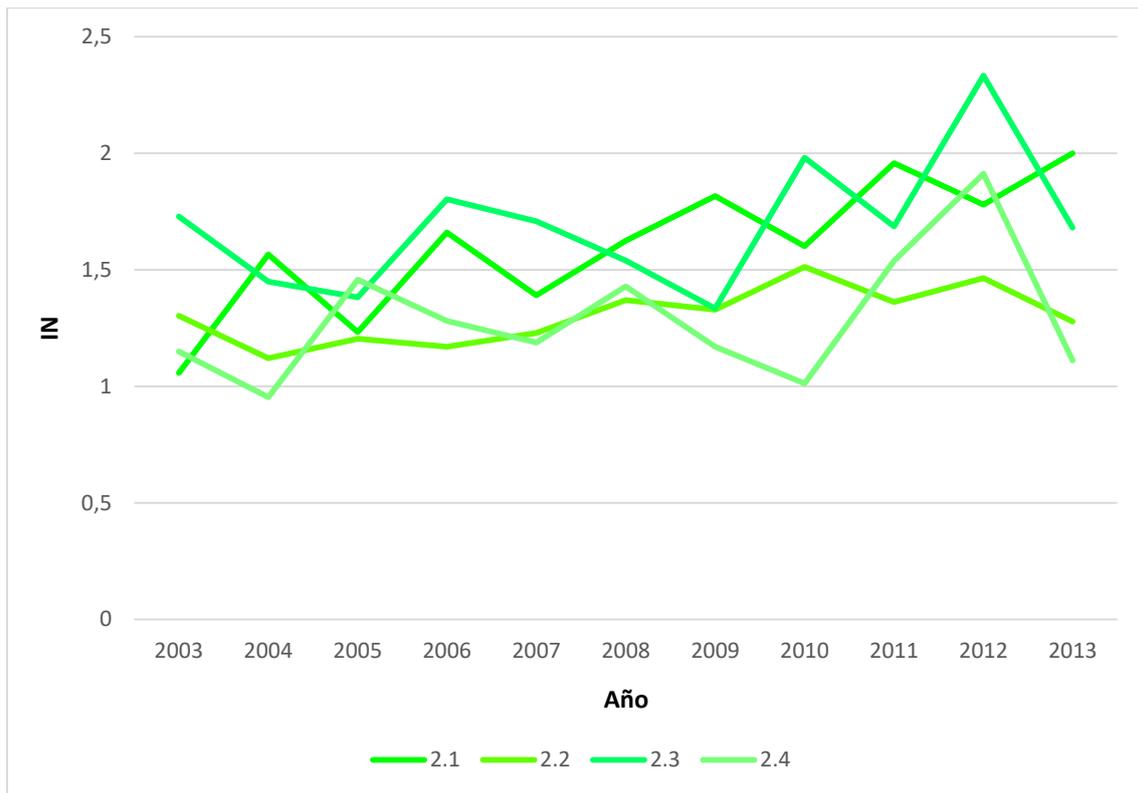


Figura 20: Evolución temporal del IN de España en los clústeres R2 del clúster 2 de R1 (Fuente: Elaboración propia).

Una clara tendencia al alza se observa en el IN de los 4 clústeres R2 mostrados en la Figura 20, aunque también es cierto que sufren bastantes altibajos durante todo el periodo. Tanto es así que al final del periodo solo el clúster 2.1 termina por encima del IN inicial llegando a casi duplicarlo, mientras que los otros 3 clústeres sufren grandes caídas en 2013, sobre todo los clústeres 2.3 y 2.4, siendo estas las causantes de terminar con un IN inferior al inicial. En todo momento se supera la media 1, excepto en 2004 donde vemos que el clúster 2.4 cae por muy poco.

Tabla 17: Periodos bursting para el caso español de las keywords del clúster 2 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
cyclodextrin	1	5,68	2007	2007	2.1
whey protein	3	3,14	2005	2007	2.1
resistant starch	4	3,52	2008	2011	2.1
functional properties	4	2,62	2004	2007	2.1
encapsulation	3	4,47	2013		2.1
dietary fibre	2	3,60	2009	2010	2.1
Î²-lactoglobulin	3	3,53	2005	2007	2.1
drying	5	4,31	2003	2007	2.2
freezing	5	3,96	2003	2007	2.2
processing	3	3,65	2003	2005	2.2
acrylamide	2	2,87	2009	2010	2.2
high pressure	3	5,49	2005	2007	2.2
kinetics	4	5,46	2003	2006	2.2
mathematical model	2	2,64	2014		2.2
browning	2	3,85	2006	2007	2.2
lactobacillus	1	3,07	2009	2009	2.3
glucose	2	2,76	2003	2004	2.3
enzymatic hydrolysis	4	3,38	2012		2.4
enzymes	5	4,39	2003	2007	2.4

Al observar el bursting del clúster general 2 (Tabla 17), se comprueba que no es un clúster que cuente con keywords con demasiada intensidad, de hecho su clúster R2 número 2.3 solo tiene 2 keywords y ninguna de ellas supera la media de intensidad 4, y aunque sus otros clústeres sí que tienen alguna keyword que supera la media, son pocas. El clúster 2.1 tiene 7 keywords con periodos bursting y solo 2 de ellas superan la media 4: “cyclodextrin” (5,68), cuyo periodo bursting empieza y acaba en 2007; y “encapsulation” la cual comienza su periodo en 2013 y sigue activa al final del periodo estudiado. El clúster 2.2 cuenta con 8 keywords con bursting y 3 de ellas tienen una intensidad superior a 4: “drying” (4,31) con un largo periodo de 5 años que va de 2003 a 2007; “high pressure” (5,49) con un periodo de 2005 a 2007; y “kinetics”, que tiene una intensidad de 5,46 y un periodo bursting de 4 años de duración, de 2003 a 2006. Como se

dijo antes ninguna de las 2 keywords del clúster 2.3 supera la intensidad 4, y en el caso del 2.4, que también tiene únicamente 2 keywords con periodos bursting, solo una de las 2, “enzymes”, supera el 4 de intensidad, 4,39 concretamente con un periodo bursting que se mantiene activo durante 5 años (2003-2007).

La representación del clúster 3, de *seguridad alimentaria*, en el mapa general de copalabras (Figura 15), muestra un clúster con una importante dispersión de sus keywords las cuales podemos encontrar de arriba abajo en la zona izquierda del mapa, aunque sobre todo en la zona inferior. Esto en principio señala poca cohesión interna, pero si se observa el mapa de densidad, en la zona inferior izquierda de este mapa se destacan 2 zonas de intensidad en rojo, una que engloba las keywords “beer”, “cereals” y “ochratoxin a”, siendo las 2 últimas del clúster 3.2, dedicado a las *micotoxinas en cereales y métodos de detección de micotoxinas en alimentos*, y otra con las keywords “staphylococcus aureus”, “listeria monocytogenes” y “bacteriocin”, todas ellas pertenecientes al clúster 3.1, dedicado a los *microorganismos patógenos en la leche y en productos lácteos*. Sabiendo todo esto, se puede afirmar que el clúster 3 en general es un clúster con poca cohesión aunque llegue a mostrar cierta intensidad entre varias keywords.

Tabla 18: Tabla resumen de los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
3.1 Microorganismos patógenos en la leche y en productos lácteos	758	15,81	1,61	31,09	32,86	18,47	1,98	94,33	5,88	3,70
3.2 Micotoxinas en cereales y métodos de detección de micotoxinas en alimentos	680	15,39	1,61	30,31	32,78	19,12	0,74	93,82	5,27	49,46
3.3 Otros contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas)	432	14,68	1,56	32,59	34,43	18,75	1,85	93,06	3,35	52,04
3.4 Agentes antimicrobianos utilizados en alimentos. Métodos de determinación de estos agentes antimicrobianos	277	17,69	1,94	117,71	31,06	25,63	5,05	94,58	2,15	9,81

En la Tabla 18 se encuentran los datos de producción, citación y excelencia del clúster 3 de R1, desglosados a través de sus clústeres R2. Es el clúster 3.1 el que mayor número de documentos acumula y por tanto el que mayor % de documentos con respecto al total tiene. En el resto de indicadores, excepto en %Var%, es el 3.4, dedicado a *agentes antimicrobianos utilizados en alimentos. Métodos de determinación de estos agentes antimicrobianos*, el que mejores resultados tiene, superando con mucho en varios de ellos al resto, especialmente en %VarIN, donde supera el 100% mientras que el resto se queda al torno al 30%, de hecho, es el que mayor %VarIN consigue de entre todo el conjunto de clústeres R2, al igual que también lo hace en el %Exc1. El %Var% en los 4 clústeres es positivo, pero es en el clúster 3.3 de, *otros contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas)*, donde obtiene su mayor diferencia positiva.

Por otro lado, el clúster 3.4 es el que menos Ndocc y % tiene; el 3.3 el que peores datos presenta en CpD, IN, situación de documentos en percentil y en %Citados; el 3.2 cuenta con los valores más bajos en %VarIN y en %Exc1, donde es el único que no alcanza la media 1; y el 3.1 en los indicadores restantes, es decir, en %Exc y en %Var%.

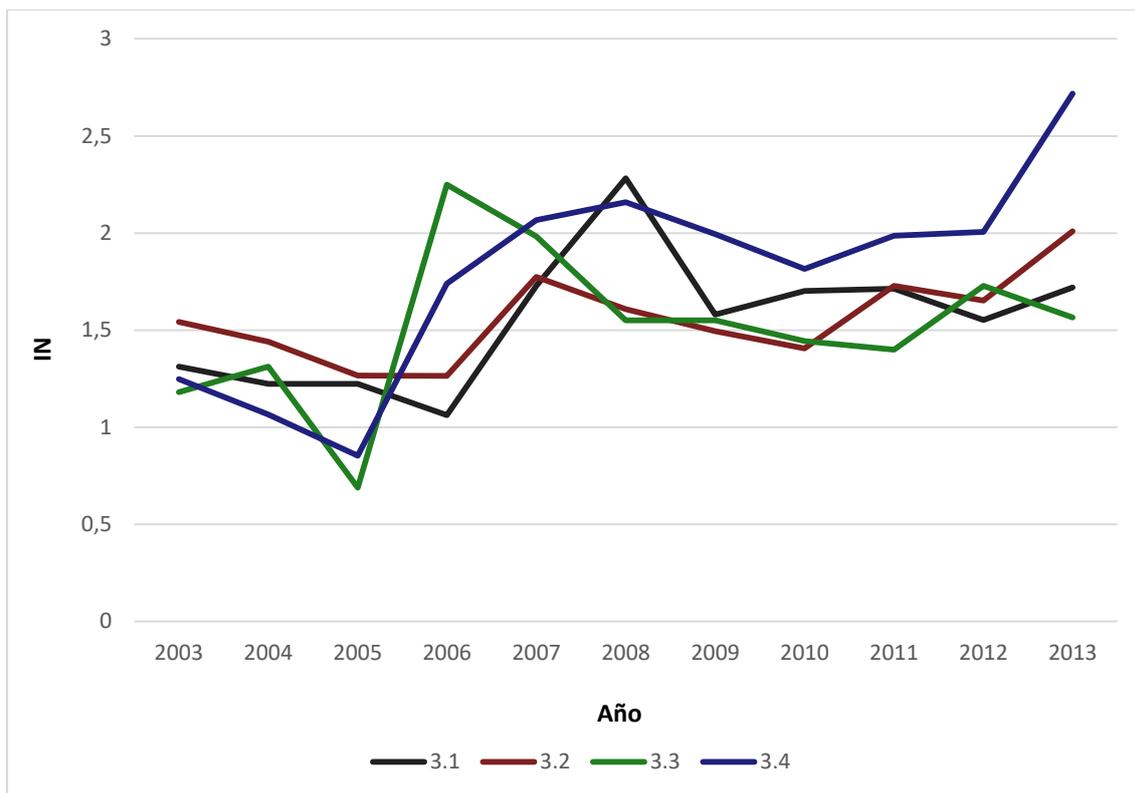


Figura 21: Evolución temporal del IN de España en los clústeres R2 del clúster 3 de R1 (Fuente: Elaboración propia).

El IN de la seguridad alimentaria (Figura 21) presenta una tendencia ascendente a partir de 2005, año en el que los clústeres 3.3 y 3.4 se situaban por debajo de la media 1, pero que suben, el 3.3 en 2006 y el 3.4 en 2007, de tal manera que llegan a duplicar esta media. Lo mismo pasa con el 1.1 y 1.2, los cuales aunque no llegan a caer por debajo de la media, sí que iban hacia abajo, en este caso, antes de 2006 e igualmente experimentan una gran crecida a partir de 2007. Toda esta crecida tiende a estabilizarse situando a todos los clústeres en torno al 1,5 de IN, excepto en el caso del clúster 3.4 que se queda en torno al 2 y al final del periodo escala hasta el 2,7 de IN. Al final del periodo estudiado, todos los clústeres cuentan con un IN superior al inicial (ver % VarIN de la Tabla 18).

Tabla 19: Periodos bursting para el caso español de las keywords del clúster 3 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
bacteriocin	2	2,88	2008	2009	3.1
mastitis	3	3,12	2003	2005	3.1
foodborne pathogen	1	4,76	2008	2008	3.1
real-time pcr	2	4,42	2011	2012	3.1
dairy cow	2	3,31	2003	2004	3.1
cheese	2	3,65	2006	2007	3.1
biofilm	4	5,44	2012		3.1
high hydrostatic pressure	1	2,89	2012	2012	3.1
allergy	3	3,32	2013		3.2
barley	2	3,23	2011	2012	3.2
lc-ms/ms	4	7,39	2012		3.2
composition	2	5,04	2010	2011	3.3
pesticides	2	2,68	2009	2010	3.3
analysis	3	5,25	2010	2012	3.3
antimicrobial activity	1	3,53	2007	2007	3.4
gc-ms	2	5,62	2011	2012	3.4

En la Tabla 19 aparecen los datos relativos al bursting de este clúster general 3 ordenados por sus clústeres R2. Los 4 clústeres que lo forman cuentan con keywords con periodos bursting y que además tienen más de 4 en intensidad. El clúster 3.1 comprende 8 keywords y 3 de ellas superan la media de intensidad: “foodborne pathogen” (4,76) cuyo periodo bursting tan solo dura el año 2008; “real-time pcr” (4,42) con un periodo que va de 2011 a 2012; y “biofilm” (5,44), que empieza en 2012 y sigue activo al final del periodo estudiado. El clúster 3.2 cuenta con 3 keywords con periodos bursting y solo una de ellas supera el 4 de intensidad y casi lo dobla: “lc-ms/ms” (7,39) con un periodo que va de 2012 a la actualidad. El clúster 3.3 tiene 3 keywords y 2 de ellas superan la media de intensidad: “composition” (5,04) con un periodo de 2010 a 2011; y “analysis” (5,25) con un periodo de 2010 a 2012. Finalmente, el clúster 3.4 solo tiene 2 keywords con periodos bursting y es “gc-ms” la que supera el 4 de intensidad con un total de 5,62 y un periodo de 2 años de duración que empieza en 2011 y acaba en 2012.

En cuanto al clúster general 4 y al 5, al principio de este apartado se explica que son los 2 clústeres que más cohesión tienen y concretamente que el 4, a través de su clúster de R2 número 4.1, es el que mayor centralidad presenta. A partir de esto, a continuación se muestran las tablas y gráficos que terminan de analizar estos 2 clústeres por completo.

Tabla 20: Tabla resumen de los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	% VarIN	Percentil	%Exc	%Exc1	%Citados	%	% Var%
4.1 Incremento de vida útil en alimentos vegetales durante su conservación. Métodos y modificaciones de calidad	1432	15,97	1,38	43,38	36,12	15,85	0,70	95,04	11,10	-31,97
4.2 Proceso de fermentación en vino y cerveza como método de conservación	1406	15,75	1,47	21,43	35,46	16,71	0,78	93,31	10,90	-27,26
4.3 Parámetros que condicionan la conservación de alimentos	295	15,80	1,26	-13,75	37,15	13,22	0,00	92,88	2,29	-41,77
4.4 Métodos de análisis estadístico	269	14,67	1,38	10,52	38,75	16,36	0,74	90,71	2,09	-15,36

La Tabla 20 muestra el resumen de los datos obtenidos para el conjunto de clústeres R2 del clúster general 4. En esta ocasión aparece el mayor número de documentos obtenidos de entre todo el conjunto de clústeres R2, así como el mayor porcentaje de documentos con respecto al total, es el clúster 4.1 el que cuenta con estos valores máximos, siendo por tanto el que también los obtiene de entre su conjunto de clústeres pertenecientes al cuarto clúster de R1. Además, también es este clúster 4.1 el que mejores resultados consigue en los indicadores CpD, % VarIN y %Citados, aunque ya sin ser el mejor de los 18 clústeres R2. En IN es el clúster 4.2 el que mayor media presenta, también este mismo clúster ubica a sus documentos en el percentil más bajo del conjunto, el que mejor porcentaje de excelencia 10 y 1 obtiene, aunque en el caso de %Exc1, aun siendo el que mejor dato consigue, no llega a superar la media 1. Por último, aunque todos los valores de % Var% son negativos, es el clúster 4.4 el que mejor resultado obtiene de los 4 clústeres.

En cuanto a los valores más bajos de la Tabla 20, estos están repartidos entre 2 clústeres, el 4.3 y el 4.4. Todos los indicadores en los que el clúster 4,3 tiene los resultados

más bajos, además de en la Tabla 20, son los más bajos de todo el conjunto de clústeres R2, siendo además el único que saca un 0 en un indicador, concretamente en %Exc1. Los otros indicadores en los que este clúster obtiene los resultados más bajos son: IN, %VarIN, %Exc, aunque en este caso sí que llega a superar la media 10 del indicador, y %Var%. El clúster 4.4 es el que menos Ndocc y % tiene, así como el que menos CpD acumula, el que coloca sus documentos en el percentil más alto, y el que menos %Citados consigue; todo esto ya a nivel del conjunto de los 4 clústeres R2 del clúster general 4.



Figura 22: Evolución temporal del IN de España en los clústeres R2 del clúster 4 de R1 (Fuente: Elaboración propia).

En la Figura 22 se observa la evolución del IN de los clústeres R2 del clúster general 4, en este caso se ve durante todo el periodo como el impacto de todos los clústeres se mantiene más o menos estable entre el 1 y el 1,5 de IN, cayendo únicamente de la media 1 el clúster número 4.4 por tan solo una centésima en el año 2007. Por otra parte, este clúster 4.4 también es el único que realmente se aleja del 1,5 de IN llegando a alcanzar un 1,99, consiguiendo así la marca más alta del conjunto de estos 4 clústeres R2. Al final del periodo, excepto el 4.3, los clústeres ven aumentado su IN.

Tabla 21: Periodos bursting para el caso español de las keywords del clúster 4 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
firmness	2	3,63	2003	2004	4.1
storage	3	6,67	2003	2005	4.1
ethylene	5	5,73	2003	2007	4.1
apple	3	2,74	2005	2007	4.1
ripening	1	4,26	2003	2003	4.1
proteolysis	4	9,64	2003	2006	4.1
beer	5	4,54	2011		4.2
lactic acid bacteria	1	4,06	2008	2008	4.2
fermentation	1	4,58	2008	2008	4.2
saccharomyces cerevisiae	1	2,58	2003	2003	4.2
water activity	4	3,17	2004	2007	4.3
ph	6	7,97	2003	2008	4.3
chemometrics	2	2,94	2005	2006	4.4

En la Tabla 21 se muestra el bursting del clúster 4 de R1. En este caso 3 de sus 4 clústeres sí tienen keywords con periodos bursting que superan la intensidad media 4, siendo el 4.4, con solo una keyword que presenta periodo bursting, el único que no lo consigue. En el clúster 4.1, 4 de las 6 keywords que incluye en la Tabla 21 superan la media de intensidad: “storage” con un total de 6,67 de intensidad y con un periodo bursting que va de 2003 a 2005; “ethylene” (5,73) con un periodo de 5 años de duración entre 2003 y 2007; “ripening” (4,26) que comienza y termina su periodo en 2003; y “proteolysis” con una intensidad que dobla a la media llegando hasta el 9,64, con su periodo entre 2003 y 2006. El clúster 4.2 tiene 3 de sus 4 keywords con una intensidad que supera la media: “beer” (4,54) con un periodo iniciado en 2011 y que sigue activo; y “lactic acid bacteria” (4,06) y “fermentation” (4,58), las cuales solo mantienen su periodo durante 2008. En el caso del 4.3 solo una de sus 2 keywords consigue superar la intensidad 4, “ph” llegando hasta el 7,97 y con un largo periodo de moda de 6 años que transcurre entre 2003 y 2008.

Tabla 22: Tabla resumen de los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).

Clúster R2	Ndocc	CpD	IN	%VarIN	Percentil	%Exc	%Exc1	%Citados	%	%Var%
5.1 Antioxidantes y sus efectos	1394	19,07	2,08	38,26	27,45	29,99	3,30	94,62	10,81	147,77
5.2 Antioxidantes vegetales	363	17,55	1,75	41,64	29,29	23,69	1,10	96,69	2,81	40,78

La Tabla 22 presenta los datos recogidos para los indicadores escogidos para este trabajo. Desde el principio se viene observando que el clúster 5 es un clúster importante entre el conjunto de los 5 clústeres R1 que conforman la estructura temática de FS, tanto en el mundo como en este caso concreto de España, algo que sigue confirmándose al comprobar los datos recogidos en esta tabla donde sus clústeres R2 consiguen obtener los mejores resultados, en la mayoría de los indicadores, de entre todo el conjunto de clústeres R2, destacando sobre todo el clúster 5.1, el cual consigue superar a todos los clústeres R2 en el número de citas por documento, en IN, en la colocación de sus documentos en el percentil más bajo, en el %Exc y en la tasa de variación %Var%, llegando hasta casi el 150%. Aparte se encuentra por encima de su compañero de clúster, el 5.2, en Ndocc, en %Exc1 y en %, quedándose el clúster 5.2 con los mejores resultados en la tasa de variación del IN, y en el porcentaje de citados donde además es el mejor de los 18 clústeres R2. Como este clúster 5, solo tiene 2 clústeres R2, los resultados más bajos en cada indicador los obtiene el clúster que no ha conseguido los más altos.

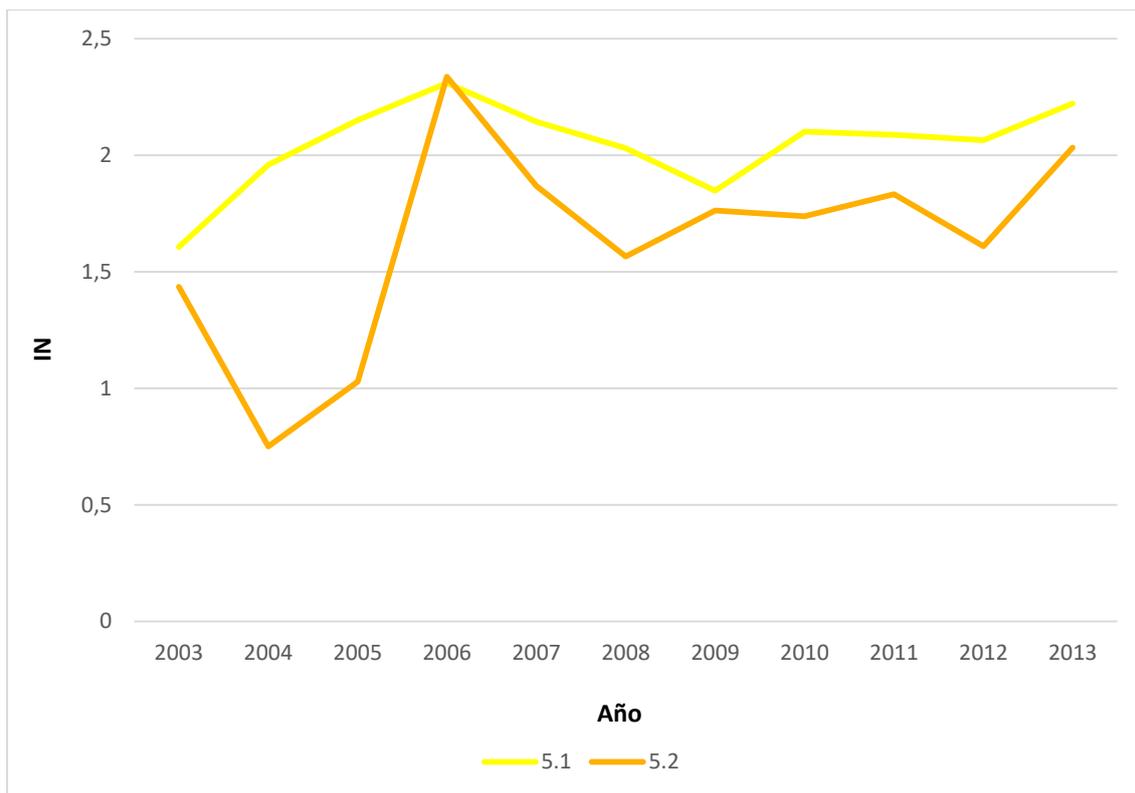


Figura 23: Evolución temporal del IN de España en los clústeres R2 del clúster 5 de R1 (Fuente: Elaboración propia).

Los 2 clústeres R2 del clúster general 5 presentan un gran IN durante todo el periodo estando por encima del 1,5 de impacto en casi todo momento (Figura 23), siendo únicamente en 2004 donde el clúster 5.2 cae por debajo de la media 1, pero volviendo a remontar rápidamente a partir de 2006 para volver a estar por encima de 1,5, llegando incluso en ese mismo 2006 hasta el 2,3 de IN. Por su lado el clúster 5.1 se mantiene durante todo el periodo por encima de 1,5 y al final de este, en 2013, ambos clústeres se encuentran por encima de 2 habiendo aumentado considerablemente su IN inicial.

Tabla 23: Periodos bursting para el caso español de las keywords del clúster 5 de R1, ordenados por los clústeres R2 (Fuente: Elaboración propia).

Keyword	Duración	Intensidad	Comienzo	Fin	Clúster R2
bioactive compounds	3	6,09	2013		5.1
phytochemicals	3	4,47	2013		5.1
antioxidant activity	2	3,65	2014		5.1
oxidative stress	2	5,44	2014		5.1
phenolic acids	6	3,24	2010		5.1
phenolics	3	3,10	2013		5.1
phenolic compounds	3	7,98	2013		5.1
antioxidant capacity	1	2,98	2011	2011	5.1
lycopene	1	2,60	2009	2009	5.2
chlorophyll	2	3,96	2014		5.2

En la Tabla 23 se muestra el bursting del clúster 5 compuesto por un total de 10 keywords de las cuales 8 pertenecen al clúster 5.1 y las otras 2 al 5.2, sin superar ninguna de estas 2 la intensidad 4, aunque “chlorophyll” no lo hace por muy poco quedándose en 3,96. El clúster 5.1, por su parte, consigue en la mitad de sus 8 keywords superar la barrera del 4 de intensidad: “bioactive compounds” (6,09) inicia su periodo en 2013 y sigue encendido al final del periodo estudiado al igual que “phytochemicals” (4,47); “oxidative stress” (5,44) empieza en 2014 y sigue activa; y por último “phenolic compounds” que es la que más intensidad presenta (7,98) y también sigue activa hasta la actualidad habiendo comenzado su periodo bursting en 2013.

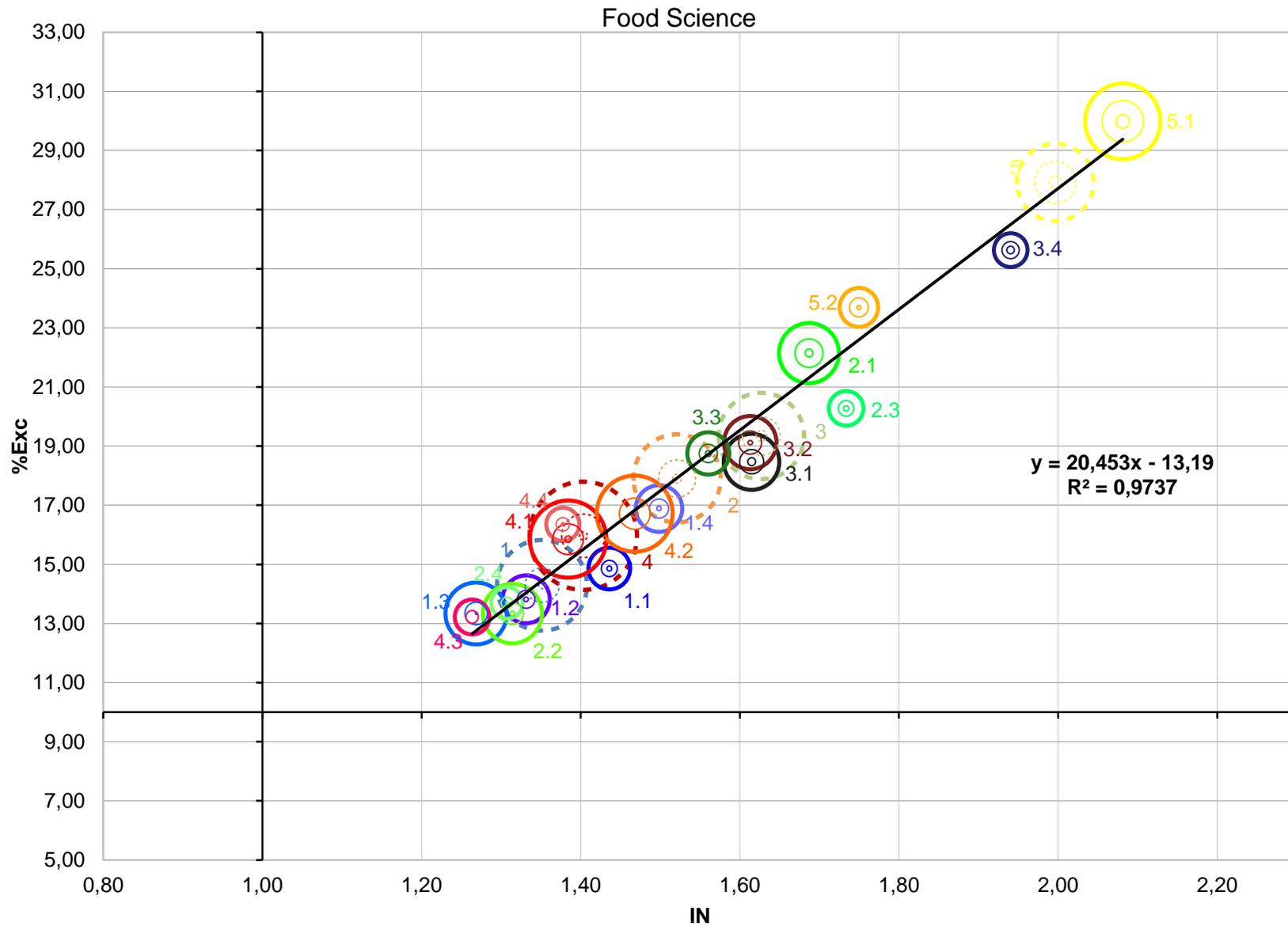


Figura 24: Porcentaje de los documentos excelentes en el 10% como “Research Guarantor” comparado con el IN de los 5 clústeres R1 de España (círculos discontinuos) y con los 18 de R2 (círculos continuos). Los 3 círculos concéntricos de cada clúster se corresponden, respectivamente, con los indicadores Ndocc, %Exc, y %Exc1 (Fuente: Elaboración propia).

Al igual que en el apartado anterior relativo al mundo, a continuación se analiza la Figura 24 correspondiente al gráfico de dispersión que compara el %Exc con el IN de los 5 clústeres R1 y los 18 clústeres R2 del caso español.

De la misma manera que en el caso mundial, aquí también se observa una correlación fuerte positiva entre las dos variables, con todo los clústeres concentrados en torno a la línea de tendencia, y aumentando el valor del eje Y a la vez que aumenta el del eje X y viceversa. A través del coeficiente de correlación de Pearson, igual a 0,9737, se confirma esta correlación fuerte, aunque algo menor que en el caso mundial.

Por otra parte, en este caso sí que todos los clústeres superan la media 1 de IN y la media 10 de %Exc.

3.4. Comparativa del caso español y mundial a través de los resultados relativos de España con respecto al mundo

A través de la visualización de las distintas figuras presentadas en los 2 apartados anteriores y de la creación de nuevas tablas y figuras realizadas en base a datos relativos entre los 2 casos, se va a llevar a cabo un análisis comparativo que muestre las diferencias y semejanzas entre los 2 casos presentados, así como los aspectos en los que uno supera al otro y viceversa.

Al comparar los 2 mapas generales de copalabras (Figura 1 y Figura 15) se observa que el clúster 5 de R1, dedicado a los *antioxidantes en alimentos*, es el que mayor cohesión presenta en los 2 casos, aunque en el caso mundial es algo mayor. La principal diferencia se encuentra en lo referente a la centralidad, mientras que en España hay una ocupación bien definida del centro del mapa, en el caso mundial no ocurre lo mismo y de hecho no se puede decir que haya ningún clúster, ni de R1 ni de R2, que destaque por su centralidad.

En lo que respecta a las relaciones entre clústeres, estas se producen cuando las keywords que los forman comparten un mismo documento, es decir, cuando en un documento aparecen keywords de 2 clústeres distintos. Mientras en más documentos aparezcan 2 keywords concretas, mayor será su tamaño en la representación de los mapas y más fuerte será la relación entre ambos. Sabiendo esto, determinar qué clústeres están mejor relacionados con otros es posible atendiendo a las keywords mejor relacionadas.

Observando las relaciones que se establecen en cada uno de los dos casos, se comprueba que las keywords que configuran los 5 clústeres, como es lógico, mantienen las relaciones más significativas con keywords de su mismo clúster. Pero dejando a un lado esto y sabiendo que todos los clústeres mantienen relaciones con todos, aquí se van a destacar las relaciones más fuertes y es donde se observarán las principales diferencias entre los dos casos:

- El clúster 1, dedicado a la *composición de alimentos y nutrición*, en el caso de España, situado principalmente a la derecha del mapa, mantiene fuertes conexiones con los clústeres relativos a la *seguridad alimentaria* y a la *conservación de alimentos y vida útil* (número 3 y 4). En el mapa del mundo las relaciones más fuertes son establecidas únicamente con el último.
- En ambos casos, en el clúster dedicado al *procesado y modificado de alimentos* (número 2), se observa una leve dispersión de sus keywords y pocas relaciones externas. Además, de un primer vistazo es el que presenta las keywords más pequeñas, dándonos a entender que puede ser el clúster con menor cantidad de documentos asociados a cada una de sus keywords.
- El clúster 3 de *seguridad alimentaria*, también en ambos casos se encuentra muy disperso por el mapa aunque en el caso español está más concentrado en la zona izquierda del mismo. Mantiene buenas relaciones, en mayor o menor medida, con el resto de clústeres excepto con el número 2 en el caso de España.
- Es en el clúster 4 dedicado a la *conservación de alimentos y vida útil*, donde encontramos las diferencias más destacadas entre uno y otro caso en lo que a relaciones con otros clústeres se refiere. Mientras que en España mantiene fuertes relaciones con el clúster 5 de *antioxidantes en alimentos*, y en menor medida con el clúster 2, en el mundo estas relaciones las mantiene con el clúster 1 de *composición de alimentos y nutrición*.
- Por último, el clúster 5 de los *antioxidantes en alimentos*, también presenta diferencias. En el caso español la buena relación con la *conservación de alimentos y vida útil* (clúster 4) es recíproca, mientras que en el mundo solo mantiene fuertes relaciones con el clúster 3 de *seguridad alimentaria* y

además únicamente a través de sus dos keywords principales. En ambos casos vemos que está bien localizado en una zona específica de su mapa correspondiente.

En cuanto a la densidad presentada en uno y otro caso, se observa claramente que el mapa de densidad del caso español es mucho más denso que el del caso mundial, presentando más zonas destacadas y por lo tanto teniendo más clústeres representados a través de esta densidad.

El siguiente mapa (Figura 25) está coloreado en función de la citación normalizada relativa de España con respecto al mundo. En él se observa que casi todas las keywords superan la media 1, esto quiere decir que la citación normalizada de España es superior, de manera relativa, a la citación normalizada del mundo, siendo las keywords que más se acercan al color rojo las que mayor diferencia positiva tienen en cuanto a impacto científico con respecto al mundo.

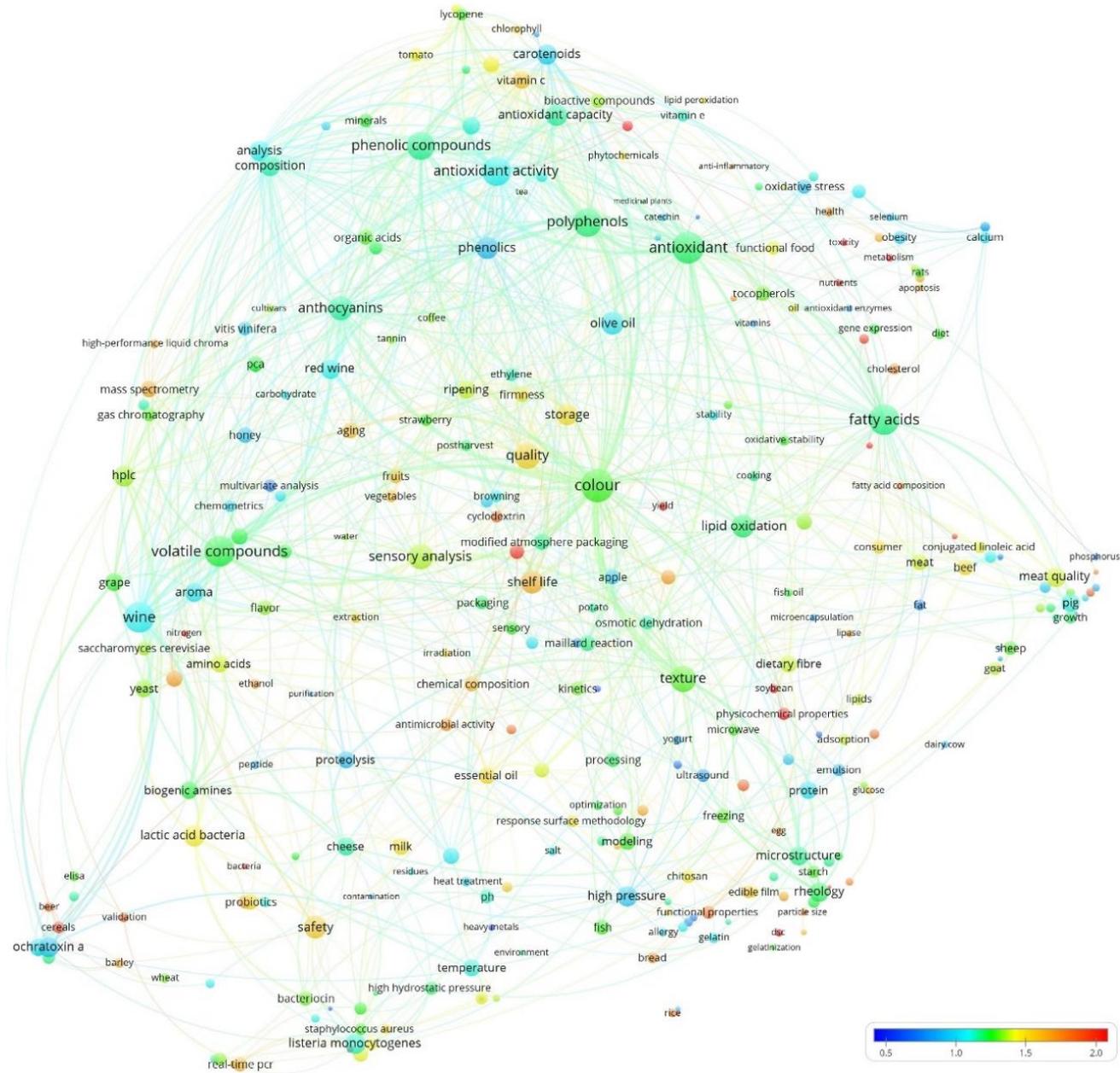


Figura 25: Vista general del mapa de copalabras donde los colores marcan la Citación Normalizada Relativa de España respecto al mundo (Fuente: Elaboración propia).

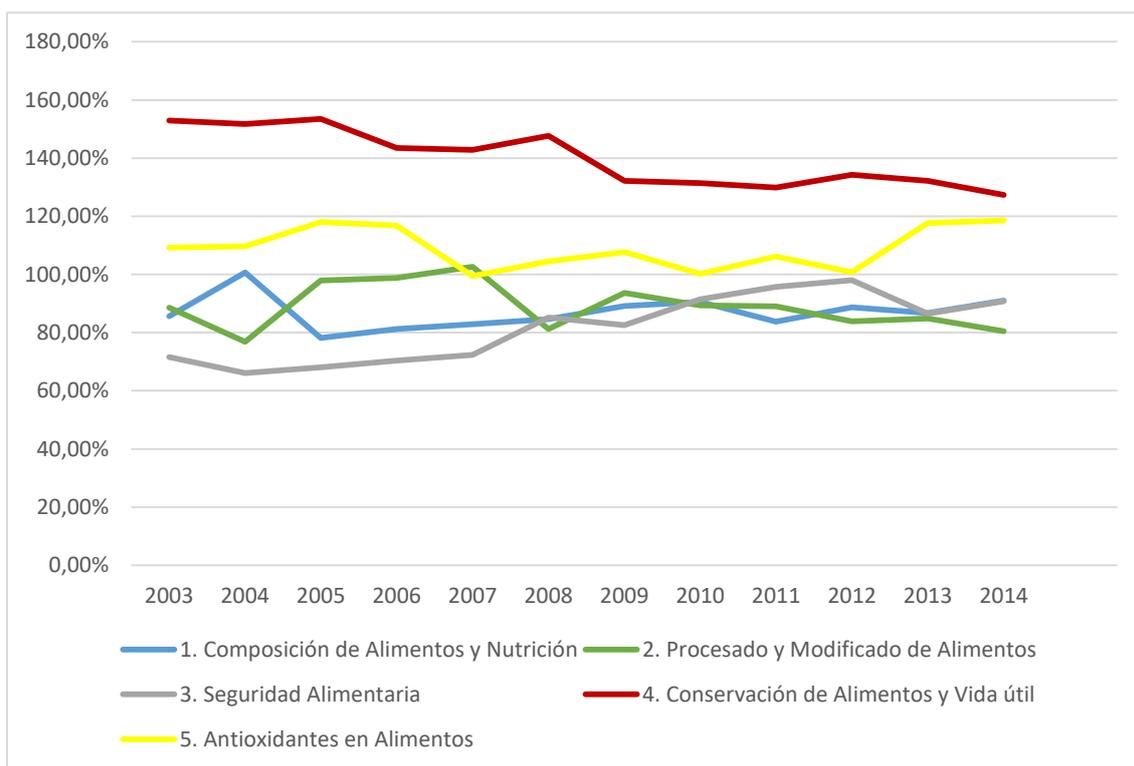


Figura 26: Evolución de la producción relativa de España respecto al mundo (Fuente: Elaboración propia).

En la Figura 26 aparece la evolución de la producción relativa de la investigación española en FS con respecto a la del mundo, a través de los 5 clústeres que conforman la disciplina. Más concretamente, la ratio de producción en cada clúster con respecto al total en FS en España, y todo ello frente a la misma ratio a nivel mundial.

Sabiendo que encontrarse en la línea del 100% del gráfico es equivalente a producir el mismo porcentaje en cada clúster, respecto al total en FS, que el conjunto de la producción mundial, todo lo que está por encima del 100% supera al mundo y viceversa. De este modo, se observa cómo los trabajos españoles que se ocupan de los temas representados por el clúster número 4 de FS, el dedicado a la *conservación de alimentos y vida útil*, suponen, de manera relativa y durante todo el periodo estudiado, una mayor producción que la generada por el conjunto de la investigación mundial. La mayor diferencia positiva para este clúster se produce en 2005, donde la producción española supera, siempre hablando de manera relativa, en un 53,45% a la mundial.

Al igual que el clúster general 4, el clúster 5, dedicado a los *antioxidantes en alimentos*, también supera a la producción mundial durante todo el periodo, exceptuando el año 2007 que se queda a un 0,59% de igualar esta producción. En esta ocasión las diferencias son algo más bajas que las del clúster 4, pero aun así la producción sigue siendo mayor. En el año 2014 se produce la mayor diferencia, superando la producción española a la mundial en un 18,52%.

Como se ve en el gráfico el resto de clústeres se encuentran por debajo de la barrera del 100% indicando que generan menor porcentaje que el conjunto mundial en sus respectivas áreas, integrantes todas de las FS. Aun así, hay dos momentos en los que estos clústeres superan levemente a la producción mundial: en 2004 el campo número 1, dedicado a la *composición de alimentos y nutrición*, lo hace por un 0,66% y en 2007 el número 2, dedicado al *procesado y modificado de alimentos*, lo hace por un 2,64%. El resto de años los clústeres 1, 2 y 3 están siempre por debajo, estableciéndose la mayor diferencia negativa en 2004, donde el clúster 3, cae en un 33,95%. Hay que tener en cuenta, cómo se compara el porcentaje de producción dentro de FS, para que unos estén por encima del mundo los otros tienen que estar por debajo. Porque en términos absolutos como podemos ver en el trabajo de 2016 de Guerrero-Bote, Olmeda-Gómez, y Moya-Anegón, la producción científica en FS en España en el periodo de estudio tiene un crecimiento sostenido similar al mundial.

Tabla 24: Tabla resumen de los clústeres R1 y R2 con los datos relativos de España respecto al mundo
(Fuente: Elaboración propia).

Clúster R1	%Ndoocc	CpD	IN	Percentil	%Exc	%Exc1	%Citados	IET
1. Composición de alimentos y nutrición	6,54	1,33	1,33	0,78	1,49	0,80	1,09	0,94
2. Procesado y Modificado de Alimentos	6,61	1,27	1,28	0,79	1,45	1,11	1,09	0,95
3. Seguridad Alimentaria	6,27	1,26	1,38	0,74	1,56	1,60	1,09	0,90
4. Conservación de Alimentos y Vida útil	10,49	1,35	1,33	0,77	1,57	0,96	1,10	1,50
5. Antioxidantes en Alimentos	7,73	1,17	1,21	0,75	1,31	1,02	1,08	1,11
Clúster R2	%Ndoocc	CpD	IN	Percentil	%Exc	%Exc1	%Citados	IET
1.1	3,97	1,38	1,35	0,81	1,41	0,96	1,09	0,57
1.2	5,54	1,38	1,46	0,77	1,82	1,75	1,08	0,79
1.3	9,06	1,22	1,24	0,81	1,35	0,16	1,09	1,30
1.4	8,47	1,26	1,33	0,75	1,47	0,86	1,08	1,21
2.1	5,81	1,26	1,34	0,78	1,62	1,22	1,07	0,83
2.2	9,31	1,19	1,20	0,82	1,27	0,80	1,08	1,33
2.3	6,22	1,29	1,33	0,77	1,39	1,42	1,09	0,89
2.4	4,62	1,28	1,18	0,82	1,15	0,38	1,11	0,66
3.1	6,80	1,20	1,30	0,78	1,40	1,94	1,08	0,97
3.2	5,76	1,40	1,51	0,69	1,74	1,21	1,12	0,83
3.3	6,71	1,24	1,35	0,75	1,58	1,33	1,09	0,96
3.4	6,37	1,10	1,36	0,73	1,54	1,80	1,08	0,91
4.1	10,19	1,37	1,37	0,75	1,61	1,04	1,12	1,46
4.2	13,24	1,31	1,31	0,79	1,52	1,20	1,08	1,90
4.3	7,41	1,42	1,28	0,77	1,56	0,00	1,11	1,06
4.4	8,73	1,20	1,11	0,88	1,22	0,44	1,07	1,25
5.1	7,72	1,14	1,20	0,75	1,31	1,02	1,07	1,11
5.2	9,28	1,27	1,35	0,70	1,54	0,80	1,10	1,33

Si de manera relativa se comparan los datos de producción, citación y excelencia de los clústeres R1 y R2 de España y el mundo, se puede saber cuán relevante puede llegar a ser la investigación científica española, en FS, a nivel global. Esto es lo que se puede ver a través de la Tabla 24 anterior.

En la primera columna de esta Tabla 24 se ve el porcentaje que supone la producción española en el conjunto del mundo, siendo el clúster dedicado a la *conservación de alimentos y vida útil* el que más aporta con un 10,49%, es decir, de todo lo que se publica en el mundo sobre este tema, el 10,49% proviene de España. Entre los

clústeres R2, el 4.2, dedicado al *proceso de fermentación en vino y cerveza como método de conservación*, es el que mayor porcentaje ostenta, un 13,24%.

Para el resto de indicadores se establece una media de 1 y tomándola como referencia se sabe que si el valor de un indicador es mayor que 1 quiere decir que España está superando al mundo para ese indicador concreto, si es menor, el mundo supera a España.

En la columna de CpD se observa que tanto los clústeres R1 como los clústeres R2 superan en todas las ocasiones la media 1, esto quiere decir que de manera relativa España genera mayor porcentaje de citas que el conjunto mundial. Por ejemplo, en el clúster 4, España genera un 35% más de citas que el mundo.

En la columna de IN pasa igual, en todos los casos se supera la media 1 implicando el mismo significado que en la columna anterior para España.

En el caso de los percentiles la cosa cambia, ya que vemos que todos los clústeres generales y específicos se encuentran siempre por debajo de la media 1, esto se debe a que el indicador tiene sentido contrario, cuanto más citado es un trabajo en un percentil más bajo se encuentra, por ello en este caso, España vuelve a superar al mundo en todos los clústeres.

En cuanto al porcentaje de documentos excelentes entre el 10% y 1% más citado se comprueba que siempre se supera la media 1 para el 10%, pero no así para el 1% donde hay varios clústeres R1 y R2 que no llegan a la media 1, siendo el primer indicador en el que el mundo supera a España en algunos clústeres. En la columna de porcentaje de documentos citados al menos en una ocasión, vemos que nuevamente se supera la media 1 en todos los casos.

Por último, en la columna dedicada al índice de especialización temática (IET) se observa que 2 de los clústeres R1, el 4 y el 5, superan la media 1, lo que significa que estos 2 clústeres presentan mayor especialización temática en España que en el mundo, y al contrario con los 3 que no superan la media. En el caso de los clústeres R2, 9 de ellos están por encima del mundo y 9 por debajo. De entre los clústeres que superan al mundo es destacable el 4.2, dedicado al *proceso de fermentación en vino y cerveza como método*

de conservación, que supera al mundo en un 90%, siendo el que mayor porcentaje de diferencia obtiene.

Finalmente, comparando las tablas bursting de los 2 casos se comprueba que, en ambos casos, los 5 clústeres generales tienen alguna keyword con intensidad superior a la media de cada caso, pero no ocurre lo mismo con los específicos. En el caso español ni el 2.3, ni el 4.4, ni el 5.2, tienen keywords con intensidad superior a 4; en el mundo, además de los 3 clústeres que no tienen keywords con intensidad superior a la media, tampoco consiguen este objetivo el 2.4, el 3.4 y el 4.2, el 3.4 ni siquiera tiene keywords con periodos bursting.

Aunque en ambos casos los periodos bursting tienen una duración similar, es cierto que en el caso español encontramos keywords que mantienen su periodo de moda durante 5 y 6 años, algo que no ocurre en el ámbito mundial donde los máximos periodos son de 4 años. Por otra parte, aunque hay una cantidad considerable de keywords que presentan periodos bursting en los 2 casos, la mayoría solo aparece en uno de los 2.

4. CONCLUSIONES

4. CONCLUSIONES

Con el análisis de copalabras y la generación de los mapas correspondientes de la “specific subject area” FS, se observa claramente una división estructural en 5 grandes clústeres, que conformarán las 5 áreas generales de FS. A su vez se observa una segunda división en otros 18 clústeres con mayor nivel de resolución, que se corresponderán con las 18 subáreas de FS.

En cuanto a las relaciones que mantienen los 5 campos generales entre sí, cabe destacar en primer lugar que sus keywords se relacionan, por regla general, de mejor manera con keywords de su mismo clúster. Pero salvando este dato que puede parecer obvio, observamos diferentes relaciones entre clústeres en los resultados de uno y otro caso, es decir, las áreas y subáreas que conforman la estructura temática de FS no se relacionan igual a nivel global que a nivel específico de España en este caso. Todas estas diferencias en las relaciones de un caso y otro, vienen a decirnos que aunque la investigación internacional en FS puede y debe estar sujeta a una misma estructura temática, según el caso específico estudiado, las relaciones estructurales pueden variar y formarse de distinta manera. Recordemos que las relaciones entre áreas, son consecuencia de los trabajos que incluyen keywords pertenecientes a 2 áreas distintas.

A simple vista, si comparamos los mapas de copalabras contruidos para el mundo y para España, se ve claramente que presenta una mayor cohesión interna el de España, con las keywords pertenecientes a las distintas áreas más concentradas en torno a sus áreas que en el caso mundial. Si además apoyamos esta observación en los mapas de densidad, nuevamente se confirma al comprobar que el mapa de densidad del caso español es mucho más intenso que el del caso mundial, algo que se aprecia fácilmente por el nivel de puntos rojos destacados en uno y otro mapa. Todo esto es indicativo de que en general la investigación en FS en España está más especializada y posiblemente goza de una mejor organización de la misma, algo que parece lógico ya que el territorio analizado es mucho menor y mucho más homogéneo que el mundo en general.

Atendiendo a la centralidad temática, en el caso mundial no podemos decir que ninguna de las áreas destaque especialmente por este hecho. Mirando la concentración de keywords situadas más al centro del mapa de copalabras, observamos que estas pertenecen a varios clústeres y que no hay ninguna que destaque especialmente. No se

puede decir que el área tenga unas bases intelectuales comunes claramente identificadas. Algo que no pasa en el caso español, ya que aquí sí que se observa que es el clúster 4 de, *conservación de alimentos y vida útil*, el que consigue esta centralidad, concretamente su clúster R2 número 4.1, dedicado al *incremento de vida útil en alimentos vegetales durante la conservación. Métodos y modificaciones de calidad*.

Y en el tema de la especialización, en ambos casos, y como se empieza a observar desde el inicio del estudio, concretamente el clúster 5.1, dedicado a los *antioxidantes y sus efectos*, es el que ostenta la superioridad en este ámbito, algo que se confirmaba mientras más datos se recogían. El tema que ocupa a este clúster 5.1, es el que mejores resultados obtiene en la mayoría de los indicadores estudiados en los 2 casos, pero sobre todo en el mundo, y todo ello es indicativo de su solidez, especialización y alta cohesión interna. En todos los gráficos y tablas del trabajo se observa esta superioridad. Para el caso español también es pertinente nombrar al clúster 4 como un clúster bien especializado, algo que se refleja especialmente a través de su clúster de R2 número 4.2 que supera en un 90% en especialización temática al mundo.

Hemos comprobado también que hay una gran relación entre el impacto normalizado y el %Exc y %Exc1, aunque hay subclústeres con distribuciones de impactos más sesgadas lo que hace que tengan más papers excelentes de los que les corresponderían por citación normalizada.

El impacto científico en el mundo, en el periodo de tiempo que hemos estudiado, goza de buena salud superando en IN y en %Exc1 la media 1, en la mayoría de los casos, y también superando la media 10 en %Exc, en la mayoría de los casos. Estar por encima de estas medias indica un impacto positivo del clúster en cuestión.

En el caso del IN también es bueno atender a su %VarIN ya que este indicador nos va a mostrar cómo ha evolucionado el IN de cada clúster durante el periodo de tiempo estudiado. Observamos que el clúster 5 del caso mundial es el que más decrece en cuanto a su tasa de variación porcentual de IN y esto puede deberse en gran medida al fuerte incremento del número de papers que acumula durante el periodo. Es decir, a principios del periodo comienza a establecerse como un tema que capta mucha atención y por tanto citación, lo que supone un gran impacto científico. Con el tiempo, esto se traduce en la captación de un gran número de papers que van difuminando ese impacto captado. El

tema de los *antioxidantes y sus efectos*, es el tema que más crece en FS yendo desde el 4%, al inicio del periodo, hasta casi 12% al final de este, además en un crecimiento constante que sólo baja mínimamente en 2012 y 2013 para volver a crecer en 2014.

En España, el impacto científico en comparación con el mundo, destaca frente a este y supera al conjunto global en lo que a IN y %Exc se refiere, además haciéndolo en todos los clústeres. En %Exc1 no todos los clústeres superan al mundo, pero sí más de la mitad de ellos.

Por último, al hablar de las modas (bursting) vemos que estas afectan, en los 2 casos, a los 5 clústeres generales, ya que todos tienen alguna de sus keywords con periodos de moda destacados. Algo que no ocurre igual en los clústeres R2, donde no todos cuentan con keywords con periodos de moda. Eso sí, si comparamos con el mundo, España presenta más clústeres R2 que el mundo con periodos de moda y además se mantienen activos durante periodos de tiempo más amplios.

La mayoría de las palabras clave aparecen en uno u otro caso, no suelen coincidir en los dos, por ello podemos afirmar que las modas de keywords no coinciden en el ámbito español y mundial.

Creemos que sería interesante realizar más estudios de este tipo en distintos países, ya que ayudarían a conocer cómo se adapta la estructura temática creada para la disciplina FS. Así mismo, si fuera necesario esta estructura podría adaptarse a las características concretas de cada territorio para ayudar mejor a los investigadores que desarrollan su trabajo en un país concreto. Además, conociendo los temas destacados de cada país, se pueden establecer nuevas metodologías y estudios que nos acerquen a conocer lo afinada que está nuestra estructura temática.

En cuanto a España, cabe destacar que sin duda es una potencia en lo que a investigación en FS se refiere y no tanto por su volumen de producción científica, que también, sino más bien por la calidad de esta. En los resultados para el caso español de este trabajo se observa como en numerosas ocasiones y para diferentes aspectos de la investigación en FS, España supera al conjunto mundial.

En definitiva, la estructura temática creada se adapta muy bien al contexto de investigación en FS que actualmente se desarrolla, estableciendo una jerarquía bien

definida formada por unos campos y subcampos que permiten su estudio pormenorizado o general. Además puede considerarse una estructura global, ya que al aplicarla al caso específico de España, se comprueba que funciona bien, no estando sujetos los resultados de este caso a los obtenidos en el mundial que podrían entenderse como un hándicap.

5. Bibliografía

5. BIBLIOGRAFÍA

- Alfaraz, P. H., y Calviño, A. M. (2004). Bibliometric study on food science and technology: Scientific production in Iberian-American countries (1991-2000). *Scientometrics*, 61(1), 89–102. doi: 10.1023/B:SCIE.0000037365.53469.91
- Blackwell, W. (2015). *World Population Prospects: the 2015 Revision, Key Findings and Advance Tables*. Recuperado de <http://www.ncbi.nlm.nih.gov/pubmed/15003161%5Cnhttp://cid.oxfordjournals.org/lookup/doi/10.1093/cid/cir991%5Cnhttp://www.scielo.cl/pdf/udecada/v15n26/art06.pdf%5Cnhttp://www.scopus.com/inward/record.url?eid=2-s2.0-84861150233&partnerID=tZOtx3y1>
- Bornmann, L. (2011). Scientific Peer Review. *Annual Review Of Information Science And Technology*, 45(1), 197–245. doi: 10.1002/aris.2011.1440450112
- Bornmann, L., y Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80. doi: 10.1108/00220410810844150
- Bornmann, L., Moya-Anegón, F., y Leydesdorff, L. (2012). The new excellence indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*, 6(2), 333–335. doi: 10.1017/CBO9781107415324.004
- Callon, M., Courtial, J. P., y Laville, F. (1991). Co-Word analysis as a tool for describing a network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22, 155–205
- Chinchilla-Rodríguez, Z., y Olmeda-Gómez, C. (2010). Producción y colaboración científica en agroalimentación. En L. Sanz-Menéndez y L. Cruz-Castro (Eds.), *Análisis Sobre Ciencia E Innovación En España (366–399)*. Madrid, España: FECYT
- Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago, Estados Unidos: University of Chicago Press
- Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 7(3), 171–176
- Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., y Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal*, 22(2), 338–342. doi: 10.1096/fj.07-9492LSF
- Frame, J. D. (1977). Mainstream research in Latin America and the Caribbean. *Interciencia*, 2(3), 143–148
- Garfield, E. (1955). Citation Indexes for Science: A new dimension in documentation through Association of Ideas. *Science*, 122(3159), 108–111
- Garfield, E., y Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, 14(3), 195–201. doi: 10.1002/asi.5090140304
- Gerland, P., Raftery, a. E., Ev Ikova, H., Li, N., Gu, D., Spoorenberg, T., ... Wilmoth, J. (2014). World population stabilization unlikely this century. *Science*, 346(6206), 234–237. doi: 10.1126/science.1257469
- Guerrero-Bote, V. P. (2015). La evaluación de la actividad científica. *Lección inaugural curso académico 2015-2016*. Universidad de Extremadura, Cáceres
- Guerrero-Bote, V. P., y Moya-Anegón, F. (2015). Analysis of Scientific Production in Food Science from 2003 to 2013. *Journal of Food Science*, 80(12), R2619-R2626. doi: 10.1111/1750-3841.13108
- Guerrero-Bote, V. P., Olmeda-Gómez, C., y Moya-Anegón, F. (2016). La Ciencia De Los

- Alimentos Georreferenciada. Aproximación Bibliométrica a Nivel Institucional. *El Profesional de La Información*, 25(1), 25–34
- Hinze, S., y Grupp, H. (1996). Mapping of R&D structures in transdisciplinary areas: New biotechnology in food sciences. *Scientometrics*, 37(2), 313–335. doi: 10.1007/BF02093627
- IFT. (2016). About Food Science and Technology. Recuperado de <http://www.ift.org/knowledge-center/learn-about-food-science/food-facts/about-fs-and-t.aspx>
- Irvine, J., Martin, B., Peacock, T., y Turner, R. (1985). Charting the decline of British science. *Nature*, 316, 587–590
- Kleinberg, J. (2003). Bursty and Hierarchical Structure in Streams. *Data Mining and Knowledge Discovery*, 7(4), 373–397. doi: 10.1023/A:1024940629314
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8), 707–710
- Leydesdorff, L., Moya-Anegón, F., y Guerrero-Bote, V. P. (2010). Journal maps on the basis of scopus data: A comparison with the journal citation reports of the ISI. *Journal of the American Society for Information Science and Technology*, 61(2), 352–369. doi: 10.1002/asi.21250
- López-Illescas, C., Moya-Anegón, F., y Moed, H. F. (2009). Comparing bibliometric country-by-country rankings derived from the Web of Science and Scopus: the effect of poorly cited journals in oncology. *Journal of Information Science*, 35(2), 244–256. doi: 10.1177/0165551508098603
- Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56–63
- Merton, R. K. (1988). The Matthew Effect in Science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623. doi: 10.1086/354848
- Miguel, S., Chinchilla-Rodríguez, Z., y Moya-Anegón, F. (2011). Open Access and Scopus: A New Approach to Scientific Visibility From the Standpoint of Access. *Journal of the American Society for Information Science and Technology*, 62(6), 1130–1145. doi: 10.1002/asi.21532
- Mingers, J., y Leydesdorff, L. (2015). A Review of Theory and Practice in Scientometrics. *European Journal of Operational Research*, 246(1), 1–19. Retrieved from doi: 10.1016/j.ejor.2015.04.002
- Moed, H. F. (2006). *Citation Analysis in Research Evaluation*. doi: 10.1007/1-4020-3714-7
- Muscio, A., y Nardone, G. (2012). The determinants of university-industry collaboration in food science in Italy. *Food Policy*, 37(6), 710–718. doi: 10.1016/j.foodpol.2012.07.003
- Nalimov, V. V., y Mulchenko, Z. M. (1971). *Measurement of Science. Study of the Development of Science as an Information Process*. Washington DC, Estados Unidos: Foreign Technology Division
- Neff, M. W., y Corley, E. A. (2009). 35 years and 160,000 articles: A bibliometric exploration of the evolution of ecology. *Scientometrics*, 80(3), 657–682. doi: 10.1007/s11192-008-2099-3
- Newman, M., y Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 26113. doi: 10.1103/PhysRevE.69.026113
- Noack, A. (2007). Energy Models for Graph Clustering. *Journal of Graph Algorithms and Applications*, 11(2), 453–480. doi: 10.7155/jgaa.00154
- Noack, A. (2009). Modularity clustering is force-directed layout. *Physical Review E*, 79(2), 26102. doi: 10.1103/PhysRevE.79.026102
- Price, D. J. D. S. (1963). *Little Science , Big Science*. New York, Estados Unidos: Columbia University Press

- Price, D. J. D. S. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515. doi: 10.1126/science.149.3683.510
- Price, D. J. D. S. (1976). A General Theory of Bibliometric and Other Cumulative Advantage Processes. *Journal of the American Society for Information Science*, 27(5), 292–306. doi: 10.1002/asi.4630270505
- Rehn, C., y Kronman, U. (2008). *Bibliometric handbook for Karolinska Institutet*. doi: 10.13140/RG.2.1.1480.9447
- Romo-Fernández, L. M., Guerrero-Bote, V. P., y Moya-Anegón, F. (2013). Co-word based thematic analysis of renewable energy (1990-2010). *Scientometrics*, 97(3), 743–765. doi: 10.1007/s11192-013-1009-5
- Seetharam, G., y Ravichandra, R. I. (1999). Growth of food science and technology literature: A comparison of CFTRI, India and the world. *Scientometrics*, 44(1), 59–79. doi: 10.1007/BF02458478
- SM. (2017). Cluster. En *Diccionario Clave*. Recuperado de <http://clave.smdiccionarios.com/app.php>
- Van-Eck, N. J. (2011). *Methodological Advances in Bibliometric Mapping of Science*. Erasmus Research Institute of Management, Rotterdam
- Van-Eck, N. J., y Waltman, L. (2007). VOS : A New Method for Visualizing Similarities between Objects. En H. J. Lenz y R. Decker (Eds.), *Advances in data analysis: Proceedings of the 30th annual conference of the German Classification Society* (299-306). Heidelberg, Alemania: Springer
- Van-Eck, N. J., y Waltman, L. (2009). VOSviewer: A Computer Program for Bibliometric Mapping. En B. Larsen y J. Leta (Eds.), *Proceedings of the 12th International Conference on Scientometrics and Informetrics* (886–897). Rio de Janeiro, Brasil: ISSI
- Van-Eck, N. J., Waltman, L., Dekker, R., y Van Den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405–2416. doi: 10.1002/asi.21421
- Waltman, L., Van-Eck, N. J., y Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629–635. doi: 10.1016/j.joi.2010.07.002
- Wouters, P. (1999). *The Citation Culture*. Universidad de Amsterdam, Amsterdam.
- Wouters, P. (2014). The Citation: From Culture to Infrastructure. En B. Cronin y C. R. Sugimoto (Eds.), *Beyond Bibliometrics: Harnessing Multidimensional Indicators Of Scholarly Impact* (47–66). Massachusetts, Estados Unidos: MIT Press
- Zhou, G., Zhang, W., y Xu, X. (2012). China's meat industry revolution: Challenges and opportunities for the future. *Meat Science*, 92(3), 188–196. doi: 10.1016/j.meatsci.2012.04.01

6. Anexos

6. ANEXOS

Anexo I. Producción absoluta de los clústeres R1 y R2 del caso mundial.

Clúster R1	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
1. Composición de alimentos y nutrición	1574	1732	1774	2049	2444	2510	2882	2992	3260	3806	4092	4138
2. Procesado y modificado de alimentos	1346	1398	1460	1867	2363	2413	2766	2702	3108	3382	3878	4487
3. Seguridad alimentaria	1338	1381	1572	1877	2371	2494	2881	2963	3215	3549	3874	4063
4. Conservación de alimentos y vida útil	1355	1408	1497	1876	2378	2385	2632	2601	2856	2992	3454	3720
5. Antioxidantes en alimentos	459	558	673	879	1427	1579	1864	2075	2472	2610	3038	3267
Clúster R2	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
1.1	382	458	488	567	723	761	902	940	1093	1309	1528	1526
1.2	586	673	609	697	745	760	842	883	895	1134	1049	1057
1.3	487	515	541	652	810	787	956	921	997	1096	1186	1230
1.4	334	338	359	389	441	491	527	558	620	685	751	732
2.1	633	677	714	872	1136	1164	1363	1251	1507	1606	1898	2249
2.2	463	469	478	637	806	766	856	814	856	901	1020	1118
2.3	187	192	221	256	303	347	407	439	493	555	593	687
2.4	176	206	201	267	339	366	396	460	527	657	752	873
3.1	552	563	610	688	839	810	978	999	1066	1285	1344	1414
3.2	467	512	596	693	880	984	1074	1073	1188	1343	1462	1524
3.3	215	204	275	340	420	523	645	681	740	752	824	816
3.4	173	188	202	267	371	333	387	415	467	454	517	575
4.1	677	759	765	994	1193	1205	1339	1182	1351	1339	1546	1705
4.2	511	472	521	644	810	856	934	980	1063	1167	1332	1327
4.3	200	202	209	253	357	332	345	370	376	410	473	452
4.4	100	108	149	147	214	230	235	280	319	364	415	519
5.1	332	423	529	705	1196	1324	1597	1794	2186	2317	2713	2931
5.2	160	169	173	225	309	326	371	395	436	438	438	472

Anexo II. Distribución de keywords por clústeres para el caso mundial.

1. Composición de alimentos y nutrición	
1.1. Nutrición y metabolismo	nutrition (1237), obesity (948), rats (734), diet (708), calcium (571), inflammation (559), adsorption (506), iron (489), bioavailability (486), diabetes (470), security (464), zinc (440), toxicity (430), metabolism (429), health (417), selenium (387), liver (384), bangladesh (381), medicinal plants (374), environment (366), children (355), cadmium (330), pregnancy (324), sustainability (322), india (321)
1.2. Nutrientes y calidad	pig (1744), growth (1035), meat quality (950), cattle (773), digestibility (731), performance (622), sheep (621), beef cattle (560), gene expression (520), nitrogen (450), nutrients (404), phosphorus (397), goat (395), broiler (394), intake (348), growth performance (347), lamb (342), behavior (319), carcass (304), heritability (303), energy (303), stress (300)
1.3. Composición de los alimentos	fatty acids (2331), protein (1425), amino acids (1078), minerals (773), cholesterol (647), oxidation (619), lipids (585), conjugated linoleic acid (570), tocopherols (557), olive oil (552), fat (486), fatty acid composition (441), oxidative stability (407), fish oil (349), proximate composition (347), vitamins (330), oil (321), carbohydrate (308)
1.4. Composición y calidad	beef (1086), lipid oxidation (1003), meat (945), fish (752), pork (573), consumer (513), tenderness (417), chicken (409), aging (407), turkey (343), poultry (338), consumer behaviour (303)
2. Procesado y modificado de alimentos	
2.1. Influencia del procesado en las características sensoriales de los alimentos	rheology (1605), starch (1243), physicochemical properties (936), microstructure (900), functional properties (837), dietary fibre (812), stability (797), viscosity (740), gelatin (704), emulsion (689), chitosan (655), polysaccharide (606), whey protein (562), pectin (544), rheological properties (515), physical properties (491), heat treatment (476), extrusion (468), mechanical properties (451), structure (441), cyclodextrin (426), nmr (406), edible film (365), particle size (363), solubility (363), dsc (362), ultrafiltration (347), β -lactoglobulin (343), encapsulation (325), resistant starch (314), gelatinization (309)
2.2. Métodos de procesado o tratamiento de alimentos	modeling (1139), drying (821), processing (761), kinetics (696), potato (618), high pressure (584), mathematical model (551), maillard reaction (513), ultrasound (490), cooking (463), microwave (455), acrylamide (446), polyphenol oxidase (434), freezing (432), osmotic dehydration (381), pulsed electric field (330), image analysis (326), simulation (315), inhibition (309), browning (305)
2.3. Utilización de microorganismos beneficiosos como cultivos iniciadores en el procesado de alimentos	probiotics (1249), functional food (690), inulin (546), lactobacillus (546), yogurt (534), spray drying (473), prebiotics (405), microencapsulation (370), glucose (369), sucrose (326)
2.4. Modificaciones de los alimentos durante el procesado y cómo determinarlas	response surface methodology (1231), extraction (850), optimization (847), enzymes (574), purification (487), characterization (477), lipase (379), enzymatic hydrolysis (350), hydrolysis (348), peptide (303)

3. Seguridad alimentaria

3.1. Microorganismos patógenos en la leche y productos lácteos	milk (1642), dairy cow (1319), listeria monocytogenes (1157), salmonella (820), cheese (797), escherichia coli (571), antimicrobial (569), pcr (568), real-time pcr (508), staphylococcus aureus (491), detection (399), bacteria (393), escherichia coli o157:h7 (389), mastitis (379), bacteriocin (376), milk production (333), high hydrostatic pressure (329), dairy cattle (324), dairy products (322), dairy (313), biofilm (308), egg (308), foodborne pathogen (305)
3.2. Micotoxinas en cereales y métodos de detección de micotoxinas en alimentos	hplc (1712), wheat (1079), mycotoxins (847), rice (838), yield (822), soybean (695), bread (661), aflatoxins (655), ochratoxin a (639), maize (597), barley (575), elisa (530), cereals (514), germination (490), coffee (425), lc-ms/ms (406), isoflavones (401), validation (378), high-performance liquid chromatography (345), residues (341), corn (326), deoxynivalenol (310), allergy (301)
3.3. Otros contaminantes importantes de alimentos en seguridad alimentaria (metales pesados, pesticidas).	safety (2055), composition (1178), analysis (1032), fruits (913), vegetables (757), heavy metals (543), risk assessment (507), pesticides (469), contamination (316)
3.4. Agentes antimicrobianos utilizados en alimentos. Métodos de determinación de estos agentes antimicrobianos	essential oil (1303), gc-ms (1054), chemical composition (868), antimicrobial activity (804), antibacterial activity (393), additive (382)

4. Conservación de alimentos y vida útil

4.1. Incremento de vida útil en alimentos vegetales durante su conservación. Métodos y modificaciones de calidad	quality (2272), colour (1930), storage (1787), texture (1772), shelf life (1491), sensory evaluation (1016), packaging (720), apple (699), sensory (682), ripening (606), proteolysis (570), irradiation (530), sensory quality (480), modified atmosphere packaging (475), postharvest (471), firmness (445), strawberry (413), ethylene (374), sensory properties (306)
4.2. Proceso de fermentación en vino y cerveza como método de conservación	fermentation (1474), lactic acid bacteria (1344), wine (1260), volatile compounds (1022), sensory analysis (896), yeast (774), flavor (701), mass spectrometry (540), gas chromatography (523), aroma (514), biogenic amines (483), saccharomyces cerevisiae (482), ethanol (476), grape (466), volatiles (437), beer (424), identification (347), vitis vinifera (334)
4.3. Parámetros que condicionan la conservación de alimentos	temperature (987), ph (766), sugar (693), organic acids (536), water activity (491), salt (343), moisture content (329), water (304)
4.4. Métodos de análisis estadístico	pca (867), honey (556), artificial neural network (463), cultivars (405), chemometrics (393), multivariate analysis (314), cluster analysis (311)

5. Antioxidantes en alimentos	
5.1. Antioxidantes y sus efectos	antioxidant (4227), antioxidant activity (3974), polyphenols (1915), anthocyanins (1454), phenolics (1419), flavonoids (1327), phenolic compounds (1317), oxidative stress (1144), antioxidant capacity (912), apoptosis (816), dpph (575), lipid peroxidation (544), total phenolics (459), phenolic acids (397), cytotoxicity (387), reactive oxygen species (363), catechin (359), tea (348), tannin (335), bioactive compounds (333), phytochemicals (331), green tea (329), quercetin (327), antioxidant enzymes (320), anti-inflammatory (313), total phenolic content (312), red wine (311), free radical (303)
5.2. Antioxidantes vegetales	carotenoids (942), ascorbic acid (889), tomato (641), vitamin c (580), Î²-carotene (484), vitamin e (430), lycopene (374), chlorophyll (305)

Anexo III. Producción absoluta de los clústeres R1 y R2 del caso español.

Clúster R1	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
1. Composición de alimentos y nutrición	121	151	131	159	177	192	229	211	229	295	287	288
2. Procesado y modificado de alimentos	107	93	135	176	212	177	231	188	232	248	266	276
3. Seguridad alimentaria	86	79	101	126	150	192	212	211	258	304	271	282
4. Conservación de alimentos y vida útil	186	185	217	257	297	318	310	266	311	351	369	362
5. Antioxidantes en alimentos	45	53	75	98	124	149	179	162	220	230	289	296
Clúster R2	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
1.1	16	21	25	33	27	40	44	41	47	68	66	78
1.2	31	40	27	49	41	42	58	60	58	69	75	55
1.3	49	75	60	62	91	85	96	87	87	110	121	116
1.4	36	29	33	42	35	53	60	53	54	79	55	66
2.1	37	38	55	61	86	67	96	92	114	108	124	129
2.2	51	33	64	85	102	78	98	75	84	90	96	97
2.3	11	17	17	28	22	23	36	22	34	44	40	42
2.4	18	14	17	18	22	21	24	22	21	29	35	38
3.1	42	30	40	47	59	68	86	76	85	129	100	119
3.2	24	33	35	53	43	72	65	62	102	99	95	98
3.3	13	13	12	17	26	45	62	64	52	76	55	54
3.4	12	9	21	13	32	23	22	27	44	36	41	36
4.1	92	103	108	130	147	153	146	110	151	140	157	171
4.2	79	63	90	109	132	141	136	132	149	184	195	157
4.3	22	20	23	26	27	38	30	27	22	31	29	35
4.4	16	21	21	22	18	27	24	26	26	28	41	37
5.1	39	45	58	77	103	117	152	141	196	215	260	264
5.2	13	11	20	32	33	42	39	38	44	36	55	50

