



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA

INGENIERÍA EN SISTEMAS

Redes Complejas del Flujo de Pasajeros en Aeropuertos Mexicanos

TESIS PARA OPTAR POR EL GRADO DE:

MAESTRA EN INGENIERÍA DE SISTEMAS

(INVESTIGACIÓN DE OPERACIONES)

PRESENTA:

Act. Olivia Sashiko Shirai Reyna

TUTOR:

Dra. Idalia Flores de la Mota

PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA INGENIERÍA EN SISTEMAS

CD. MX. , SEPTIEMBRE 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

JURADO ASIGNADO:

Presidente: M.I. Soler Anguiano Francisca Irene

Secretario: Dra. Segura Pérez Esther

Vocal: Dra. Flores De La Mota Idalia

1^{er.} Suplente: Dra. Obregón Quintana Bibiana

2^{do.} Suplente: Dra. Huerta Barrientos Aida

Lugar o lugares donde se realizó la tesis: Ciudad Universitaria, UNAM

TUTOR DE TESIS:

Dra. Idalia Flores De La Mota

FIRMA

(Segunda hoja)

Agradecimientos

Quiero agradecer a mis papás José Akio Shirai Inouye y Olivia Reyna Pérez por siempre estar conmigo, apoyarme en mis decisiones, por ser tan tolerantes y pacientes conmigo. De igual forma quiero agradecer a mi hermana Yoshi y a toda mi familia por su apoyo incondicional.

Quiero agradecer al Posgrado de Ingeniería de Sistemas de la UNAM por darme esta oportunidad en especial a mi tutora Dra. Idalia Flores de la Mota y a mis sinodales, M.I. Francisca Irene Soler Anguiano, Dra. Aida Huerta Barrientos, Dra. Esther Segura Pérez, en especial a la Dra. Bibiana Obregón Quintana por su paciencia, consejos y sus acertados comentarios.

Quiero agradecer a todas las personas involucradas en este proceso en especial a todos mis amigos.

Por último, agradezco a CONACYT por la beca otorgada para realizar mis estudios.

Índice

Introducción.....	8
Capítulo 1 Descripción de la Problemática.....	12
1.2 Estadísticas Actuales	12
Capítulo 2 Metodología	17
2.1 Primeros pasos	17
2.2 Algoritmo de visibilidad.....	18
2.3 Redes Complejas	21
2.4 Modelos de Redes Complejas	31
2.4.1 Modelo de Redes Aleatorias	31
2.4.2 Modelo de Mundo Pequeño (Modelo WS)	32
2.4.3 Modelo de Redes de Libre Escala (Modelo BA)	33
2.5 Series de Tiempo	35
Capítulo 3 Análisis de Resultados.....	58
Capítulo 4 Conclusiones	77
Anexos	79
Anexo 1. Gráficas series de tiempo pasajeros nacionales	79
Anexo 2. Gráficas series de tiempo pasajeros internacionales.....	81
Anexo 3. Gráficas series de tiempo pasajeros totales	83
Anexo 4. Redes complejas resultado del algoritmo de visibilidad para pasajeros nacionales, internacionales y totales	85
Anexo 5. Distribución de grado	89
Anexo 6. Descomposición series de tiempo	93
Anexo 7. Función de Auto Correlación.....	97
Anexo 8. Función de Auto Correlación Parcial	101
Anexo 9. Pronóstico a 24 meses	105
Anexo 10 Resultados métricas pasajeros nacionales.....	108
Anexo 11 Tablas métricas pasajeros internacionales	109
Anexo 12 Tablas métricas pasajeros totales	110
Bibliografía	113

Tabla de figuras

Figura 1 Mapa aeropuertos México	9
Figura 2 Participación de mercado de los aeropuertos en transportación de pasajeros domésticos.....	13
Figura 3 Participación de mercado de los aeropuertos en transportación de pasajeros internacionales.....	14
Figura 4 Metodología.....	17
Figura 5 Representación algoritmo de visibilidad	19
Figura 6 Ejemplo de algoritmo de visibilidad	20
Figura 7 Transformación de series de tiempo	21
Figura 8 Representación Puentes de Königsberg	22
Figura 9 Representación Red dirigida y Red no dirigida	24
Figura 10 Representación matriz de adyacencia.....	24
Figura 11 Ejemplo definiciones.....	25
Figura 12 Representación Modelo de Redes Aleatorias	32
Figura 13 Representación Modelo de Redes de Mundo Pequeño	33
Figura 14 Representación Modelo de Redes de Libre Escala.....	34
Figura 15 Comparación Modelos de Redes Complejas	34
Figura 16 Ejemplos de series de tiempo	35
Figura 17 Serie de tiempo pasajeros aéreos (1949-1960)	36
Figura 18 Gráfica función de auto correlación	40
Figura 19 Gráfica función de autocorrelación parcial	41

Figura 20 Descomposición aditiva de series de tiempo.....	44
Figura 21 Descomposición multiplicativa de series de tiempo	45
Figura 22 Tipos de suavizamiento.....	46
Figura 23 Ejemplo suavizamiento.....	47
Figura 24 Representación etapas modelación de series de tiempo	53
Figura 25 Ejemplos de series de tiempo Pasajeros Nacionales de Acapulco y Monterrey	58
Figura 26 Ejemplos de series de tiempo Pasajeros Internacionales de Acapulco y Monterrey	59
Figura 27 Ejemplos de series de tiempo Pasajeros Totales de Acapulco y Monterrey	59
Figura 28 Ejemplos del algoritmo de visibilidad en aeropuertos de Acapulco y Monterrey	60
Figura 29 Ejemplos de redes de los aeropuertos de Acapulco y Monterrey.....	61
Figura 30 Ejemplos distribución de grado aeropuertos de Acapulco y Monterrey	63
Figura 31 Ejemplo descomposición series de tiempo Pasajeros de Acapulco	64
Figura 32 Ejemplo descomposición de series de tiempo Pasajeros de Monterrey	65
Figura 33 Ejemplos funciones de autocorrelación y autocorrelación parcial Pasajeros Acapulco.....	66
Figura 34 Ejemplos funciones de autocorrelación y autocorrelación parcial Pasajeros Acapulco.....	67
Figura 35 Ejemplos funciones de autocorrelación y autocorrelación parcial Pasajeros Monterrey	68
Figura 36 Pronóstico a 24 meses pasajeros Nacionales, Internacionales y Totales de Acapulco.....	71
Figura 37 Pronóstico a 24 meses pasajeros Nacionales, Internacionales y Totales de Monterrey	72
Figura 38 Comparación redes real contra pronóstico Acapulco Nacionales	73

Figura 39	Comparación redes real contra pronóstico Acapulco Internacionales .	73
Figura 40	Comparación redes real contra pronóstico Acapulco Totales	73
Figura 41	Comparación Distribución de Grado Acapulco	75

Tabla de tablas

Tabla 1	Estadística de pasajeros domésticos por principales aeropuertos	12
Tabla 2	Estadística de pasajeros internacionales por principales aeropuertos	13
Tabla 3	Número de pasajeros y porcentaje de cambio	14
Tabla 4	Rutas más transitadas aeropuerto Acapulco 2016.....	15
Tabla 5	Rutas más transitadas aeropuerto Monterrey 2016	16
Tabla 6	Resumen modelos.....	50
Tabla 7	Resultados métricas	61

Introducción

Hoy en día, el tráfico aéreo se ha incrementado cada vez más, siendo una opción viable y accesible para un gran número de personas. Actualmente, México es el tercer país con el mayor número de aeropuertos, aeródromos y pistas de vuelo con un total de 1714, esto según datos publicados por la CIA (Central Intelligence Agency) World Factbook en 2013. Esto ocasiona que los aeropuertos requieran una organización eficiente para dar servicio tanto a las empresas que utilizan las instalaciones, como a los usuarios finales. Además, es importante considerar que se genera demasiada información que puede no ser fácil de analizar. Por este motivo es importante hacer un diagnóstico sobre la operación actual de los aeropuertos mexicanos, detectar los principales problemas y con base en ello proponer alternativas de mejora tanto en sus operaciones como en su infraestructura. Por lo anterior, en esta tesis se realiza un análisis de la información obtenida de algunos aeropuertos nacionales e internacionales de la República Mexicana, ya que no se conocen.

El problema principal es que no se conocen ni se han identificado las fortalezas y debilidades de este sistema, ocasionando, entre otras cosas, pérdidas en las oportunidades de negocio. Por esa razón, es importante realizar un diagnóstico inicial con aplicación de modelos matemáticos a datos reales para su modelación con redes complejas, a fin de proponer mejoras para las estrategias de mercado. Por lo que se requiere la aplicación de modelos estadísticos y modelos matemáticos para analizar, encontrar patrones, identificación de tendencias y predecir el comportamiento de los pasajeros tanto nacionales como internacionales y totales que es resultado de sumar el número de pasajeros nacionales (domésticos) e internacionales.

Desde el año 2000, se ha llevado un registro de la información generada en aeropuertos nacionales e internacionales de la república mexicana, la cual no ha sido analizada. Con dicha información, el primer paso será realizar un análisis estadístico de los datos, después, construiremos la serie de tiempo del flujo de los pasajeros (nacionales, internacionales y totales), posteriormente, aplicaremos el algoritmo de visibilidad, el cual nos permite realizar una transformación de la serie de tiempo en una red compleja. Una vez obtenida dicha red, analizaremos la estructura o topología de la red mediante métricas de redes como las medidas de centralidad (centralidad intermedia, centralidad de grado, el grado de cercanía) que nos dan información sobre la importancia de los nodos dentro de la red, distancia media, diámetro, asortatividad, clustering (local y global), distribución de grado, entre otras. Con lo que se podrá analizar y comparar las redes de los pasajeros nacionales, internacionales y totales, además, se podrá analizar la topología de

estas redes y compararlas con las de los demás aeropuertos, con lo que concluimos a qué tipo de modelo de redes pertenecen y qué características y propiedades específicas comparten.

Para este estudio nos enfocamos en el análisis de los pasajeros nacionales, internacionales y totales atendidos por distintos aeropuertos mexicanos desde enero del 2000 hasta marzo del 2017. Los aeropuertos que utilizaremos en este estudio son: Acapulco, Chihuahua, Culiacán, Durango, Mazatlán, Monterrey, Reynosa, San Luis Potosí, Tampico, Torreón, Zacatecas y Zihuatanejo.



Figura 1 Mapa aeropuertos México, Fuente: <http://www.oma.aero/es/aeropuertos/mapas.htm> (Se modificó la imagen para eliminar al aeropuerto de Ciudad Juárez)

Realizaremos el estudio a las ciudades antes mencionadas, pero nos enfocaremos principalmente en los aeropuertos de Acapulco y Monterrey, ya que Acapulco es uno de los mayores centros turísticos de la república atrayendo turismo tanto nacional como internacional. Por otro lado, Monterrey es la tercera ciudad más grande de México y uno de los centros industriales más productivos de la república, por lo que el aeropuerto de Monterrey ocupa el cuarto lugar a nivel nacional por el número de operaciones y pasajeros atendidos.

Acapulco es uno de los destinos turísticos más importantes de la Costa del Pacífico Mexicano al recibir anualmente a más de 5 millones de turistas. Desde los tiempos remotos cuando se traían bienes y especies de Acapulco, la ciudad y su bahía han sido una parte importante de la historia de México. La ciudad es ahora sede de grandes e importantes eventos culturales, deportivos y de espectáculos. Es un destino turístico privilegiado con dos espléndidas bahías, Santa Lucía y Puerto Marqués, poseedoras de bellezas naturales e impresionantes acantilados, hermosas y variadas playas de diferentes oleajes, rodeadas por exótica vegetación tropical.

Acapulco cuenta principalmente con 3 temporadas de alta ocupación: La primera va de enero a abril (incluyendo Semana Santa); la segunda es la temporada de verano, que abarca los meses de Julio y Agosto; y la tercera es la de invierno, a partir de la segunda semana de diciembre. Durante estas épocas las tarifas de hoteles, restaurantes y centros de entretenimiento varían. Actualmente operan 11 aerolíneas ofreciendo 10 destinos.

Se tiene previsto la creación de una nueva Terminal Aérea para el Aeropuerto Internacional General Juan N. Álvarez de Acapulco. Son \$547 millones de pesos destinados para la construcción de la Nueva Terminal Aérea, la cual contará con 2 niveles, dicha rehabilitación se desarrollará en 2 fases y tendrá 18 mil metros con capacidad de 1.3 millones de pasajeros por año (con posibilidad de 2 millones en mediano plazo). La obra dio inicio en el mes de julio de 2016 y concluirá durante el segundo trimestre del 2018. [<http://www.oma.aero/es/aeropuertos/acapulco/>].

Por su parte, Monterrey, capital del estado de Nuevo León y unas de las tres ciudades más importantes del país gracias a su gran actividad comercial e industrial la cual se caracteriza por ser una ciudad industrial y financiera, es una de las ciudades más desarrolladas de México, se posiciona como la segunda mejor desarrollada en el país, sólo por detrás de la Ciudad de México. Tiene una gran cultura e historia, sin mencionar que es además es el principal centro de negocios, industrial y económico muy importante para México. Fue nombrada como una de las 10 ciudades más felices del mundo. Es la ciudad con mejor calidad de vida en México.

El aeropuerto de Monterrey está dividido en tres terminales. La Terminal A se divide en dos zonas la Norte (A) donde parten vuelos nacionales, en las puertas 1-15; y la zona Sur (B) desde donde parten los internacionales, puertas 3-8. En las Terminales B y C se ofrecen tanto destinos nacionales como internacionales. Actualmente operan 13 aerolíneas ofreciendo 47 destinos. [<http://www.oma.aero/es/aeropuertos/monterrey/>].

Esta tesis se desarrolla de la manera siguiente en el capítulo 1 se encuentra la introducción al objetivo de la tesis, así como la descripción de la problemática actual, en el capítulo 2 se hace una descripción de la metodología y se hace una descripción de las herramientas que se utilizarán en la tesis, en el capítulo 3 se desarrolla la metodología con las herramientas descritas en el capítulo 2 y se realiza un análisis de los resultados obtenidos, el capítulo 4 corresponde a las conclusiones, en la sección de anexos se encuentran todas las gráficas, tablas y figuras adicionales y por último tenemos la referencias.

Capítulo 1 Descripción de la Problemática

1.2 Estadísticas Actuales

En este primer capítulo Se mostrará el panorama actual que se vive en el país en cuestión de operaciones aeroportuarias, además, se mostrará la problemática actual que tienen los pasajeros tanto nacionales como internacionales en los aeropuertos mexicanos.

Como un primer acercamiento a los aeropuertos mexicanos se realizó un estudio de las principales estadísticas de los aeropuertos en general, además de que se realizó un comparativo del número de pasajeros nacionales (domésticos) del año 2016 contra el año 2017 y encontramos los siguientes resultados.

Tabla 1 Estadística de pasajeros domésticos por principales aeropuertos

Estadística de pasajeros domésticos por principales aeropuertos Domestic air passenger traffic by main airports Dic 2016 vs Dic 2017 / Jan-Dec 2016 vs Jan-Dec 2017 miles/ thousand								
Top 10	dic-16	dic-17	var	% var	Ene-Dic 16 / Jan-Dec 16	Ene-Dic 17 / Jan-Dec 17	var	% var
Ciudad De México/Mexico City	2,581.8	2,474.0	-107.9	-4.2%	27,354.2	28,769.9	1415.8	5.2%
Guadalajara	735.9	842.9	107.1	14.5%	7,693.3	9,021.6	1328.3	17.3%
Monterrey	681.2	708.1	26.9	3.9%	7,917.6	8,420.4	502.7	6.3%
Cancun	650.0	683.4	33.4	5.1%	6,844.2	7,808.4	964.2	14.1%
Tijuana	661.0	668.8	7.9	1.2%	6,260.7	7,046.5	785.7	12.5%
Culiacan	169.6	187.0	17.3	10.2%	1,781.1	1,947.1	166.0	9.3%
Merida	172.7	182.5	9.8	5.7%	1,699.6	1,828.1	128.4	7.6%
Hermosillo	130.2	142.1	11.9	9.1%	1,450.4	1,532.7	82.4	5.7%
San Jose Del Cabo	122.1	133.3	11.3	9.2%	1,264.2	1,461.6	197.3	15.6%
Bajo	107.8	126.8	19.0	17.6%	1,267.6	1,389.6	122.0	9.6%
Otros/ Others	1,877.5	1,884.6	7.1	0.4%	19,722.8	20,737.1	1014.4	5.1%
Total / Total	7,890	8,034	143.7	1.8%	83,256	89,963	6,707.2	8.1%

Fuente: SCT, SST, DGAC, DDE. Cifras proporcionadas por los grupos aeroportuarios.
Source: SCT, SST, DGAC, DDE. Information provided by the airports groups
Nota: No se están contabilizando pasajeros en tránsito ni en aviación general

Es importante conocer cuáles son los aeropuertos con mayor número de flujo de pasajeros nacionales dentro del país y encontramos los siguientes resultados.



Figura 2 Participación de mercado de los aeropuertos en transportación de pasajeros domésticos

Tabla 2 Estadística de pasajeros internacionales por principales aeropuertos

Estadística de pasajeros internacionales por principales aeropuertos								
International air passenger traffic by main airports								
Dic 2016 vs Dic 2017/ Jan-Dic 2016 vs Jan-Dec 2017								
miles/ thousand								
Top 10	dic-16	dic-17	var	% var	Ene-Dic 16 / Jan-Dec 16	Ene-Dic 17 / Jan-Dec 17	var	% var
Ciudad De México/Mexico City	1,355.4	1,529.5	174.2	↑ 12.9%	14,053	15,750	1696.7	↑ 12.1%
Cancun	1,362.3	1,474.8	112.5	↑ 8.3%	14,572	15,793	1221.5	↑ 8.4%
Guadalajara	373.0	368.1	-4.8	↓ -1.3%	3,669	3,758	89.0	↑ 2.4%
Puerto Vallarta	325.5	341.7	16.2	↑ 5.0%	2,723	3,044	320.7	↑ 11.8%
San Jose Del Cabo	282.1	297.8	15.6	↑ 5.5%	2,825	3,240	415.3	↑ 14.7%
Monterrey	117.0	119.8	2.8	↑ 2.4%	1,261	1,263	1.9	↑ 0.1%
Bajio	61.2	73.2	12.1	↑ 19.7%	596	645	49.2	↑ 8.3%
Mazatlan	36.4	38.1	1.7	↑ 4.5%	279	293	14.3	↑ 5.1%
Cozumel	35.2	35.8	0.6	↑ 1.7%	397	405	7.6	↑ 1.9%
Morelia	25.0	28.9	3.9	↑ 15.5%	268	296	28.2	↑ 10.5%
Otros/Others	214	233	18.6	↑ 8.7%	2,125	2,242	116.8	↑ 5.5%
Total / Total	4,187	4,541	353.35	↑ 8.4%	42,767	46,729	3,961.19	↑ 9.3%

Fuente: SCT, SST, DGAC, DDE. Cifras proporcionadas por los grupos aeroportuarios
 Source: SCT, SST, DGAC, DDE. Information provided by the airports groups



Figura 3 Participación de mercado de los aeropuertos en transportación de pasajeros internacionales

Principalmente nos enfocaremos en conocer la problemática de los aeropuertos de Acapulco y Monterrey, por lo que, en las siguientes tablas encontraremos información relevante al tráfico de pasajeros en estos aeropuertos. En la Tabla 3, se muestran el número de pasajeros totales de desde el año 2000 hasta el 2017, además, se muestra el porcentaje de cambio con respecto al año anterior.

Tabla 3 Número de pasajeros y porcentaje de cambio

AÑO	Pasajeros totales	% cambio	Pasajeros totales	% cambio
	ACA		MTY	
2000	1 018 288	-	3,680,397	-
2001	940 197	▼ 7.6%	3,479,221	▼ 5.46%
2002	793 420	▼ 15.6%	3,446,469	▼ 0.94%
2003	774 349	▼ 2.4%	3,703,288	▲ 7.45%
2004	821 301	▲ 6.0%	4,293,816	▲ 15.94%
2005	880 190	▲ 7.1%	4,660,138	▲ 8.53%
2006	994 286	▲ 13.0%	5,253,600	▲ 12.73%
2007	1 057 347	▲ 6.3%	6,559,613	▲ 24.85%

Continuación de **Tabla 3**.

AÑO	Pasajeros totales ACA	% cambio	Pasajeros totales MTY	% cambio
2008	1 087 974	▲ 2.9%	6,586,190	▲ 0.40%
2009	839 048	▼ 22.9%	5,199,895	▼ 21.04%
2010	736 878	▼ 12.2%	5,380,412	▲ 3.47%
2011	596 326	▼ 19.1%	5,582,794	▲ 3.76%
2012	546 951	▼ 8.2%	6,105,910	▲ 9.37%
2013	617 079	▲ 12.9%	6,417,755	▲ 5.10%
2014	631 570	▲ 2.3%	7,128,531	▲ 11.07%
2015	730 382	▲ 15.65%	8,461,917	▲ 18.71%
2016	718 493	▼ 1.65%	9,178,533	▲ 8.46%
2017	685 125	▼ 4.64%	9,771,630	▲ 6.46%

Por otro lado, las Tablas 4 y 5 muestran un resumen de las rutas nacionales más transitadas, número de pasajeros por ruta y las aerolíneas que proporcionan dicho servicio, todo con respecto al año 2016.











Tabla 4 Rutas más transitadas aeropuerto Acapulco 2016

Número	Ciudad	Pasajeros	Ranking	Aerolínea
1	 Ciudad de México	253,728	—	Aeromar, Aeroméxico, Aeroméxico Connect, Interjet, Volaris
2	 Tijuana, Baja California	34,369	▲ 1	Volaris
3	 Monterrey, Nuevo León	28,350	▲ 1	Magnicharters, Viva Aerobus, Volaris
4	 Toluca, México	13,447	▼ 2	Interjet, TAR
5	 Querétaro, Querétaro	3,282	—	TAR

Continuación **Tabla 4**

Número	Ciudad	Pasajeros	Ranking	Aerolínea
6	 Ixtapa-Zihuatanejo, Guerrero	2,922	—	TAR
7	 Guadalajara, Jalisco	689	▲5	TAR
8	 Oaxaca, Oaxaca	26	▼1	

Tabla 5 Rutas más transitadas aeropuerto Monterrey 2016

Número	Ciudad	Pasajeros	Ranking	Aerolínea
1	 Ciudad de México	1,586,219	—	Aeroméxico, Aeroméxico Connect, Interjet, Viva Aerobus, Volaris
2	 Cancún, Quintana Roo	473,818	—	Aeroméxico, Interjet, Magnicharters, Viva Aerobus, Volaris
3	 Guadalajara, Jalisco	320,544	—	Aeroméxico Connect, Interjet, TAR, Viva Aerobus, Volaris
4	 León, Guanajuato	127,325	—	Aeroméxico Connect, Interjet, Viva Aerobus, Volaris
5	 Tijuana, Baja California	123,819	▲8	Aeroméxico Connect, Viva Aerobus, Volaris
6	 Hermosillo, Sonora	111,681	▼1	Aeroméxico Connect, Viva Aerobus, Volaris
7	 Chihuahua, Chihuahua	106,085	▼1	Aeroméxico Connect, TAR, Viva Aerobus, Volaris
8	 Querétaro, Querétaro	103,712	▼1	Aeroméxico Connect, TAR, Volaris
9	 Toluca, Estado de México	99,361	—	Aeroméxico Connect, Interjet, Volaris
10	 Veracruz, Veracruz	97,407	▼2	Aeroméxico Connect, Viva Aerobus, Volaris

Capítulo 2 Metodología

En esta sección se realizará una descripción de los pasos a seguir para el análisis de los datos obtenidos, el proceso para transformar los datos a series de tiempo y poder utilizar las diferentes herramientas estadísticas para su análisis. Además, de la aplicación del algoritmo de visibilidad para poder hacer la transformación de las series de tiempo en redes con complejas, esto con el fin de poder utilizar las diferentes métricas descritas para poder identificar su estructura, topología y comportamiento.

2.1 Primeros pasos

Para realizar este análisis utilizamos un software de libre acceso R y RStudio que son un entorno y lenguaje de programación con un enfoque al análisis estadístico. Para llevar acabo todo nuestro análisis utilizamos distintas paqueterías o librerías, que nos permiten hacer diferentes cálculos, gráficas, entre muchas más funciones. Las librerías que utilizamos son *sand*, *igraph*, *TSA*, *timeSeries*, *forecast*, *ggplot2*, *ggnetwork*, *GGally*, además de las que ya trae por determinado R.

Como se comentó en la introducción, contamos con los datos de 12 aeropuertos de la República Mexicana (Acapulco, Chihuahua, Culiacán, Durango, Monterrey, Mazatlán, San Luis Potosí, Reynosa, Tampico, Torreón, Zacatecas y Zihuatanejo), los datos que tenemos de estos aeropuertos son: el número de pasajeros nacionales, internacionales y totales desde enero del 2000 hasta marzo del 2017.

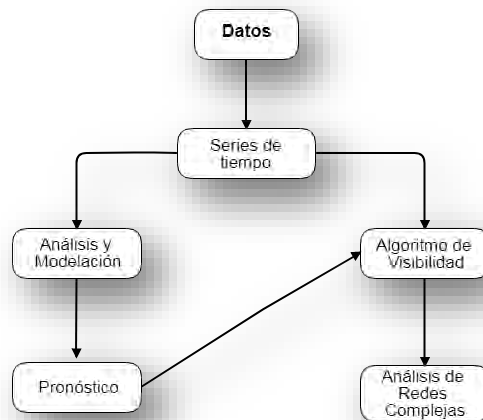


Figura 4 Metodología

2.2 Algoritmo de visibilidad

En los últimos años se han desarrollado varios algoritmos que asocian una red compleja a una serie de tiempo, lo que permite la aplicación de muchas herramientas topológicas para analizar la dinámica del sistema subyacente que se generó a partir de las series de tiempo. Esta conexión se estableció por primera vez en el contexto de series de tiempo pseudo-periódicas por Zhang y Small, donde los ciclos en las series de tiempo están asociados a nodos en la red compleja. Los nodos correspondientes a ciclos que ocurren en cierta proximidad están relacionados. Esto resultó ser útil para distinguir entre ciertos sistemas dinámicos caóticos y series de tiempo periódicas con ruido. Posteriormente, Zhang, Lue y Small desarrollaron una generalización para aplicarlo a series de tiempo generales. Esta generalización requiere una incrustación (a través de la transformación de coordenadas de tiempo con retraso u otros medios) de la serie de tiempo en un espacio de fase más grande, y se requiere una serie de tiempo lo suficientemente grande ($O(10^4)$). Este método fue explorado más a fondo con cierto progreso utilizando el enfoque de comportamiento de la superfamilia de motivos reticulares. Otro método basado en gráficas de recurrencia donde la matriz de recurrencia se interpretó como la matriz de adyacencia de una red compleja asociada al sistema dinámico.

Ambos métodos se basan en el estudio de algún tipo de recurrencia del sistema o de la serie de tiempo. El tercer método se basa en el análisis de la convexidad de las observaciones sucesivas, es decir, cada punto de la serie de tiempo se asocia con un nodo en la gráfica o red de visibilidad correspondiente, y los arcos se establecen de acuerdo con varios algoritmos de visibilidad.

El algoritmo de visibilidad se ha introducido recientemente como un mapeo entre series de tiempo y redes complejas. Este procedimiento nos permite aplicar métodos de teoría de redes complejas para caracterizar series de tiempo. El algoritmo de visibilidad es un algoritmo computacional fácil y rápido.

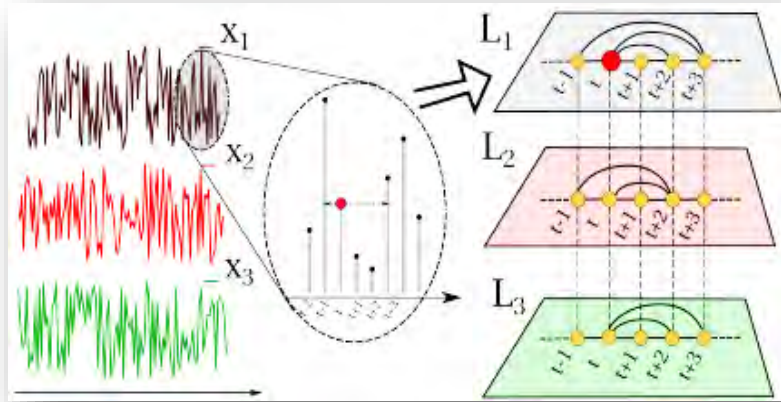


Figura 5 Representación algoritmo de visibilidad

Lo que hace este algoritmo es transformar la serie de tiempo en una red compleja, donde los nodos corresponden en el mismo orden a los datos de la serie, mientras que dos nodos están conectados si existe una visibilidad entre ellos, es decir, si hay una línea que conecta a los datos de la serie, a esta línea se le conoce como línea de visibilidad.

Por lo que la formulación matemática de este algoritmo es la siguiente; sean dos valores arbitrarios (t_a, y_a) y (t_b, y_b) , tendrán visibilidad y sus nodos estarán conectados si otro dato (t_c, y_c) está situado entre ambos y se cumple:

$$y_c < y_b + (y_a - y_b) \frac{t_b - t_c}{t_b - t_a} \quad (2.1)$$

En la siguiente Figura se muestra cómo funciona el algoritmo de visibilidad aplicado a una serie de tiempo con 20 datos.

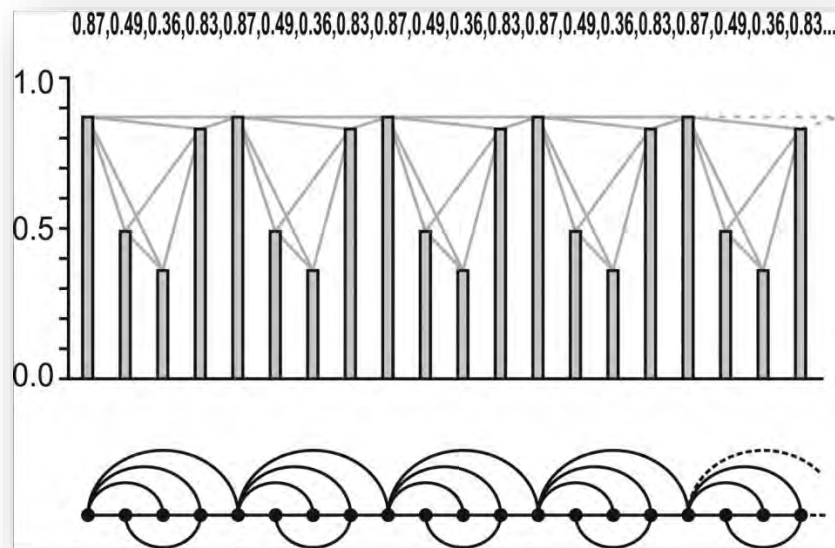


Figura 6 Ejemplo de algoritmo de visibilidad, (Fuente: Lacasa et al 2008 [14])

Las redes obtenidas de la aplicación del algoritmo de visibilidad a series de tiempo tienen propiedades específicas de mucha importancia, todas estas propiedades fueron expuestas por Lacasa et al 2008 [14], en su artículo *From time series to complex networks: the visibility graph*.

- Conexa: todos los nodos de la red están conectados, ya que cada nodo puede ver a sus vecinos más cercanos (izquierda y derecha).
- No dirigida: por la forma en la que el algoritmo está construido, no se tiene una dirección definida en los arcos.
- Invariante ante transformaciones a la serie de tiempo: el criterio de visibilidad es invariante ante reescalamientos verticales y horizontales en los ejes además de traslaciones de los datos horizontales y verticales.

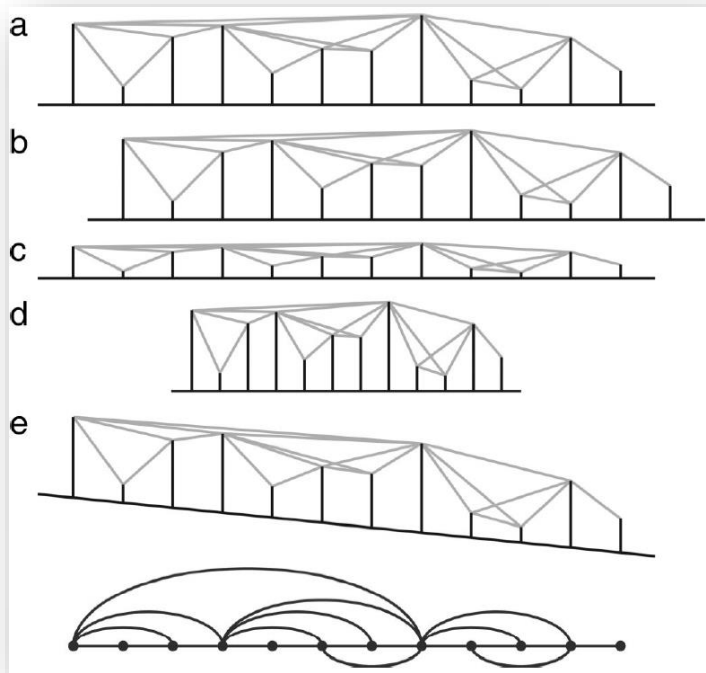


Figura 7 Transformación de series de tiempo, (Fuente: Lacasa et al 2008 [14])

En la Figura anterior se ve gráficamente las transformaciones de la serie de tiempo, en el inciso a) es la serie original con los arcos de visibilidad, b) traslación de los datos, c) reescalamiento vertical, d) reescalamiento horizontal, e) se añade una tendencia a la serie de tiempo.

Se han hecho diferentes estudios [14, 15], principalmente a la topología de las redes obtenidas de diversas series de tiempo se llegó a la conclusión de que si se tienen series de tiempo periódicas el algoritmo de visibilidad las transforma en redes regulares, las series de tiempo no periódicas o aleatorias el algoritmo las convierte en redes aleatorias, mientras que las series de tiempo fractales el algoritmo de visibilidad las convierte en redes de libre escala. Por lo que para nuestros fines es importante hacer una introducción a qué son las redes complejas.

2.3 Redes Complejas

La sociedad, el desarrollo de las ciudades, los avances tecnológicos, incluso, nosotros mismos, dependemos del funcionamiento de una gran diversidad de sistemas, los cuales pueden ser modelados mediante redes. En los últimos años el estudio de las redes complejas ha tenido un gran auge debido a la gran cantidad de problemas de la vida real que se pueden estudiar mediante redes, es decir, por sus

propiedades estructurales y dinámicas. Por lo que el uso de redes se ha hecho muy popular en diferentes áreas como la física, biología, economía, fisiología, sociología, medicina, neurología, entre otras. Podemos encontrar redes biológicas, redes genéticas, redes neuronales, redes metabólicas, redes de comunicación, redes sociales, redes de transporte, redes de información, entre muchas más. Por lo que es importante comenzar con la idea de ¿qué es una red?, una red es un conjunto de entidades (nodos), unidos por una relación específica (arcos). Específicamente nos centraremos en redes de transporte, en particular, en redes de aeropuertos (flujo de pasajeros en los aeropuertos) donde los nodos corresponden a los años, los arcos al flujo de pasajeros y las capas o niveles corresponden a cada aeropuerto.

Los primeros estudios datan de 1736, con un problema conocido como “*Los Puentes de Königsberg*” en Prusia (actualmente, Kaliningrado, Rusia). El problema consistía en encontrar un camino que recorriera los siete puentes que atravesaban el río Pregel, sin repetir ninguno y regresando al punto de origen. Dicho problema fue resuelto por el matemático suizo Leonard Euler, quien estableció que el problema no tenía solución, sin embargo, se dio el inicio de la teoría de gráficas porque Euler representó el problema de una forma abstracta, donde la tierra la representó con puntos y los puentes con líneas, ver fig. 1. Podemos observar que la gráfica tiene 3 nodos de grado 3 y un nodo de grado 5.



Figura 8 Representación Puentes de Königsberg, (Fuente: <http://bosoneando.blogspot.mx/2015/07/solucion-la-pregunta-de-la-ardilla-de.html>)

Euler estableció que el problema de los Puentes de Königsberg no tenía solución, pero no fue hasta 1873 cuando Carl Hierholzer prueba que este problema tiene solución. Él determinó que este tipo de caminos existen sí y solo sí, cada vértice de la gráfica tiene un número par de aristas. En general, el progreso en el estudio de las gráficas fue muy lento, en parte debido a la dificultad de trabajar con redes grandes.

En 1950, dos grandes matemáticos Paul Erdős y Alfred Rényi, propusieron un

modelo de redes aleatorias, lo que da inicio a la “Teoría de Redes”. La aportación principal que hicieron Erdős y Rényi fue que, dada una probabilidad de interconexión entre nodos, al irse incrementando, llega un momento en que la mayoría de los nodos están unidos en un componente gigante.

En 1998, los matemáticos Duncan Watts y Steven Strogatz llevaron a cabo un estudio en el cual se centraron en propiedades de conectividad, con lo que mostraron que las redes se podían clasificar con base en dos propiedades: la distancia y el coeficiente de aglomeración (en inglés Clustering). Cabe mencionar que, la distancia media es pequeña en la mayoría de las redes, esta propiedad recibe el nombre de Mundo Pequeño. Además, introducen un nuevo modelo, el Modelo de Mundo Pequeño. Estas investigaciones fueron publicadas en el artículo “*Collective Dynamics of Small-World Networks*” (Dinámica colectiva de las redes de mundo pequeño). Un año más tarde, en 1999, Albert-Lászlo Barabási y Reka Albert estudiaron redes que presentan nodos con mayor conectividad que los demás nodos, por lo tanto, tienen un mayor grado de conexión que el grado promedio de los nodos de la red. A estos nodos con una conectividad muy alta, se les llaman *hubs*. Estas investigaciones fueron publicadas en el artículo “*Emergence of scaling in random networks*” (La emergencia de la escala en las redes aleatorias).

Sea una red $G = (N, A)$, donde N es el conjunto de nodos o vértices, y A es el conjunto de arcos o aristas o enlaces. En el contexto de una red social, los nodos reciben el nombre de actores y los arcos de vínculos, mientras que en teoría de gráficas a los nodos se les conoce como vértices y a los arcos como aristas. Se han desarrollado una gran cantidad de métricas y herramientas para clasificar y analizar las redes, una medida básica es el grado de un nodo k , el cual consiste en el número de enlaces incidentes en dicho nodo. En ocasiones una red puede tener una dirección privilegiada donde los arcos se representan con flechas (ver fig. 2), en este caso, el grado se divide en grado interior (el arco entra al nodo), y grado exterior (el arco sale del nodo). La distribución de grado $P(k)$ es fundamental en el conocimiento de la topología de una red, se determina graficando el grado de los nodos vs el número de nodos. Las redes con topología homogénea presentan una distribución tipo Poisson para N grande, debido a que todos los nodos poseen más o menos el mismo grado. En cambio, las redes heterogéneas presentan una distribución libre de escala, la cual en un plano log-log representa una línea recta con pendiente negativa, debido a que existen nodos con un alto grado de conectividad llamados hubs, es decir, unos cuantos nodos atraen la mayoría de las conexiones y la gran parte de los nodos restantes tienen que compartir los enlaces que sobran.

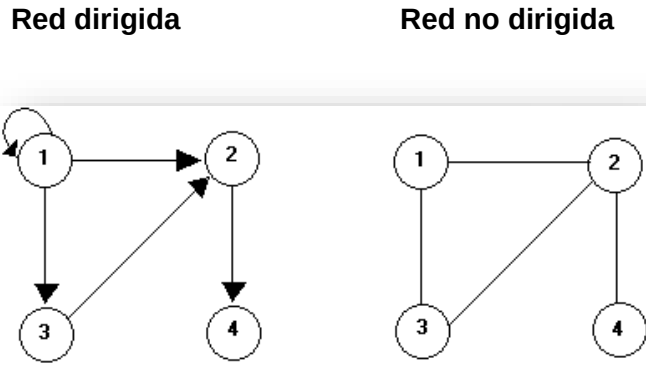


Figura 9 Representación Red dirigida y Red no dirigida, (Diseño propio)

Una de las maneras que tenemos para representar a las redes es por medio de matrices, la matriz de adyacencia nos da una representación de la relación nodo-nodo que se tiene en una red, de forma que tenemos una matriz de $N \times N$, en donde la entrada ij toma el valor de 1 si existe un arco del nodo i al nodo j , en otro caso es 0. En la fig. 3 se muestra una red con su respectiva matriz de adyacencia. Este tipo de matrices representa quién está cerca de quién.

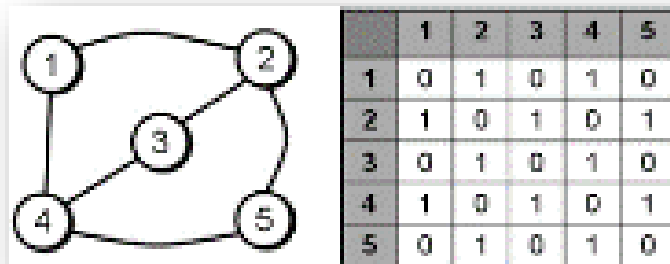


Figura 10 Representación matriz de adyacencia, (Diseño propio)

La pareja ordenada (i, j) , representa el arco que va del nodo inicial i al nodo final j , además, es común representar el arco con una flecha, $i \rightarrow j$. Un camino entre el nodo i y el nodo j , es una secuencia alternada de nodos y arcos, la cual comienza en el nodo i y termina en el nodo j . De hecho, hemos escuchado las palabras camino, trayectoria, ruta y podríamos pensar que se refieren a lo mismo o que son sinónimos, pero en el caso de las redes estas palabras nos llevan a conceptos distintos, es decir, cuando hablamos de una ruta nos referimos a un camino en el cual los nodos son distintos, en cambio cuando nos referimos a una trayectoria

estamos hablando de un camino en el cual los arcos son distintos. Además, una ruta o trayectoria cerrada, en las cuales no es posible identificar el nodo inicial del nodo final, se conocen como ciclo y circuito, respectivamente. Se dice que una red es conexa si existe un camino entre cada par de nodos que conforman la red, es decir, todos los nodos están conectados, no hay nodos aislados.

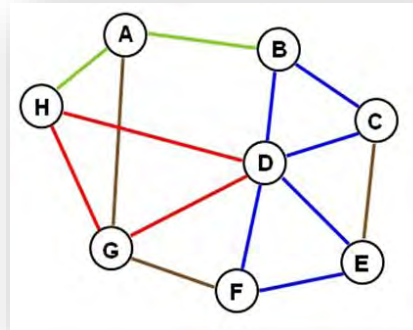


Figura 11 Ejemplo definiciones, fuente:
https://commons.wikimedia.org/wiki/File:Graph_cycle.gif#/media/File:Graph_cycle.gif

En la Figura 4, mostramos gráficamente las definiciones anteriores, por ejemplo, si seguimos los nodos $A \rightarrow B \rightarrow D \rightarrow G \rightarrow H \rightarrow D \rightarrow F \rightarrow E \rightarrow C$, (En teoría de gráficas se utiliza la siguiente notación (A, B) , (B, D) , (D, G) , (G, H) , (H, D) , (D, F) , (F, E) , (E, C) . Aunque la notación que estamos utilizando es más utilizada en teoría de redes.) Tenemos un camino, pero en el caso que siguiéramos la secuencia de nodos $A \rightarrow G \rightarrow H \rightarrow D \rightarrow F \rightarrow E \rightarrow C \rightarrow B$ tenemos una ruta, si tuviéramos esta misma ruta, solo que al final se conectará con A , es decir, $A \rightarrow G \rightarrow H \rightarrow D \rightarrow F \rightarrow E \rightarrow C \rightarrow B \rightarrow A$ estaríamos hablando de un ciclo. Ahora bien, en el caso de que tuviéramos lo siguiente $A \rightarrow H \rightarrow G \rightarrow D \rightarrow F \rightarrow E \rightarrow D \rightarrow C \rightarrow B \rightarrow D$ estamos hablando de una trayectoria, pero si tuviéramos la misma trayectoria, pero añadiendo de $D \rightarrow G \rightarrow A$, es decir, quedaría como $A \rightarrow H \rightarrow G \rightarrow D \rightarrow F \rightarrow E \rightarrow D \rightarrow C \rightarrow B \rightarrow D \rightarrow G \rightarrow A$ estamos hablando de un circuito. Si queremos conocer un poco sobre las métricas de una red, podemos empezar por calcular la longitud de ruta, esta nos dice el número de arcos que contiene una ruta, sin embargo, no solo podemos saber cuántos arcos están involucrados en nuestra ruta, sino que además podemos conocer tanto la distancia como la distancia media (p). La distancia entre dos nodos i y j , se define como el mínimo número de arcos que conectan a los nodos i y j , en tanto que la distancia media nos da la distancia de forma más general, ya que nos dice el promedio de las distancias entre todo par de nodos. Por último, podemos hablar del

diámetro de una red, el cual es la ruta más larga de entre todas las rutas más cortas entre todo par de nodos.

Por otro lado, cuando hablamos de la vecindad de un nodo, nos referimos a los nodos que se encuentran a la misma distancia de un determinado nodo, es decir, existe un número de arcos fijo de separación hacia un determinado nodo. La eficiencia de una red mide el intercambio de información dentro de la red, se calcula mediante el inverso de la ruta más corta ($\varepsilon_{ij} = \frac{1}{d_{ij}}$) y se refiere, a cómo se mueve la información dentro de una red. Se puede calcular la eficiencia global en la cual se considera toda la red o se puede calcular la eficiencia local de un nodo en específico, en el cual se considera la primera vecindad del nodo en cuestión. En el caso de que no exista la ruta entre el nodo i y el nodo j entonces, la distancia entre el nodo i y j es infinito, por lo tanto, la eficiencia es 0.

Además, existen diferentes tipos de redes como son redes regulares, redes vacías y, redes completas. Las redes regulares, son aquellas en las que los nodos que conforman la red tienen el mismo número de arcos o grado, mientras que las redes vacías, son aquellas redes las cuales no tienen arcos, por lo que es una red no conexa, y, por último, las redes completas, son aquellas redes en las cuales existe un arco entre cada par de nodos. Por otra parte, los árboles tienen características especiales ya que son acíclicos (no presentan ciclos), son conexos, es una red no dirigida, cuenta con N nodos y $N - 1$ arcos, existe una ruta única entre cada par de nodos y, además, si eliminamos un arco se desconecta, pero si añadimos un arco se forman ciclos, y en ambos casos ya no estaríamos hablando de un árbol. La densidad de una red, es una medida expresada en porcentaje del cociente entre el número de relaciones existentes con los posibles enlaces, es decir, nos muestra la alta o baja conectividad de la red.

Una red está constituida por un conjunto de componentes, estos componentes constan de nodos y arcos, incluso podemos hablar de un solo nodo y decir que es un componente, utilizamos el término componente gigante cuando la mayoría de los nodos se encuentran conectados. Es importante definir los componentes fuertemente conexos y débilmente conexos. El componente fuertemente conexo se define en una red dirigida como, el subconjunto de nodos en que para cualquier par de nodos (ij) existe una ruta directa desde el nodo i hasta el j y viceversa, además, puede consistir de un solo nodo. Cuando hablamos de componente débilmente conexo, nos referimos a que dos o más nodos están conectados por una o más rutas sin importar la dirección.

Por otro lado, los cliques son importantes dentro de las redes, si aplicamos la definición formal, ésta nos dice que es un subconjunto máximo de nodos dentro de una red no dirigida, en la que cada uno de los nodos está conectado a todos los demás nodos, es decir, es una subred o subgráfica completa dentro de la red. El k -core de una red, es la subgráfica más grande en la cual todos sus nodos tienen al menos grado k . La colección completa de los k -cores, es decir, $k = 1, 2, 3$, nos da una descripción más detallada de la red, que cuando únicamente se consideran los componentes, en el caso de un árbol (red acíclica y conexa), solo existe el 1-core.

Un concepto que utilizamos con frecuencia es el coeficiente de aglomeración (Clustering), el cual es una forma de calcular la tendencia a formar triángulos dentro de la red, también conocido como transitividad de un nodo. Para determinarlo, se calcula el cociente del número de arcos entre los vecinos del nodo i , y el número máximo de los posibles enlaces entre los vecinos del nodo i , esto se puede ver como la cantidad de triángulos conectados al nodo dividido entre los triángulos incidentes a él, es decir, nos indica la tendencia que tiene la red de formar triángulos. Podemos calcular el coeficiente de aglomeración local (la vecindad del nodo) y general (toda la red). Cuando el clustering es igual a 0 nos dice que hay nodos aislados, es decir, que los nodos no están conectados por ningún arco. Si el clustering es igual a 1, esto nos indica que todos los nodos están conectados por un arco directo. Los nodos que presentan un grado bajo tienden a pertenecer a comunidades locales bien interconectadas, mientras que los hubs se conectan con muchos nodos que no están directamente interconectados. Un ejemplo rápido del coeficiente de aglomeración dentro de las redes sociales es, si dos personas A y B se conocen y B conoce a C, entonces existe una probabilidad alta de que A y C se conozcan, por lo que se dice también que se tiene una relación transitiva y se forma un triángulo.

Clustering local:

$$C_i = \frac{2|C_3(i)|}{k_i(k_i - 1)} \quad (2.2)$$

Clustering global:

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i \quad (2.3)$$

Donde:

$C_3(i)$ = número de triángulos que conectan al nodo i

k_i = grado del nodo i

n = número de nodos

La centralidad intermedia (Betweenness centrality), es una medida de distancia por medio de la cual se calcula la centralidad de cada uno de los nodos, es decir, nos dice qué tan importante es un nodo dentro de la red. La idea es que un nodo es central siempre y cuando esté situado en la ruta más corta para llegar a todo par de nodos, si existen varias rutas más cortas, se asume que todos tienen la misma probabilidad de ser usadas. Los nodos con alta intermediación juegan el rol de conectar a los diferentes grupos. La siguiente fórmula nos muestra la manera en la que se define:

$$C_B(i) = \sum_{j,k \neq i} \frac{g_{jk}(i)}{g_{jk}} \quad (2.4)$$

Donde:

$g_{jk}(i)$ = longitud de las rutas más cortas que conectan al nodo j y al nodo k , que pasan a través del nodo i

g_{jk} = ruta más corta entre los nodos j y k

Una medida de centralidad más sofisticada es el grado de cercanía (Closeness centrality), debido a que mide el inverso de la distancia media desde un nodo hacia los demás nodos de la red. El grado de cercanía puede ser considerado como una forma de medir cuánto tiempo tomará para que la información se propague de un

determinado nodo a los otros nodos de la red, razón por la cual es muy utilizado en el análisis de redes sociales. Se calcula de la siguiente manera:

$$C_c(n_i) = \sum_j \frac{1}{p_{ij}} \quad (2.5)$$

Donde:

p_{ij} = distancia media entre el nodo i y el nodo j

n_i = nodo i

Mark Newman [20] propuso una clasificación importante conocida como asortatividad (Assortativity), la cual nos indica la preferencia de los nodos de una red por unirse a otros nodos que son similares en alguna característica. Es decir, los nodos con un grado alto tienden a conectarse a otros nodos con un grado alto, por lo que la comunicación fluye más rápido. En caso contrario, una red es disortativa (Disassortative) cuando nodos de bajo grado se conectan con hubs, por lo que hay una tendencia a la generación de cuellos de botella, esto significa que la comunicación no fluye tan rápido por lo que se puede llegar a saturar la red. Estudios realizados han encontrado que las redes tecnológicas y biológicas, presentan una estructura negativa, es decir, son disortativas, en cambio en las redes sociales presentan un número positivo, por lo tanto, son asortativas. El coeficiente de asortatividad depende de la correlación, se calcula con la siguiente fórmula:

Coeficiente de Asortatividad:

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j q_k)}{\sigma_q^2} \quad (2.6)$$

Donde:

e_{jk} = es el exceso de la probabilidad conjunta del exceso de grado de j y del exceso de grado de k (el exceso de grado de un nodo también conocido como el grado restante, es igual al grado del nodo menos 1)

$q_k = \frac{(k+1)p_{k+1}}{\sum_j j p_j}$ es la distribución normalizada del exceso del grado D_r de un nodo elegido aleatoriamente

σ_k = desviación estándar de q_k

La fórmula anterior nos da valores de r para la cual $-1 < r < 1$, esto significa que cuando r nos da un valor de 1 o muy cercano, se dice que la red es asortativa, cuando r es igual a 0 esto nos indica que la red es neutra y cuando r nos da un valor de -1 o muy cercano, nos dice que la red es disortativa.

Por otro lado, la robustez es un indicador el cual nos ayuda a saber qué tan vulnerable es nuestra red a perturbaciones. Existen dos tipos de perturbaciones, la primera se refiere a fallas internas que se presentan dentro de la red, esto es, debido a una posible falla o a un problema interno, a esto se le conoce como error interno, mientras que el segundo tipo de perturbaciones son los ataques, en este caso la falla o el problema está fuera de la red, se pueden tener ataques dirigidos a un componente específico. Si el componente gigante continúa, sin ningún daño grave se dice que la red resiste estos errores o ataques por lo que la red es robusta.

La propiedad de mundo pequeño consiste en el hecho de que la distancia media entre cualesquiera dos nodos es muy pequeña, es decir, si un nodo dentro de una red tiene pocos nodos conectados a él y pocos alejados de él, entonces la mayoría de los nodos están a una distancia media que es relativamente corta. Lo anterior se refiere a que, a partir de un determinado nodo, casi todos los nodos están a muy pocos pasos de él, por lo que, aunque añadamos muchos nodos a la red con sus respectivos arcos, la distancia media no aumentaría mucho, ya que existe una ruta

más corta promedio entre los nodos, es decir, que, aunque no todos los nodos son vecinos entre sí, se puede llegar a la mayoría de los nodos en un número de pasos pequeño.

2.4 Modelos de Redes Complejas

2.4.1 Modelo de Redes Aleatorias

El primer modelo de redes complejas que se conoció fue el modelo de redes aleatorias, el cual fue propuesto por los matemáticos húngaros Erdős y Rényi en 1959 [12]. En este modelo se inicia con un número de nodos N , los cuales se enlazan con la misma probabilidad, por lo que en este tipo de redes todos los nodos tienen más o menos el mismo número de enlaces. Las redes que poseen un número de nodos N pequeño siguen una distribución de grado binomial, pero en el caso en el que el número de nodos N es grande siguen una distribución de grado Poisson, es decir, presentan una conectividad media porque la mayoría de los nodos tienen el mismo grado. En estas redes la distancia media tiende a $p \sim \log N$, mientras que el clustering generalmente es bajo, ya que tiende a $\frac{k}{N}$, donde k es el grado promedio de los nodos. En redes aleatorias, los vecinos de un determinado nodo se eligen al azar, por lo que no hay correlación entre el grado de nodos vecinos. Por último, estas redes son más robustas a ataques dirigidos, pero a su vez son vulnerables ante errores internos. A continuación, en la Figura 5 se da un ejemplo gráfico del modelo de redes aleatorias.

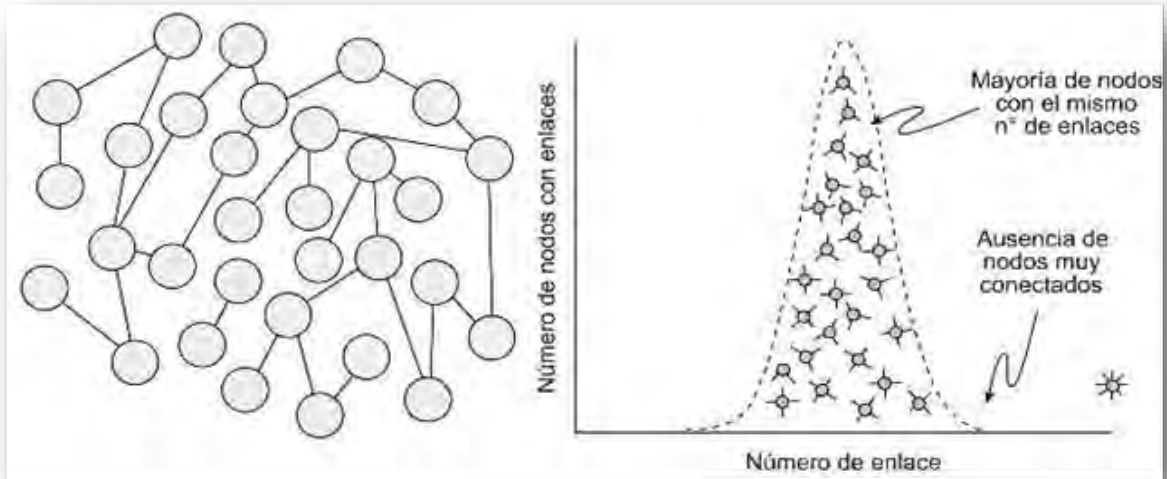
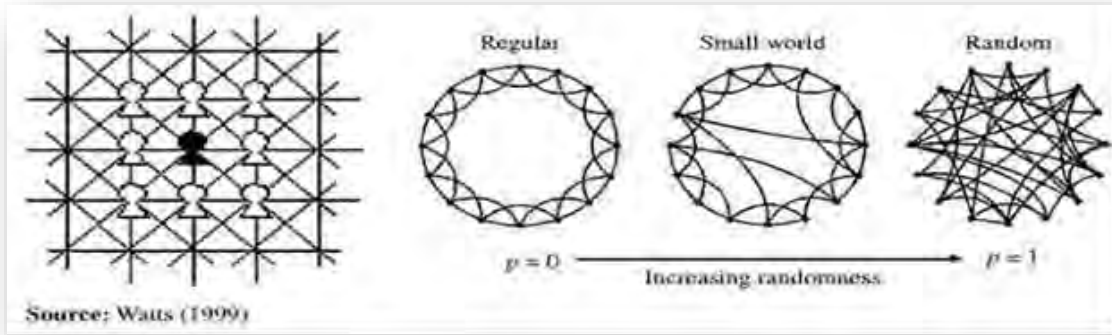


Figura 12 Representación Modelo de Redes Aleatorias (Fuente: <http://www.fotoseimagenes.net/red-libre-de-escala>)

2.4.2 Modelo de Mundo Pequeño (Modelo WS)

El modelo de Mundo Pequeño fue introducido por Watts y Strogatz en 1998 [23], en este tipo de redes la distancia entre dos nodos elegidos aleatoriamente crece proporcionalmente al logaritmo del número de nodos N que contiene la red. El modelo consiste en la interpolación de una red regular y una red aleatoria (ver Figura 6), el ejemplo más sencillo consiste en un anillo con N nodos que se conectan con sus primeros k vecinos, los cuales tienen una misma probabilidad de conexión, en este procedimiento se evita tener multiarcos y rizos o lazos (multiarcos se refiere a dos o más arcos que son incidentes a dos nodos, rizos o lazos son arcos que conectan un nodo consigo mismo). Este modelo ha sido observado en redes, que presentan a nivel global una baja distancia media (su distancia media tiende a $p \sim \log N$), y a nivel local presentan un clustering alto. Además, estas redes siguen una distribución Poisson, es decir, presentan una conectividad media porque la mayoría de los nodos tienen el mismo grado. Las redes de mundo pequeño son muy eficientes con respecto al intercambio de información local y globalmente. Por otro lado, son más robustas a ataques dirigidos, pero a su vez son vulnerables ante errores internos. Este tipo de redes se han observado en redes sociales.



Source: Watts (1999)

Figura 13 Representación Modelo de Redes de Mundo Pequeño, (Fuente: <http://teOriasociocultural.blogspot.mx/2014/03/la-ciencia-de-las-redes-jose-luis-molina.html>)

2.4.3 Modelo de Redes de Libre Escala (Modelo BA)

En 1999, Barabási y Albert [1] plantearon un modelo en el cual buscaban la representación de redes reales, a este modelo se le conoce por el nombre de modelo de redes de libre escala. Este es un modelo del crecimiento de una red, en el cual desarrollaron el concepto de preferencia de enlace, la cual empieza con un conjunto pequeño de nodos (incluso sólo dos o tres nodos) conectados de forma aleatoria. Posteriormente se añaden nuevos nodos con un ritmo constante, en el cual cada uno de estos nuevos nodos lleva un número determinado de nuevos enlaces. La regla para añadir los nuevos nodos es muy simple y es la siguiente, los nodos entrantes prefieren conectarse con los nodos antiguos los cuales ya tienen muchos enlaces (ver Figura 7). En principio los nuevos nodos se pueden conectar a cualquier nodo, pero cuanto mayor sea el grado de un nodo antiguo, mayor es la probabilidad de que atraiga a un nuevo nodo, generando una tendencia a la formación de hubs. Cabe señalar que los nodos menos conectados recibirán nuevos enlaces, pero en la mayoría de los casos los hubs serán mucho más atractivos. Por lo tanto, surge una jerarquía de nodos, la cual va desde los menos conectados hasta los hubs, dando como resultado una red heterogénea, con una distribución de grado la cual corresponde a la ley de potencia, $P(k) \sim k^{-\gamma}$; donde γ (gamma) es el exponente del grado, esto indica que un número pequeño de nodos tiene alto grado de conectividad, en un plano *log-log* se observa que la gráfica tiene una pendiente negativa. Además, debido a la presencia de los hubs, la distancia media de estas redes es ultra pequeña y tiende a $p \sim \log \log N$. Por último, estas redes son más

vulnerables a ataques dirigidos a hubs, pero a su vez son más robustas a errores internos.

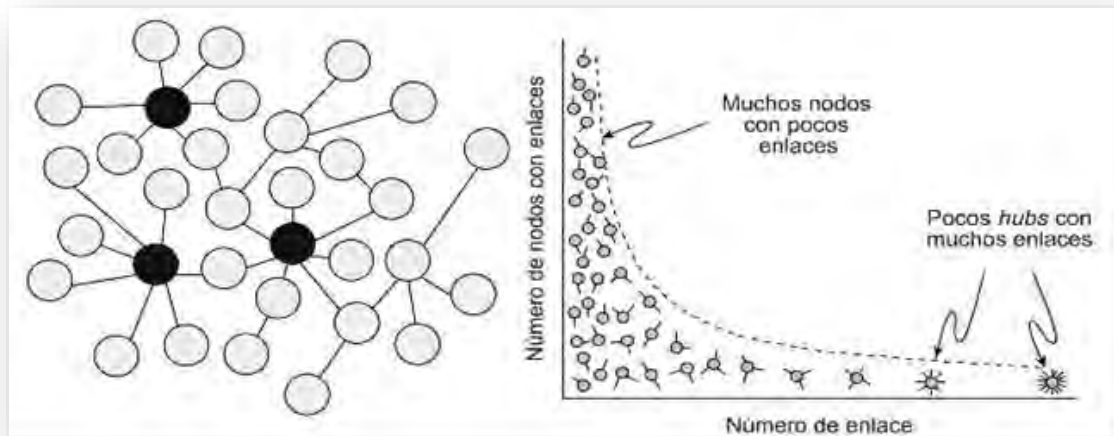
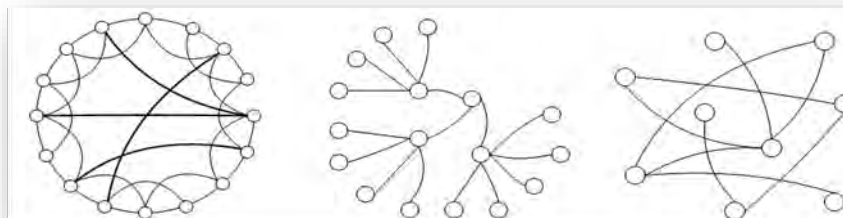


Figura 14 Representación Modelo de Redes de Libre Escala, (Fuente: <http://www.fotoseimagenes.net/red-libre-de-escala>)

Finalmente, en la Figura 15 mostramos una forma sencilla de visualizar los 3 diferentes modelos.



a) Red de Mundo Pequeño b) Red de Libre Escala c) Red Aleatoria

Figura 15 Comparación Modelos de Redes Complejas, (Fuente: <http://jasss.soc.surrey.ac.uk/8/4/8.html>)

Es de suma importancia conocer las definiciones anteriores con el fin de poder estudiar la topología de las redes complejas e identificar la estructura y comportamiento de las redes y con ayuda de los modelos de redes descritos

anteriormente podremos clasificar a las redes en dichos modelos y así poder conocer más a fondo sus características y comportamiento.

2.5 Series de Tiempo

En nuestro día a día nos topamos con series de tiempo sin que realmente nos demos cuenta. La propiedad característica de una serie de tiempo es el hecho de que los datos no se generan de forma independiente, generalmente, hay dependencia entre observaciones adyacentes, es decir, su dispersión varía en el tiempo, a menudo se rige por una tendencia y tienen componentes cíclicos. El análisis de series de tiempo consiste en el uso de técnicas estadísticas para analizar esta dependencia, que requiere el desarrollo de modelos estocásticos y dinámicos para los datos de series de tiempo y su uso en diversas áreas de aplicación como: IPC, PIB, inflación, precio del petróleo, precio de una acción bursátil, precipitación pluvial, niveles diarios de algún contaminante, nivel de exportaciones de un país, número de robos diarios, temperatura de una región, consumo de un producto, nivel de agua de un río o presa, entra e otros. Es importante recalcar que para nuestro estudio utilizamos series de tiempo ya que contamos con los datos del flujo de pasajeros tanto nacionales como internacionales en aeropuertos mexicanos desde enero 2000 hasta marzo 2017. En la Figura 16 se muestran diversos ejemplos de series de tiempo, los datos y las gráficas fueron obtenidos de las bases de datos que trae por default R en las paqueterías TimeSeries, TSA, tseries.

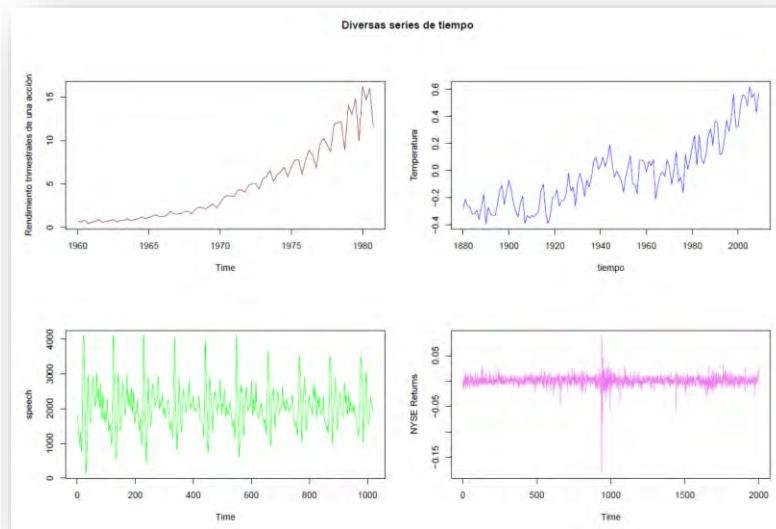


Figura 16 Ejemplos de series de tiempo, (Diseño propio utilizando datos predeterminados en R)

Para este capítulo estaremos utilizando la base de datos AirPassengers que viene por default en R en la paquetería *timeSeries*. La base de datos AirPassengers es conocida como la base de datos de aerolíneas Box & Jenkins, cuenta con datos mensuales de los pasajeros internacionales de 1949 a 1960.

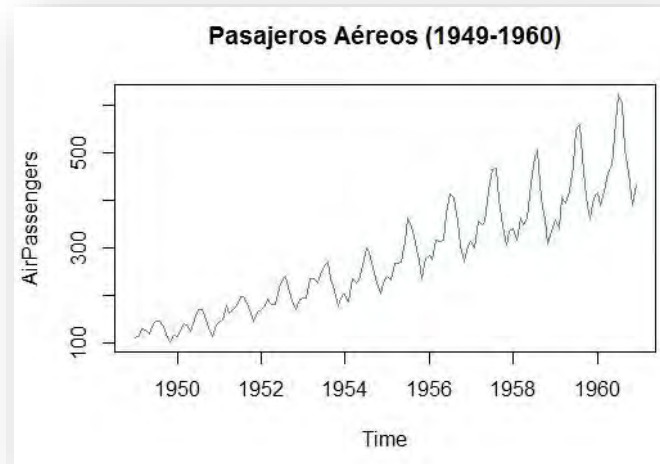


Figura 17 Serie de tiempo pasajeros aéreos (1949-1960), (Diseño propio)

Con lo que podemos definir a una serie de tiempo como una colección de observaciones realizadas de forma secuencial a lo largo de un periodo de tiempo.

Debido a la naturaleza de las observaciones podemos tener dos tipos diferentes de series de tiempo, las series de tiempo discretas y continuas. En las series de tiempo discretas las observaciones se toman en un conjunto de tiempos discretos, pueden ser espaciados. En las series de tiempo continuas las observaciones se hacen de manera continua a lo largo de un intervalo de tiempo.

Una serie de tiempo se considera discreta cuando las observaciones se toman en un conjunto de tiempos discretos, generalmente, igualmente espaciados. El término discreta se usa para series con esta característica, no obstante que la variable medida sea continua. Por otro lado, una serie de tiempo se considera continua cuando las observaciones se hacen de manera continua a lo largo de un intervalo de tiempo. El adjetivo “continua” se usa para este tipo de series, no obstante que la variable medida sólo tome valores en un conjunto discreto.

Los objetivos principales del análisis de series de tiempo son:

- Descripción de los datos
- Interpretación de los datos
- Pronóstico (uno de los objetivos usualmente más importantes)
- Control
- Modelación / Pruebas de hipótesis

Matemáticamente, una serie de tiempo se define de la siguiente manera.

Sea un conjunto de observaciones $X_{t_1}, X_{t_2}, \dots, X_{t_n}, X_{t_i} : i \in T$ realizadas de forma secuencial a lo largo del tiempo. Donde X_{t_i} es el valor de la variable X en el instante i .

La función media de X_t se define como:

$$\mu_x = \mathbb{E}(X_t) \quad (2.7)$$

Mientras que la función de covarianzas de X_t se define como:

$$\gamma_X(r, s) = \text{Cov}_X(X_r, X_t) = \mathbb{E}[(X_r - \mu_r)(X_s - \mu_s)] \quad (2.8)$$

Ahora bien, hablaremos de la estacionariedad, una serie de tiempo $X_t, t = 1, \dots, n$ se dice que es estacionaria si no tiene cambios sistemáticos en la media, es decir, no hay tendencia en la serie, si no tiene cambios sistemáticos en la varianza, es decir, la varianza es constante a lo largo del tiempo y si no hay variaciones periódicas, es decir, no tiene ciclos. Hay dos tipos de estacionariedad, la

estacionariedad débil y estacionariedad estricta:

- Estacionariedad estricta: Una serie de tiempo $X_t, t = 1, \dots, n$ se dice estrictamente estacionaria si la distribución conjunta de X_1, X_2, \dots, X_k es la misma que $X_{1+\tau}, X_{2+\tau}, \dots, X_{k+\tau}$ para todo $k > 0, \tau > 0$. Una manera simple y útil es caracterizar una serie a través de sus momentos, en particular de los primeros dos.
- Estacionariedad débil: Una serie de tiempo $X_t, t = 1, \dots, n$ se dice débilmente estacionaria o estacionaria de segundo orden si:
 - $\mu_x(t)$ es independiente de t
 - $\gamma_X(\tau) = \gamma_X(t + \tau, t)$ es independiente de t

Esto significa que la media y la varianza son constantes y la función de autocovarianza solo depende del retraso (lag) de la serie.

Los valores de la función de autocovarianza dependen de las unidades de X_t , por lo que lo más conveniente es considerar la función de autocorrelación que se define como:

$$\rho_X(\tau) = \frac{\gamma_X(\tau)}{\gamma_X(0)} \quad (2.9)$$

Donde:

$$\gamma_X(0) = \sigma_X^2 \text{ (la varianza de la serie)}$$

La función de autocorrelación cumple que:

- $\rho_X(0) = 1$

- $\rho_X(\tau) = \rho_X(-\tau)$
- $|\rho_X(\tau)| \leq 1$

La interpretación usual del conjunto de coeficientes de autocorrelación se hace a través del correlograma, en el cual se grafican los coeficientes estimados de ρ_k contra su correspondiente retraso k con $k = 1, 2, \dots, M$ con M menor que n . Para $\rho_k(0) = 1$ ya que la correlación de la serie con ella misma es uno. Suele ser muy útil realizar una inspección visual del correlograma, ya que se puede apreciar el grado de asociación que guarda una observación con sus adyacentes, además de su orden (cuantas observaciones atrás se tiene una correlación importante). También nos ayuda a ver si la serie tiene ciclos. Por lo que para determinar cuántos coeficientes de correlación son importantes es necesario realizar las pruebas de hipótesis correspondientes:

$$\mathbb{H}_0: \rho_k = 0 \quad vs \quad \mathbb{H}_a: \rho_k \neq 0 \quad (2.10)$$

La forma correcta de saber que correlaciones son estadísticamente significativas es graficar las bandas de confianza del correlograma. Las correlaciones que se salen de las bandas de confianza se consideran significativamente distintas de cero. (ACF por sus siglas en inglés Auto Correlation Function).

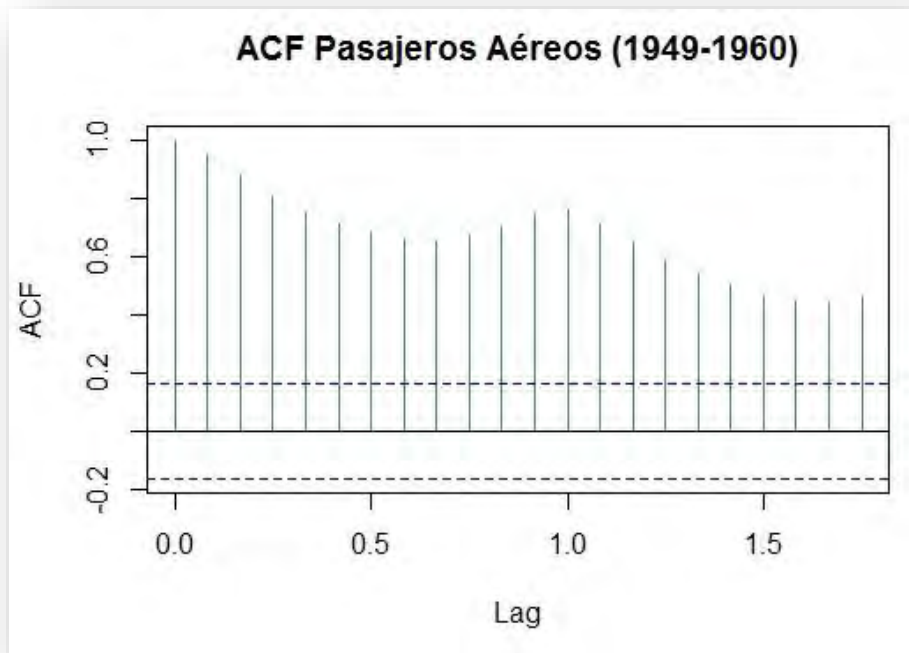


Figura 18 Gráfica función de auto correlación (Diseño propio)

En muchas ocasiones la correlación entre dos variables se debe a que están correlacionadas con una tercera, y no necesariamente entre sí. La correlación parcial es una medida de asociación que remueve la correlación que puede existir entre las variables de interés con otras variables coleccionadas, esto con el fin de mostrar solo la correlación entre estas variables, sin considerar las que tienen con el resto del conjunto. El correlograma con correlaciones parciales es útil para determinar el orden de la serie de tiempo considerada.

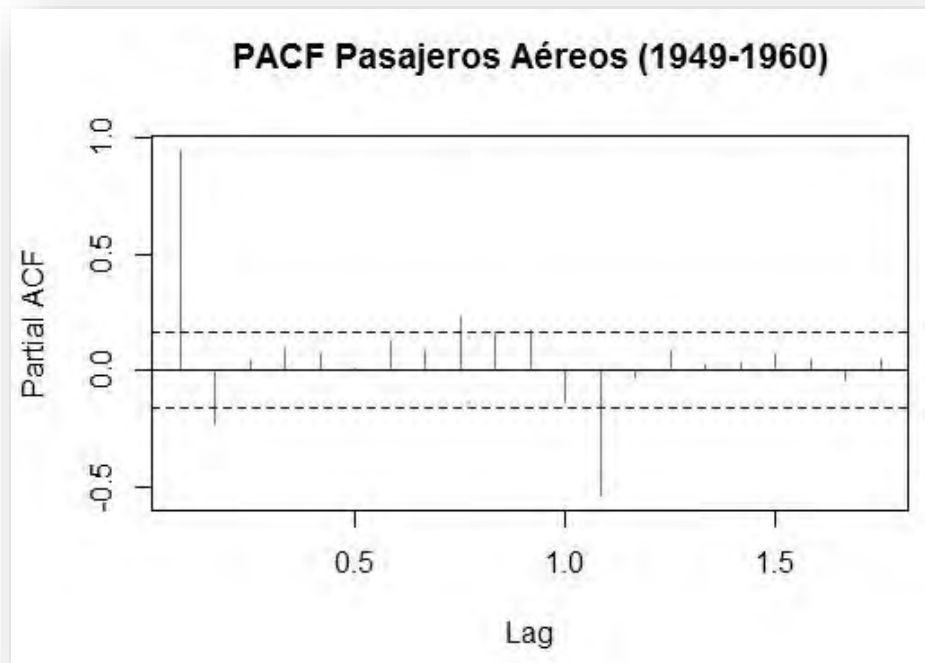


Figura 19 Gráfica función de autocorrelación parcial (Diseño propio)

Una serie de tiempo está compuesta por cuatro componentes principales, los cuales contribuyen a los cambios observados en un periodo de tiempo y es lo que le da a la serie su aspecto errático. Además, existe una relación multiplicativa entre estas componentes, es decir, cualquier valor de una serie es el producto de factores que se pueden atribuir a los componentes.

Los cuatro componentes de la serie son:

- Tendencia (T_t)
- Ciclo (ciclos económicos, recesión, expansión \- ciclos largos) (C_t)
- Estacionalidad (ciclos cortos) (E_t)
- Aleatoriedad (A_t)

La tendencia a largo plazo de una serie de tiempo caracteriza el patrón gradual y consistente de las variaciones de la propia serie, las cuales se consideran

consecuencias de fuerzas persistentes que afectan ya sea el crecimiento o la reducción de la misma, tales como: cambios en la población, cambios en los ingresos, cambio en el precio de las acciones, entre otros. Las tendencias a largo plazo se ajustan a diversos esquemas. Algunas se mueven continuamente hacia arriba, otras declinan, y otras más permanecen igual en un cierto período o intervalo de tiempo.

Las series de tiempo con frecuencia presentan ciclos que son secuencias alternas de puntos abajo y arriba de la línea de tendencia que duran más de un año, esta variación se mantiene después de que se han eliminado las variaciones o tendencias estacional e irregular. Un ejemplo de este tipo de variación son los ciclos comerciales cuyos períodos recurrentes dependen de la prosperidad, recesión, depresión y recuperación, las cuales no dependen de factores como el clima o las costumbres sociales.

La estacionalidad es el componente de la serie de tiempo que representa la variabilidad en los datos debida a influencias de las estaciones. Esta variación corresponde a los movimientos de la serie que recurren año tras año en los mismos meses del año más o menos con la misma intensidad. Por ejemplo: Un fabricante de albercas inflables espera poca actividad de ventas durante los meses de otoño e invierno y tiene ventas máximas en los de primavera y verano, mientras que los fabricantes de equipo para la nieve y ropa de abrigo esperan un comportamiento anual opuesto al del fabricante de albercas. Este componente también depende del clima, costumbres sociales y región.

La aleatoriedad también denominada residual o ruido blanco, se debe a factores a corto plazo, imprevisibles y no recurrentes que afectan a la serie de tiempo. Como este componente explica la variabilidad aleatoria de la serie, es impredecible, es decir, no se puede esperar predecir su impacto sobre la serie de tiempo. Usualmente se hace la suposición que el componente aleatorio tiene media cero y varianza constante. Existen dos tipos de variación irregular:

- Las variaciones que son provocadas por acontecimientos especiales, fácilmente identificables, como las elecciones, inundaciones, huelgas, terremotos.
- Variaciones aleatorias o por casualidad, cuyas causas no se pueden señalar en forma exacta, pero que tienden a equilibrarse a la larga.

Dado que los componentes de una serie no se observan aisladamente, se necesita aplicar hipótesis que representen el proceso generador de los datos y nos ayudan a estimar la parte estacional:

- Aditivo: $Y_t = T_t + C_t + E_t + A_t$
- Multiplicativo: $Y_t = T_t * C_t * E_t * A_t$
- Mixto: $Y_t = T_t * C_t * E_t + A_t$

Los residuos son:

$$Y_t = \frac{E_t}{A_t} \quad (2.11)$$

Donde:

Y_t = serie observada en el instante t

T_t = componente de tendencia

C_t = componente de ciclo

E_t = componente estacional

A_t = componente aleatorio

El modelo aditivo asume que los componentes de la serie son independientes, es decir, la amplitud de la estacionalidad independiente del nivel de la tendencia o ciclo, un aumento en el nivel de la tendencia o ciclo no ocasiona un aumento en la amplitud estacional. Este modelo se puede aplicar en todo momento, incluso cuando la serie presenta valores negativos o ceros.

El modelo multiplicativo asume que los componentes están interrelacionados, es decir, un aumento en el nivel de la tendencia o ciclo ocasiona un aumento en la amplitud estacional. Los componentes estacional y aleatoria están expresados en porcentajes. La mayoría de las series de tiempo económicas siguen un modelo

multiplicativo, no se puede aplicar si la serie presenta valores negativos o ceros. Para saber si la serie sigue un modelo aditivo o multiplicativo, se debe analizar la amplitud del ciclo, si este aumenta a medida que lo hace la tendencia, el modelo es multiplicativo, si permanece constante es aditivo. En las Figuras 20 y 21 se muestra el ejemplo de las descomposiciones aditiva y multiplicativa de los datos que hemos estado utilizando como ejemplo.

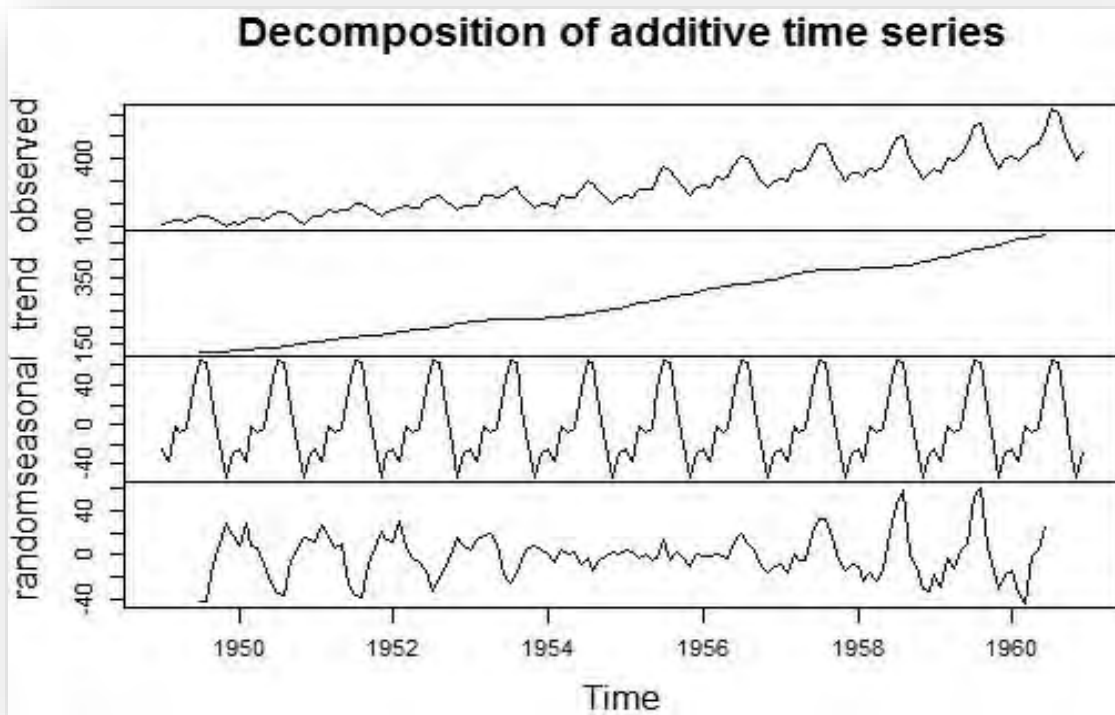


Figura 20 Descomposición aditiva de series de tiempo, (Diseño propio)

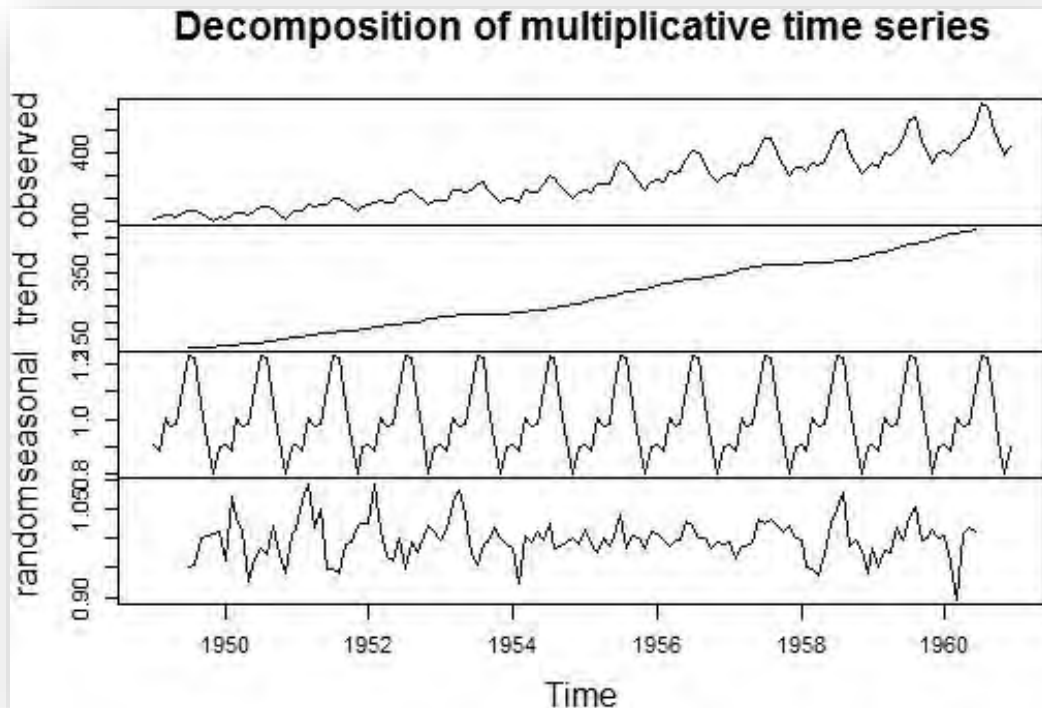


Figura 21 Descomposición multiplicativa de series de tiempo (Diseño propio)

Existen varios métodos para estimar la tendencia (T_t), los cuales consisten en:

- Ajustar una función del tiempo, como un polinomio, una exponencial u otra función suave de t .
- Suavizar (o filtrar) los valores de la serie.
- Utilizar diferencias.

A continuación, podemos observar diferentes tipos de suavizamiento o filtros que le podemos aplicar a nuestras series de tiempo.

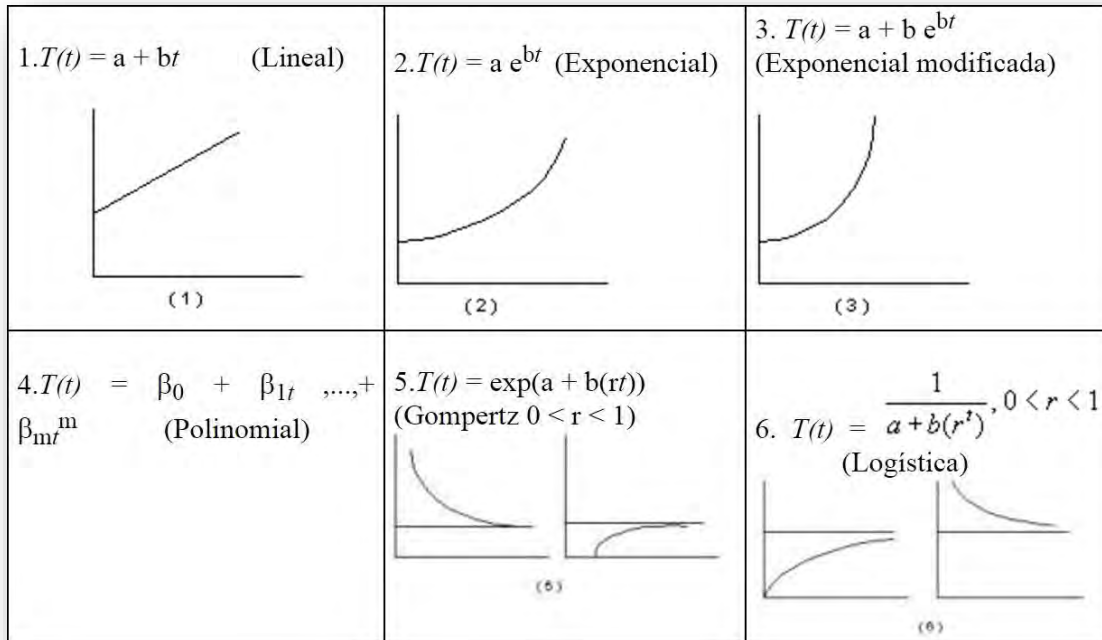


Figura 22 Tipos de suavizamiento, (Fuente: <https://www.gestiopolis.com/automatizando-operaciones-minitab-series-tiempo/>)

Es importante señalar que la curva de tendencia debe cubrir un periodo relativamente largo para ser una buena representación de la tendencia a largo plazo. La tendencia rectilínea y exponencial son aplicables a corto plazo, puesto que una curva S a largo plazo puede parecer una recta en un período restringido de tiempo.

Una forma de visualizar la tendencia es mediante el suavizamiento de la serie. La idea central es definir a partir de la serie observada una nueva serie que suaviza los efectos ajenos a la tendencia (estacionalidad, efectos aleatorios o de ruido blanco), de manera que podamos determinar la dirección de la tendencia.

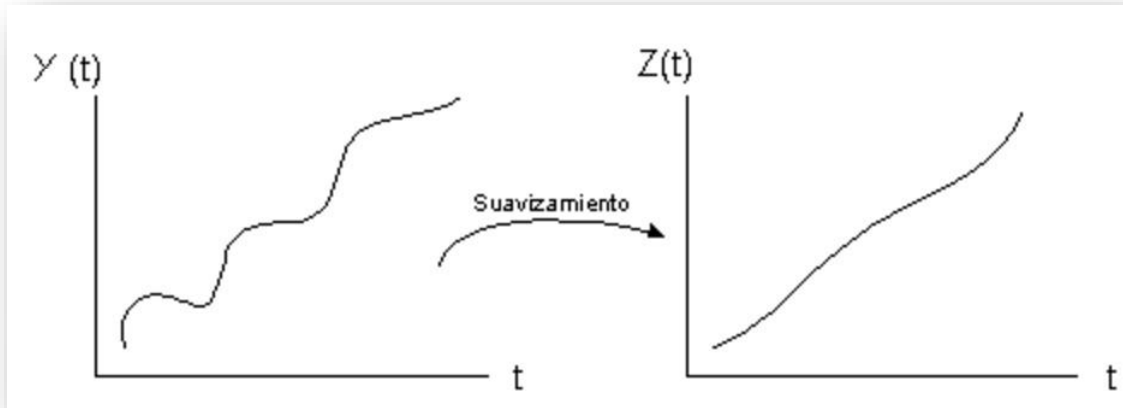


Figura 23 Ejemplo suavizamiento (Fuente: <https://www.gestiopolis.com/automatizando-operaciones-minitab-series-tiempo/>)

Lo que hacemos es usar una expresión lineal que transforma la serie Y_t en una serie suavizada $Z(t)$: $Z(t) = F(X(t))$, $t = 1, \dots, n$ de tal modo que $F(X(t)) = T_t$. La función F se denomina Filtro Lineal. El filtro lineal más usado es el promedio móvil. El objetivo principal de los promedios móviles es eliminar de la serie los componentes estacionales y aleatorios. La serie se suaviza cuando existen muchos cambios bruscos o movimientos irregulares.

Por otro lado, es importante definir algunos operadores aplicables a las series de tiempo:

- Operador de retraso: definimos el operador de retraso B como:

$$B^j X_t = X_{t-j} \quad \forall j \in \mathbb{Z} \quad (2.12)$$

- Operador diferencia: definimos el operador diferencia ∇ como:

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t \quad (2.13)$$

2.5.1 Modelos de Series de Tiempo

En el análisis de series de tiempo se utilizan diferentes modelos que nos ayudan a identificar como es nuestra serie y a pronosticar los posibles valores que puede tomar nuestra serie en el futuro. Para eso existen diferentes modelos:

- AR(p)- Modelo Autorregresivo
- MA(q)- Modelo Medias móviles
- ARMA (p, q)- Modelo Autorregresivo con Medias móviles
- ARIMA (p, q)- Modelo Autorregresivo Integrado de Medias móviles

Modelo Medias Móviles (MA(q))

Sea Z_t un proceso puramente aleatorio con media cero y varianza σ^2 . Un proceso X_t se dice que es un proceso de medias móviles de orden q (MA(q)), si:

$$X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \dots + \beta_q Z_{t-q} \quad (2.14)$$

Con β_i constantes. Las Z 's usualmente se escalan para tener $\beta_0 = 1$.

Modelo Autorregresivo (AR(p))

Sea Z_t un proceso puramente aleatorio con media cero y varianza σ^2 . Un proceso X_t se dice que es un proceso autorregresivo de orden p (AR(p)), si:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + Z_t \quad (2.15)$$

Similar al modelo de regresión lineal múltiple, solo que X_t tiene como regresores valores anteriores de ella misma y no variables de otra naturaleza. El proceso AR (1) también es conocido como proceso de Markov.

Modelo Autorregresivo con Medias Móviles (ARMA (p, q))

Los modelos AR(p) son una clase general de procesos estocásticos que son prácticamente suficientes para las aplicaciones reales. No obstante, una forma de ampliar esta clase de modelos para series de tiempo es incluir un proceso de medias móviles MA(q) para generar un proceso mixto conocido como proceso autorregresivo de medias móviles ARMA (p, q). Entonces, se dice que X_t es un proceso autorregresivo de medias móviles ARMA (p, q) si

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q} \quad (2.16)$$

Modelo Autorregresivo Integrado de Medias Móviles (ARIMA (p, d, q))

Muchas series presentan dos componentes, un componente de tendencia no estacionario y otro componente de media cero estacionarios. El modelo ARMA integrado mejor conocido como ARIMA permite ampliar la clase de los modelos ARMA para incluir diferenciación.

Se dice que el proceso X_t es ARIMA (p, d, q) si:

$$\nabla^d X_t = (1 - B)^d X_t \quad (2.17)$$

Es un modelo ARMA (p, q). Escribiremos de forma general estos modelos como:

$$\phi(B)(1 - B)^d X_t = \theta(B)Z_t \quad (2.18)$$

Si $d = 0$ tenemos un proceso ARMA (p, q). Los modelos ARIMA son apropiados para modelar datos con tendencia. Estos modelos pueden extenderse para incluir términos estacionales (ciclos), proporcionando un ARIMA no estacionario estacional, conocido como SARIMA. Un modelo ARIMA estacional utiliza diferencias iguales al periodo del retraso (ciclo) que se quiere remover. Una serie puede ser no estacionaria también por cambios en la varianza, conocida como

heteroscedasticidad, que usualmente da como resultado periodos de volatilidad, e los que hay un claro cambio en la varianza de la serie.

Tabla 6 Resumen modelos

Proceso	FAC	FACP	FAS	FAP
AR(p)	Convergencia a cero, con comportamiento dictado por la ecuación $\Phi(B)\rho_k = 0$, para $k \geq p$.	Solamente las primeras p autocorrelaciones parciales son distintas de cero.	Muchos coeficientes no nulos.	Primeros p no nulos, resto igual a cero.
MA(q)	Solo las primeras q autocorrelaciones son distintas de cero.	Sucesión infinita convergente a cero.	Primeros p no nulos, resto igual a cero.	Muchos coeficientes no nulos.
ARMA(p, q)	Comportamiento irregular de las primeras q autocorrelaciones y después convergencia a cero de acuerdo a $\Phi(B)\rho_k = 0$, para $k > q$.	Sucesión infinita convergente a cero.	Muchos coeficientes no nulos.	Muchos coeficientes no nulos.

Especificación del modelo para una serie de tiempo

Ahora debemos estudiar cómo implementar estos modelos, además de la manera de realizar inferencias con ellos, por eso nuestros objetivos serán:

- Elegir de manera apropiada los valores (p, d, q) para los datos de una serie real.
- Estimar los parámetros del modelo ARIMA (p, d, q) especificado en el punto anterior.
- Verificar qué tan apropiado es el ajuste del modelo, así como sus supuestos.

La estrategia general para lograr estos objetivos es iniciar con la elección razonable, aunque tentativa, de los valores de (p, d, q) . Una vez realizado este paso deberemos estimar de la forma más eficiente posible, los parámetros ϕ, θ y σ_z^2 , involucrados en el modelo. Finalmente, debemos juzgar de manera crítica el ajuste del modelo que obtuvimos para verificar lo adecuado del modelo. Si el modelo no parece adecuado, debemos considerar en qué sentido no es adecuado para que nos ayude a proponer otro modelo. Entonces procederemos de forma semejante con el nuevo modelo hasta obtener un modelo que ajuste adecuadamente a nuestros datos y cumpla con los supuestos sobre los que se construyó.

Identificación

Supondremos que la serie de tiempo ya es estacionaria, ya sea por origen por transformaciones o por diferenciación para quitar tendencia o estacionalidad (ciclos). Estos procesos se hacen de un análisis de la gráfica de la serie. Dada la serie de tiempo estacionaria el siguiente paso consiste en asociar la función de autocorrelación (FAC o ACF) muestral con un posible modelo ARIMA. Por lo que para determinar cuántas autocorrelaciones pueden considerarse estadísticamente significativas a un nivel de significancia α . Barlett (1946) desarrolló aproximaciones para las varianzas y covarianzas de las autocorrelaciones muestrales, en caso de que el proceso sea generado a partir de ruido blanco con distribución normal.

Ahora hay que sustituir los valores desconocidos de ρ_k por sus correspondientes valores estimados de r_k que proporciona una buena estimación para muestras grandes ($n > 50$).

Para probar si algunas de las correlaciones son iguales a cero a partir de cierto retraso q es necesario hacer la prueba de hipótesis correspondiente.

$$\mathbb{H}_0: \rho_k = 0 \quad vs \quad \mathbb{H}_a: \rho_k \neq 0 \quad (2.19)$$

Generalmente cuando se grafica el correlograma, los paquetes grafican las bandas de confianza (intervalos de confianza para ρ_k) con lo que se puede juzgar esta hipótesis. Si la correlación es estadísticamente significativa al nivel de confianza seleccionado, entonces la barra que la representa en el correlograma rebasará cualquiera de las líneas que delimitan el intervalo.

En general, el orden de un $AR(p)$ no es posible detectarlo únicamente con el uso de la función de autocorrelación muestral. En este caso es necesario el uso de la función de autocorrelación parcial para hacer la identificación de un modelo $AR(p)$. Entonces para decidir el orden del proceso AR subyacente, debemos verificar si una correlación parcial, equivalentemente, si un coeficiente del modelo AR es estadísticamente significativo. Quenouille (1949) desarrolló una aproximación para la media y varianza de las autocorrelaciones muestrales.

Finalmente, es importante decir que mientras un proceso $AR(p)$ tiene solo las primeras p autocorrelaciones parciales distintas de cero, un proceso $MA(q)$ (que es equivalente a un proceso $AR(1)$), tendrá todas las autocorrelaciones parciales distintas de cero, aunque la función de autocorrelación parcial muestre convergencia a cero. De forma similar, un proceso $ARMA(p, q)$ tendrá asociada una función de autocorrelación parcial que no desaparece después de un número finito de retrasos. La idea es que se proponga un modelo sencillo (criterio de parsimonia “pocos parámetros”).

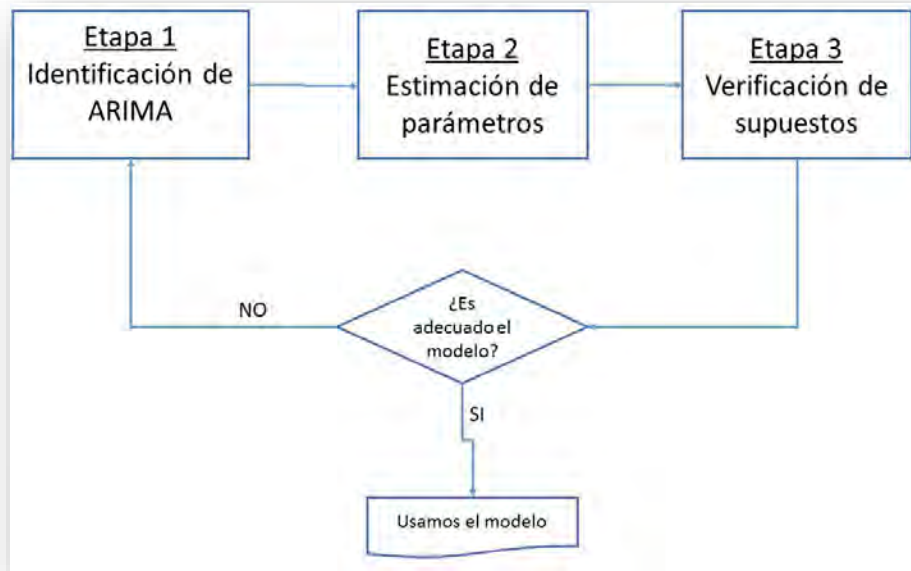


Figura 24 Representación etapas modelación de series de tiempo, (Diseño propio)

Estimación

En esta etapa, se asume que ya se ha identificado un modelo razonable para nuestra serie de tiempo y que lo que resta es encontrar los valores de los parámetros que lo determinan. Una vez propuesto los órdenes de los modelos autorregresivos y de promedios móviles (p, q) , además del grado de diferenciación requerido (d) , postulamos el modelo:

$$\phi(B)\nabla^d T(X_t) = \theta(B)Z_t \quad (2.20)$$

Por lo que debemos de estimar los valores de $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ a partir de los datos de la serie de tiempo bajo estudio.

En otras palabras, una vez elegido el modelo, se estiman los parámetros y se

verifican si cada parámetro es significativo. Un parámetro será significativo, si los siguientes intervalos no contienen al cero. Para el caso donde si lo contiene nos indicará que el cero es un valor posible del parámetro.

Verificación

La etapa de verificación parte de la idea de que todos los modelos son erróneos, puesto que son sólo representaciones simplificadas de la realidad, entonces, si hay que elegir entre varios modelos, habrá que elegir aquel que presente menos fallas, o bien, fallas menos importantes. Por tal motivo, habrá que poner en tela de juicio a los posibles modelos para detectar sus fallas, que se evidencian como violaciones a los supuestos sobre los que se construyeron.

El análisis de residuos es una de las formas más simples para detectar violaciones a los supuestos de los modelos, en donde, como residuos se entiende aquella parte de las observaciones que no es explicada por el modelo. Los residuos para el modelo ARMA(p, q) ($\phi(B)W_t = \theta(B)Z_t$) se definen como:

$$\hat{Z}_t = \hat{\theta}(B)^{-1}\hat{\phi}(B)W_t = \hat{\psi}W_t \quad (2.21)$$

Por lo que los residuos miden la discrepancia entre los valores observados de la serie y sus valores estimados, buscamos evidencia que nos permita rechazar la hipótesis en las que nos hemos basado para estimarlo. Queremos:

- Su media muestral no es significativamente distinta de cero o es cero.
- Su varianza constante.
- Si tiene distribución normal.
- Sus elementos son mutuamente independientes (errores no auto correlacionados).
- Detección de datos atípicos.

Otra forma de verificar el modelo es mediante una prueba de bondad de ajuste, usualmente utilizamos el criterio de AIC, el criterio de Akaike (1974) nos dice a grandes rasgos que si tenemos la distribución verdadera de una variable y una familia paramétrica de modelos donde representa un vector de parámetros la idea es escoger aquel modelo que minimiza la cantidad de información requerida. Por lo que el criterio consiste en tener un AIC menor.

Pronósticos con series de tiempo.

Una vez que hemos propuesto, ajustado y verificado un modelo, la siguiente etapa consiste en hacer uso del modelo, para lograr los fines para lo que construido. Uno de los objetivos más comunes es el de realizar pronósticos, que no es otra cosa más que estimar los valores W_{t+h} con $h = 1, 2, \dots$ a partir de los valores observados $W = \{W_1, W_2, \dots, W_n\}$, debemos de hacer un pronóstico que sea óptimo. El criterio que usaremos para lograrlo será el del error cuadrático medio mínimo, es decir, si denotamos $\widehat{W}_t(h)$ como el valor pronosticado por el modelo para W_{t+h} , entonces se deberá cumplir la condición de:

$$\mathbb{E}[W_{t+h} - \widehat{W}_t(h)]^2 = \min_{\widehat{W}_t(h)} \mathbb{E}[W_{t+h} - \widehat{W}_t(h)]^2 \quad (2.22)$$

Donde \mathbb{E} se denota como la esperanza condicional, dada toda la información (valores de la serie) hasta el momento t , es decir

$$\mathbb{E}[W_{t+h} - \widehat{W}_t(h)]^2 = \mathbb{E}\{[W_{t+h} - \widehat{W}_t(h)]^2 | X_t, X_{t-1}, \dots\} \quad (2.23)$$

Otra forma de plantear el error cuadrático medio; supongamos que y_t es un proceso estocástico estacionario con $\mathbb{E}(y_t) = 0$, por lo que estamos interesados en predecir los valores de este proceso para varios periodos en el futuro, basándonos en su historia observada, esta historia está contenida en el llamado conjunto de información o conjunto de datos. En la práctica este conjunto siempre es finito $\{y_t, y_{t-1}, \dots, y_{t-p}\}$ que representa el pasado reciente del proceso, pero en el desarrollo de la teoría de predicción es usual considerarlo infinito $\{y_t, y_{t-1}, \dots, y_{t-p}, \dots\}$ representando el pasado completo.

Denotemos por \hat{y}_{t+h} a la predicción de y_{t+h} que implica que se hace el pronóstico h periodos adelante. El criterio usual para realizar la predicción \hat{y} , de una variable aleatoria X , es su error cuadrático medio definido como:

$$\mathbb{E}\{(y - \hat{y})^2\} \tag{2.24}$$

Si toda la información disponible sobre y está contenida en su distribución marginal, entonces el error cuadrático medio mínimo de predicción es simplemente el valor esperado de y , $\mathbb{E}(y)$. Sin embargo, si está estadísticamente relacionado con otra variable aleatoria x , cuyo valor pueda observarse, y si se conoce la función de distribución conjunta de x y y , entonces este error cuadrático medio mínimo de predicción de y es la esperanza condicional $\hat{y} = \mathbb{E}(y|x)$.

En el caso de un predictor óptimo que combina valores anteriores de la serie, se sigue por el principio de ortogonalidad, que los errores de predicción no están correlacionados con las predicciones previas. En ejemplo resultante de este tipo utilizada por economistas es la llamada hipótesis de eficiencia del mercado. Un mercado es eficiente si los precios de los activos negociados constituyen previsiones óptimas de sus rendimientos futuros descontados, que consisten en pagos de intereses y dividendos y las ganancias de capital.

De acuerdo con la hipótesis, los cambios en los precios de los activos estarán correlacionados con los anteriores o actuales niveles de precios, lo que nos dice que los precios de los activos siguen caminatas aleatorias. Por otro lado, no deberá ser posible que alguien valorando solo la historia pasada de precios de los activos pueda cosechar una ganancia especulativa de manera sistemática y regular.

Error de pronóstico

Ya que se ha encontrado la forma de calcular de forma óptima la predicción de la serie de tiempo, es necesario calcular su error de pronóstico y ver algunas de sus características. El error de pronóstico está dado por:

$$e_t(h) = W_{t+h} - \widehat{W}_{t+h} = - \sum_{j=0}^{h-1} \psi_j Z_{t+h-j} \quad (2.25)$$

De aquí se deduce que al hacer un pronóstico óptimo mientras más alejado del tiempo actual t , se desee hacerlo, para h grande, mayor será la varianza del mismo y menor su precisión.

Capítulo 3 Análisis de Resultados

En este capítulo se plasmará la parte más fuerte de esta tesis que es la aplicación de metodología seguida para el análisis de los datos obtenidos, con el fin de obtener resultados y, posteriormente hacer un análisis de dichos resultados. Se utilizaron todas las herramientas vistas anteriormente que son series de tiempo, el algoritmo de visibilidad y análisis de redes complejas, todo esto se llevó acabo con el software R y RStudio como ya se había comentado en el capítulo2.

En esta sección únicamente mostraremos las gráficas más importantes, que, en este caso, decidimos que serán los aeropuertos de Acapulco y Monterrey. Por lo que, este primer acercamiento a los datos se puede representar en forma de serie de tiempo, proponemos dividir las series en nacionales, internacionales y totales como se mostrará a continuación:

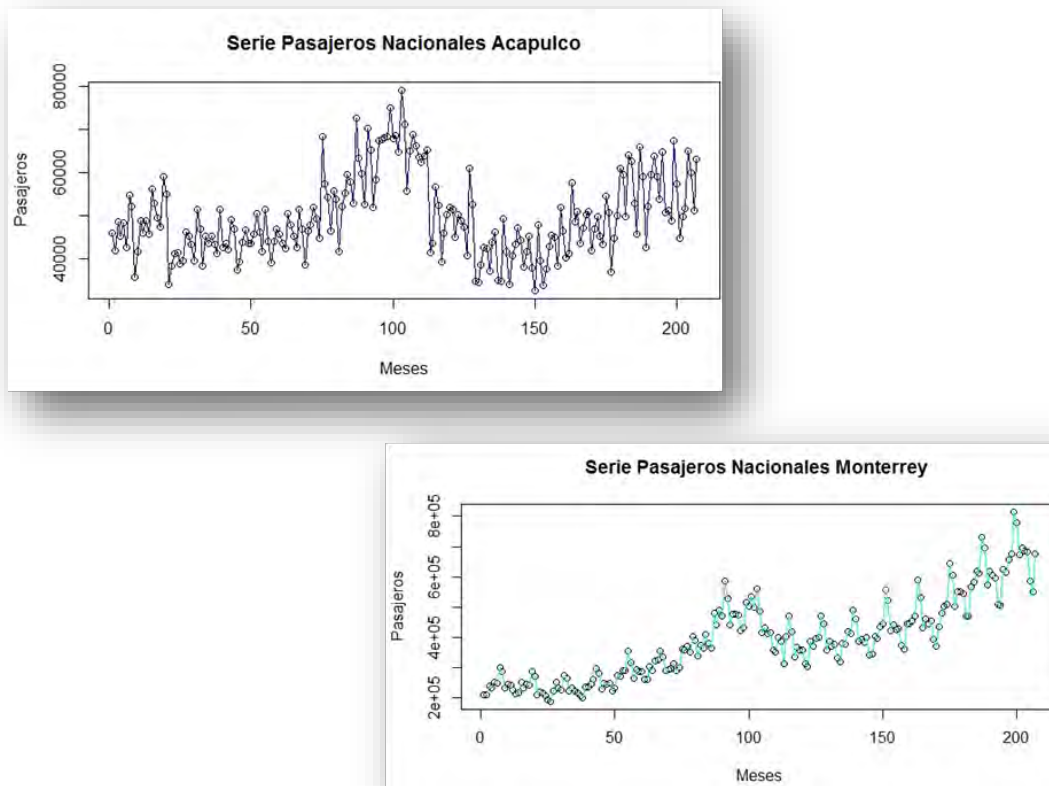


Figura 25 Ejemplos de series de tiempo Pasajeros Nacionales de Acapulco y Monterrey

Se realiza este mismo proceso tanto para las series de tiempo de pasajeros internacionales como totales y obtenemos series de tiempo de la siguiente forma:

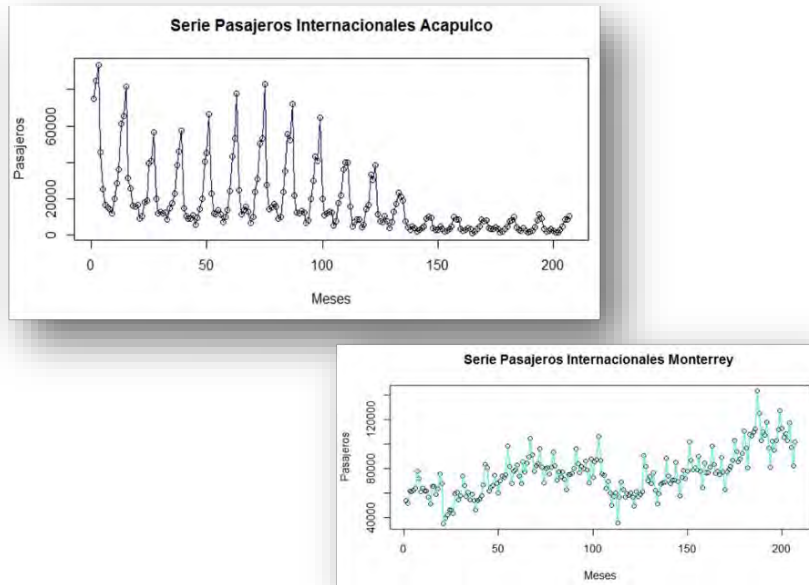


Figura 26 Ejemplos de series de tiempo Pasajeros Internacionales de Acapulco y Monterrey

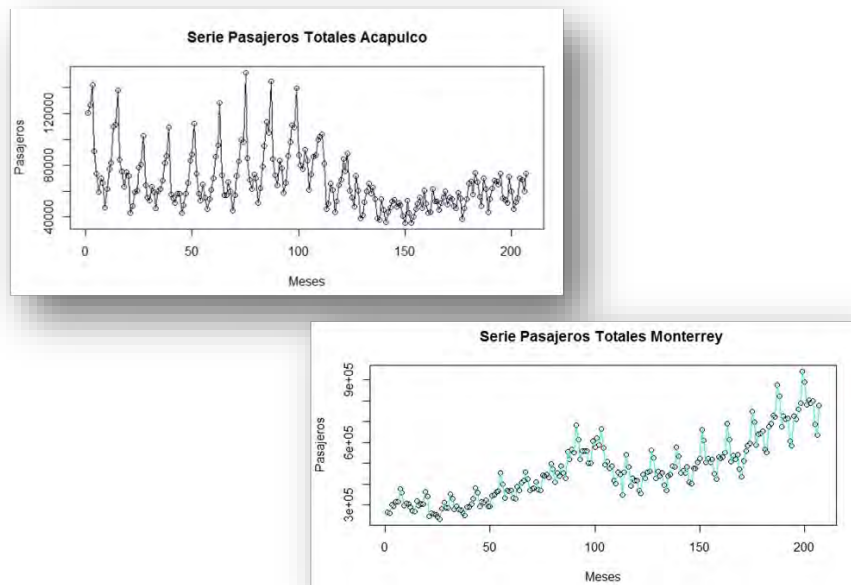


Figura 27 Ejemplos de series de tiempo Pasajeros Totales de Acapulco y Monterrey

Una vez que obtuvimos las series de tiempo correspondientes procedemos a aplicar el algoritmo de visibilidad, para que podamos obtener las redes complejas que posteriormente estudiaremos.

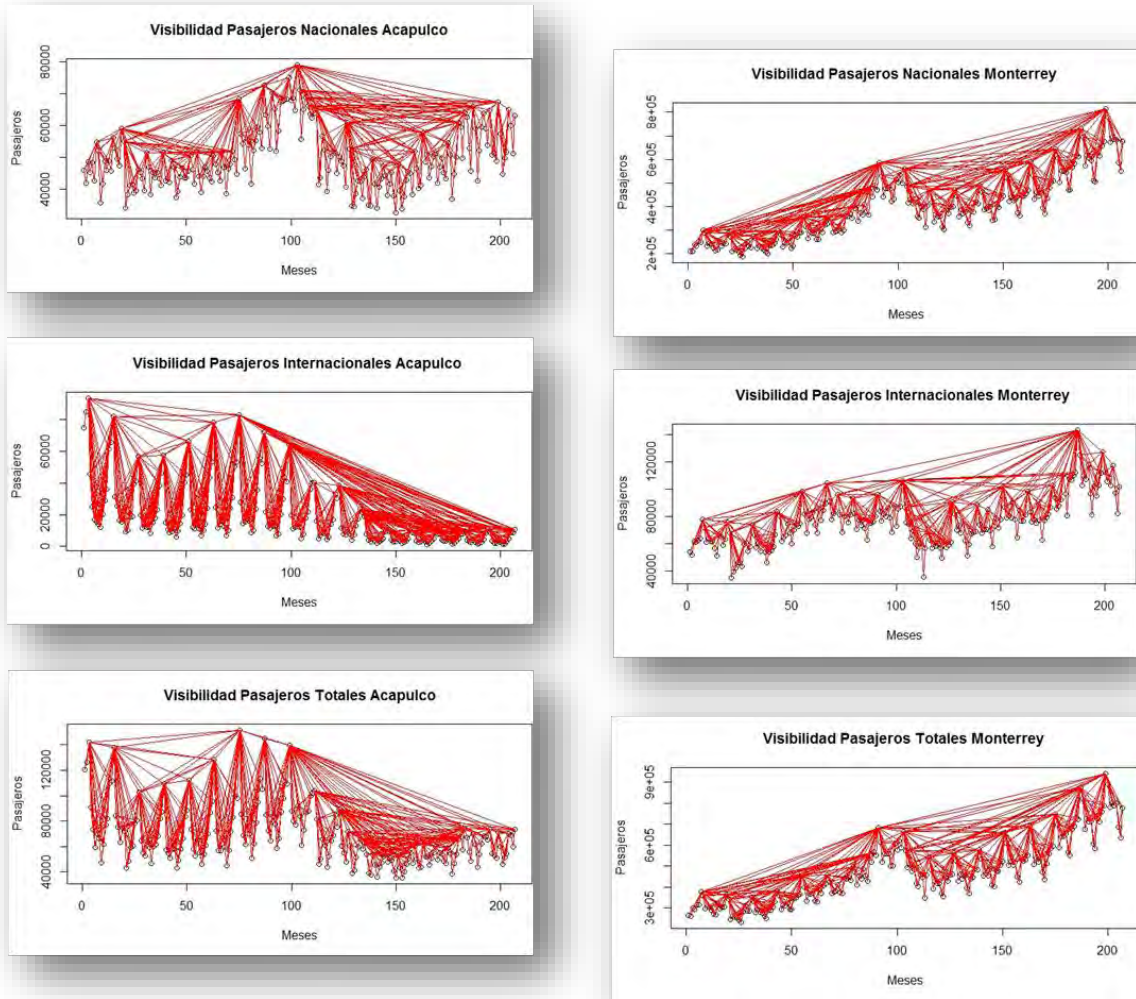


Figura 28 Ejemplos del algoritmo de visibilidad en aeropuertos de Acapulco y Monterrey

Una vez aplicado el algoritmo de visibilidad a las series de tiempo, simplemente realizamos la extracción de la red obtenida con dicho algoritmo, está en solo una forma visualización ya que los nodos se pueden acomodar como se requieran o se necesiten.

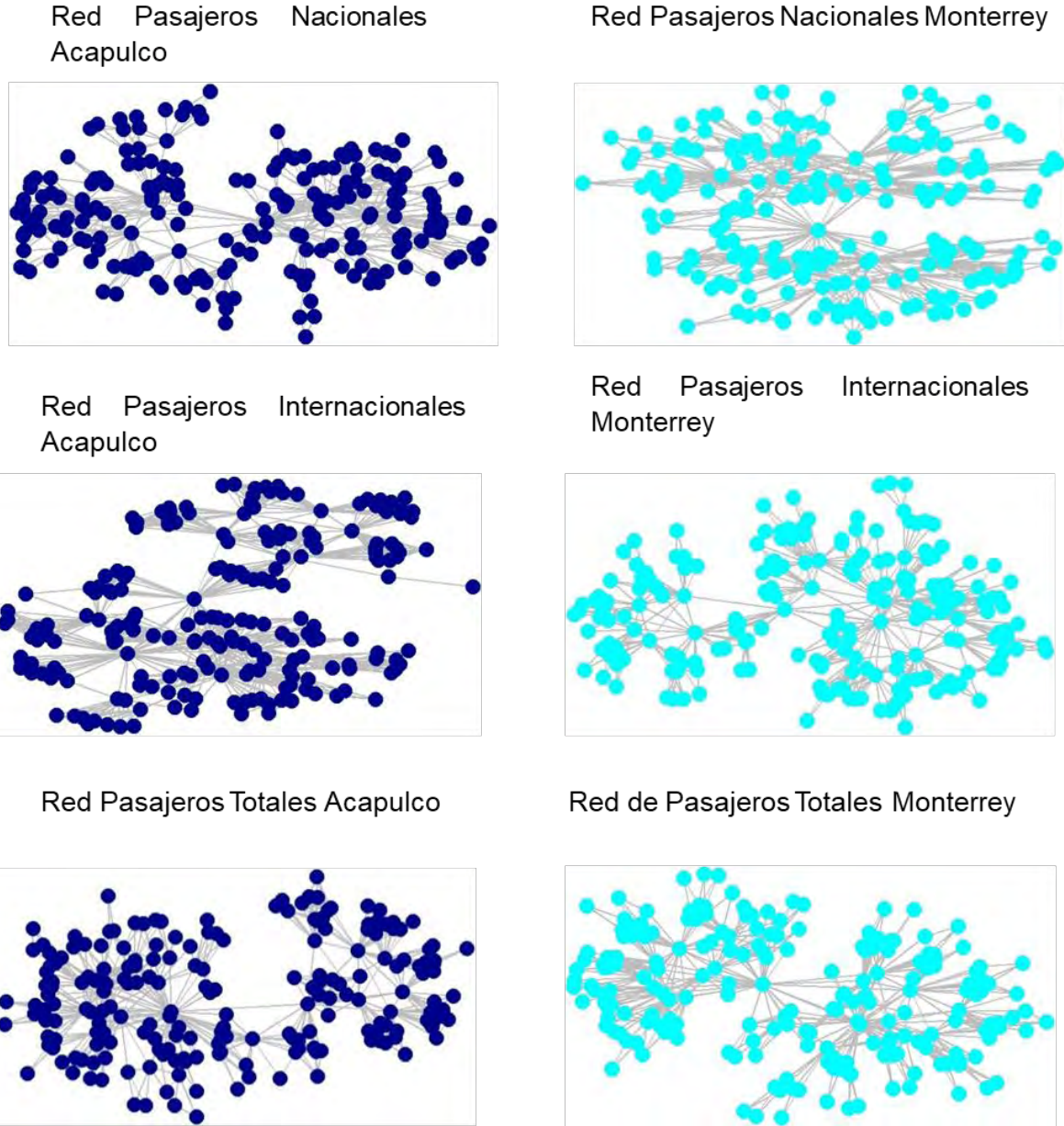


Figura 29 Ejemplos de redes de los aeropuertos de Acapulco y Monterrey

Una vez obtenidas las redes realizamos los cálculos necesarios para obtener las distintas métricas de redes ya que nos interesa conocer su topología.

Tabla 7 Resultados métricas

Acapulco

Monterrey

Resultados	Nacional	Internacional	Total	Nacional	Internacional	Total
Nodos (Orden)	207	207	207	207	207	207
Arcos (Tamaño)	638	1013	801	791	691	774
Max. Grado	32	43	44	40	31	34
Min. Grado	2	1	2	2	2	2
Grado Promedio	6.164251	9.78744	7.73913	7.642512	6.676329	7.478261
Diámetro	8	7	7	7	8	7
Distancia Media	3.663337	3.415224	3.597908	3.381314	3.878946	3.487454
Cliques	8	9	9	8	8	9
Densidad	0.02992355	0.04751184	0.03756859	0.03709957	0.03240936	0.03630224
Asortatividad	0.1387072	0.01292899	0.03675553	0.09718336	0.04117615	0.08005149
Clustering Global	0.3931777	0.504932	0.4548834	0.4516678	0.4439925	0.4624291
Clustering Local Promedio	0.7447704	0.7810761	0.7675662	0.7705321	0.7700375	0.7809941

Continuación de **Tabla 7**

Además, se graficaron las distribuciones de grado correspondientes a las redes.

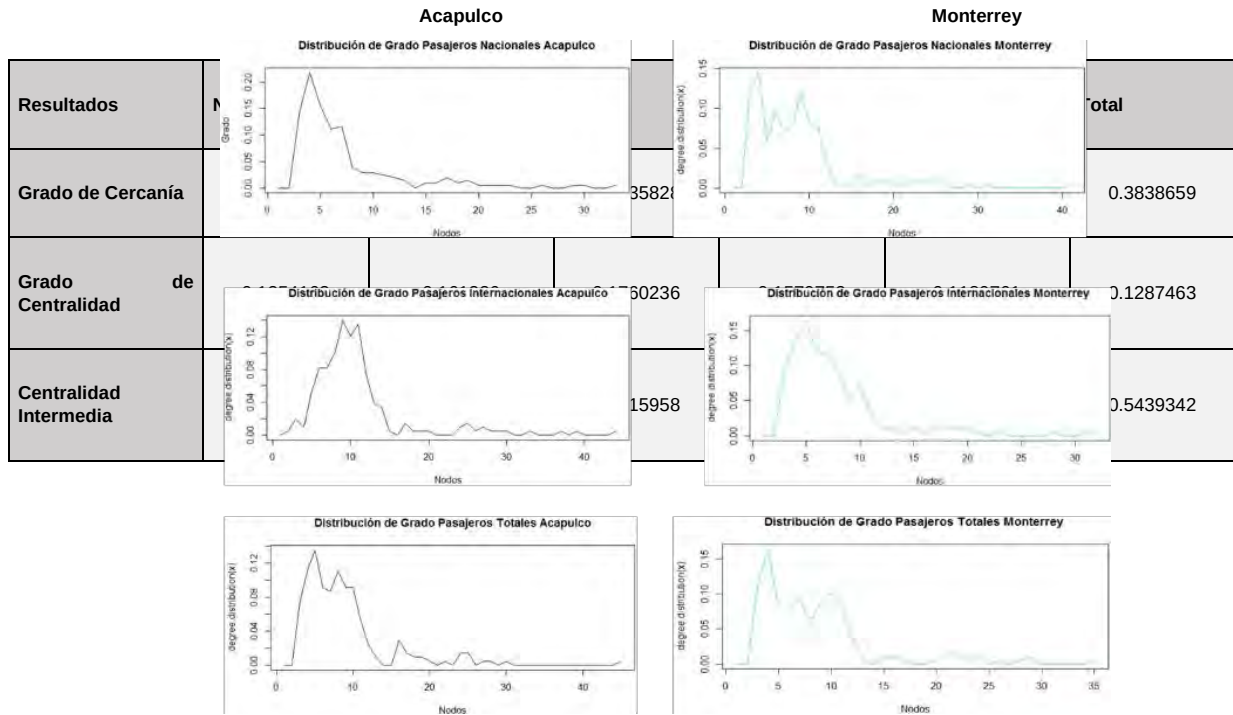


Figura 30 Ejemplos distribución de grado aeropuertos de Acapulco y Monterrey

De las gráficas anteriores y con ayuda de las métricas obtenidas se puede concluir que ambas redes, tanto del aeropuerto de Acapulco como las del aeropuerto de Monterrey, se comportan como redes de libre escala, es decir, se observa que las redes presentan una estructura de hubs entonces cada vez que se añaden más meses (nodos) se va a seguir manteniendo la misma estructura o modelo de redes.

Con esto terminamos el análisis desde el enfoque de redes complejas y procederemos a realizar el análisis de series de tiempo. Para esta parte realizamos todo el análisis en R y utilizamos las paqueterías *TSA*, *TimeSeries* y *forecast*.

Primero obtendremos la descomposición aditiva de las series en sus distintos componentes: ciclos, tendencia y la parte aleatoria. Posteriormente, realizaremos el análisis de la función de autocorrelación (ACF) y de la función de autocorrelación

parcial (PACF). Por último, con toda esta información ajustamos un modelo de series de tiempo de tipo ARIMA (p, d, q) el cual utilizaremos para crear la predicción del flujo de pasajeros en los próximo 24 meses (a partir de abril 2017).

A continuación, mostramos la descomposición aditiva de las series de tiempo.

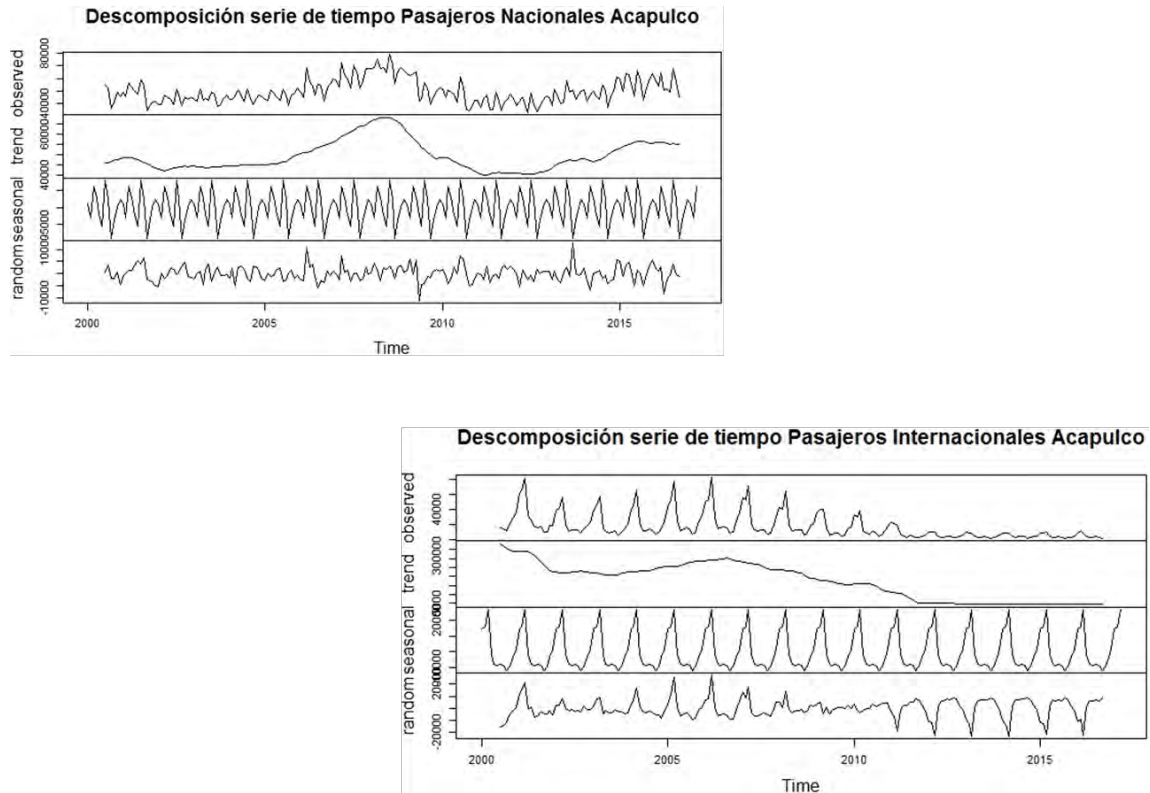
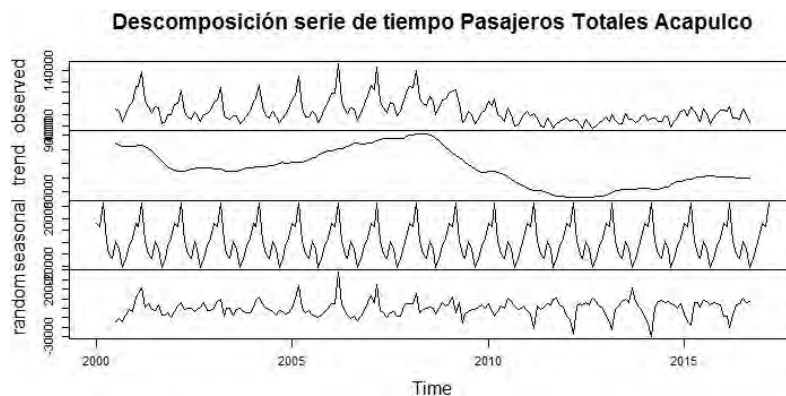


Figura 31 Ejemplo descomposición series de tiempo Pasajeros de Acapulco



Continuación **figura 31**

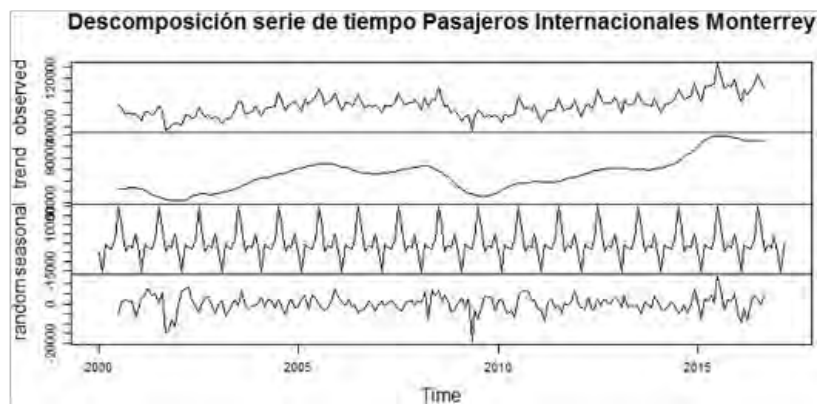
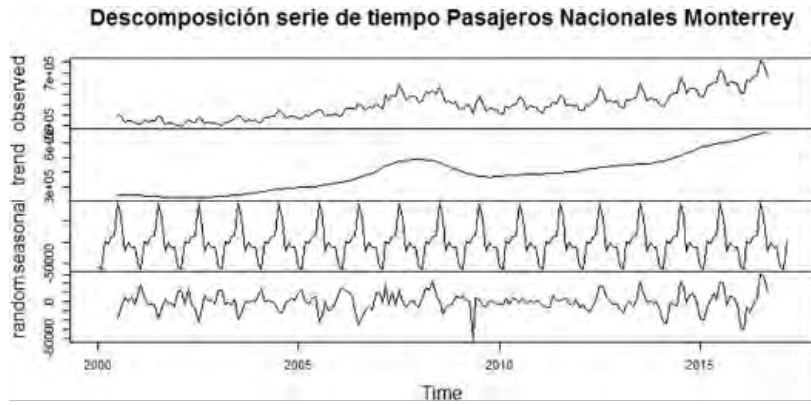
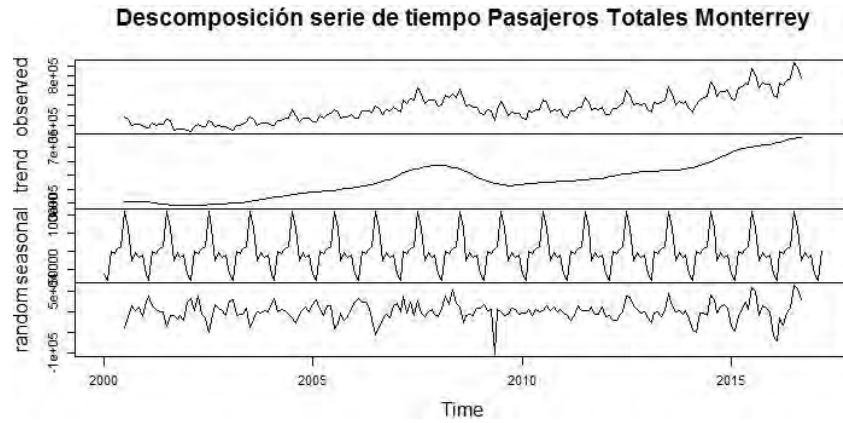


Figura 32 Ejemplo descomposición de series de tiempo Pasajeros de Monterrey

Continuación **Figura 32**



Continuamos con el análisis de series de tiempo y realizamos el análisis de las funciones de autocorrelación y autocorrelación parcial, esto con el fin de hacer un mejor ajuste para cada una de las series de tiempo.

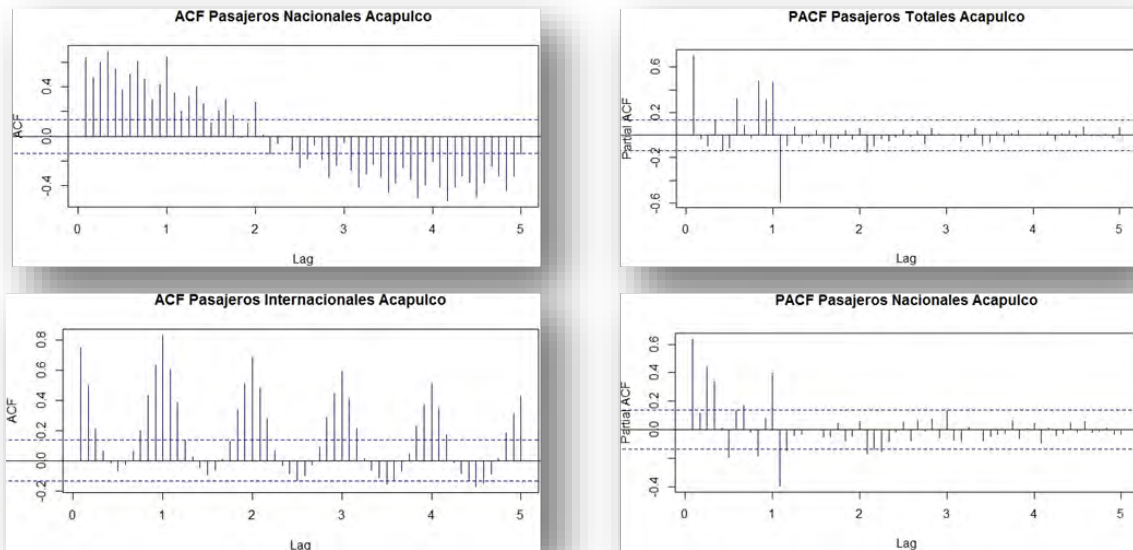
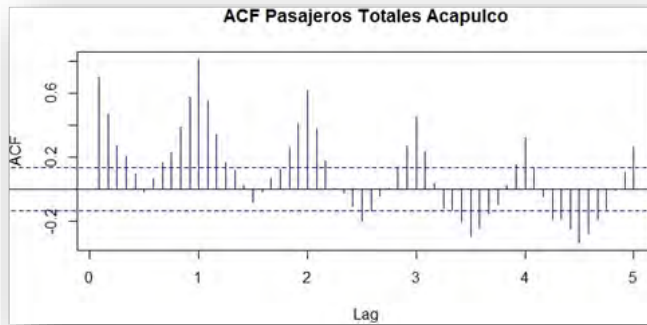
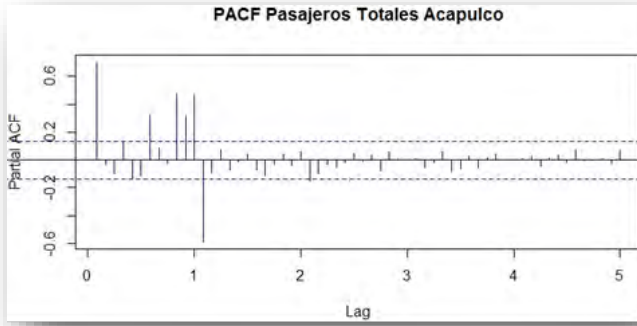


Figura 33 Ejemplos funciones de autocorrelación y autocorrelación parcial Pasajeros Acapulco



Continuación de **Figura 33**

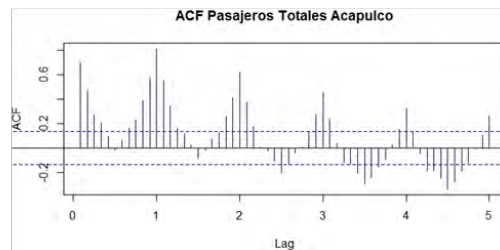
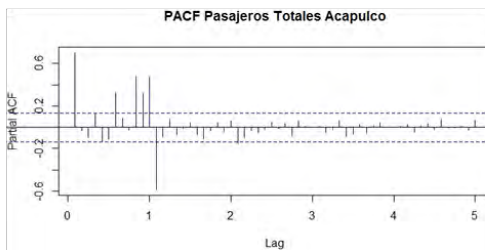
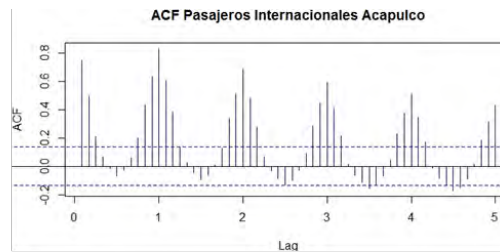
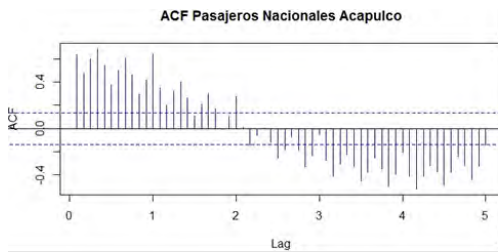
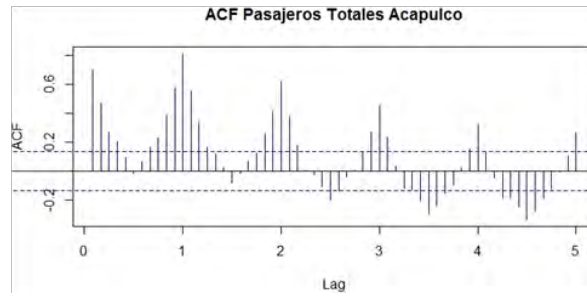
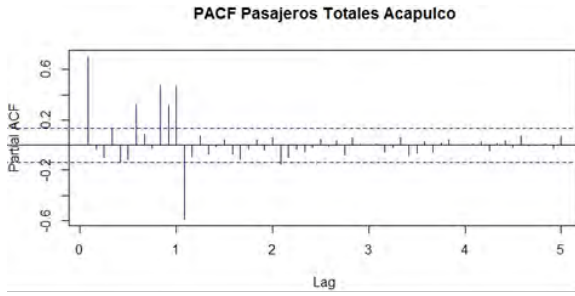


Figura 34 Ejemplos funciones de autocorrelación y autocorrelación parcial Pasajeros Acapulco



Continuación de **Figura 33**

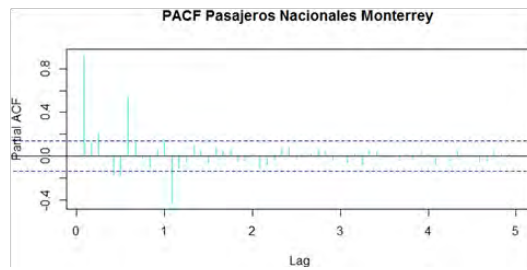
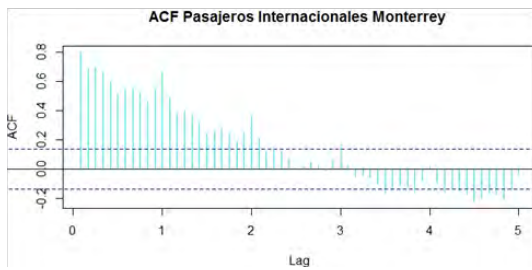
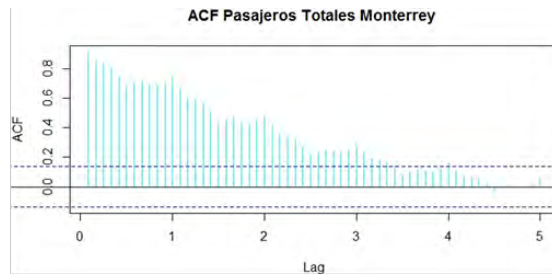
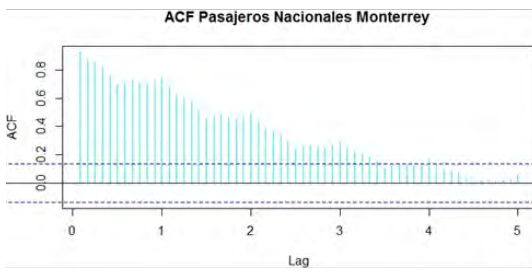
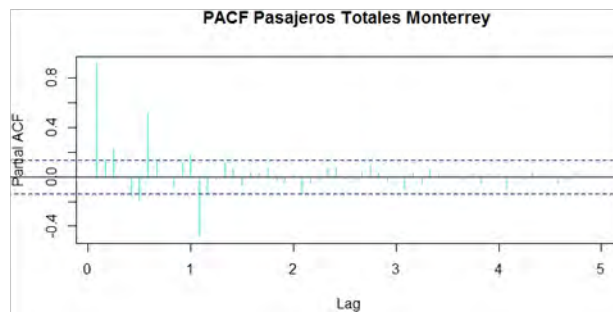
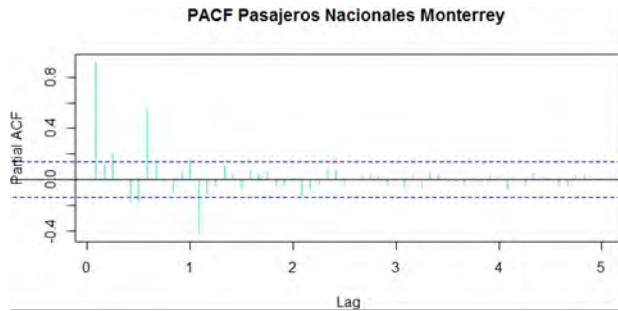


Figura 35 Ejemplos funciones de autocorrelación y autocorrelación parcial Pasajeros Monterrey



Continuación de **Figura 34**

Una vez realizado el análisis de descomposición de la serie, función de autocorrelación y función de autocorrelación parcial procedemos a realizar el ajuste de un modelo de series de tiempo. Ya que como se vio anteriormente tanto las FAC como las PACF nos ayuden a ajustar los mejores modelos para cada una de las series ya que su comportamiento es distinto.

Ajuste modelo de series de tiempo para Pasajeros Nacionales Acapulco

```
Series: databs
ARIMA(0,1,1)(1,0,0)[12]
```

```
Coefficients:
      ma1      sar1
    -0.4991  0.7427
s.e.   0.0634  0.0471
```

```
sigma^2 estimated as 22506667: log likelihood=-2039.97
AIC=4085.95  AICc=4086.06  BIC=4095.93
```

Ajuste modelo de series de tiempo para Pasajeros Internacionales Acapulco

```

Series: databs
ARIMA(0,1,1)(0,0,1)[12] with drift

Coefficients:
      ma1      sma1      drift
    -0.0849  0.7508  -289.8006
s.e.   0.0701  0.0408   905.6337

sigma^2 estimated as 70068570:  log likelihood=-2156.46
AIC=4320.93  AICc=4321.13  BIC=4334.24

```

Ajuste modelo de series de tiempo para Pasajeros Totales Acapulco

```

Series: databs
ARIMA(0,1,1)(0,0,1)[12] with drift

Coefficients:
      ma1      sma1      drift
    -0.2546  0.7325  -163.1702
s.e.   0.0765  0.0475  1009.8962

sigma^2 estimated as 133697374:  log likelihood=-2222.68
AIC=4453.36  AICc=4453.56  BIC=4466.67

```

Ajuste modelo de series de tiempo para Pasajeros Nacionales Monterrey

```

Series: databs
ARIMA(0,1,1)(0,1,1)[12]

Coefficients:
      ma1      sma1
    -0.3101  -0.5409
s.e.   0.0631  0.0604

sigma^2 estimated as 393193848:  log likelihood=-2196.01
AIC=4398.01  AICc=4398.14  BIC=4407.82

```

Ajuste modelo de series de tiempo para Pasajeros Internacionales Monterrey

```

Series: databs
ARIMA(2,0,0)(2,1,0) [12]

Coefficients:
      ar1      ar2      sar1      sar2
      0.7105  0.2383 -0.8040 -0.4105
s.e.  0.0706  0.0700  0.0671  0.0687

sigma^2 estimated as 26537165:  log likelihood=-1946.5
AIC=3903  AICC=3903.32  BIC=3919.36

```

Ajuste modelo de series de tiempo para Pasajeros Totales Monterrey

```

Series: databs
ARIMA(0,1,1)(0,1,1) [12]

Coefficients:
      ma1      sma1
      -0.3103 -0.5762
s.e.  0.0606  0.0570

sigma^2 estimated as 513660915:  log likelihood=-2222.28
AIC=4450.55  AICC=4450.68  BIC=4460.35

```

Por lo tanto, con estos modelos realizamos la predicción del número de pasajeros para los próximos 24 meses con sus respectivos intervalos de confianza.

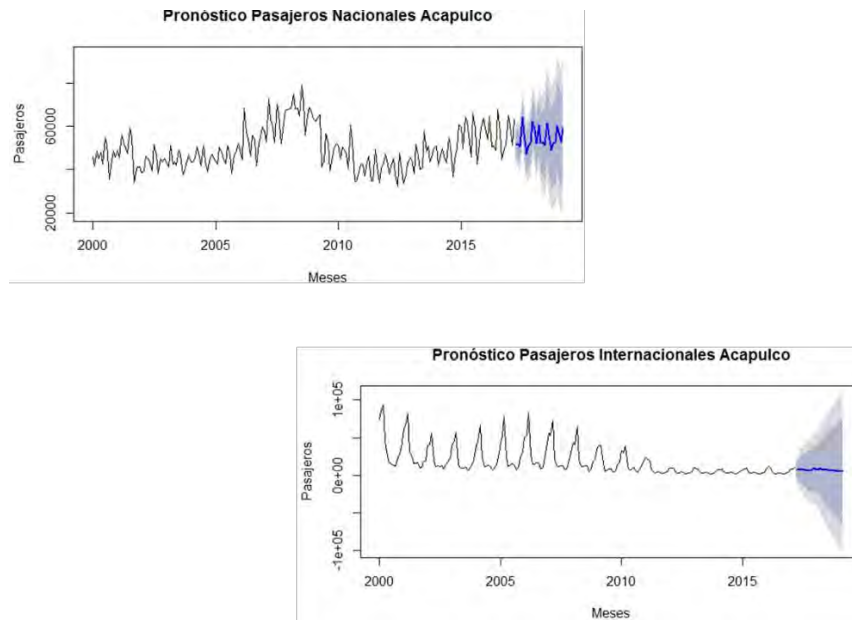


Figura 36 Pronóstico a 24 meses pasajeros Nacionales, Internacionales y Totales de Acapulco

Continuacion **Figura 35**

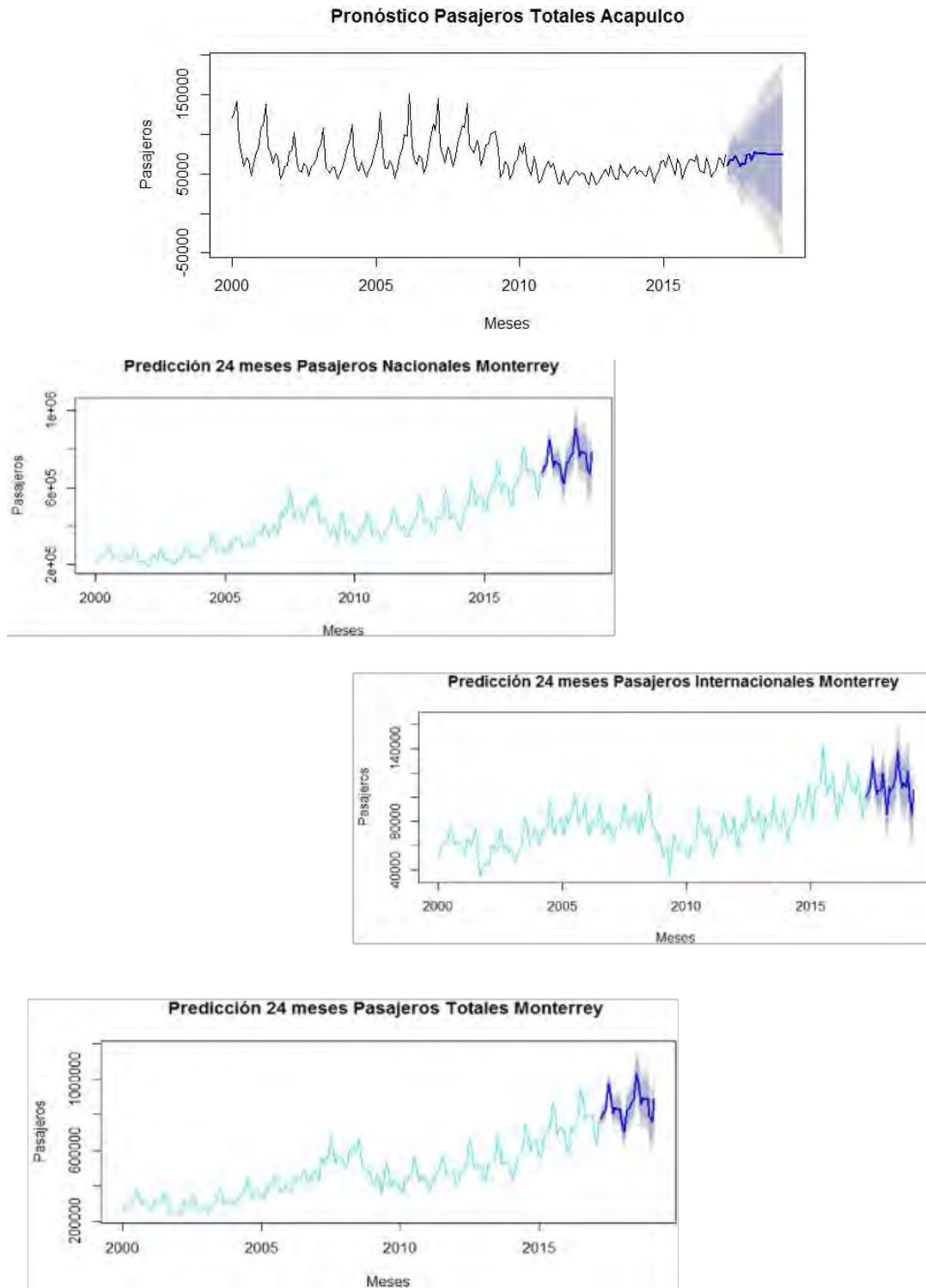


Figura 37 Pronóstico a 24 meses pasajeros Nacionales, Internacionales y Totales de Monterrey

Una vez realizado el pronóstico aplicamos la metodología nuevamente pero ahora sobre las series de tiempo tomando en cuenta la series más la parte del pronóstico,

esto con el fin de comparar las redes complejas de las series contra las series con el pronóstico.

Iniciamos comparando las redes complejas obtenidas mediante el algoritmo de visibilidad solo para los aeropuertos de Acapulco y Monterrey.

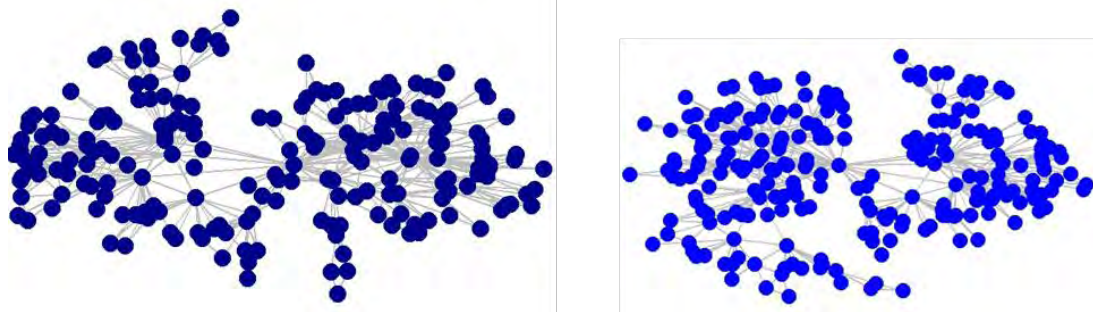


Figura 38 Comparación redes real contra pronóstico Acapulco Nacionales

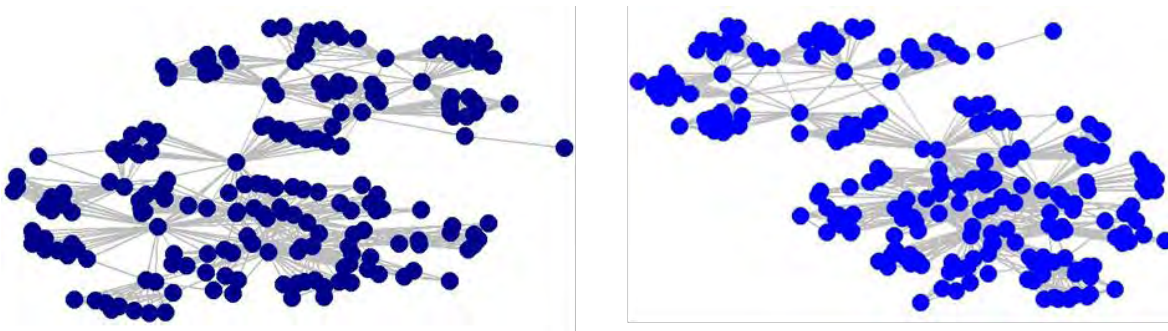


Figura 39 Comparación redes real contra pronóstico Acapulco Internacionales

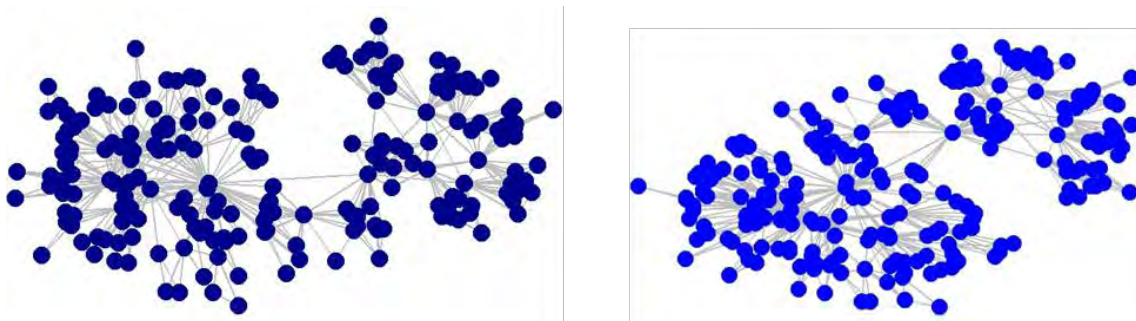


Figura 40 Comparación redes real contra pronóstico Acapulco Totales

A simple vista las redes se comportan de forma similar, es decir, que al agregarle los 24 meses parece que no se modifica la estructura, pero se corroborará con el análisis más a fondo de las redes y esto se hace a través de las métricas que ya se calcularon, es por eso, que en la Tabla 8 se muestra el comparativo de las métricas tanto las reales contra las métricas de las redes obtenidas después del pronóstico. Las primeras tres columnas tanto de Acapulco (Acapulco, Acapulco2 y Acapulco3) como de Monterrey corresponden a las redes reales, mientras que las siguientes 3 columnas de Acapulco (Acapulco32, Acapulco33 y Acapulco34) y Monterrey corresponden a las redes del pronóstico.

Tabla 8 Tabla comparativa métricas reales contra pronóstico Acapulco y Monterrey

Aeropuerto	Acapulco	Acapulco2	Acapulco3	Acapulco32	Acapulco33	Acapulco34	Monterrey	Monterrey	Monterrey	Monterrey	Monterrey	Monterrey3
Resultados	Nacional	International	Total	Nacional	International	Total	Nacional	International	Total	Nacional	International	Total
Nodos	207	207	207	231	231	231	207	207	207	231	231	231
Arcos	638	1013	801	697	1148	879	791	691	774	895	771	869
Max. Grado	32	43	44	32	59	52	40	31	34	42	32	36
Min. Grado	2	1	2	2	1	2	2	2	2	2	2	2
Grado Promedio	6.164251	9.78744	7.73913	6.034632	9.939394	7.61039	7.642512	6.676329	7.478261	7.748918	6.675325	7.52381
Diámetro	8	7	7	9	7	7	7	8	7	7	8	7
Distancia Media	3.663337	3.415224	3.597908	3.902503	3.314813	3.571579	3.381314	3.878946	3.487454	3.450405	3.947186	3.535893
Cliques	8	9	9	8	9	9	8	8	9	8	8	9
Densidad	0.02992355	0.04751184	0.03756859	0.02623753	0.04321476	0.03308865	0.03709957	0.03240936	0.03630224	0.03369095	0.02902315	0.03271222
Asortatividad	0.1387072	0.01292899	0.03675553	0.1590976	-0.04778021	0.02095854	0.09718336	0.04117615	0.08005149	0.1099698	0.0420205	0.08372664
Clustering Global	0.3931777	0.504932	0.4548834	0.3957882	0.4496091	0.426846	0.4516678	0.4439925	0.4624291	0.460479	0.4415493	0.464485
Clustering Local Promedio	0.7447704	0.7810761	0.7675662	0.7463036	0.7832875	0.7638752	0.7705321	0.7700375	0.7809941	0.7751273	0.7700946	0.7829297
Grado de Cercanía	0.3800197	0.3303979	0.3635828	0.3641112	0.3487378	0.3811166	0.408025	0.2979166	0.3838659	0.4098317	0.3196855	0.3911445
Grado de Centralidad	0.1254163	0.161226	0.1760236	0.1128929	0.213307	0.1929983	0.1570752	0.1180761	0.1287463	0.1489177	0.1101073	0.1238095
Centralidad Intermedia	0.5545298	0.503284	0.5915958	0.5498444	0.486264	0.5603983	0.5472721	0.4535664	0.5439342	0.536457	0.4418242	0.5479397

Por último, es de suma importancia analizar las gráficas de distribución de grado de las redes ya que estas nos brindan información sobre una de las características importantes para poder clasificar a nuestras redes en alguno de los modelos de redes complejas (redes aleatorias, redes de mundo pequeño y redes de libre escala). Por lo que en las Figuras 40 y 41 se muestran del lado izquierdo las gráficas de la distribución de grado de las redes reales y del lado derecho se observan las gráficas de distribución de grado de las redes del pronóstico.

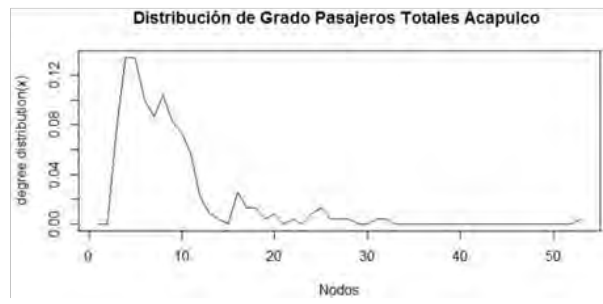
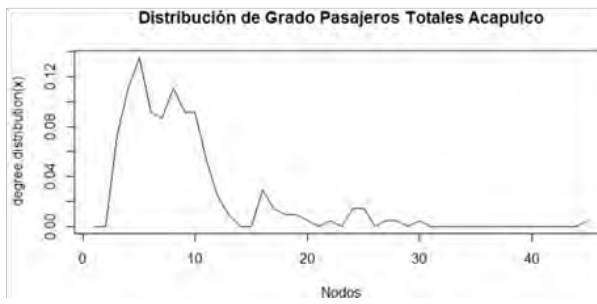
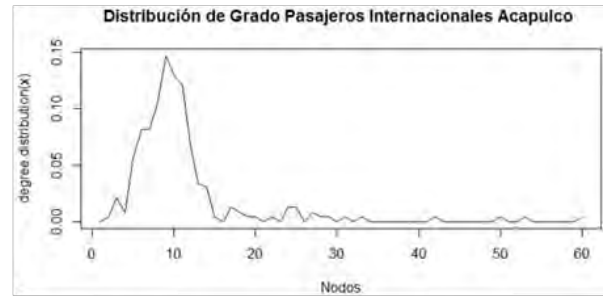


Figura 41 Comparación Distribución de Grado Acapulco

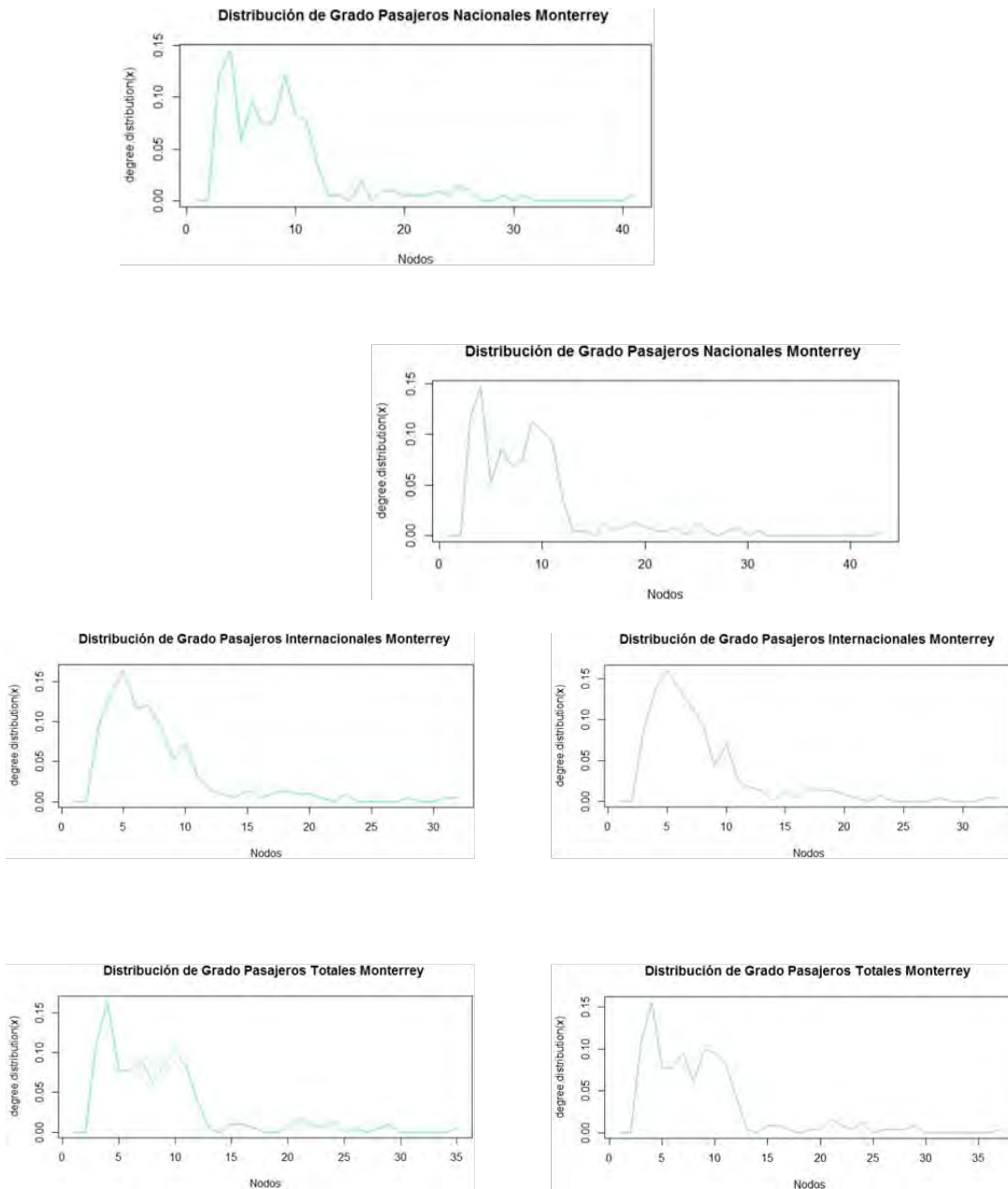


Figura 42 Comparación Distribución de Grado Monterrey

Cabe señalar que las distribuciones de grado de las redes de pronóstico de Acapulco son similares a las redes reales, sin embargo, se observa que están un poco más suaves, es decir, no se observan los picos tan marcados, eso significa que hay una mayor acumulación de nodos con el mismo grado en esas zonas. Para el caso de las redes de Monterrey, el comportamiento que se observa en la distribución de grado es muy parecido, prácticamente igual, a excepción de los pasajeros internacionales que vemos un comportamiento similar al descrito anteriormente, es decir, que se observa un suavizamiento en los picos de la gráfica

Capítulo 4 Conclusiones

Por último, hablaremos de forma general de las conclusiones a las que llegamos. Al hablar de operaciones de aeropuertos no podemos hacer generalizaciones, ya que todos los aeropuertos se comportan y se relacionan de maneras distintas. Lo anterior se debe a distintos factores como la localización, capacidad, aerolíneas y otros factores externos, pero todo esto conlleva distintas características, propiedades y estructuras.

Algo importante que se puede observar es un decremento importante en el número de pasajeros en el aeropuerto de Acapulco, debido a factores externos, como la inseguridad que se ha presentado en los últimos años. Por otro lado, se observa un incremento en el número de pasajeros en el aeropuerto de Monterrey, ya que hay un incremento significativo en la industria desde hace un par de años.

Al realizar el comparativo entre las redes actuales contra las redes que incluyen el pronóstico, la primera observación obvia es que el número de nodos aumentó en 24 es decir, pasó de 207 a 231, de igual forma el número de arcos se incrementó y ahora la red de pasajeros internacionales de Acapulco cuenta con 1148 arcos, también se observa que el grado máximo es de 59 contra el grado mínimo que es de 1 y el grado promedio máximo es de 9, esto nos dice que en promedio los nodos tienen 9 arcos, y para nuestro análisis significa que en promedio un mes se conecta con otros nueve meses, lo cual puede requerir un análisis más a fondo de las vulnerabilidades de la red y ver la robustez y cómo afecta a los meses futuros. En general, se tiene una distancia media promedio muy baja, en este caso de 3, esto nos habla de que existe una buena conexión entre los nodos. Además, la mayoría de las redes son asortativas y significa que los nodos con un grado alto tienden a conectarse con otros nodos de grado alto por lo que la comunicación fluye de manera más eficiente. Sin embargo, una de las redes disortativa (Acapulco doméstico con pronóstico), nos indica que los nodos de grado alto se conectan con nodos de grado bajo y esto puede generar cuellos de botella. En todas las redes se observa un clustering local promedio muy alto lo que nos refuerza la idea de que nuestros nodos están muy bien conectados. Al analizar las métricas de centralidad se observa que todas las redes tienen un mismo comportamiento, tanto en el grado de cercanía como en el grado de centralidad, y la centralidad intermedia, realmente sus valores varían muy poco, por lo que tienen una topología muy similar.

Con respecto a la distribución de grado, hay una tendencia a que se comportan como modelos de redes de libre escala, es decir, el tráfico de pasajeros es representado como un modelo de redes de libre escala, es decir, tiende a presentar hubs, además de una distancia media muy baja y un clustering local muy alto. También, podemos especificar que el funcionamiento de la red es muy bueno porque tiene buen nivel de conexión, los números de cliques que se forman es muy similar, y esto se refuerza ya que las series de tiempo son muy estacionales y cíclicas y esto tiene mucho sentido, ya que la mayor parte de los pasajeros viajan en ciertas estaciones del año como en semana santa, vacaciones de verano, navidad y año nuevo.

Los pronósticos realizados nos muestran que el tráfico de los pasajeros de Monterrey sigue una tendencia a la alza y con el mismo comportamiento que ha seguido en los últimos años con ciertas estacionalidades. Mientras que el flujo de pasajeros en el aeropuerto de Acapulco el comportamiento sigue una tendencia igual a la que ha seguido en los últimos años, donde se alcanza a ver un pequeño incremento es en los pasajeros nacionales, mientras que para los pasajeros internacionales la línea de pronóstico se asemeja a una línea horizontal sin muchos cambios, por lo que este aeropuerto no tiene problemas de capacidad pero se debería de implementar acciones para mejorar el turismo.

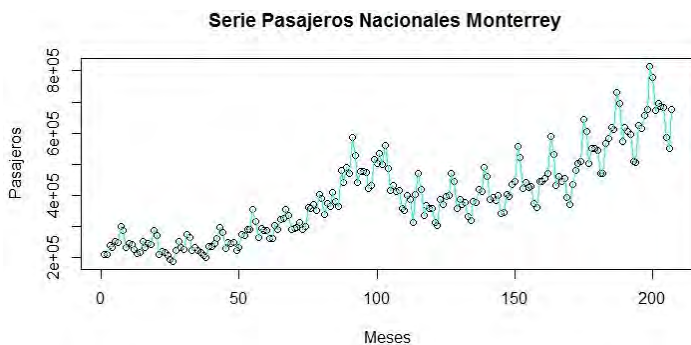
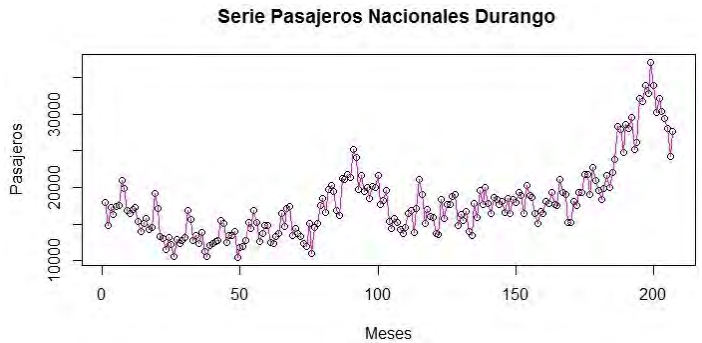
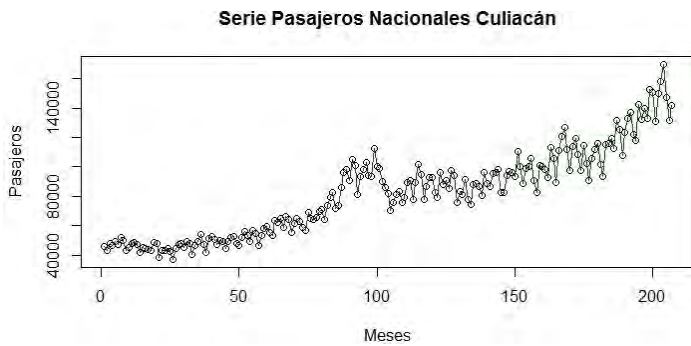
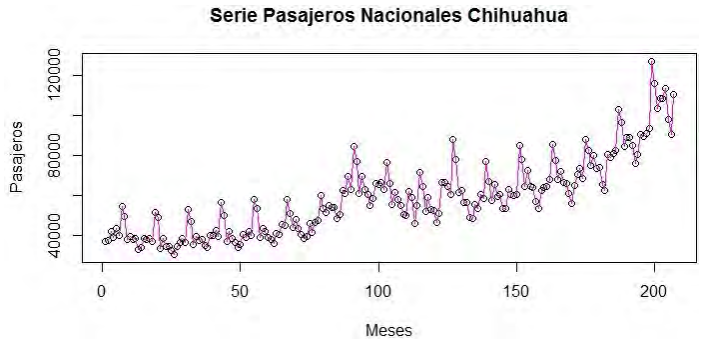
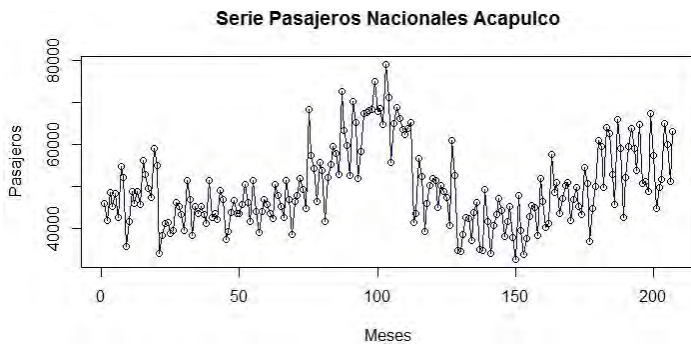
Este tipo de análisis no sólo nos sirve para analizar redes de transporte, sino que se pueden realizar diferentes tipos de análisis tanto a redes de áreas específicas como tecnología, biología, redes sociales, lingüística, entre otras, como análisis comparativos entre las distintas áreas. En general, La finalidad de los análisis y la investigación puede ser comprender mejor las estructuras y la topología, las propiedades locales y globales, así como características específicas de sistemas reales modelados como redes.

En la actualidad es muy importante realizar este tipo de análisis ya que los resultados obtenidos son útiles para diversos fines como mejoras en la planeación de las operaciones en general de los aeropuertos, esto incluye tanto la parte física de los aeropuertos (infraestructura, espacios, capacidad) como en la parte de operaciones (itinerarios/horarios, aerolíneas, seguridad, equipaje). Lo anterior con la finalidad de tomar decisiones adecuadas y oportunas para mejorar todos los servicios brindados por los aeropuertos para poder llevar un mejor control y satisfacer todas las necesidades de todos los involucrados.

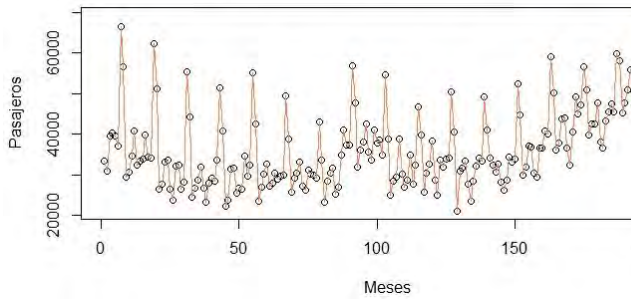
Anexos

En esta sección se presentan las series de tiempo de pasajeros nacionales, internacionales, totales, las redes complejas resultantes de aplicar el algoritmo de visibilidad a las series de tiempo anteriores, gráficas de distribución de grado, descomposición de las series de tiempo, gráficas de funciones de autocorrelación y autocorrelación parcial, pronóstico de los siguientes 24 meses y tablas de resultados de métricas de redes complejas; todo esto para los aeropuertos de Acapulco, Chihuahua, Culiacán, Durango, Monterrey, Mazatlán, San Luis Potosí, Tampico, Reynosa, Torreón, Zacatecas y Zihuatanejo. Además, se anexa los códigos utilizados en R.

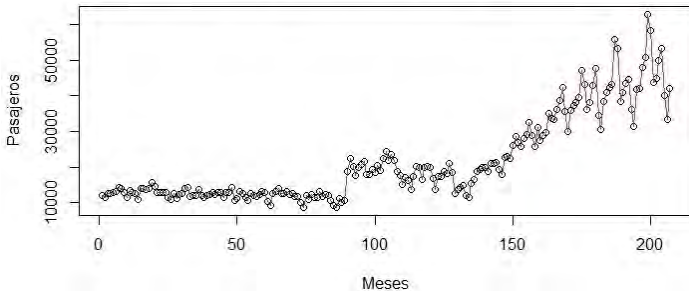
Anexo 1. Gráficas series de tiempo pasajeros nacionales



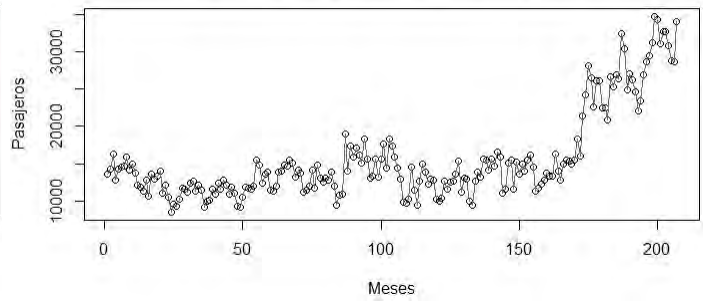
Serie Pasajeros Nacionales Mazatlán



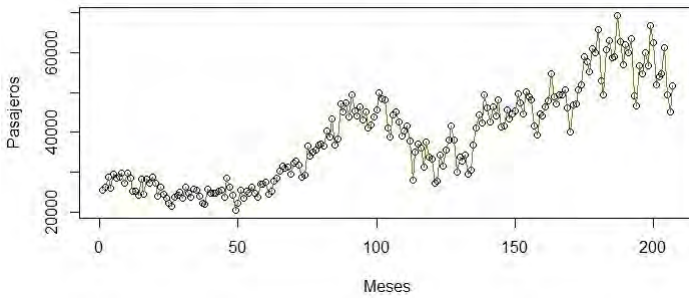
Serie Pasajeros Nacionales Reynosa



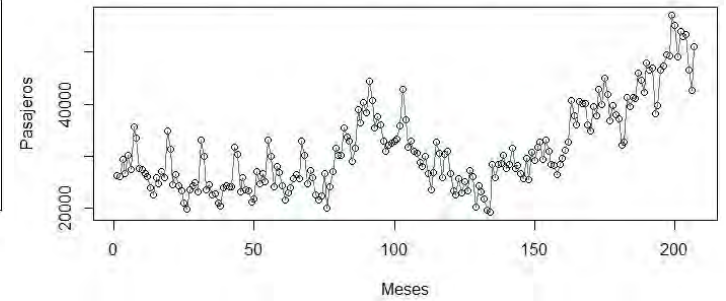
Serie Pasajeros Nacionales San Luis Potosí



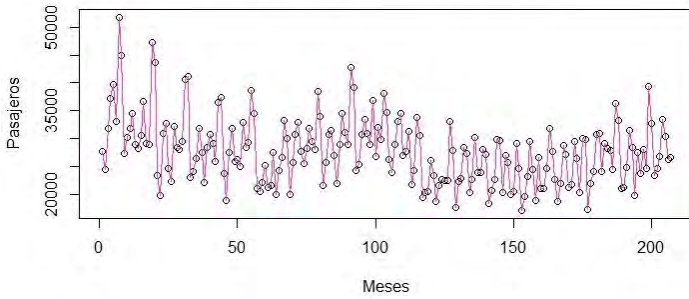
Serie Pasajeros Nacionales Tampico



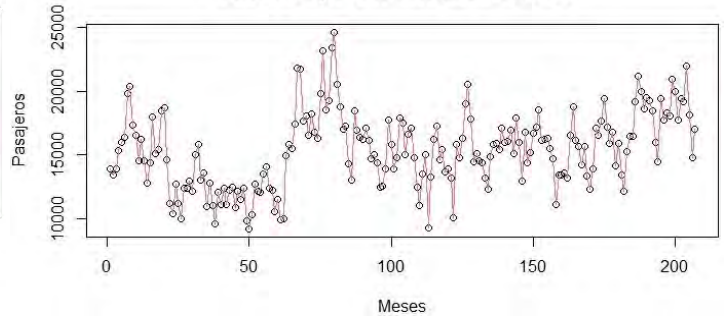
Serie Pasajeros Nacionales Torreón



Serie Pasajeros Nacionales Zihuatanejo

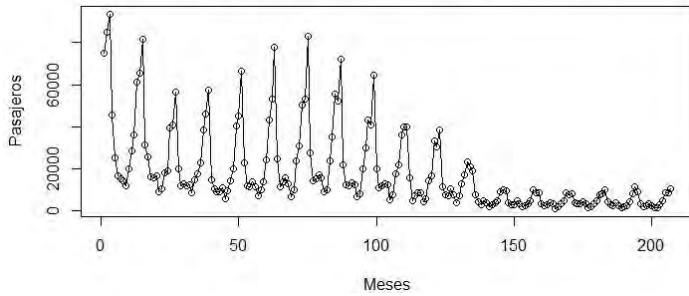


Serie Pasajeros Nacionales Zacatecas

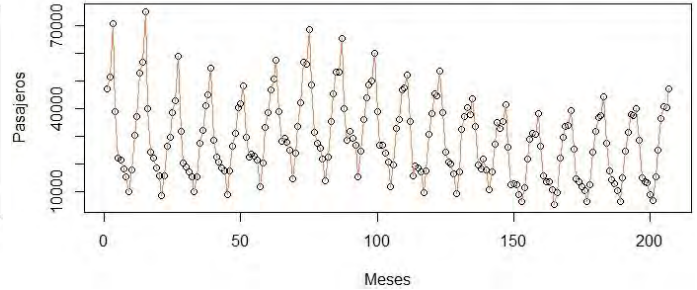


Anexo 2. Gráficas series de tiempo pasajeros internacionales

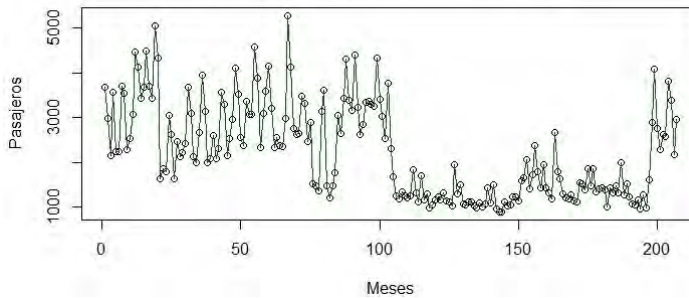
Serie Pasajeros Internacionales Acapulco



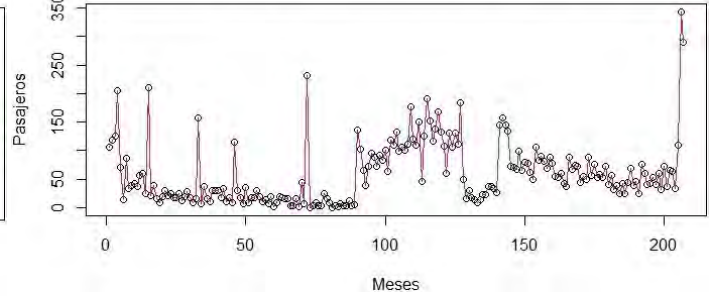
Serie Pasajeros Internacionales Mazatlán



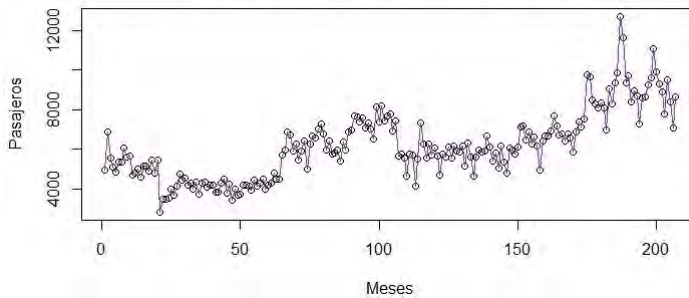
Serie Pasajeros Internacionales Culiacán



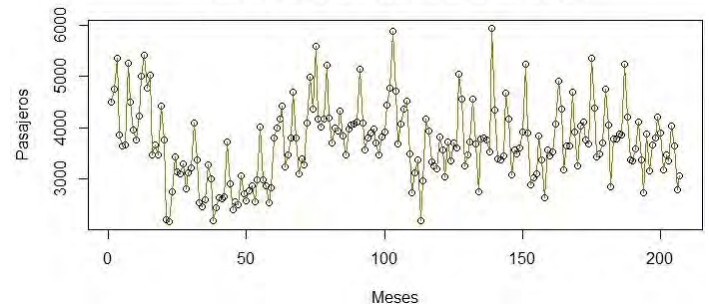
Serie Pasajeros Internacionales Reynosa



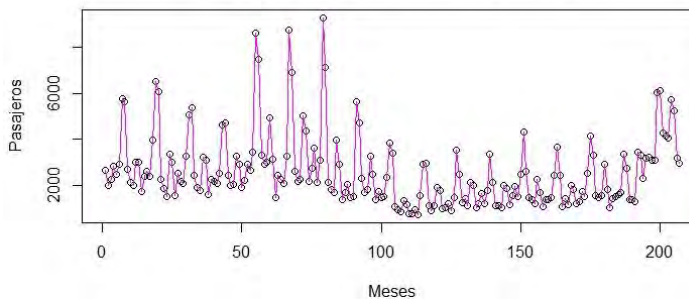
Serie Pasajeros Internacionales Chihuahua



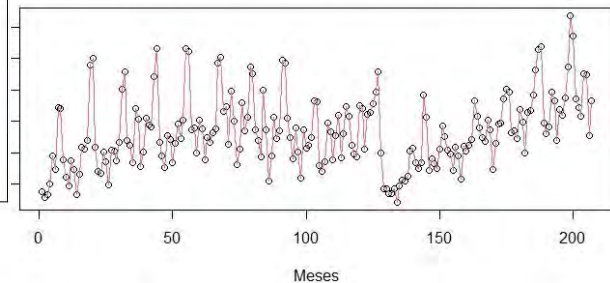
Serie Pasajeros Internacionales Tampico



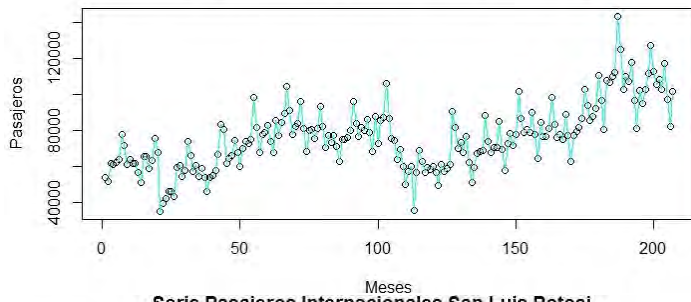
Serie Pasajeros Internacionales Durango



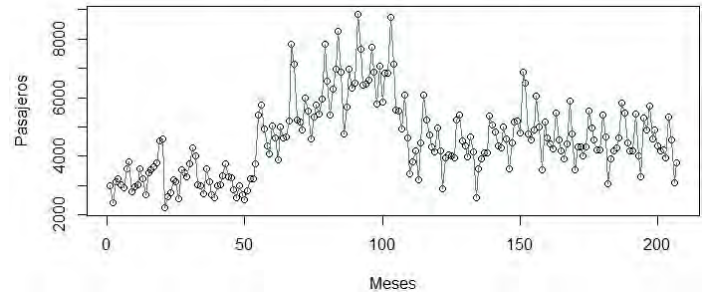
Serie Pasajeros Internacionales Zacatecas



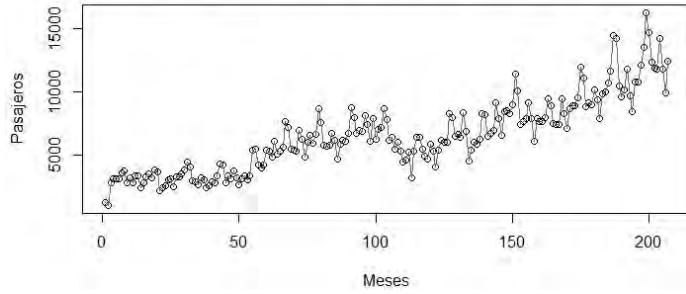
Serie Pasajeros Internacionales Monterrey



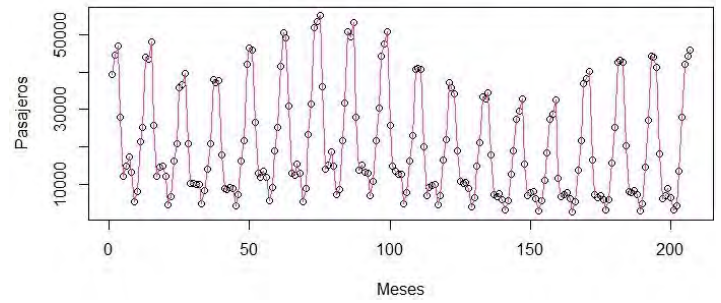
Serie Pasajeros Internacionales Torreón



Serie Pasajeros Internacionales San Luis Potosi

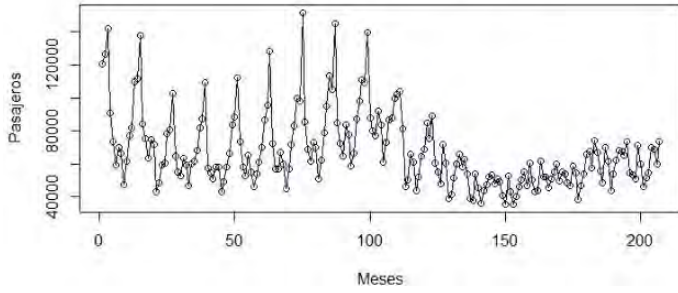


Serie Pasajeros Internacionales Zihuatanejo

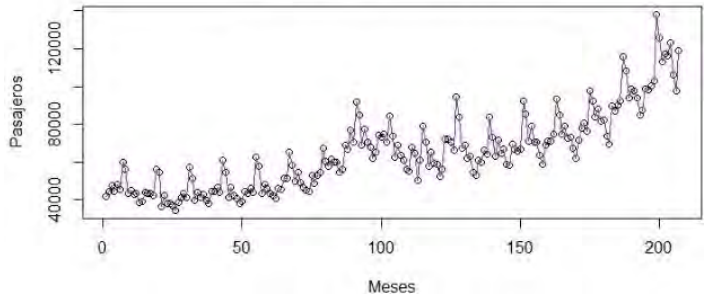


Anexo 3. Gráficas series de tiempo pasajeros totales

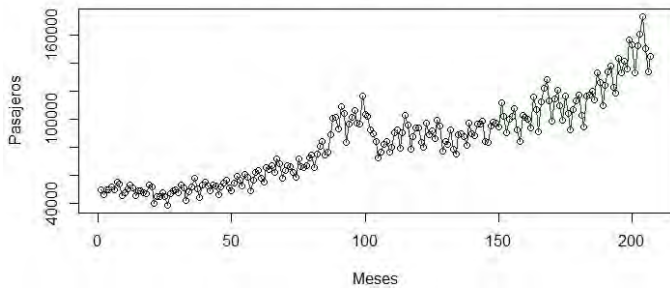
Serie Pasajeros Totales Acapulco



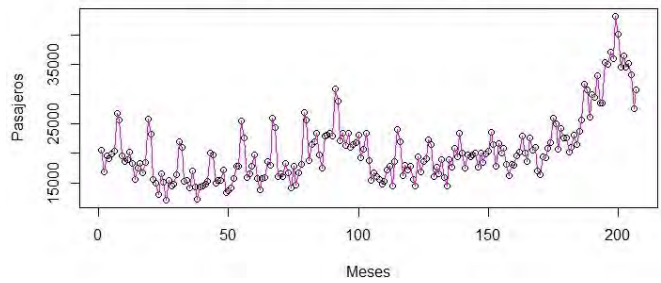
Serie Pasajeros Totales Chihuahua



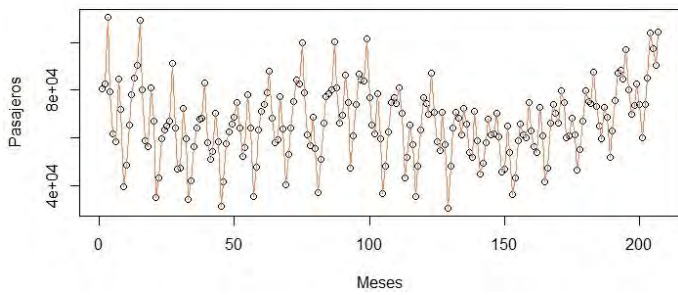
Serie Pasajeros Totales Culiacán



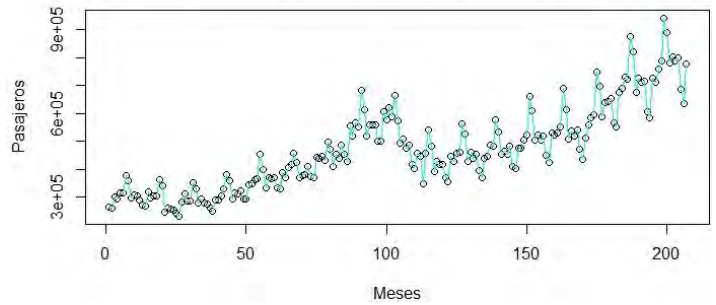
Serie Pasajeros Totales Durango



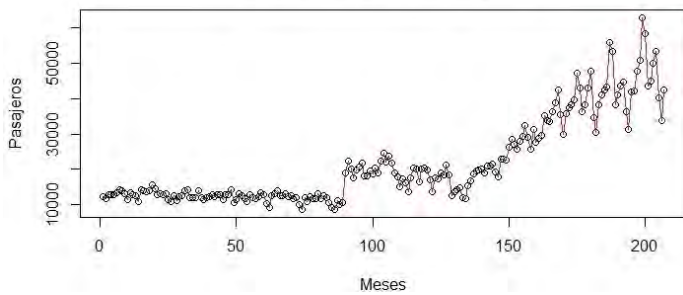
Serie Pasajeros Totales Mazatlán



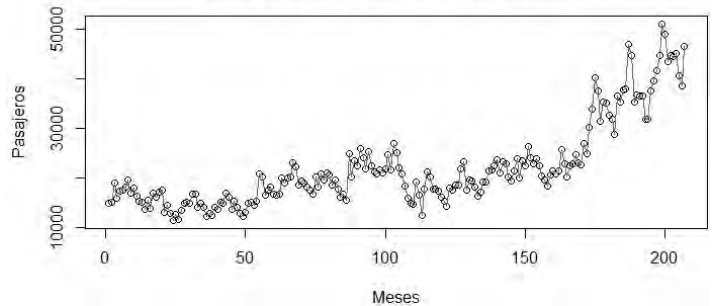
Serie Pasajeros Totales Monterrey



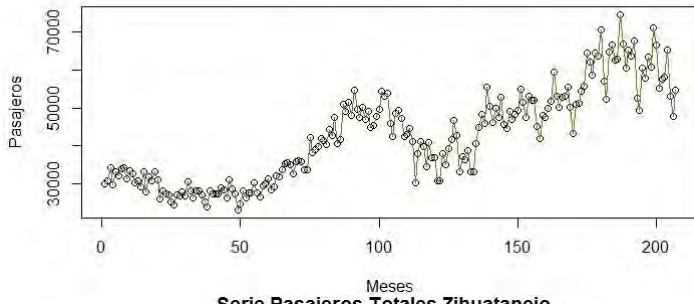
Serie Pasajeros Totales Reynosa



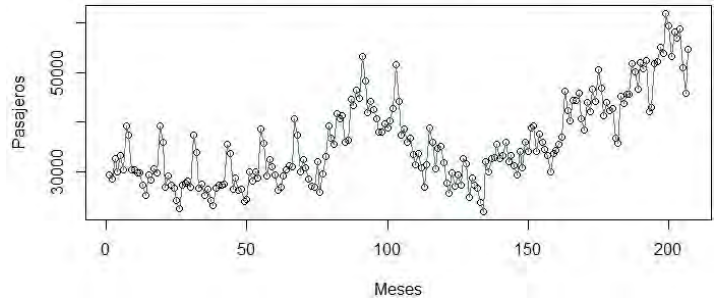
Serie Pasajeros Totales San Luis Potosí



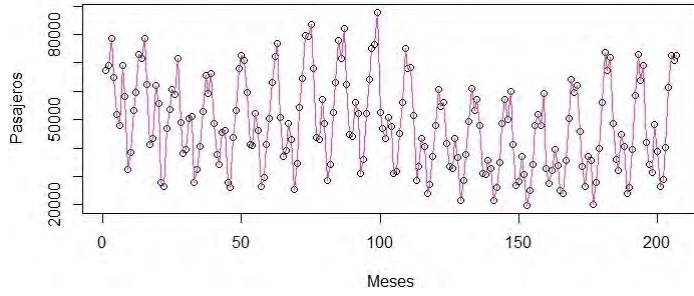
Serie Pasajeros Totales Tampico



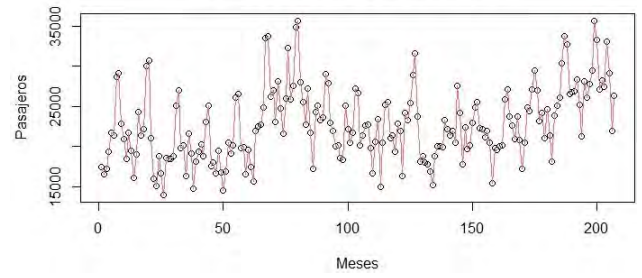
Serie Pasajeros Totales Torreón



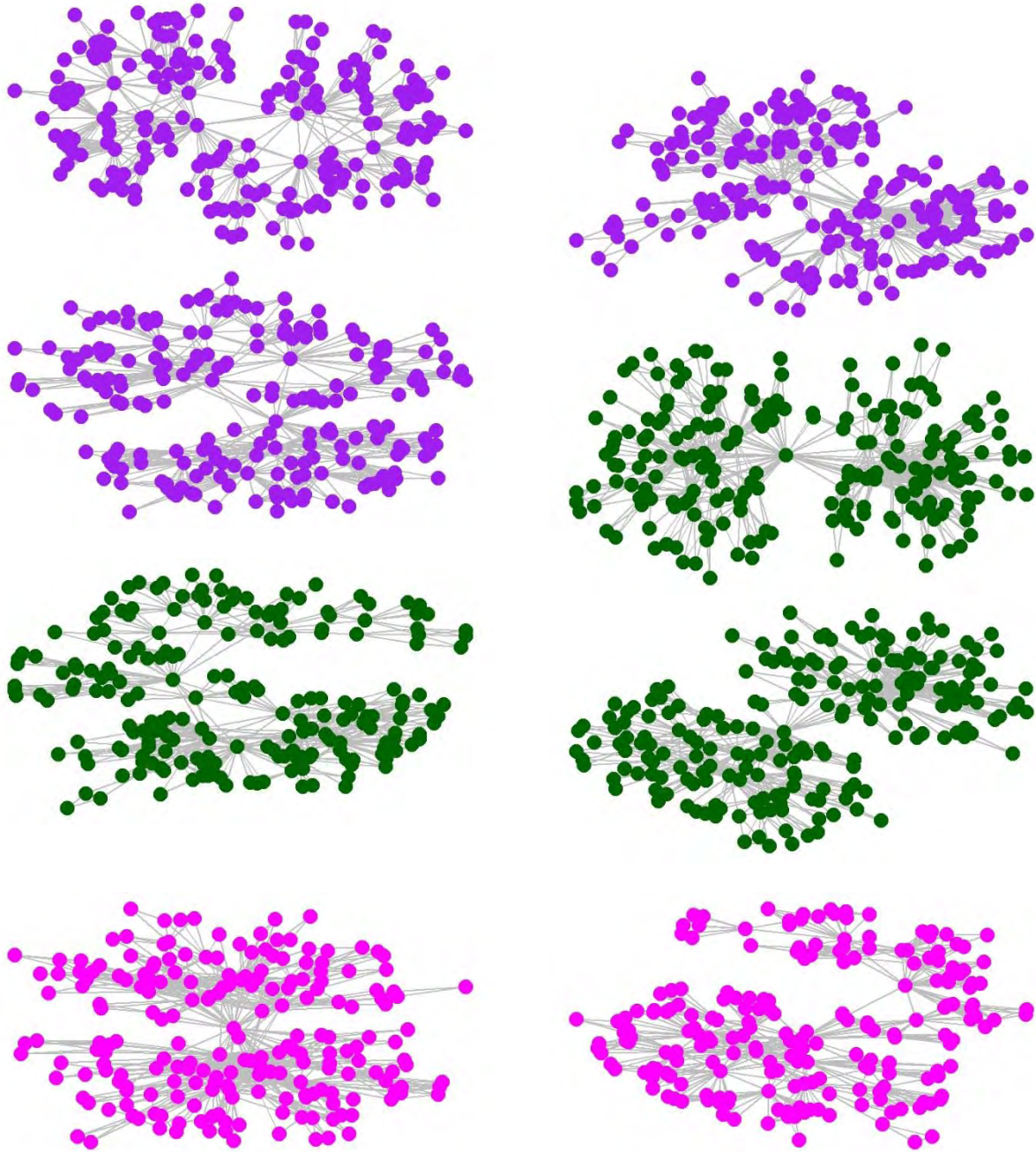
Serie Pasajeros Totales Zihuatanejo

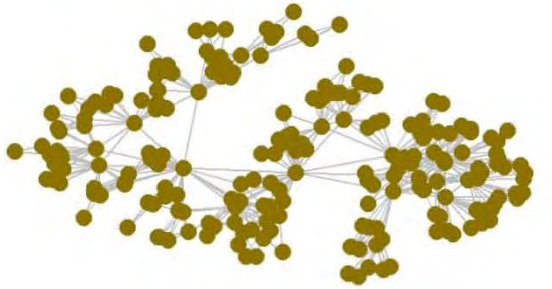
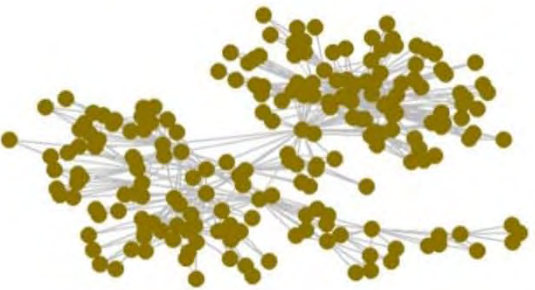
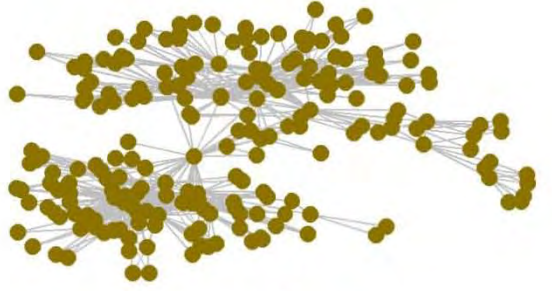
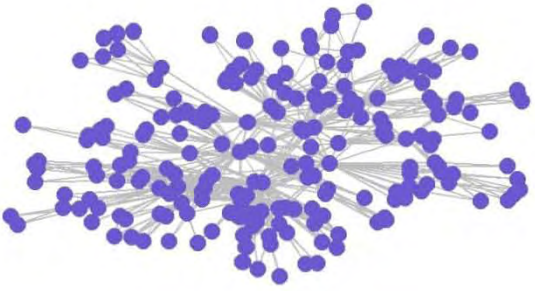
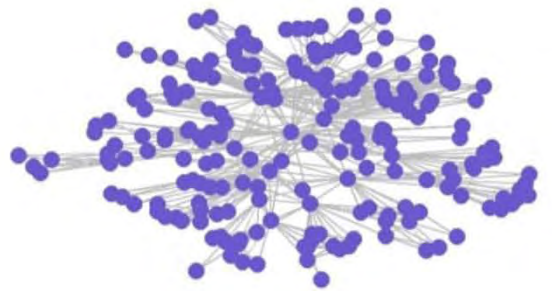
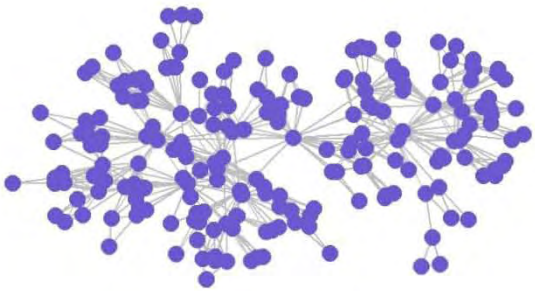
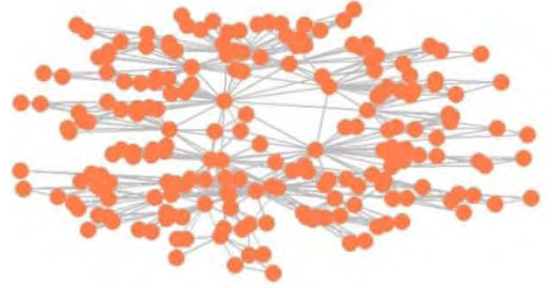
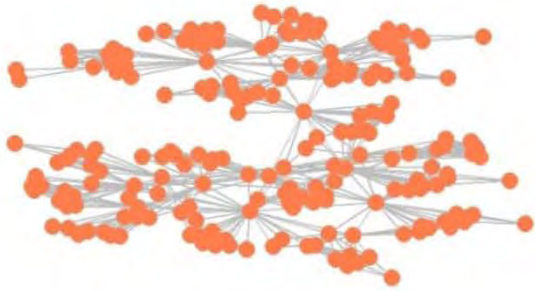
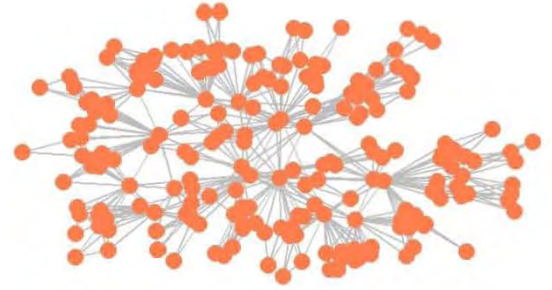
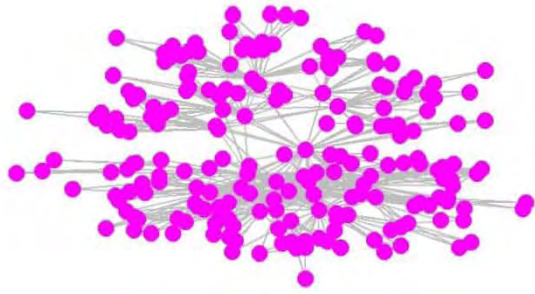


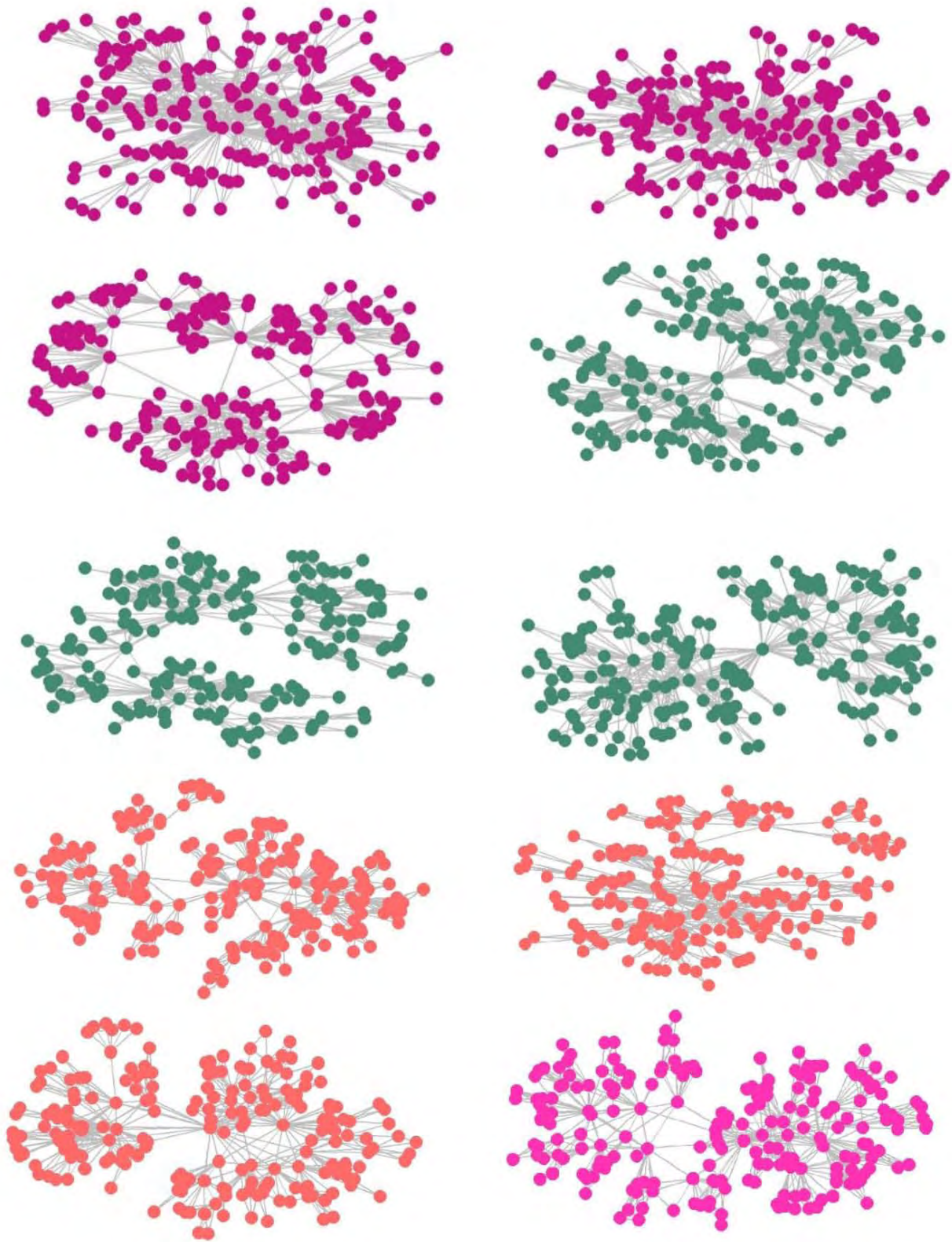
Serie Pasajeros Totales Zacatecas

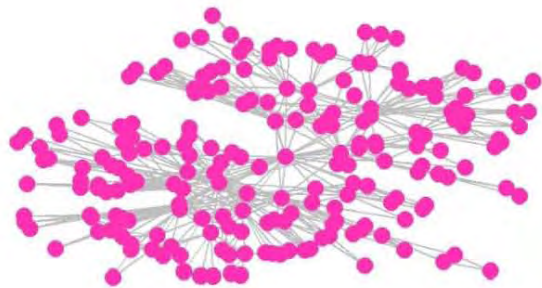
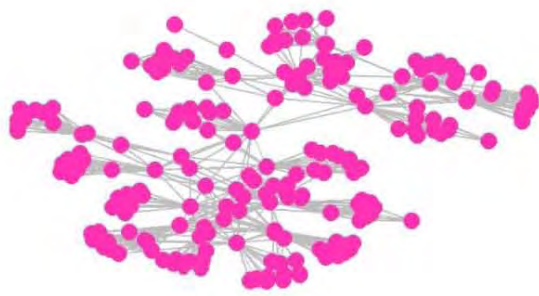


Anexo 4. Redes complejas resultado del algoritmo de visibilidad para pasajeros nacionales, internacionales y totales

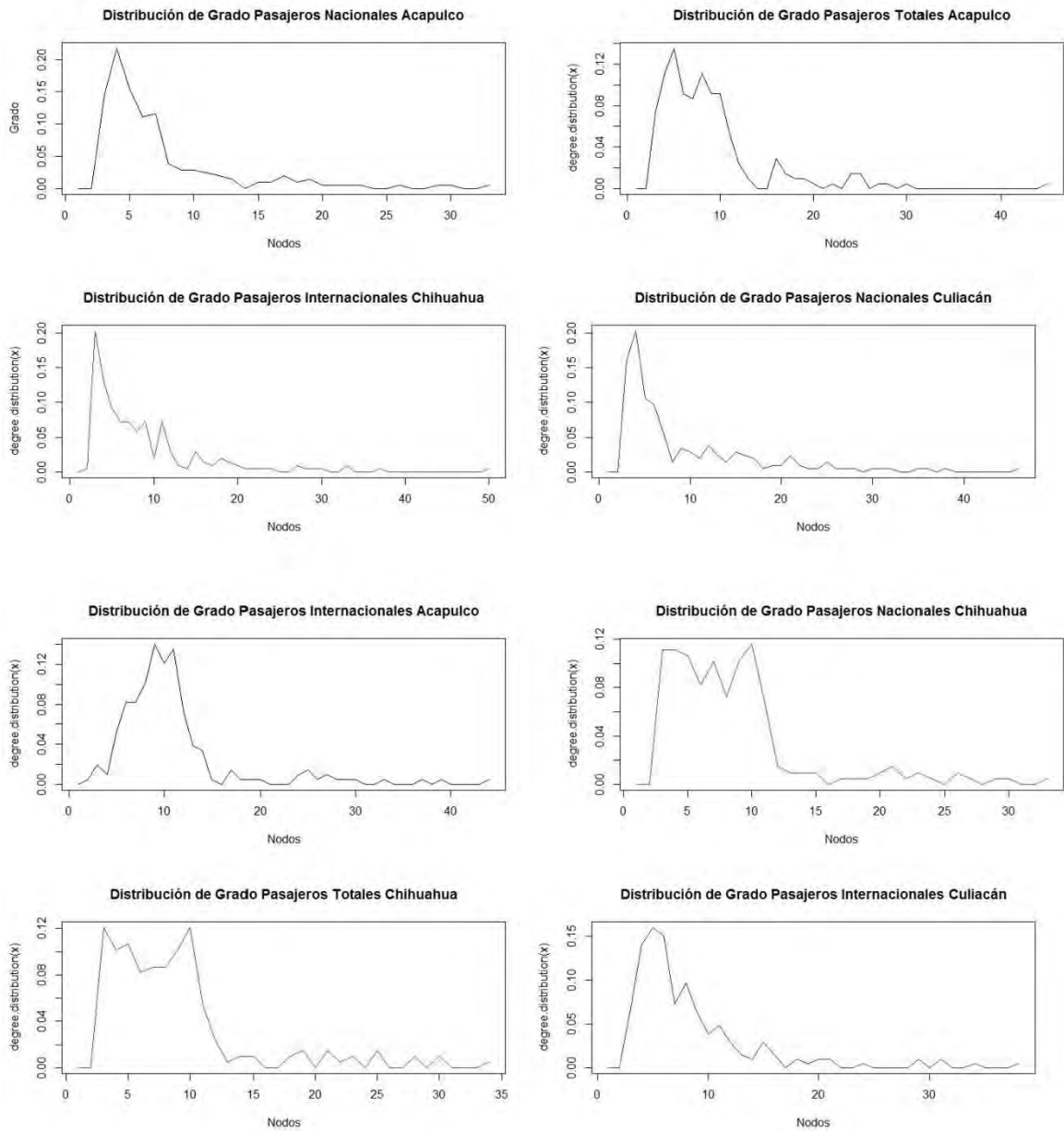


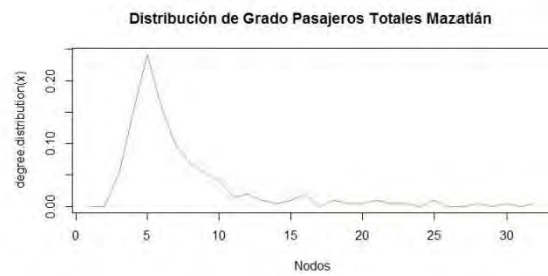
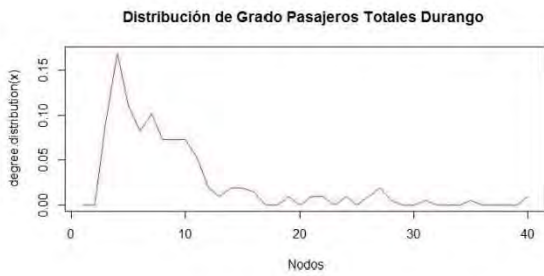
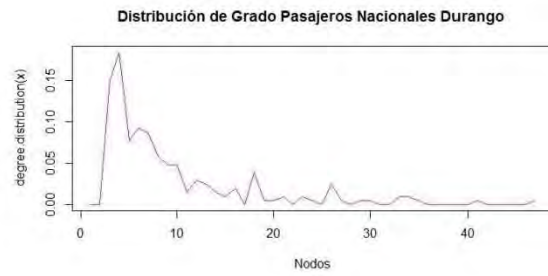
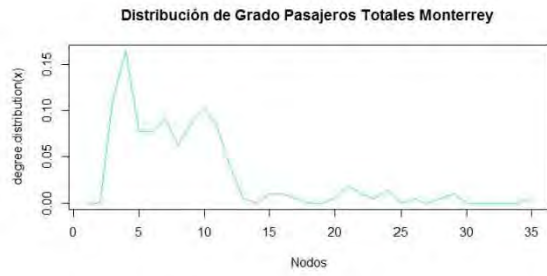
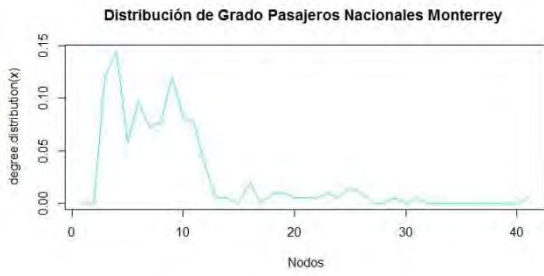
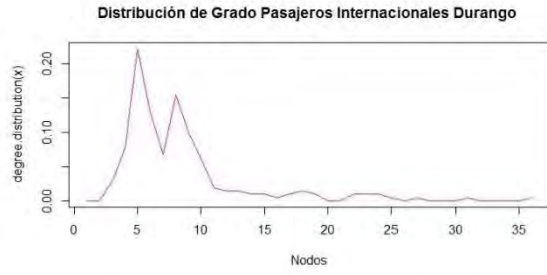
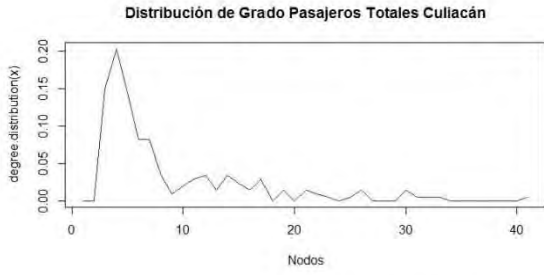


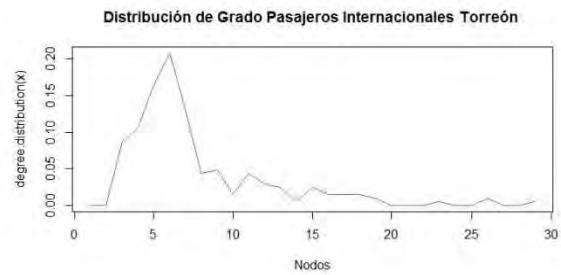
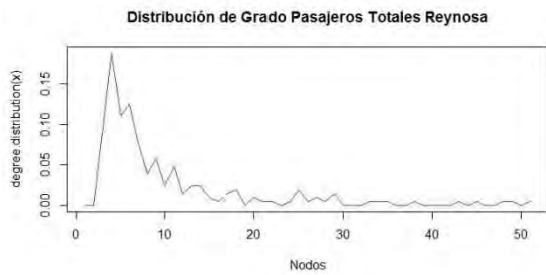
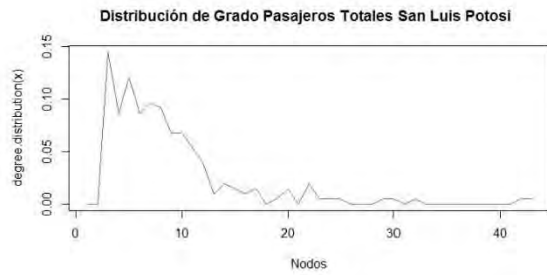
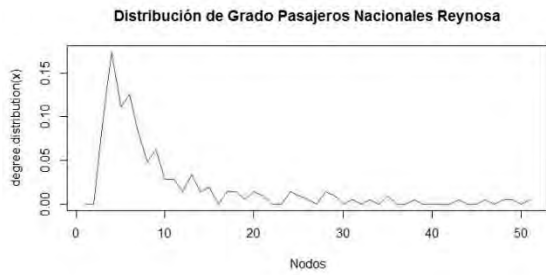
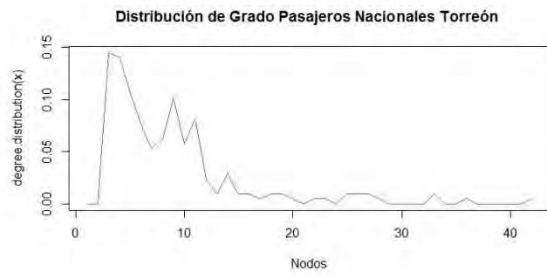
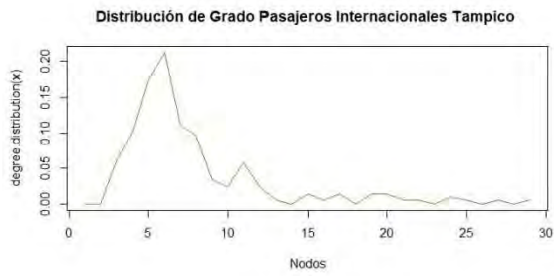
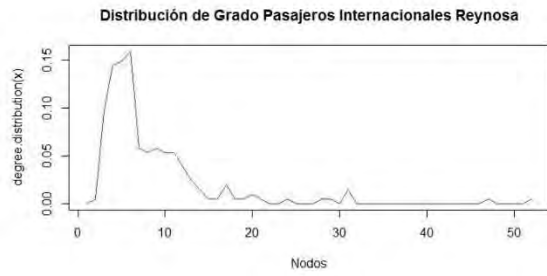
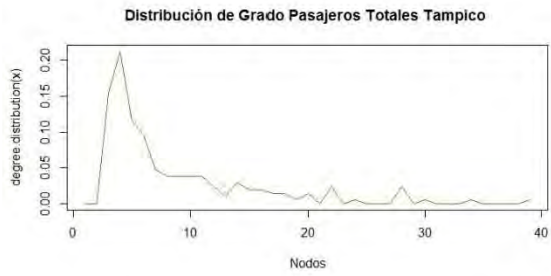
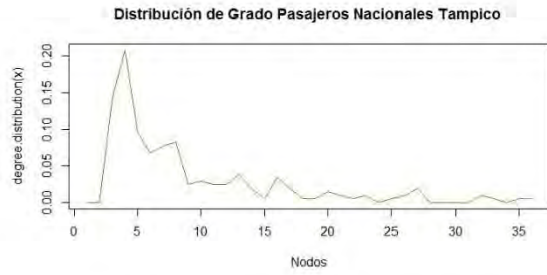


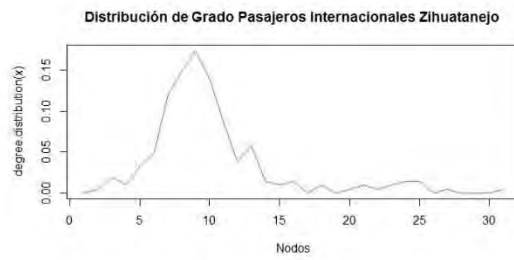
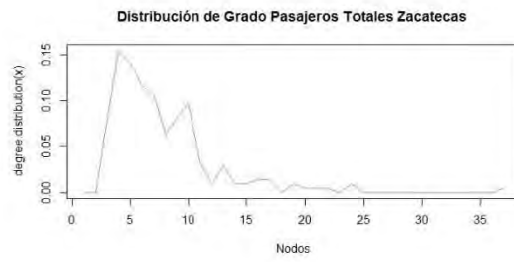
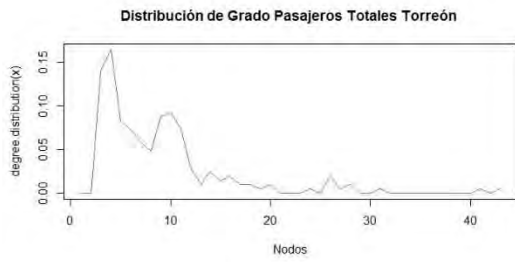


Anexo 5. Distribución de grado

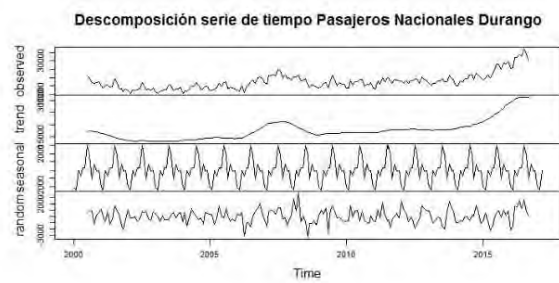
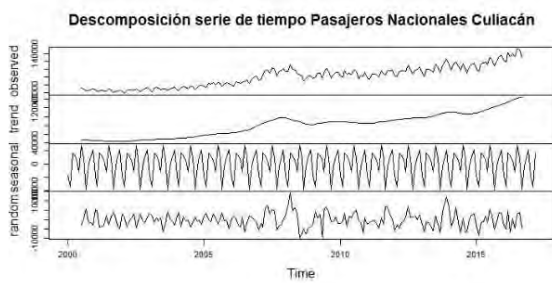
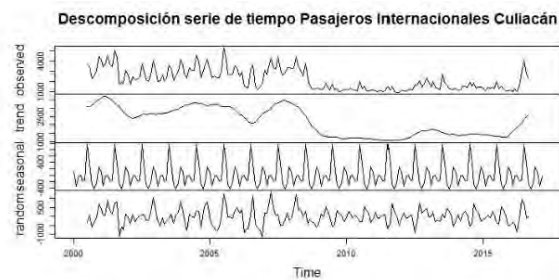
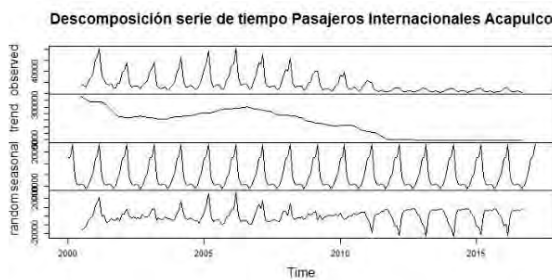
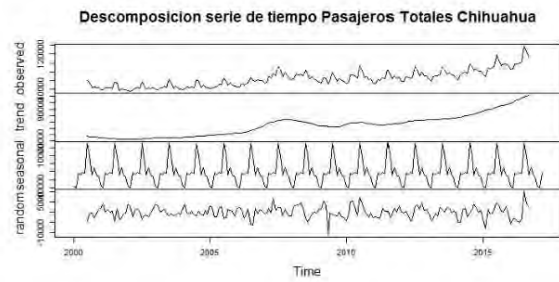
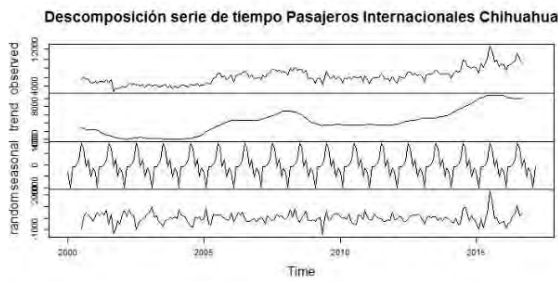
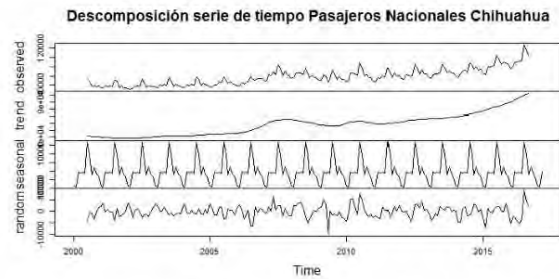
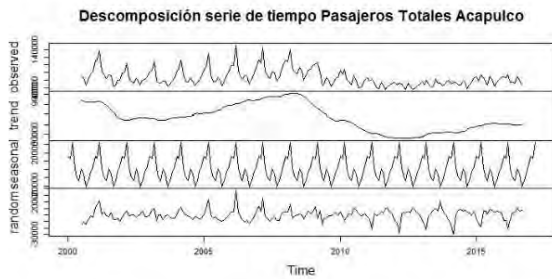
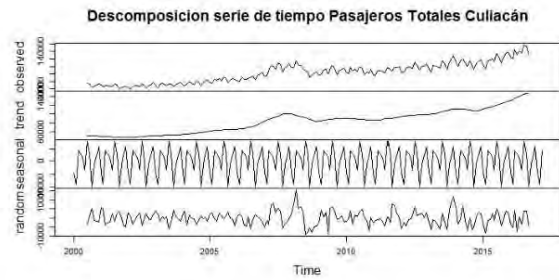
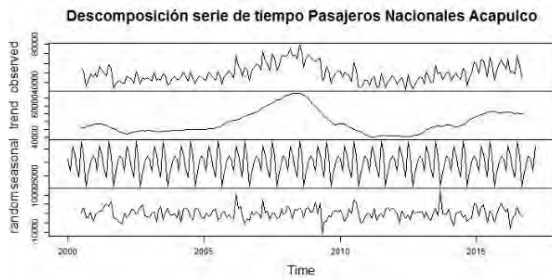




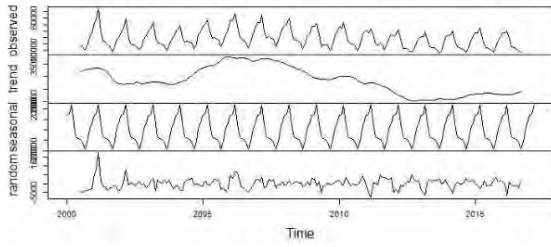




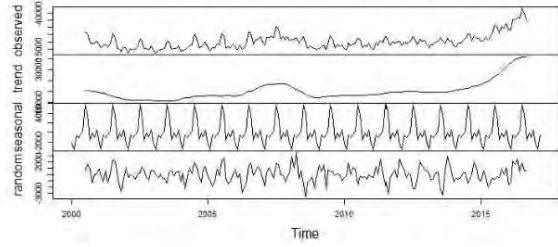
Anexo 6. Descomposición series de tiempo



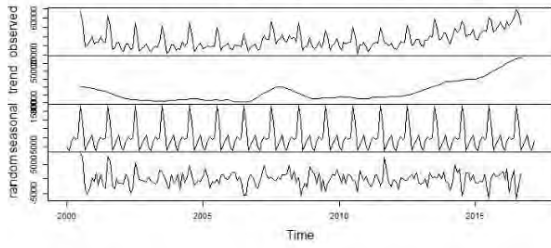
Descomposición serie de tiempo Pasajeros Internacionales Mazatlán



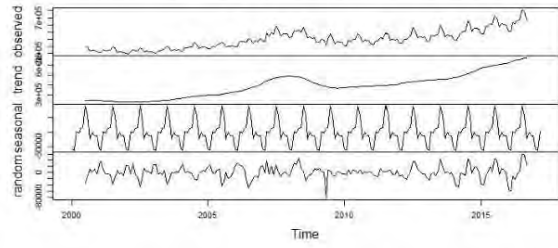
Descomposición serie de tiempo Pasajeros Totales Durango



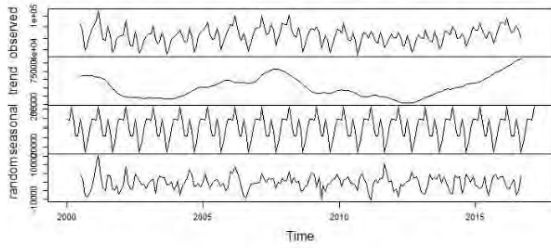
Descomposición serie de tiempo Pasajeros Nacionales Mazatlán



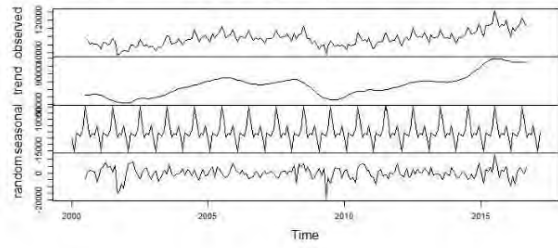
Descomposición serie de tiempo Pasajeros Nacionales Monterrey



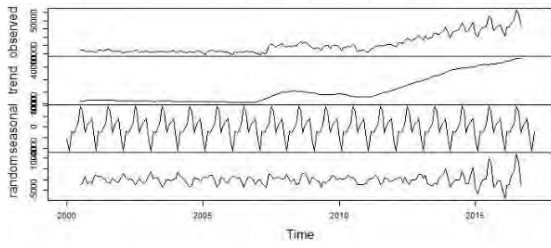
Descomposición serie de tiempo Pasajeros Totales Mazatlán



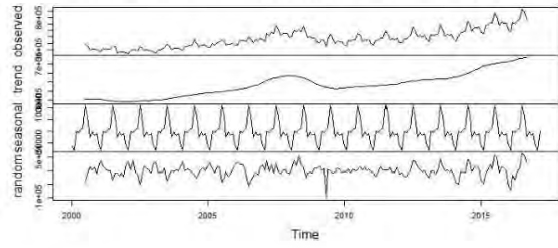
Descomposición serie de tiempo Pasajeros Internacionales Monterrey



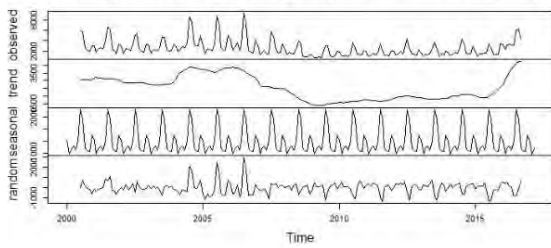
Descomposición serie de tiempo Pasajeros Nacionales Reynosa



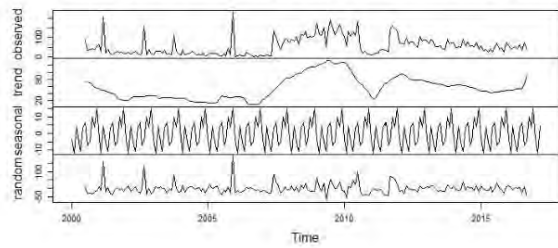
Descomposición serie de tiempo Pasajeros Totales Monterrey



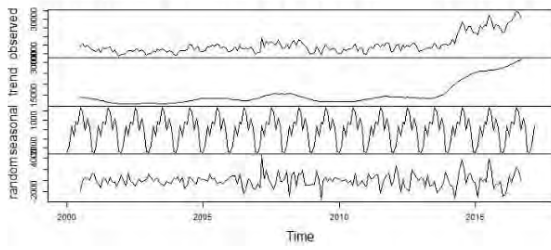
Descomposición serie de tiempo Pasajeros Internacionales Durango



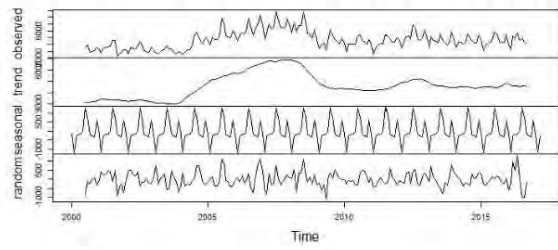
Descomposición serie de tiempo Pasajeros Internacionales Reynosa



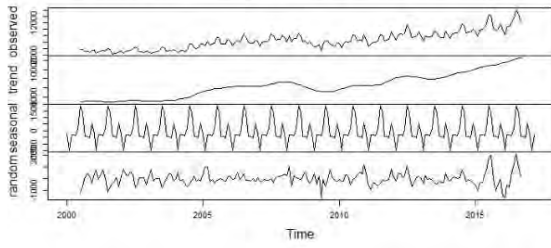
Descomposición serie de tiempo Pasajeros Nacionales San Luis Potosí



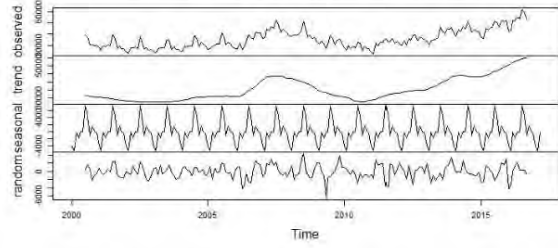
Descomposición serie de tiempo Pasajeros Internacionales Torreón



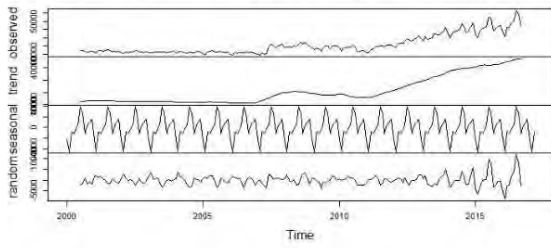
Descomposición serie de tiempo Pasajeros Internacionales San Luis Potosí



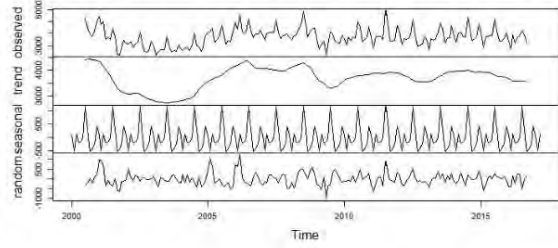
Descomposición serie de tiempo Pasajeros Totales Torreón



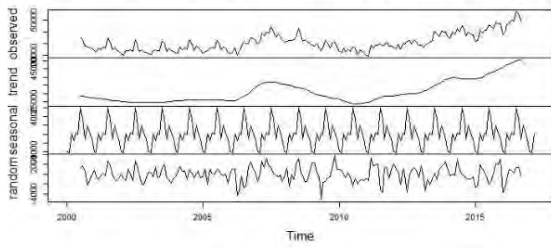
Descomposición serie de tiempo Pasajeros Totales Reynosa



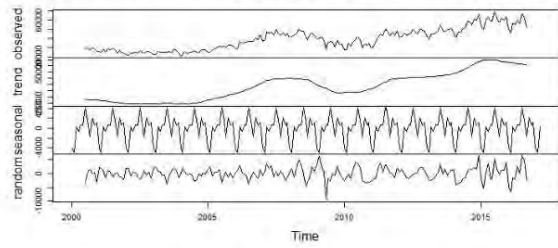
Descomposición serie de tiempo Pasajeros Internacionales Tampico



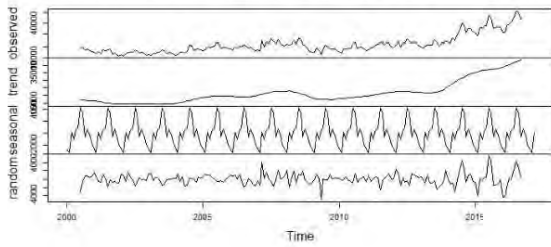
Descomposición serie de tiempo Pasajeros Nacionales Torreón



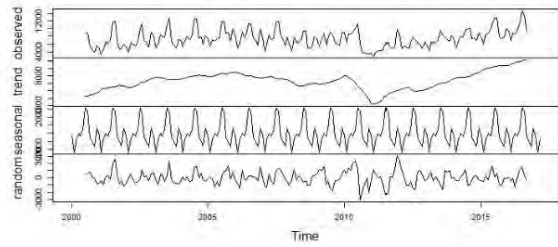
Descomposición serie de tiempo Pasajeros Nacionales Tampico



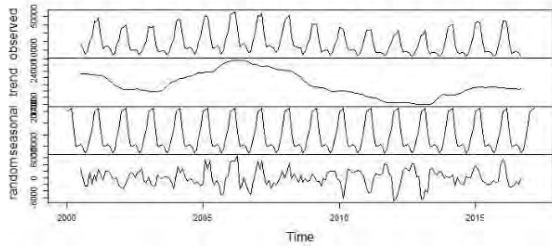
Descomposición serie de tiempo Pasajeros Totales San Luis Potosí



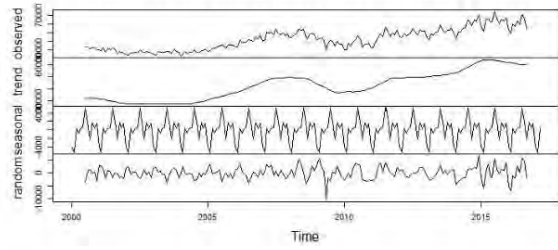
Descomposición serie de tiempo Pasajeros Internacionales Zacatecas



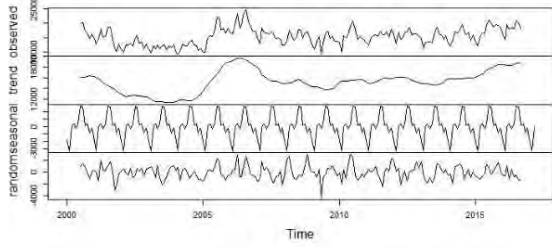
Descomposición serie de tiempo Pasajeros Internacionales Zihuatanejo



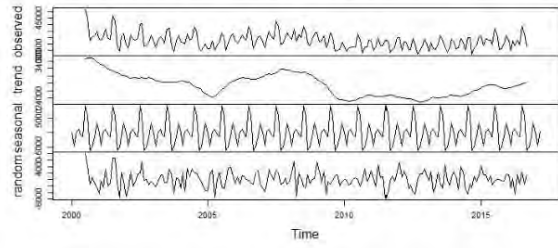
Descomposicion serie de tiempo Pasajeros Totales Tampico



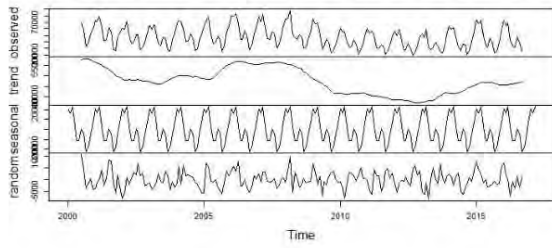
Descomposición serie de tiempo Pasajeros Nacionales Zacatecas



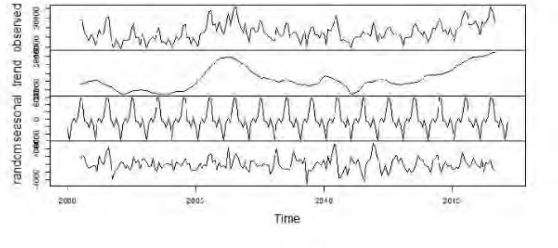
Descomposición serie de tiempo Pasajeros Nacionales Zihuatanejo



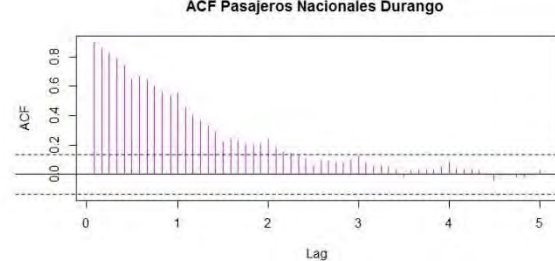
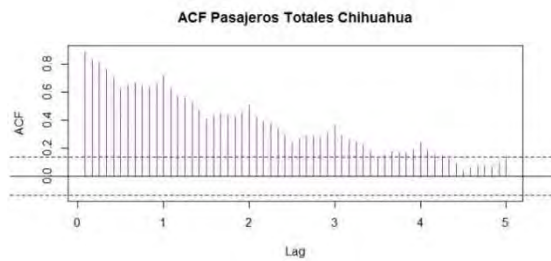
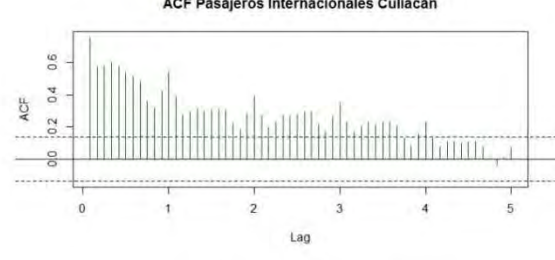
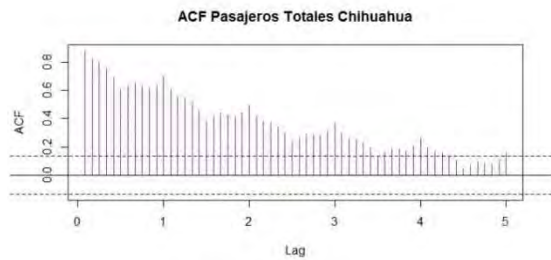
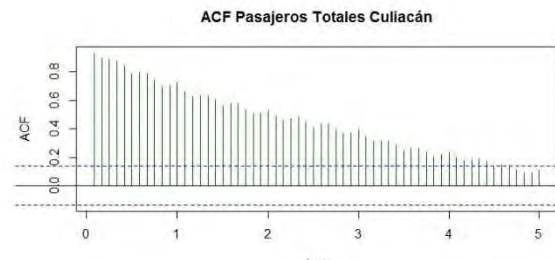
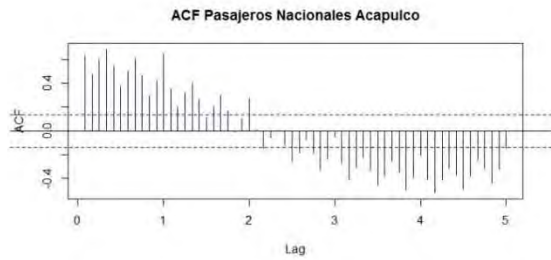
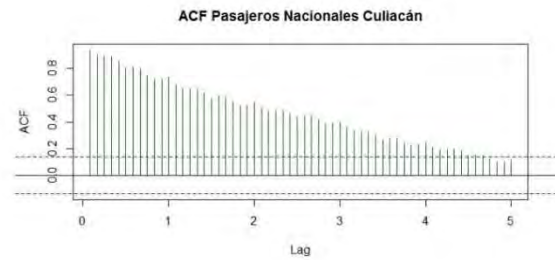
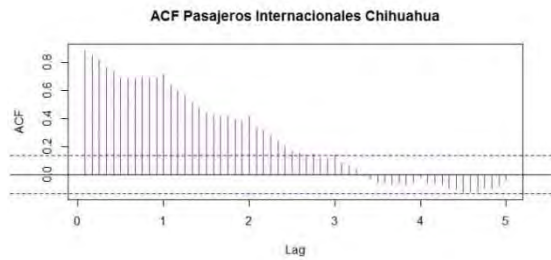
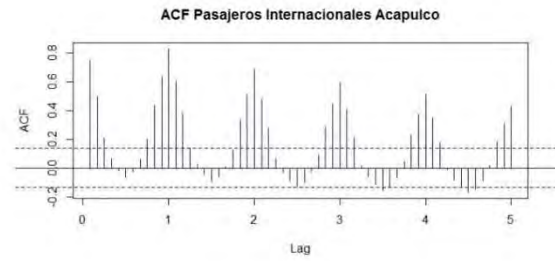
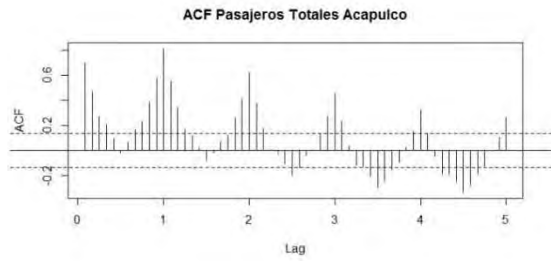
Descomposición serie de tiempo Pasajeros Totales Zihuatanejo



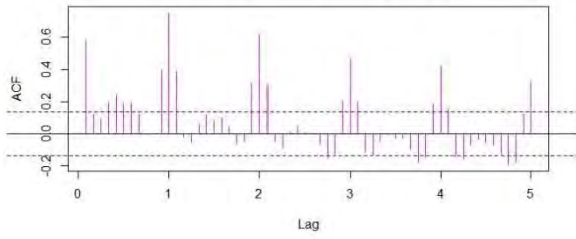
Descomposicion serie de tiempo Pasajeros Totales Zacatecas



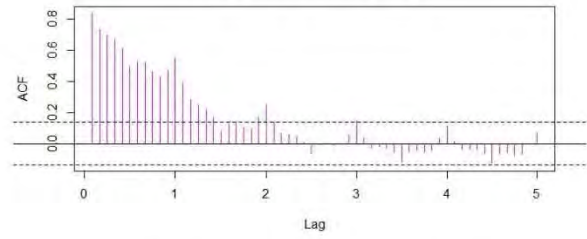
Anexo 7. Función de AutoCorrelación



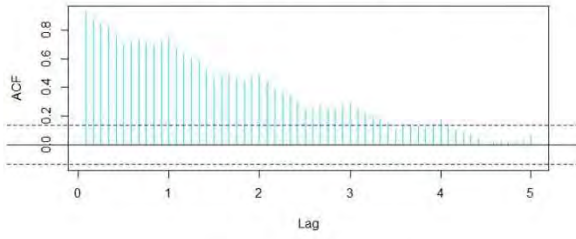
ACF Pasajeros Internacionales Durango



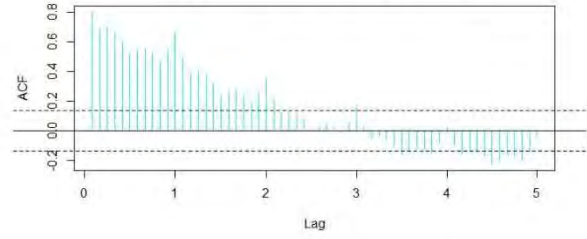
ACF Pasajeros Totales Durango



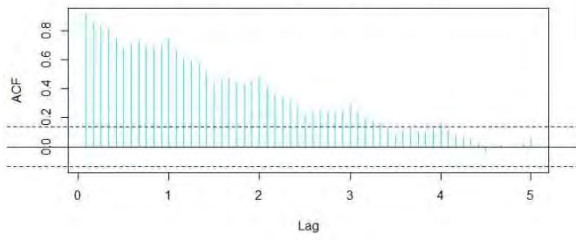
ACF Pasajeros Nacionales Monterrey



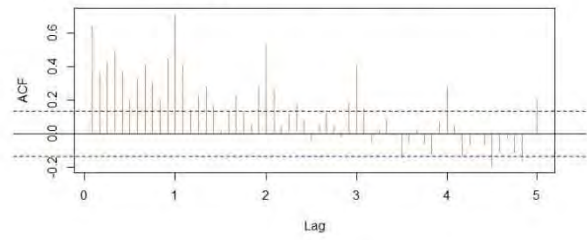
ACF Pasajeros Internacionales Monterrey



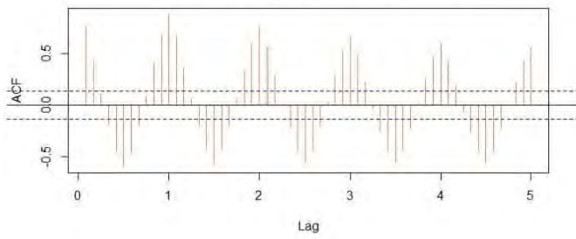
ACF Pasajeros Totales Monterrey



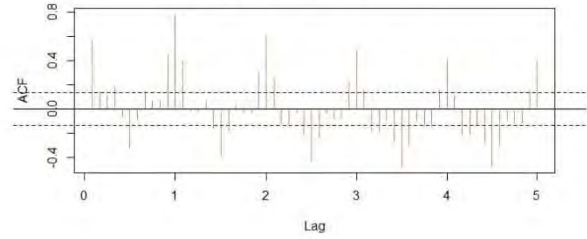
ACF Pasajeros Nacionales Mazatlán



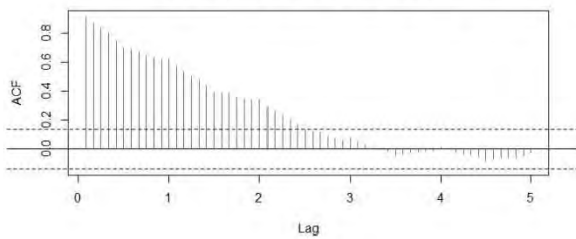
ACF Pasajeros Internacionales Mazatlán



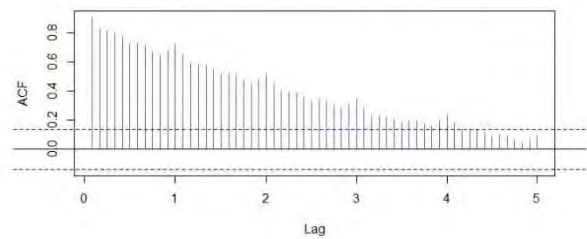
ACF Pasajeros Totales Mazatlán

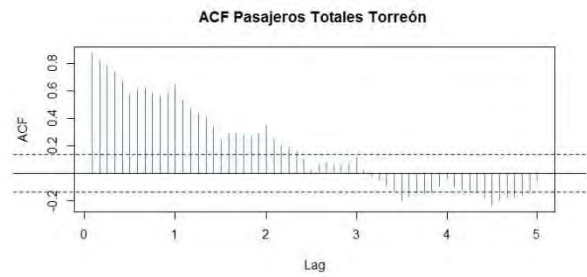
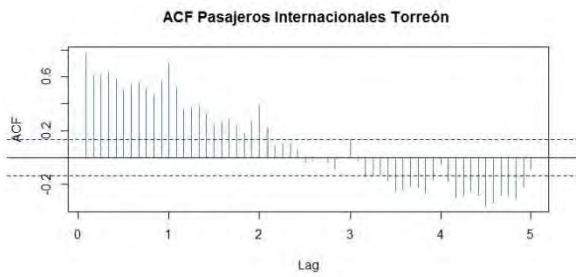
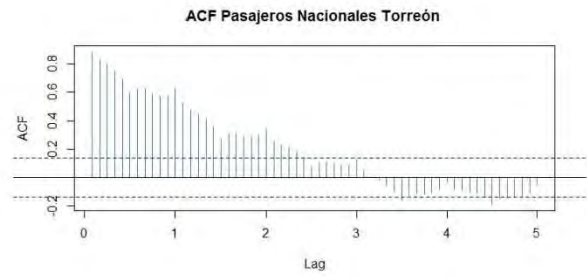
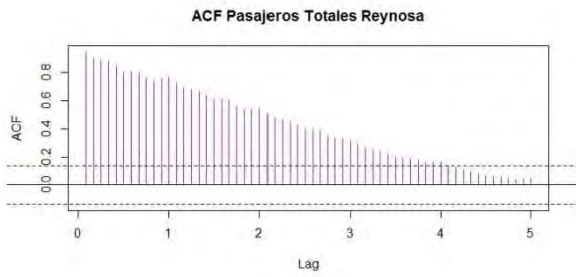
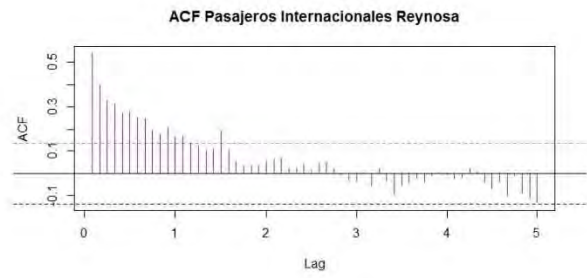
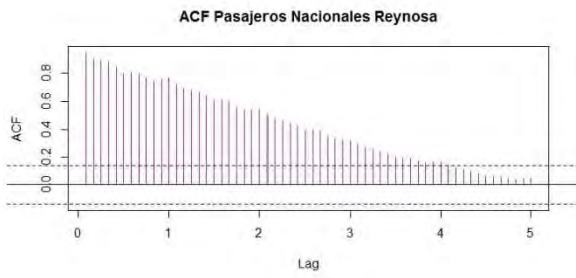
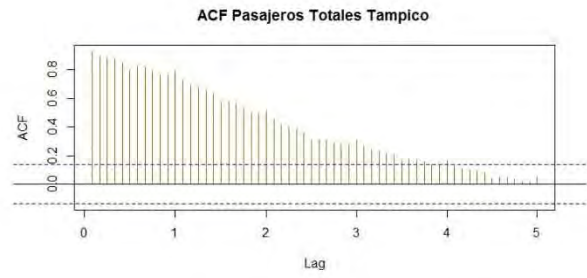
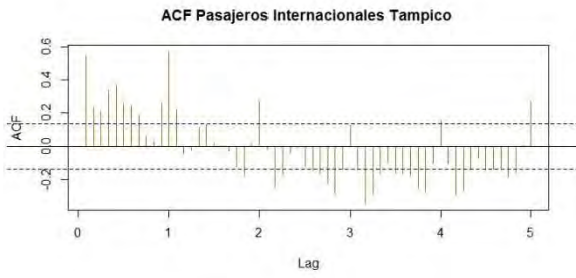
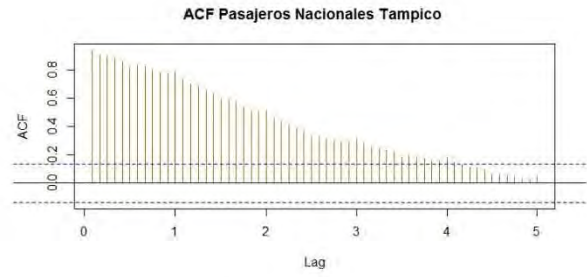
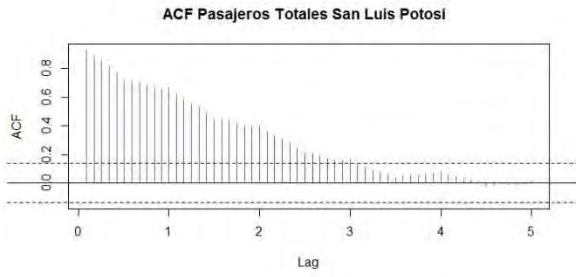


ACF Pasajeros Nacionales San Luis Potosi

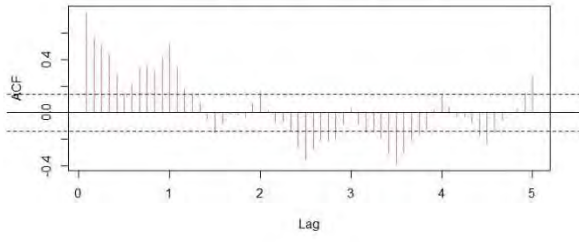


ACF Pasajeros Internacionales San Luis Potosi

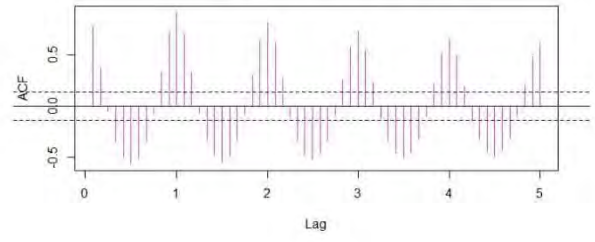




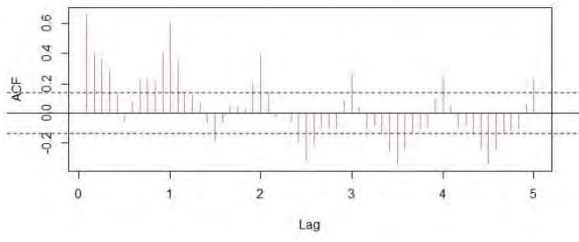
ACF Pasajeros Nacionales Zacatecas



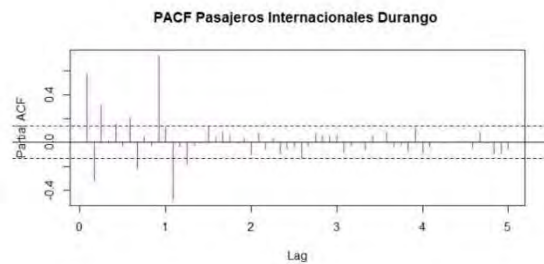
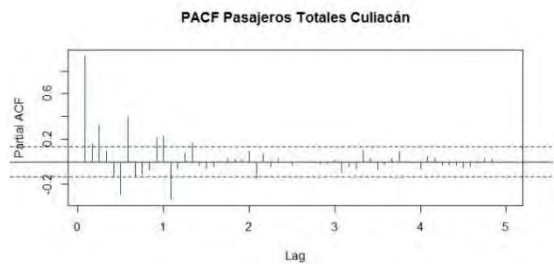
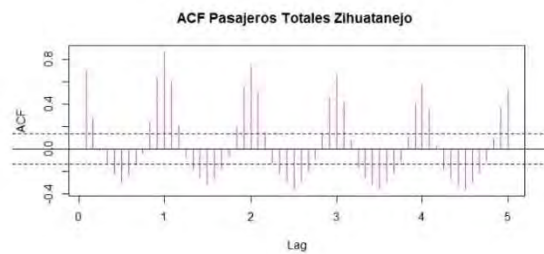
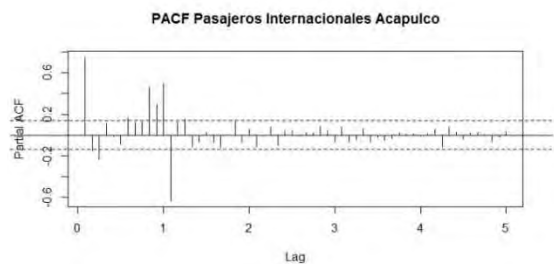
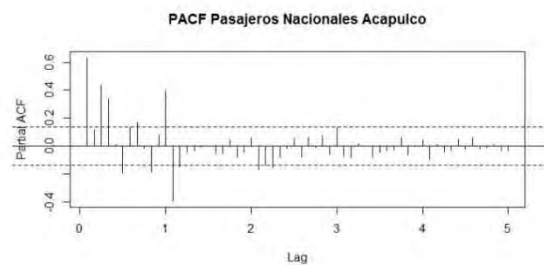
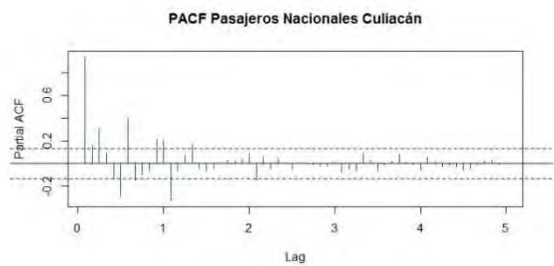
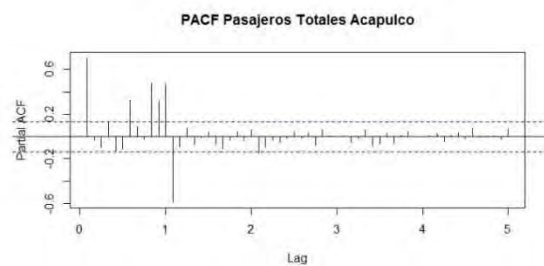
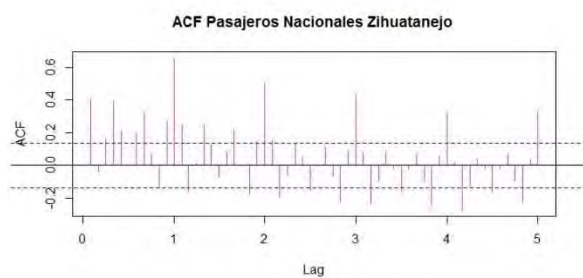
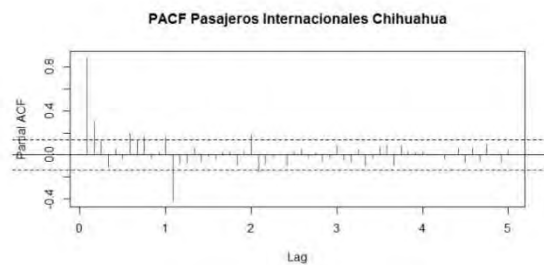
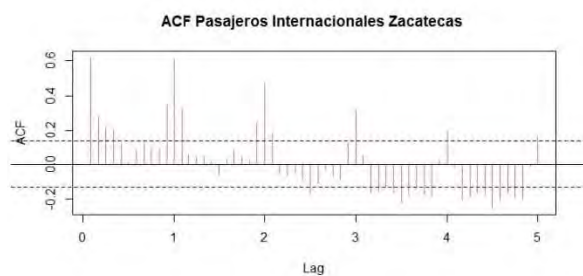
ACF Pasajeros Internacionales Zihuatanejo



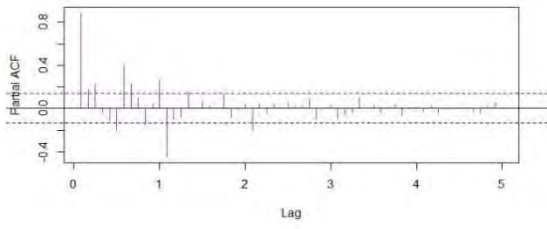
ACF Pasajeros Totales Zacatecas



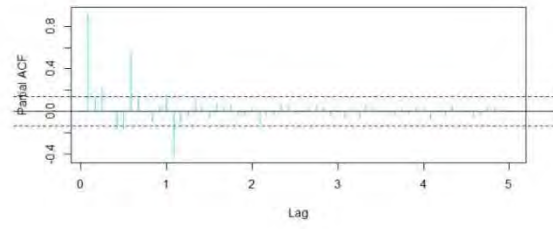
Anexo 8. Función de AutoCorrelación Parcial



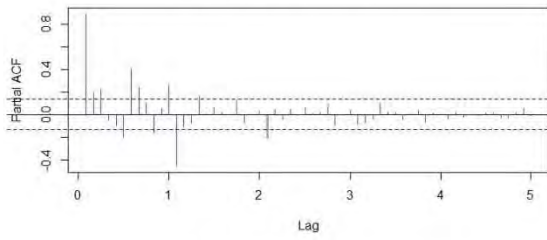
PACF Pasajeros Totales Chihuahua



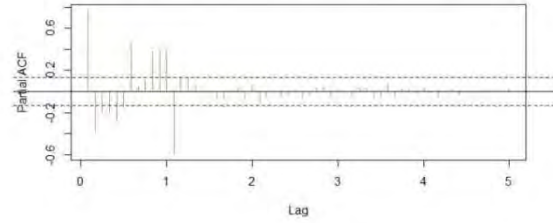
PACF Pasajeros Nacionales Monterrey



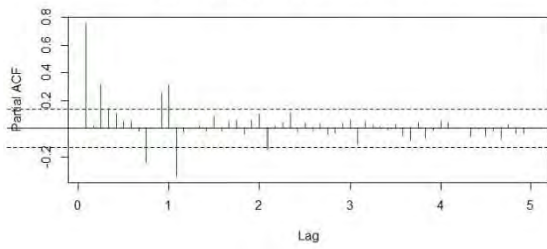
PACF Pasajeros Totales Chihuahua



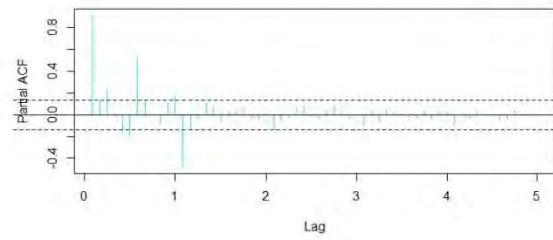
PACF Pasajeros Internacionales Mazatlán



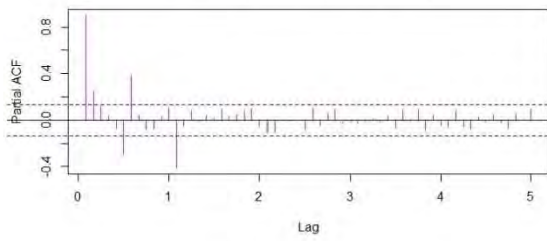
PACF Pasajeros Internacionales Culiacán



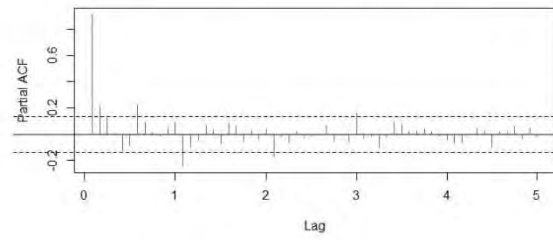
PACF Pasajeros Totales Monterrey



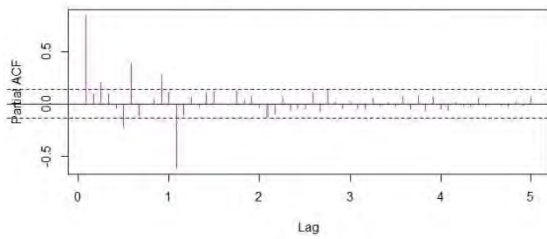
PACF Pasajeros Nacionales Durango



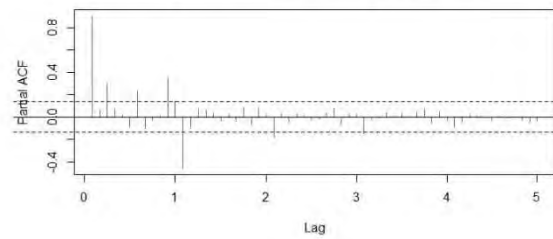
PACF Pasajeros Nacionales San Luis Potosi



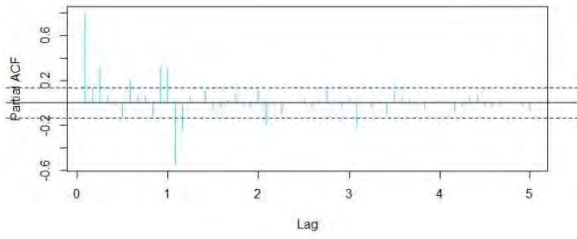
PACF Pasajeros Totales Durango



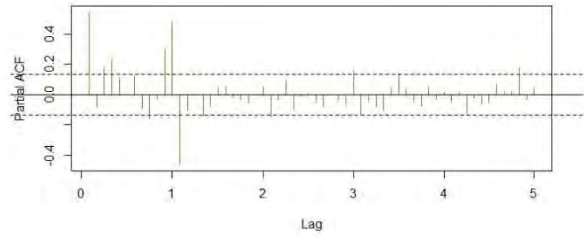
PACF Pasajeros Internacionales San Luis Potosi



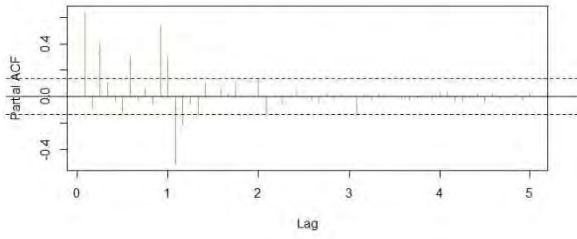
PACF Pasajeros Internacionales Monterrey



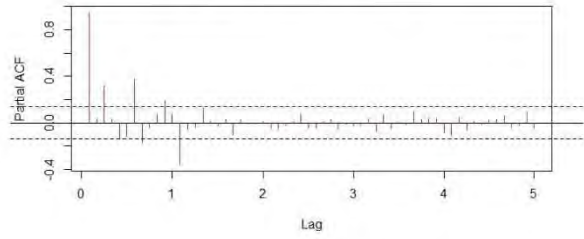
PACF Pasajeros Internacionales Tampico



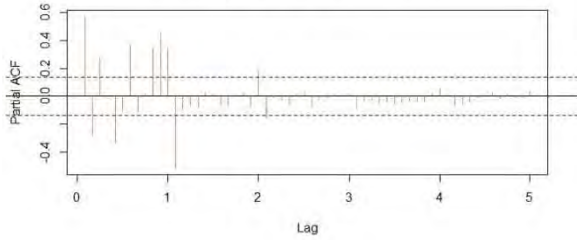
PACF Pasajeros Nacionales Mazatlán



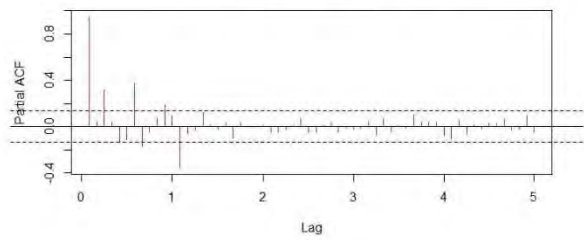
PACF Pasajeros Nacionales Reynosa



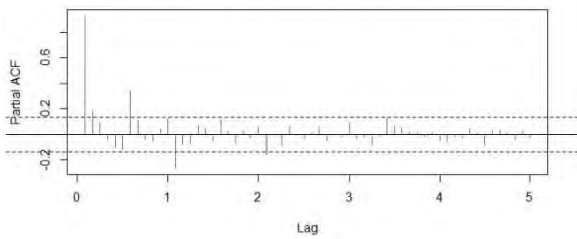
PACF Pasajeros Totales Mazatlán



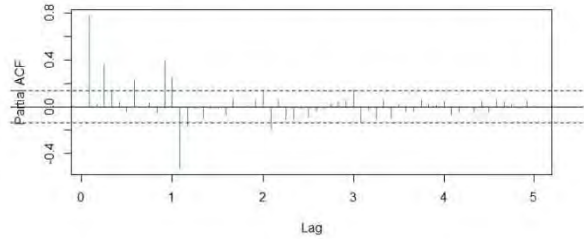
PACF Pasajeros Totales Reynosa



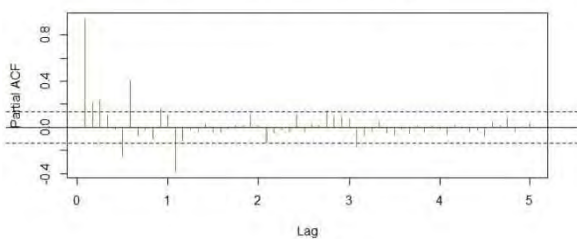
PACF Pasajeros Totales San Luis Potosi



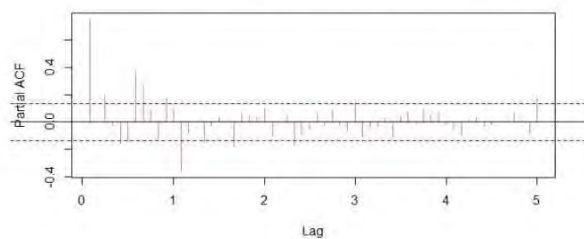
PACF Pasajeros Internacionales Torreón



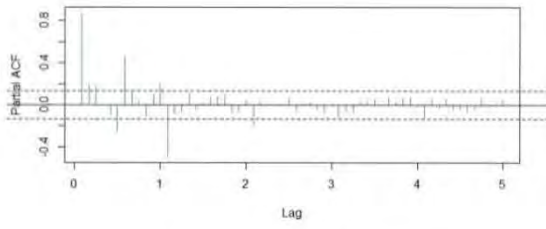
PACF Pasajeros Nacionales Tampico



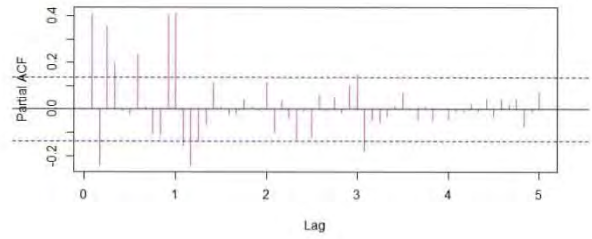
PACF Pasajeros Nacionales Zacatecas



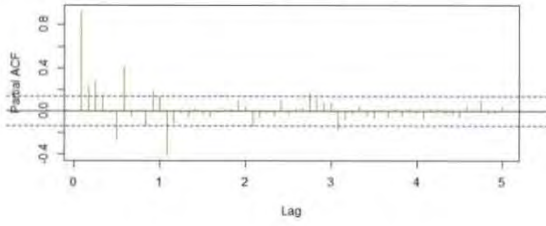
PACF Pasajeros Totales Torreón



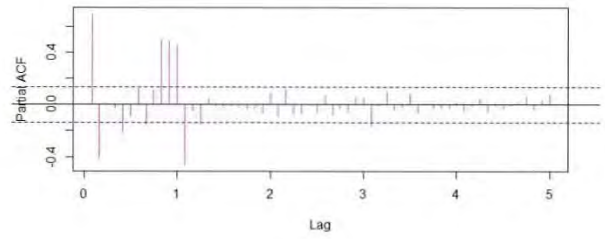
PACF Pasajeros Nacionales Zihuatanejo



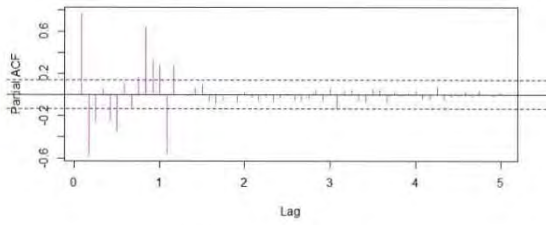
PACF Pasajeros Totales Tampico



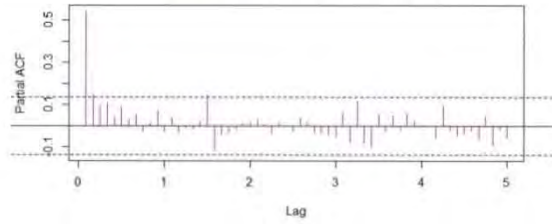
PACF Pasajeros Totales Zihuatanejo



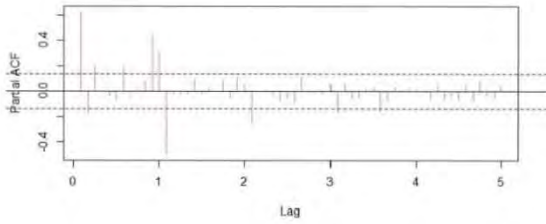
PACF Pasajeros Internacionales Zihuatanejo



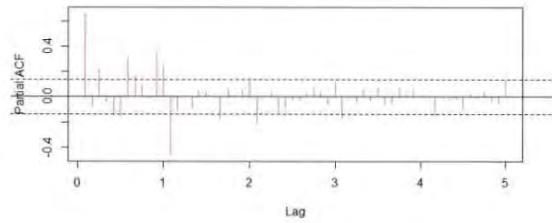
PACF Pasajeros Internacionales Reynosa



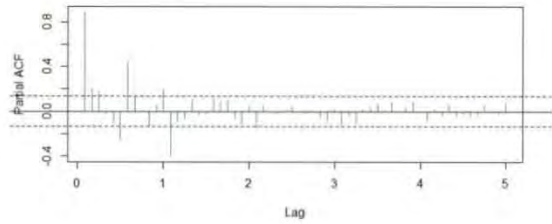
PACF Pasajeros Internacionales Zacatecas



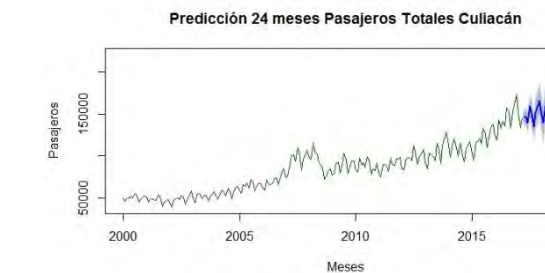
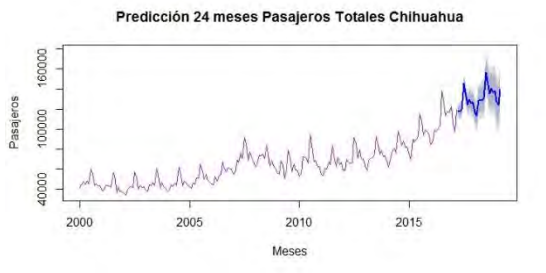
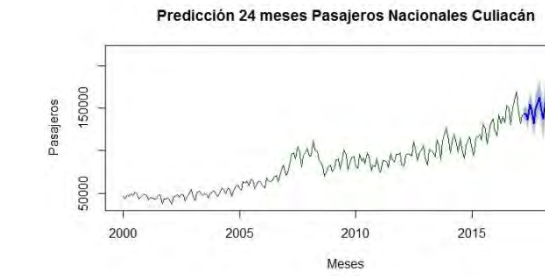
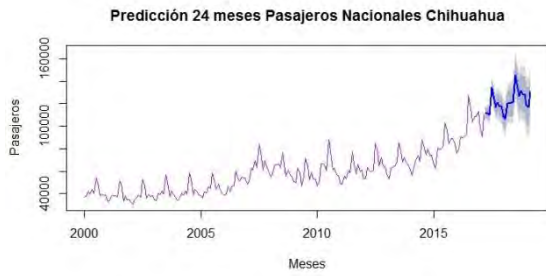
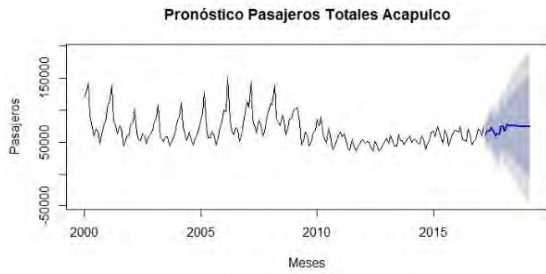
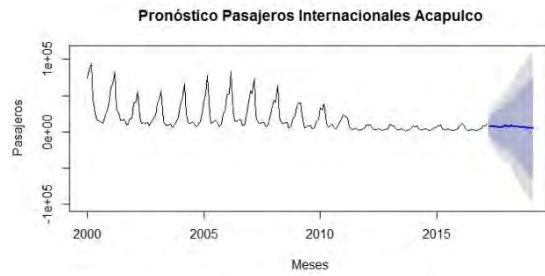
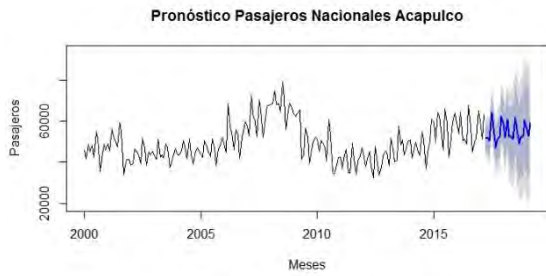
PACF Pasajeros Totales Zacatecas



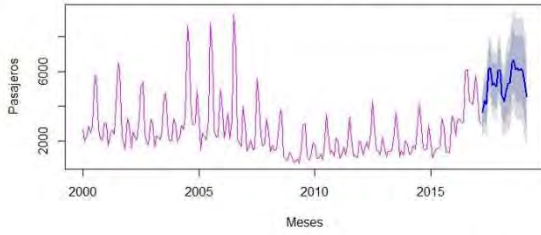
PACF Pasajeros Nacionales Torreón



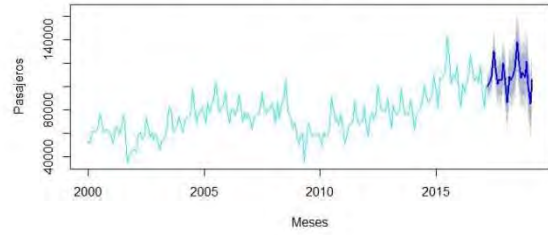
Anexo 9. Pronóstico a 24 meses



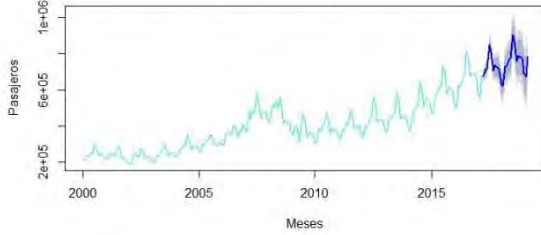
Predicción 24 meses Pasajeros Internacionales Durango



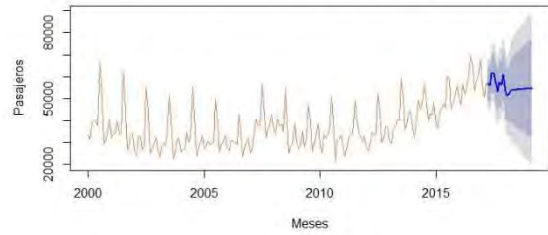
Predicción 24 meses Pasajeros Internacionales Monterrey



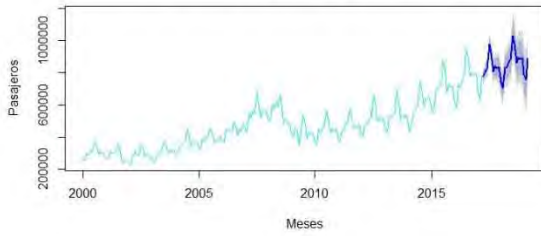
Predicción 24 meses Pasajeros Nacionales Monterrey



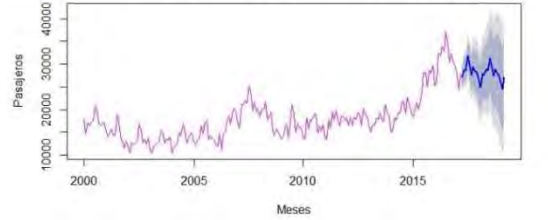
Predicción 24 meses Pasajeros Nacionales Mazatlán



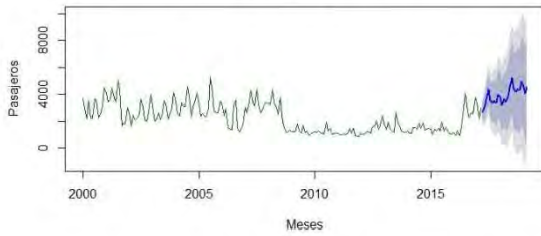
Predicción 24 meses Pasajeros Totales Monterrey



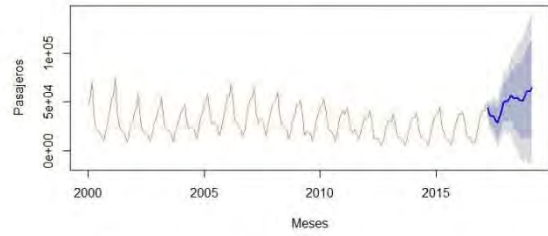
Predicción 24 meses Pasajeros Nacionales Durango



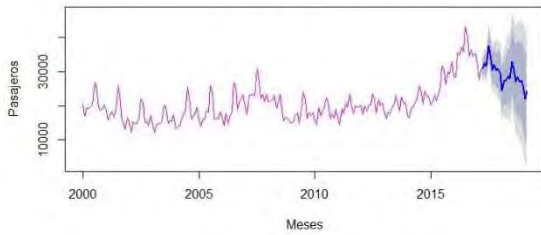
Predicción 24 meses Pasajeros Internacionales Culiacán



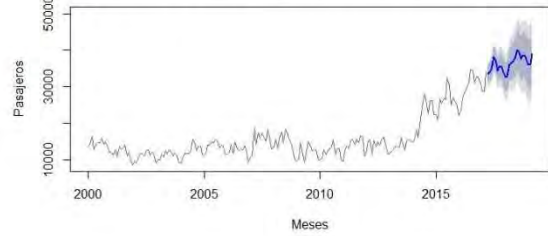
Predicción 24 meses Pasajeros Internacionales Mazatlán

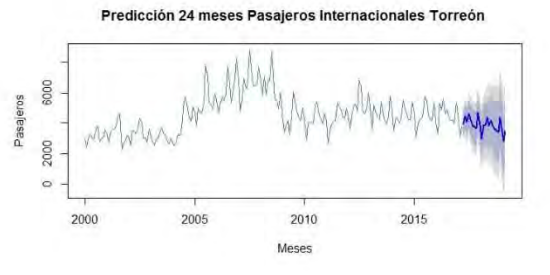
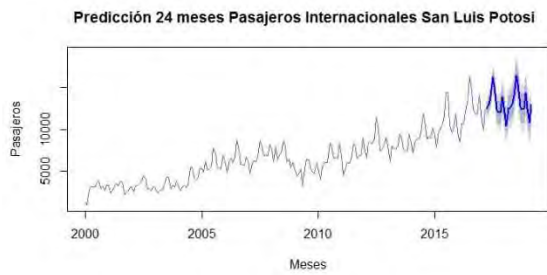
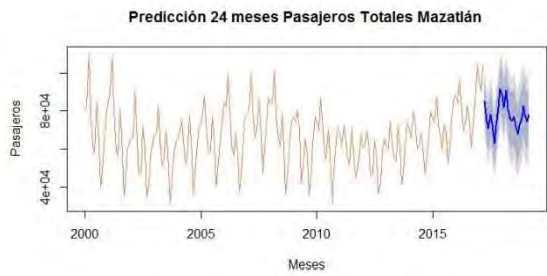
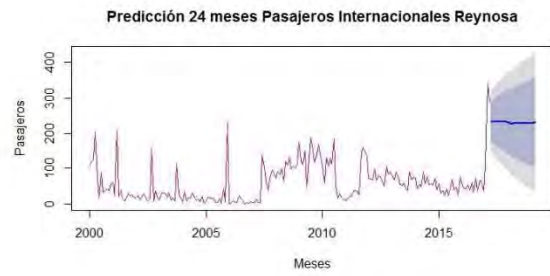
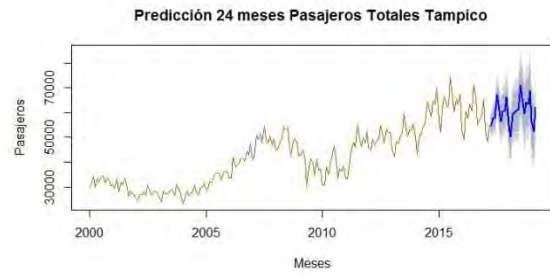
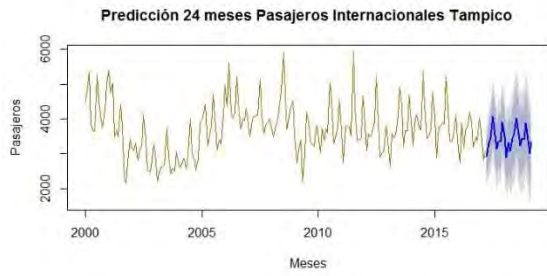
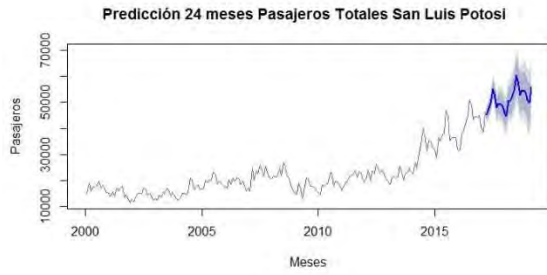


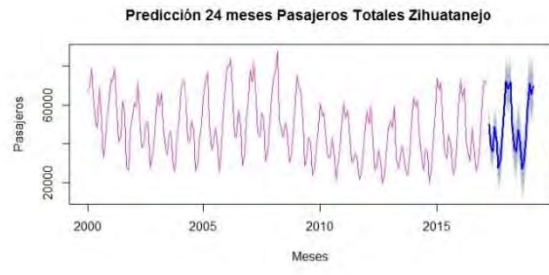
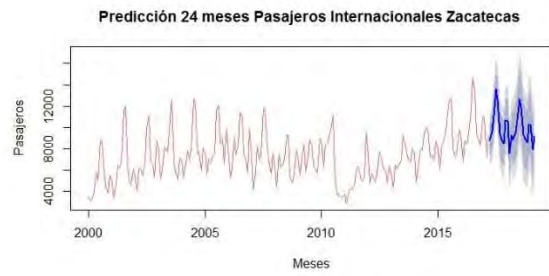
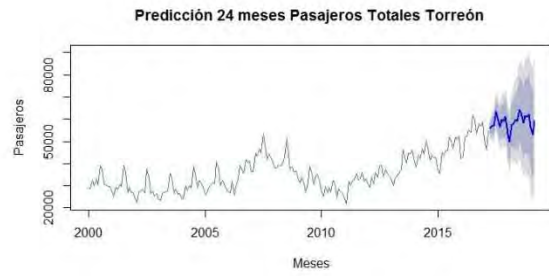
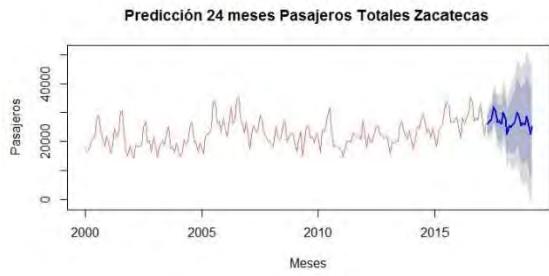
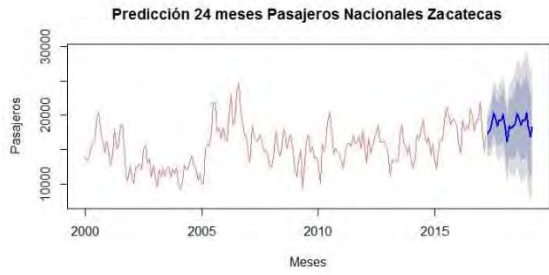
Predicción 24 meses Pasajeros Totales Durango



Predicción 24 meses Pasajeros Nacionales San Luis Potosí







Anexo 10 Resultados métricas pasajeros nacionales

Resultados Nacionales	Acapulco	Chihuahua	Culiacán	Durango	Monterrey	Mazatlán	SLP	Tampico	Reynosa	Torreón	Zacatecas	Zihuatanejo
Nodos	207	207	207	207	207	207	207	207	207	207	207	207
Arcos	638	778	851	859	791	709	789	811	932	799	741	620
Max. Grado	32	32	45	46	40	27	45	35	50	41	35	30
Min. Grado	2	2	2	2	2	2	2	2	2	2	2	2
Grado Promedio	6.164251	7.516908	8.222222	8.299517	7.642512	6.850242	7.623188	7.835749	9.004831	7.719807	7.15942	5.990338
Diámetro	8	7	6	6	7	7	6	7	6	6	6	7
Distancia Media	3.663337	3.430608	3.409831	3.203743	3.381314	3.428826	3.315182	3.651939	3.057549	3.346231	3.60152	3.600863
Cliques	8	8	10	10	8	7	9	8	10	9	9	7
Densidad	0.02992355	0.03648985	0.0399137	0.04028892	0.03709957	0.0332536	0.03700577	0.03803762	0.04371277	0.03747479	0.03475447	0.02907931
Asortatividad	0.1387072	0.03460537	0.1505972	0.1522678	0.09718336	0.06460644	0.1593228	0.1341088	0.08805418	0.07548244	0.1501778	0.2180259
Clustering Global	0.3931777	0.4603271	0.4224741	0.4160138	0.4516678	0.4340068	0.4051226	0.4120414	0.3814401	0.4250179	0.4563094	0.3993722
Clustering Local Promedio	0.7447704	0.7753178	0.7303383	0.7512943	0.7705321	0.7726576	0.7443887	0.7481026	0.7569745	0.7636464	0.7613358	0.7477119
Grado de Cercanía	0.3800197	0.3718219	0.3934506	0.4083466	0.408025	0.3591728	0.3733999	0.342245	0.3674094	0.3849608	0.40143	0.3766944
Grado de Centralidad	0.1254163	0.11885	0.1785329	0.1830121	0.1570752	0.09781436	0.1814408	0.1318653	0.1990057	0.1615543	0.1351484	0.1165518
Centralidad Intermedia	0.5545298	0.4434218	0.5462464	0.3530526	0.5472721	0.2963521	0.2396553	0.5112985	0.1860085	0.4465044	0.5597011	0.3944264

Resultados Internacionales	Acapulco	Chihuahua	Culiacán	Durango	Monterrey	Mazatlán	SLP	Tampico	Reynosa	Torreón	Zacatecas	Zihuatanejo
Nodos	207	207	207	207	207	207	207	207	207	207	207	207
Arcos	1013	796	752	758	691	920	696	700	756	682	734	944
Max. Grado	43	49	37	35	31	30	32	28	51	28	37	30
Min. Grado	1	1	2	2	2	2	6	2	1	2	2	1
Grado Promedio	9.78744	7.690821	7.2657	7.323671	6.676329	8.888889	6.724638	6.763285	7.304348	6.589372	7.091787	9.120773
Diámetro	7	6	8	8	8	8	8	8	6	9	9	7
Distancia Media	3.415224	3.302847	3.611697	3.662305	3.878946	3.625721	3.811829	4.047887	3.158388	4.357863	3.912809	3.72914
Cliques	9	9	8	7	8	9	8	8	9	7	8	10
Densidad	0.04751184	0.03733408	0.03527039	0.0355518	0.03240936	0.04314995	0.03264387	0.03283148	0.035458	0.03198724	0.03442615	0.0442756
Asortatividad	0.01292899	0.09336615	0.08541535	0.1313137	0.04117615	0.01907834	0.1233343	0.08779612	-0.1124236	0.09581404	0.1678235	0.1508116
Clustering Global	0.504932	0.4268662	0.4214998	0.4577831	0.4439925	0.5662788	0.4499064	0.4687352	0.3813288	0.4579376	0.4637783	0.5835449
Clustering Local Promedio	0.7810761	0.7602181	0.7660479	0.7761492	0.7700375	0.7853795	0.7554931	0.7862108	0.7880651	0.7637492	0.7577699	0.7882539
Grado de Cercanía	0.3303979	0.4220163	0.3425525	0.3729423	0.2979166	0.3290794	0.3229193	0.3068517	0.404093	0.2607192	0.3142012	0.3157362
Grado de Centralidad	0.161226	0.20053	0.1443413	0.1343511	0.1180761	0.1024811	0.1226959	0.1030908	0.2121148	0.1039351	0.1451855	0.1013555
Centralidad Intermedia	0.503284	0.5204397	0.5534362	0.5314784	0.4535664	0.4879979	0.5039856	0.5445958	0.489614	0.5330167	0.3315665	0.5074025

Resultados Totales	Acapulco	Chihuahua	Culiacán	Durango	Monterrey	Mazatlán	SLP	Tampico	Reynosa	Torreón	Zacatecas	Zihuatanejo
Nodos	207	207	207	207	207	207	207	207	207	207	207	207
Arcos	801	787	784	812	774	675	794	767	921	802	697	722
Max. Grado	44	33	40	39	34	31	42	38	80	42	36	29
Min. Grado	2	2	2	2	2	2	2	2	2	2	2	2
Grado Promedio	7.73913	7.603865	7.574879	7.845411	7.478261	6.521739	7.671498	7.410628	8.898551	7.748792	6.7343	6.975845
Diámetro	7	7	7	6	7	8	6	7	6	6	8	7
Distancia Media	3.597908	3.384222	3.502228	3.27888	3.487454	3.62891	3.29417	3.537076	3.073777	3.364007	4.008395	3.820834
Cliques	9	9	9	10	9	9	9	8	11	9	7	7
Densidad	0.03756859	0.03691196	0.03677126	0.03808452	0.03630224	0.03165893	0.03724028	0.03597392	0.04319685	0.0376155	0.03269077	0.03386333
Asortatividad	0.03675553	0.05686128	0.143146	0.1646557	0.08005149	0.1258798	0.1169604	0.1597576	0.08883111	0.06912177	0.1016958	0.1629528
Clustering Global	0.4548834	0.4561629	0.4139344	0.4320165	0.4624291	0.4109068	0.4089318	0.4136622	0.3801761	0.4316502	0.4650899	0.484918
Clustering Local Promedio	0.7675662	0.7737922	0.7350116	0.7660147	0.7809941	0.7533366	0.7466265	0.7515967	0.757164	0.7641049	0.754838	0.7479327
Grado de Cercanía	0.3635828	0.3866741	0.3756692	0.4063534	0.3838659	0.3380337	0.3766335	0.3744411	0.3681372	0.3971153	0.3246542	0.3633273
Grado de Centralidad	0.1760236	0.1232822	0.1574035	0.1512359	0.1287463	0.1188265	0.1666432	0.1484921	0.1995216	0.166268	0.1420665	0.1069134
Centralidad Intermedia	0.5915958	0.4493342	0.5408647	0.3826214	0.5439342	0.3499375	0.2346736	0.5184047	0.1834915	0.5272701	0.5316618	0.5710251

Anexo 13 Código en R

Visibilidad

```
##Visibilidad
setwd("C:/Users/sshirai/Documents")
data<-read.csv("pasajeros-aca.csv",header=T,na.strings="NA",quote="", sep=",")
datab<-c(data$Nacionales)
plot(datab,xlab="Meses",ylab="Pasajeros", main="Serie Pasajeros Nacionales Acapulco ")
plot(datab,xlab="Meses",ylab="Pasajeros", main="Visibilidad Pasajeros Nacionales Acapulco ")
n2 <- cbind(datab, nuevaCol=c(1:length(datab)))
n3<-n2[,c(2,1)]
lines(datab, col="indianred1")
ady<-matrix(0,nrow=length(datab),ncol=length(datab))

#visibilidad
for(i in 1:(length(n3)/2-1)){
  for(j in i:(length(n3)/2)){
    r<-c(0)
    for(k in (i+1):j){
      ta<-n3[i,1]
      ya<-n3[i,2]
      tb<-n3[j,1]
      yb<-n3[j,2]
      tc<-n3[k,1]
      yc<-n3[k,2]
      if((yc<(yb+(ya-yb)*((tb-tc)/(tb-ta)))) & (tb-ta)>0 & (tb-tc)>0){
        r<-r+1
      }
      if(r==(j-i-1)){
        segments(ta,ya,tb,yb,col = "red")
        ady[i,j]<-1
      }
    }
  }
}
```

Redes

```
#redes
library(sand)
x<-graph.adjacency(ady, mode=c("undirected"))

library(igraph)

summary(x)
max(igraph::degree(x, mode="all")) # Grado máximo total
min(igraph::degree(x)) # Grado mínimo
mean(igraph::degree(x)) #Gradopromedio
diameter(x) #Diámetro
mean_distance(x, directed=FALSE, unconnected=FALSE) # Distancia media
clique.number(x) # Número de cliques
graph.density(x) # Obtenemos la densidad
assortativity_degree(x, directed=FALSE) #Asortatividad
transitivity(x) # Clustering global
transitivity(x, type="average") # Clustering promedio
centr_clo(x, mode = "all")$centralization #Grado de cercanía
centr_degree(x)$centralization #Grado de centralidad
centr_betw(x, directed = FALSE)$centralization #Centralidad intermedia
centr_eigen_tmax(x) #Centralidad eigenvector

plot(degree.distribution(x), type="l", xlab="Nodos" ,col="indianred1", main="Distribución de Grado Pasajeros
Nacionales Acapulco ") #Gráfica distribución de grado

library(ggplot2)
library(ggnetwork)
library(GGally)

ggnet2(x, size = 7, color = "indianred1", edge.size = 1, edge.color = "gray", main="Red") #Gráfica red
```

Series de Tiempo

```

##Series de tiempo
#Pronóstico
library(timeSeries)
library(TSA)
library(forecast)
data<-read.csv("pasajeros-aca.csv",header=T,na.strings="NA",quote="", sep=",")
datab<-c(data$Nacionales)
databs<-ts(datab, start=c(2000,1), end=c(2017,3), frequency = 12) #serie de tiempo

plot.ts(databs, xlab="Meses",ylab="Pasajeros", main="Pasajeros Nacionales Acapulco ",col="indianred1") #Gráfica serie
de tiempo

plot(decompose(databs)) #Gráfica descomposición serie de tiempo

# Descomposición estacional
fit <- stl(databs, s.window="period")
plot(fit,main="Pasajeros Nacionales Acapulco ",col="indianred1")

# Gráfica meses y estacional
monthplot(databs)
seasonplot(databs)

acf(databs, lag.max=60, main=c("ACF Pasajeros Nacionales Acapulco "),col="indianred1") #ACF
pacf(databs, lag.max=60, main=c("PACF Pasajeros Nacionales Acapulco "),col="indianred1") #PACF

mod<-auto.arima(databs) #Ajuste modelo
mod

plot.ts(residuals(mod), main=c("Residuales Pasajeros Nacionales Acapulco "),col="indianred1") #Gráfica residuales
acf(residuals(mod), lag.max=60, main=c("ACF Residuales Pasajeros Nacionales Acapulco ")) #Gráfica ACF residuales
pacf(residuals(mod), lag.max=60, main=c("PACF Residuales Pasajeros Nacionales Acapulco ")) #Gráfica PACF
residuales

plot(forecast(mod, h=24), main="Predicción 24 meses Pasajeros Nacionales Acapulco ",xlab=c("Meses"),
ylab=c("Pasajeros"),col="indianred1") #Gráfica Pronóstico a 24 meses

```

Bibliografía

- [1] Barabási A.-L., Albert R., 1999. Emergence of scaling in random networks. *Science* 286 (5439): 509-512.
- [2] Barabási A.-L., Bonabeau E., 2003. Scale-Free Networks. *Scientific American*. 50-59.
- [3] Barrat A., Barthélemy M., Pastor-Satorras R., Vespignani A., 2004. The architecture of complex weighted networks. *PNAS* Vol.1 No.11. 3747-3752.
- [4] Boccaletti S., Latora V., Moreno Y., Chavez., Hwang D.-U., 2006. Complex networks: Structure and dynamics. *Physics Reports*. 424. 175-308.
- [5] Boccaletti S., Bianconi G., Criado R., del Genio C.I., Gómez-Gardeñes J., Romance M., Sendiña-Nadal I., Wang Z., Zanin M., 2014. The structure and dynamics of multilayer networks. *Physics Reports*. 544. 1-122.
- [6] Bollobás B., 1998. *Modern Graph Theory*. Springer.
- [7] Buchanan M., 2002. *Nexus: Small Worlds and the Groundbreaking Science of Networks*. New York, N.Y.: W.W. Norton.
- [8] Caldarelli G., Catanzaro M., 2012. *Networks: A Very Short Introduction*. Oxford University Press, Oxford.
- [9] Caldarelli G., Vespignani A., 2007. *Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science*. World Scientific Publishing Co., Inc., USA.
- [10] Dalgaard P., 2008. *Introductory Statistics with R*. Springer.
- [11] Dorogovtsev S.N., 2010. *Lectures on Complex Networks*. Oxford University Press, Inc. New York, NY, USA.
- [12] Erdős P., Rényi A., 1959. *On random graphs, Publ. Math. Debrecen*, 6, pp. 290–297.
- [13] Kolaczyk E.D., Csárdi G., 2014. *Statistical Analysis of Network Data with R*. Springer, New York, USA
- [14] Lacasa L., Luque B., Ballesteros F., Luque J., Nuño J.C., 2008. From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences of the United States of America, PNAS*. Vol. 5. No, 13. 4972-4975.
- [15] Lacasa, L. et al., 2015. Network structure of multivariate time series. *Sci. Rep.* 5,15508; doi: 10.1038/srep15508.
- [16] Lütkepohl H., 2005. *New Introduction to Multiple Time Series Analysis*. Springer.
- [17] Mood A., 1974. *Introduction to the Theory of Statistics*. McGraw-Hill.

- [18] Mujica M., Flores I., Guimarans D., 2015. *Applied Simulation and Optimization in Logistics, Industrial and Aeronautical*. Springer.
- [19] Newman M.E.J., 2010. *Networks: An Introduction*. Oxford University Press, Oxford.
- [20] Newman M.E.J., 2003. The Structure and Function of Complex Networks. *SIAM Review* 45,167-256.
- [21] Newman M.E.J., Barabási A.-L., Watts D.J., 2006. *The Structure and Dynamics of Networks*. Princeton University Press, Princeton, NJ, USA.
- [22] Newman M.E.J., Peixoto T.P., 2015. Generalized Communities in Networks. *Physical Review Letters* 115, 08870. American Physical Society.
- [23] Watts D. J., 2003. *Six Degrees: The Science of a Connected Age*, Norton, New York.
- [24] Watts D. J., Strogatz S. H., 1998. *Collective dynamics of "small-world" networks*, *Nature*, 393, pp. 440–442
- [25] Wei W. W. S., 2006. *Time Series Analysis Univariate and Multivariate Analysis*. Pearson.
- [26] Zanin M., Lacasa L., Cea M., Cristobal S., 2009. A dynamical model for the air transportation network. *Proceedings 23rd European Conference on Modelling and Simulation*.
- [27] Zanin M., Lillo F., 2013. *Eur. Phys. J. Spec. Top.* 215:5. doi:10.1140/epist/e2013-01711-9
- [28] Zumbach G., 2013. *Discrete Time Series*. Springer.
- [29] www.oma.aero/es/ [consultada en Mayo 2017]
- [30] <https://www.gob.mx/sct> [consultada Marzo 2018]