

Estudio de la Colaboración Científica de la región de Valparaíso de Chile desde la Perspectiva de las Redes Complejas

J.P. Cárdenas, G. Olivares, R. Alfaro, F. Cabrera,
J. Gibert, D. Goya, H. Samaniego, A. Palacios

*Instituto de Sistemas Complejos de Valparaíso
Subida Artillería 470, Valparaíso, Chile*

Resumen

En este trabajo estudiamos la colaboración científica de la región de Valparaíso de Chile ocurrida durante las dos últimas décadas mediante técnicas basadas en grafos. Usando la base de datos de Web of Science[®] (WoS) de Thomson Reuters, construimos redes de colaboración en las cuales los nodos corresponden a los autores y las publicaciones a los enlaces que los relacionan. Nuestros resultados muestran que desde el año 2000 comienza un cambio drástico en la actividad de científica de la región. La llegada de nuevos autores, la colaboración con el extranjero y una mayor productividad son aspectos que han variado positivamente durante la última década. Sin embargo, las principales disciplinas responsables de este cambio no están vinculadas directamente al sector productivo de la región de Valparaíso. Nuestros análisis muestran también que la colaboración es un fenómeno que ha cambiado, si bien la cantidad de autores ha crecido, así como las publicaciones, los grafos de colaboración son cada vez menos densos y con un mayor número de grupos aislados de colaboración. Por otro lado, un gran porcentaje de los autores nuevos que llegan a Valparaíso colaboran con autores que también son nuevos, y además en grupos pequeños.

1. Introducción

Los sistemas complejos son aquellos que se caracterizan por presentar un rico comportamiento, difícil de describir, poco intuitivo y por lo general impredecible. Lo más importante es que dicho comportamiento no está descrito en los elementos que componen estos sistemas, sino que emerge como resultado de la intrincada relación entre estos [2].

Desde hace más de una década este tipo de sistemas está siendo estudiado

mediante un formalismo denominado Teoría de Redes Complejas [3,4]. Las llamadas redes complejas, nombre que se refiere a su organización poco intuitiva y para nada trivial, son abstracciones puramente topológicas de sistemas de este tipo en las cuales los nodos representan a sus componentes y los enlaces sus relaciones. La evidencia empírica tras el estudio de un sinnúmero de redes complejas reales sugiere que la complejidad del sistema, aquellas propiedades topológicas emergentes, aparecen de forma ubicua independiente de la naturaleza del sistema. La universalidad con la que este tipo de organización se manifiesta en sistemas biológicos, sociales, tecnológicos y otros, sería el resultado espontáneo de la evolución de estos sistemas mediante mecanismos físicos fundamentales que subyacen dicho proceso.

Actualmente estas ideas están cada vez más arraigadas en la comunidad científica, y muchas de las investigaciones se están haciendo desde esta perspectiva. De hecho, esto fue lo que nos motivó como grupo a estudiar la colaboración de los investigadores en la región de Valparaíso de Chile desde la perspectiva de la complejidad.

Existen muchos trabajos que han hecho algo parecido. De hecho hay investigaciones bastante antiguas como la realizada por Derek de Solla Price [5,6], que aportaba ya para la actual Teoría de Redes Complejas describiendo por primera vez, y sin saberlo, la complejidad topológica de este tipo sistemas. Otros trabajos actuales [7] sugieren que el surgimiento de nuevos paradigmas científicos, así como el desechar otros, está estrechamente relacionado con la topología de las redes sociales de investigadores.

Nuestro trabajo va en la misma línea pero tiene un objetivo más simple, aunque no por eso menos ambicioso, el cual es obtener una primera “fotografía” de la estructura de colaboración científica entre investigadores de la región de Valparaíso, los que representan cerca del 8% del total de investigadores nacionales [1] detrás de la región Metropolitana y del Biobío. Para esto transformamos la colaboración en una red de autores conectados por publicaciones a partir del procesamiento de la base de datos de la Web of Science[®] (WoS) (isiknowledge.com) de Thomson Reuters. En particular, trabajamos con una parte de dicha base de datos que contenía sólo publicaciones (artículos, proceeding papers, meeting abstracts, etc.) en que al menos un autor aparecía afiliado a alguna institución de Valparaíso entre el periodo 1991-2011. Utilizamos esta base de datos por ser un referente en la ciencia mundial para estudios bibliométricos y porque cubre las mejores revistas en todas las áreas temáticas, de hecho más de 10.000 revistas y 100.000 actas de congresos (aproximadamente 40 millones de registros). Además su orientación hacia las ciencias básicas favorecía nuestra investigación. Otras bases de datos, tales como Scopus y Google Scholar, presentaban ciertos problemas. La primera de ellas, competencia directa de WoS, no es tan selectiva y además en ella prima la cantidad de fuentes sobre la “calidad” de las mismas. En Google Scholar los

datos no tienen ninguna normalización, consecuencia de la amplia cobertura, la variedad de fuentes de información y el procesamiento automático de ésta, lo que dificultaba mucho la extracción y procesamiento.

Así, transformamos la base de datos WoS, construida en formato .bib, en un grafo $G(N, E)$ que describe la relación entre N autores distintos a través de E publicaciones. Este grafo corresponde a la representación unimodal del original bipartito compuesto por autores enlazados a publicaciones. Para colapsar el grafo bipartito a G decimos que *dos autores colaboran (están enlazados) si es que han participado en forma conjunta en una o más publicaciones en alguno de los años del periodo 1991-2011*.

La teoría aporta con una serie de algoritmos que permiten estudiar propiedades globales del grafo así como otras locales que describen propiedades particulares de los nodos en el contexto de G . Respecto a las propiedades globales destacan la llamada distribución de grados, aquella que nos dice cómo están distribuidos los nodos según conectividad. Esta es una medida clásica que ha permitido distinguir redes en la cuales los enlaces tienen un origen sin sentido de otras con enlaces preferentes. Las llamadas redes libres de escala, caracterizadas por presentar una distribución de grado sumamente heterogénea, totalmente distinta a la observada en modelos de redes con conexiones aleatorias, caracterizan a los sistemas complejos. Otras medidas globales importantes son la distancia y el clustering medios. La primera hace referencia a la distancia que separa en promedio a cualquier par de nodos del grafo, mientras que la segunda indica la proporción media de vecinos de un nodo cualquiera que son vecinos entre sí.

Dentro de las propiedades locales de los nodos destaca obviamente la cantidad de enlaces, el llamado grado k , que intuitivamente nos da la idea de importancia del nodo en G . Sin embargo, existen otras medidas como la centralidad y el betweenness que nos indican qué tan bien situado está el nodo en la red.

Aplicamos estas y otras medidas en cada uno de los grafos de colaboración extraídos desde la base de datos. Un ejemplo de grafo G es el que se muestra en la Figura 1. Esta representación corresponde a un subgrafo de publicaciones del año 2011. El tamaño de los nodos tiene que ver con la posición estratégica de los autores en la red de colaboración, la forma de los nodos con el país de afiliación de los autores (triángulo=Chile, círculo=otro), mientras que el tipo de enlace refleja la disciplina científica del artículo.

A partir de grafos de este tipo nuestro trabajo buscó obtener, en primer lugar, una caracterización de la investigación en la región que nos permitiera luego generar hipótesis que explicaran su estructura.

El trabajo se presenta de la siguiente manera. En la siguiente sección explicamos la metodología utilizada para el procesamiento de la base de datos,

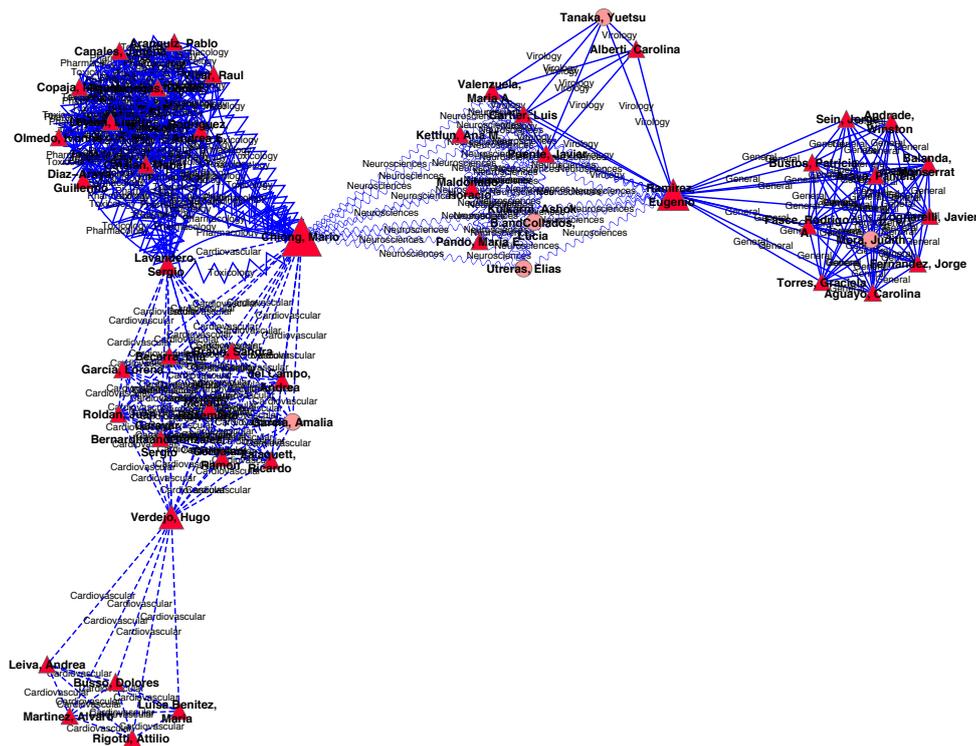


Figura 1. Parte de la red de colaboración para publicaciones asociadas a la región de Valparaíso durante 2011. Nodos caracterizados por posición estratégica y afiliación de autores. Enlaces diferenciados por disciplina científica.

especialmente por los problemas de formato de ésta. En la sección 3 se presentan los resultados más importantes de la investigación y en la última sección las conclusiones más relevantes de nuestro trabajo.

2. Metodología de extracción y procesamiento de la base de datos

El procesamiento de la base de datos para la construcción de las redes presentó problemas en algunos casos, especialmente a la hora de extraer la afiliación de los autores en aquellas publicaciones anteriores a 2008. Dentro de los problemas más importantes que encontramos podemos mencionar:

1. autores firmando sus publicaciones con diversos formatos de nombre,
2. ausencia en la base de datos de un formato claro que indique específicamente a qué afiliación pertenece cada autor,
3. en la gran mayoría de los casos sólo se indica la afiliación del primer autor,

4. presencia de publicaciones con un gran número de autores¹,
5. nombre de instituciones con distinto formato.

Debido a los problemas anteriores, fue necesario realizar ciertos supuestos o simplificaciones para la generación de las redes y extracción de atributos de los nodos y enlaces. Estas simplificaciones se detallan a continuación:

1. Para no considerar más de una vez a un autor con distinta firma, es decir representar a una misma persona como nodos distintos, se cambió la firma original por la primera inicial de su o sus nombres y su apellido completo (Juan Pablo Cárdenas → JP Cárdenas).
2. En las publicaciones en donde no era clara la afiliación de un autor, se rellenó el campo como no afiliado.
3. Toda publicación con más de 20 autores fue ignorada.
4. El nombre de las universidades estatales y particulares de la región fue uniformado manualmente para un análisis más preciso.
5. Se consideró sólo la primera afiliación por autor.

Teniendo claras las medidas adoptadas para el tratamiento de la base de datos, a continuación presentamos los resultados más importantes. Algunos de ellos están relacionados con la topología de las redes, mientras que otros son de naturaleza meramente descriptiva de la colaboración científica de la región de Valparaíso de Chile.

3. Resultados

En primer lugar estudiamos la participación de las cuatro principales universidades tradicionales de la región²: Pontificia Universidad Católica de Valparaíso (PUCV), Universidad Técnica Federico Santa María (UTFSM), Universidad de Valparaíso (UV) y Universidad de Playa Ancha de Ciencias de la Educación (UPLA) respecto al número de afiliados para todas las disciplinas científicas de la base de datos extraída. La Figura 2 muestra esto considerando las dos últimas décadas. La participación de cada universidad está medida como número de autores afiliados, como primera opción, a cada una de ellas. La forma de leer este mapa de disciplinas y autores es la siguiente. Para cada disciplina existe una franja vertical en donde aparecen las universidades que tienen publicaciones en ella, su “altura” refleja el número de autores afiliados a la universidad.

¹ Esto se daba especialmente en artículos relacionados con la Astronomía donde equipos de trabajo de gran tamaño eran considerados por completo en el campo de autores de la base de datos.

² Otras universidades no fueron consideradas por la baja cantidad de publicaciones en el periodo total de estudio.



Figura 2. Participación de cuatro universidades de la región en todas las disciplinas científicas de los artículos publicados en las dos últimas décadas. El tamaño de los puntos tiene que ver con el número de autores afiliados.

Es posible ver que la disciplina de la Ingeniería es la que tiene más autores y estos pertenecen a la UTFSM. La misma disciplina tiene también autores en la PUCV y muy por debajo en la UV, posiblemente porque su Facultad de Ingeniería nació recién el año 2011 como tal, antiguamente pertenecía a la Facultad de Ciencias. Por otro lado, disciplinas como la Química, y la Biotecnología y Microbiología Aplicada, tienen a la PUCV dominando en términos de investigadores. Es interesante notar el carácter pluridisciplinario de la PUCV.

El inset de la figura muestra aquellas categorías en donde hay más de 60 afiliados en las dos últimas décadas. Se aprecia que las universidades PUCV y UTFSM dominan el escenario, sin embargo, la UV tiene casi exclusividad en categorías como la Medicina General e Interna y en Neurociencia y Neurología.

La UPLA tiene una participación muy baja en las dos últimas décadas en comparación a las otras tres universidades. Esto podría deberse a que ésta tiene unas pocas carreras de ingeniería (Industrial, Ambiental, Informática y Estadística) y otras pocas en el área de la salud (Kinesiología, Terapia Ocupacional y otras), el grueso de las carreras de la UPLA son humanistas cuyas publicaciones no están bien representadas en la base de datos de Thomson Reuters.

Considerando que existen disciplinas en las cuales algunas universidades dominan (Ingeniería en la UTFSM y Medicina en la UV por ejemplo), restringimos nuestro análisis sólo a cuatro y que denominamos Ciencias Básicas: Matemática, Física, Química y Biología, en un intento por capturar la participación de las cuatro principales universidades tradicionales de la región. Además estas disciplinas concentran, según [1], el mayor número de investigadores en la región y son, en conjunto, áreas prioritarias de investigación en tres de las cuatro universidades tradicionales de la región.

En la gráfica superior de la Figura 3 se observa que a partir del año 2004 comienza una tendencia al aumento en el número de autores para las Ciencias Básicas vinculados a la región de Valparaíso. Este aumento se traduce en una curva productiva (publicaciones, gráfica inferior Fig. 3) muy similar en su forma, destacando el año 2009 en ambos casos. Lo mismo sucede con la curva de publicaciones totales que incluye todas las categorías científicas de la base de datos. La diferencia entre ambas curvas de publicaciones (negra y gris) son el resto de disciplinas no consideradas en este gráfico. Respecto a esto se puede ver que al pasar el tiempo, el aporte de otras disciplinas científicas al total de publicaciones es mayor.

En la misma figura se representa la dinámica del número de autores y publicaciones para disciplinas como la Ingeniería y otras relacionadas con la investigación marina. El caso de las “Ciencias del Mar”, debido a su poca visibilidad, y en la Ingeniería, por el dominio casi exclusivo de la UTFSM, tal

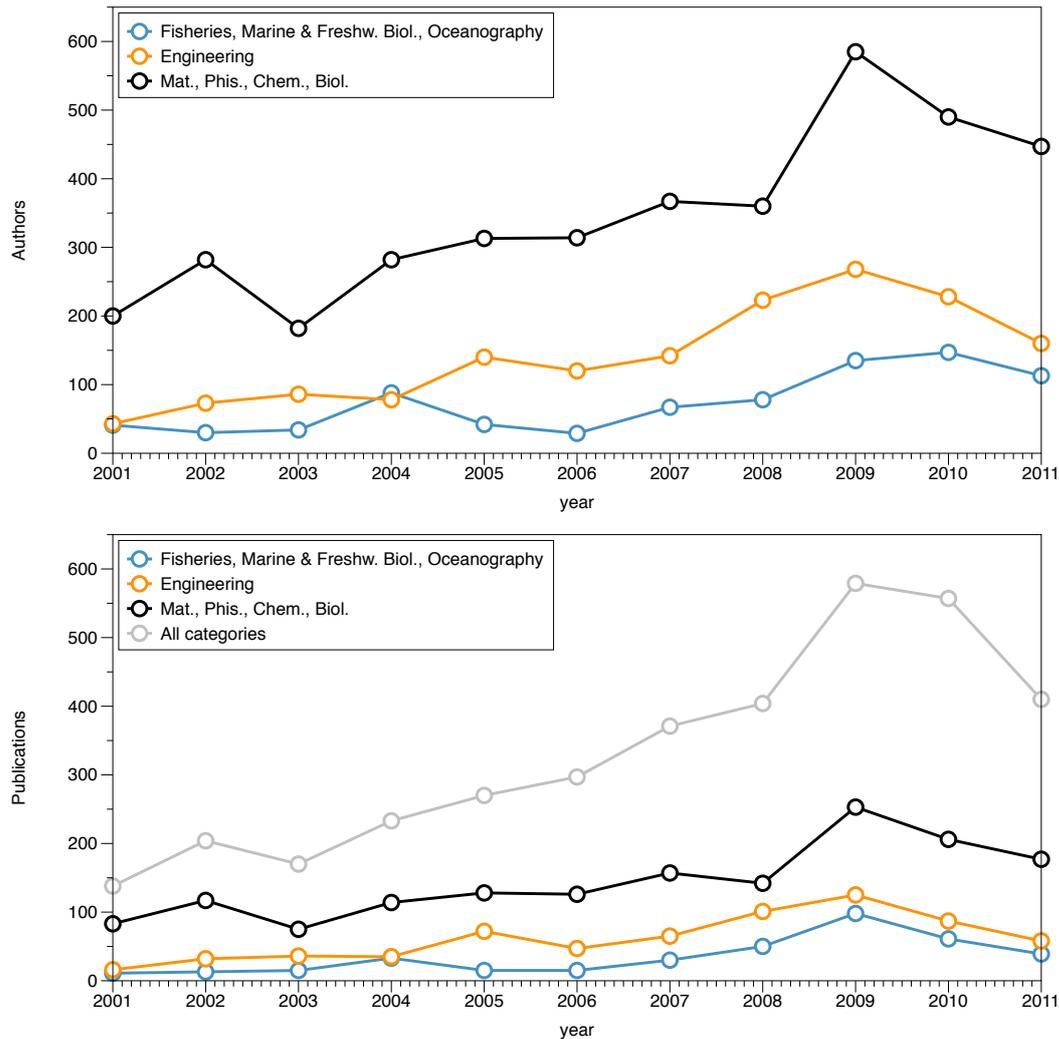


Figura 3. Variación en el número de autores (arriba) y publicaciones (abajo) durante la última década para las disciplinas de las Ciencias Básicas (negro), la Ingeniería (naranja) y Ciencias del Mar (azul).

como se mencionó antes, no han sido incorporadas al análisis que se presenta a continuación.

En términos de la colaboración (enlaces de nuestras redes) se observa un aumento en el número de componentes conectados³ a medida que pasan los años (Fig. 4). Esto indicaría que a medida que el número de autores crece aparecen más grupos aislados de investigación, tal como se ve en la gráfica superior de la Figura 5.

También podemos estudiar la colaboración utilizando otras métricas clásicas

³ Conjunto de autores pertenecientes al grafo de colaboración que están conectados entre sí pero que no están conectados con el resto. El tamaño mínimo de un componente es de 2 nodos.

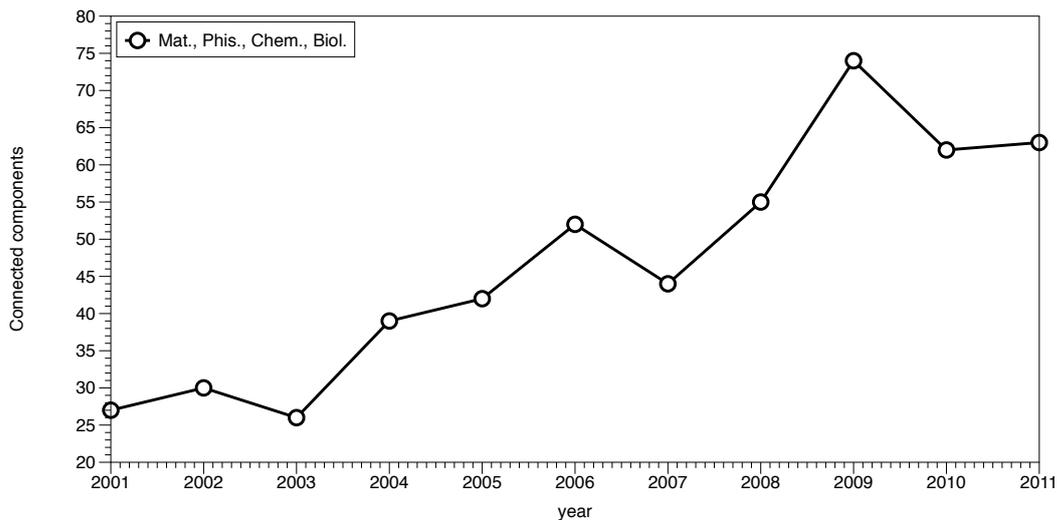


Figura 4. Número de componentes conectados anuales durante la última década para publicaciones en Ciencias Básicas vinculadas a la región de Valparaíso.

para evaluar grafos. Así, definimos la vecindad de un autor como el conjunto de todos sus vecinos. El número medio de vecinos indica la conectividad media de los autores en la red. Una versión normalizada de este parámetro es lo que llamamos densidad y puede tomar un valor entre 0 y 1 indicando qué tan densamente poblada está la red de colaboraciones. Por ejemplo, una red que no tiene colaboración tiene una densidad igual a 0. Por el contrario, la densidad de un clique (conjunto de autores en que todos están colaborando entre sí) tendría un valor de 1.

La gráfica inferior de la Figura 5 muestra la densidad de colaboración durante la última década. Claramente se observa que ésta cae durante este periodo. Si bien el número de autores aumenta y obviamente, por la forma en que construimos las redes, el número de enlaces también lo hace, la población de enlaces disminuye notoriamente, indicando una disminución en la colaboración.

Sin embargo, un aspecto muy interesante tiene que ver con el análisis de la variación en el tamaño de la Componente Gigante (CG), aquel conjunto mayor de autores del grafo de colaboración no conectado al resto pero en el cual es posible conectar a un autor, escogido al azar, desde cualquier otro. Esta definición es clave porque nos dice que dos autores que pertenezcan a la componente, a pesar de no haber publicado juntos (si es que no hay un enlace directo entre ellos), están siempre conectados por otros autores, en promedio cerca de 6 según nuestros análisis. Es decir, que pertenecen a un mismo sistema social cerrado.

Realizamos un estudio del tamaño acumulado de la componente desde 1991, año a año, hasta 2011, también para las Ciencias Básicas. De esta manera,

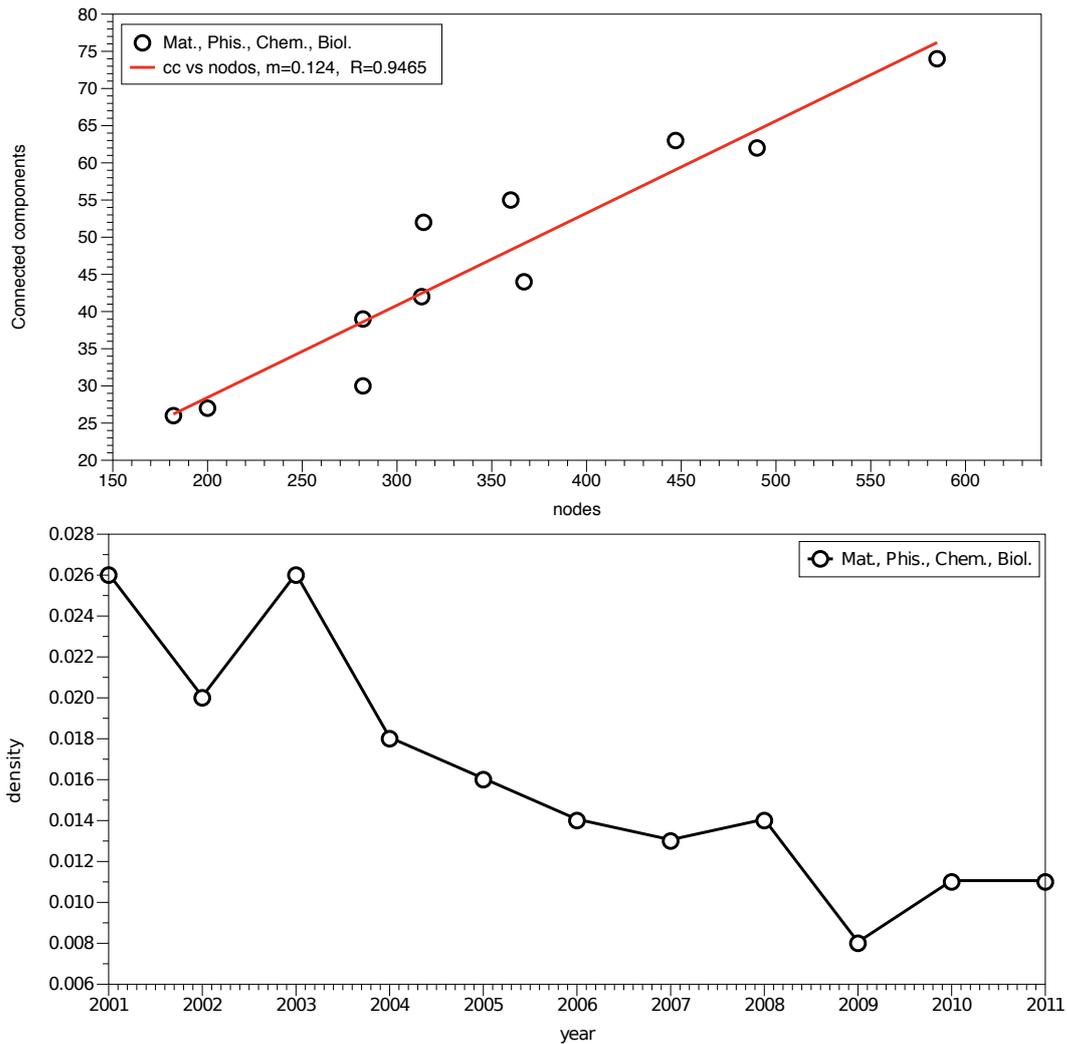


Figura 5. Arriba: Número de componentes conectadas respecto al número de nodos. **Abajo:** Densidad de colaboración de la red.

estudiamos su tamaño en el periodo 1991-1992, 1991-1993 y así sucesivamente hasta el periodo total de las dos décadas 1991-2011. Este tipo de análisis nos permitió observar los cambios de la comunidad social de investigadores con el paso del tiempo.

La Figura 6 muestra la dinámica del tamaño de la componente gigante en términos absolutos y porcentuales. Es posible observar que durante la primera década el tamaño de la componente es pequeño, no superando los 200 autores. Durante este periodo su tamaño representa entre el 35 % y el 20 % del total de los nodos, con una clara tendencia a la disminución. Solo los periodos 1991-1995 y 1991-1996 presentan valores porcentuales más altos debido a la menor cantidad de autores en esos momentos (datos no mostrados). Es interesante destacar que en el periodo 1995-2001 no hay variación en el tamaño de la componente. Las razones pueden ser dos; no hay nuevos autores que entren a

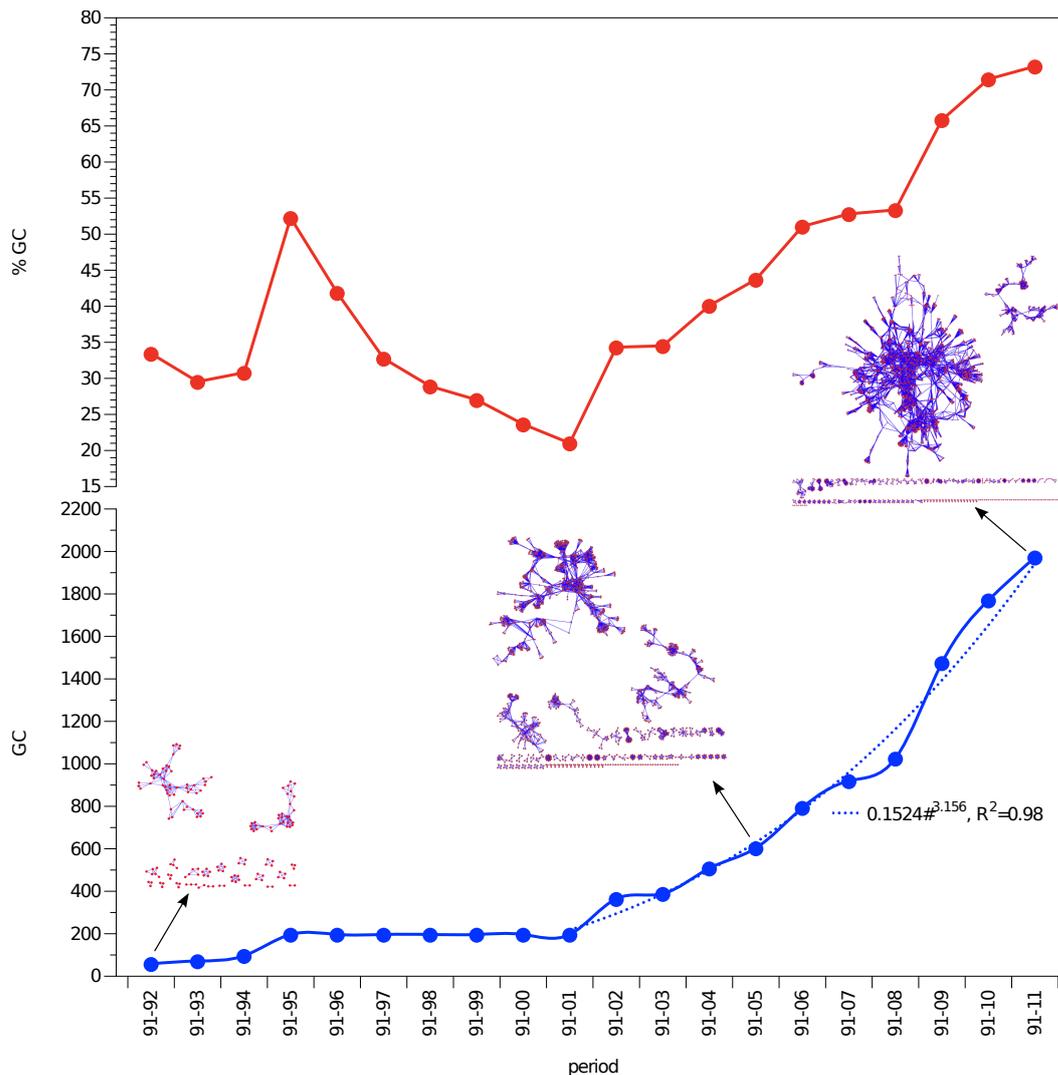


Figura 6. Variación en el tamaño de la componente gigante de colaboración en las dos últimas décadas para artículos de Ciencias Básicas vinculados a la región de Valparaíso. Línea azul representa el tamaño de la componente en cada uno de los años. Línea roja indica el porcentaje que representa la componente gigante respecto al total de autores de la red en cada uno de los años.

formar parte de este gran sistema, o bien, que año a año se renueva totalmente esta componente, lo que parece mucho más difícil. Lo que sí está claro es que durante este periodo de casi una década no hay un aumento en la complejidad de la red, entendiendo ésta como un enriquecimiento de la red mayor de colaboración en cuanto a autores y enlaces.

Lo interesante es que a partir del año 2002 comienza, abruptamente, un crecimiento de la componente gigante de colaboración. Este resultado sugiere que podríamos estar en presencia de una transición de fase con el año 2002 como punto crítico. Lo que esto nos indica es que a partir de este año comienza a

formarse una gran comunidad de colaboración en Ciencias Básicas en donde es muy posible conectar a cualquier par de autores del grafo. Esta comunidad llega a agrupar el año 2011 cerca del 75 % del total de autores que han publicado en las dos últimas décadas artículos vinculados a la región de Valparaíso. Los efectos de un cambio tan drástico en la topología del grafo de colaboración son evidentes dejando claro que esta última década ha sido un periodo fructífero para la investigación en la región. Algo muy interesante de este crecimiento explosivo es que corresponde a un fenómeno de escalado, muy típico de sistemas complejos. La figura muestra el ajuste a una ley de potencias de este crecimiento con $\gamma=3.1$ como exponente de escalado.

Por último, la Figura 7 muestra el aporte anual de cada una de las cuatro disciplinas de las Ciencias Básicas, durante las dos últimas décadas. Se puede ver que hasta el año 2001 había una repartición más bien pareja de publicaciones para estas cuatro disciplinas, sin embargo, con el paso del tiempo esta distribución fue haciéndose más sesgada en la región tomando un papel protagónico la Física y perdiendo protagonismo la Química. Al respecto, según el INE los principales sectores productivos de la región, durante 2008 están relacionados directamente con la Química.

Además, durante esta última década lo común es que todas las disciplinas aumenten año a año su número de publicaciones, no así en la década pasada donde no había mayor variación.

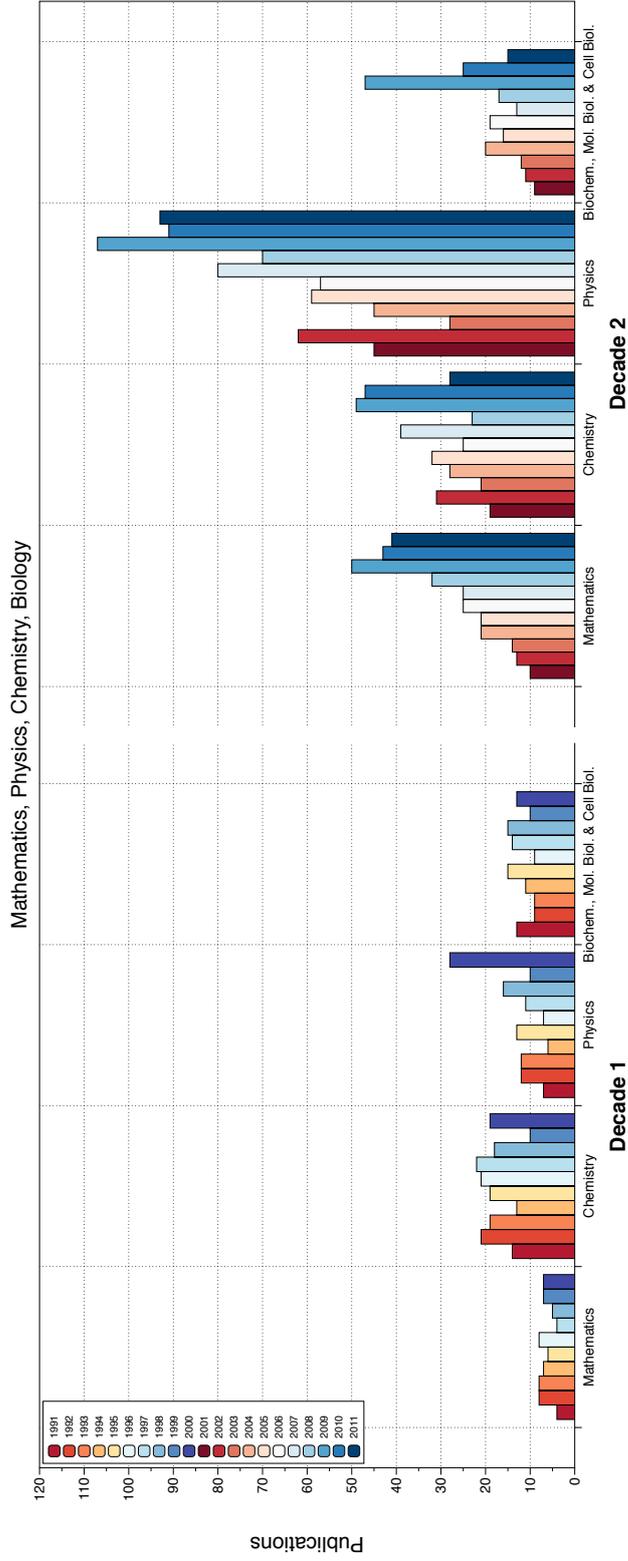


Figura 7. Número de publicaciones por cada disciplina científica de las Ciencias Básicas desde 1991 a 2011. Cada barra representa el número de publicaciones en un año determinado. Cada conjunto de barras corresponde a una disciplina. Notar que el gráfico ha sido partido en dos: Década 1 y Década 2.

3.1. Autores Colaboradores

Dejando de lado la topología de la red nos centramos en el análisis de los autores. Una forma de caracterizarlos tiene que ver también con la colaboración.

Así, definimos “autores colaboradores” como aquellos que han publicado al menos dos artículos con autores distintos (al menos uno de los autores que pertenece a un artículo no está en el otro). A partir de esta definición, que no es otra cosa que un valor de betweenness distinto de 0 para el autor i , $B_i > 0$, en la red, generamos un ranking de colaboradores. La manera de posicionar a los autores en este ranking fue considerando su valor de betweenness y el número de conexiones de colaboración que tiene, su grado k en la red.

El betweenness mide el número de caminos que pasan a través de un nodo. Dicho de otra forma, el número de veces que un autor cualquiera requiere de otro para alcanzar un autor a través del camino más corto. Más precisamente, supongamos que $g_i^{(st)}$ es el número de pasos desde el autor s hasta t que pasan a través de i , además supongamos que n_{st} es el número total de pasos desde s a t en la red. De esta forma, según [8], el betweenness b del autor i se puede definir como,

$$b_i = \frac{\sum_{s < t} g_i^{(st)} / n_{st}}{(1/2)n(n-1)} \quad (1)$$

donde n es el número total de autores en la red.

De esta forma, si un autor tiene muchas conexiones de colaboración, un alto grado k , pero no conecta a grupos distintos de colaboración, $B_i = 0$, no entra en este ranking. Con la siguiente medida de colaboración para un autor i ,

$$C_i = B_i k_i, \quad (2)$$

castigamos el número de enlaces de colaboración si estos no representan conexiones entre distintos grupos. Por otro lado, potenciamos el carácter colaborativo del autor mientras más grupos conecte y enlaces de colaboración tenga.

La Figura 8 muestra la dinámica del número de colaboradores durante las dos últimas décadas así como la afiliación de estos a las cuatro universidades tradicionales de la región. No incluimos en este trabajo el ranking de colaboradores.

Claramente se puede observar que la cantidad de colaboradores comienza a

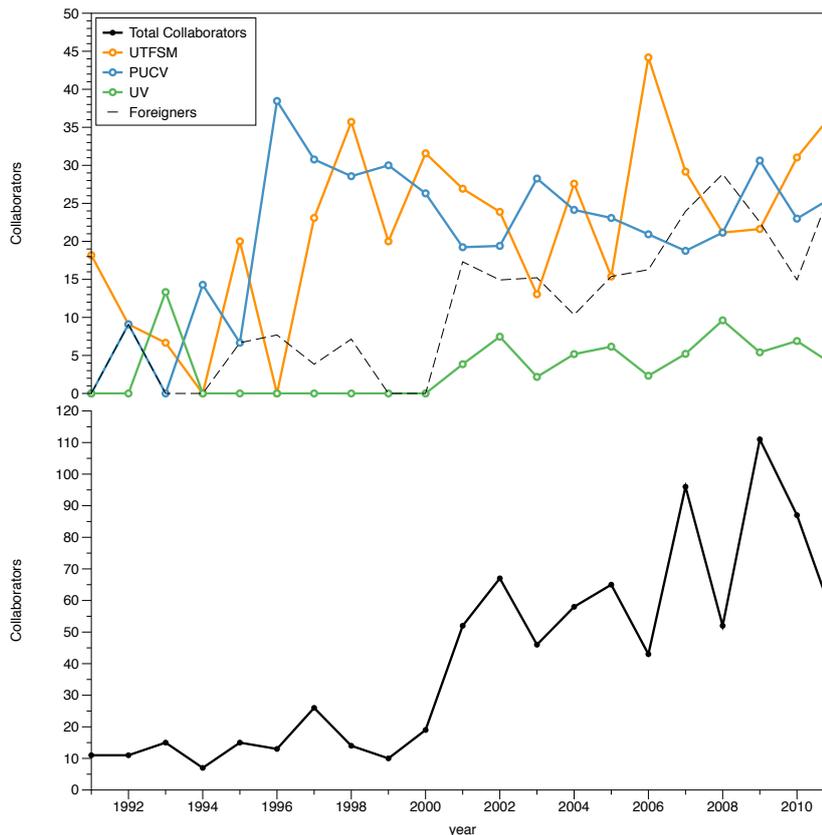


Figura 8. Dinámica de los “colaboradores” en las dos últimas décadas para artículos de Ciencias Básicas vinculados a la región de Valparaíso.

aumentar a partir de la segunda década (gráfico inferior). Algo que apoya lo planteado anteriormente respecto a que desde 2000 hubo un cambio radical en el escenario científico de la región. Sin embargo, es necesario decir que las variaciones en el número de colaboradores son bastante erráticas durante la última década.

Al momento de ver a qué universidad pertenecen estos colaboradores, observamos que principalmente son de la PUCV y UTFSM (gráfico superior). Ambas instituciones logran en 1996 y 1997, respectivamente, alcanzar un número de colaboradores que se ha mantenido más o menos constante hasta 2011. La UV presenta un ligero aumento en el aporte de colaboradores a partir de 2000 aunque no es significativo. Esto se debería a que no son estas las disciplinas científicas en las que domina la UV, tal como se vio en la Figura 2. Es necesario destacar que estas son las tres únicas universidades de la región que aportan colaboradores.

El caso de los extranjeros también es interesante. Precisamente, a partir del año 2000, comienzan a ser más los extranjeros “colaboradores” en artículos vinculados a la región de Valparaíso.

3.2. Nuevos autores

Otro aspecto de estas redes de colaboración tiene que ver con el renuevo de investigadores. En la Figura 9 se presenta la dinámica de incorporación de nuevos autores en términos absolutos y también como aporte porcentual. Se hace la distinción entre nuevos autores sin considerar afiliación (totales), aquellos de Chile y aquellos afiliados a alguna institución de Valparaíso.

La figura muestra una tendencia al aumento en el número de autores nuevos totales. Este tendencia se manifiesta sólo desde el año 2000, lo que sugiere nuevamente un cambio en el panorama científico a partir de ese año. El mismo fenómeno, más marcado hacia finales de la segunda década, se observa en nuevos autores chilenos y de la región de Valparaíso. En términos porcentuales se aprecia un aumento progresivo (pero lento) desde 1995 de nuevos autores chilenos; en el caso de nuevos autores de la región de Valparaíso se mantiene más o menos constante cerca del 20 %.

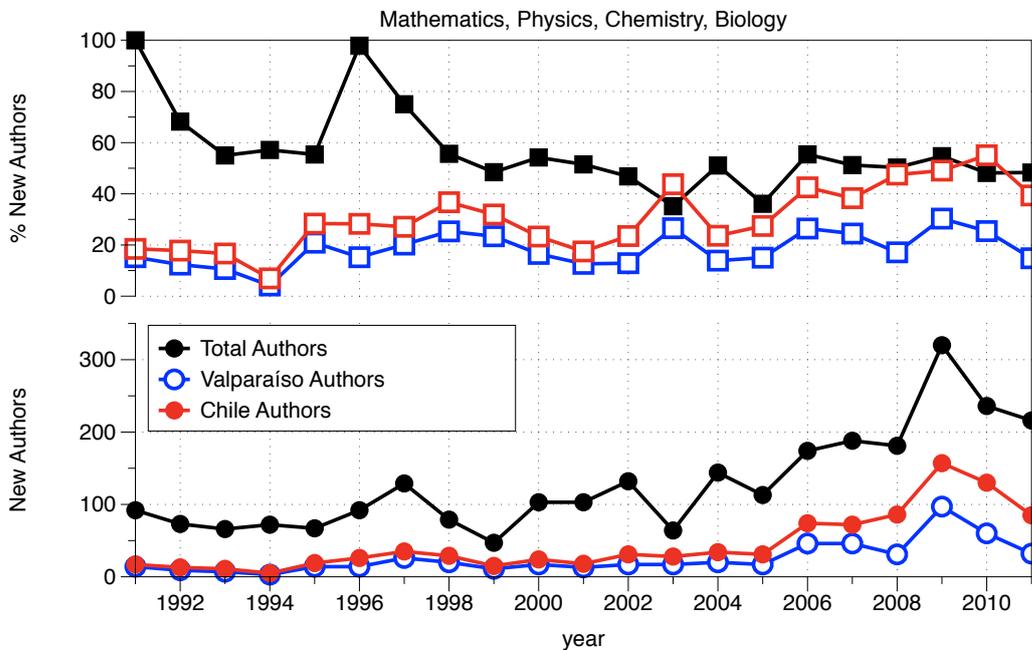


Figura 9. Incorporación de autores nuevos en cada año durante los últimos dos décadas. Se presentan las incorporaciones como números brutos y también como porcentajes: autores nuevos totales con respecto a los autores totales y chilenos, y de la región de Valparaíso como porcentaje de los autores nuevos.

La Figura 10 muestra la red de colaboración del año 2008 en la cual están distinguidos con color rojo todos aquellos autores nuevos.

Tal como se aprecia, la cantidad de nuevos autores vinculados directa o indirectamente a la región en estos últimos años es bastante significativa. Sin

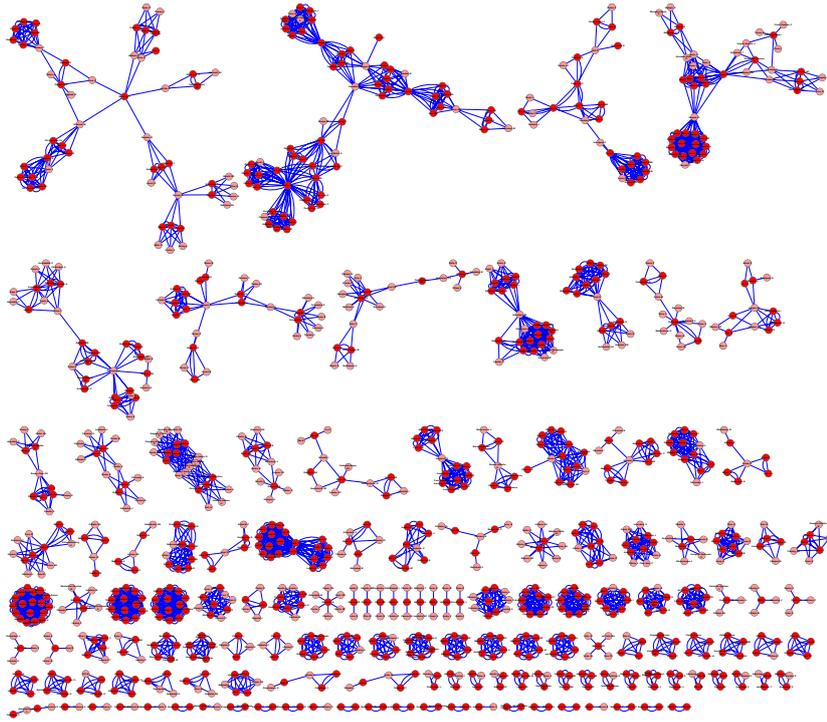


Figura 10. Red de colaboración del año 2008. En color rojo autores nuevos que publicaron un artículo vinculado a Valparaíso.

embargo, centramos nuestra atención en un subconjunto de esos autores nuevos, aquellos afiliados a alguna institución de la región de Valparaíso. La Figura 11 muestra la misma red de colaboración anterior en donde aparecen diferenciados en color rojo sólo los autores nuevos afiliados a la región. La cantidad de nuevos autores con esta restricción disminuye dramáticamente, lo que indica que un gran porcentaje de autores nuevos en cada año corresponde a investigadores extranjeros o de otras regiones del país. . . un signo positivo si hablamos de colaboración pero no tan bueno si pensamos en la investigación regional.

La dinámica de incorporación de nuevos autores afiliados a alguna institución de la región de Valparaíso, desde 2008 a 2011, se presenta en la Figura 12. La razón para considerar estos últimos cuatro años fue analizar la realidad actual. Es necesario remarcar que para este análisis consideramos todas las disciplinas científicas de la base de datos.

En la Figura 12 se puede apreciar que entre el año 2008 y 2009 hay un significativo aumento de nuevos autores afiliados a la región. Esta tendencia no se mantiene en el tiempo, incluso decae fuertemente en los dos últimos años, aunque es necesario decir que el muestreo del año 2011 no es completo (la base de datos contempla publicaciones hasta agosto de 2011).

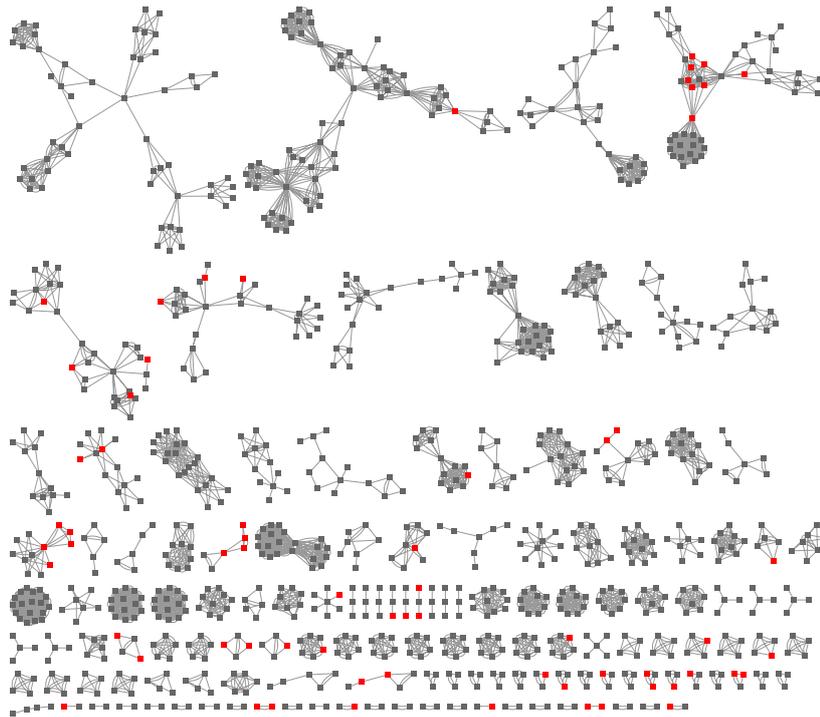


Figura 11. Red de autores que publicaron artículos vinculados a Valparaíso el año 2008. En color rojo autores nuevos afiliados a la V región.

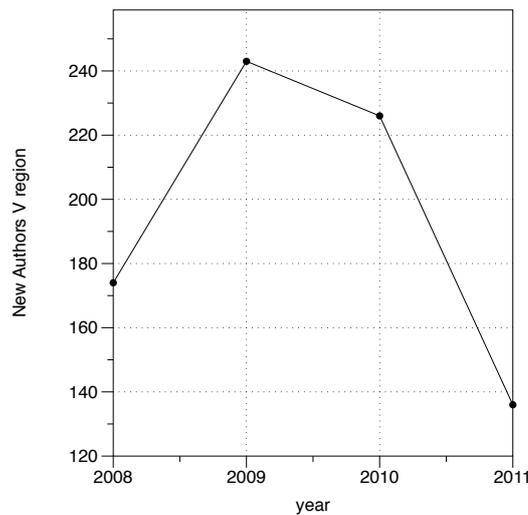


Figura 12. Dinámica de nuevos autores afiliados a la región de Valparaíso en los últimos cuatro años.

3.2.1. Sobre la afiliación de los nuevos autores

Todos estos nuevos autores afiliados a la región aparecen formalmente vinculados a una institución. Sin embargo, debemos señalar que la base de datos no nos permite saber si estos nuevos autores son estudiantes de paso o investigadores formalmente afiliados a alguna institución de la región. La distribución

de autores nuevos por institución en los años 2008, 2009 y 2010 se muestra en Figura 13.

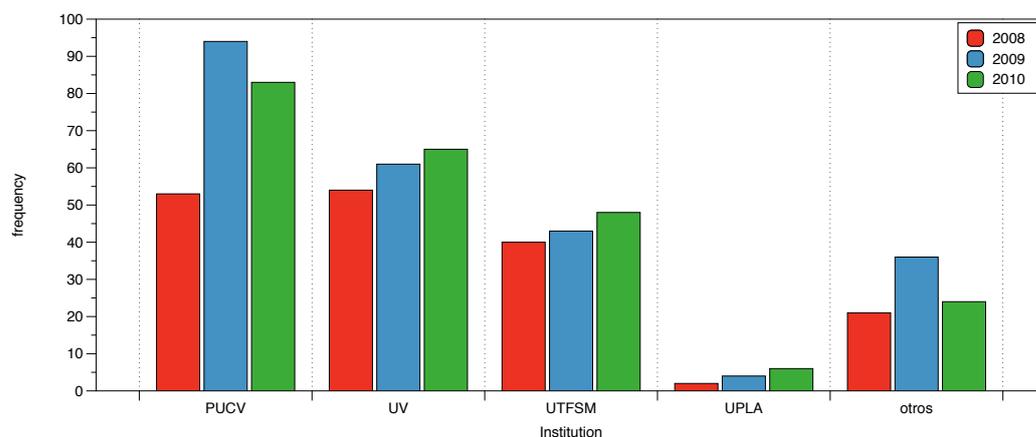


Figura 13. Universidades y otras instituciones como polo de atracción para nuevos investigadores.

Se puede apreciar que durante los años 2009 y 2010 la PUCV fue la institución que funcionó como mayor polo de atracción para nuevos investigadores, seguida por la UV, luego por la UTFSM y finalmente por la UPLA. Otras instituciones, principalmente Hospitales y el Instituto de Fomento Pesquero (IFOP), capturan también nuevos autores. La atracción de la PUCV no era tal en el año 2008 en donde comparte protagonismo con la UV.

3.2.2. Sobre la productividad de los nuevos autores

Un aspecto interesante de este análisis tiene que ver con el seguimiento de la carrera de investigador de estos nuevos autores afiliados a la región. Consideramos como “investigadores productivos” aquellos que publicaron al menos un trabajo en los años posteriores, después de incorporarse. Sabemos que la carrera de investigador es más que las publicaciones (congresos, proyectos, otros), pero es un buen punto de partida.

La Figura 14 muestra la distribución de la productividad de estos nuevos autores. Es interesante notar que la productividad=0 es la que domina en cada año, muy por sobre el resto. En especial en el año 2010. Es decir, que de los cerca de 230 nuevos autores, más de 200 no publicaron nada el 2011. Podemos ver que el año 2008 es un poco más productivo, lo más posible porque han tenido más tiempo. De los 174 nuevos autores de 2008 cerca de un 40 % publicó algún trabajo en años posteriores. El año 2009 es interesante. Ya hemos visto que es un año sumamente productivo en el cual además muchos nuevos autores se vincularon a la región, sin embargo cerca de un 30 % de estos nuevos autores no publicaron nada en los dos años siguientes. Para todos los años, de los pocos autores que vuelven a publicar, la gran mayoría lo hacen

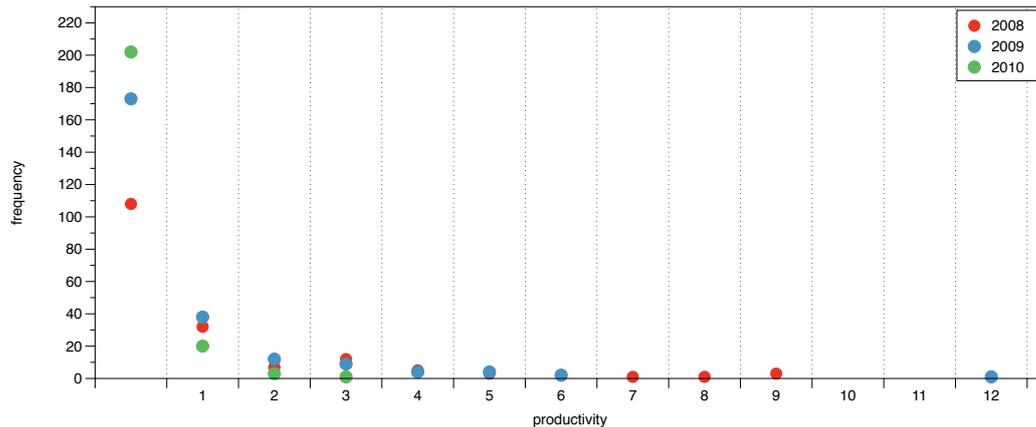


Figura 14. Distribución de productividad (artículos) en años posteriores para los nuevos autores afiliados a la V región.

sólo una vez. Estos resultados ponen de manifiesto un fenómeno indeseado. De hecho, la región de Valparaíso está bajo la media nacional respecto a la cantidad de personal académico con grado de Doctor que realiza investigación (Valparaíso=37 %, Nacional=50 %) [1].

El año 2011 no se consideró porque la base de datos está incompleta para ese año.

3.2.3. Sobre los que acogen a los nuevos autores

También analizamos los atributos de los autores que acogen a los nuevos. Para esto estudiamos propiedades de conjunto del grupo de autores con los que llega a trabajar (el nuevo autor) en ese año de incorporación a la región.

El motivo para analizar el grupo de acogida es ver si la génesis de los enlaces que conectan un nuevo autor con antiguos responde a algún enlace preferencial por parte del nuevo.

En primer lugar analizamos el tamaño de este grupo que “acoge” al nuevo autor. Tal como se mencionó, dicho grupo es el conjunto de autores que aparece trabajando con el nuevo autor en ese primer año de incorporación, pudiendo ser autores antiguos y nuevos. Es necesario decir que este grupo puede corresponder en realidad a distintos grupos de trabajo (por ejemplo si el nuevo autor publica más de un artículo con distintas personas); en nuestro caso no hacemos tal distinción y todos forman parte del mismo grupo de acogida.

La Figura 15 muestra la distribución de tamaños de los grupos que acogen a los nuevos autores en cada uno de los últimos cuatro años.

Se ve que para los años 2009, 2010 y 2011 el tamaño típico del grupo de

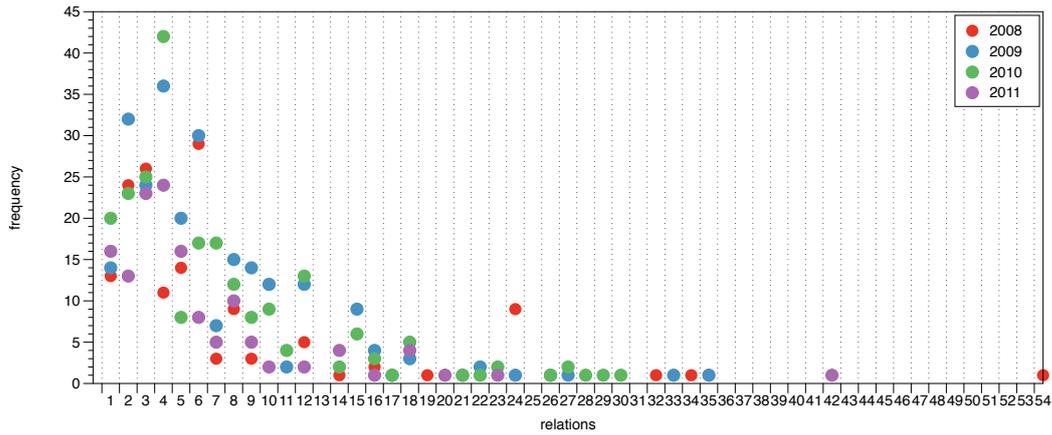


Figura 15. Distribución de tamaño de los grupos de acogida para los nuevos autores afiliados a la región de Valparaíso.

acogida es de 4 personas. Para el año 2008 fue 6. Sin embargo, en la gran mayoría de los casos, los autores nuevos llegan a trabajar con grupos de entre 3 y 6 personas, siendo menos los que llegan a grupos pequeños o a grupos muy grandes, aunque también existen.

Por otro lado, decimos que un grupo tiene experiencia cuando tiene entre sus miembros al menos a una persona “no nueva”. Según esto calificamos a un grupo de acogida como “inexperto” si todos sus miembros son nuevos. La Figura 16 muestra el número de grupos de acogida inexpertos para los últimos cuatro años. Se puede ver que el año 2008 cerca del 40 % de los grupos de acogida estaban formados sólo por personas nuevas. Este porcentaje cae los años siguientes pero en 2011 vuelve a aumentar.

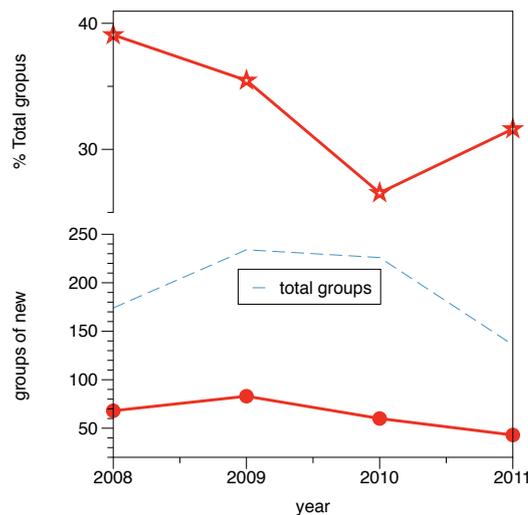


Figura 16. Número de grupos de acogida “inexpertos” (abajo) y su porcentaje respecto al total (arriba).

3.2.4. Sobre la calidad de los que acogen a los nuevos autores

Otro aspecto interesante de analizar es la “calidad” del grupo de acogida. Nuevamente tomamos ciertos atributos (que sabemos que no son suficientes) para catalogar a los grupos de acogida como calificados o no. Decimos que un grupo de acogida es de calidad de acuerdo a dos dimensiones: su productividad posterior y las conexiones de colaboración históricas de sus miembros.

Para esto presentamos las distribuciones de ambas dimensiones como histogramas de sus valores medios.

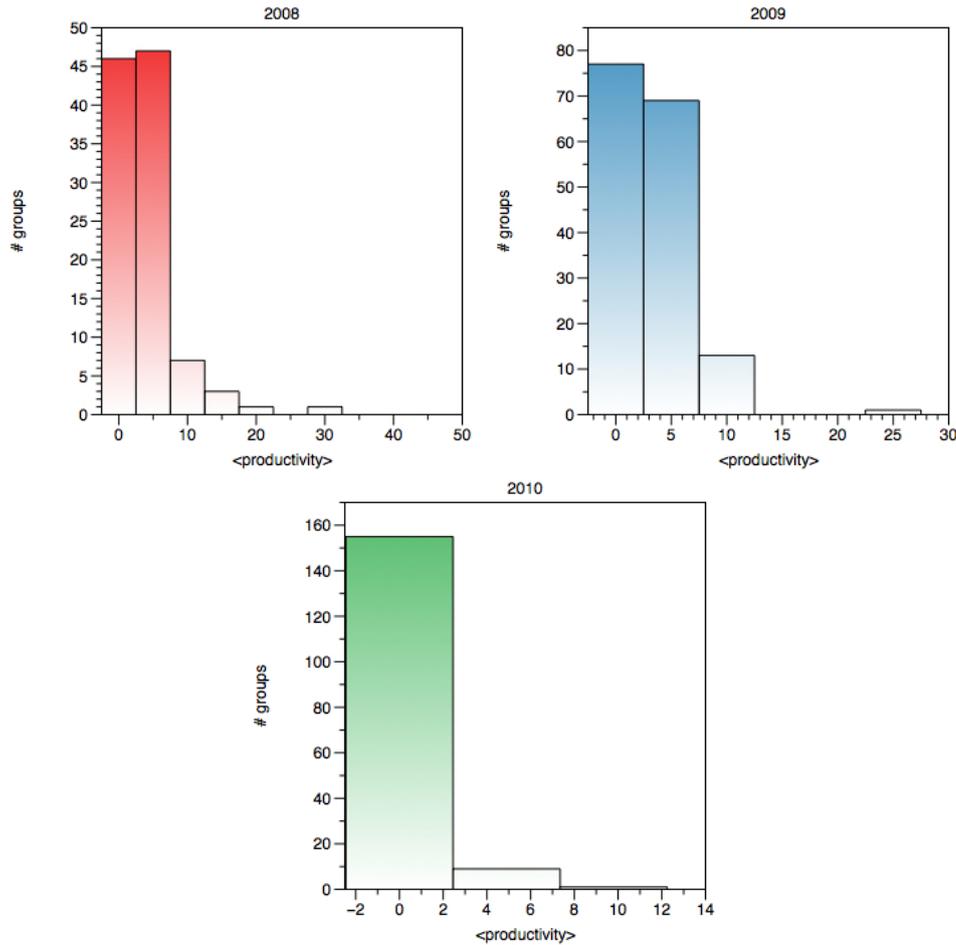


Figura 17. Histograma de productividad media para los grupos de acogida. Notar que los histogramas tienen distintos intervalos.

Es necesario recalcar que estos grupos no necesariamente son grupos constituidos de trabajo, y por lo tanto, no necesariamente continúan los años posteriores trabajando juntos. Esto requiere de un análisis más profundo.

Al igual que la productividad de los autores nuevos, en este caso decimos que el grupo de acogida es más productivo mientras más artículos haya publicado desde la incorporación del nuevo autor, en términos de publicaciones promedio

del grupo.

La Figura 17 muestra las distribuciones de productividad de los grupos de acogida para los años 2008, 2009 y 2010. El año 2011 no se representa.

Se ve que para los años 2008 y 2009 la productividad media del grupo se concentra entre 0 y 7 publicaciones posteriores a la incorporación del nuevo autor. El año 2010 entre 0 y 2, lo más probable porque han tenido menos tiempo. Sin embargo, es interesante notar que existen grupos de acogida de alta productividad aunque son muy pocos.

Si bien esta medida no es del todo precisa, por no saber si estamos hablando o no de grupos formalmente constituidos, nos puede dar una idea si es que el nuevo autor prefiere trabajar con investigadores de calificados (que publican mucho) o no.

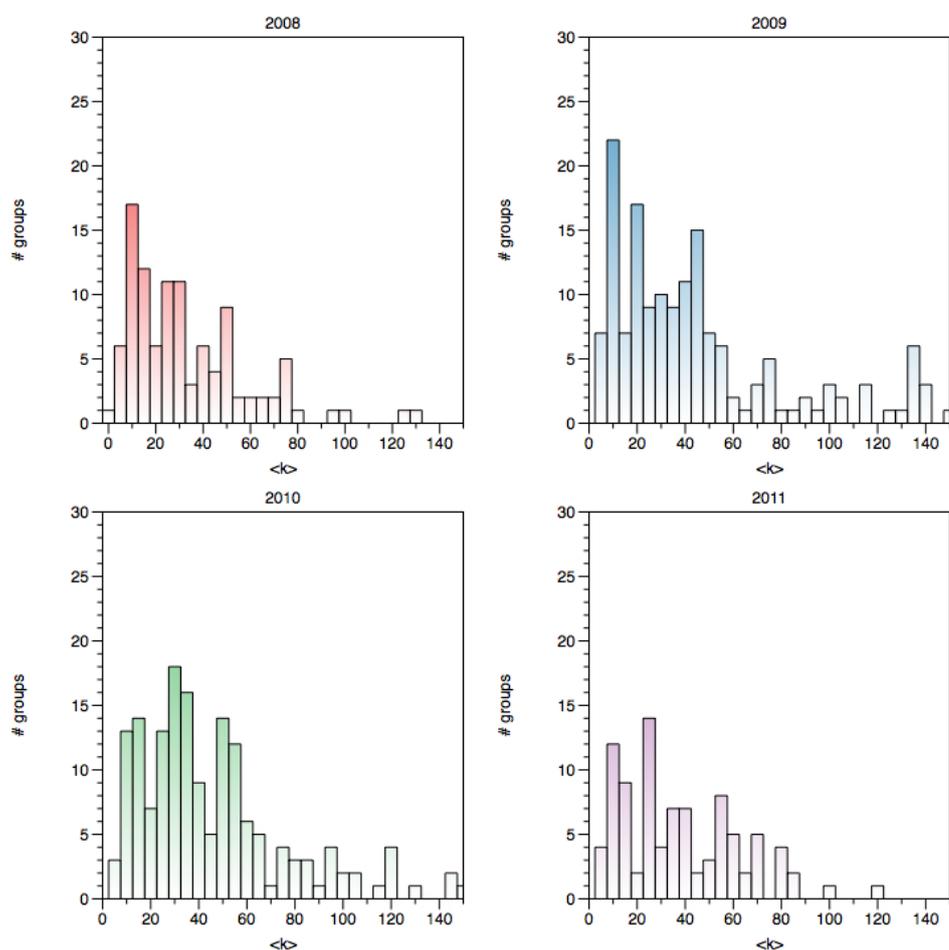


Figura 18. Histograma de conectividad media para los grupos de acogida. Notar que los histogramas tienen distintos intervalos.

Siguiendo la línea anterior, estudiamos si los nuevos autores prefieren enlazarse a personas con muchas conexiones-históricas (consideramos sus colaboraciones

desde 1991 al año de incorporación). Para esto, nuevamente tomamos la media de conectividad histórica del grupo de acogida. La Figura 18 muestra las distribuciones la conectividad media para los últimos cuatro años.

Tal como se ve, entre los años 2008 y 2009 la colaboración histórica media de los grupos de acogida está concentrada cerca de 10. En 2010 es mayor y se concentra cerca de las 30 colaboraciones. En 2011 la colaboración histórica media vuelve a bajar concentrándose entre 10 y 30 colaboraciones.

3.3. Colaboración entre Universidades de la región de Valparaíso

Por último, estudiamos la colaboración ocurrida entre las universidades tradicionales de la región y también entre éstas con instituciones nacionales de otras regiones de Chile y extranjeras. Para esto utilizamos grafos pesados en los cuales el grosor de los enlaces indica la cantidad de colaboración entre instituciones, medida como número de autores colaborando. La Figura 19 muestra esto para los últimos cuatro años.

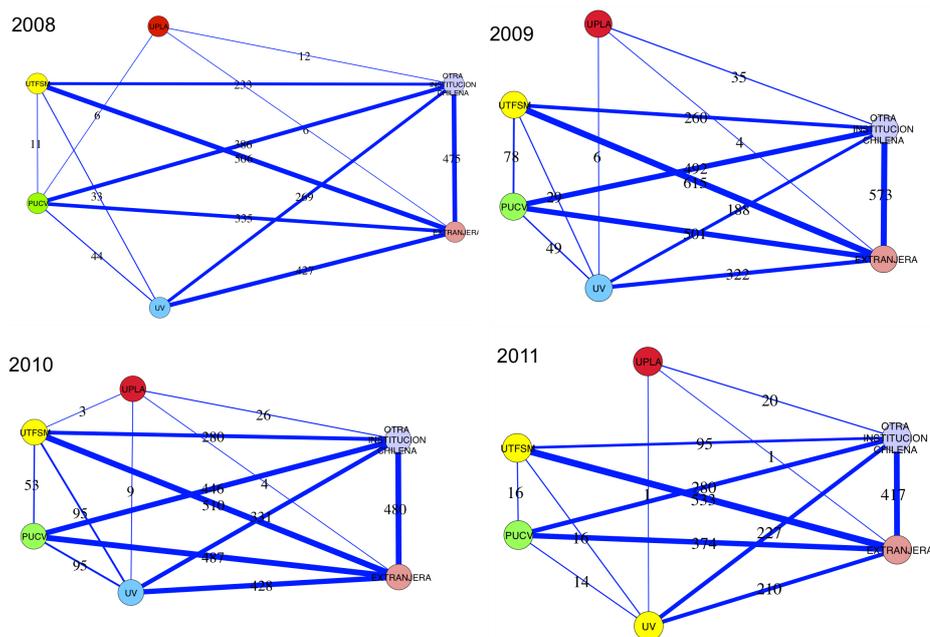


Figura 19. Colaboración entre Universidades tradicionales de la región de Valparaíso y también entre éstas con instituciones nacionales de otras regiones de Chile y extranjeras. Cada enlace está caracterizado por el número de autores que han colaborado.

Se puede ver que la colaboración entre las universidades de la región es mucho menor que la que hay entre estas universidades y otras del país y el extranjero. Destacan en este aspecto la PUCV y la UTFSM, seguida por la UV y en menor magnitud la UPLA.

Durante los cuatro años la PUCV tiene una colaboración similar, en magnitud, con universidades chilenas no pertenecientes a la región y extranjeras. Sólo el año 2011 la PUCV aparece colaborando más con el extranjero. La UTFSM por su parte colabora más con instituciones extranjeras que nacionales en todos los años. Lo mismo, aunque en menor magnitud, sucede con la UV, aunque el año 2011 colabora más con las nacionales. La UPLA colabora principalmente con instituciones chilenas externas a la región y muy poco con el extranjero.

Respecto a la colaboración entre universidades de la región podemos ver que ocurre principalmente entre tres de las cuatro universidades de la región: PUCV, UTFSM y UV. Entre éstas la colaboración es más bien pareja en los años, aunque destaca la colaboración entre la PUCV y la UTFSM con la UV. El año 2010 es el que presenta una mayor cantidad de colaboración entre las tres instituciones.

4. Conclusiones

En este trabajo estudiamos una parte de la colaboración científica de la región de Valparaíso de Chile, en particular aquella registrada en publicaciones de la base de datos de la ISI Web of Knowledge (isiknowledge.com) de Thomson Reuters. Para esto construimos redes de colaboración en las cuales los autores eran sus nodos y las publicaciones los enlaces que los relacionaban.

Después de un proceso arduo para lograr uniformidad en la base de datos logramos construir redes de colaboración con una serie de atributos para sus nodos y enlaces.

Nuestros resultados sugieren que la red de colaboración científica de la región de Valparaíso está determinada, casi exclusivamente, por las publicaciones que tres (PUCV, UTFSM y UV) de las cuatro universidades tradicionales de la región realizan en disciplinas como Física, Química y Matemática. Otras disciplinas tales como las Médicas y la Ingeniería si bien aportan con muchas publicaciones son disciplinas más bien exclusivas de algunas instituciones.

Obtuvimos evidencia de que la red de colaboración es un sistema que presenta propiedades similares a las de otros sistemas complejos, tales como distribuciones de conectividad libres de escala y propiedades de redes tipo *small world*. Sin embargo, no quisimos detenernos en tales resultados, ya descritos en sistemas similares, y nos concentramos en aquellos más interesantes desde el punto de vista del desarrollo de políticas que puedan implementarse para intervenir este sistema a futuro.

Al respecto, observamos que la colaboración, definida como publicaciones en

conjunto, es un fenómeno que ha cambiado con el tiempo. Si bien la red social de autores se ha transformado en una gran componente desde el año 2000, agrupando a más del 70% de todos los autores que han publicado artículos en Ciencias Básicas vinculados a Valparaíso hasta el año 2011, la densidad de colaboración ha disminuido con el paso de los años y la producción es cada vez más atomizada en grupos pequeños de colaboración. Esto podría responder a una nueva forma de hacer ciencia gracias a las nuevas tecnologías de la información y a la especialización de la investigación.

Nuestros resultados muestran además que a partir del año 2000 ocurrió un cambio en el escenario científico en la región que ha impulsado las publicaciones, la incorporación de nuevos autores y la colaboración con más extranjeros. Destaca durante este periodo el año 2009. Las razones para esto pueden ser varias. Por ejemplo, cerca de 2001 comenzó el Programa de Desarrollo e Innovación Tecnológica (PDIT), financiado por el Banco Interamericano de Desarrollo. Aunque esto iba más enfocado a las empresas, puede haber permeado a las Universidades. Por otro lado está el Programa Bicentenario de Ciencia y Tecnología (PBCT), un programa del Banco Mundial que entregó fondos via CORFO, CONICYT y FIA, según el área del proyecto. Los consorcios, orientados a aumentar la colaboración entre empresas y universidades, son parte del PBCT. Además, desde 2005 hubo incentivos institucionales para formar grupos de investigación intra universidades. Sin embargo, creemos que una de las razones fundamentales tras este cambio es la penetración de la Internet al ambiente académico.

Por otro lado, hemos visto que las universidades PUCV y UV son las que atraen a más autores nuevos durante los últimos años. Estos autores llegan con una alta probabilidad a publicar con grupos de poca o nula experiencia (publicando) y de tamaño reducido. Esta realidad no alentadora, se suma a la baja productividad (publicando) que tienen los nuevos autores. No obstante, debemos hacer hincapié en que no tenemos certeza de cuántos de estos nuevos autores son sólo estudiantes de paso en la institución.

Nuestros resultados nos muestran además que no existe una clara relación entre el número de publicaciones y el sector productivo de la región. A pesar de que áreas como la Pesca y la Acuicultura son prioritarias en investigación aplicada en la región, vía proyectos FONDEF, estos no se traducen en un número significativo de publicaciones. Estudiar la relación entre la productividad científica y el sector productivo de la región es uno de los temas que como grupo intentaremos abordar en trabajos futuros.

La región de Valparaíso tiene una concentración de instituciones e investigadores que le permitiría desarrollarse sobre la base de sus capacidades. Además, los programas de doctorado han consolidado líneas de investigación y aumentado la masa crítica de investigadores. Por otro lado existen áreas del conocimiento

donde la región posee ventajas y que además es posible establecer un gran potencial a partir de la interdisciplinariedad. Las ciencias físicas y químicas, con la biotecnología y las ciencias de la Ingeniería pueden tener el potencial para obtener desarrollo de vanguardia, siempre y cuando exista colaboración entre ellas. Tal como mostraban nuestros resultados, si bien existe dicha colaboración, esta es menor en comparación a la que generan las universidades de la región con otras instituciones chilenas o extranjeras.

Aunque la forma tradicional de ver cómo la Ciencia y la Tecnología aporta al desarrollo de la región es pensar en cómo los sectores productivos pueden apoyarse en las capacidades científicas, también es necesario que el sector científico sea un sector productivo más y que pueda, en si mismo, ser un foco de desarrollo de la región.

En la actualidad, en la región existe un reconocimiento de sus capacidades universitarias y científicas, las que con una vinculación y fortalecimiento, permitirían proyectar mayor generación de valor y desarrollo.

Referencias

- [1] R. Alfaro, (2008), *Estrategia regional de Ciencia, Tecnología e Innovación para la Región Valparaíso*. Informe final Kawax.
- [2] J.P. Cárdenas, (2009), *Modelo de Redes Complejas Mediante Enlace Compatible. Aplicación a Sistemas Reales*. Tesis de Doctorado. Universidad Politécnica de Madrid.
- [3] M.E.J. Newman, (2003), *The Structure and Functions of Complex Networks*: SIAM Review, 45, (2):167-256.
- [4] M. E. J. Newman, A-L. Barabási y D. Watts, (2006), *The Structure and Dynamics of Networks*, Princeton University Press.
- [5] D. J. S., Price, (1965), *Networks of scientific papers*, Science 149: 510-515.
- [6] D. J. S. Price, (1976), *A general theory of bibliometric and other cumulative advantage processes*, Journal of the American Society for Information Science and Technology 27,: 292-306.
- [7] Bettencourt, L.M.A., Kaiser D. I., Kaur J., (2009), *Scientific discovery and topological transitions in collaboration Networks*, Journal of Informetrics 3.
- [8] M. E. J. Newman, (2005), *A measure of betweenness centrality based on random walks*, Social Networks, 27, 1: 39-54.