

## **Tarea 2**

Carolina Concha Ramírez

### **1. Estadísticas descriptivas**

La red estudiada corresponde a la proyección de una red bipartita que relaciona solicitantes de derechos de agua con las cuencas correspondientes a dichos derechos. En este caso, se trabajará con la proyección correspondiente a los solicitantes, donde cada nodo corresponde a un solicitante y cada enlace representa la existencia de derechos compartidos sobre una misma cuenca.

Al calcular las métricas básicas de la red proyectada se obtuvieron los siguientes resultados:

**Tabla I. Resultados métricas básicas**

<b>Nodos</b>	4 412
<b>Aristas</b>	543 890
<b>Diámetro</b>	113
<b>Longitud de camino medio</b>	7.247295
<b>Densidad</b>	0.0559
<b>Clusterización media</b>	0.85863
<b>Clusterización global</b>	0.500
<b>Asortatividad</b>	-0.35522

El número de nodos (4 412), correspondiente al número total de solicitantes de derechos de agua, evidencia la gran cantidad de actores involucrados. Por otro lado, el número de aristas (543 890) representa el alto número de derechos concedidos. Este valor es considerablemente mayor que el número de nodos, sin embargo, si consideramos el número total de enlaces posibles entre todos los nodos, el valor de enlaces presentes es relativamente bajo. Esto se ve reflejado, además, en el bajo valor de densidad obtenido (0.056), que indica que solo un 5,6% de todos los enlaces posibles en la red estarían presentes. Esto posiblemente sea indicio de una alta presencia de nodos periféricos, con pocos enlaces, que estarían disminuyendo el valor de la densidad de la red.

El valor del diámetro (113), correspondiente a la distancia más larga entre dos nodos, es considerablemente alto en comparación al número de enlaces. De la mano con lo anterior, esto sugiere una alta presencia de nodos o conjuntos de nodos marginales muy poco conectados, los que requerirían una gran cantidad de nodos intermediarios para conectarse con otros. Por otro lado, si comparamos este valor con el obtenido para el camino medio (7.25), este último es considerablemente menor, confirmando la presencia de solicitantes más periféricos y sugiriendo la existencia de un conjunto de nodos centrales muy interconectado, con pocos pasos de distancia entre ellos.

Por otro lado, el valor de clusterización media presentado (0.86) estaría indicando una fuerte tendencia de los nodos de la red a formar grupos densamente conectados. En este contexto particular, correspondería a una alta formación de grupos de solicitantes

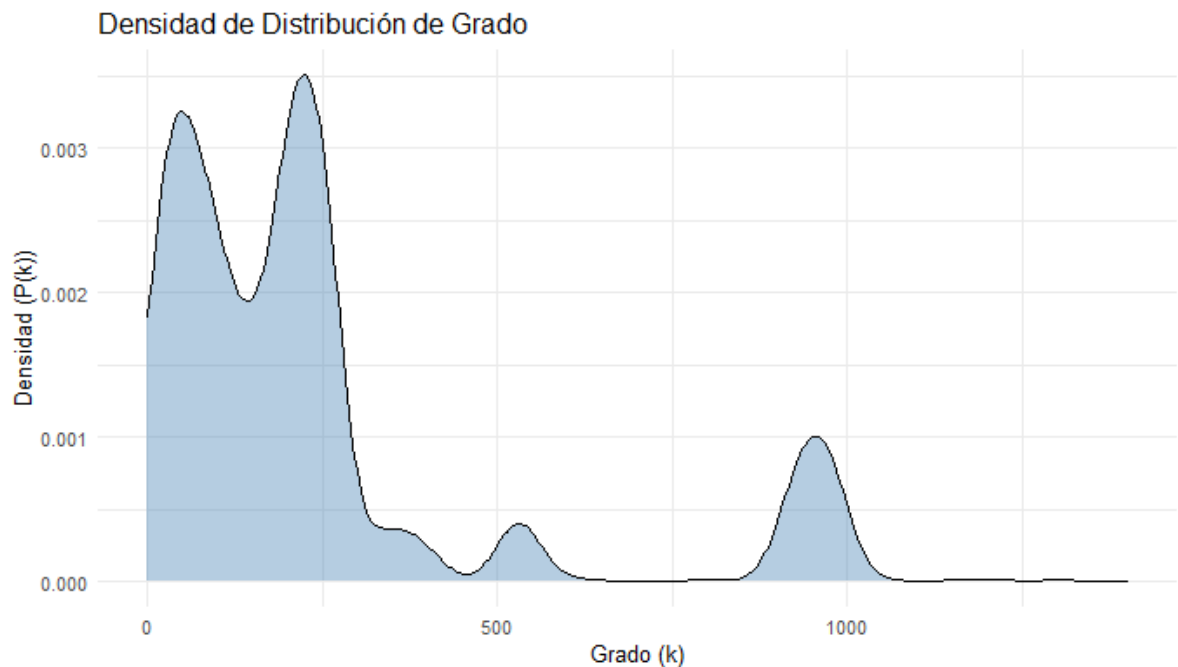
en torno a ciertas cuencas comunes. Sin embargo, el valor de clusterización global (0.5) es bastante menor, lo que indicaría que estos grupos locales no necesariamente se unirían entre sí para formar una red globalmente cohesionada.

Finalmente, la asortatividad presenta un valor negativo (-0.355), representando una red disasortativa, es decir, con la presencia de nodos muy conectados que van a tender a unirse con nodos poco conectados. En este caso, los solicitantes que compartan derechos sobre muchas cuencas (muy conectados con otros nodos) se enlazarán con muchos solicitantes pequeños (que solo tengan derechos sobre un número pequeño de cuencas), indicando que posiblemente la red presente una estructura con diversos hubs.

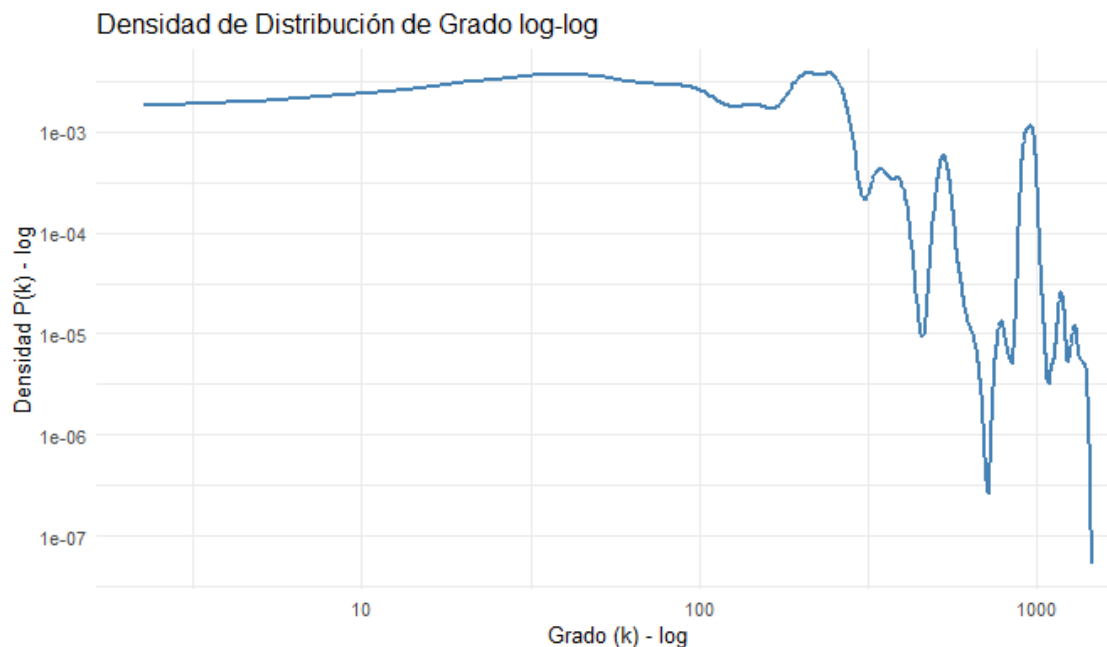
Considerando las métricas obtenidas se puede concluir que la red muestra una estructura central altamente interconectada, con conjuntos locales muy cohesionados, pero con un gran número de actores periféricos con baja conectividad, donde solo pocos actores estarían actuando como puentes entre estos y grupos más conectados, lo que se refleja principalmente por el alto valor del diámetro y la asortatividad presente. Esta configuración estaría reflejando una alta desigualdad existente en la distribución de derechos de agua entre los distintos solicitantes, donde ciertos actores estarían concentrando gran cantidad de estos en diversas cuencas, concentrando vínculos estratégicos, mientras que una gran cantidad solo tendría derechos sobre un número menor de estas.

### **Distribución de grado**

Los resultados obtenidos para la distribución de grados se muestran en las figuras 1 y 2, para escala lineal y escala logarítmica, respectivamente.



**Fig. 1 Distribución de grado**



**Fig. 2 Distribución de grado log-log**

En este caso se puede apreciar que la distribución de grado no muestra una distribución típica monótona, sino más bien una distribución multimodal, con varios peaks de grados. Específicamente, se logran apreciar 2 grupos con una cantidad de grados baja, un pequeño grupo intermedio y un grupo con un nivel importante de grados. Esto confirma lo encontrado por las métricas respecto a las distribuciones de los nodos dentro de la red. Lo que además se refleja en la distribución en escala logarítmica, donde de igual manera se logran apreciar las fluctuaciones presentes. Esta multimodalidad sugiere que la red podría estar influenciada por diversos procesos, como institucionales, legales o geográficos, lo que genera diferentes estratos de solicitantes.

Particularmente, las diferencias entre los grupos con número de grado bajo y alto indicarían una desigualdad estructural, representando una diferencia importante entre los tipos de solicitantes, pudiendo distinguir entre solicitantes pequeños y solicitantes grandes y estratégicos, los que abarcarían un mayor número de conexiones, además, el alto número de grado que presenta este grupo (cercano a 1000) representa posiblemente a actores centrales, importantes en términos de poder e influencia. De la mano de los resultados obtenidos para asortatividad, esto también sugiere la presencia importante de hubs dentro de la red.

### **Medidas de centralidad**

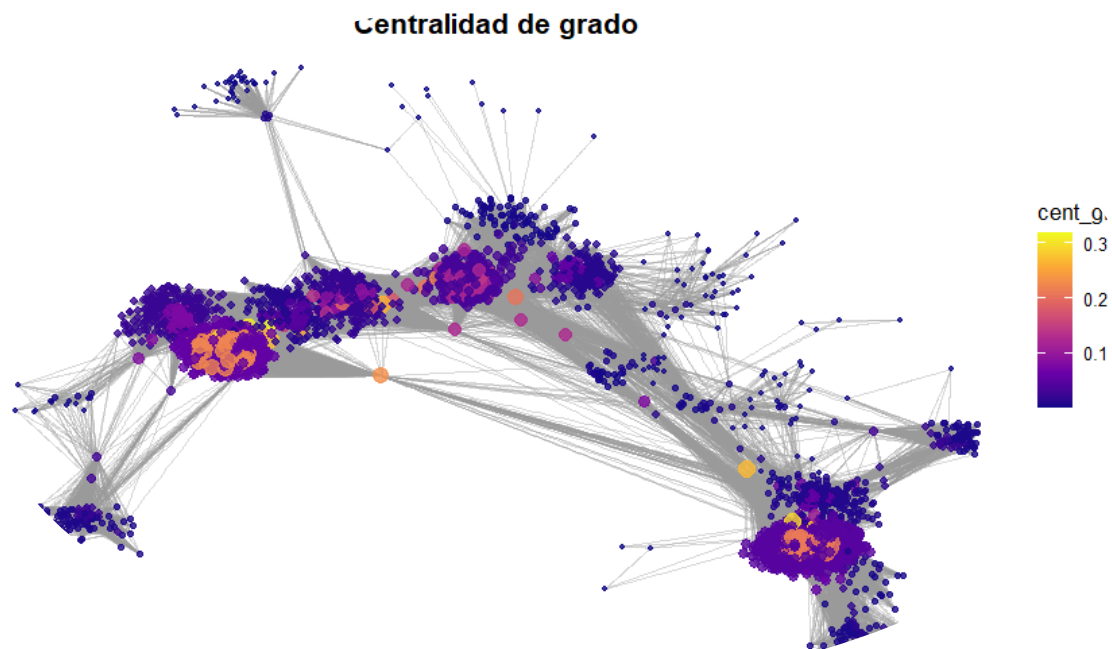
Dentro de las medidas de centralidad se incorporaron en el análisis: centralidad de grado, centralidad de intermediación (betweenness) y centralidad eigenvector. La consideración de estas centralidades se basa en las características estructurales observadas en la red y, sobre todo, en la distribución de grados presente, siendo importante identificar qué nodos son los que presentan un mayor número de

conexiones y analizar si estos se encuentran principalmente vinculados entre sí o con nodos de menor grado, además de conocer la distribución de estos últimos dentro de la red y lograr detectar la presencia de nodos intermediarios entre distintos conjuntos.

#### Centralidad de grado:

Corresponde al número de conexiones de un nodo con otros nodos, en este caso, al ser una red no dirigida, coincide además con el número de enlaces que presenta cada nodo.

La visualización de los resultados para la centralidad de grado se muestra la figura 3.



**Fig. 3 Visualización centralidad de grado**

Los resultados muestran de manera clara la existencia de grupos de nodos altamente conectados, es decir, con una alta centralidad de grado, confirmando lo señalado previamente por las métricas. Dentro de estos grupos, a su vez, destacan algunos actores claves, los que estarían concentrando una mayor cantidad de conexiones internas.

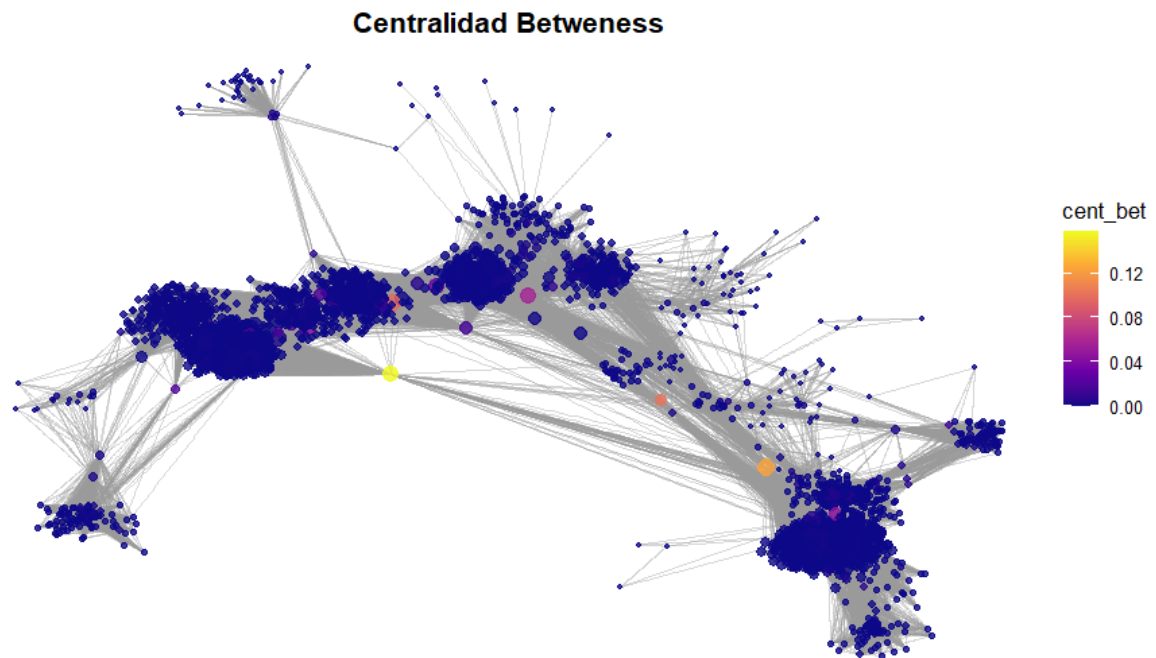
Paralelamente, se logra evidenciar una alta presencia de nodos periféricos con baja centralidad, los que estarían presentando un número menor de conexiones, así como también se presentan ciertas comunidades donde la mayoría de sus nodos presentan baja centralidad de grado, lo que podría sugerir estructuras de menor importancia o más marginales dentro de la red.

#### Centralidad de intermediación (betweness):

Corresponde a la cantidad de veces que cada nodo se presenta como puente en los caminos más cortos entre otros nodos, siendo una medida de intermediación donde

nodos con alta centralidad de intermediación tendrían una alta influencia en la conectividad de la red.

La figura 4 muestra los resultados de centralidad de intermediación obtenidos para la red.



**Fig. 4 Visualización centralidad de intermediación (betweness)**

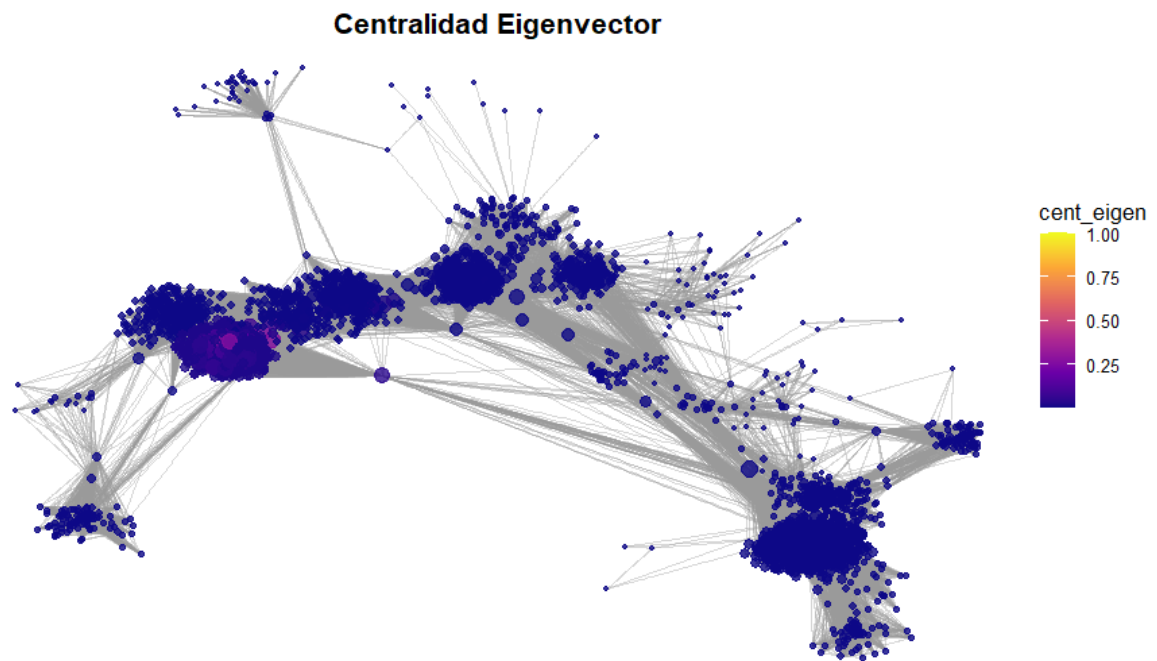
La visualización de centralidad de intermediación representa un patrón específico donde un número reducido de nodos presentan valores altos, mientras que la mayoría de nodos muestra una centralidad de intermediación baja. Esto indica la existencia de ciertos nodos específicos que actuarían como puente en los diferentes grupos formados, siendo fundamentales para mantener la unidad de la red, sin embargo, al ser una cantidad escasa de nodos, también le confieren cierta vulnerabilidad estructural a la red.

Por otro lado, comparando los resultados con los obtenidos para centralidad de grado, se logra identificar que, en general, los nodos con alta centralidad de grado tienden a presentar valores bajos de intermediación. Esto indica que, si bien pueden representar actores importantes de sus grupos debido al alto número de conexiones internas que presentan, presentan una baja conectividad con otros grupos, por lo que su influencia se limitaría localmente.

### Centralidad eigenvector:

Corresponde a la importancia de los nodos a los que un nodo particular está conectado, en ese sentido es útil para identificar conexiones con nodos influyentes y, a su vez, nodos influyentes que pueden estar conectados entre sí.

La figura 5 muestra los resultados obtenidos para esta centralidad en la red.



**Fig. 5 Visualización centralidad eigenvector**

En este caso se observa una baja presencia de nodos conectados con nodos de alta influencia, con excepción de algunos nodos puntuales dentro de los grupos que presentan nodos con alto grado. En ese sentido se puede interpretar que, a pesar de la estructura en comunidades con alto grado de interconexión en la red, la influencia de los distintos nodos se encuentra distribuida de forma relativamente homogénea a través de esta.

Considerando la naturaleza de los datos, los resultados de centralidad sugieren que, en general, solicitantes tienden a compartir derechos sobre las mismas cuencas, formando clústeres locales. Esto hace sentido considerando la distribución geográfica de las cuencas, siendo esperable que solicitantes de una misma zona geográfica van a tender a tener derechos sobre cuencas cercanas y, por lo tanto, compartir más conexiones entre sí. A su vez, dentro de cada clúster se identifica la presencia de solicitantes con mayor poder, dado por un mayor número de conexiones con otros, lo que podría indicar solicitantes con un alto poder local sobre los derechos de agua en cada zona.

Considerando esta perspectiva, los nodos con alta centralidad de intermediación corresponderían a solicitante que, al poseer derechos distribuidos en distintas zonas geográficas actuarían como vínculos entre las distintas comunidades presentes.

La presencia de nodos periféricos con baja centralidad de grado refleja la presencia de solicitantes con un menor número de derechos, lo que representa además la distinta naturaleza y alcance entre los distintos actores involucrados en el sistema de derechos de agua.

## **2. Planteamiento comparativo**

Como se mencionó anteriormente, la red en estudio corresponde a una proyección de una red bipartita donde los nodos (solicitantes) están conectados si comparten derechos sobre al menos una cuenca común.

Considerando lo anterior, dentro de los fenómenos que podrían dar origen a la formación de la red se incluyen mecanismos de “preferential attachment”, en donde hay una preferencia de los nodos por conectarse con aquellos que ya presentan un alto grado de conexiones, lo que podría estar detrás de la generación de los hubs observados. Específicamente en este caso, este mecanismo se justificaría en la preferencia por ciertas cuencas específicas en la solicitud de derechos de agua, lo que podría estar asociado tanto a factores geográficas, de accesibilidad, estratégicos, económicos y/o, incluso, regulatorios, lo que generaría que un mayor número de actores obtengan derechos sobre ciertas cuencas particulares, las que, a su vez, nuevos solicitantes también tendrán una mayor preferencia por solicitar.

Para poder evaluar estos posibles mecanismos se plantea la comparación de la red observada con un modelo libre de escala (Barbasi-Albert), con el fin de verificar si la formación de la red responde a procesos de crecimiento y “preferential attachment”.

Complementariamente, se plantea la comparación de la red observada con un modelo aleatorio (Erdos-Renyi), al ser un punto de referencia básico para el análisis de redes. Dado que en una red aleatoria los enlaces se generan de manera independiente y sin estructura, esta comparación permite confirmar si las conexiones observadas en la red responden a un mecanismo estructural, no siendo producto del azar.

En ese sentido se plantea como hipótesis que la red observada difiere significativamente de una red aleatoria (Erdos-Renyi) con el mismo número y densidad de enlaces, donde su mecanismo de formación se asemeja más a una red libre de escala (Barabasi-Albert).

## **3. Implementación de simulación/bootstrapping**

Para la comprobación de la hipótesis se realizó una comparación de métricas (densidad, clusterización y distribución de grado) entre la red observada y redes generadas mediante modelos aleatorios y libres de escala. Para esto se procedió a realizar un bootstrapping utilizando 500 iteraciones para cada tipo de modelo nulo y luego se realizó una comparación de los resultados con los obtenidos para la red observada.

Los resultados de densidad y clusterización global obtenidos para cada modelo y los valores de la red observada se muestran en la tabla II.

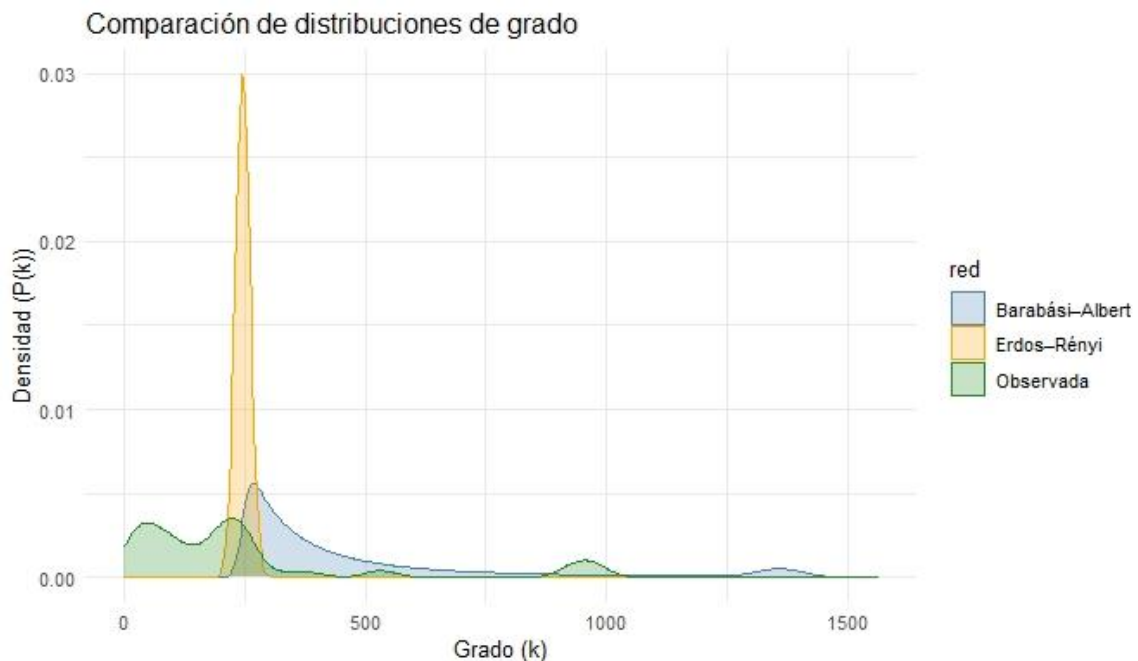
**Tabla II. Comparación valores de densidad y clusterización**

	Densidad	Clusterización global
<b>Observada</b>	<b>0.05589443</b>	<b>0.5000992</b>
<b>E-R</b>	<b>0.05587552</b>	<b>0.0558865</b>
<b>B-A</b>	<b>0.05542663</b>	<b>0.1349984</b>

En el caso de la densidad, las 3 redes muestran similitud en los resultados, lo que refleja que la comparación es válida y controlada, indicando que las principales diferencias con los modelos nulos se presentan en la distribución de los enlaces entre los nodos y no en el número de estos, permitiendo identificar diferencias entre los mecanismos de formación de estos.

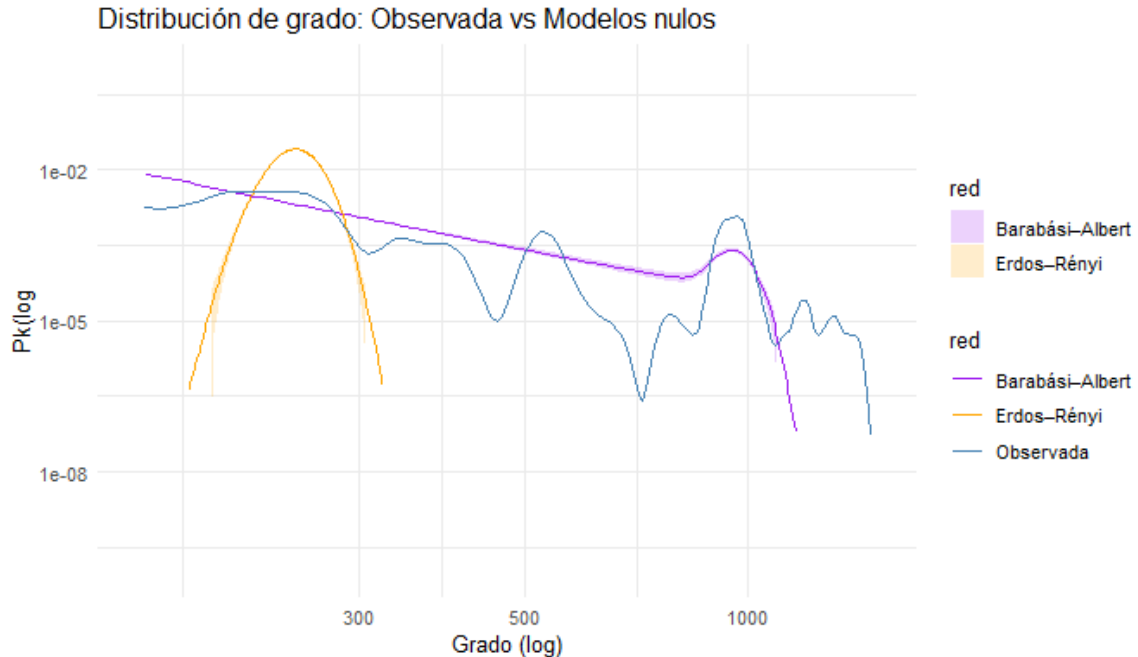
Por otro lado, los valores obtenidos para clusterización global varían significativamente en cada caso. La red observada presenta un valor de clusterización elevado (0.50), mientras que los valores presentes en los modelos nulos son más bajos. Esto indica que la red presenta una gran cantidad de clústers, los que no son capaces de explicarse a partir de ninguno de los 2 modelos considerados.

Las figuras 6 y 7 muestran la comparación en la distribución de grados de los 3 casos, correspondientes a escala lineal y log-log, respectivamente.



**Fig. 6 Comparación distribución red observada – modelos nulos**





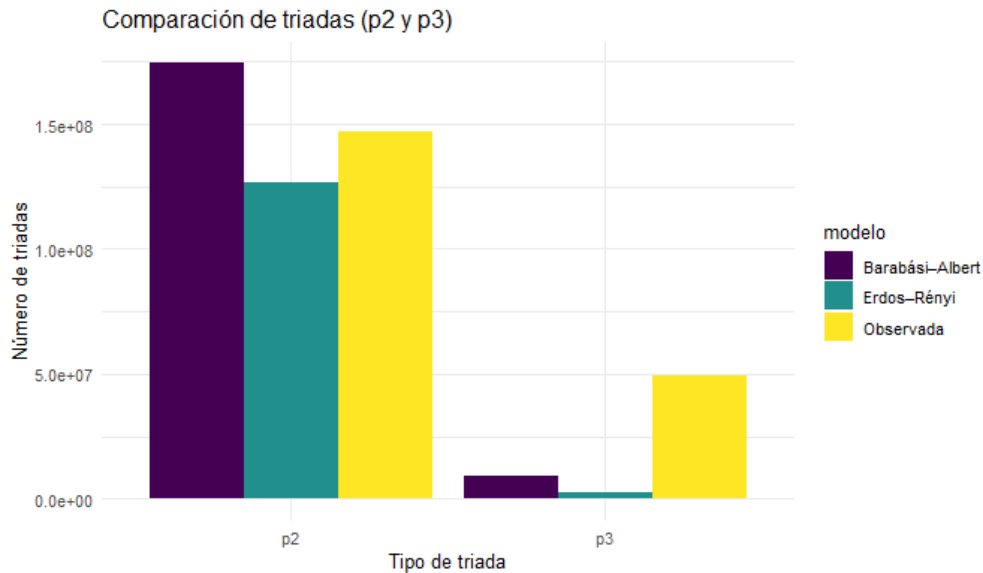
**Fig. 7 Comparación distribución escala log-log red observada – modelos nulos**

De las distribuciones de grado se puede apreciar que para el caso del modelo aleatorio y libre de escala las distribuciones obtenidas corresponden a sus distribuciones características. En el caso de la red observada se observan ciertos patrones característicos ya discutidos previamente (punto 1).

Al comparar las distribuciones de grado, la red observada muestra ciertas similitudes con el modelo libre de escala, como su tendencia negativa y la mayor concentración de nodos con grado bajo. Sin embargo, la red observada muestra una mayor variabilidad en comparación al modelo nulo, representando fluctuaciones locales, así como también la presencia de una cola más pesada y una mayor concentración de nodos con grado muy bajo, además de una estructura multimodal. Estos resultados indican que este modelo podría representar una parte de los mecanismos presentes en la formación de la red, pero no estaría representando el total de estos. El hecho de que la red observada no siga una línea recta en escala log-log confirma que no es una red libre de escala clásica, sino algo más complejo.

Por otro lado, con relación al modelo aleatorio, se aprecian diferencias notables entre esta y la red observada, donde la distribución de grado se aleja significativamente de la presentada por la red real, confirmando que un mecanismo aleatorio no sería adecuado para explicar la formación de esta.

Finalmente, se realizó un censo triádico entre los distintos modelos y la red observada, en este caso se consideraron los nodos p2 y p3 para realizar la comparación (figura 8).



**Fig. 8 Resultado censo triádico**

En el caso de p2, es decir, nodos que forman triadas abiertas, la red observada presentaría un número de nodos intermedio entre ambos modelos nulos, estando más cercana al valor obtenido para el modelo aleatorio. Esto estaría indicando que el número total de triadas abiertas no se desvía de forma significativa en comparación a los procesos generativos nulos.

Por otro lado, en el caso de p3, es decir, nodos que forman triadas cerradas, se presenta una diferencia importante entre el número presente en la red observada en comparación a los modelos nulos. Esto confirmaría una alta clusterización en la red que no logra ser explicada por mecanismos de formación aleatoria ni de “preferential attachment”.

En conclusión, a partir de la información obtenida de la comparación de la red con los modelos nulos, se evidencia que la red cuenta con propiedades estructurales únicas, reflejando procesos organizacionales complejos. En ese sentido, los mecanismos generadores no estarían siendo aleatorios, si no que se aproximarían en mayor medida a mecanismos de “preferential attachment”, validando la hipótesis planteada. Sin embargo, estos mecanismos tampoco lograrían explicar la totalidad de los procesos detrás de la formación de la red real, específicamente, no estaría siendo capaces de representar la clusterización presente, por lo que la red real presentaría una naturaleza híbrida y estaría respondiendo a mecanismos de formación complejos, propios del fenómeno de estudio.